IL NUOVO CIMENTO **46 C** (2023) 137
DOI 10.1393/ncc/i2023-23137-3

COMMUNICATIONS: SIF Congress 2022

Deep learning applied to medical image analysis: Epistemology and data

F. LIZZI
INFN, Sezione di Pisa - Pisa, Italy

received 31 January 2023

Summary. — Deep learning (DL) is changing the way we analyze medical images. What do these changes consist of? According to some interpretations, AI epistemology is moving towards a new paradigm called the "fourth paradigm", in which theory, hypothesis and experiment are going to be unified through data. In this context, the famous statement "Correlation is enough" seems to be an effective way to describe this evolution. This change raises several issues related to medical image analysis. One of the most interesting is about data. Data are a main part of DL development especially because DL algorithms are not explainable. Medical images data sets are scarce, underpopulated, and usually do not contain acquisition and reconstruction parameters. It is only by knowing data characteristics that we can define the boundaries in which an algorithm can properly work.

1. – Introduction

The use of data-driven algorithms to infer, predict, evaluate and build models appears to be a real revolution in the epistemology field and it is changing the rules of the classical scientific method. Artificial Intelligence-based (AI) methods, and especially deep learning ones, have been and will be applied to many fields, from social sciences to theoretical physics. One of the most interesting point of this change is to study and understand how the scientific method is evolving and, in particular, which are the possible scenarios for science and technology. The field of medical image analysis is undergoing a deep change due to these new methodologies. The interdisciplinary and applied nature of such field forces a discussion on the epistemological premises on this kind of study. In order to help the discussion, it is useful to introduce the definition of paradigm as elaborated by Kuhn in "The structure of scientific revolutions" [1]. According to Kuhn, a paradigm is an accepted model or pattern whose fundamental components, for a certain period, remain substantially undisputed. In science, it assumes the shape of an object for further articulations and specifications under new and stricter conditions. When a new paradigm arises within a disciplinary field, according to Kuhn, it is very limited both in its scope and precision. It gains a dominant status when it is more successful than other paradigms. Kuhn considers the success of a paradigm its capability to be able to solve problems considered extremely important by the relevant scientific community. However, at the beginning, it is mostly a "promise of success discoverable in selected and still incomplete examples". Normal science, the sort of scientific activity deployed after the establishment of a paradigm is, according to Kuhn, the actualization of this promise and it allows expanding knowledge in many directions. Even if Kuhn recognizes that normal science usually investigates a limited area of knowledge, he affirms that those same restrictions were born from the confidence in a paradigm and they are essential to the development of the discipline. This way, science is able to do research in a very detailed and deep way. Moreover, when a paradigm works, the nature of the objects of inquiry changes: the paradigm broaden its scope and it is applied to issues beyond its initial reach. Finally, according to Kuhn, a part of these achievements proves to be permanent. A classical critique to the Kuhnian approach is that, in some academic domains, there is only little evidence of this modus operandi. Furthermore, taking into account just a paradigmatic approach produces too clean and linear stories on how disciplines evolve, deleting the pluralism of the history of science. However, the definition of paradigm elaborated by Kuhn has been very influential and allows for more clarity in the discussion on the epistemology of AI and DL. In fact, big data and deep learning algorithms introduce a new epistemological approach, testing a theory by analysing relevant data and inferring the theory itself from data.

2. – The fourth paradigm and medical image analysis

According to Kitchin [2] and Hey *et al.* [3], we can delineate a simplified scheme made of four steps to describe the evolution of scientific paradigms. The first one is called "Experimental Science", dates back to the pre-Renaissance and consists of pure empiricism based on the observation of natural phenomena. The second paradigm is the so-called "Theoretical Science", dates back to pre-computer era and consists in moving towards a broader generalization through the theoretical modelling. The third one is referred to as "Computational Science" and it is based on the simulation of complex phenomena. The "Fourth Paradigm" [3] is the "exploration science" which is the paradigm that the intense use of big data and the data mining techniques are designing nowadays. Exploration science, also called e-science, is mainly based on unifying theories, experiments and simulations using data taken by instruments or simulations and analyzing them with some software. For example, in the medical images analysis domain, especially for diagnosis, prognosis and follow up studies, data are usually taken by instruments from hospitals and, with some exceptions, simulations are usually used for dosimetric studies and evaluation. Gray affirms that e-science is changing the world of science itself, arguing that techniques and technologies are defining a data-intensive science, which is a radical extension of the established scientific method. Kitchin also states that there are others that look at the Fourth Paradigm as the new era of empiricism, underlining that the main difference between a pure empiricist approach and other kinds of approach concerns the domains where AI is used, *i.e.*, industry versus academy. There are many voices and opinions which try to define the e-science paradigm, its boundaries and its potentials. In 2008, Anderson [4] states that "Correlation is enough", so that, in the Big Data era, correlation overcomes causation. Prensky [5], similarly to Anderson, affirms that data mining techniques can extract the complete set of patterns and effects, producing scientific conclusions without any further experiment. In 2013, Steadman [2] comes to affirm that data analysts should not propose or even "bother" themselves with hypothesis anymore. Even if these positions about e-science are typical of industry, its critical discussion should be taken into account even within the academic context.

There are two possible interpretations of the Fourth Paradigm:

- an inductivist approach: big data are able to capture an entire domain of knowledge; there is no need of *a priori* model, theory or hypothesis; the application of data mining is agnostic, *i.e.*, data are inherently meaningful and truthful; meaning transcends the specific domain;

- a constructivist approach: data are constructed and not natural; data must contain the information we are looking for (hypothesis); data selection and sampling always introduce a bias; interpretation needs expertise.

