Review Article

Journal of Advances in Medical Education & Professionalism

Study of Deep Learning in Medical Education: Opportunities, Achievements and Future Challenges

HOSSEIN MORADIMOKHLES[1*], PhD; GWO-JEN HWANG[2], PhD; HOSSEIN ZANGENEH[1], PhD; MARYAM POURJAMSHIDI[1], PhD; AMIR HOSSEIN AMOOEIRAZANI[3], MSc

[1]*Department of Educational Sciences, Faculty of Humanities, Bu-Ali Sina University, Hamedan, Iran;* [2]*Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taipei, Taiwan;* [3]*Bu-Ali Sina University Faculty of Humanities, Hamedan, Iran*

Abstract

Introduction: In this era of progress, interest has developed regarding advancing deep learning (DL) in medicine. However, there has been reluctance to use deep learning, particularly among medical educators. The limitations of previous research were examined in this study, along with the extent to which DL can be used in medical education and its potential impact on educational quality. We were interested in discussing DL's prospects, and determining whether we could benefit from it in medical education.

Methods: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol was used to manage this review procedure. Six databases were searched carefully to obtain relevant studies. Our search identified 981 articles from the database based on our standards. After filtering the duplicated articles, 11 studies were included in the systematic review.

Results: The results showed that DL applications attracted researchers' attention in the medical and education technology owing to their effectiveness to provide the personalized assistance and feedback. Furthermore, the majority of research concentrated on teaching medical students how to utilize DL applications in the classroom, and all of them tried to improve medical students' proficiency with DL instruments in practical applications. Deep learning components in medical learning environments have two segments—in the educational settings like speech recognition or Video content analysis for affecting students' learning, and in the medical settings, applying deep learning from diagnosis to prevention. An integration of them can work better in medical education.

Conclusion: Medical education uses DL to improve the students' education. DL is a powerful instrument which has become more famous in terms of superb outcomes. Besides, using DL in medical education is likely to continue as a hotly debated area of research and a well-known classroom strategy.

Keywords: Deep learning, Medical education, Data mining, Online learning, Literature review

**Corresponding author:*
Hossein Moradimokhles, PhD;
Department of Educational Sciences, Faculty of Humanities,
Bu-Ali Sina University,
Hamedan, Iran
Tel: +98-9120249086
Email: moradimokhles@basu.ac.ir

Please cite this paper as:
Moradimokhles H, Hwang GJ, Zangeneh H, Pourjamshidi M, Amooeirazani AH. Study of Deep Learning in Medical Education: Opportunities, Achievements and Future Challenges. J Adv Med Educ Prof. 2024;12(3):148-162. DOI: 10.30476/JAMP.2024.99740.1853.

Received: 5 August 2023
Accepted: 17 March 2024

Introduction

The term "deep learning" has incorrect interpretations in education and technology disciplines. It describes a method of learning in which the objective is to achieve an in-depth understanding of topic rather than succumb to surface learning. "Non-surface learning" is undeniably related to deep information clarification, which depends on the level of need for cognition (1). The non-surface learning approach accentuates Bloom's high cognitive, emotional, and psychomotor learning levels, for instance, combining data and organizing values (2). In contrast, "deep learning" allows machines to determine broad classification by different processing layers like a human neural network (3). However, there is a similarity between them in that they both seek to provide more accurate forecasts or useful support by considering more basic types of data. We define "deep learning" as feature classification that is strongly based on "artificial intelligence" and "machine learning."

One of the aspects of "artificial intelligence", and one of the machine learning algorithms is deep learning [DL]. Figure 1 (4) shows the differences between Machine learning, and Deep Learning. "Deep learning" refers to learning methods that incorporate multi-level processing, and achieving via discrete but simple steps (5). This formation allows very complicated operations to be received, and used over procedures. In the classification steps, stages cannot alter significant aspects of input and include extraneous diversities (3). DL technology indicates learning from extensive data to problem-solving, such as high multidimensionality, complication, and hidden depths (6). Deep Learning provides a link by making forecasts based on narrative/text, image-based data, and quantitative as inputs. As shown in Figure 1, by developing technology, feature extraction that humans perform will be provided by machines from now on.

Deep Learning divides features, and this process is performed in several steps (5). Analyzing shallow-level learning at a high level increases learning strategies by excerpting advantageous components from the inputs (6). Classifying data is not entered by the humans; machines learn from each process by a general-purpose learning method. By excerpting this deep arrangement of features from the inputs, they use this collection to make an ultimate decision. Deep learning is characterized by search queries, data sorting, e-commerce, object recognition, speech-to-text, and many more (3). Figure 2 (7) illustrates the components of DL (Input floor, process floor, and output floor). Extracting features and classification is performed on the process floor. In current years, there is developing interest integrate artificial intelligence (AI) into medicine, still there is a long way to use it in different clinical fields (8).

Although it is imperative to address the deficiency in the medical educational systems, extant research shows that limited systematic reviews pertaining to this issue and related themes have been conducted so far. Only Carin (5) and Sanal, Paul (9) have done literature reviews in this area. The objective of Carin (5) was to conduct an evaluation of the capacity of this technology to mechanize specific aspects of the medical profession. Following this, a discussion is given concerning the integration of deep learning principles into the domain of medical education, with the conclusion that while DL has significant potential to improve medical practice, astute healthcare professionals will also have a thorough understanding of its limitations. Sanal, and Paul (9) focused on The integration of Artificial Intelligence (AI), and effective utilization of vast amounts of "Big Data," in conjunction with the ever-evolving techniques in the biotechnology, portends to bring about significant transformative changes in medicine and concluded that the transformation of the medical education system is a necessary endeavour so as to appropriately equip healthcare professionals, and proactively sensitize the populace, thereby facilitating a seamless transition.

