

Artificial intelligence in ophthalmology and healthcare: An updated review of the techniques in use

Jatinder Bali, Ojasvini Bali¹

Artificial intelligence (AI) refers to “the ability of a digital machine or computer to accomplish tasks that traditionally have required human intelligence.” These days, artificial intelligence is becoming popular in healthcare and more so in ophthalmology. It has shown promising results in diabetic retinopathy detection and referral. Recently, Indian data has depicted that the new algorithms can be generalized to the Indian population as well. An increased understanding of the tools is required especially by the practitioners and medical researchers so that they can contribute meaningfully to the development of the technology and not become mere data providers and data labelers. While AI is extensively being used by finance, marketing and travel industry, its application is more recent in medicine. The applications based on artificial intelligence have the potential to benefit all stakeholders in the healthcare industry.

Key words: Artificial intelligence, machine learning, neural networks, support vector machine, techniques

Access this article online

Website:

www.ijo.in

DOI:

10.4103/ijo.IJO_1848_19

Quick Response Code:



Artificial intelligence (AI) refers to “the ability of a digital machine or computer to accomplish tasks that traditionally have required human intelligence.” Machine learning deals with the computing machine’s or system’s ability to teach or improve itself without explicit programming for each improvement, using experience and methods of forward chaining of algorithms derived from backward derivation of functions from training data. Within the domain of machine learning there is a niche called deep learning (DL), which answers highly abstract problems using self-learning algorithms like artificial neural networks.^[1] This article briefly discusses the main artificial intelligence tools used in healthcare.

Methods

A search for literature was made using keywords “Artificial Intelligence, techniques, tools, healthcare, ophthalmology, algorithms” in PubMed, Web of Science Core Collection and Google Scholar. The relevant articles which discussed different techniques in use in relation to ophthalmology were shortlisted. Thereafter, the main techniques in use were tabulated. The papers which discussed the same or overlapping research were removed. From a total of 72 articles only 17 were found to be of consequence to be included. Then the grey literature was manually searched for additions. If a firm recommendation of the use of the AI technique existed in peer-reviewed literature, only then it was added to the discussion. The results were subsequently checked against the facts from other industry reports. In case of discordance of reports about the use of the technique, the medical literature was to gain precedence over literature from engineering and other branches as per decided protocol; the discordance would have been highlighted in the discussion. However, the need for reporting such an event did not arise.

Ophthalmologist I/C, Narela Polyclinic and Health Complex, Delhi,
¹Maulana Azad Medical College, Delhi, India

Correspondence to: Dr. Ojasvini Bali, 55-D, Third Floor, DDA Flats, Kalidas Road, Gulabibagh, Delhi - 110007, India. E-mail: drjatinderbali@gmail.com

Received: 12-Oct-2019

Revision: 05-Jul-2020

Accepted: 05-Jul-2020

Published: 15-Dec-2020

Application of AI

The main areas where AI is being applied in healthcare are:

- Mass screening
- Diagnostic imaging
- Laboratory data
- Electro-diagnosis
- Genetic diagnosis
- Clinical data
- Operation notes
- Electronic health records
- Records from wearable devices^[2,3]

AI Devices

AI devices are broadly of two main types:

- Machine Learning (ML) Techniques analyzing structured data like imaging, genetic and EP data and
- Natural Language Processing (NLP) Methods extracting information from unstructured data like clinical notes, medical journals and other unstructured medical data [Fig. 1].^[3,4]

Machine Learning Algorithms

The Machine Learning algorithms can be broadly divided into: unsupervised learning and supervised learning. Unsupervised learning helps feature extraction while supervised learning is used for predictive analytics after decreasing the principle

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

Cite this article as: Bali J, Bali O. Artificial intelligence in ophthalmology and healthcare: An updated review of the techniques in use. Indian J Ophthalmol 2021;69:8-13.

components for analysis. A semi-supervised mode is also proposed in recent times which bridges the two.^[4,5]

Increased computing power, larger amounts of data, real-time-online availability of databases and high availability of fast internet allows predictive-algorithm-development. Today, driverless cars are a distinct business opportunity. Other vistas have been opened by these developments.