The first one can be properly called a pure empiricist and inductive approach, which really mirrors pre-Renaissance empiricism. This approach may be dangerous for several reasons.

3. – Medical images and epistemological approaches

In the medical images domain, the empiricist approach is risky and raises several issues. First, it poses a scientific problem. Neglecting bias and data sampling limits lead to the inability to define the boundary conditions in which algorithms we develop can properly work. For example, we can have an algorithm which is able to diagnose a certain disease without specifying on which population it has been trained and tested. The correct stratification of population needs attention and there are many issues that still need to be addressed, such as the role of ethnicity [6] or how to correctly sample a comprehensive disease spectrum. There are also technical conditions that should be included in the analysis , such as the specific imaging system or image acquisition characteristics. However, datasets that contain medical images are usually small and technical information is usually not preserved to minimize privacy risk, making an in-depth analysis impossible.

Another issue that deals with medical images is the reliability of the ground truth. The ground truth is usually given by one or more than one physicians and depends on the task of the algorithm. When given in a consensus by many physicians, a variability in the assessment can always be observed despite the specific task of the algorithm. Guidelines proved to be effective in reducing this variability, but they cannot completely eliminate it [7]. Discussing the epistemological address of health research is important to face the problem of variability. We can frame it in the empiricist approach or in the constructivist one. The most recent research adopts the first one with the fundamental premise that there exists a single truth [8]. In this view, there is the underlying assumption that the gold standard can be made by the opinion of an expert. Several studies showed that the reliance on an expert opinion is not reliable [8] while the lack of a true gold standard makes an absolute comparison difficult [9]. We can conclude that variability exists and there is no way to delete it. Approaching this problem in a constructivist frame means to assume that there are several "truths" that depend on inherent biases, experience or judgments, not only among different individuals, but also within the same individual at different time points. Bridge *et al.* [8] state that clinicians should embrace the variability in the constructivist approach, suggesting that, instead of deleting it, they should study what is the acceptable variability amount for the specific clinical task.

Furthermore, AI applied to medical image analysis needs to be substantiated by a scientific hypothesis: we are always assuming that, given the constructed data and the context, there is a model which may solve the task we want to study. This means that we are using data that already contain the solution. In this context, the necessity of different expertise to interpret the model and also to evaluate the risks of the use of algorithms in the medical domain is clear. However, the collaboration among so many disciplines is not straightforward, since we need a shared language, methods and cooperative approach. Affirming that everyone is able to interpret the results without expertise in that specific domain is a reductionist and functionalist approach that ignores the socio-political context of the technoscientific practice. Finally, developing a tool or an algorithm in the medical images domain is not unrelated to social context, since it deals with hospitals, physicians availability, financial support, privacy management and also with how institutions that menage health, academic research and technology in general interact among themselves.

4. – Discussion and conclusions

The process towards a precision medicine able to personalize the diagnostic path and therapy is still long and we need to discuss about epistemological problems and issues. It is crucial in this field, because the lack of epistemological assumptions and, hence, of a consolidated method that integrates the high performance possibilities and limitations of DL techniques, turns immediately in ethical issues. In this context, the deepening of Explainable Artificial Intelligence (XAI) methods is mandatory to highlight possible biases and to limit the risk in clinical applications [10]. We should refuse the empiricist approach, since we cannot consider data as inherently truthful and meaningful, and medicine and health processes are complex.

REFERENCES

[1] Kuhn T. S., *The Structure of Scientific Revolution*, Vols. **I,II** (University of Chicago Press, Chicago) 1970.
[2] Kitchin R., *Big Data Soc.*, **1** (2014) 1.
[3] Hey T., Tansley S. and Tolle K. M., *Fourth Paradigm* (Microsoft Research) 2009.
[4] Anderson C., *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, in *Wired Magazine* (2008); URL `https://www.webcitation.org/5eBIPrc60`.
[5] Prensky M., *Innovate J. Online Educ.*, **5** (2009) 1.
[6] Gichoya J. W., Banerjee I., Bhimireddy A. R., Burns J. L., Celi L. A., Chen L. C., Correa R., Dullerud N., Ghassemi M., Huang S. C., Kuo P. C., Lungren M. P., Palmer L. J., Price B. J., Purkayastha S., Pyrros A. T., Oakden-Rayner L., Okechukwu C., Seyyed-Kalantari L., Trivedi H., Wang R., Zaiman Z. and Zhang H., *Lancet Digit. Health*, **4** (2022) e406.
[7] Khoo E. L., Schick K., Plank A. W., Poulsen M., Wong W. W., Middleton M. and Martin J. M., *Int. J. Radiat. Oncol. Biol. Phys.*, **82** (2012) 1923.
[8] Bridge P., Fielding A., Rowntree P. and Pullar A., *J. Med. Imaging Radiat. Sci.*, **47** (2016) 217.
[9] Van De Velde J., Vercauteren T., De Gersem W., Wouters J., Vandecasteele K., Vuye P., Vanpachtenbeke F., D'Herde K., Kerckaert I., De Neve W. and Van Hoof T., *Strahlenther. Onkol.*, **190** (2014) 628.
[10] Krenn M., Pollice R., Guo S. Y., Aldeghi M., Cervera-Lierta A., Friederich P., dos Passos Gomes G., Häse F., Jinich A., Nigam A. K., Yao Z. and Aspuru-Guzik A., *Nat. Rev. Phys.*, **4** (2022) 761.