Machine Learning in Medicine

With an increased number of medical data generated every day, there is a strong need for reliable, automated evaluation tools. With high hopes and expectations, machine learning has the potential to revolutionize many fields of medicine, helping to make faster and more correct decisions and improving current standards of treatment. Today, machines can analyze, learn, communicate, and understand processed data and are used in health care increasingly.

1. Introduction

Living in the big data era, with billions of terabytes of data generated every year, it might be challenging for humans to proceed with all the information. However, Artificial Intelligence (AI) can lend a helping hand. In the past, machines have gained an advantage over humans in physical work, where automation contributed to industry and agriculture's rapid development. Nowadays, machines are gaining an advantage over humans in typically human cognitive skills like analyzing and learning. Moreover, their communication and understanding skills are improving quickly. There are numerous examples where AI already achieves much better results than humans in analyzing [1][2][3].

The AI focuses on exploiting calculation techniques with advanced investigative and prognostic facilities to process all data types, which allows for decision-making and the mimicking of human intelligence. Such computational systems usually operate on large amounts of data and often integrate different types of input. AI is a broader field of science, and one of the most significant branches of AI in medicine is machine learning (ML). ML means understanding and processing information from a given dataset by the algorithm, namely machine. The word "learning" stands here as the machine's ability to become more effective with training experience. Such a machine can quickly draw novel conclusions from the data that may be omitted by humans. Machines' potential increases year by year, making them more autonomous. However, human interference is necessary and still has the final word about taking or not particular actions. At least for now. Will it change in the future? Will we let the AI perform actions itself, or will it remain only as a human tool? One thing is unquestionable—we must start accustoming ourselves to live alongside the machines that begin to equal or even surpass people in the processes of analyzing and deciding.

2. Application of Machine Learning in Medicine

Techniques based on ML started to step into medicine in the 1970s, but over time, the possibilities for their use began to multiply [4][5]. The first-ever ML-based diagnostic system was already approved by the U.S. Food and Drug Administration (FDA) in 2018 [6]. The system implements "in silico clinical trials", which helps develop more efficient clinical trial strategies. It allows investigators to detect safety and effectiveness signals earlier in the new drug development process and contributes to costs reduction [7].

With many hopes and expectations, ML has the capacity to revolutionize many fields of medicine, helping to make faster and more correct decisions and improving current standards of treatment.

2.1. Imaging in Medicine

With an increased number of images taken every day, e.g., magnetic resonance imaging (MRI), computer tomography (CT), or X-rays, there is a strong need for a reliable, automated image evaluation tool. An interesting example is a tool created by Kermany et al., which, when adequately trained, has the potential of numerous applications in medical imaging [8]. It uses a neural network to analyze optical coherence tomography (OCT) images of the retina, allowing to diagnose macular degeneration or diabetic retinopathy with high accuracy and sensitivity. Moreover, this model could also indicate the cause of bacterial or viral pediatric pneumonia, making it a universal radiological tool. ML also allows creating images

with better quality. In reconstructing the noisy image, the automated transform by manifold approximation (AUTOMAP) framework is used to obtain better resolution and quality [9]. As more details can be recognized, the diagnosis can be faster and more accurate.

The accuracy of imaging and its assessment is essential, especially in (the case of) detecting and diagnosing abnormalities in the development of the fetus. Parental diagnosis of fetal abnormalities has markedly benefited from the advances in ML. ML algorithms have been widely used to predict the risk of chromosomal abnormalities (i.e., euploidy, trisomy 21) or preterm births. The latest technological advances in ML also improve the diagnosis of fetal acidemia or hypoxia based on CTG analysis [10].

ML also progresses in imaging methods. In silico staining technique provides an excellent solution to microscopy problems, such as the need for additional staining to visualize some cells or tissue structures [11]. Based on patterns invisible to the human eye, the algorithm can accurately predict the cell nuclei's location and size, cell viability, or recognize neurons among mixed cell populations.

Recent advances and using DL-based techniques enabled to read more information from various images. It is now possible to improve the transplantation process by using CNN [12]. The approach created by Altini et al. analyzes kidney histological slides and determines the global glomerulosclerosis (ratio between sclerotic glomeruli and an overall number of glomeruli), which is one of the necessary steps in the pre-transplantation process. By using DL, it can be assessed faster and with high accuracy, and therefore has the potential to quicken the whole transplantation process. Using automatic semantic segmentation of patients with autosomal dominant polycystic kidney disease enables noninvasive disease monitoring [13]. The introduction of the latest ML techniques also enables predicting less obvious information from microscopic section images. Two interesting examples determine RNA expression [14] and predict patient survival after tumor resection [15]. Schmauch et al. created the HE2RNA model, which correctly predicted transcriptome of different cancer types, detected molecular and cellular modifications within cancer cells and was able to spatialize differentially expressed genes specifically by T cells or B cells [15]. A different study developed two CNN models that could predict survival from histological slides after the surgical resection of hepatocellular carcinoma. Both models outperformed a composite score incorporating all baseline variables associated with survival [15].

