

Derin Öğrenme Araştırma Alanlarının Literatür Taraması

M. Mutlu YAPICI^a, Adem TEKEREK^{*b}, Nurettin TOPALOĞLU^c

^a. Ankara University, Computer Technologies Department, Ankara, 06100, Turkey

^{b,*} Gazi University, Information Technology Department, Ankara, 06500, Turkey

^c Gazi University, Technology Faculty, Computer Engineering Department, Ankara, 06500, Turkey

MAKALE BİLGİSİ

Alınma: 07.09.2019
Kabul: 23.11.2019

Anahtar Kelimeler:
Derin Öğrenme,

Yapay Sinir Ağları,

Evrişimsel Sinir
Ağları,

Makine Öğrenme

***Sorumlu Yazar:**
e-posta:
atekerek@gazi.edu.tr

ÖZET

Derin öğrenme (Deep Learning-DL), birçok alanda önemli başarılar elde etmiş güçlü bir makine öğrenmesi yöntemidir. Özellikle son on yılda, bilgisayarlı görü, nesne tanıma, konuşma tanıma, doğal dil işleme gibi birçok araştırma alanında başarılı sonuçlar elde ederek, yapay zekanın derin uykudan uyanmasına yol açmıştır. Günümüzde, çeşitli alanlardaki birçok araştırmacı, DL yöntemlerini kullanarak alanlarında en iyi sonucu almaya çalışmaktadır. Bu tarama çalışmasında, DL modelleri ve DL ile çalışılabilecek önemli araştırma konuları hakkında bilgiler vererek araştırmacılara rehberlik etmeyi hedefliyoruz. Çalışmada Özerk Araçlar (Autonomous Vehicles), Doğal Dil İşleme (Natural Language Processing), El Yazısı Karakter Tanıma (Handwritten Character Recognition), İmza Doğrulama (Signature Verification), Ses ve Video Tanıma (Voice and Video Recognition), Tıbbi Görüntü İşleme (Medical Image Processing), Büyük Veri (Big Data) gibi dünyanın en popüler ve en zorlu alanlarında yapılan DL çalışmalarını inceliyoruz. Ayrıca, araştırmacılara yardımcı olmak için, incelediğimiz bu alanlardaki DL ile çalışılabilecek, henüz çalışılmamış veya yeterince iyi sonuçlar elde edilememiş problemlere dikkat çekerek olası araştırma konularını listeliyoruz. Günümüzde, Derin öğrenme yöntemleri nesne tanıma alanında insanlardan daha iyi sonuçlara ulaşmıştır. DL üzerine yapılan çalışmalar dikkate alındığında, bu başarının otonom araçlar, tıbbi görüntü işleme, büyük veri analizi ve karakter tanıma gibi alanlarda da gerçekleşeceği öngörülmektedir.

<https://dx.doi.org/10.30855/gmbd.2019.03.01>

Literature Review of Deep Learning Research Areas

ARTICLE INFO

Received: 07.09.2019
Accepted: 23.11.2019

Keywords:
Deep learning,

Artificial Neural
Networks,

Convolutional Neural
Network,

Machine Learning

***Corresponding
Authors**
e-mail:
atekerek@gazi.edu.tr

ABSTRACT

Deep learning (DL) is an important machine learning field that has achieved considerable success in many research areas. In the last decade, the state-of-the-art studies on many research areas such as computer vision, object recognition, speech recognition, and natural language processing were especially led to the awakening of the artificial intelligence from deep sleep. Nowadays, many researchers try to find solutions to many problems in various fields under the light of DL methods. In this study, it is presented important knowledge to guide about DL models and challenging topics that can be used in DL for researchers. This study investigated DL studies which are made in the most popular and challenging fields such as autonomous vehicles, natural language processing, handwritten character recognition, signature verification, voice and video recognition, medical image processing, and big data. Furthermore, this study points out the remaining challenges of these research areas that can be solved by DL, and discusses future topics to help the researchers. In the present day, Deep learning methods have reached better results than humans in object recognition. According to the literature studies on DL, It is foreseen that this success will be achieved in areas such as autonomous vehicles, medical image processing, big data analysis, and character recognition.

<https://dx.doi.org/10.30855/gmbd.2019.03.01>

1. INTRODUCTION (*Giriş*)

In last decades, Artificial Neural Network (ANN) based Machine Learning (ML) approaches have driven advances in many different research areas [1–24]. The use of Deep Learning (DL) methods, which are the multi-layered structure of ANNs along with the improvements of GPU technology, have accelerated these advances. Furthermore, DL approaches have significantly outperformed state-of-the-art approaches in many fields such as object recognition [1, 3, 7, 9, 25, 26], image processing [11, 27–32], computer vision [33–36], speech recognition [37–39], natural language processing (NLP) [10, 21, 27, 40–42], character recognition [5, 30, 43–46], signature verification [2, 6, 47–51]. Although the foundations of DL were based on ANN proposed by McCulloch and Pitts in 1943 [52], the real popularity has increased in 2012. Multilayer neural networks have been unsuccessful for a long time as they have consistently achieved a local optimal solution. In addition, interest in multi-layered neural networks has not been more significance for a long time because the processing power of big data had increased too much. In 2006, Hinton et al. [53] proposed a two-stage strategy for training DL effectively, which are pre-training, and fine-tuning with respectively. This was the first step to increased interest in DL. In 2012, Krizhevsky et al. [54] achieved a great success by improving the Top-5 error rate from 26.2% to 15.3% in the ImageNet competition which is the most important competition in the object recognition field. This success achieved great reactions in the academy, and increased interest in DL. Besides the academic communities, many technology companies also contribute to the development of DL approaches by supporting. Many companies, such as Google [55, 56], Facebook [57], Microsoft [58, 59] and NVIDIA [60, 61] have developed their own DL frameworks and released them as open source software so that researchers working on this area can develop new DL models. The success of DL is attributed to its high representational ability of input data, by using various layers such as Fully Connected (FC), Dropout, and Pooling. DL discovers intricate structure in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation between layers [62, 63]. In literature, there are various DL architectures such as Deep Neural Network (DNN), Convolutional Neural Network (CNN), Deep Belief Network (DBN), Sparse Auto-Encoder (SAE), and

Recurrent Neural Network (RNN). Although the essential structures of these frameworks are same, there are also some differences between DL models. Fig. 1. shows the approximate number of published DL articles by year. As it can be seen clearly in the Fig. 1, interest in DL has increased rapidly after 2012.

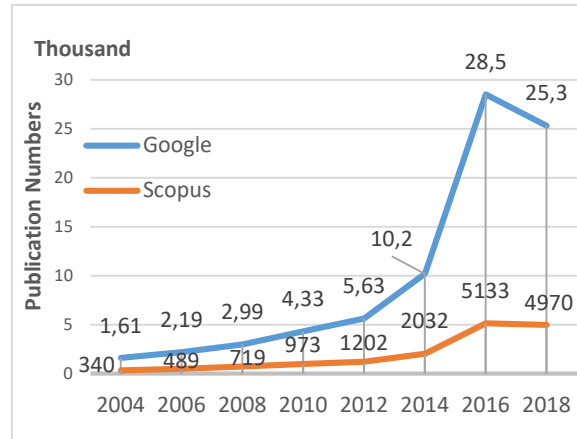


Figure 1. Number of published deep learning articles by year. The numbers of articles were obtained from the search results on Scopus and Google Scholar with the query of ‘Deep learning’.

The basic structure of DNNs consists of an input layer, hidden layers and an output layer as in ANNs. The difference between DNNs and ANNs is the number of hidden layers which is more than one and directly affects the depth of the algorithm in DNNs. Firstly, raw data is given to the input layer, and secondly, all values are calculated sequentially along with the network layers as output. The obtained output value is used as input data for each next layer. The output values are calculated via the sum of the multiplication of the input and the weight for each unit in the current layer. Then, a non-linear function such as a rectified linear unit (ReLU), hyperbolic tangent or sigmoid is applied to compute the output values of the layer. The output data obtained as it progresses towards the last layer contains slightly more abstract representations than the state of the raw input data. It allows a more successful classification with the data [64]. According to the developments in technology and the rapid increase in the use of DL in Figure 1 many more researchers working in different fields may find their way to generate different DL approaches.

Therefore, this review aims to guide the researchers who want to use the DL approaches in their studies. In the literature, although there are some review studies using DL on the subjects such as autonomous

vehicles[65], medical image processing [29, 66–69], big data [70, 71], natural language processing (NLP) [10, 19, 72], signature recognition [6, 32, 73], and many other fields [64, 69, 74–87], there is no comprehensive review covering all these fields. In this paper, we provide a review of DL which covers many challenging areas such NLP, medical image processing, mig data, hand-written character recognition, signature recognition, audio and video processing, and autonomous vehicles. This study investigates more than 230 articles related to DL applications in the selected areas. Main contributions of this study are summarized below:

- to review of DL methods on many challenging areas.
- to implement DL approaches in the most popular areas.
- to introduce the challenging problems in popular topics.

The manuscript is organized as follows. In Section 2, comprehensive information on the DL applications is presented for the most popular areas. In Section 3, the important DL issues are proposed. Finally, the conclusion is given in Section 4.

2. DEEP LEARNING RESEARCH AREAS (DERİN ÖĞRENME ARAŞTIRMA ALANLARI)

In the last decade, with the development of ANN, many researchers have tried to develop further studies using DL methods. They achieved the-state-of-art results in many research areas such as autonomous vehicles, medical image processing, handwritten character recognition, and natural language processing. In this section, DL studies are investigated in popular areas to illuminate the paths of researchers working in DL. We visualized the researches in tables according to study fields, titles, methods, used tests, and training databases in order to make it easier for researchers to analyze. In literature, there are many study fields using DL approaches achieved high success, but in this review, seven research areas such as Autonomous Vehicles, Natural Language Processing, Handwritten Character Recognition, Signature Verification, Voice and Video Processing, Medical Image Processing, and Big Data are selected.

2.1. Autonomous Vehicles

In this section, the recent DL studies are investigated in autonomous vehicles. With the

development of DL, technologies in autonomous vehicles have seen significant advances in recent years. DL become indispensable in application of autonomous vehicles. In the literature, autonomous vehicle studies for the aerial, underwater, and lander vehicles have focused on vehicle detection, obstacle detection, control strategies, autonomous driving, traffic and collision prediction. Vehicle and obstacle detection is one of the most important part of autonomous driving. Hence, so many researchers have studied to overcome these problems. To solve this problem, DL algorithms generally consist of two steps: region proposal and region classification. Shi et al. [88] proposed a Fast Recurrent-CNN to detect vehicle. They tested the algorithm on the public KITTI data set, self-collected BUU-T2Y data set, and mixed data set. In another study, Du et al. [89] proposed a DL framework for car detection, which fuses LIDAR and camera. Soin and Chahande [90] proposed moving vehicles detection on-road for driverless car assistance systems based on R-CNN. They claimed that the method was achieved 100% accuracy with respect to detection accuracy on the CIFAR-10 dataset. Obstacle detection and classification on-road is one of the key tasks in the perception system of self-driving vehicles. Nguyen et al. [91] used the DL for the classification of obstacles that were detected by using adaptive U-V disparity. Prabhakar et al. [36] developed a DL system that is region-based CNN trained with the PASCAL VOC image dataset for the detection and classification of on-road obstacles such as vehicles, pedestrians and animals. A fully convolutional network is proposed to predict pixel-wise semantic labeling of on-road unexpected obstacles such as lost cargo by Ramos et al. [24]. Deepika and Variyar [92] used the SegNet encoder-decoder architecture for pixel-wise semantic segmentation of the video frame followed by an obstacle detection algorithm. Dairi et al. [93] designed a hybrid encoder that integrates deep Boltzmann machines (DBM) and auto-encoders (AE) for obstacle detection. They combined the unsupervised hybrid model with the one-class support vector machines (OCSVM) to visually monitor an urban scene. In literature, collision and traffic predictions are other challenging tasks for self-driving cars. Vehicle taillights detection is an important topic in collision avoidance and autonomous vehicles. For this reason, Vancea et al. [31] addressed this problem by using FCN based on VGG16 architecture for vehicle rear light segmentation. Chen et al. [94] proposed a BP neural network optimized by GA to estimate the probability of collision. In this neural network, it has taken into account the effects of many factors,

including drivers, environment, vehicles, and path. In the study, the selection of the thresholds in the BP neural network is optimized by the GA. Liu et al. [95] proposed a decentralized multi-agent collision avoidance method based on deep reinforcement learning. Obtaining accurate and timely traffic flow information is essential for the success of intelligent transportation systems. Hence, traffic flow prediction is one of the major tasks of self-driving cars. In the literature, there are many studies to address this task by using DL algorithms such as DBN, SAE, CNN, RCNN, RNN. DBN based architecture was used widely for traffic flow prediction. Huang et al. [96] developed a deep architecture that consists of two parts, i.e., DBN for unsupervised feature learning at the bottom and a multitask regression layer (MTL) at the top. For short-term traffic speed information prediction, Jia et al. [97] are proposed a DBN based model that was trained in a greedy unsupervised method, and fine-tuned by labeled data. In the other study, Jin et al. [98] used DBN and LSTM to conduct short-term traffic speed prediction with the consideration of rainfall impact as a non-traffic input. They reported that the R-LSTM has the best accuracy value according to R-DBN, R-BPNN, and R-ARIMA. Koesdwiady et al. [99] proposed a prediction architecture that incorporates DBN and data fusion to derive traffic flow prediction by using traffic flow history and weather data. Wang et al. [100] proposed an error-feedback recurrent CNN (eRCNN) model for the continuous traffic speed prediction. In another study, Liu, Polson and their teams an LSTM based RNN architecture was used to improve traffic flow prediction accuracy. Du et al. [101] implemented an LSTM based regression model to predict the historical traffic data by revealing the pattern of traffic volume. They claimed that the LSTM model has the ability to remember the former state of time point and give a more accurate estimation of the next time points. Liu et al. [102] proposed a combination of Convolution and LSTM for traffic flow prediction. First, they used a Conv-LSTM module to extract the spatial-temporal feature of traffic flow, and then a Bi-LSTM module to obtain the period feature of traffic flow. For traffic flows prediction, Polson et al. [103] developed a DL architecture that combines a linear model fitted using regularization and a sequence of layers. In the other studies related to traffic flow prediction, Lv, Duan and their colleagues used SAE architecture. Lv et al. [104] explored a DL approach to predict traffic flow at different time ranges such as 15-min flow, 30-min flow, 45-min flow, and 60-min flow. They reported the best SAE architecture for every time ranges. Duan et al. [105] evaluated the performance of the SAE

model for traffic flow prediction at daytime and nighttime. They also reported the hyper parameters of the SAE model at different times. In another study, Stacked Denoising Auto-Encoder (SDAR) based DL architecture was used for outdoor vehicle tracking by Xin et al. [106]. They trained the SDAE and then added k-sparse constraint to SDAE (kSSDAE) for classification. Hadsell et al. [107] described a self-supervised DL based DBN approach for the classification complex terrain at distances up to the horizon by identifying trees, paths, man-made obstacles, and ground. Object recognition and pedestrian detection algorithm was developed with DL for autonomous driving by Ucar et al. [26]. In this study, feature extraction was carried out by using both CNN and Bag of Visual Words (BOW) approach. Al-Qizwini et al. [22] introduced the Direct Perception DL algorithm by comparing different CNN architecture for autonomous driving. According to obtained results, GoogLeNet is the best performing architecture for autonomous driving in road feature extraction. Xia and Rausch separately, introduced a different DL model to control autonomous vehicles. Rausch et al. [108] used CNN architecture in the application vehicles, and control strategy of Xia [109] is deep Q-learning with filtered experiences (DQFE) model, which combined the classic neural fitted Q-iteration and DL technique to obtain a control strategy with less resource consumption.