From an educational viewpoint, it was essential

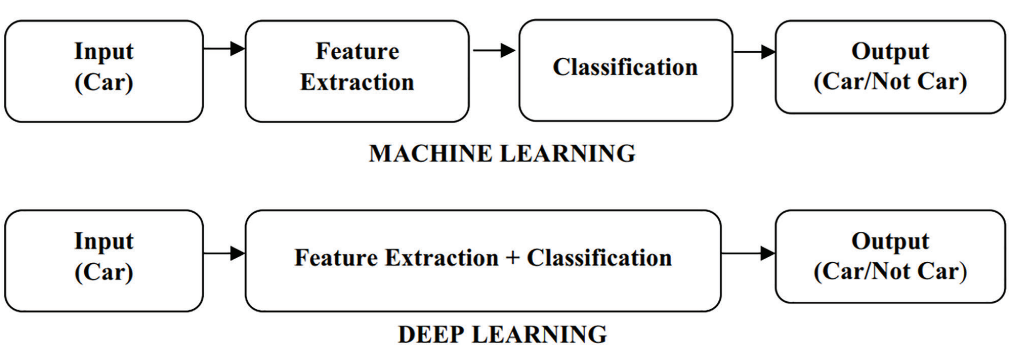

Figure 1: Machine learning vs. Deep Learning

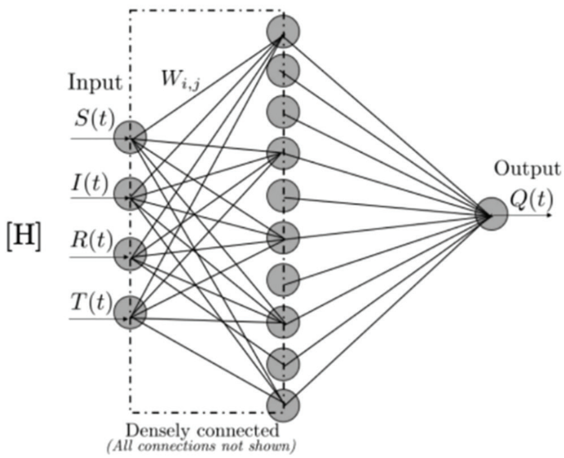

Figure 2: Steps of features dividing in DL

to develop a teaching process with the components of DL to assist the conception of students and make such classrooms more enjoyable and effective. DL technology has a significant opportunity to reform educational settings in medicine, yet we are at the first phase of the approval of DL technology. Interventions thereafter demonstrated that this model has a greater influence on medical students (5). The practical application of classification in the classroom resulted in the most noticeable improvement in the students' operational abilities during construction; thus, DL enhanced the mechanical abilities of the required students (10). One of the central claims of AltSchool, which asserts that it will be the forthcoming of learning when DL reconstructs education, is the focus on individualized studying (11) —regarded as putting the principles, and standards established during the membership speech process into practice (12). If the students could apply the standards mentioned in the speech to their works, it would show they were successful (10). The results of this review accommodate the realization of opportunities and challenges that face both physician-teachers and students when using it.

This study investigated the shortcomings of present studies and discusses how far DL can be used in medical education, and how much it can affect educational quality. Thus, this article introduces the subsequent research questions:

RQ1: What is the evidence for the effectiveness of "deep learning" in medical education environments?

RQ2: What are the components of "deep learning" in the field of medical education environments?

Literature of Deep Learning and Education Framework

Three waves make up the evolution of deep learning. In 1940s–1960s, cybernetics with biological learning theories emerged as the first wave of DL. The following phase is connectionism, which used back-propagation to develop a human brain with various complexities in 1980s and early 1990s. The current circumstance under the term "deep learning" started in 2006 and was first presented in 2016 (11). Deep neural networks' root in the late 1940s and the construction of the first statistical model of nerve cells by McCulloch and Pitts (13), where DL genesis located were the fundamental base of deep learning (12, 14). Perception-related shallow-neural networks, which Rosenblatt (15) first described, consist of a single-layer neural network. The development attributed to Rosenblatt (15) serves as a cornerstone of deep learning. Ivakhnenko (16) introduced the Group Method of Data Handling (GMDH) as a technique set for training neural networks. Neocognitro, the first convolutional neural network (CNN), was not developed during that time period, according to Fukushima (17). Soon afterward, Hop (18) submitted a recurrent neural network (RNN) that became the main competitor of CNN (14).

The recognition of handwritten digits was the first use of CNN by LeCun, Boser (19). Regarding insufficient teaching strategies and computing tools, there was an "AI winter" in late 1990s and early 2000s (5). Fortunately, Hinton (20) announced a new step-by-step teaching method for data classification networks to beginners. Chellapilla, and Puri (21) proposed applying the graphics processing units (GPUs) in DL, so all barriers are lost. Alexnet (22) emerged as the preeminent image classification model in the 2012 Imagenet Challenge. A multitude of deep neural network models were developed in response to the evolving circumstances, as depicted in Figure 3 (14), organized by year. Nevertheless, a considerable number of authorities have lauded the fact that despite the abundance of research on the application of DL in education, there are few studies investigating its effects on medical patients. Therefore, we should explain models of deep learning in this section.

Convolutional Neural Network (CNN) Model

This model shows extensive execution by taking the arrangement features from the inputs. Layers with combinatorial structures are the primary guideline of this model. Complexity is a significant math operation of this method. Understanding the issues with picture composition, and labelling is the fundamental utilization of this approach accomplished by continuously scanning the inputs, and processing them as nested and complex

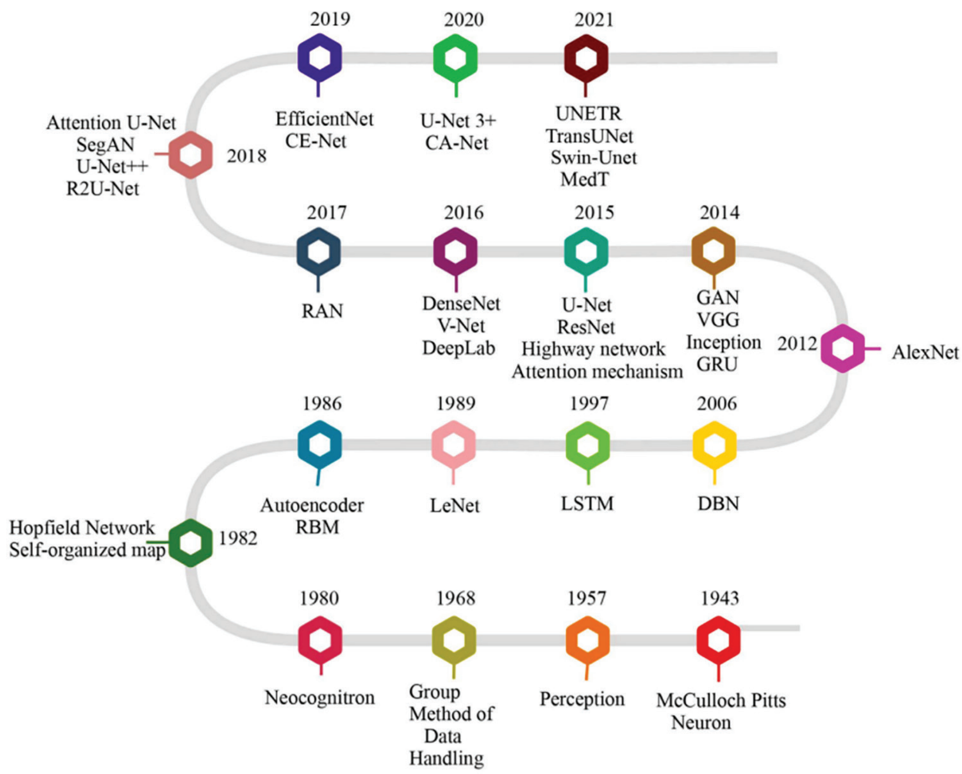

Figure 3: Models of Deep learning

with a sigmoid function. Figure 4 (23) illustrates a CNN conceptual character, which is constructed by superimposing numerous blocks that extract features on a floor-by-floor basis. A convolutional floor, a nonlinear refreshing layer, and a merge floor comprise each section. As shown in Figure 4, lineaments near the merge floor summarize characters by grading or maximizing, eliciting fundamental components, lowering character dimensionality, and then defending all floors by an incidental floor (24).

Stacked Autoencoder Model

Unsupervised component revolutions are stacked autoencoders. An autoencoder comprises three floors: the input floor, the hidden floor, and the output floor. In the case of a stacked autoencoder, the hidden floor operates as a non-straight revolution process. To improve classification, the authentic input characters are modified floor by floor, as illustrated in Figure 5 (25), by roughly arranging numerous encoders in the concealed floor. Nevertheless, all floors retain data similar to the authentic input (26).

