Futuristic Machine Learning Techniques for Diabetes Detection

Pavan kumar Panakanti¹, Sammulal Porika ², SK Yadav³

¹Research Scholar Shri JJT University Rajasthan,
²Associate professor JNTUH Hyderabad,
³Research Director Shri JJT University, Rajasthan

http://doi.org/10.26782/jmcms.2019.10.00034

Abstract

Diabetes detection has become an important task for medical practitioners in India and abroad. Researchers and scientists have been working on this problem actively. Machine learning has been contributing majorly to systems, techniques and solutions for diabetes detection problem. Yet there are challenges which remain to be addressed.

Recently convolution based machine learning techniques have evolved to give efficient results in various domains. They have shown applicability over range of problems. So here recent architectures of Convolution based machine learning models like Convolutional Neural Networks (CNN) and Capsule Networks (CapsNet) are discussed. Also, application of these recent models is presented here.

Additionally, challenges faced by current Diabetes detection systems are discussed. Along with these challenges CapsNet architecture for text analytics is presented. This CapsNet architecture is closest to Diabetes detection problem in terms of structure and arrangement of data to be handled. Thus in future this architecture and its variants can be applied for Diabetes detection.

Keywords: Diabetes detection, Convolutional Neural Networks, CNN, Capsule Networks, CapsNet

I. Introduction

Diabetes detection has become an important task for medical practitioners in India and abroad [VI] [VIII]. Latest advances in Machine Learning come to their rescue. In past, researchers have explored various techniques regarding diabetes detection. Additionally they have explored various behavioral and medical parameters actively contributing to diabetic scenario in a human body [VII]. Also improvements in Machine Learning promise accurate results consistently. In many medical facilities such Machine Learning algorithms or systems are implemented to provide patients with predictive alerts about patient’s health.
This work aims to survey existing models for diabetes detection, their advantages and shortcomings. Also, latest Machine Learning models like CapsNet which are proved to be successful in various applications are analyzed here. In future such model can be applied to diabetes detection for better results. Another interesting model, Convolutional Neural Networks was recently used for diabetes detection [XXXXVII]. It was a successful experiment and has shown significant improvement in results. Also, application of latest models on Text processing is discussed in this survey. In text processing problems, CNN and CapsNet architectures are modified to address inputs other than images. These modifications are useful for application of CNN or CapsNet to diabetes detection problem. Also [XXXXVII] has described details of implementation of CNN for diabetes. So, such latest models can be used for diabetes detection in near future.

II. Diabetes Detection using Machine Learning

Machine learning has shown its efficient decision making in various medical applications. Diabetes detection is also one of them. Many researchers have worked on this problem in last few decades. Type two diabetes detection is discussed by Work [XIX]. Here improved k-means algorithm coupled with logistic regression is used for diabetes detection. Here first k-means algorithm performs on given data to form two clusters. One cluster represents positive class and other represents negative class. Records which were not properly clustered were separated. Then tuned data was given to logistic regression. This approach was compared with various existing algorithms. This novel approach has proved better than others in comparison. In work [IV] shows application and use of feature transformation techniques like Principle Component Analysis [I] in various disease predictions. This work covers varieties of disease prediction problems in medical domain. Feature transformation helps transform data into independent features resulting in better performance. Additionally, in [XXXI] recent survey of machine learning algorithms for medical field is presented. Work [XXXXIV] discusses classification problem of diabetic images for maculopathy. Here images are classified using fuzzy logic and clustering method based techniques. Here images are classified to detect problems in eye retina. Similar problem is tackled by authors in [XXXIII]. Here deep learning methods applied for diabetic retinopathy are studied. Deep learning has sufficiently improved to provide state of the art results for diabetic retinopathy. Another aspect of diabetes is explored by [XVIII]. Here, family history data is studied for probability of diabetes. Historical and family data together is used for type 2 diabetes detection. For this purpose this work uses altering decision trees. Such weak classifiers can be improved in efficiency using bagging or boosting techniques. Work [XXVIII] presents a hybrid approach involving Least squares Support Vector Machines (LS-SVM) and Generalized Discriminant Analysis (GDA). Least squares SVM (LS-SVM) have typical feature of linear equation based training differing from SVMs having quadratic equation based training. This enables swift and efficient calculations of support vectors. Another major difference is in LSSVM lagrangian multipliers can be either positive or negative while in standard SVMs they can be
only positive. This new model has proved better compared to SVM for diabetes detection. Also here authors have clubbed it with Generalized Discriminant Analysis (GDA). GDA is used as method for feature selection. Most important features depicting latent patterns are selected using GDA. Then they are processed by LSSVM for diabetes detection.