In literature, DL has also used so much in aerial and underwater autonomous vehicles. For the control of autonomous aerial vehicles, Zhang et al. [110] trained a DL algorithm by generated data according to MPC-guided policy. They claimed that the algorithm was able to learn complex policies, such as high-speed obstacle avoidance, using raw sensor inputs, and low-level rotor command outputs. For docking of autonomous underwater vehicle, Yahya and Arshad [111] proposed a R-CNN based DL application which detected the markers on a docking station. Yu et al. [112] tackled the control problem of trajectory tracking of Autonomous Underwater Vehicles (AUVs) based on the DL algorithm. They used a combination of two different DL algorithms to control AUVs. The first algorithm was used to select the action, and then the second was used to evaluate the accuracy of the selected action.

2.2. Natural Language Processing

Big data has grown rapidly over the last few decades. Therefore, semantic data extraction from raw data has become important necessity. Natural Language Processing (NLP) plays a major role in

acquiring relevant information of importance to business and intelligence. Studies in the field of NLP has great importance for this purpose. With the development of DL and artificial intelligence (AI), many researchers used neural networks in NLP tasks such as call-routing classification, spoken language understanding (SLU) systems, and semantic data extraction.

Cambria and White [19] introduced a review on recent developments in NLP research to look at the past, present, and future of NLP technology. It is a helpful study to learn about research field of NLP. Application of DBNs to natural language call routing was proposed by Sarikaya et al. [113]. They compared between DBNs and other methods such as Boosting, MaxEnt, SVM. It was reported that DBNs produce better classification results than other methods even though it currently uses an impoverished representation of the input. In other study, Sarikaya et al. [40] applied DBNs to overcome natural language understanding problem. In a PhD thesis, Richard Socher [21] it was studied RNN to solve multiple language tasks involving word and sentence-level predictions of both continuous and discrete nature [21]. In this study, different RNN architectures such as SU-RNN, MV-RNN, C-RNN, DT-RNN, and RNTN were used and compared the performance of them.

The most important limitation of NLP research on DL is insufficient data. For this reason, Galinsky et al. [41] proposed data augmentation approaches that are replacing words with their synonyms, reshuffling the words and adding new adjectives. After augmentation, they used CNN for addressing the sentiment analysis problem on Russian language. Natural Language Generation (NLG) systems have not been generalized across domains, and this is heavily restricting their usability beyond a single application. Therefore, Dethlefs [114] proposed a LSTM-Encoder-Decoder based RNN architecture for addressing the domain adaptation problem of NLG. In the study, Two LSTM model was used. The first one learns a latent representation of a semantic input, and the second one learns to decode it to a sequence of words. It is claimed that the learned representations can be transferred across domains and leveraged effectively to improve training on new unseen domains. Zhuang et al. [42] used a CNN model with chunk-max pooling by the leverage Chinese stroke for learning the continuous representation of Chinese character [42]. They built an NLP service for Chinese text classification for use in future studies. In another study, Zhang et al. [27] proposed a novel CNN and RNN based DL method which generates natural language descriptions of

remote sensing images. In this study, CNN detected the main objects of the remote sensing images, and RNN generated the natural language descriptions of the objects. Al-Ayyoub, Sharma and their colleagues published a survey paper about DL studies in NLP. Sharma and Kaushik [72] have emphasized the importance of DL in NLP studies, and also recommended reinforcement learning as an extension to neural networking which is widely used in gaming for NLP. The survey of Al-Ayyoub et al. [10] was about DL studies in Arabic NLP. In this survey, they aimed to serve a guide for the young and growing ANLP community to help bridge the huge gap between ANLP literature and richer and more mature English NLP literature.

2.3. Handwritten Character Recognition

Character recognition is still one of the most challenging fields for researchers. Especially, the handwritten character recognition is an arduous task since the character usually has various appearances according to a different writer, writing style and noise. Handwriting recognition is composed of two types such online (stroke trajectory-based) and offline (image.-based). Due to the inadequate number of features that characterize, the offline handwriting character recognition problem is more difficult than online. With the great developments on the DL, many researchers try to tackle this problem by using DL based methods.

One of the first DL-based studies on handwriting character recognition is hand-written digit recognition published by Y. LeCun et al. [115]. They proposed a fully connected multilayer neural network called LeNet for hand-written digit recognition. In 1995, Bengio, LeCun et al. [116] introduce a new Artificial Neural Network (ANN) and Hidden Markov Model (HMM) based hybrid approach for online recognition of handwritten words. They used CNN for recognition and used HMM to tag every word with an error score. They globally trained an entire system to minimize word-level errors. The different HMM and ANN models were used by Espana-Boquera et al. [117] for recognizing unconstrained offline handwritten texts. For the handwritten Arabic character recognition task, Al Jawfi [5] designed a LeNet based network that consists of two stages. The first stage is to recognize the main shape of the character, and the second stage is to determinate the dots. Lamsa-at and Horata [43] compared the recognition ability of Deep-Learning Feedforward-Backpropagation Neural Network (DFBNN) and Extreme Learning Machine (ELM) on three different data sets of the handwritten character.

They compared the classifiers in two categories: non-extracted and extracted features according to Histograms of Oriented Gradients (HOG). They reported that features extracted by HOG can improve recognition rates of both DFBNN and ELM. According to the obtained results, DFBNN provided higher recognition rates than ELM. Yin et al. [118] presented a paper to define the tasks, methods, and results of the Chinese handwriting recognition competition held at the 12th International ICDAR13. They reported that ten groups joined with 27 systems for achieving five tasks: classification on extracted feature data (Task 1), offline isolated character recognition (Task 2), online isolated character recognition (Task 3), offline handwritten text recognition (Task 4), and online handwritten text recognition (Task 5). In 2014 Szal et al. [119] proposed a DBN based DL architecture to overcome Bangla handwritten character recognition. They applied the algorithm in two stages: supervised and unsupervised. They tested the algorithm on the dataset of Bangla numerals and other basic characters collected from the Indian Statistical Institute. As a result, they reported that the DBN approach can achieve a higher recognition rate even though no handcrafted feature was used. Pham, Doetsch and their colleagues used RNN to overcome the problems of offline handwriting recognition. Pham et al. [120] used LSTMs on three different handwriting datasets which contain French, English and Arabic text. They used the dropout technique with LSTM and reported that the performance of the word recognition networks could be improved if dropout is applied at multiple LSTM layers. Doetsch et al. [121] introduced a trainable modification of the activation functions inside the gating units of an LSTM-RNN in offline handwriting recognition. They showed that the frame-wise recognition performance can be improved using the proposed gate scaling method.

In the literature, CNN is the more preferred than different types of DL architectures such as DBN, RNN, etc. for the hand-written character recognition tasks. In 2011, to improve handwritten character recognition error rates, Cireşan et al. [122] proposed a group of CNN classifiers for each error on various parts of the training set. They reported 0.27% error rates for seven deep CNNs trained with graphics cards, narrowing the gap to human performance. In 2012, Multi-column DNN (MCDNN) was the first to achieve human competitive performance on the famous MNIST handwritten digit recognition task. MCDNN was designed with combining several DNN columns by Cireşan et al. [30]. They also show how MCDNN further decreases the error rate by 30-40%.

In 2015, another study was published by Cireşan and Meier [123] for the character classification. In this study, they used the same MCDNN model in their previous paper for offline handwritten Chinese character classification. It was claimed that MCDNN could classify handwritten Chinese character dataset with almost human performance. CNN based handwritten Hangul recognition application was presented by Kim and Xie [45]. They also improved the training speed and recognition performance through GPU-based parallel processing and elastic distortion. They reported that the recognition rate on two public Hangul image databases, 95.26 % on SERI95a and 92.92 % on PE92. Zheng et al. [46] used CNN to train the recognition model of handwritten character images which are extracted from the original sunspot drawings. They tested the method on sunspot drawings provided by the Chinese Academy Yunnan Observatory. Wu et al. [15] proposed a hand-writing recognition method based on Relaxation CNN (R-CNN) and Alternately Trained Relaxation CNN (ATR-CNN). The difference between R-CNN and CNN is whether neurons within a feature map share the same convolutional kernel. In the traditional CNN, neurons share the same weight matrix, but in the R-CNN, neurons use different weight matrices. ATR-CNN was used for the training procedure because R-CNN increased the total number of parameters sharply. Chen et al. [124] designed a multi-supervised training model. The separately trained CNN models can be seen as different classifiers, though the training dataset is the same. Hence, they used the multi model voting method to decide for the final recognition result. They reported that the performed method was better than humans in handwritten digit (MNIST) and Chinese character (CASIA) recognition. A new training method to enhance Deep Convolutional Neural Networks (DCNN) in handwritten Chinese character recognition was proposed by Yang et al. [125]. For better performance, they used various types of domain knowledge, including deformation transformation, non-linear normalization, imaginary stroke maps, 8-directional feature maps, and path signature feature maps. They also used a new training method called DropSample for avoiding interference due to noisy samples. In 2017, Xiao et al. [44] proposed a new method for building fast and compact CNN model for large scale handwritten Chinese character recognition (HCCR). In the model, they used a Global Supervised Low-rank Expansion (GSLRE) method and an Adaptive Drop-weight (ADW) technique to solve the problems of speed and storage capacity. They claimed that compared with the state-of-the-art CNN model for HCCR, their model is

approximately 30 times faster, yet 10 times more cost-efficient. Huang et al. [126] designed the inception font network (IFN) model which composed of two additional CNN structure elements for Chinese font recognition (CFR). In the model, they used the DropRegion method which randomly removes several elastically sized regions from the characters instead of Dropout. By using DropRegion, they aimed at data augmentation, thus improved the generalized applicability of the CNN-based network model and prevented model overfitting. CNN based DL algorithms are widely used in other languages as much as Chinese for achieving handwritten character recognition tasks. Boufenar et al. [127] designed a new DCNN architecture for off-line handwritten Arabic character recognition. The architecture was built around the AlexNet CNN and uses the transfer learning paradigm with and without preprocessing step. The Genetic Algorithm (GA) assisted CNN model was used to cope with the Devanagari handwritten numeral recognition problem. In this hybrid model, Trivedi et al. [128] proposed to use the genetic algorithm on the training phase of CNN to cope with the expensive computational cost of training. They reported that the GA implementation improves the accuracy of CNN. Soomro et al. [129] published a comparative study of hand-written character recognition. They compared the two-state of art CNN architectures such as AlexNet and Google Net. The architectures were tested on Chars74K and local datasets, which include English hand-written characters. As a result, they reported that the GoogleNet architecture outperforms the Caffe

AlexNet architecture on varying test cases, on the application of handwritten character recognition. A combination of models of CNN and sequence to sequence (Seq2Seq) was used by Sueiras et al. [130] for offline hand-written Latin character recognition. They applied the CNN over a sequence of image patches obtained from the character image using a horizontal sliding window, and then they used the feature which obtained from CNN in the Seq2Seq network for character recognition.

2.4. Signature Verification

Signatures are the most commonly used biometric authentication technique for hundreds of years. The signatures have signed on paper with the pen for hundreds of years. Nowadays, with the development of technology, also online signature can be used via electronic devices such as tablets and computers. Therefore, today there are two types of signatures such as offline (static) and online (dynamic).

In offline signature verification, a signature that is used as input data for process is available on a document and it is scanned to obtain its digital image representation. The offline signature has very few features for classification. The online signature is signed by special hardware, such as a digitizing tablet or a pressure-sensitive pen, to record the pen movements during writing. Therefore, it has many features for classification process [28, 49, 131, 132]. Table 6 presents a comparison of common features for offline and online signatures [133–135].