In ophthalmology, interpretation of complex images has been achieved. In 2009, Retinopathy Online Challenge used competition fundus photographic sets from 17,877 patient visits of 17,877 people with diabetes who had not previously been diagnosed with DR consisting of two fundus images from each eye. These were compared using a single rater to that of a large computer-aided early DR detection project EyeCheck. The fundus photograph set of every visit was analyzed by single retinal expert. 792 out of these 17,877 sets had more than minimal DR which was the threshold for patient referral. Two algorithmic lesion detectors were used on the datasets separately and compared by standard statistical measures (area under the ROC curve as the main performance indicator). The two computerized lesion detectors demonstrated high agreement. At 90% sensitivity, the specificity of the EyeCheck algorithm was 47.7%. The specificity of the ROC-2009 winner algorithm was 43.6%. On comparing this with interobserver variability of the employed experts it was concluded that DR detection algorithms demonstrated maturity and the detection performance was not too different from the prevailing best clinical practices having reached the human intrareader variability limit. A combination of blood vessel parameters, microaneurysm detection, exudates, texture and distance between the exudates and fovea were accepted to be the most important features to detect the different stages of diabetic retinopathy.^[6] In 2008 Nayak *et al.* used area of the exudates, blood vessels and texture parameters analyzed through neural network to classify the fundus image into normal, non-proliferative DR (NPDR) and proliferative DR (PDR).^[7] The detection accuracy of 93% with sensitivity of 90% and specificity of 100% were reported. Support vector machine (SVM) classifier classified fundus images into normal, mild, moderate, severe and prolific DR classes with detection accuracy of 82% and sensitivity of 82% and specificity of 88%. Different software to grade the severity of hemorrhages and microaneurysms, hard exudates and cotton-wool spots of DR to classify DR have been developed and evaluated were able to identify hemorrhages and microaneurysms, hard exudates, and cotton wool spots.^[8]

Adjudication by experts has further improved the algorithms. Deep neural networks trained and validated using Gulshan *et al.*'s methods gave algorithms to grade retinal fundus photography images according to the International Clinical Diabetic Retinopathy (ICDR) severity scale. In this prospective study conducted with data from 2 tertiary eye care centers in South India, Aravind Eye Hospital and Sankara Nethralaya, the investigators trained the model to make a multiway classification of the 5-point ICDR grade. The algorithm was trained to make the various 5 point predictions. However, only 2 outputs, referable DR and referable DME, of the model were used to demonstrate that the automated DR system's findings generalized to this population of Indian patients in a prospective setting.^[9] The feasibility of using automated DR grading and referral system to screening programs was thus further proved in developing countries. Already cardiology has developed automated electrocardiographic analysis and

ophthalmology has used wave front analysis in implementing expert systems delivering results at par or beyond the capability of the human experts with years of clinical experience.

AI and AI-Enabled Machines

AI and AI-enabled machines are classified into seven main categories by two different types of classifications. The machines simulate human mind's thinking. Thus these machines can be:

1. Reactive Machine Systems e.g., Deep Blue chess playing system which defeated the world champion Kasparov in 1997.
2. Limited Memory Machine Systems which improve with experience e.g., chatbots like Tay.
3. Mind Theory Systems recognizing the need for other domains
4. Self-Aware AI that can actually plan for self-preservation.
5. Artificial Narrow Intelligence (ANI) which is focused on a narrow-range of abilities and processes tasks related to one single narrow task. All AI tools right now belong to this weak AI or ANI category e.g., Cortana, Siri or Google Assistant.
6. Artificial General Intelligence (AGI) can transfer knowledge from one domain into another on its own. It is also called strong AI or full AI. It can do "general intelligent action" and can also experience consciousness. We are a long distance away from something like this.
7. Artificial Super Intelligence (ASI) is the future of machine learning. It will surpass humans in all domains and all types of pursuits. Theoretically, it would be able to demonstrate creativity, emotions, engage in relationships, practice different art forms and take "bounded-rationality-decisions" with limited sets of information. Some glimpses in narrow domains can be seen of these even today. As of today, the integration and transference of domain expertise is not there. For example, the Chess or Go playing machines cannot scarcely do other things. But that has begun changing. However, we are still a long distance from anything as powerful as Artificial Super Intelligence.^[10,11]

AI Tools in Healthcare

Neural networks are not the only tools used for Healthcare AI. The main tools being used in the healthcare industry are briefly discussed below. This is not an exhaustive list as only the most common ones are being discussed here.