Novel technique called General Regression Neural Networks is proposed in [XXVI]. It is evolved from kernel regression. As standard regression algorithm this also provides an output in continuous variable form. For a classification problem this output needs to be processed further to decide predicted class. In training it doesn't require iterative handling like in other neural networks. This model provides estimate of joint probability distribution function for independent and dependent variable. On other hand, work [XXVII] presents Adaptive neuro fuzzy inference system (ANFIS) for detection of diabetes. In this system neural network is clubbed together with fuzzy logic. First preprocessed data is applied with fuzzy logic and then supplied to the neural network [XXIII]. In this work preprocessing and component formation is done using Principle Component Analysis (PCA).

In recent times many other machine learning models have been proposed for various applications. These models have achieved significant results in respective domains.

III. Latest Machine Learning Models

III.i. Naive Bayes with Genetic Computing

This approach is used by [X]. Here Diabetes detection system uses Genetic algorithm for feature selection. Then after feature selection, Naive Bayes approach is used to predict diseased patients. This is a novel application of Genetic Computing and Naive Bayes together for diabetes detection. One such model was applied on skin disease detection previously [V]. Results show significant improvement in feature selection due to Genetic computing which results in overall better performance of algorithm.

III.ii. Reinforcement Learning Approach

In survey [XXIV] highlights use of Reinforcement learning in diabetes detection field. Detailed description is provided in [XIV]. These systems independently interact with environment. These interactions lead to either achievement of positive or negative learning, i.e. reward or penalty. Based on this model system modifies its approach to interact with an environment [XXIX].

III.iii. Recurrent Neural Network (RNN) Approach

Recurrent Neural Networks (RNN) [XXXVIII] and Long or Short Term Memory (LSTM) [XXXXIII] has different approaches in recurrent learning. Here recurrent connections in neural networks allow finding deeper hidden patterns in input data. In work [XXXXVII] ICU sensor data is collected from patients. These inputs are processed using LSTM and RNN to predict critical illness pattern in a patient. Here it is shown that LSTM with minimal training outperforms all other models.
III.iv. Convolutional Neural Networks (CNN)

LeCun proposed this model for the first time as LeNet [XXXXXI]. Then subsequently it was modified and used in various applications. This model has been one of the pioneers in advancement of deep learning for image processing over GPUs. In recent times use GPU based CNNs have significantly captured various application area of image processing. CNNs have been remarkable in applications like face detection, object detection, object classification, and video processing [XXXXXII].

III.v. Capsule Network

Capsule Networks (CapsNet) were introduced recently as novel architecture in Neural Networks [XXXXV]. Sequence of multiple convolutional kernels bundled together is known as a capsule. Capsule Networks consists of large number of these capsules [XV]. Each capsule operates at local level of features. Global understanding of features is done by communication between various capsules using routing paths [XXXXXVI]. Figure 1 shows a sample basic architecture of CapsNet as mentioned in [XXXXV]. This architecture overcomes some drawbacks of CNNs keeping all its advantages. So, in very less time this model has become popular in varieties of applications. Also it can be applied in any domain where CNNs are applied.

![Figure 1: Basic CapsNet Architecture](image)

IV. Futuristic approaches for Diabetes Detection

Current diabetes detection models have limited support for doctors. They suffer from class imbalance problems and poor explainability problems [XXXXXII]. So, better and more robust approaches for diabetes detection are required. So, here latest approaches of CNN and CapsNet are presented which can be applied easily to Diabetes Detection problem.

IV.i. Latest Convolution Based Architectures

Convolution based Neural Networks have consistently proved their worth. In various applications they have been successful so application of such architectures for Diabetes detection is an interesting filed to be explored. There are very few nearly
none Convolution based implementations directed towards Diabetes detection. In figure 2 Convolution based architectures of neural network are listed. These approaches are discussed here in details.

IV.i.a. Latest CapsNet based Architectures and Applications

In Work [III], CapsNets are studied for better transparency of operations. It is argued that vectorized output of a capsule has likelihood information in regards to features. This enables ease in explainability of capsule outputs. Also, authors propose how to improve transparency of CNNs using capsules in CNNs.