Table 1 Comparison of common features for offline and online signatures

List of common features			
Offline Signatures		Online Signatures	
f1	Aspect Ratio	f1 -f13	All common features of offline signature
f2	Out Ratio	f14	Coordinate $x(t)$
f3	Vertical and Horizontal Center	f15	Coordinate $y(t)$
f4	Base line shift	f16	Pressure $p(t)$
f5	Moments Features	f17	Time stamp
f6	Vertical and Horizontal Projection Peaks	f18	Absolute position, $r(t) = \sqrt{x^2(t) - y^2(t)}$
f7	Number of Edge and Cross Points	f19	Velocity in $x, v_x(t)$
f8	Number of closed loops	f20	Velocity in $y, v_y(t)$
f9	Maximum Vertical and Horizontal Projection	f21	Absolute velocity, $v(t) = v_x^2(t) - v_y^2(t) $
f10	Grid Features	f22	Velocity of $v(t), v_r(t)$
f11	Texture Features	f23	Acceleration in $x, a(t)$
f12	Height	f24	Acceleration in $y, a_y(t)$

f13	Width	f25	Absolute acceleration, $a(t) = a_x^2(t) - a_y^2(t) $
		f26	Pen down Time $P_d(t)$
		f27	Pen up Time $P_u(t)$

Signatures, which legally impose financial and moral liabilities, are an authentication technique that is still widely used today especially in legal documents, banking, and commercial transactions. Hence, signature verification/recognition is one of the most important fields for researchers. To tackle this problem, signature competitions which is named as SigComp2011 [136], 4NSigComp2012 [137] and SigWiComp2013 [138] were organized. In addition, many researchers presented [28, 32, 73, 139] the-state-of-the-art studies and surveys about signature verification and recognition systems. Researchers have tried to solve this problem by using algorithms such as Support Vector Machines (SVM), Dynamic Time Wrap (DTW), Principle Component Analysis (PCA), Fuzzy Systems Methods, Probabilistic Neural Network (PNN), Deep Multitask Metric Learning (DMML). In literature, the best results have been taken by DL algorithms. Along with the promising developments of DL, researchers have begun to develop new DL-based algorithms to address the signature verification and recognition.

In 2011, for offline signature recognition, Ribeiro et al. [6] proposed a two-step hybrid classifier system, the first, identifies the owners of the signatures while, the second, determine its authenticity. They compared performances of the fourteen-different feature extraction algorithms in the study. They also extracted a high-representation of the signature images through multi-layers in a deep hierarchical structure and pointed out that the DL algorithm will perform good success in this area together with the development of the GPU technology. In the literature, it has clearly seen that many researchers using multi-model algorithms that are composed of Writer-Independent (WI) and Writer-Dependent (WD) for signature verification/recognition [2, 35, 140, 141]. Zhang et al. [2] proposed a new Deep Convolutional Generative Adversarial Network (DCGANs) model for offline signature verification with multi-phase architecture which works together with a hybrid Writer-Independent-Writer-Dependent classifier. In the study, they reported that the method is promising, even though their system does not achieve performance close to the-state-of-the-art for GPDS, because it combines conveniences and robustness. Soleimani et al. [140] developed a Deep Multitask

Metric Learning (DMML) method for offline signature verification. In the method, they used the knowledge extracted from the similarities and dissimilarities between the genuine and forged samples of other classes for achieving better results. They used a mixed approach of WI and WD to obtain the knowledge data. As a result, they reported that DMML outperforms SVM and DDML whether WI or WD. Hafemann et al. [35] used CNN with WI and WD model for signature verification. In the model, they used the WI method for feature extraction in the learning phase and then used the WD method by using the extracted features for classification. They reported the method is promising and achieving performance close to the state-of-the-art. In another study, Hafemann et al. [141] compared two different CNN architecture that consists of AlexNet and VGG for increasing the success of signature verification. They used the architectures with different parameters (number of hidden, pooling and fully connected layers, filter size, number of stride and padding) for fine-tuning. They also compared these architectures with dropout and without dropout. As a result, they reported that the best performing models were the AlexNet architecture (with 8 trainable layers) and the AlexNetreduced (with 6 trainable layers) when using the features to training linear SVMs and SVMs with RBF kernel, respectively. They also reported that features learned in a writer-independent format can be very effective and writer-dependent classifiers trained with these features can perform very well even with the limited number of samples per user (e.g. 5 samples) and linear classifiers. In 2017 Hafemann et al. [142] presented different formulations based on the CNN model which was shown above by them, for learning representations for offline signature verification. They claimed that Equal Error Rate (EER) 1.72% have been obtained in the GPDS-160 dataset. Tayeb et al. [143] proposed a CNN based signature verification application. They tested the application on the SIGCOMP 2011 dataset and reported that the application was able to correctly detect the validity of the inputted signature approximately 83 percent.

In the literature, it is seen that almost all studies are about the offline signature verification by the CNN methods. Similarly, DL algorithms are more

successful on verification of online signatures, since online signatures have high and robust features. Thus, this field is promising study area for researchers. In 2018, Tolosana et al. [144] proposed multiple RNN systems for online handwritten signature verification. They used both Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) systems with a Siamese architecture. They reported the proposed recurrent Siamese networks have outperformed the state-of-the-art online signature verification systems.

We noticed that there is not comprehensive literature view about signatures verification and signatures recognition studies based on DL, although there are two papers introduced by Sanmorino, Hafemann and their teams. This proves the importance of review will contribute greatly to the researchers. In 2012, Sanmorino et al. [73] published a survey on handwriting signature verification in three main sections which composed of data acquisition method, preprocessing, and verification technique (Artificial neural networks-ANN, hidden Markov model- HMM, and support vector machine-SVM). Although the survey does not contain DL methods, they reported that ANN had the best performance and the easiest implementation. Another literature review paper on handwriting signature verification was published by Hafemann et al. [32] in 2017. This review consists of five sections such as problem statement, datasets, preprocessing, feature extraction and training model. In the training model section, they discussed performance of DL models in addition to ANN. As a result, they reported that in the future works on signature verification will continue increasingly and in particular learning representations from signature images by using DL methods will increase the performance.

In our opinion, signature verification studies will gain more importance in the future together with the serious development of DL technologies. DL methods will show especially out performs on feature extraction and data augmentation parts. So, this will make better performance on signature verification and recognition.

2.5. Voice and Video Processing

DL has used frequently in video and audio processing areas as well as in many image processing areas. For this reason, DL was used in audio and video applications such as audio-visual voice activity detection, phonological features detection and classification, speaker emotion recognition, speech enhancement and recognition, noise separation,

human action recognition, emotion, expression, and object recognition in videos.

DL studies conducted in the field of audio processing are mostly studied in speech recognition applications. For achieving this process, it was seen that DL architectures which are mostly used are Recurrent Neural Network (RNN), Deep Denoising Autoencoders (DDAE), Deep Belief Networks (DBN). Wei et al. [145] built an articulatory model using hybrid deep neural networks (DNNs) and hidden Markov models (HMMs) methods to map both the acoustic and voicing features into articulatory spaces for Mandarin speech recognition. Tamura et al. [39] developed a DNN application for Audio-Visual Speech Recognition (AVSR) using deep bottleneck features and high-performance lipreading. They reported that the application achieved 73.66% lipreading accuracy and 90% AVSR accuracy on average in noisy environments. In another study, Wu et al. [146] used DNN for speech enhancement and speech recognition. They adopted two techniques, namely DNN-based regression to enhance reverberant and noisy speech, followed by DNN-based multi-condition training for recognition. Becerra et al. [37] introduced two new variations of the frame-level cost function for training a DNN in order to achieve better speech recognition. Due to the profound differences between acoustic characteristics of neutral and whispered speech, the performance of traditional automatic Speech Recognition (ASR) systems trained on neutral speech degrades significantly when the whisper is applied. To tackle this problem, Grozdić et al. [38] proposed the new system based on Deep Denoising Autoencoders (DDAE). Speech enhancement is one of the most important preprocessing step. Therefore, many researchers [147, 146, 148] work on speech enhancement applications to enhance the background noise and interfering speech in a speaker dependent. Another challenging task in audio processing based on DL is Speech Emotion Recognition (SER). Fayek, Manneppall and their teams [149, 150] proposed to the DL methods to overcome this challenging task. Fayek et al. [149] focused on a frame-based formulation for SER, aiming to achieve a system with a simple pipeline and low latency by modeling intra-utterance emotion dynamics. They proposed the SER system to empirically explore feed-forward and RNN architectures and their variants. Manneppall et al. [150] developed a new novel algorithm, called Adaptive Fractional DBN (AFDBN) for robust SER. They claimed that the proposed AFDBN achieved the higher accuracy as 99.17% and 97.74% on Berlin and Telugu database which ensured better emotion

recognition performance. Zhang et al. [151] proposed a DL framework to extract and recognize emotional information from respiration (RSP) signals. The DL framework includes a sparse auto-encoder (SAE) to extract emotion-related features and two logistic regression with one for arousal classification and the other for valence classification. Bhowmik et al. [152] designed deep-structured models which are pre-trained by stacked autoencoder for detection and classification of the phonological feature of Bengali continuous speech. Another DL based speech classification study proposed by Deng et al. [153]. Achanta and Gangashetty [154] compared deep Elman RNN (ERNN) against deep gated RNNs for statistical parametric speech synthesis (SPSS). They claimed that Deep Elman RNNs are better suited for acoustic modeling in SPSS when compared to DNNs and perform competitively to deep SLSTMs, GRUs, and LSTMs. They remained Elman RNN based duration model is better than the DNN based counterpart. Ariav et al. [155] performed a deep architecture based on RNN for audio-visual voice activity detection.

Video processing has been in the center of computer vision community for a long time. Therefore, video processing is one of the most challenging area for researchers. In this area, there are many important problems as object recognition, face recognition, emotion recognition, and action recognition in video. Learning by deep CNNs has shown outstanding effectiveness in a variety of image-based classification tasks. Deep CNNs have also begun to be used to tackle video problems, along with the understanding that the computational efficiency of CNNs can be used in real-time applications. Object recognition is one of the most studied areas in the literature. Roman et al. [1] proposed a general DL framework for real-time object recognition in egocentric videos guided by visual salience to help grasping actions for neuro-prostheses. They reported that computational time at generalization is lower than a time of visual fixation. CNN knowledge transfer framework was described by Xin et al. [7] for underwater object recognition. The developed system can work for real-time live object recognition from underwater videos. They reported that discriminative features can be extracted from relatively low contrast images via the system. For security, detection of a hazardous substance as a bomb, gun, knife etc. is crucial problem. Real time live object recognition from videos is also very important for security. To tackle this problem, Olmos et al. [156] presented a novel automatic pistol detection system in videos appropriate for both, surveillance and control purposes. They pointed out that the most promising

results have been obtained with Faster R-CNN based model which has a VGG-16 classifier. It is seen that face and emotion recognition are commonly studied in other areas for object recognition from video. It is seen that face recognition and emotion recognition are other challenging tasks in object recognition from video. CNN based Face Recognition (FR) from the video which has only a single sample (i.e., still ROI) per person (SSPP) is proposed by Parchami et al. [157]. The developed Canonical Face Representation CNN (CFR-CNN) method is based on a deep supervised autoencoder that can represent the divergence between the source (still ROI) and target (video ROI) domains encountered in still-to-video FR. Rao et al. [158] proposed a local temporal representation for video face recognition by combining the CNN feature with recurrent layers with locality constraints to make better use of temporal information. They presented a new attention-aware deep reinforcement learning (ADRL) method for video face recognition. Another face recognition study based on DL for spoof face detection published by Akbulut et al. [159]. They used two different DL models, namely (Local Receptive Field) LRF-ELM and CNN. The LRF-ELM model contains a convolution layer, a pooling layer, and a fully connected layer. They reported that the LRF-ELM method yielded better results against CNN. Studies of DL based emotion recognition from the video is presented by many researchers. A multimodal approach for video-based emotion recognition in the wild was proposed by Kaya et al. [4], DL framework for achieving emotion recognition from low bit rate video was published by Cheng et al. [160], in another study Gupta et al. [161] presented an architecture based on DL for expression recognition in videos, which are invariant to local scaling. Classification of human actions is very challenging and important in many video-based applications. The challenge in human action analysis is to properly characterize spatial-temporal information and facilitate subsequent comparison/recognition tasks. To address this problem, some approaches would use the body movements for building various action syntactic primitives to represent and recognize actions. The ConvNets largely promoted the development of action recognition in video. A new deeper two-stream ConvNets for action recognition was proposed by Han et al. [162]. Another ConvNets based video classification was proposed to recognize human actions based on motion sequence information in RGB-D video using DL by Ijjina and Chalavadi [163]. Alghyaline et al. [18] proposed a novel video representation to improve human action recognition

based on symmelets, IDT and DL. Parisi et al. [164] introduce a deep neural architecture for the lifelong learning of the body. The architecture consists of a series of hierarchically arranged self-organizing neural networks for learning action representations from pose–motion features. Ma et al. [165] proposed Region-sequence based six-stream CNN features for fine-grained human action recognition in videos. In the method, they divided the human body into six regions and used six-stream CNN for human action recognition. Herath et al. [82] provide a detailed review study that also includes the impact of Deep-nets on human action recognition over the past decade. The human targets need to be localized first before actual analysis is conducted to analyze her/his behavior or activities. For this reason, a novel deep multi-channel residual networks-based metric learning method is introduced for the first time to realize moving human localization in video surveillance by Huang et al. [166]. Video compressive sensing (CS) aims at increasing the temporal resolution of a sensor. In another study, Iliadis et al. [167] present the first deep learning architecture for temporal video CS reconstruction approach, based on fully connected neural networks, which learns to map directly temporal CS measurements to video frames. Li and Wang [12] proposed a novel deep residual CNN to predict a high-resolution image using multiple motion-compensated observations. In the study, they used an optical flow algorithm for motion estimation and motion compensation.