2.2. Personalized Decision Making

Fast and personalized decisions are crucial in almost every field of medical sciences. Moreover, detecting and predicting life-threatening conditions before their full clinical manifestation is a highly significant issue. Cardiology's main goals focus on developing tools predicting cardiovascular disease risk [16] and the mortality rate in heart failure patients [17]. AI can also be applied for prognosis in acute kidney injury [18]. Physicians can be informed about the injury before changes, detectable with current methods, occur. AI uses a recurrent neural network trained with big datasets of over 700,000 adult patients. It can predict kidney function deterioration up to 48 h in advance, giving some extra time to improve patients' condition.

The auspicious direction of AI is an individualized prediction of genetic disease occurrence based on the patient's genome screening. Integrating genomic data with parameters such as lifestyle or previous conditions established the tool, which may be used in the early screening of abdominal aortic aneurysm [19].

Another useful tool, created for personalized nutrition, processes data (e.g., blood tests, gut microbiome profile, physical activity, or dietary habits) and predicts postprandial blood glucose level [20]. The evaluation indicated a high correlation between predicted and measured glycemic response, indicating high fidelity of ML application. Such an approach may be the beginning of the personalized nutrition era to program diet in other metabolic disorders.

As data suggest, the microbiome is strictly related to cancer, affecting tumorigenesis's natural course. Specific microbial signatures promote cancer development and affect many aspects of cancer therapies, such as the treatment's efficacy or safety. Hence, ML-driven gut microbiota analysis seems to be extremely useful in oncology to prevent cancer development, make an appropriate diagnosis, and finally treat cancer [111].

Early diagnosis is a crucial but often also challenging task. Here once again, ML proves useful. It is now possible to detect abnormalities in patients' handwriting. Using ANN, the algorithm can determine whether a person may be affected by Parkinson's Disease or how much the disease has already developed [21]. Often the symptoms of a particular condition are subtle and therefore difficult to observe. That is what happens with blepharospasm, which is caused by orbicularis oculi muscle contractions and, in most problematic cases, may result in complete closure of the eyelids and blindness.

Based on ANN, AI software was created to deal with diagnosis making [22]. It analyzes recorded videos, recognizes facial landmarks, and can detect even subtle blinks and around the eye area movement, which are necessary for diagnosing this dystonia.

2.3. Drug Design

The traditional approach of new drug design is based on numerous wet-lab experiments and is costly and time-consuming. Solutions to these problems are combining traditional synthesis methods with ML techniques [23]. Granda et al. applied the algorithm to analyze obtained data and classify reagents as reactive or non-reactive, faster, and with high precision. The used approach is the beginning of creating an automated tool for chemicals discovery, contributing to new therapeutic compounds development. Screening big datasets of compounds to find ligands with target proteins is a very long part of the drug design process, even with utilizing ML. The fast-screening compounds tool, which uses traditional support vector ML and a graphics processing unit (GPU), was created to face this challenge. A GPU divides all the data into small parts and analyzes them simultaneously in smaller subsets, shortened screening time. The multi-GPU computers might reduce this time even more [24]. Applying a deep neural network enabled to screen of over 107 million molecules and identified a new antibiotic [25]. This compound, named halicin, differs in structure from previously known antibiotics and exhibits broad-spectrum activity in a mouse model, including pan-resistant bacteria.

Another big problem in the field of pharmacology is to identify the compounds' mechanism of action. Yang et al. proposed a "white-box" ML approach that could identify new drugs and antibiotics mechanisms of action, contribute to overcoming antibiotic resistance, and design new therapeutics [26].

2.4. Infectious Diseases

Almost all global media in the first part of 2020 were dominated by information about the SARS-CoV2 outbreak. With a promptly increasing number of cases and COVID-19-related deaths, there is a strong need for tools to fast diagnoses, estimate epidemic trends, and determine viruses' evolutionary history. Taking all of these needs into account, ML comes in handy. Combined ML techniques, such as neural network, support vector machine, random forest, and multilayer perception, were used to create a tool for rapid, early detection of SARS-CoV2 patients. This algorithm analyzes computed tomography (CT) chest scans and clinical information such as leucocyte count, symptomatology, age, sex and travel, and exposure history [27].