SegCaps, capsule network with Convolution-Deconvolution Capsules is presented in [XXXIX]. In this work, high demand for resources by CapsNets is handled. Proposed model uses routing within limited parent capsules as one way to restrict computations. On other hand it reduces size of shared matrices as well. All together affects drastically on memory and computational demand along with number of parameters being processed within layers. Deconvolution capsules use transformational convolution for keeping the global context for routes and parameters. Comparison shows how SegCaps operate on LUNA16 dataset with same effect with minimal parameters and resources for object segmentation.

Next architecture is Sparse Unsupervised Capsules from [XIII]. Supervised capsule networks have limitations in depth. Unsupervised capsule networks can be formed deeper layers. Here sparsely connected unsupervised capsules are used for classification problem. System is tested on AFFNIST dataset without training on it. This has given better benchmark results.

Another interesting advancement is brought out by [XXXV]. Here unknown entities in text are handled efficiently. Bidirectional gated recurrent units are used before applying CapsNet. Here recurrent units find out possibility of relationship between entities. Then CapsNet uses this information to decide whether there exists hypernymy in compound entity or not.

Fast CapsNet architecture is proposed in [II]. Here performance of CapsNet is improved using limited routing coefficients in each primary capsules. Single pixel in a primary capsule corresponds to single routing path. So, number of routing paths is drastically reduced. Also, another important contribution of this work is convolutional decoder. This decoder helps in decoding samples with high visual variability in a restricted region.

Next, CapsuleGAN architecture [XXI] is novel step in generative adversarial networks (GAN). Here GAN objective function is also updated. Proposed model outperforms traditional convolution based GAN in semi-supervised learning for MNIST and CIFAR10 datasets.

Another work [IX] proposes CapsNet with Hit-or-Miss layer called as Hit Net. Here new loss function called centripetal loss is proposed here. Hit or Miss Capsules incorporate this loss function and they reduce losses when capsule is far from or too close to a class of sample supplied. Also, this architecture makes sure that gradient goes to zero only when class of capsule and sample is matched. Here authors claim that batch normalization process works better than method proposed for original CapsNet.
Work [XXXXVIII] is another variation of CapsNet proposed recently. Here Long Short Term Memory Neural Network is applied for evaluation. LSTM helps in predicting future position or transformation of the object based on history. Additionally new routing is used to approximate an object's positions. Work has added another valuable part of evaluating transfer of learning for different datasets. Next is Recurrent Neural Networks (RNN) based Capsule network from [XXXXXIII]. Here capsules are formed based on RNN. This architecture allows interpreting time series data like language words. Also, it has proved efficient on sentiment analysis without prior knowledge of sentiment of words.

Brain tumor type classification is considered as one of the toughest tasks in medical field. In [XXXIV] CapsNet are applied for this problem. This work uses CapsNets as they are thorough in handling various angles or scaling or other affine transformations on images. This work adapts CapsNet for brain tumor detection. Also, addresses over fitting problem occurring on real dataset. Here CNN and CapsNets are compared over complete brain and segmented tumor images. In both type of images CapsNet perform better than CNN. Another work [II] applies latest architecture of CapsNet for lung cancer detection.

Capsule Network Performance on Complex Data is studied in [XV]. Authors here address problem of complex data from datasets like CIFAR10. Various models are formed here. Models of capsule networks are based on Ensemble layer, number of convolution layers and reconstruction scaling. Stacking of capsule layer is also applied as one model. All these models are compared with standard Convolution Neural Network and CpasNet. Here 7-model ensemble with additional convolutional layer is best with capsule network.

In thesis [XXX], detailed study of CapsNet is presented. Then it is applied for object comprehension in various views. Here 3D objects with various view points are given for training. Then in training familiar and non-familiar view points are provided. CapsNet show significant reduction in error rate compared to traditional CNN. On other hand [XXXXXIII] applied CapsNet for Object localization and motion transfer. Sentiment Analysis is another area where CapsNet is applied [XXXXXXXIII]. CapsNet has proved to be very efficient in this application on three datasets. Here CapsNet is modified according to RNN architecture to support sentiment analysis application.

**IV.i.b. Latest CNN Based Architectures**

Deep convolutional approach is used for speech recognition [XX]. Traditional Hidden Markov Approach is considered for comparison with proposed system. Also, limitations of decision trees to solve this problem are explained here. Input text features and speech features are correlated to find patterns. In results, Hidden Markov Models with various alpha values and Deep Neural Network with 4 layers and varying number of neurons are compared. Deep Neural Network is found to be better. Also, p-value is significant for the given test.