Recurrent neural networks Model

An RNN processes the input arrangement characters by maintaining hidden units with a situation floor that inherently consists of data that shows the history of classification, then examines the outputs with the invisible operation at different times. Backpropagation was most effectively utilized at the time by RNNs processing continuous input, such as narrative/text. As depicted in Figure 6 (27), this model represents a dynamical connection of legitimate significance. Through process improvement, each floor strives to extricate more suitable features. Nevertheless,

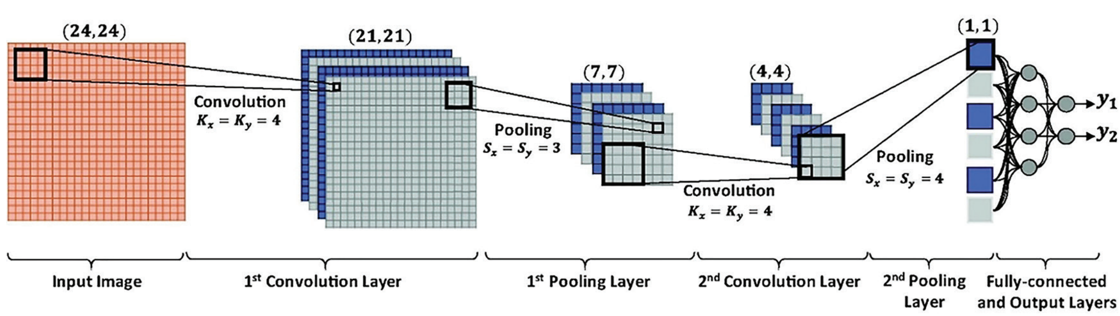

Figure 4: Concept of Convolutional Neural Network (CNN) Model

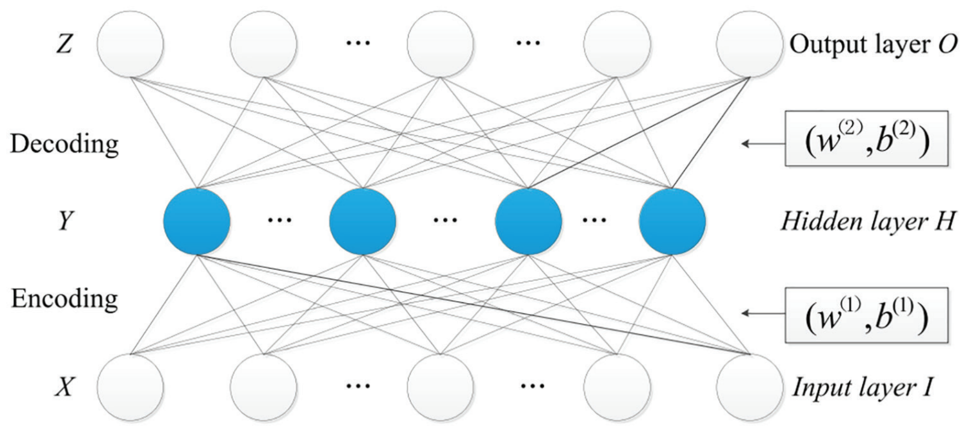

Figure 5: Input floor, a hidden floor, and an output floor, and the hidden floor functions as a non-straight revolution process in Stacked Autoencoder Model

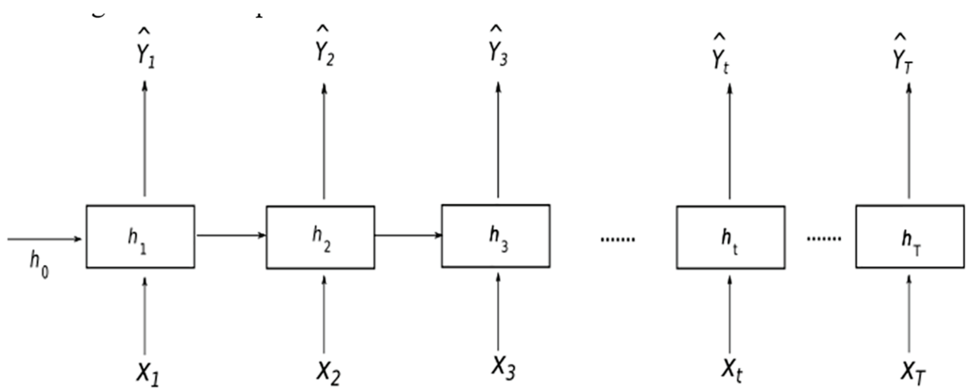

Figure 6: Input arrangement characters of Recurrent neural networks Model

their endeavours have been equivocal in light of their progress in reverse propagation. Recurrent neural networks model can be used for complex processes, such as forecasting the coming character in the books, deducting words in a text, finding out the subject from the first word, and many more (28). These models have various applications.

In 2006, hierarchical learning was the old name of DL, which considered study areas as pattern recognition. Supervised models generally require tags for the activity objectives. On the other hand, unsupervised models do not require a classification for learning. The indicative activities are well suited for supervised learning in terms of their complication and truthfulness requirements. Currently, the majority of presented DL applications in medicine are modified CNNs. Constructing CNN are a large number of concealed, interconnected brain cells with distinct input and output floors. According to deep learning algorithms, deep networks have two-fold advantages over outdated theories due to the absence of suitable illustrations. The function of architecture is the reason for these benefits, and depends on the component construction of the hidden data generation arrangement. Initially, getting to know an allotted illustration permits us to combine the found characteristic values to new combos now no longer visible for the duration of exercising (29). There is no internal structure associated with its use. Moreover, inferring a symbol requires connecting it to a variable by carefully chosen inference rules (30). On the other hand, nets use only large activity vectors, large weight matrices, and scalar non-linearities to draw rapid "intuitive" conclusions that underpin intuitive thinking (31). Moreover, we apply the framework to use deep learning in education.

Deep learning Education Framework

A collection of the pedagogical guidelines designed for the courses on Artificial Intelligence:

• It is imperative to advocate a systems perspective that elucidates the interplay of mechanisms in generating intelligence. This approach will challenge the perception that artificial intelligence is comprised of the disjointed algorithms.

• This statement proposes providing the students with an opportunity to gain expertise in the process of encoding representational content that can be interpreted by various mechanisms to produce an array of behavioral responses.

• The approach of sequentially presenting subjects in a cumulative manner, with subsequent material built upon preceding content, is analogous to the manner in which calculus relies on algebra, which in turn is grounded in arithmetic.

• Educating students ought to encompass not only applying artificial intelligence (AI) techniques, but also the process of constructing these techniques by assembling elementary constituents.

• It is essential to acknowledge and prioritize critical skills and capabilities that manifest within the human cognition, even when they pose difficulties in the realm of formalization.

An inaugural course would adhere to these guiding principles (32). The commencement tasks pertaining to pattern matching, and decision-making would provide a foundation for subsequent assignments on conceptual inference, sequential control, and plan generation. On the other hand, the methods aimed at managing uncertain information would be incorporated as adaptations of traditional symbolic methodologies (33). There is an expectation that the implementation of a methodical and cumulative strategy for artificial intelligence (AI) education at the system level will yield more skilled practitioners and researchers in comparison to the prevalent algorithm-focused methodologies that are presently being utilized (32).

Methods

Our study is a controlled systematic review. We reclaim accepted evidence for a specific question or question, and evaluate and blend the findings to notify training, strategy, and forward study in deep learning and medical education. We selected this particular review methodology due to its emphasis on gathering literature that is both accurate and candid regarding the implementations of deep learning in medical education (34). We followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol in order to conduct our systematic literature review (35). This review of literature and content analysis provide pertinent information on the subjects, approaches, and conceptual framework of deep learning and medical educational research (36). We systematically investigate the issue to enhance medical education via DL.