Linear regression

This models a linear relationship between a dependent variable or scalar response and one or more explanatory or independent variables. In simple linear regression, the relationship between the dependent and one explanatory variable is studied. In multiple linear regression, it is the relation with more than one explanatory variable. In multivariate linear regression, multiple correlated dependent variables are predicted using different explanatory variables. This relationship can be used for predictive modeling in a very narrow sense in statistics and is one of the simplest tools used for developing functions or equations that explain the results or dependent variable based on independent variables. The result can be viewed as

Dependent Variable = Constant+ [Slope x Independent Variable] + Error.^[3,4,12]

Any number of independent variables can be studied and the effort is to reduce the error.

Logistic regression

When the data has a binomial distribution, which means that it can be separated into two mutually exclusive groups like yes/no, pass/fail, alive/dead or healthy/sick, then the logistic model or logit model is used. It assigns a probability between 0 and 1 to the factors with the sum adding to one. Logistic regression statistical model uses a logistic function to model a binary dependent variable but other more complex analysis and permutations are possible. The logarithm of the odds for the dependent variable labeled "A" is a linear combination of one or more independent variables or "predictors". These independent variables can be continuous (any real value) or binary (yes/no) variables. The corresponding probability of the value labeled "B" can vary between 0 and 1.00 i.e. between 0-100%. This function that converts log-odds to probability is the logistic function. The unit of measurement for the log-odds scale is called logit. Similar rendition of models with a different sigmoid function is called the probit model. It is of use where categorical variables are used.^[3,4,12]

Naïve Bayes

Naive Bayes is used for constructing classifiers or models that assign class labels to examples of the problems like referable and non-referable. These examples get assigned as vectors of feature values with the class labels drawn from a finite set of feature values. Naive Bayes requires only a small number of training data to estimate the parameters for classification. There is a family of algorithms in naïve Bayes which is dependent on likelihood, probability before and probability after. The common principle for all naïve Bayes classifiers is that it assumes independence of features and each feature contributes to the classification regardless of any possible correlations between the features. A naïve Bayes classifier considers each of the features to contribute independently to the probability of its classification. In plain English, using Bayesian probability terminology, the above equation can be written as

$$\text{Posterior probability} = \frac{\text{prior probability} \times \text{likelihood}}{\text{evidence}}$$

Practically only the numerator of this fraction is important because the denominator the denominator is effectively constant (as it does not depend on C and the values of the individual features say x_i are given). The numerator is equivalent to the joint probability model given above. Naive Bayes is a probabilistic machine learning algorithm with wide application in heterogeneous classification tasks like labeling referable and even email spam. It is called 'Naive' because it assumes the features that go into the model are independent of each other. Changing the value of one feature does not directly influence or change the value of any of the other features used in the algorithm. Rev. Thomas Bayes (1702–61) gave us the elements of this and, therefore, it is named after him. It is popular because it can be coded easily and runs almost real time.^[3,4,13] It is scalable and responds to user's requests instantaneously as the calculations are relatively straight.

Decision tree analysis

This is a schematic representation of several decisions having two or more outcomes followed by the probability of the occurrence of each of them. This gives a tree-shaped graphical representation of decisions and the nodes or the chance points that help to investigate the possible outcomes [Fig. 2]. There are broadly 6 steps in a decision tree analysis:

- Definition of the problem in structured terms listing all the factors relevant to the solution. The probability distributions of the conditional future behavior of those factors are also then estimated.

- Modelling of the decision process listing all the alternatives in the problem is constructed. The entire decision process is presented schematically and in an organized step-by-step fashion.
- The application of appropriate probability values to all the branches and sub-branches of the decision tree.
- The "solution of decision tree" by finding the particular branch of the tree which has the largest expected value or that maximizes the solution (or vice versa depending on the definition of the problem.)
- Sensitivity analysis can be performed to see how the solution reacts to changes in inputs. This can show how the model behaves when run in real world situations.