Another work describes how CNN can be used for text classification is [XXXVII]. In this work, text data with information about term order is processed using CNN. This work has presented details of how to use text as input to CNN. It also describes text pooling in CNN.
Text recognition from natural images using CNN is presented in [XXXII]. It works on regional processing. Various regions from image segments are processed and ranked for recognition of text. Highest ranked regions are used for text recognition. Uniqueness of this work is use of complete synthetic training data. Detailed experiments are done here to show effectiveness of proposed system. In work [XII], sensitivity analysis of one layer CNN is done with details. It deals with problem of hyper parameter tuning. Experiments are focused on sensitivity of CNN model towards various parameters. Even effect of pooling strategy, dropout rate is studied here. One major observation here is about filter region size, which affects performance vividly. Novel activation function is discussed in work [XXXXIX]. Here, activation function is made closer to biological activation function. Also, it is compared with state of the art neural network setups. This activation function significantly contributes to improvement in neural network performance. This activation function also addresses sparsely in signal zero.

Figure 2: Latest Convolution Based Architectures

In [XXXXX] work, named AOGNet, grammar approach is combined with Convolutional networks. AND-operation represents concatenation of lower node outputs and OR-operation provides alternative combination paths. Additionally, terminal nodes are defined which take only sliced input. Proposed model is tested on

Copyright reserved © J. Mech. Cont.& Math. Sci.

Pavan kumar Panakanti et al
benchmark datasets CIFAR-10, CIFAR-100 and ImageNet-1K. They have shown better performance than Residual Nets and Dense Net over classification and object detection tasks.

Work [XXXXVI], proved that CNN can be trained to adapt to any domain of image by training on single image of that domain. Here once training is done using soriensis image data. Then other domain data is trained only using single image and comparable results are achieved.

Dilated convolutional networks or atrous convolution Networks are proposed in [XI]. Dilation refers to inserting holes in intermediate layer outputs to increase resolution. Better feature extraction is possible using this method.

Squeezenet architecture is proposed by authors of [XVII]. Here intermediate layers use 1 by 1 convolution to reduce size of input. Then again 3 by 3 convolutions for expanding. By this process memory required and number of parameters required are reduced by 50 times. Results show squeezenet performs with same accuracy as convolutional networks.

Recurrent residual CNN with U-nets is proposed in [XXXXXV]. It uses power of recurrent, residual and U-nets neural networks together in CNN for image segmentation. Residual layer plays key role in training. Better feature recognition is possible using recurrent residual convolution layer. Additionally this allows better U-net architecture which outperforms simple U-net and Residual U-nets. The model is tested on various datasets for segmentation task.

Another work is [XXII] Images with poor ambient light are improved by using CNN. First input image is converted to a map and then ratio of image is used along with the map to dehaze the image using residual CNN.

Novel idea of Regional CNN is presented in [XXXXXIV]. Ultrasonic images of various sizes and orientations are successfully handled here for thyroid cancer detection. Here Regional Convolutional networks are improved by using multi-layer linking and multi-scale input. Experiments conclude with best number of iterations for high number of training samples, comparison of models based on MAP, and ROC comparison. Proposed fast detection of thyroid cancer using region based CNN has highest area under ROC curve compared to other CNN models.

Global sum pooling (GSP) is novel method for pooling layer in CNNs [XXXXXI]. Results here show that GSP performs better than traditional Global Average Pooling.

Evolving CNN is another architecture proposed by authors of [XXXVI]. Initially with small set of manually tagged dataset CNN is trained. Then by artificial methods images are collected from various sources. Using unsupervised ways they are tagged. Later using these new images CNN is trained. If there is significant increase in performance beyond threshold then CNN weights are updated to accept new training. In all this process a standard validation image set is kept constant for verification benchmark. This framework of evolution is applied over four different CNN models. Over all models it has affected positively.

Another novel proposal is HSI-CNN for Hyper spectral Images from [XVI]. Authors here have developed an additional preprocessing layer before convolutional neural network. This layer transforms hyper spectral image using one dimensional convolution. Then features obtained after one dimensional convolution are arranged
in form of a two dimensional matrix. This matrix is finally fed as input to convolutional layer. This method has obtained promising results onto various datasets.