Search Strategy

Six databases were carefully searched for relevant studies, including Springer, Wiley online library, Web of Science, Science Direct, Scopus, and ProQuest. This search was handled in October 2022, and upgraded in November 2022. Google Scholar required no systematic searches for reports and other gray articles to add to the data set a collection of keywords belonging to the fields of (a) deep learning; (b) medical students; (c) education:

("deep learning" OR "machine learning) AND (medi*) AND (edu* OR ins* OR learn* OR teach* OR student*)

Moreover, in the coming journals, a non-systematic search was organized: Medical Teacher, Expert Systems with Applications, BMC Medical Education, and Academic Medicine. Bibliographic and digital library searches returned most results from publications in these journals, which were therefore elected. We manually located supplementary research articles, reference lists, and full-text reviews through Google Scholar. The majority of research studies rely on participant self-reports as opposed to precise skill levels. In 128 studies, self-reports were utilized, whereas performance evaluations were employed in 12 studies. We used the Scopus journal classification for subject classification to categorize research areas. The most convenient field was selected when there were multiple categories. Table 1 shows that education is the most studied field. Besides, much research had been done in Biological Sciences. Examples include engineering and education and computer sciences and education. Thus, the trends of deep learning dimensions shown in Figure 7, is created by Microsoft Excel.

Inclusion and Exclusion Criteria

The database combined articles focusing on fostering medical education through deep learning in the educational settings, published between 2018 and 2022. We selected the articles in the following phases. First, we screened the article titles for eligibility based on our criteria. Then, the abstracts of all relevant studies were evaluated for eligibility, using uniform criteria. Finally, we examined the full texts of remaining articles. The study publication date, age, design, participant language, or location were all left unrestricted. The document excluded editorials, commentaries, book chapters, and newspaper dispatches. In addition, conference proceedings and papers were omitted. By agreement of two reviewers, articles that did not report effect size or number of available studies were excluded. Two reviewers separately made research selections. However, the "publish or perish" desktop application estimated score for article quality

Table 1: The study type included articles (n=138).

		Operational	Conceptual	n
Study type	Case-study	5	1	6
	Comparative-analysis	1	2	3
	Content-analysis	1	3	4
	Delphi-study	1	1	2
	Experiment	7	10	17
	Mixed method	13	13	26
	Performance test	1	1	2
	Review	2	13	15
	Survey	4	10	14
	Theoretical		11	12
	Total			138
Study Field	Biological Sciences, Education	9	1	10
	Arts and Humanities	1		1
	Business and International Management	2	3	5
	Business, Management, and Accounting	1	1	2
	Computer Networks, Communications		1	1
	Computer Science, Education	1	2	3
	Communication		1	1
	Developmental and Educational Psychology	7	8	15
	Education	2	2	4
	Engineering, Education	1	1	2
	Human-Computer Interaction			
	Language and Linguistics	1		1
	Library and Information Sciences			
	Management of Technology and Innovation, Education	4	13	17
	Psychology	4	8	12
	Strategy, Management, Education	1	26	36
	Total			138

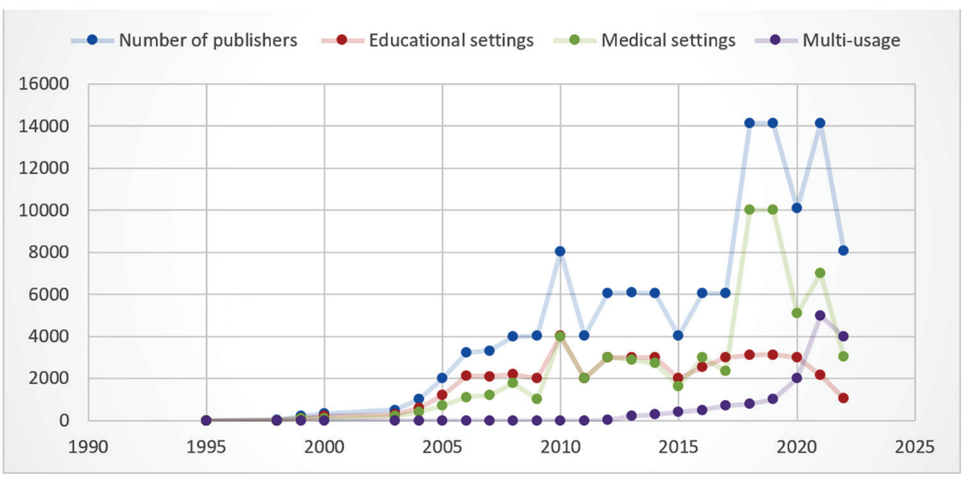

Figure 7: Trends of deep learning dimensions.

using hA number, g-index, citations, h-index, hI-index, and l-hI-index as the primary endpoint.

PRISMA Flowchart

We identified 981 articles from the database based on our standards. Nevertheless, we reviewed 138 separate studies, of which 113 were duplicates. Then, 11 studies matched our standards. The process of selecting articles is shown in Figure 8 based on PRISMA 2020. Other sources did not reveal any more records.

Our causes for screening full-text articles:
(1) Not containing our matter.
(2) They only focused on deep learning in the medical education.
(3) Not quoting as part of teachers' knowledge.
(4) Not an original journal study.
(5) No complete textual content is accessible online.
(6) Duplicating by the first author.

We selected the most recent obtainable articles if there were multiple accepted articles with a