2.6. Medical Image Processing

In recent years, the developments of deep learning algorithms have had a major impact on medical image analysis studies. Using the DL algorithms on medical images, very successful results have been obtained in the fields such as classification, object detection, segmentation, registration, and retrieval. Object classification which is usually focused on the classification of a small (previously identified) part of the medical image into two or more classes is not enough for accurate classification. In the literature, we have seen that many researchers have used multi-model CNN architectures to tackle this problem. For skin lesion classification, Kawahara and Hamarneh [16] presented a novel CNN architecture that is composed of multiple tracts, which is each tract analyzes the image at a different resolution simultaneously. For skin lesion classification designed to learn based on information from multiple image resolutions while leveraging pre-trained CNNs. While traditional CNNs are generally trained on a

single resolution image. Setio et al. [14], presented a CAD system for pulmonary nodule detection in CT scans based on multi-view CNNs. They showed that ConvNets-CAD is highly suited to be used as a decision aid in a lung cancer screening scenario by promising results and low computation time make. Mohsen et al. [168] used a Deep Neural Network classifier for classifying brain MRIs into 4 classes e.g. normal, glioblastoma, sarcoma, and metastatic bronchogenic carcinoma tumors. The classifier was combined with the discrete wavelet transform (DWT), the powerful feature extraction tool, and principal components analysis (PCA). According to the results, the performance of the classifier was high. Gao et al. [169] proposed a fused CNN architecture to featuring both automatic and selective DL networks for the classification of echocardiography videos of eight viewpoint classes. They point out that CNN architecture of two-strand networks performs the best with classification results up to 92.1% of accuracy.

The detection of objects of interest or lesions in medical images is a key part of diagnosis and is one of the most labor-intensive for clinicians. So, this is a very popular research area in the literature. Kooi et al. [170] provided a head-to-head comparison between the latest technology in the mammography CAD system and the CNN to obtain a system that can read mammography independently. As a result, they showed that the CNN model trained on a large data set of mammographic lesions outperforms. Another efficient DL model was developed by Saha et al. [171] for mitosis detection using breast histopathology images. Wang and Yang [172] developed a context-sensitive deep neural network (DNN), aimed to take into account both the local image features of a microcalcification (MC) and its surrounding tissue background, for MC detection in mammograms. As a result, they claimed that the proposed approach could achieve significantly higher FROC curves when compared to two MC-based detectors. The earliest sign of diabetic retinopathy (DR) which is one of the leading causes of preventable blindness in the world is red lesions. Orlando et al. [173] proposed a novel method for red lesion detection based on combining both deep learned and domain knowledge. In the study, they have used CNN to learn features and then augmented by incorporating hand-crafted features. Finally, they used a random forest classifier to identify true lesion candidates.

Segmentation studies on such as brain tumors, vessels, sclerosis lesion are so important tasks in medical image processing. Thus, one of the most studied fields on the medical image is segmentation. In most of the study about medical segmentation, it was used CNNs. Havaei et al. [174] presented a fully

automatic brain tumor segmentation method based on Deep Neural Networks (DNNs) which are tailored to glioblastomas (both low and high grade) pictured in MR images. The CNN exploits both local features as well as more global contextual features simultaneously, and also uses a final layer that is a convolutional implementation of a fully connected layer which allows a 40-fold speed up. Pereira et al. [8] proposed an automatic segmentation method based on CNN to augment the efficiency of brain tumor segmentation in MRI images. The study has been validated in the Brain Tumor Segmentation Challenge 2013 database (BRATS 2013), and has also been second in BRATS 2015. Işın et al. [29] presented a review article containing more detailed studies on brain tumor segmentation using DL models up to 2015. Milletari et al. [33] proposed Hough-CNN that is a novel segmentation approach based on a voting strategy for the segmentation of deep brain regions in MRI and ultrasound. In the study, 3D CNN which is based on the Caffe framework is compared to the more common 2D convolution, as well as to a recent 2.5D approach. They claimed that Hough-CNN delivers results outperforms to methods based on voxel-wise, semantic classification. Hussain et al. [175] proposed, five types of nexus architectures are proposed namely: linear, two-path, two-path linear, inception, and inception linear for segmentation of glioma tumors in brain. In the study, a patch-based approach along with an inception module is used for training the deep network by extracting two co-centric patches of different sizes from the input images. The newest study that has been built upon the DL model about brain tumor segmentation up to now was developed by integrating Fully CNN (FCNNs) and Conditional Random Fields (CRFs) in a unified framework to obtain segmentation results with the appearance and spatial consistency. The model which was developed Zhao et al. developed a model [176] to evaluate the image dataset provided by the Multimodal Brain Tumor Image Segmentation Challenge (BRATS) 2013, BRATS 2015, and BRATS 2016. Apart from brain tumor segmentation, vessel segmentation and sclerosis lesion segmentation studies are also common in the field of medical imaging. Valverde et al. [177] present a novel automated method based on a cascade of two 3D patch-wise CNN for White Matter (WM) lesion segmentation of Multiple Sclerosis (MS) patient images. Birenbaum and Greenspan [178] proposed a CNN based method for multiple sclerosis lesion segmentation. They used longitudinal data to improve segmentation accuracy significantly. Nasr-Esfahani et al. [23] proposed the DL method using CNN for detecting vessel regions in angiography

images. Barkana et al. [179] have evaluated the performance of descriptive statistical features in retinal vessel segmentation by using fuzzy logic, an artificial neural network classifier (ANN), a support vector machine (SVM), and classifier fusion. We figured out that there is not so many studies about retinal vessel segmentation by using DL methods in the literature.

Image registration is the process of geometrically align two images or the process of overlaying images (two or more) of the same scene taken at different times, from different viewpoints, and/or by sensors [74]. Registration is very important for medical image analysis as well as for other image processing areas. Miao et al. [180] present a CNN regression approach which is used on the appearances of the digitally reconstructed radiograph and X-ray images to tackle slow computation and small capture range which are the two major limitations of existing intensity-based 2-D/3-D registration technology. Yang et al. [181] proposed a deep regression model to predict 3D deformation parameters using the MR image in a time-efficient manner. Jia et al. [182] proposed a registration-based coarse segmentation to the pre-processed prostate MRI images to get the potential boundary region by using DNN which is based on a VGG-19 network model.

It is seen that medical image processing fields such as image retrieval, generation, and enhancement are also studied frequently. For medical images, using with and without class predictions, content-based image retrieval approach proposed by Qayyum et al. [183]. In another study, parallel deep solutions for image retrieval from imbalanced medical imaging archives were proposed by Khatami et al. [20]. For medical image analytics, a comprehensive review which also includes DL models was published for large-scale retrieval by Li et al. [67].

In the literature, it is seen that DL algorithms have a wide usage area in medical image processing. In addition to the topics mentioned above, cancer, protein subcellular localization, and toxicity prediction topics are also studied. In literature, multi-deep learning models are used to make a robust cancer prediction. Yuan et al. [184] introduced a regularized ensemble framework of DL to address the imbalanced, multi-class learning problems for cancer detection. They have claimed that their method improved the maximum accuracy by 24.7%. Xiao et al. [185] proposed a DL-based ensemble method for cancer prediction. The proposed method was tested on three public RNA-seq dataset of three kinds of cancers, Lung Adenocarcinoma, Stomach Adenocarcinoma, and Breast Invasive Carcinoma. In

another study, Chougrad et al. [186] proposed to increase the success of mammography mass lesion classification by fine-tuning CNN methods that have been studied previously. The identification of interactions between compounds and proteins plays an important role in network pharmacology and drug discovery. Hence recently many researchers study to predict protein-protein interaction. In literature, a lot of studies based on DL about this area have been achieved by proving the advantages of DL techniques [187]–[191]. A comprehensive study of recent applications of DL in drug discovery research published by Chen et al. [87]. In the study, they introduced some popular DL architectures and discussed the future development of DL in drug discovery. Some products and medicines such as food additives, cleaning and cosmetic contain toxic substances. The prediction of the toxicity of chemical compounds is an important problem to solve. Therefore, there are many studies on toxicity prediction. Even a challenge named Toxicology in the 21st Century (Tox21) <https://tripod.nih.gov/tox21/challenge/> held every year to solve this problem. Unterthiner et al. [192] proposed a DL based method to learn automatically features resembling well-established toxicophores. There are various studies introduced in Tox21 by using DL for toxic prediction [11, 193, 194].

Consequently, there are more comprehensive studies on DL algorithms used in medical image processing field. Litjens et al. [66] published a survey on DL in medical image analysis and another survey about DL in medical image analysis was introduced by Shen et al. [68].

2.7. Big Data

Recently, together with the social network systems, sensor networks, and communication technologies, have made great progress which makes it possible to the collection of big data [195]. Big data is defined as a 3V model, which is described as high volume, high speed and high diversity [196]. In the following years, the 4V model was introduced by adding high value. This model can be even extended to 5V (L. Kuang) if the concept of Veracity is incorporated into the big data definition [70, 197], [198]. For big data analyzing, deep learning models can be examined in four aspects as DL models for huge amounts of data, DL models for heterogeneous data, DL models for real-time data and DL models for low-quality data. First and foremost, big data includes so many samples for analysis and so, it poses a big challenge on DL models. In order to extract features and classification

of big data, some large-scale DL models which have a few hidden layer and billions of parameters and parallel DL algorithms have been developed.

Deng et al. [153] presented DSN to cope with the problem of parallelizing learning algorithms for DL architectures. Hutchinson et al. [199] presented CPU clusters to improve the training efficiency of DSN and use an efficient and scalable parallel implementation. A Tensor DSN (T-DSN) that is a software framework called DistBelief was developed by Dean et al. to overcome the problem of training a deep network which uses tens of thousands of CPU cores with billions of parameters [200]. Therefore, it is very suitable to learn big data feature learning since DistBelief is able to scale up over many computers [201]. NVIDIA's large-scale DL models with 16 GPU servers are another impressive DL system for huge amounts of data analyzing. The system is able to train 1 billion parameter networks on just 3 machines in a couple of days, and it can scale to networks with over 11 billion parameters using just 16 machines [202]. There are also FPGA based solutions for large scale DL. In order to improve the performance of large scale DL models as well as to maintain the low power cost, it was designed the DL accelerator unit (DLAU), which is a scalable accelerator architecture for large-scale DL networks using field-programmable gate array (FPGA) [203]. Another FPGA model was developed to improve the performance of large scale DL models by Wang et al. [174].

One of the biggest problems in large data analysis is the data format differences. The data can be in different formats such as image, sound, raw data, MR, and X-Ray. For example, in data mining of social media such as Twitter and Flickr etc., the data may contain texts, images, audio, and videos. Similarly, various multi-modal information such as X-ray, CT, MRI, PET, SPECT, and fMRI etc. is collected in medical analysis. These kinds of heterogeneous data have so many complexities in data analysis. To overcome these problems, the multi-model DL model has been developed in literature. For example, a Deep Boltzmann Machine Model (DRBM) is developed for learning a generative model of data that consists of multiple and diverse input modalities by Ngiam et al. They used the DL model for audio-video objects feature learning [204]. Another multi-modal Restricted Boltzmann Machines (RBM) model was developed by Srivastava and Salakhutdinov [205]. In this model, they obtained a joint representation from the two given modalities that are image and text. Ouyang et al. built a multi-source deep model to extract non-linear representation from the score, appearance mixture type and deformation which are

three important information sources for human pose estimation. With the deep model, they extracted high-order human body articulation patterns [34]. To reach this goal, Zhao et al. [206] proposed a novel framework which is composed of two modules: multi-modal deep neural network, and feature selection with sparse group LASSO. Wang et al. [25] first construct deep CNN layers for color and depth separately, which are then connected with a designed multi-modal learning framework for RGB-D object recognition. Zhang et al. [207] proposed a deep computation model for feature learning on big data, which uses a tensor to model the complex correlations of heterogeneous data.

Even though today's technologies have accelerated DL algorithms, their classification performance is not good enough. Therefore, features extraction and representations for big data are difficult for traditional DL models in real-time classification problem. In recent years, for high-velocity data feature extraction, many incremental learning methods have presented. Wan and Banta [208] developed online training algorithm for neural networks, called parameter incremental learning (PIL) algorithm with high speed and accuracy. Zhao et al. [209] proposed a novel real-time and semi-supervised DL framework that integrates the strengths of computational epidemiology and social media mining techniques. Their aim was to learn the social media users' health states and intervention actions in real-time, which are regularized by the underlying disease model and contact network. Yu et al. [210] presented real-time (online) novel offline and online three-dimensional (3D) DL integration framework by leveraging the 3D fully convolutional network (3D-FCN) for automated polyp detection. Many researchers proposed an incremental learning method based on the structural modification to tackle real-time big data analytics overtime problem. Zhou et al. [211] proposed an incremental feature extraction algorithm to determine the optimal model complexity for large-scale online dataset based on the denoising auto encoder. The performed algorithm is effective in recognizing new patterns when the data distribution changes over time in the massive online data stream.

Low-quality, noisy and lack of data are other challenges in big data analytics. These challenges make so difficult to extract features and representations from big data by DL models. In literature, some studies used DL models on corrupted and noisy data. Vincent et al. [212] introduced a new training principle for unsupervised learning, which learns the representations with the partial corruption of the input pattern. In another study, they also

proposed an original strategy for building deep networks, based on stacking layers of denoising autoencoders which are trained locally to denoise corrupted versions of their inputs [213]. Wang and Tao [214] developed a (stacked) non-local auto-encoder, which exploits self-similar information in natural images for stability. In the study, they proposed that similar inputs should induce similar network propagation.