In [XXXXVII], authors propose innovative approach of adaptive Fuzzy system with CNN. Here data fuzzification is applied before providing data to CNN. Also, work has used private data for its experiments. It is applied on solving Diabetes detection problem. Results show significant improvement over previous methods of diabetes detection.

V. Scope of Application of CapsNet (Convolutional Approach) in Diabetes Detection

Survey [XXXXXII] presents detailed challenges faced by latest models of machine learning in Diabetes detection. Mainly these challenges include handling large size data, explainability of machine learning algorithms and class imbalance problem. Large size data or big data is available due to IoT like technologies. Challenge is to have systems cope up these huge sums of data.

As per [XXXXXII] many existing machine learning models fail to overcome this challenge. CapsNets have proved to handle huge sizes of image data [XXXXV]. Also, as discussed previously CapsNet has been applied in various domains successfully. Next challenge is about explainability of machine learning models. In [III] CapsNet is modified to have better explainability. Authors have shown how proposed novel approach improves applicability and explainability of CapsNet. So, CapsNet can address this challenge as well.

Another challenge is linking historical records of same patient with current findings. Here problem is composite of time series data and normal statistical data. So, for handling data simulated by time series CapsNet can be enabled with RNNs as presented by [XXXXXIII]. Such novel architecture of CapsNet hand easily handle hybrid problems in disease detection field.

Data imbalance or data unavailability challenge is not only limited to Diabetes detection, but its concern of complete medical field [XXV][XXXX]. In this case data about certain class is present in ample amount and other class or few classes of dataset have limited samples. Such data can be used by re-sampling or sub-sampling but intern it loses its value. This is a challenge which needs to be addressed by future models.

VI. Conclusions

In this paper advances in Diabetes detection and challenges are analyzed. Also, CapsNet a novel convolution based approach and its applications are discussed. CapsNet's application in text analytics is discussed in details in light that such architectures can easily be applied on Diabetes Detection data due to structural resemblance. Finally, scope of application of CapsNet to address current challenges in Diabetes detection presented. In future, CapsNet may be applied over Diabetes Detection problem. Also, in future machine learning models in Diabetes Detection system need to have better explainability. This will improve their usability from point of view of doctors.

Copyright reserved © J. Mech. Cont.& Math. Sci.
Pavan kumar Panakanti et al

495
References

I. A. Mackiewicz and W. Ratajczak, "Principal components analysis (pca)," Computers and Geosciences, vol. 19, pp. 303-342, 1993.

II. A. Mobiny and H. Van Nguyen, "Fast capsnet for lung cancer screening," arXiv preprint arXiv:1806.07416, 2018.

III. A. Shahroudnejad, A. Mohammadi, and K. N. Plataniotis, "Improved explainability of capsule networks: Relevance path by agreement," arXiv preprint arXiv:1802.10204, 2018.

IV. B. D. Kanchan and M. M. Kishor, "Study of machine learning algorithms for special disease prediction using principal of component analysis," in Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), 2016 International Conference on, pp. 5-10, IEEE, 2016.

IV. B. Sierra and P. Larranaga, "Predicting survival in malignant skin melanoma using bayesian networks automatically induced by genetic algorithms. an empirical comparison between different approaches," Artificial Intelligence in Medicine, vol. 14, no. 1-2, pp. 215-230, 1998.

V. C. Bennett, M. Guo, and S. Dharmage, "Hba1c as a screening tool for detection of type 2 diabetes: a systematic review," Diabetic medicine, vol. 24, no. 4, pp. 333-343, 2007.

VI. C. Willi, P. Bodenmann, W. A. Ghali, P. D. Faris, and J. Cornuz, "Active smoking and the risk of type 2 diabetes: a systematic review and meta-analysis," Jama, vol. 298, no. 22, pp. 2654 -2664, 2007.

VII. D. B. Carr and S. Gabbe, "Gestational diabetes: detection, management, and implications," Clinical Diabetes, vol. 16, no. 1, pp. 4-12, 1998.

IX. Deliege, A. Cioppa, and M. Van Droogenbroeck, "Hitnet: a neural network with capsules embedded in a hit-or-miss layer, extended with.

X. D. K. Choubey, S. Paul, S. Kumar, and S. Kumar, "Classification of pima indian diabetes dataset using naive bayes with genetic algorithm as an attribute selection," in Communication and Computing Systems: Proceedings of the International Conference on Communication and Computing System (ICCCCS 2016), pp. 451-455, 2017.
XI. Dolz, X. Xu, J. Rony, J. Yuan, Y. Liu, E. Granger, C. Desrosiers, X. Zhang, I. B. Ayed, and H. Lu, "Multi-region segmentation of bladder cancer structures in mri with progressive dilated convolutional networks," arXiv preprint arXiv:1805.10720, 2018.