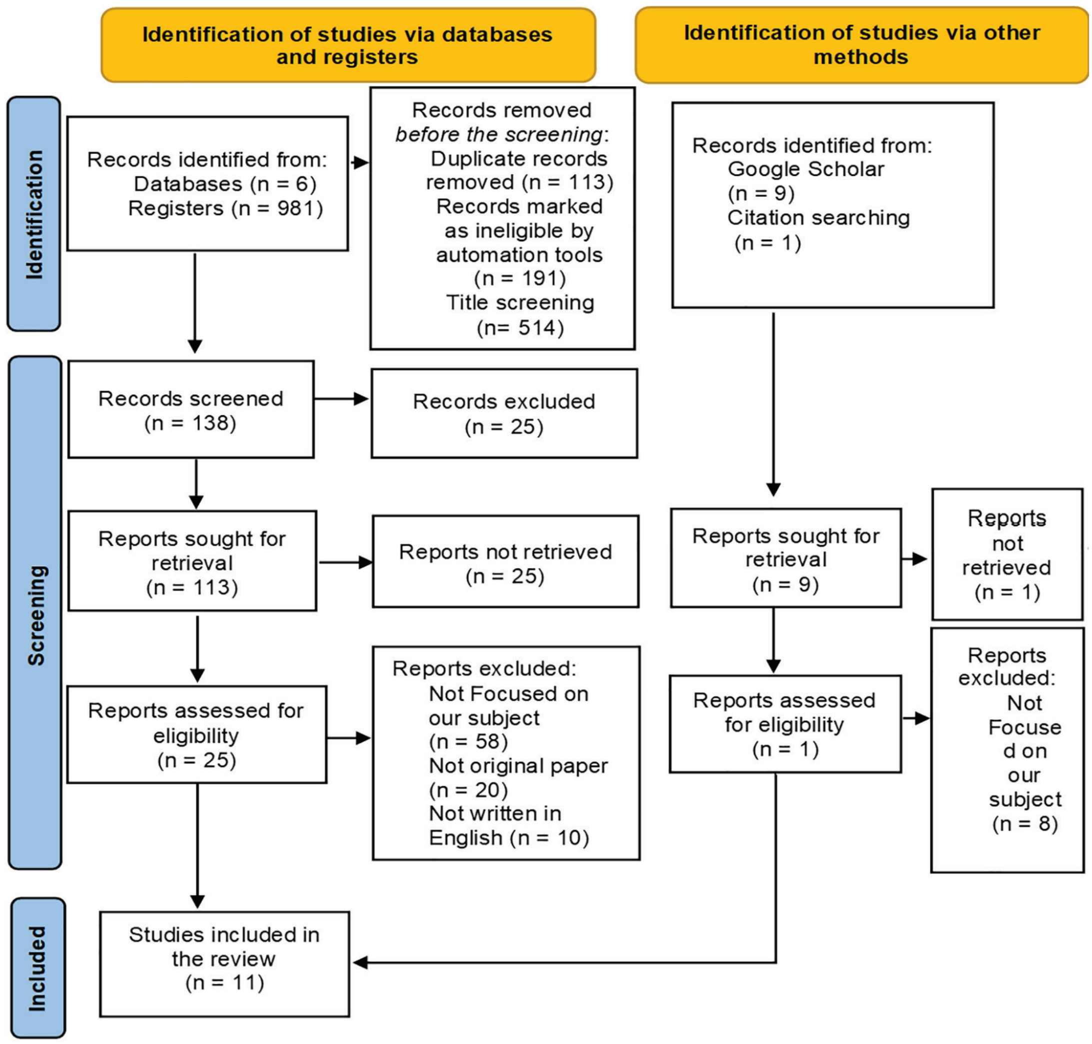

Figure 8: PRISMA flowchart

replication first author.

Selection Bias

A second coder independently coded a selection of studies to determine their quality. In a systematic review, publication bias occurs during the selection process, and a selection procedure significantly reduces publication bias (35). For the precise search, we utilized all databases. A second coder, in accordance with our criteria, executed the search operation and incorporated the title, abstract, and full-text evaluation stages of the study selection procedure. The second coder selected each study for examination in accordance with our criteria. If the full text had the justification for excluding the study, it was not selected. We used an inter-ranking statistics kappa coefficient as a unit of measure. To solve any doubt, we stated the standards clearly. The statistic inter-ranking measuring unit for 138 selected articles was 0,682, indicating that two coders had good agreement after determining the requirements. The second reviewer randomly selected 7 of the 25 articles in screening phase to ensure the validity of the coding by rechecking. Because the two coders agreed highly, we reached 0,755 of Cohen's kappa after resolving all issues to achieve the best outcomes through discussion following the coding sessions.

Data Collection Process

We used Microsoft Excel to input data and standard forms to extract it. A first reviewer extracted the information, and a second reviewer verified that it was accurate and comprehensive. We categorized the data according to our study questions. We combined descriptive, quantitative, and mixed methods with the semantic application of keyword research for the studies (37). The number of studies covering multiple keywords in DL is illustrated in Table 2.

Moreover, the number of studies covering multiple keywords of DL is illustrated in Table 3. One hundred thirty-eight articles were covered, addressing eight dimensions in two categories.

Results

The descriptive data of the included articles is shown in Table 4.

RQ1: What is the evidence for the effectiveness of "deep learning" in medical education environments?

The supplementary articles and Table 5 cover the basic concepts and applications of deep learning. The research employed a variety of methodologies and analytical tools to address a wide range of themes and applications. Then, we described, analyzed, and explored each question except the third one that showed the study trend.

All researches emphasize the effectiveness of deep learning in the medical education environments, like Multi-usage by Carin (5), Behrad and Abadeh (14), Medical settings usage by Ejaz, McGrath (38), Therapy usage by Wei, Yu (6), Diagnosis usage by Liu, Wang (39), and data mining usage by Yang, Zheng (40) as described in Table 6. All of them reported having a better effect on students' learning. Moreover, their similar dimension of features, especially hidden affection for students, is better compared to other methods. It is a big chance that DL technology will change how medical students are practicing. Since this technology is still in its early stages of application, medical educators must carefully evaluate its potential

Table 2: Keywords used included articles (n=138)

Term	n
Deep learning	54
Artificial Intelligence	13
Machine Learning	27
Multiliteracy	7
Total	138

Table 3: Deep learning dimensions (n=138).

Term		n
Educational settings	Narrative Review	3
	Speech Recognition	1
	Video content analysis	7
	Quantitative Assessment	3
	Data mining	19
Medical settings	Diagnosis	87
	Therapy	13
	Health (Prevention)	1
Multi-usage		4
Total		138

Table 4: Characteristics of included studies (n=11).

Author (year)	Country	Participants (N)	Age	Discipline	Usage	Evaluation tool	Measured Outcomes	Quality
Carin, L. (5)	USA	NA	NA	Medical students	Multi-usage	NA	Students	%77
Ejaz, McGrath. (38)	UK	28	18-22	Medical students	Medical settings	Qualitative and quantitative analysis	Students	%77
Wei, et al. (6)	China	NA	NA	Medical students & patients	Therapy	Weighting method	Students & patients	%77
Liu, Wang. (39)	China	NA	NA	NA	Diagnosis	NA	Students & patients	%77
Yang, Zheng. (40)	China	200 postgraduates	NA	30 disciplines	Data mining	Interview	Students	%77
Behrad and Abadeh (14)	Iran	NA	NA	NA	Multi-usage	NA	Students	%77
Pinnock, McDonald. (41)	New Zealand	NA	NA	Medical undergraduate students	Diagnosis	NA	Students	%77
Harmon, Pitt. (42)	Australia	NA	NA	Nursing students	Therapy	NA	Students	%77
Hwang, Tang. (43)	Taiwan	NA	NA	medical students	Multi-usage	NA	Students	%77
Ranchon, Chanoine. (44)	France	NA	NA	Medical undergraduate students	Multi-usage	NA	Students	%77
Li, Liang. (45)	Australia	Five groups of physicians	NA	Graduated physicians	Educational-usage	NA	Students	%77

Table 5: Concept of included studies (n=11).

Author (year)	Deep learning dimensions	Main concept
Carin, L. (5)	This chapter discusses the deep learning as a subset of machine learning, and artificial intelligence.	Multi-usage
Ejaz, McGrath. (38)	As future healthcare leaders and clinicians, medical students play a crucial role in the clinical implementation of AI-driven technologies. By investigating the medical students' perspectives, the authors hope to publish the world's first report on the state of AI in medical education.	Medical settings
Wei, et al. (6)	The diagnosis accuracy is statistically proved by studying a method to show developed DL algorithm for the medical college studying. Moreover, the systematic information classification and the impact of fusion methods are compared and assessed using the Alzheimer's disease public dataset, which includes medical education information.	Therapy
Liu, Wang. (39)	After providing an overview of a few well-known deep learning architectures, the authors estimate how these architectures can be used for different specific US image analysis tasks like classification, diagnosis, and segmentation.	Diagnosis
Yang, Zheng. (40)	As scientific research becomes more data-driven, collaborative, and quantitative, there is an enormous potential for medical postgraduates with information skills. Incorporating data mining skills into postgraduate medical education is crucial because of using big data in medicine and health care.	Data mining
Behrad and Abadeh (14)	Introducing the most common paradigms, fusion techniques, and deep learning frameworks initially in this study. Furthermore, they describe several learning techniques, such as multitasking training, transfer learning, and end-to-end studying. They provide a summary of the deep learning techniques for interpreting multidimensional health information.	Multi-usage
Pinnock, McDonald. (41)	Artificial intelligence in undergraduate medical reasoning curricula for students should handle the advantages, and hidden problems of advanced access to artificial intelligence in the medical diagnoses. We provide our curriculum with an educational structure for this ambitious transition.	Diagnosis
Harmon, Pitt. (42)	Nurses play an essential role in pain estimate, and management in many therapy settings that require extensive intelligence and medical abilities. Experimental studies of therapy designs are needed to instruct nursing education, practice, and policy and support progress in the pain nursing discipline.	Therapy
Hwang, Tang. (43)	This article verified that AI is gambling a chief function in profiling and forecast in nursing analysis. Intellectual retailers were the maximum customarily used AI gadget in nursing. Quantitative approaches were the primary analysis method; the most appropriate analyses were in medical education. Healthcare workers and patients are its two introductory research models for the problem of selection and measurement. A middle dimension cognizance becomes comparing the implementation of AI-associated gear and procedures.	Multi-usage
Ranchon, Chanoine. (44)	The primary aim of this systematic literature review was to ascertain and examine the quantitative studies that make use of artificial intelligence (AI) for clinical pharmacy services. It is of utmost importance to undertake considerable endeavours in conjunction with data scientists to effectively determine whether the utilization of AI-driven applications and tools offer any tangible benefits to the provision of the clinical pharmacy services in the actual practice setting.	Educational-usage
Pelly, Fatehi. (46)	The perspective of nurses and healthcare professionals with regards to health learning programs using Artificial Intelligence (AI), which play a crucial role in facilitating effective implementation, is inadequately comprehended.	Multi-usage