Big data analytics require new and sophisticated algorithms based on machine and DL techniques to process data in real-time with high accuracy and efficiency. In recent years, DL algorithms seem to fill this gap. Jan et al. [215] showed that DL techniques can be built by introducing a number of methods in combination with supervised and unsupervised training techniques for processing huge amount of data with the different number of neurons and hidden layers. Kiral-Kornek et al. [13] presented a seizure prediction system that is accurate, fully automated, patient-specific, and tunable for an individual's needs by using intracranial electroencephalography (iEEG) data. Qi et al. [216] proposed a fault-diagnosis system using machine-learning techniques to data analysis and fault diagnosis for the reciprocating compressors. They used the support vector machine (SVM) for recognizing and classifying. With the development of sensors and Internet of Things (IOT) technologies, there is an increasing need for big data analytics. For this reason, DL provides advanced analytics tools for processing and analyzing big data. In this section, we highlighted the success of DL methods on big data analytics and tried to draw attention to commonly used DL algorithms for big data analytics and discussed applications. In the study, we focused on advantages, methods and future studies about DL overall research areas. Hence, we did not much emphasis on mathematical details and theories of studies which are done with DL methods for big data analytics. In literature, there are deeper studies on DL algorithms used in big data analytics [71, 201, 207, 215]. Zhang et al. [71] presented a paper to point out the remaining challenges of big data analysis by using DL and discuss future topics about the emerging researches of DL models for big data feature learning.

3. SUGGESTIONS ON DEEP LEARNING (*DERİN ÖĞRENME ÖNERİLERİ*)

In this section, some popular subjects are proposed for researchers who want to study in DL. Many topics are identified in the direction of the information obtained from more than 230 articles in many research

areas and investigated from medical image processing to the signature verification. According to the reviewed papers, DL methods may produce more successful results than the other methods applied to the topics described in this paper.

Fault-diagnosis systems for reciprocating compressors have studied with different approaches as SVM [216–221]. But there are no many DL based studies for these systems. Qi et al. [216] proposed a fault-diagnosis system using SVM to data analysis and fault diagnosis for reciprocating compressors. DL approaches in such data analysis and classification processes will be more brittle. Some studies have done using the DL approach supports us [220], [222].

Although DL approaches have been widely used in many Medical Image Processing and achieved successful results [8, 13, 14, 16, 20, 66, 170, 171, 176, 180, 183, 223], it seems that they have not been used much in vessel segmentation yet. In literature, we observed that there are only a few studies to determine the eye vessels [224–226]. We believe that the DL models can obtain the successful results in vessel segmentation as it is in other medical fields.

Ariav et al. [155] first suggested that bidirectional RNNs may be used to exploit the temporal context from future frames, and long short-term memory (LSTM) networks may facilitate learning even longer-term dependencies between inputs. Second, it should be performed a fine-tuning of the entire network from end to end in a supervised manner, while simultaneously updating the weights of the auto-encoder and the RNN via backpropagation. According to obtained results, the proposed models can use for future works in DL [155].

DL approaches have achieved considerable success in face recognition systems as well as in many other areas. However, determining the vitality of the faces is as important as the face recognition in security [227–232]. Hence, it should be used the DL algorithms in spoof face detection method to determine both the identity and liveness of the face.

Although on-road obstacle detection and classification is one of the key tasks for self-driving vehicles, we noticed that the performance of studies is not high for the Indian road scenario [36, 91–93]. Another important task for autonomous car is collision prediction [24, 94, 95, 110]. However, there is no remarkable study in the literature. For this reason, we suppose that studies to overcome self-driving and collision prediction problems will be important in the future.

In literature, handwriting character recognition studies are widespread. In last decades, DL approaches are commonly used in handwriting

recognition applications in last decades [5, 10, 15, 43–46, 127, 117, 119, 122–125, 129]. There are at least 3,866 languages using an established spelling system in the world [233]. For this reason, it is concluded that DL algorithms can be applied to many different writing languages for handwriting character recognition in this review.

Signature recognition and verification is a promising area which must be addressed for researchers. Despite some studies on offline signature verification, there is not satisfied solution yet in literature [2, 28, 32, 35, 133, 136, 138–142, 234]. Therefore, signature verification is still one of the most challenging areas. It is recommended to researchers that the performance of the DL algorithms on signature verification can be increased by using feature extraction and augmentation methods. In literature, almost all studies on signature are about offline signature verification and most of them are based on CNN method. DL algorithms can be more successful in verification of online signatures since online signatures have high and robust features. Hence, online signature verification by DL models may be an important research area in the future.

4. CONCLUSION

In this study, we reviewed the deep learning studies related to popular seven research areas: Autonomous Vehicles, Natural Language Processing, Handwritten Character Recognition, Signature Verification, Voice and Video Recognition, Medical Image Processing, and Big Data. This review also determined the most challenging issues that can be studied with DL in these fields. Furthermore, we have pointed out the remaining challenges in these research areas that could be solved by using DL to help the researchers. Finally, we have presented the promising research topics which are unresolved yet.

The success of deep learning methods is increasing with new techniques developed. Nowadays, DL achieves better success than humans in many areas such as object recognition. It is anticipated that this success will be achieved in areas such as autonomous vehicles, medical image processing, and character recognition. DL methods can obtain promising results in solving problems such as obstacle detection and collision prediction. Big data analysis is another promising area where high success can be achieved by DL.

REFERENCES (KAYNAKLAR)

- [1] P. P. de San Roman, J. Benois-Pineau, J.-P. Domenger, F. Paquet, D. Cataert, and A. de Ruyg, "Saliency Driven Object recognition in egocentric videos with deep CNN: toward application in assistance to Neuroprostheses," *Comput. Vis. Image Underst.*, vol. 164, pp. 82–91, 2017.
- [2] Z. Zhang, X. Liu, and Y. Cui, "Multi-phase Offline Signature Verification System Using Deep Convolutional Generative Adversarial Networks," in *2016 9th International Symposium on Computational Intelligence and Design (ISCID)*, 2016, vol. 2, pp. 103–107.
- [3] J. Maria, J. Amaro, G. Falcao, and L. A. Alexandre, "Stacked Autoencoders Using Low-Power Accelerated Architectures for Object Recognition in Autonomous Systems," *Neural Process. Lett.*, vol. 43, no. 2, pp. 445–458, 2016.
- [4] H. Kaya, F. Gürpınar, and A. A. Salah, "Video-based emotion recognition in the wild using deep transfer learning and score fusion," *Image Vis. Comput.*, vol. 65, pp. 66–75, 2017.
- [5] R. Al-Jawfi, "Handwriting Arabic character recognition LeNet using neural network.," *Int. Arab J. Inf. Technol.*, vol. 6, no. 3, pp. 304–309, 2009.
- [6] B. Ribeiro, I. Gonçalves, S. Santos, and A. Kovacec, "Deep learning networks for off-line handwritten signature recognition," in *Iberoamerican Congress on Pattern Recognition*, 2011, pp. 523–532.
- [7] X. Sun *et al.*, "Transferring deep knowledge for object recognition in Low-quality underwater videos," *Neurocomputing*, vol. 275, pp. 897–908, 2018.
- [8] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in MRI images," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1240–1251, 2016.
- [9] M. Ribeiro, A. E. Lazzaretti, and H. S. Lopes, "A study of deep convolutional auto-encoders for anomaly detection in videos," *Pattern Recognit. Lett.*, vol. 105, pp. 13–22, 2018.
- [10] M. Al-Ayyoub, A. Nuseir, K. Alsmearat, Y. Jararweh, and B. Gupta, "Deep learning for Arabic NLP: A survey," *J. Comput. Sci.*, 2017.
- [11] A. Mayr, G. Klambauer, T. Unterthiner, and S. Hochreiter, "DeepTox: Toxicity Prediction using Deep Learning," *Front. Environ. Sci.*, vol. 3, p. 80, 2016.
- [12] D. Li and Z. Wang, "Video Superresolution via Motion Compensation and Deep Residual Learning," *IEEE Trans. Comput. Imaging*, vol. 3, no. 4, pp. 749–762, 2017.
- [13] I. Kiral-Kornek *et al.*, "Epileptic Seizure Prediction Using Big Data and Deep Learning: Toward a Mobile System," *EBioMedicine*, 2017.
- [14] A. A. A. Setio *et al.*, "Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1160–1169, May 2016.
- [15] C. Wu, W. Fan, Y. He, J. Sun, and S. Naoi, "Handwritten Character Recognition by Alternately Trained Relaxation Convolutional Neural Network," in *2014 14th International Conference on Frontiers in Handwriting Recognition*, 2014, pp. 291–296.
- [16] J. Kawahara and G. Hamarneh, "Multi-resolution-Tract CNN with Hybrid Pretrained and Skin-Lesion Trained Layers," in *Machine Learning in Medical Imaging*, 2016, pp. 164–171.
- [17] X. Song, T. Rui, S. Zhang, J. Fei, and X. Wang, "A road segmentation method based on the deep auto-encoder with supervised learning," *Comput. Electr. Eng.*, vol. 68, pp. 381–388, 2018.
- [18] S. Alghyaline, J. W. Hsieh, and C. H. Chuang, "Video action classification using symmelets and deep learning," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp. 414–419.
- [19] E. Cambria and B. White, "Jumping NLP Curves: A Review of Natural Language Processing Research [Review Article]," *IEEE Comput. Intell. Mag.*, vol. 9, no. 2, pp. 48–57, May 2014.
- [20] A. Khatami, M. Babaie, A. Khosravi, H. R. Tizhoosh, and S. Nahavandi, "Parallel deep

- solutions for image retrieval from imbalanced medical imaging archives,” *Appl. Soft Comput.*, vol. 63, pp. 197–205, 2018.
- [21] R. Socher, “Recursive deep learning for natural language processing and computer vision,” Citeseer, 2014.
- [22] M. Al-Qizwini, I. Barjasteh, H. Al-Qassab, and H. Radha, “Deep learning algorithm for autonomous driving using GoogLeNet,” in *2017 IEEE Intelligent Vehicles Symposium (IV)*, 2017, pp. 89–96.
- [23] E. Nasr-Esfahani *et al.*, “Segmentation of vessels in angiograms using convolutional neural networks,” *Biomed. Signal Process. Control*, vol. 40, pp. 240–251, 2018.
- [24] S. Ramos, S. Gehrig, P. Pinggera, U. Franke, and C. Rother, “Detecting unexpected obstacles for self-driving cars: Fusing deep learning and geometric modeling,” in *2017 IEEE Intelligent Vehicles Symposium (IV)*, 2017, pp. 1025–1032.
- [25] A. Wang, J. Lu, J. Cai, T. J. Cham, and G. Wang, “Large-Margin Multi-Modal Deep Learning for RGB-D Object Recognition,” *IEEE Trans. Multimed.*, vol. 17, no. 11, pp. 1887–1898, 2015.
- [26] A. Uçar, Y. Demir, and C. Güzeliş, “Moving towards in object recognition with deep learning for autonomous driving applications,” in *2016 International Symposium on Innovations in Intelligent Systems and Applications (INISTA)*, 2016, pp. 1–5.
- [27] X. Zhang, X. Li, J. An, L. Gao, B. Hou, and C. Li, “Natural language description of remote sensing images based on deep learning,” in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017, pp. 4798–4801.
- [28] I. Abroug and N. E. Ben Amara, “Off-line signature verification systems: Recent advances,” in *International Image Processing, Applications and Systems Conference*, 2014, pp. 1–6.
- [29] A. Işın, C. Direkoğlu, and M. Şah, “Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods,” *Procedia Comput. Sci.*, vol. 102, pp. 317–324, 2016.
- [30] D. Ciregan, U. Meier, and J. Schmidhuber, “Multi-column deep neural networks for image classification,” in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 3642–3649.
- [31] F. I. Vancea, A. D. Costea, and S. Nedevschi, “Vehicle taillight detection and tracking using deep learning and thresholding for candidate generation,” in *2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, 2017, pp. 267–272.
- [32] L. G. Hafemann, R. Sabourin, and L. S. Oliveira, “Offline handwritten signature verification #x2014; Literature review,” in *2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 2017, pp. 1–8.
- [33] F. Milletari *et al.*, “Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound,” *Comput. Vis. Image Underst.*, vol. 164, pp. 92–102, 2017.
- [34] W. Ouyang, X. Chu, and X. Wang, “Multi-source Deep Learning for Human Pose Estimation,” in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2337–2344.
- [35] L. G. Hafemann, R. Sabourin, and L. S. Oliveira, “Writer-independent feature learning for Offline Signature Verification using Deep Convolutional Neural Networks,” in *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016, pp. 2576–2583.
- [36] G. Prabhakar, B. Kailath, S. Natarajan, and R. Kumar, “Obstacle detection and classification using deep learning for tracking in high-speed autonomous driving,” in *2017 IEEE Region 10 Symposium (TENSymp)*, 2017, pp. 1–6.
- [37] A. Becerra, J. I. de la Rosa, E. González, A. D. Pedroza, J. M. Martínez, and N. I. Escalante, “Speech recognition using deep neural networks trained with non-uniform frame-level cost functions,” in *2017 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*, 2017, pp. 1–6.
- [38] Đ. T. Grozdić, S. T. Jovičić, and M. Subotić, “Whispered speech recognition using deep denoising autoencoder,” *Eng. Appl. Artif. Intell.*, vol. 59, pp. 15–22, 2017.