XII. dos Santos and M. Gatti, "Deep convolutional neural networks for sentiment analysis of short texts," in Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, pp. 69-78, 2014.

XIII. D. Rawlinson, A. Ahmed, and G. Kowadlo, "Sparse unsupervised capsules generalize better," arXiv preprint arXiv:1804.06094, 2018.

XIV. E. Alpaydin, "Introduction to machine learning", MIT press, 2014.

XV. E. Xi, S. Bing, and Y. Jin, "Capsule network performance on complex data," arXiv preprint arXiv:1712.03480, 2017.

XVI. F. Liang, H. Liu, X. Wang, and Y. Liu, "Hyperspectral image recognition based on artificial neural network," NeuroQuantology, vol. 16, no. 5, 2018.

XVII. F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: Alexnet-level accuracy with 50x fewer parameters and< 0.5 mb model size," arXiv preprint arXiv:1602.07360, 2016.

XVIII. H. Uemura, A. A. Ghaibeh, S. Katsuura-Kamano, M. Yamaguchi, T. Bahari, M. Ishizu, H. Moriguchi, and K. Arisawa, "Systemic inflammation and family history in relation to the prevalence of type 2 diabetes based on an alternating decision tree," Scientific reports, vol. 7, p. 45502, 2017.

XIX. H. Wu, S. Yang, Z. Huang, J. He, and X. Wang, "Type 2 diabetes mellitus prediction model based on data mining," Informatics in Medicine Unlocked, vol. 10, pp. 100-107, 2018.

XX. H. Ze, A. Senior, and M. Schuster, "Statistical parametric speech synthesis using deep neural networks," in Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, pp. 7962-7966, IEEE, 2013.

XXI. Jaiswal, W. AbdAlmageed, and P. Natarajan, "Capsulegan: Generative adversarial capsule network," arXiv preprint arXiv:1802.06167, 2018.

XXII J. Li, G. Li, and H. Fan, "Image dehazing using residual-based deep cnm," IEEE Access, 2018.

XXIII J.-S. Jang, "Anfis: adaptive-network-based fuzzy inference system," IEEE transactions on systems, man, and cybernetics, vol. 23, no. 3, pp. 665-685, 1993.
XXIV  Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, and I. Chouvarda, "Machine learning and data mining methods in diabetes research," Computational and structural biotechnology journal, vol. 15, pp. 104-116, 2017.

XXV  Khan, I. Yaqoob, I. A. T. Hashem, Z. Inayat, M. Ali, W. Kamaleldin, M. Alam, M. Shiraz, and A. Gani, "Big data: survey, technologies, opportunities, and challenges," The Scientific World Journal, vol. 2014, 2014.

XXVI  K. Kayaer and T. Yldrm, "Medical diagnosis on pima indian diabetes using general regression neural networks," in Proceedings of the international conference on artificial neural networks and neural information processing (ICANN/ICONIP), pp. 181-184, 2003.

XXVII  K. Polat and S. Gunes, "An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease," Digital Signal Processing, vol. 17, no. 4, pp. 702-710, 2007.

XXVIII  K. Polat, S. Gunes, and A. Arslan, "A cascade learning system for classification of diabetes disease: Generalized discriminant analysis and least square support vector machine," Expert systems with applications, vol. 34, no. 1, pp. 482-487, 2008.

XXIX  L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," Journal of artificial intelligence research, vol. 4, pp. 237-285, 1996.

XXX  M. Engelin, "Capsnet comprehension of objects in different rotational views," divaportal.org, 2018.

XXXI  M. Fatima and M. Pasha, "Survey of machine learning algorithms for disease diagnostic," Journal of Intelligent Learning Systems and Applications, vol. 9, no. 01, p. 1, 2017.

XXXII  M. Jaderberg, K. Simonyan, A. Vedaldi, and A. Zisserman, "Reading text in the wild with convolutional neural networks," International Journal of Computer Vision, vol. 116, no. 1, pp. 1-20, 2016.

XXXIII  M. Singh, "Classification of diabetic retinopathy stages using deep learning," Scientific reports, 2018.