effects and investigate them via educational research. Background and knowledge of "AI Winters" beginning from the early stages of AI, and overstate of this technology capabilities, should be part of this education. Although DL has a possibility for the future of medicine, informed physicians will be aware of its restrictions (5).

Numerous research endeavors were conducted to ascertain the efficacy of deep learning in the context of medical education settings. An illustration of this can be found in the research of Kononowicz et al. A novel methodology was implemented in 2020 to generate real-time responses to student input through the utilization of deep learning techniques to generate a virtual patient. This endeavor aimed to provide personalized feedback and encourage active learning (14). The empirical study concluded that the virtual patient, which operates on a deep learning-based architecture, showed noteworthy efficacy in enhancing learners' engagement, and academic achievements. A study conducted by Wang et al. yielded similar results. In 2018, a deep learning algorithm was formulated by researchers, exhibiting significant precision in its ability to perform a target task.

The viability of profound learning in therapeutic instruction situations can be ascribed to a few components. Profound learning calculations can be gotten from expansive and different datasets, which can offer assistance to make strides in the precision of recalling and comprehension of restorative instruction (47). Moreover, profound learning can be utilized to create personalized learning encounters that are custom-made to the wants of person

Table 6: Concept of included studies (n=11).

Deep learning dimensions	**Conceptual definition with operational components**
Narrative review	First-year college students typically study narrative review as a standard technique. Their purpose is to find some research that explains an exciting topic. The subject of interest provides pre-defined research question or search strategy in the narrative review. They are not orderly and do not follow any particular protocol. Any policy or protocol does not provide these reviews. Reviewers learn about the problem but should fully understand the current state of art. In Hospice and Palliative Care, Fins and his colleagues provide examples of the narrative review.
Speech recognition	The primary advantage of speech recognition is the ability to search. Speech recognition is an intriguing area of machine learning, and computational linguistics that constructs approaches and technology to facilitate the recognition and translation of speech into text by technology. It is often called speech-to-text (STT), automated voice recognition, or automatic speech recognition (ASR). It draws on expertise and study from systems engineering, computer science, and linguistics. Word recognition is the opposite process. A speaker should practice some speech recognition systems by reading text or a small vocabulary to the device and improving the correctness of voice recognition by the system's analysis of the individual's speech and using that information.
Video content analysis	The capacity to automatically analyze video to find and identify both temporally and spatial occurrences are called video content analysis (VCA), often referred to as video analysis or video analytics (VA). Different industries utilize these technological capabilities, such as entertainment, education and healthcare, marketing, industrial, transportation, smart home, fire and smoky detecting, security, and protection. The methods are available as hardware in specialist video processing chips or programs on overall computers. VCA applies a wide range of functions. One of the fundamental methods of motion sensors that involve a static background is a video motion analyzer. Video tracking and ego-motion analysis are other features.
Data mining	Data mining refers to recognizing trends, and extracting information from big data sets using techniques combining machine learning, analytics, and database management systems. Data mining is a cross-disciplinary area of computer science that involves taking data collection, and organizing the data into a format that is understood. The analytical stage of extracting knowledge in datasets procedure is known as data mining. The distinction between data analysis and data mining is that, regardless of the volume of data, data analysis is available to evaluate theories and models on the database, as in the study of an advertising campaign's efficacy.
Diagnosis	US analysis requires recognizing exciting items in US pictures or video sequences, such as cancers, lesions, and nodules. Tumor or lesion identification, in particular, can offer substantial help for object segmentation and separate benign from cancerous tumors. It admitted that positioning anatomical objects, such as the fetal standard plane, organs, tissues, or landmarks, is crucial before segmentation tasks or the clinical diagnostic process for image-based treatment and cure.
Therapy	We have numerous therapy methods; e.g. Tumors or lesions One of the most time-consuming processes in healthcare setting is identifying and positioning tumors or lesions, which is essential for designing, and implementing therapy. Various anatomical constructions show several distinctions in the lesions. This process usually entails locating and identifying tiny lesions over the whole picture area. The diagnosis and rating of prostate cancer were recently effectively performed by Azizi, Mousavi (49), combining advanced abstract properties collected from temporal-enhanced US by a DBN and the architecture of tissue from pathogen detection. Yap, Hachiuma (50) characterized three different deep learning approaches; a patch-based LeNet, a U-net, and a transfer learning approach with a pre-trained FCN(AlexNet) for identifying breast lesions from two US image datasets obtained by two different US systems to perform an extensive comparison. Cardiac The estimate of numerous cardiac parameters, including stroke volume, ejection fraction, and end-diastolic volume, involves detecting details of cardiac cycle phases (end-diastolic (ED) and end-systolic (ES)) in echocardiograms. A deep residual recurrent neural network (RRN) was suggested by Dezaki, Dhungel (51) automatically identify cardiac cycle stages. RRNs incorporate the benefits of ResNet, which manages the disappearing or erupting gradient problem as the CNN deepens, and the RNN (LSTM), which can prepare a model of the temporal relationships among sequential frames. RRNs comprise residual neural networks (ResNet), two blocks of LSTM units, and a fully linked floor.
Health (Prevention)	One of the newest health techniques is smartwatches, a biological test called an "epigenetic clock" calculated longevity. The method assesses the buildup of methyl groups on a person's DNA molecules and focuses on DNA methylation degrees. What DNA methylation age actually measures is still unknown. Although specifics are uncertain, scientists proposed that DNA methylation age estimates the long-term effect of an epigenetic maintenance system. It claimed that the DNA methylation age of blood is related to a mechanism that affects aging since it forecasts all-cause death in adulthood.

understudies. This may offer assistance to make strides understudy engagement, and learning results (47).

RQ2: What are the components of deep learning in the field of medical education learning environments?

Deep learning components in the medical learning environments have two segments—in the educational settings like speech recognition or Video content analysis for impacting students' learning, and in medical settings, applying deep learning from diagnosis to prevention. Nevertheless, integrating education and medicine is more effective. Deep learning can be used in the domain of restorative instruction to enhance traditional pedagogical approaches, facilitate individualized learning experiences, and advance clinical decision-making. The components of profound learning in restorative instruction incorporate information procurement, information preprocessing, show advancement, show preparing, and demonstrate assessment (48-51).

Notwithstanding, as shown in Table 6, they contributed to creating a theoretical deep learning framework with significant strategic elements targeted at skilled professionals, and determining the essential dimensions of DL by systematically evaluating papers that describe and evaluate them.

The components of deep learning in medical education can be encourage analyzed as takes after (52):

• Information procurement: This includes the collection, and organization of expansive datasets of restorative pictures, persistent information, and other restorative data.