- [39] S. Tamura *et al.*, “Audio-visual speech recognition using deep bottleneck features and high-performance lipreading,” in *2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, 2015, pp. 575–582.
- [40] R. Sarikaya, G. E. Hinton, and A. Deoras, “Application of Deep Belief Networks for Natural Language Understanding,” *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 4, pp. 778–784, 2014.
- [41] R. Galinsky, A. Alekseev, and S. I. Nikolenko, “Improving neural network models for natural language processing in russian with synonyms,” in *2016 IEEE Artificial Intelligence and Natural Language Conference (AINL)*, 2016, pp. 1–7.
- [42] H. Zhuang, C. Wang, C. Li, Q. Wang, and X. Zhou, “Natural Language Processing Service Based on Stroke-Level Convolutional Networks for Chinese Text Classification,” in *2017 IEEE International Conference on Web Services (ICWS)*, 2017, pp. 404–411.
- [43] S. Iamsa-at and P. Horata, “Handwritten Character Recognition Using Histograms of Oriented Gradient Features in Deep Learning of Artificial Neural Network,” in *2013 International Conference on IT Convergence and Security (ICITCS)*, 2013, pp. 1–5.
- [44] X. Xiao, L. Jin, Y. Yang, W. Yang, J. Sun, and T. Chang, “Building fast and compact convolutional neural networks for offline handwritten Chinese character recognition,” *Pattern Recognit.*, vol. 72, pp. 72–81, 2017.
- [45] I.-J. Kim and X. Xie, “Handwritten Hangul recognition using deep convolutional neural networks,” *Int. J. Doc. Anal. Recognit.*, vol. 18, no. 1, pp. 1–13, Mar. 2015.
- [46] S. Zheng *et al.*, “Sunspot drawings handwritten character recognition method based on deep learning,” *New Astron.*, vol. 45, pp. 54–59, 2016.
- [47] H. Feng and C. C. Wah, “Online signature verification using a new extreme points warping technique,” *Pattern Recognit. Lett.*, vol. 24, no. 16, pp. 2943–2951, 2003.
- [48] D. Bertolini, L. S. Oliveira, E. Justino, and R. Sabourin, “Reducing forgeries in writer-independent off-line signature verification through ensemble of classifiers,” *Pattern Recognit.*, vol. 43, no. 1, pp. 387–396, 2010.
- [49] G. Rigoll and A. Kosmala, “A systematic comparison between on-line and off-line methods for signature verification with hidden Markov models,” in *Proceedings. Fourteenth International Conference on Pattern Recognition (Cat. No.98EX170)*, 1998, vol. 2, pp. 1755–1757 vol.2.
- [50] P. Porwik, R. Doroz, and T. Orczyk, “Signatures verification based on PNN classifier optimised by PSO algorithm,” *Pattern Recognit.*, vol. 60, pp. 998–1014, 2016.
- [51] R. Kumar, J. D. Sharma, and B. Chanda, “Writer-independent off-line signature verification using surroundedness feature,” *Pattern Recognit. Lett.*, vol. 33, no. 3, pp. 301–308, 2012.
- [52] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *Bull. Math. Biophys.*, vol. 5, no. 4, pp. 115–133, 1943.
- [53] G. E. Hinton, S. Osindero, and Y.-W. Teh, “A fast learning algorithm for deep belief nets,” *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2006.
- [54] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [55] Google, “TensorFlow.” [Online]. Available: <https://www.tensorflow.org/>.
- [56] M. Abadi *et al.*, “Tensorflow: Large-scale machine learning on heterogeneous distributed systems,” *arXiv Prepr. arXiv1603.04467*, 2016.
- [57] Facebook, “FAIR,” <https://research.fb.com/fair-open-sources-deep-learning-modules-for-torch/>.
- [58] C. Microsoft, “Computational Network Toolkit (CNTK),” 2016. [Online]. Available: <https://www.microsoft.com/en-us/cognitive-toolkit/>.
- [59] D. S. Banerjee, K. Hamidouche, and D. K.

- Panda, "Re-Designing CNTK Deep Learning Framework on Modern GPU Enabled Clusters," in *2016 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)*, 2016, pp. 144–151.
- [60] NVIDIA, "Caffe2 Deep Learning Framework," <https://developer.nvidia.com/caffe2>, 2017. .
- [61] Y. Jia *et al.*, "Caffe: Convolutional architecture for fast feature embedding," in *Proceedings of the 22nd ACM international conference on Multimedia*, 2014, pp. 675–678.
- [62] S. Shi, Q. Wang, P. Xu, and X. Chu, "Benchmarking State-of-the-Art Deep Learning Software Tools," in *2016 7th International Conference on Cloud Computing and Big Data (CCBD)*, 2016, pp. 99–104.
- [63] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.
- [64] S. Min, B. Lee, and S. Yoon, "Deep learning in bioinformatics," *Brief. Bioinform.*, vol. 18, no. 5, pp. 851–869, 2017.
- [65] V. N. Nguyen, R. Jenssen, and D. Roverso, "Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning," *Int. J. Electr. Power Energy Syst.*, vol. 99, pp. 107–120, 2018.
- [66] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.
- [67] Z. Li, X. Zhang, H. Müller, and S. Zhang, "Large-scale retrieval for medical image analytics: A comprehensive review," *Med. Image Anal.*, vol. 43, pp. 66–84, 2018.
- [68] D. Shen, G. Wu, and H.-I. Suk, "Deep Learning in Medical Image Analysis," *Annu. Rev. Biomed. Eng.*, vol. 19, no. 1, pp. 221–248, 2017.
- [69] Z. Hu, J. Tang, Z. Wang, K. Zhang, L. Zhang, and Q. Sun, "Deep learning for image-based cancer detection and diagnosis – A survey," *Pattern Recognit.*, vol. 83, pp. 134–149, 2018.
- [70] H. Fang, Z. Zhang, C. J. Wang, M. Daneshmand, C. Wang, and H. Wang, "A survey of big data research," *IEEE Netw.*, vol. 29, no. 5, pp. 6–9, 2015.
- [71] Q. Zhang, L. T. Yang, Z. Chen, and P. Li, "A survey on deep learning for big data," *Inf. Fusion*, vol. 42, pp. 146–157, 2018.
- [72] A. R. Sharma and P. Kaushik, "Literature survey of statistical, deep and reinforcement learning in natural language processing," in *2017 International Conference on Computing, Communication and Automation (ICCCA)*, 2017, pp. 350–354.
- [73] A. Sanmorino and S. Yazid, "A survey for handwritten signature verification," in *2012 2nd International Conference on Uncertainty Reasoning and Knowledge Engineering*, 2012, pp. 54–57.
- [74] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, 2003.
- [75] P. Wang, W. Li, P. Ogunbona, J. Wan, and S. Escalera, "RGB-D-based human motion recognition with deep learning: A survey," *Comput. Vis. Image Underst.*, 2018.
- [76] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: A Survey," *Pattern Recognit. Lett.*, 2018.
- [77] S. Purushotham, C. Meng, Z. Che, and Y. Liu, "Benchmarking deep learning models on large healthcare datasets," *J. Biomed. Inform.*, vol. 83, pp. 112–134, 2018.
- [78] P. Meyer, V. Noblet, C. Mazzara, and A. Lallement, "Survey on deep learning for radiotherapy," *Comput. Biol. Med.*, vol. 98, pp. 126–146, 2018.
- [79] P. Li, D. Wang, L. Wang, and H. Lu, "Deep visual tracking: Review and experimental comparison," *Pattern Recognit.*, vol. 76, pp. 323–338, 2018.
- [80] S. Khan and T. Yairi, "A review on the application of deep learning in system health management," *Mech. Syst. Signal Process.*, vol. 107, pp. 241–265, 2018.
- [81] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput. Electron. Agric.*, vol. 147, pp. 70–90, 2018.
- [82] S. Herath, M. Harandi, and F. Porikli, "Going

- deeper into action recognition: A survey,” *Image Vis. Comput.*, vol. 60, pp. 4–21, 2017.
- [83] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, “Deep learning for visual understanding: A review,” *Neurocomputing*, vol. 187, pp. 27–48, 2016.
- [84] P. S. Grewal, F. Oloumi, U. Rubin, and M. T. S. Tennant, “Deep learning in ophthalmology: a review,” *Can. J. Ophthalmol.*, 2018.
- [85] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, P. Martinez-Gonzalez, and J. Garcia-Rodriguez, “A survey on deep learning techniques for image and video semantic segmentation,” *Appl. Soft Comput.*, vol. 70, pp. 41–65, 2018.
- [86] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, “Deep learning for healthcare applications based on physiological signals: A review,” *Comput. Methods Programs Biomed.*, vol. 161, pp. 1–13, 2018.
- [87] H. Chen, O. Engkvist, Y. Wang, M. Olivecrona, and T. Blaschke, “The rise of deep learning in drug discovery,” *Drug Discov. Today*, 2018.
- [88] K. Shi, H. Bao, and N. Ma, “Forward Vehicle Detection Based on Incremental Learning and Fast R-CNN,” in *2017 13th International Conference on Computational Intelligence and Security (CIS)*, 2017, pp. 73–76.
- [89] X. Du, M. H. Ang, and D. Rus, “Car detection for autonomous vehicle: LIDAR and vision fusion approach through deep learning framework,” in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 749–754.
- [90] A. Soin and M. Chahande, “Moving vehicle detection using deep neural network,” in *2017 International Conference on Emerging Trends in Computing and Communication Technologies (ICETCCT)*, 2017, pp. 1–5.
- [91] V. D. Nguyen, H. Van Nguyen, D. T. Tran, S. J. Lee, and J. W. Jeon, “Learning Framework for Robust Obstacle Detection, Recognition, and Tracking,” *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 6, pp. 1633–1646, 2017.
- [92] N. Deepika and V. V. S. Variyar, “Obstacle classification and detection for vision based navigation for autonomous driving,” in *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2017, pp. 2092–2097.
- [93] A. Dairi, F. Harrou, M. Senouci, and Y. Sun, “Unsupervised obstacle detection in driving environments using deep-learning-based stereovision,” *Rob. Auton. Syst.*, vol. 100, pp. 287–301, 2018.
- [94] C. Chen, H. Xiang, T. Qiu, C. Wang, Y. Zhou, and V. Chang, “A rear-end collision prediction scheme based on deep learning in the Internet of Vehicles,” *J. Parallel Distrib. Comput.*, 2017.
- [95] Y. F. Chen, M. Liu, M. Everett, and J. P. How, “Decentralized non-communicating multiagent collision avoidance with deep reinforcement learning,” in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 285–292.
- [96] W. Huang, G. Song, H. Hong, and K. Xie, “Deep Architecture for Traffic Flow Prediction: Deep Belief Networks With Multitask Learning,” *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2191–2201, 2014.
- [97] Y. Jia, J. Wu, and Y. Du, “Traffic speed prediction using deep learning method,” in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2016, pp. 1217–1222.
- [98] Y. Jia, J. Wu, M. Ben-Akiva, R. Seshadri, and Y. Du, “Rainfall-integrated traffic speed prediction using deep learning method,” *IET Intell. Transp. Syst.*, vol. 11, no. 9, pp. 531–536, 2017.
- [99] A. Koesdwiady, R. Soua, and F. Karray, “Improving Traffic Flow Prediction With Weather Information in Connected Cars: A Deep Learning Approach,” *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9508–9517, 2016.
- [100] J. Wang, Q. Gu, J. Wu, G. Liu, and Z. Xiong, “Traffic Speed Prediction and Congestion Source Exploration: A Deep Learning Method,” in *2016 IEEE 16th International Conference on Data Mining (ICDM)*, 2016, pp. 499–508.
- [101] X. Du, H. Zhang, H. V. Nguyen, and Z. Han,

- “Stacked LSTM Deep Learning Model for Traffic Prediction in Vehicle-to-Vehicle Communication,” in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 2017, pp. 1–5.
- [102] Y. Liu, H. Zheng, X. Feng, and Z. Chen, “Short-term traffic flow prediction with Conv-LSTM,” in *2017 9th International Conference on Wireless Communications and Signal Processing (WCSP)*, 2017, pp. 1–6.
- [103] N. G. Polson and V. O. Sokolov, “Deep learning for short-term traffic flow prediction,” *Transp. Res. Part C Emerg. Technol.*, vol. 79, pp. 1–17, 2017.
- [104] Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Y. Wang, “Traffic Flow Prediction With Big Data: A Deep Learning Approach,” *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 865–873, 2015.
- [105] Y. Duan, Y. Lv, and F. Y. Wang, “Performance evaluation of the deep learning approach for traffic flow prediction at different times,” in *2016 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI)*, 2016, pp. 223–227.
- [106] J. n. Xin, X. Du, and J. Zhang, “Deep learning for robust outdoor vehicle visual tracking,” in *2017 IEEE International Conference on Multimedia and Expo (ICME)*, 2017, pp. 613–618.
- [107] R. Hadsell, A. Erkan, P. Sermanet, M. Scoffier, U. Muller, and Y. LeCun, “Deep belief net learning in a long-range vision system for autonomous off-road driving,” in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008, pp. 628–633.
- [108] V. Rausch, A. Hansen, E. Solowjow, C. Liu, E. Kreuzer, and J. K. Hedrick, “Learning a deep neural net policy for end-to-end control of autonomous vehicles,” in *2017 American Control Conference (ACC)*, 2017, pp. 4914–4919.
- [109] W. Xia, H. Li, and B. Li, “A Control Strategy of Autonomous Vehicles Based on Deep Reinforcement Learning,” in *2016 9th International Symposium on Computational Intelligence and Design (ISCID)*, 2016, vol. 2, pp. 198–201.
- [110] T. Zhang, G. Kahn, S. Levine, and P. Abbeel, “Learning deep control policies for autonomous aerial vehicles with MPC-guided policy search,” in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 528–535.
- [111] M. F. Yahya and M. R. Arshad, “Detection of markers using deep learning for docking of autonomous underwater vehicle,” in *2017 IEEE 2nd International Conference on Automatic Control and Intelligent Systems (2CACIS)*, 2017, pp. 179–184.
- [112] R. Yu, Z. Shi, C. Huang, T. Li, and Q. Ma, “Deep reinforcement learning based optimal trajectory tracking control of autonomous underwater vehicle,” in *2017 36th Chinese Control Conference (CCC)*, 2017, pp. 4958–4965.
- [113] R. Sarikaya, G. E. Hinton, and B. Ramabhadran, “Deep belief nets for natural language call-routing,” in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 5680–5683.
- [114] N. Dethlefs, “Domain Transfer for Deep Natural Language Generation from Abstract Meaning Representations,” *IEEE Comput. Intell. Mag.*, vol. 12, no. 3, pp. 18–28, 2017.
- [115] Y. LeCun *et al.*, “Handwritten digit recognition with a back-propagation network,” in *Advances in neural information processing systems*, 1990, pp. 396–404.
- [116] Y. Bengio, Y. LeCun, C. Nohl, and C. Burges, “LeRec: A NN/HMM Hybrid for On-Line Handwriting Recognition,” *Neural Comput.*, vol. 7, no. 6, pp. 1289–1303, 1995.
- [117] S. Espana-Boquera, M. J. Castro-Bleda, J. Gorbe-Moya, and F. Zamora-Martinez, “Improving Offline Handwritten Text Recognition with Hybrid HMM/ANN Models,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 4, pp. 767–779, 2011.
- [118] F. Yin, Q. F. Wang, X. Y. Zhang, and C. L. Liu, “ICDAR 2013 Chinese Handwriting Recognition Competition,” in *2013 12th International Conference on Document Analysis and Recognition*, 2013, pp. 1464–1470.