XXXIV  P. Afshar, A. Mohammadi, and K. N. Plataniotis, "Brain tumor type classification via capsule networks," arXiv preprint arXiv:1802.10200, 2018.
XXXV  Q. Wang, T. Ruan, Y. Zhou, D. Gao, and P. He, "An attention-based bi-gru-capsnet model for hypernymy detection between compound entities," arXiv preprint arXiv:1805.04827, 2018.

XXXVI  Q. Xuan, H. Xiao, C. Fu, and Y. Liu, "Evolving convolutional neural network and its application in fine-grained visual categorization," IEEE Access, 2018.

XXXVII  R. Johnson and T. Zhang, "Effective use of word order for text categorization with convolutional neural networks," arXiv preprint arXiv:1412.1058, 2014.

XXXVIII  R. J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks," Neural computation, vol. 1, no. 2, pp. 270-280, 1989.

XXXIX  R. LaLonde and U. Bagci, "Capsules for object segmentation," arXiv preprint arXiv:1804.04241, 2018.

XL  Safran, M. Bloomrosen, W. E. Hammond, S. Labkoff, S. Markel-Fox, P. C. Tang, and D. E. Detmer, "Toward a national framework for the secondary use of health data: an american medical informatics association white paper," Journal of the American Medical Informatics Association, vol. 14, no. 1, pp. 1-9, 2007.

XLI  S. Aich and I. Stavness, "Object counting with small datasets of large images," arXiv preprint arXiv:1805.11123, 2018.

XLII  Schmidhuber, "Deep learning in neural networks: An overview," Neural networks, vol. 61, pp. 85-117, 2015.

XLIII  S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735-1780, 1997.

XLIV  S. Ibrahim, P. Chowriappa, S. Dua, U. R. Acharya, K. Noronha, S. Bhandary, and H. Mugasa, "Classification of diabetes maculopathy images using data-adaptive neuro-fuzzy inference classifier," Medical & biological engineering & computing, vol. 53, no. 12, pp. 1345-1360, 2015.

XLV  S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," in Advances in Neural Information Processing Systems, pp. 3856-3866, 2017.

XLVI  S. Valverde, M. Salem, M. Cabezas, D. Pareto, J. C. Vilanova, L. Ramio-Torrenta, A. Rovira, J. Salvi, A. Oliver, and X. Llado, "One-shot domain adaptation in multiple sclerosis lesion segmentation using convolutional neural networks," arXiv preprint arXiv:1805.12415, 2018.

XLVII  T. Deshmukh and H. Fadewar, "Fuzzy deep learning for diabetes detection," in Computing, Communication and Signal Processing, pp. 875-882, Springer, 2019.
XLVIII. W. Liu, E. Barsoum, and J. D. Owens, "Object localization and motion transfer learning with capsules," arXiv preprint arXiv:1805.07706, 2018.

XLIX. X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in Proceedings of the fourteenth international conference on artificial intelligence and statistics, pp. 315-323, 2011.

L. X. Li, T. Wu, X. Song, and H. Krim, "Aognets: Deep and-or grammar networks for visual recognition," arXiv preprint arXiv:1711.05847, 2017.

LI. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, 1998.

LII. Y. Sebastian, X. T. Tiong, V. Raman, A. Y. Y. Fong, and P. H. H. Then, "Advances in diabetes analytics from clinical and machine learning perspectives," International Journal of Design, Analysis and Tools for Integrated Circuits and Systems, vol. 6, no. 1, pp. 32-37, 2017.

LIII. Y. Wang, A. Sun, J. Han, Y. Liu, and X. Zhu, "Sentiment analysis by capsules," in Proceedings of the 2018 World Wide Web Conference on World Wide Web, pp. 1165-1174, International World Wide Web Conferences Steering Committee, 2018.

LIV. Y. Wang, W. Ke, and P. Wan, "A method of ultrasonic image recognition for thyroid papillary carcinoma based on deep convolution neural network," NeuroQuantology, vol. 16, no. 5, 2018.

LV. Zahangir Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, "Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation," arXiv preprint arXiv:1802.06955, 2018.

LVI. Z. Chen and D. Crandall, "Generalized capsule networks with trainable routing procedure," arXiv preprint arXiv:1808.08692, 2018.

LVII. Z. C. Lipton, D. C. Kale, C. Elkan, and R. Wetzel, "Learning to diagnose with lstm recurrent neural networks," arXiv preprint arXiv:1511.03677, 2015.