• Information preprocessing: This includes the cleaning, normalization, and change of the information to plan it for using within the profound learning demonstrate.

• Demonstrate improvement: This includes the plan, and creation of the neural organize engineering, which incorporates the number and sort of layers, enactment capacities, and other parameters.

• Demonstrate preparing: This includes using arranged dataset to prepare the profound learning demonstrate, which includes altering the weights and the inclinations of neural organize to play down the blunder between the anticipated and genuine results.

• Demonstrate assessment: This includes the testing and approval of profound learning show employing a partitioned dataset to survey its exactness and execution.

The components of deep learning in medical education are basic to the improvement of compelling profound learning models. Information procurement is fundamental to guarantee that we can get to an expansive and different dataset to memorize from. The preprocessing of information is crucial in order to ensure that the data is accurate and structured in a way that is suitable for the deep learning model (53). Illustrate progress encompasses the architecture of the neural network structure, which is critical for acquiring knowledge from the data. Demonstrate preparing includes utilizing the arranged dataset to prepare the profound learning model, which permits it to memorize designs, and connections within the information. At long last, demonstrate assessment is fundamental to survey the exactness and execution of profound learning show (54).

Discussion

The advent of deep learning, a subset of artificial intelligence (AI), had a significant impact on various fields, including medical education. Significant promise has been demonstrated by this technology in augmenting conventional pedagogical approaches, facilitating individualized learning encounters, and bolstering the clinical decision-making capabilities of healthcare practitioners. Deep learning models, for instance, can be trained on extensive datasets comprising medical images, lab results, and other pertinent data. This enables students to enhance their diagnostic capabilities and gain insights from practical scenarios. Moreover, these models can analyze vast amounts of information, and identify patterns that human instructors may miss, allowing for more accurate diagnoses and treatment plans. However, integrating deep learning techniques into medical education poses several challenges that must be addressed (53).

One major hurdle is the need for high-quality, diverse, and abundant datasets to train deep learning models effectively. Furthermore, ensuring that these models operate transparently and comprehensively is crucial, as the healthcare professionals must understand how the models arrive at their conclusions. Ethical concerns regarding data privacy, bias, and the potential for AI systems to replace human judgment require careful consideration (54). Despite these challenges, the benefits of harnessing deep learning in the medical education are substantial. Through the provision of customized learning trajectories, the augmentation of clinical decision-making capabilities, and ultimately the enhancement of patient outcomes, deep learning possesses the capacity to fundamentally transform the training of aspiring healthcare professionals. To gain a comprehensive understanding of the benefits and challenges associated with the integration of deep learning into medical education, additional research is required. Such studies will provide valuable insights into the most effective ways to integrate AI technologies into medical curricula and ensure that they complement human expertise rather than replacing it (47).

Limitations

Automatic learning of features, multi-layer learning of features, high potential to create more features, high accuracy, high maintenance, and hardware and software support is the positive side of deep learning applications in the medical education. Besides, it can only be as innovative or effective as the quality of data provides. A few of the limitations of deep learning applications in medical education are algorithmic biases, the “black box” nature of such applications, data sourcing and privacy violations, and unclear legal liability. The basis of present algorithms has many collapses that cannot use them over the transfer resembles. Moreover, the models of DL demand high computing ability, and it takes time to compute. Physicians should improve the first side and the outcome second side.

Future research directions

Future research fields for the question “What is the evidence for the effectiveness of ‘deep learning’ in the medical education environments?” could include investigating the long-term impact of deep learning on medical student performance, and career success, comparing the effectiveness of deep learning with other educational interventions, such as simulation-based training or problem-based learning, and exploring the use of deep learning in continuing professional development for practicing physicians. Investigating the ethical ramifications of incorporating deep learning into medical education could be an additional area of focus. This would involve concerns such as mitigating potential biases in dataset composition and guaranteeing impartiality and transparency in algorithmic decision-making.

Regarding the question “What are the components of ‘deep learning’ in the medical education environments?”, future research could focus on identifying the key features of successful deep learning models in medical education, such as the importance of dataset size, image resolution, and annotation quality. Researchers could study the optimal balance between breadth and depth of knowledge in deep learning models for the medical education, as well as the role of transfer learning and domain adaptation in adapting deep learning models to new medical domains or tasks. Finally, exploring the integration of deep learning with other educational technologies, such as virtual reality or gamification, could provide insight into how to create engaging and immersive learning experiences for medical students and practitioners.

Conclusion

From the information included in the multiple databases, this study is a snapshot of the development and current value of deep learning in medical education during the previous 28 years. The results indicate that DL has the potential to influence science, a burgeoning academic discipline that has experienced a meteoric rise in prominence over the last decade. Diverse studies published in specialized fields are impacted by DL due to the information it combines. Ultimately, the findings show that the experts in this area have investigated various subjects, dividing them into two parts: aspects of deep learning cultivation and facts to support the usefulness of deep learning in cultivating physicians. Medical education utilizes deep learning to improve the learning of students. DL is a powerful instrument that has become more famous in terms of superb outcomes. A detailed overview of the latest articles on the deep learning models used for physicians’ education is shown, which shows how multi-applications develop neural nets’ efficiency. We categorize relevant research into two primary groups: the application of medical factors and the application of educational criteria. We come to the conclusion that better use of medical facilities is required. Deep learning models and methodologies, as well as preventative strategies, are discussed in this article. We cannot use them in the classroom in terms of generalization, class inequality, information privacy matters, and many other causes. Notwithstanding, we assume that the usage of DL in medicine will extend to be a sharply challenged field of research in the coming years based on its severe rising interest in future outcomes.

Authors’ Contributions

All authors contributed to the discussion, read and approved the manuscript and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated resolved.

Conflict of Interest

The authors declare no conflicts of interest.

References

1. Chonkar SP, Ha TC, Chu SSH, Ng AX, Lim MLS, Ee TX, et al. The predominant learning approaches of medical students. BMC Medical Education. 2018;18(1):17-32.
2. Jhala M, Mathur J. The association between deep learning approach and case based learning. BMC Medical Education. 2019;19(1):106-27.
3. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature.