The underlying assumptions are listed and ideally should be found to be possible and plausible.^[3,4,13,14]

Nearest Neighbor analysis

Nearest Neighbor Analysis evaluates the distances between the given point and the point closest to it. The analysis is done for every point. The algorithm then compares these values to expected values for a random sample of points from a complete spatial randomness (CSR) pattern. CSR is calculated by two assumptions:

- All points have same likelihood of receiving or not receiving a positive event. Or, as a corollary, are equally likely to have a negative case or negative event
- All positive events or cases are located independently of one another.

The null hypothesis of complete spatial randomness is tested using the standard normal variate (Z statistic). In such a situation, a negative Z score demonstrates clustering while a positive score correlates with dispersion or evenness. The mean nearest neighbor distance

$$\bar{d} = \frac{\sum_{i=0}^N d_i}{N}$$

Where N is the number of points. d_i is the nearest neighbor distance for point i.

The expected value of the nearest neighbor distance in a random pattern

$$E(d_i) = 0.5\sqrt{\frac{A}{N}} + (0.0514 + \frac{0.041}{\sqrt{N}})\frac{B}{N}$$

Where A is the area and B is the length of the perimeter of the study area.

The variance

$$\sigma^2 = 0.0070\frac{A}{N^2} + 0.037B\sqrt{\frac{A}{N^5}}$$

The above equations have a correction factor to counteract the boundary effect.

Z statistic

$$Z = \frac{\bar{d} - E(d_i)}{\sqrt{\sigma^2}}$$

The output file in nearest neighbor analysis gives:

- Input data points
- Total number of points
- The minimum and maximum of the X and Y coordinates
- Size of the study area,
- Observed mean nearest neighbor distance
- Variance
- Z statistic (standard normal variate).

In this method, the study area is a regular rectangle or square and cannot be used for irregularly shaped study areas.

Random Forest Decision Trees

Decision trees are the building blocks of the random forest model which act together as an ensemble. Each individual tree in the random forest gives a class prediction and the class with the most votes becomes prediction from the model. It is a very powerful tool. The random forest model outperforms many more sophisticated tools in making a prediction because of the random effects and the central limiting theorem operating together. This is also called the wisdom of crowds. A large number of relatively uncorrelated models (trees) operating as a committee will outperform any of the individual constituent models [Fig. 3]. The low correlation between models is the key. The trees protect each other from errors unless all of them err in same direction. By probability, some trees are wrong but other trees are right, so the group’s probability moves in the correct direction. But an essential precondition is the absence of multicollinearity or correlation with each other such that the predictions don’t err in the same direction together.^[3,4,13]

Discriminant analysis

Discriminant Analysis is a statistical tool to assess the adequacy of a classification, given the group memberships; or to assign objects to one group among a number of groups. Discriminant Analysis is called Discriminant Function Analysis (DFA) when it is used to separate two groups and Canonical Varieties Analysis (CVA) when more than two groups are involved.

Discriminant Analysis can be used to determine predictor variables related to the dependent variable. It can also be used to predict the value of the dependent variable when values of the predictor variables are available. It is often used in combination with Cluster Analysis and allows for determination of the subject’s location in cases or controls if the risk factors or predicting factors are known.^[13,14]

Support vector machine (SVM)

SVM classifies the subjects into two or more groups. The outcome is used as a classifier. It works on mutually exclusive groups of subjects separable into two or more groups through decision boundaries defined by the traits. The goal of training is to assign optimal weight *w* to each factor so that the sum of

weights comes to 1 so that the weights acting with traits explain the outcomes. This can be done by minimizing quadratic loss function or OLS.¹⁷ The main tuning parameters used are kernel, regularization, gamma and margin. The learning of the hyperplane in linear SVM involves transformation of the problem using a linear equation. For linear kernel the equation for prediction for a new input using the dot product between the input (*x*) and each support vector (*x_i*) is calculated by $f(x) = B(0) + \text{sum}[a_i * (x, x_i)]$.^[15]

Larger C value makes a hyperplane which attempts to classify the training points correctly even if that line or plane has to curve repeatedly. A small C value makes the optimizer define a larger margin separating hyperplane at the cost of misclassifying more points.