- [119] M. M. R. Sazal, S. K. Biswas, M. F. Amin, and K. Murase, "Bangla handwritten character recognition using deep belief network," in *2013 International Conference on Electrical Information and Communication Technology (EICT)*, 2014, pp. 1–5.
- [120] V. Pham, T. Bluche, C. Kermorvant, and J. Louradour, "Dropout Improves Recurrent Neural Networks for Handwriting Recognition," in *2014 14th International Conference on Frontiers in Handwriting Recognition*, 2014, pp. 285–290.
- [121] P. Doetsch, M. Kozielski, and H. Ney, "Fast and Robust Training of Recurrent Neural Networks for Offline Handwriting Recognition," in *2014 14th International Conference on Frontiers in Handwriting Recognition*, 2014, pp. 279–284.
- [122] D. C. Cireşan, U. Meier, L. M. Gambardella, and J. Schmidhuber, "Convolutional Neural Network Committees for Handwritten Character Classification," in *2011 International Conference on Document Analysis and Recognition*, 2011, pp. 1135–1139.
- [123] D. Cireşan and U. Meier, "Multi-Column Deep Neural Networks for offline handwritten Chinese character classification," in *2015 International Joint Conference on Neural Networks (IJCNN)*, 2015, pp. 1–6.
- [124] L. Chen, S. Wang, W. Fan, J. Sun, and S. Naoi, "Beyond human recognition: A CNN-based framework for handwritten character recognition," in *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, 2015, pp. 695–699.
- [125] W. Yang, L. Jin, D. Tao, Z. Xie, and Z. Feng, "DropSample: A new training method to enhance deep convolutional neural networks for large-scale unconstrained handwritten Chinese character recognition," *Pattern Recognit.*, vol. 58, pp. 190–203, 2016.
- [126] S. Huang, Z. Zhong, L. Jin, S. Zhang, and H. Wang, "DropRegion training of inception font network for high-performance Chinese font recognition," *Pattern Recognit.*, vol. 77, pp. 395–411, 2018.
- [127] C. Boufenar, A. Kerboua, and M. Batouche, "Investigation on deep learning for off-line handwritten Arabic character recognition," *Cogn. Syst. Res.*, 2017.
- [128] A. Trivedi, S. Srivastava, A. Mishra, A. Shukla, and R. Tiwari, "Hybrid evolutionary approach for Devanagari handwritten numeral recognition using Convolutional Neural Network," *Procedia Comput. Sci.*, vol. 125, pp. 525–532, 2018.
- [129] M. Soomro, M. A. Farooq, and R. H. Raza, "Performance Evaluation of Advanced Deep Learning Architectures for Offline Handwritten Character Recognition," in *2017 International Conference on Frontiers of Information Technology (FIT)*, 2017, pp. 362–367.
- [130] J. Sueiras, V. Ruiz, A. Sanchez, and J. F. Velez, "Offline continuous handwriting recognition using sequence to sequence neural networks," *Neurocomputing*, vol. 289, pp. 119–128, 2018.
- [131] A. K. Jain, F. D. Griess, and S. D. Connell, "On-line signature verification," *Pattern Recognit.*, vol. 35, no. 12, pp. 2963–2972, 2002.
- [132] Y. Qi and B. R. Hunt, "Signature verification using global and grid features," *Pattern Recognit.*, vol. 27, no. 12, pp. 1621–1629, 1994.
- [133] M. A. Shouman, N. Lashin, and H. M. Hamza, "OFFLINE SIGNATURE VERIFICATION BASED ON DIFFERENT SETS OF FEATURES."
- [134] R. Doroz, P. Porwik, and T. Orczyk, "Dynamic signature verification method based on association of features with similarity measures," *Neurocomputing*, vol. 171, pp. 921–931, 2016.
- [135] M. Fayyaz, M. Hajizadeh Saffar, M. Sabokrou, and M. Fathy, "Feature representation for online signature verification," *arXiv Prepr. arXiv1505.08153*, 2015.
- [136] M. Liwicki *et al.*, "Signature Verification Competition for Online and Offline Skilled Forgeries (SigComp2011)," in *2011 International Conference on Document Analysis and Recognition*, 2011, pp. 1480–1484.

- [137] M. Liwicki, M. I. Malik, L. Alewijnse, E. v. d. Heuvel, and B. Found, "ICFHR 2012 Competition on Automatic Forensic Signature Verification (4NsigComp 2012)," in *2012 International Conference on Frontiers in Handwriting Recognition*, 2012, pp. 823–828.
- [138] M. I. Malik, M. Liwicki, L. Alewijnse, W. Ohyama, M. Blumenstein, and B. Found, "ICDAR 2013 Competitions on Signature Verification and Writer Identification for On- and Offline Skilled Forgeries (SigWiComp 2013)," in *2013 12th International Conference on Document Analysis and Recognition*, 2013, pp. 1477–1483.
- [139] Y. M. Al-Omari, S. N. H. S. Abdullah, and K. Omar, "State-of-the-art in offline signature verification system," in *2011 International Conference on Pattern Analysis and Intelligence Robotics*, 2011, vol. 1, pp. 59–64.
- [140] A. Soleimani, B. N. Araabi, and K. Fouladi, "Deep Multitask Metric Learning for Offline Signature Verification," *Pattern Recognit. Lett.*, vol. 80, pp. 84–90, 2016.
- [141] L. G. Hafemann, R. Sabourin, and L. S. Oliveira, "Analyzing features learned for Offline Signature Verification using Deep CNNs," in *2016 23rd International Conference on Pattern Recognition (ICPR)*, 2016, pp. 2989–2994.
- [142] L. G. Hafemann, R. Sabourin, and L. S. Oliveira, "Learning features for offline handwritten signature verification using deep convolutional neural networks," *Pattern Recognit.*, vol. 70, pp. 163–176, 2017.
- [143] S. Tayeb *et al.*, "Toward data quality analytics in signature verification using a convolutional neural network," in *2017 IEEE International Conference on Big Data (Big Data)*, 2017, pp. 2644–2651.
- [144] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "Exploring Recurrent Neural Networks for On-Line Handwritten Signature Biometrics," *IEEE Access*, vol. 6, pp. 5128–5138, 2018.
- [145] Y.-W. Tan, W.-J. Liu, W. Jiang, and H. Zheng, "Integration of articulatory knowledge and voicing features based on DNN/HMM for Mandarin speech recognition," in *2015 International Joint Conference on Neural Networks (IJCNN)*, 2015, pp. 1–8.
- [146] B. Wu *et al.*, "An End-to-End Deep Learning Approach to Simultaneous Speech Dereverberation and Acoustic Modeling for Robust Speech Recognition," *IEEE J. Sel. Top. Signal Process.*, vol. 11, no. 8, pp. 1289–1300, 2017.
- [147] Y.-H. Tu, J. Du, Q. Wang, X. Bao, L.-R. Dai, and C.-H. Lee, "An information fusion framework with multi-channel feature concatenation and multi-perspective system combination for the deep-learning-based robust recognition of microphone array speech," *Comput. Speech Lang.*, vol. 46, pp. 517–534, 2017.
- [148] T. Gao, J. Du, L.-R. Dai, and C.-H. Lee, "A unified DNN approach to speaker-dependent simultaneous speech enhancement and speech separation in low SNR environments," *Speech Commun.*, vol. 95, pp. 28–39, 2017.
- [149] H. M. Fayek, M. Lech, and L. Cavedon, "Evaluating deep learning architectures for Speech Emotion Recognition," *Neural Networks*, vol. 92, pp. 60–68, 2017.
- [150] K. Mannepalli, P. N. Sastry, and M. Suman, "A novel Adaptive Fractional Deep Belief Networks for speaker emotion recognition," *Alexandria Eng. J.*, vol. 56, no. 4, pp. 485–497, 2017.
- [151] Q. Zhang, X. Chen, Q. Zhan, T. Yang, and S. Xia, "Respiration-based emotion recognition with deep learning," *Comput. Ind.*, vol. 92–93, pp. 84–90, 2017.
- [152] T. Bhowmik, A. Chowdhury, and S. K. Das Mandal, "Deep Neural Network based Place and Manner of Articulation Detection and Classification for Bengali Continuous Speech," *Procedia Comput. Sci.*, vol. 125, pp. 895–901, 2018.
- [153] L. Deng, D. Yu, and J. Platt, "Scalable stacking and learning for building deep architectures," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, 2012, pp. 2133–2136.
- [154] S. Achanta and S. V Gangashetty, "Deep Elman recurrent neural networks for statistical parametric speech synthesis," *Speech Commun.*, vol. 93, pp. 31–42, 2017.

- [155] I. Ariav, D. Dov, and I. Cohen, "A deep architecture for audio-visual voice activity detection in the presence of transients," *Signal Processing*, vol. 142, pp. 69–74, 2018.
- [156] R. Olmos, S. Tabik, and F. Herrera, "Automatic handgun detection alarm in videos using deep learning," *Neurocomputing*, vol. 275, pp. 66–72, 2018.
- [157] M. Parchami, S. Bashbaghi, E. Granger, and S. Sayed, "Using deep autoencoders to learn robust domain-invariant representations for still-to-video face recognition," in *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2017, pp. 1–6.
- [158] Y. Rao, J. Lu, and J. Zhou, "Attention-Aware Deep Reinforcement Learning for Video Face Recognition," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 3951–3960.
- [159] Y. Akbulut, A. Şengür, Ü. Budak, and S. Ekici, "Deep learning based face liveness detection in videos," in *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*, 2017, pp. 1–4.
- [160] B. Cheng *et al.*, "Robust emotion recognition from low quality and low bit rate video: A deep learning approach," in *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2017, pp. 65–70.
- [161] O. Gupta, D. Raviv, and R. Raskar, "Illumination invariants in deep video expression recognition," *Pattern Recognit.*, vol. 76, pp. 25–35, 2018.
- [162] Y. Han, P. Zhang, T. Zhuo, W. Huang, and Y. Zhang, "Going deeper with two-stream ConvNets for action recognition in video surveillance," *Pattern Recognit. Lett.*, 2017.
- [163] E. P. Ijjina and K. M. Chalavadi, "Human action recognition in RGB-D videos using motion sequence information and deep learning," *Pattern Recognit.*, vol. 72, pp. 504–516, 2017.
- [164] G. I. Parisi, J. Tani, C. Weber, and S. Wermter, "Lifelong learning of human actions with deep neural network self-organization," *Neural Networks*, vol. 96, pp. 137–149, 2017.
- [165] M. Ma, N. Marturi, Y. Li, A. Leonardis, and R. Stolkin, "Region-sequence based six-stream CNN features for general and fine-grained human action recognition in videos," *Pattern Recognit.*, vol. 76, pp. 506–521, 2018.
- [166] W. Huang, H. Ding, and G. Chen, "A novel deep multi-channel residual networks-based metric learning method for moving human localization in video surveillance," *Signal Processing*, vol. 142, pp. 104–113, 2018.
- [167] M. Iliadis, L. Spinoulas, and A. K. Katsaggelos, "Deep fully-connected networks for video compressive sensing," *Digit. Signal Process.*, vol. 72, pp. 9–18, 2018.
- [168] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty, and A.-B. M. Salem, "Classification using Deep Learning Neural Networks for Brain Tumors," *Futur. Comput. Informatics J.*, 2017.
- [169] X. Gao, W. Li, M. Loomes, and L. Wang, "A fused deep learning architecture for viewpoint classification of echocardiography," *Inf. Fusion*, vol. 36, pp. 103–113, 2017.
- [170] T. Kooi *et al.*, "Large scale deep learning for computer aided detection of mammographic lesions," *Med. Image Anal.*, vol. 35, pp. 303–312, 2017.
- [171] M. Saha, C. Chakraborty, and D. Racoceanu, "Efficient Deep Learning Model for Mitosis Detection using Breast Histopathology Images," *Comput. Med. Imaging Graph.*, 2017.
- [172] J. Wang and Y. Yang, "A context-sensitive deep learning approach for microcalcification detection in mammograms," *Pattern Recognit.*, 2018.
- [173] J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, "An ensemble deep learning based approach for red lesion detection in fundus images," *Comput. Methods Programs Biomed.*, vol. 153, pp. 115–127, 2018.
- [174] C. Wang, L. Gong, Q. Yu, X. Li, Y. Xie, and X. Zhou, "DLAU: A scalable deep learning accelerator unit on FPGA," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 36, no. 3, pp. 513–517, 2017.