2015;30(2):521-47.
4. Dargan S, Kumar M, Ayyagari MR, Kumar G. A survey of deep learning and its applications: a new paradigm to machine learning. Archives of Computational Methods in Engineering. 2020;27(1):1071-92.
5. Carin L. On Artificial Intelligence and Deep Learning. Acad Med. 2020;95(5):45-57.
6. Wei L, Yu Z, Qinge Z. Medical College Education Data Analysis Method Based on. Mobile Information Systems. 2022;67(1):172-199.
7. Dandekar R, Barbastathis G. Quantifying the effect of quarantine control in Covid-19 infectious spread using machine learning. MedRxiv. 2020:4(3):200-84.
8. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56.
9. Sanal MG, Paul K, Kumar S, Ganguly NK. Artificial Intelligence and Deep Learning: The Future of Medicine and Medical Practice. The Journal of the Association of Physicians of India. 2019;67(4):71-3.
10. Caires S, Almeida L, Vieira D. Becoming a teacher: Student teachers' experiences and perceptions about teaching practice. European journal of teacher education. 2012;35(2):163-78.
11. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. Communications of the ACM. 2020;63(11):139-44.
12. Peters MA. Deep learning, education and the final stage of automation. Educational Philosophy and Theory. 2018;42(18):201-239.
13. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics. 1943;5:115-33.
14. Behrad F, Abadeh MS. An overview of deep learning methods for multimodal medical data mining. Expert Systems with Applications. 2022;11(7):607-19.
15. Rosenblatt F. The perceptron, a perceiving and recognizing automaton Project Para. USA: Cornell Aeronautical Laboratory; 1957. pp. 61-23.
16. Ivakhnenko AG. The group method of data handling A rival of stochastic approximation. Soviet Automatic Control. 1968;13(7):43-55.
17. Fukushima K. A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern. 1980;36(3):193-202.
18. Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. Proc Natl Acad Sci U S A.1982;79(2):554-8.
19. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, et al. Backpropagation applied to handwritten zip code recognition. Neural computation. 1989;1(4):541-51.
20. Hinton GE, Osindero S, The YW. A fast learning algorithm for deep belief nets. Neural Computation. 2006;18(7):1527-54.
21. Chellapilla K, Puri S, Simard P, editors. High performance convolutional neural networks for document processing. Tenth international workshop on frontiers in handwriting recognition. Université de Rennes 1, Oct 2006, La Baule (France). ffinria-00112631.
22. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Communications of the ACM. 2017;60(6):84-90.
23. Kiranyaz S, Avci O, Abdeljaber O, Ince T, Gabbouj M, Inman DJ. 1D convolutional neural networks and applications: A survey. Mechanical Systems and Signal Processing. 2021;15(1):197-228.
24. McKinley R, Rebsamen M, Meier R, Wiest R, editors. Triplanar ensemble of 3D-to-2D CNNs with label-uncertainty for brain tumor segmentation. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 5th International Workshop, BrainLes 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Revised Selected Papers, Part I. 2019. pp. 379-87.
25. Yu M, Quan T, Peng Q, Yu X, Liu L. A model-based collaborate filtering algorithm based on stacked AutoEncoder. Neural Computing and Applications. 2022;51(4):1-9.
26. Khamparia A, Saini G, Pandey B, Tiwari S, Gupta D, Khanna A. KDSAE: Chronic kidney disease classification with multimedia data learning using deep stacked autoencoder network. Multimedia Tools and Applications. 2020;79(35):425-40.
27. Hewamalage H, Bergmeir C, Bandara K. Recurrent Neural Networks for Time Series Forecasting: Current status and future directions. International Journal of Forecasting. 2021;37(1):388-427.
28. Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. Computer Science. 2014;82(6):937-53.
29. Montúfar GF, Morton J. When does a mixture of products contain a product of mixtures? SIAM Journal on Discrete Mathematics. 2015;29(1):321-47.
30. Collobert R, Weston J, Bottou L, Karlen M, Kavukcuoglu K, Kuksa P. Natural language processing (almost) from scratch. Journal of machine learning research. 2011;12(3):2493−537.
31. Cho SE, Geem ZW, Na KS. Prediction of suicide among 372,813 individuals under medical check-up. Journal of Psychiatric Research. 2020;131(9):9-14.
32. Chaudhri VK, Lane HC, Gunning D, Roschelle J. Intelligent learning technologies Part 2: applications of artificial intelligence to contemporary and emerging educational challenges. Ai Magazine. 2013;34(4):10-2.
33. Langley P, editor. An integrative framework for artificial intelligence education. Proceedings of the AAAI Conference on Artificial Intelligence. 2019;33(1):9670-7.
34. Aromataris E, Pearson A. The systematic review: an overview. AJN The American Journal of Nursing. 2014;114(3):53-8.
35. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic reviews. 2015;4(1):1-9.
36. Elliott JH, Synnot A, Turner T, Simmonds M, Akl EA, McDonald S, et al. Living systematic review: 1. Introduction—the why, what, when, and how. Journal of clinical epidemiology. 2017;91(4):23-30.
37. Knoke D, Yang S. Network fundamentals. Social

network analysis. 2008;154(1):3-14.
38. Ejaz H, McGrath H, Wong BL, Guise A, Vercauteren T, Shapey J. Artificial intelligence and medical education: A global mixed-methods study of medical students' perspectives. Digital Health. 2022;8(20):559-621.
39. Liu S, Wang Y, Yang X, Lei B, Liu L, Li SX, et al. Deep learning in medical ultrasound analysis: a review. Engineering. 2019;5(2):261-75.
40. Yang L, Zheng S, Xu X, Sun Y, Wang X, Li J. Medical Data Mining Course Development in Postgraduate Medical Education: Web-Based Survey and Case Study. JMIR Medical Education. 2021;7(4):220-47.
41. Pinnock R, McDonald J, Ritchie D, Durning SJ. Humans and machines: Moving towards a more symbiotic approach to learning clinical reasoning. Med Teach. 2020;42(3):246-51.
42. Harmon J, Pitt V, Summons P, Inder KJ. Use of artificial intelligence and virtual reality within clinical simulation for nursing pain education: a scoping review. Nurse Education Today. 2021;97(10):47-83.
43. Hwang GJ, Tang KY, Tu YF. How artificial intelligence (AI) supports nursing education: profiling the roles, applications, and trends of AI in nursing education research (1993–2020). Interactive Learning Environments. 2022;51(1):1-20.
44. Ranchon F, Chanoine S, Lambert-Lacroix S, Bosson JL, Moreau-Gaudry A, Bedouch P. Development of artificial intelligence powered apps and tools for clinical pharmacy services: A systematic review. International Journal of Medical Informatics. 2023;172(9):874-913.
45. Li Y, Liang S, Zhu B, Liu X, Li J, Chen D, et al. Feasibility and effectiveness of artificial intelligence-driven conversational agents in healthcare interventions: A systematic review of randomized controlled trials. International Journal of Nursing Studies. 2023;143(5):391-412.
46. Pelly M, Fatehi F, Liew D, Verdejo-Garcia A. Artificial intelligence for secondary prevention of myocardial infarction: A qualitative study of patient and health professional perspectives. International Journal of Medical Informatics. 2023;173(3):105-41.
47. Alam A, editor. Possibilities and apprehensions in the landscape of artificial intelligence in education. 2021 International Conference on Computational Intelligence and Computing Applications (ICCICA), Nagpur, India, 2021. pp. 1-8.
48. Pedro F, Subosa M, Rivas A, Valverde P. Artificial intelligence in education: Challenges and opportunities for sustainable development. France: United Nations Educational; 2019. pp. 12-48.
49. Azizi S, Mousavi P, Yan P, Tahmasebi A, Kwak JT, Xu S, et al. Transfer learning from RF to B-mode temporal enhanced ultrasound features for prostate cancer detection. International journal of computer assisted radiology and surgery. 2017;12(1):111-21.
50. Yap MH, Hachiuma R, Alavi A, Brüngel R, Cassidy B, Goyal M, et al. Deep learning in diabetic foot ulcers detection: a comprehensive evaluation. Computers in Biology and Medicine. 2021;135(8):45-96.
51. Dezaki FT, Dhungel N, Abdi AH, Luong C, Tsang T, Jue J, et al., editors. Deep residual recurrent neural networks for characterisation of cardiac cycle phase from echocardiograms. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings. 2017. pp. 100-8.
52. Mirchi N, Bissonnette V, Yilmaz R, Ledwos N, Winkler-Schwartz A, Del Maestro RF. The Virtual Operative Assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine. PloS one. 2020;15(2):98-135.
53. Goksel N, Bozkurt A. Artificial intelligence in education: Current insights and future perspectives. Handbook of Research on Learning in the Age of Transhumanism. USA: IGI Global; 2019. pp. 1-448.
54. Wong GK, Ma X, Dillenbourg P, Huan J. Broadening artificial intelligence education in K-12: where to start? ACM Inroads. 2020;11(1):20-9.