A margin is a separation of line from closest class points. A good margin has large separation for both the classes. Gamma parameter gives influence of a single training example. A low gamma means ‘far’ points from separation line get considered in calculation and high values mean ‘close’ points near separation line are considered.^[13,15,16]

Neural network

Neural networks refers to a set of algorithms designed to recognize patterns. A neural network can be compared to a network or circuit of neurons. Artificial neural network is a network of artificial neurons or nodes. There are hidden layers

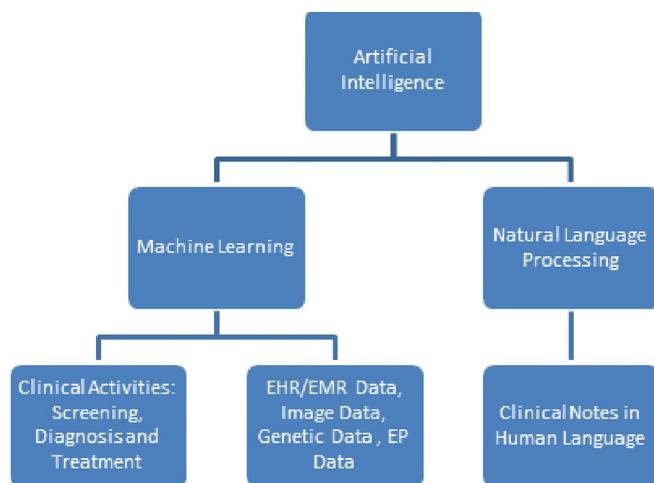


Figure 1: Clinical data to clinical decision making using Natural Language Processing and Machine Learning Analysis

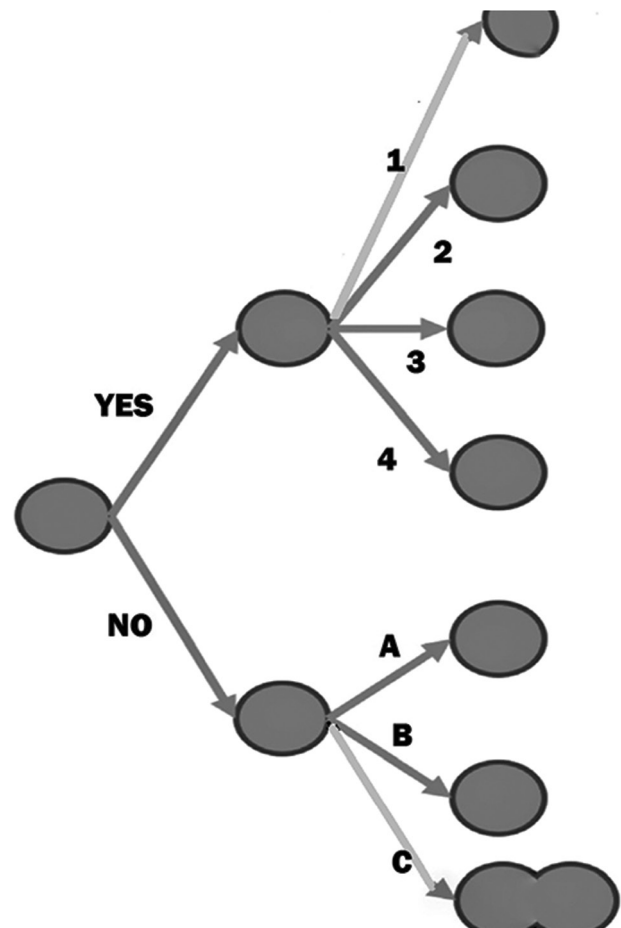


Figure 2: Decision Tree

between the input and the output as shown in Fig. 4. These hidden layers have weights attached to different inputs and can have complex mathematical functions modeled on them. The connections of the biological neuron can also be modeled as weights. A positive weight represents an excitatory connection and a negative weight signifies an inhibitory one. All inputs are acted upon by the weights attached to the hidden layers and summed to get an output. This is called linear combination.

Predictive modeling, adaptive control and training using a dataset can be done with neural networks. Experiential-self-learning using neural networks help draw conclusions from complex and unrelated set of information. They can pick up information from images, sound, text or time series. These are converted into vectors from which numerical signals about all real-world data are picked up.