ML techniques are also beneficial for virologists and epidemiologists. Supervised learning with digital processing was used for the rapid classification of novel pathogens [28]. The authors created an alignment-free tool, which analyzes viral genomic sequences and enables tracking the evolutionary history of viruses and detecting their origin. Modeling epidemic trends is significant from the public health and health care system's point of view. A combination of ML algorithms and mathematical models can reliably predict the number of confirmed cases, deaths, and recoveries in the peak of an epidemic several months earlier. What is more, it can estimate the number of additional hospitalizations, which gives the hospitals and health care facilities time to prepare [29].

References

1. Ernest, N.; Carroll, D. Genetic Fuzzy based Artificial Intelligence for Unmanned Combat Aerial Vehicle Control in Simulated Air Combat Missions. J. Def. Manag. 2016, 6.

2. Morando, M.M.; Tian, Q.; Truong, L.T.; Vu, H.L. Studying the Safety Impact of Autonomous Vehicles Using Simulation-Based Surrogate Safety Measures. J. Adv. Transp. 2018, 2018, 6135183.

3. Palmer, C.; Angelelli, L.; Linton, J.; Singh, H.; Muresan, M. Cognitive Cyber Security Assistants–Computationally Deriving Cyber Intelligence and Course of Actions; AAAI: Menlo Park, CA, USA, 2016.

4. Chu, K.C.; Feldmann, R.J.; Shapiro, B.; Hazard, G.F.; Geran, R.I. Pattern Recognition and Structure-Activity Relation Studies. Computer-Assisted Prediction of Antitumor Activity in Structurally Diverse Drugs in an Experimental Mouse Brain Tumor System. J. Med. Chem. 1975, 18, 539–545.

5. Shortliffe, E.H. Computer-Based Medical Consultations: MYCIN; Elsevier: New York, NY, USA, 1976; ISBN 978-0444569691.

6. Finlayson, S.G.; Bowers, J.D.; Ito, J.; Zittrain, J.L.; Beam, A.L.; Kohane, I.S. Adversarial attacks on medical machine learning. Science 2019, 363, 1287–1289.

7. FDA's Comprehensive Effort to Advance New Innovations: Initiatives to Modernize for Innovation | FDA. Available online: https://www.fda.gov/news-events/fda-voices/fdas-comprehensive-effort-advance-new-innovations-initiatives-

modernize-innovation (accessed on 6 January 2021).

8. Kermany, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.S.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell 2018, 172, 1122–1131.e9.

9. Zhu, B.; Liu, J.Z.; Cauley, S.F.; Rosen, B.R.; Rosen, M.S. Image reconstruction by domain-transform manifold learning. Nature 2018, 555, 487–492.

10. Garcia-Canadilla, P.; Sanchez-Martinez, S.; Crispi, F.; Bijnens, B. Machine Learning in Fetal Cardiology: What to Expect. Fetal Diagn. Ther. 2020, 47, 363–372.

11. Christiansen, E.M.; Yang, S.J.; Ando, D.M.; Javaherian, A.; Skibinski, G.; Lipnick, S.; Mount, E.; O'Neil, A.; Shah, K.; Lee, A.K.; et al. In Silico Labeling: Predicting Fluorescent Labels in Unlabeled Images. Cell 2018, 173, 792–803.e19.

12. Altini, N.; Cascarano, G.D.; Brunetti, A.; Marino, F.; Rocchetti, M.T.; Matino, S.; Venere, U.; Rossini, M.; Pesce, F.; Gesualdo, L.; et al. Semantic Segmentation Framework for Glomeruli Detection and Classification in Kidney Histological Sections. Electronics 2020, 9, 503.

13. Bevilacqua, V.; Brunetti, A.; Cascarano, G.D.; Guerriero, A.; Pesce, F.; Moschetta, M.; Gesualdo, L. A comparison between two semantic deep learning frameworks for the autosomal dominant polycystic kidney disease segmentation based on magnetic resonance images. BMC Med. Inform. Decis. Mak. 2019, 19, 1–12.

14. Schmauch, B.; Romagnoni, A.; Pronier, E.; Saillard, C.; Maillé, P.; Calderaro, J.; Kamoun, A.; Sefta, M.; Toldo, S.; Zaslavskiy, M.; et al. A deep learning model to predict RNA-Seq expression of tumours from whole slide images. Nat. Commun. 2020, 11, 1–15.

15. Saillard, C.; Schmauch, B.; Laifa, O.; Moarii, M.; Toldo, S.; Zaslavskiy, M.; Pronier, E.; Laurent, A.; Amaddeo, G.; Regnault, H.; et al. Predicting survival after hepatocellular carcinoma resection using deep-learning on histological slides. Hepatology 2020, 72, 2000–2013.