- [175] S. Hussain, S. M. Anwar, and M. Majid, "Segmentation of glioma tumors in brain using deep convolutional neural network," *Neurocomputing*, 2017.
- [176] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, and Y. Fan, "A deep learning model integrating FCNNs and CRFs for brain tumor segmentation," *Med. Image Anal.*, vol. 43, pp. 98–111, 2018.
- [177] S. Valverde *et al.*, "Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach," *Neuroimage*, vol. 155, pp. 159–168, 2017.
- [178] A. Birenbaum and H. Greenspan, "Multi-view longitudinal CNN for multiple sclerosis lesion segmentation," *Eng. Appl. Artif. Intell.*, vol. 65, pp. 111–118, 2017.
- [179] B. D. Barkana, I. Saricicek, and B. Yildirim, "Performance analysis of descriptive statistical features in retinal vessel segmentation via fuzzy logic, ANN, SVM, and classifier fusion," *Knowledge-Based Syst.*, vol. 118, pp. 165–176, 2017.
- [180] S. Miao, Z. J. Wang, and R. Liao, "A CNN Regression Approach for Real-Time 2D/3D Registration," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1352–1363, May 2016.
- [181] X. Yang, R. Kwitt, M. Styner, and M. Niethammer, "Quicksilver: Fast predictive image registration – A deep learning approach," *Neuroimage*, vol. 158, pp. 378–396, 2017.
- [182] H. Jia, Y. Xia, Y. Song, W. Cai, M. Fulham, and D. D. Feng, "Atlas registration and ensemble deep convolutional neural network-based prostate segmentation using magnetic resonance imaging," *Neurocomputing*, vol. 275, pp. 1358–1369, 2018.
- [183] A. Qayyum, S. M. Anwar, M. Awais, and M. Majid, "Medical image retrieval using deep convolutional neural network," *Neurocomputing*, vol. 266, pp. 8–20, 2017.
- [184] X. Yuan, L. Xie, and M. Abouelenien, "A regularized ensemble framework of deep learning for cancer detection from multi-class, imbalanced training data," *Pattern Recognit.*, vol. 77, pp. 160–172, 2018.
- [185] Y. Xiao, J. Wu, Z. Lin, and X. Zhao, "A deep learning-based multi-model ensemble method for cancer prediction," *Comput. Methods Programs Biomed.*, vol. 153, pp. 1–9, 2018.
- [186] H. Chougrad, H. Zouaki, and O. Alheyane, "Deep Convolutional Neural Networks for Breast Cancer Screening," *Comput. Methods Programs Biomed.*, 2018.
- [187] Y. Wang, H. Mao, and Z. Yi, "Protein secondary structure prediction by using deep learning method," *Knowledge-Based Syst.*, vol. 118, pp. 115–123, 2017.
- [188] T. Du, L. Liao, C. H. Wu, and B. Sun, "Prediction of residue-residue contact matrix for protein-protein interaction with Fisher score features and deep learning," *Methods*, vol. 110, pp. 97–105, 2016.
- [189] S. Wang, Z. Li, Y. Yu, and J. Xu, "Folding Membrane Proteins by Deep Transfer Learning," *Cell Syst.*, vol. 5, no. 3, p. 202–211.e3, 2017.
- [190] L. Wei, Y. Ding, R. Su, J. Tang, and Q. Zou, "Prediction of human protein subcellular localization using deep learning," *J. Parallel Distrib. Comput.*, 2017.
- [191] K. Tian, M. Shao, Y. Wang, J. Guan, and S. Zhou, "Boosting compound-protein interaction prediction by deep learning," *Methods*, vol. 110, pp. 64–72, 2016.
- [192] T. Unterthiner, A. Mayr, G. Klambauer, and S. Hochreiter, "Toxicity prediction using deep learning," *arXiv Prepr. arXiv1503.01445*, 2015.
- [193] R. Huang *et al.*, "Tox21Challenge to Build Predictive Models of Nuclear Receptor and Stress Response Pathways as Mediated by Exposure to Environmental Chemicals and Drugs," *Front. Environ. Sci.*, vol. 3, p. 85, 2016.
- [194] A. Koutsoukas, J. St. Amand, M. Mishra, and J. Huan, "Predictive Toxicology: Modeling Chemical Induced Toxicological Response Combining Circular Fingerprints with Random Forest and Support Vector Machine," *Front. Environ. Sci.*, vol. 4, p. 11, 2016.
- [195] L. Kuang, L. Yang, and Y. Liao, "An Integration Framework on Cloud for Cyber

- Physical Social Systems Big Data,” *IEEE Trans. Cloud Comput.*, 2015.
- [196] S. K. Bansal and S. Kagemann, “Integrating Big Data: A Semantic Extract-Transform-Load Framework,” *Computer (Long Beach, Calif.)*, vol. 48, no. 3, pp. 42–50, Mar. 2015.
- [197] G. Bello-Orgaz, J. J. Jung, and D. Camacho, “Social big data: Recent achievements and new challenges,” *Inf. Fusion*, vol. 28, pp. 45–59, 2016.
- [198] B. T. Hazen, C. A. Boone, J. D. Ezell, and L. A. Jones-Farmer, “Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications,” *Int. J. Prod. Econ.*, vol. 154, pp. 72–80, 2014.
- [199] B. Hutchinson, L. Deng, and D. Yu, “Tensor Deep Stacking Networks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1944–1957, 2013.
- [200] J. Dean *et al.*, “Large scale distributed deep networks,” in *Advances in neural information processing systems*, 2012, pp. 1223–1231.
- [201] X.-W. Chen and X. Lin, “Big data deep learning: challenges and perspectives,” *IEEE access*, vol. 2, pp. 514–525, 2014.
- [202] A. Coates, B. Huval, T. Wang, D. Wu, B. Catanzaro, and N. Andrew, “Deep learning with COTS HPC systems,” in *International Conference on Machine Learning*, 2013, pp. 1337–1345.
- [203] C. Zhang, P. Li, G. Sun, Y. Guan, B. Xiao, and J. Cong, “Optimizing fpga-based accelerator design for deep convolutional neural networks,” in *Proceedings of the 2015 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays*, 2015, pp. 161–170.
- [204] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A. Y. Ng, “Multimodal deep learning,” in *Proceedings of the 28th international conference on machine learning (ICML-11)*, 2011, pp. 689–696.
- [205] N. Srivastava and R. R. Salakhutdinov, “Multimodal learning with deep boltzmann machines,” in *Advances in neural information processing systems*, 2012, pp. 2222–2230.
- [206] L. Zhao, Q. Hu, and W. Wang, “Heterogeneous Feature Selection With Multi-Modal Deep Neural Networks and Sparse Group LASSO,” *IEEE Trans. Multimed.*, vol. 17, no. 11, pp. 1936–1948, 2015.
- [207] Q. Zhang, L. T. Yang, and Z. Chen, “Deep Computation Model for Unsupervised Feature Learning on Big Data,” *IEEE Trans. Serv. Comput.*, vol. 9, no. 1, pp. 161–171, 2016.
- [208] S. Wan and L. E. Banta, “Parameter incremental learning algorithm for neural networks,” *IEEE Trans. Neural Networks*, vol. 17, no. 6, pp. 1424–1438, 2006.
- [209] L. Zhao, J. Chen, F. Chen, W. Wang, C. T. Lu, and N. Ramakrishnan, “SimNest: Social Media Nested Epidemic Simulation via Online Semi-Supervised Deep Learning,” in *2015 IEEE International Conference on Data Mining*, 2015, pp. 639–648.
- [210] L. Yu, H. Chen, Q. Dou, J. Qin, and P. A. Heng, “Integrating Online and Offline Three-Dimensional Deep Learning for Automated Polyp Detection in Colonoscopy Videos,” *IEEE J. Biomed. Heal. Informatics*, vol. 21, no. 1, pp. 65–75, 2017.
- [211] G. Zhou, K. Sohn, and H. Lee, “Online incremental feature learning with denoising autoencoders,” in *Artificial Intelligence and Statistics*, 2012, pp. 1453–1461.
- [212] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, “Extracting and Composing Robust Features with Denoising Autoencoders,” in *Proceedings of the 25th International Conference on Machine Learning*, 2008, pp. 1096–1103.
- [213] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, “Stacked denoising autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion,” *J. Mach. Learn. Res.*, vol. 11, pp. 3371–3408, 2010.
- [214] R. Wang and D. Tao, “Non-Local Auto-Encoder With Collaborative Stabilization for Image Restoration,” *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2117–2129, May 2016.

- [215] B. Jan *et al.*, “Deep learning in big data Analytics: A comparative study,” *Comput. Electr. Eng.*, 2017.
- [216] G. Qi, Z. Zhu, K. Erqinhu, Y. Chen, Y. Chai, and J. Sun, “Fault-diagnosis for reciprocating compressors using big data and machine learning,” *Simul. Model. Pract. Theory*, vol. 80, pp. 104–127, 2018.
- [217] H. Cui, L. Zhang, R. Kang, and X. Lan, “Research on fault diagnosis for reciprocating compressor valve using information entropy and SVM method,” *J. Loss Prev. Process Ind.*, vol. 22, no. 6, pp. 864–867, 2009.
- [218] A. Widodo and B.-S. Yang, “Support vector machine in machine condition monitoring and fault diagnosis,” *Mech. Syst. Signal Process.*, vol. 21, no. 6, pp. 2560–2574, 2007.
- [219] G.-M. Xian and B.-Q. Zeng, “An intelligent fault diagnosis method based on wavelet packer analysis and hybrid support vector machines,” *Expert Syst. Appl.*, vol. 36, no. 10, pp. 12131–12136, 2009.
- [220] B.-S. Yang, W.-W. Hwang, D.-J. Kim, and A. C. Tan, “Condition classification of small reciprocating compressor for refrigerators using artificial neural networks and support vector machines,” *Mech. Syst. Signal Process.*, vol. 19, no. 2, pp. 371–390, 2005.
- [221] N. K. Verma, A. Roy, and A. Salour, “An optimized fault diagnosis method for reciprocating air compressors based on SVM,” in *2011 IEEE International Conference on System Engineering and Technology*, 2011, pp. 65–69.
- [222] V. T. Tran, F. AlThobiani, and A. Ball, “An approach to fault diagnosis of reciprocating compressor valves using Teager–Kaiser energy operator and deep belief networks,” *Expert Syst. Appl.*, vol. 41, no. 9, pp. 4113–4122, 2014.
- [223] M. Havaei *et al.*, “Brain tumor segmentation with deep neural networks,” *Med. Image Anal.*, vol. 35, pp. 18–31, 2017.
- [224] Z. Jiang, H. Zhang, Y. Wang, and S.-B. Ko, “Retinal blood vessel segmentation using fully convolutional network with transfer learning,” *Comput. Med. Imaging Graph.*, vol. 68, pp. 1–15, 2018.
- [225] A. F. M. Oliveira, S. R. M. Pereira, and C. A. B. Silva, “Retinal Vessel Segmentation based on Fully Convolutional Neural Networks,” *Expert Syst. Appl.*, p. , 2018.
- [226] S. Moccia, E. De Momi, S. El Hadji, and L. S. Mattos, “Blood vessel segmentation algorithms — Review of methods, datasets and evaluation metrics,” *Comput. Methods Programs Biomed.*, vol. 158, pp. 71–91, 2018.
- [227] G. Pan, Z. Wu, and L. Sun, “Liveness detection for face recognition,” in *Recent advances in face recognition*, InTech, 2008.
- [228] J. Li, Y. Wang, T. Tan, and A. K. Jain, “Live face detection based on the analysis of fourier spectra,” in *Biometric Technology for Human Identification*, 2004, vol. 5404, pp. 296–304.
- [229] Y. A. U. Rehman, L. M. Po, and M. Liu, “LiveNet: Improving features generalization for face liveness detection using convolution neural networks,” *Expert Syst. Appl.*, vol. 108, pp. 159–169, 2018.
- [230] W. Bao, H. Li, N. Li, and W. Jiang, “A liveness detection method for face recognition based on optical flow field,” in *Image Analysis and Signal Processing, 2009. IASP 2009. International Conference on*, 2009, pp. 233–236.
- [231] G. Pan, L. Sun, Z. Wu, and S. Lao, “Eyeblink-based anti-spoofing in face recognition from a generic webcam,” in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, 2007, pp. 1–8.
- [232] K. Kollreider, H. Fronthaler, and J. Bigun, “Non-intrusive liveness detection by face images,” *Image Vis. Comput.*, vol. 27, no. 3, pp. 233–244, 2009.
- [233] Wikipedia, “List of writing systems,” https://en.wikipedia.org/wiki/List_of_writing_systems. .
- [234] G. Alvarez, B. Sheffer, and M. Bryant, “Offline Signature Verification with Convolutional Neural Networks,” 2016.

M. Mutlu YAPICI

He is a Instructor at Ankara University. He graduated from Faculty of Technical Education, Electronics and Computer Education Department in 2010. As a second bachelor degree, he completed computer engineering program in 2016. He graduated from MSc program of Informatics Institute in 2012. His master thesis subject is Development of Course Scheduling Software Using Genetic Algorithm. He is a student at Faculty of Technology Computer Engineering PhD. His research interests lie in image processing and machine and deep learning and their applications on signature verification, face recognition, autonomus vehichles and medical image analysis.

Adem TEKEREK

He is a Doctor Instructor at Gazi University, Department of Information Technology. He graduated from Faculty of Technical Education, Electronics and Computer Education Department in 2007. He graduated from MSc program of Informatics Institute in 2010. His master thesis is about Content Management Systems. He graduated from PhD program of Informatics Institute in 2016. His PhD thesis is about Web Application Firewall algorithms. He has published 12 papers on computer sciences. His research is data mining, machine learning and their applications especially on information security.

Nurettin TOPALOĞLU

Nurettin Topaloglu is currently an Associate Professor of Computer Engineering Department of Technology Faculty at Gazi University in Turkey. He received BSc in Electronics, MSc in Electronics and Computer Education and PhD in Electric Education. His research interests are computer architecture and organization, informatics technologies and information security. He developed the educational software Visual 6502 Microprocessor Simulator for teaching computer architecture. He is a writer of Microprocessors and Assembly Language, and x86 Microprocessor Architecture and Assembly Language in Turkish.