Neural Networks are used as a clustering and classification layer on top of the stored data. They help to group unlabeled data according to similarities among the example inputs, and they classify data when they have a labeled dataset to train on. Neural networks can extract features from other algorithms for clustering and classification working as components of larger machine-learning applications for reinforcement learning, classification and regression.^[3,4,15-17] Examples of publicly available deep neural networks like convolutional neural networks are GoogleNet, AlexNet and VGGNet. Software like Caffe and Tensorflow can also be used.

Hidden Markov

These statistical models help to get the hidden information from observed sequential attributes or symbols. Hidden Markov Models (HMMs) derive their name from Russian mathematician Andrey Andreyevich Markov. They have been used in speech recognition, biological nucleotide sequences, predict exons and introns in DNA, identify functional motifs (domains), in proteins (profile HMM) and align two sequences (pair HMM). A good HMM simulates the real world source by converting the real world's observed data to symbols. Machine Learning techniques based on HMMs have solved problems including speech recognition, optical character recognition, bioinformatic needs like genetic analysis and computational biology problems. In HMM, a discrete stochastic process progresses through a series of states 'hidden' from the observer to generate the output which is the solution to the problem. Each hidden state generates a symbol representing an elementary unit of the modeled data. This is a powerful technique used when a probability for a sequence of observable events can be computed. Some of the events of interest are hidden. They are not observed directly. A Hidden Markov Model (HMM) allows us to talk about both observed events and hidden events. It is like the hidden layers in the neural

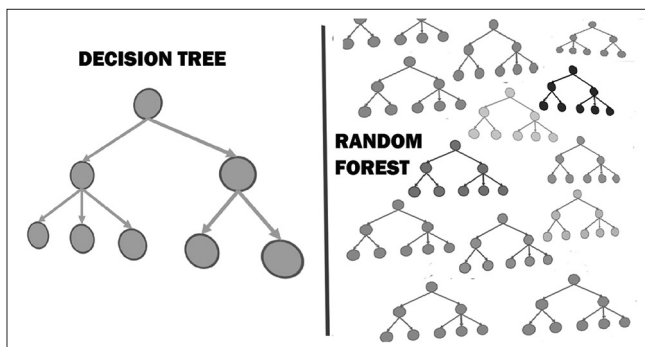


Figure 3: Random Forest Analysis and Relation to Decision Trees

networks. A transition probability matrix is first constructed representing the probability of moving from first state to second state.^[13] The variables of interest and computations include a sequence of observations drawn from a vocabulary, a sequence of observation likelihoods called emission probabilities. Each emission probability expresses the probability of an observation getting generated from a given state from the initial probability distributions over both states.^[3,4,15] A first-order HMM assumes that the probability of a particular state depends only on the previous state and is not affected by any other state. Other techniques can be modeled for more complex scenarios.

IDx-DR, an artificial intelligence algorithm analyzing retina images from Topcon NW400 camera uploaded to the cloud, became the first medical device to be approved by the United States Food and Drug Administration for using artificial intelligence to detect greater than mild diabetic retinopathy in adults with diabetes in April 2018.^[18] Intra-Ocular-Lens (IOL) 'super formula' was introduced as a 3-D framework using similarities in IOL formulas to develop IOL 'super surface' by amalgamating the modern formulae-- Hoffer Q, Holladay I, Holladay I with the Koch adjustment and Haigis formulae. This super formula calculates IOL power in all types of eyes.^[19] Ecstatic corneal conditions and glaucoma are also seeing a large number of algorithms being developed. Lietman *et al.*