16. Jamthikar, A.; Gupta, D.; Khanna, N.N.; Saba, L.; Araki, T.; Viskovic, K.; Suri, H.S.; Gupta, A.; Mavrogeni, S.; Turk, M.; et al. A low-cost machine learning-based cardiovascular/stroke risk assessment system: Integration of conventional factors with image phenotypes. Cardiovasc. Diagn. Ther. 2019, 9, 420–430.

17. Adler, E.D.; Voors, A.A.; Klein, L.; Macheret, F.; Braun, O.O.; Urey, M.A.; Zhu, W.; Sama, I.; Tadel, M.; Campagnari, C.; et al. Improving risk prediction in heart failure using machine learning. Eur. J. Heart Fail. 2019, 22, 139–147.

18. Tomašev, N.; Glorot, X.; Rae, J.W.; Zielinski, M.; Askham, H.; Saraiva, A.; Mottram, A.; Meyer, C.; Ravuri, S.; Protsyuk, I.; et al. A clinically applicable approach to continuous prediction of future acute kidney injury. Nature 2019, 572, 116–119.

19. Li, J.; Pan, C.; Zhang, S.; Spin, J.M.; Deng, A.; Leung, L.L.K.; Dalman, R.L.; Tsao, P.S.; Snyder, M. Decoding the Genomics of Abdominal Aortic Aneurysm. Cell 2018, 174, 1361–1372.

20. Zeevi, D.; Korem, T.; Zmora, N.; Israeli, D.; Rothschild, D.; Weinberger, A.; Ben-Yacov, O.; Lador, D.; Avnit-Sagi, T.; Lotan-Pompan, M.; et al. Personalized Nutrition by Prediction of Glycemic Responses. Cell 2015, 163, 1079–1094.

21. Cascarano, G.D.; Loconsole, C.; Brunetti, A.; Lattarulo, A.; Buongiorno, D.; Losavio, G.; Di Sciascio, E.; Bevilacqua, V. Biometric handwriting analysis to support Parkinson's Disease assessment and grading. BMC Med. Inform. Decis. Mak. 2019, 19, 252.

22. Trotta, G.F.; Pellicciari, R.; Boccaccio, A.; Brunetti, A.; Cascarano, G.D.; Manghisi, V.M.; Fiorentino, M.; Uva, A.E.; Defazio, G.; Bevilacqua, V. A neural network-based software to recognise blepharospasm symptoms and to measure eye closure time. Comput. Biol. Med. 2019, 112, 103376.

23. Granda, J.M.; Donina, L.; Dragone, V.; Long, D.L.; Cronin, L. Controlling an organic synthesis robot with machine learning to search for new reactivity. Nature 2018, 559, 377–381.

24. Jayaraj, P.B.; Jain, S. Ligand based virtual screening using SVM on GPU. Comput. Biol. Chem. 2019, 83, 107143.

25. Stokes, J.M.; Yang, K.; Swanson, K.; Jin, W.; Cubillos-Ruiz, A.; Donghia, N.M.; MacNair, C.R.; French, S.; Carfrae, L.A.; Bloom-Ackerman, Z.; et al. A Deep Learning Approach to Antibiotic Discovery. Cell 2020, 180, 688–702.e13.

26. Yang, J.H.; Wright, S.N.; Hamblin, M.; McCloskey, D.; Alcantar, M.A.; Schrübbers, L.; Lopatkin, A.J.; Satish, S.; Nili, A.; Palsson, B.O.; et al. A White-Box Machine Learning Approach for Revealing Antibiotic Mechanisms of Action. Cell 2019, 177, 1649–1661.e9.

27. Mei, X.; Lee, H.-C.; Diao, K.-Y.; Huang, M.; Lin, B.; Liu, C.; Xie, Z.; Ma, Y.; Robson, P.M.; Chung, M.; et al. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. Nat. Med. 2020, 26, 1224–1228.

28. Randhawa, G.S.; Soltysiak, M.P.M.; El Roz, H.; de Souza, C.P.E.; Hill, K.A.; Kari, L. Machine learning using intrinsic genomic signatures for rapid classification of novel pathogens: COVID-19 case study. PLoS ONE 2020, 15, e0232391.

29. Liu, Z.; Huang, S.; Lu, W.; Su, Z.; Yin, X.; Liang, H.; Zhang, H. Modeling the trend of coronavirus disease 2019 and restoration of operational capability of metropolitan medical service in China: A machine learning and mathematical model-based analysis. Glob. Heal. Res. Policy 2020, 5, 1–11.

29. Liu, Z.; Huang, S.; Lu, W.; Su, Z.; Yin, X.; Liang, H.; Zhang, H. Modeling the trend of coronavirus disease 2019 and restoration of operational capability of metropolitan medical service in China: A machine learning and mathematical model-based analysis. Glob. Heal. Res. Policy 2020, 5, 1–11.