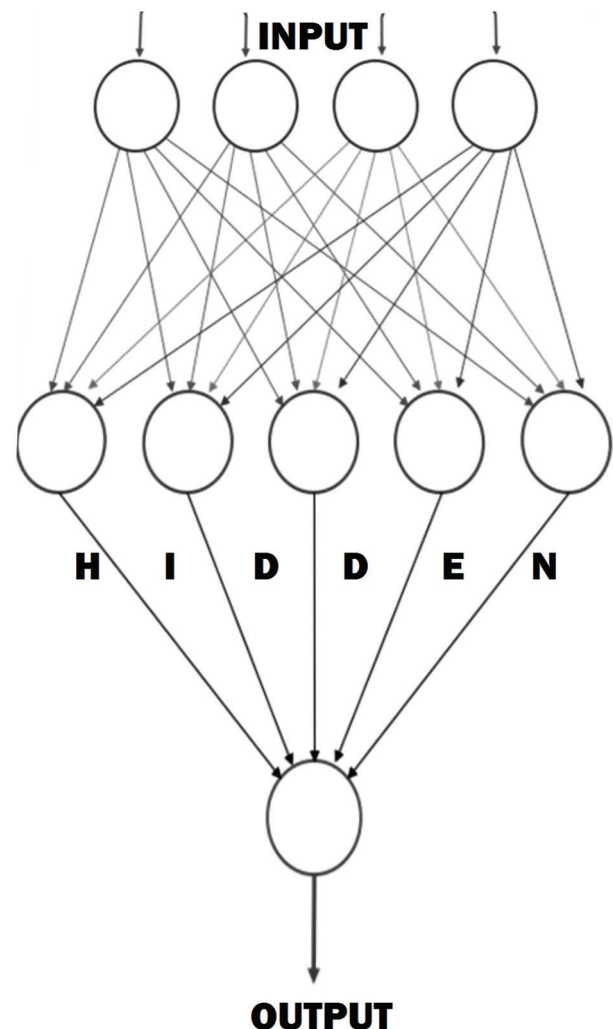


Figure 4: Neural Network Analysis

who used artificial neural networks on 106 glaucoma patients and 249 controls for diagnosing glaucoma based on visual fields reported that the algorithm outperformed global indices at high specificities (90%–95%).^[20] Li *et al.* used deep learning algorithm on 4012 pattern deviation images for functional glaucoma diagnosis with a reported accuracy of 87.60% (Sensitivity = 93.20%, Specificity = 82.60%).^[21] Yousefi *et al.* in another cross-sectional study, 677 patients and 1146 controls used unsupervised learning methods of visual fields for prognostication% (Sensitivity = 87%, Specificity = 96%) where unsupervised machine learning consistently detected the progression of glaucoma much earlier than conventional methods.^[22] Prediction of progression using Humphrey's Visual Fields even 24-2 algorithm with deep learning can be made up to five and a half years before conventional methods.^[23] Mardin *et al.* combined confocal laser scanning ophthalmoscope images with visual fields using a machine learning classifier to get area under the curve (AUROC) of 0.977 (Sensitivity = 95%, Specificity = 91%).^[24] The advantage of AI is that it can use data of great variety and variability to model the outcomes and predict them. Even genetic data can be used for risk stratification once the mapping can be completed.^[24,25]

The basic purpose of this article was to focus on the different techniques being used in healthcare and more so in ophthalmology rather than provide an exhaustive list of the reported and ongoing studies.

Conclusion

AI applications in healthcare can have tremendous potential and usefulness. However, the success of healthcare AI depends on the availability of clean healthcare data of high quality which can come only with careful execution and liberal funding. It is critical to consider data capture, storing, preparation and mining. Standardization of clinical vocabulary and the sharing of data across platforms is imperative for future growth. It is also important to ensure that bioethical standards are maintained in collection and use of the data. There is a need to develop strong foundations for computational bioethics.^[26] The authors hope that this paper helps the stakeholders to realize their potential and make a contribution to the artificial intelligence in healthcare literature as well as practice.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

- Marr B. The Key Definitions of Artificial Intelligence (AI) That Explain Its Importance [Internet]. Forbes.com. 2019 [cited 27 September 2019]. Available from: <https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/#2ec26df44f5d>.
- Administration UFaD. Guidance for industry: Electronic source data in clinical investigations. 2013 [cited 27 September 2019]. Available from: <https://www.fda.gov/downloads/drugs/guidances/ucm328691.pdf>.
- Darcy AM, Louie AK, Roberts LW. Machine learning and the profession of medicine. *JAMA* 2016;315:551-2.
- Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA* 2013;309:1351-2.
- Murff HJ, FitzHenry F, Matheny ME, Gentry N, Kotter KL, Crimin K, *et al.* Automated identification of postoperative complications within an electronic medical record using natural language processing. *JAMA* 2011;306:848-55.
- Krause J, Gulshan V, Rahimy E, Karth P, Widner K, Corrado G, *et al.* Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy. *Ophthalmology* 2018;125:1264-72.
- Nayak J, Bhat P, Acharya UR, Lim C, Kagathi M. Automated identification of diabetic retinopathy stages using digital fundus images. *J Med Syst* 2007;32:107-15.
- Kapoor R, Walters S, Al-Aswad L. The current state of artificial intelligence in ophthalmology. *Surv Ophthalmol* 2019;64:233-40.
- Gulshan V, Rajan RP, Widner K, Wu D, Wubbels P, Rhodes T, *et al.* Performance of a Deep-Learning algorithm versus manual grading for detecting diabetic retinopathy in India. *JAMA Ophthalmol* 2019;137:987-93.
- Mahendra S. 7 Types Of Artificial Intelligence (AI) [Internet]. www.cisin.com. 2019 [cited 28 September 2019]. Available from: <https://www.cisin.com/coffee-break/technology/7-types-of-artificial-intelligence-ai.html>.
- Tweedie M. 3 Types of AI: Narrow, General, and Super AI [Internet]. Codebots. 2019 [cited 10 September 2019]. Available from: <https://codebots.com/ai-powered-bots/the-3-types-of-ai-is-the-third-even-possible>.
- Bali J, Kant A. Basics of Biostatistics. 1st ed. New Delhi: Jaypee Brothers Medical Publishers; 2017.
- Parsian M. Data Algorithms. 1st ed. O'Reilly Media, Inc.; 2015. [cited 27 September 2019]. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/18461814/>.
- Acuallos T. Discriminant Analysis and Applications. 1st ed. New York: Acad. Pr.; 1973.
- Ma Y, Guo G. Support Vector Machines Applications. 1st ed. Switzerland: Springer, Cham; 2014.
- Patel VL, Shortliffe EH, Stefanelli M, Szolovits P, Berthold MR, Bellazzi R, *et al.* The coming of age of artificial intelligence in medicine. *Artif Intell Med* 2009;46:5-17.
- Shahid N, Rapon T, Berta W. Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PLoS One* 2019;14:e0212356. doi: 10.1371/journal.pone.0212356.
- Stark A. FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems [Internet]. U.S. Food and Drug Administration. 2018 [cited 18 June 2020]. Available from: <https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-based-device-detect-certain-diabetes-related-eye>.
- Ladas JG, Siddiqui AA, Devgan U, Jun AS. A 3-D 'Super Surface' combining modern intraocular lens formulas to generate a 'super formula' and maximize accuracy. *JAMA Ophthalmol* 2015;133:1431-6.
- Lietman T, Eng J, Katz J, Quigley HA. Neural networks for visual field analysis: How do they compare with other algorithms? *J Glaucoma* 1999;8:77-80.
- Li F, Wang Z, Qu G, Song D, Yuan Y, Xu Y, *et al.* Automatic differentiation of glaucoma visual field from nonglaucoma visual field using deep convolutional neural network. *BMC Med Imaging* 2018;18:35.
- Yousefi S, Kiwaki T, Zheng Y, Sugiura H, Asaoka R, Murata H, *et al.* Detection of longitudinal visual field progression in glaucoma using machine learning. *Am J Ophthalmol* 2018;193:71-9.
- Wen JC, Lee CS, Keane PA, Xiao S, Rokem AS, Chen PP, *et al.* Forecasting future Humphrey visual fields using deep learning. *PLoS One* 2019;14:e0214875.
- Mardin CY, Peters A, Horn F, Jünemann AG, Lausen B. Improving glaucoma diagnosis by the combination of perimetry and HRT measurements. *J Glaucoma* 2006;15:299-305.
- Burdon KP, Mitchell P, Lee A, Healey PR, White AJ, Rochtchina E, *et al.* Association of open-angle glaucoma loci with incident glaucoma in the Blue Mountains eye study. *Am J Ophthalmol* 2015;159:31-6.
- Bali J, Garg R, Bali R. Artificial intelligence (AI) in healthcare and biomedical research: Why a strong computational/AI bioethics framework is required? *Indian J Ophthalmol* 2019;67:3-6.