Why Can't AI Understand Images as Man Does?

Aurel CODOBAN¹

¹ Professor PhD, University of Art and Design of Cluj-Napoca, "Dimitrie Cantemir" Cristian University Bucurest, and Philosophy Doctoral School at "Babes-Bolyai" University, aurelteodor@yahoo.com

Abstract: AI can identify images, but cannot understand them as man does. The problem of understanding the iconic signs is the analogy, which cannot be clearly operationalized. Nothing guarantees signification by analogy, because it is neither the necessary effect of a cause, as in the indicative signs, nor the obligatory consequence of a rule, as of symbols (words). But the analogy is also fundamental to the human condition because our Ego implies the presence of Other. Or, just as the images, the understanding of Other implies the analogy: he is a self like me, but another self than myself, that is, an analogous self. That is, you can understand the behavior of the other's activities and actions, even what he communicates as messages because you interpret them as if it were about yourself. Different from the human being who is an existent, in AI the essence precedes existence. Even if the algorithms of the analogy process will be infinitely perfected, that analogy will miss the interpretation that comes from the life of the existing one. AI knows digital, man understands analog; AI understands from knowledge, man knows from understanding.

Keywords: *AI; image; analogy; understand; knowledge; Other; Ego.*

How to cite: Codoban, A. (2020). Why Can't AI Understand Images as Man Does? *Postmodern Openings*, 11(4), 174-182. doi:10.18662/po/11.4/228

1. Difference between the way people and AI perceive images

The problem of the difference between the way people and AI perceive images has become very current in August - September months of 2019. "The way that the human mind, the human visual system, understands shape is a mystery that has baffled people for many generations, partly because it is so intuitive and yet it's very difficult to program" says Jacob Feldman, professor of psychology at the University Rutgers. (Feldman, 2019) Indeed, for some time now researchers have tested the ability of image recognition by intelligent machines, either on the basis of pixels or on the basis of neural networks (systems based on artificial intelligence). The problem is not that AI cannot identify the emblematic image, defining an object like a concept. Not that it cannot identify images offered from different perspectives on the same object when moving around it, spatially integrating the image with the movement: if it moves around the same bus, AI can identify the image by recognizing the identity of the object with all the difference between the images offered by different perspectives. (And we could dare to compare, as has already been done, probably to the horror of phenomenologists, this capacity of AI with the "eidetic intuition".) But if he receives different images in spatial and temporal discontinuity of the same object - a bus with wheels up or a bus lying on one side - he will think that they are completely different objects: snow plow, garbage car, etc. (Loclair, 2019) Just as he will think, that a kettle with a smooth surface and a design according to the pattern of golf balls is a golf ball; or that the image of a human face with a few changed pixels will be taken as a completely different object. "Researchers at Auburn University trained a neural network to fool Google's best image-recognition system, Inception, by rotating objects in space to novel positions. The lesson is that today's AI doesn't really identify objects very well". (Ray, 2018)

Researchers at Emory University, led by psychologist Stella Lourenco, said in a text from June 2019 that people do not evaluate an object as a pixel-processing computer, but they rely on an imagined internal skeleton. People recognize objects by visualizing such "skeletons" of them. (Ayzenberg & Lourenco, 2019). But James Elder, a professor at York University in Toronto, a specialist in comparing the way human and computer perceive images, has a relevant objection to the importance of skeletons for modeling and perceiving the object: "we can't really answer that question by only looking at the perception of skeleton-generated shapes. Because obviously in a world of skeleton-generated shapes, skeletons are

probably fairly important because that's the way those shapes were made". (Feldman, 2019) That is, people can sense the shapes that govern the configurations of the images of objects thanks to the ability to guess into them the intentions of others people, who made those objects or images.

Therefore, although AI can identify images, it cannot understand them as human intelligence does. This inability, which may explain the weirdness of the misrecognitions that AI makes, we think it marks the difference between AI and people, thus is worth deepening. It repeats on another level the problem that Searle answered in 1980, when Shank claimed that the computer program can simulate human ability to understand stories and it can provide answers to questions about what happens in the story. Searle's (1980) answer since then has remained an appropriate one: a powerful AI doesn't have much to say about thinking, because it's not about machines, but about programs and no program *per se* is enough for thinking. He imagined an interesting mental experiment: The Chinese Chamber. In an isolated room there is an English-speaking person with a rule book linking one set of Chinese symbols to another. A set of questions written in Chinese is entered through a window, and the person who does not know Chinese offers a series of answers. Someone outside the room may think that the person can understand and provide answers in Chinese, but this appearance is actually false.

2. We understand only analogically

The problem of understanding images by AI is similar, but a bit more complex. I would like to try to solve the problem making use of the terms of continental philosophy. But in order to understand it and deepen this problem, we must begin by putting the problem of the image as a sign. If we stick to the general landmarks of Pierce's semiotics, without losing ourselves in its initial entanglements and complications, but rethinking it from the current perspective of semiotics, it can be useful in understanding the complexity and complications of the sign that is the image. (Bergman & Paavola, 2019). Pierce had established, in principle, three types of signs: indexical, iconic and symbolic. Accepting the proposed simplifications we could say that the relation of significance for the indicative and symbolic signs is quite well defined operationally, while for the iconic signs - and especially for the image - it is more difficult to formulate precisely. Thus, for indices the relation is causal, based, in principle, on the relation cause-effect: most often the indicial sign is, in a broad sense, the effect of a cause or the consequence of a condition. And for symbols - and here Pierce's American

English is deceptive because it refers, quite differently from German romanticism, to mathematical, chemical or physical words and formulas - it is ultimately normative, meaning dictated by rules. Thus, if causality and normativity are the basis of meaning for the two types of signs, for the iconic signs there remains only the vague resemblance. We know the causes, we decode the rules, but to understand, we understand analogically! An old commentary of Paul Watzlawick's on the difference between digital and analogue computers, given that man is the only body that uses both digital and analog communication may be interesting: analog correspondents used instead of real sizes are inaccurate and this inaccuracy is amplified by the operation of analog computers themselves. While the digital computer is of perfect accuracy (within the limits of its memory) and is not only an arithmetic machine, but also a logical one. While analog computers cannot represent negative values, negation or any other truth function. (Watzlawick et al., 1967, paragraph 2.54).

The iconic sign - portrait, diagram, icon, technical image, medical image, visual image or even poetic image, i.e. metaphor - denotes its object only by resemblance or similarity. That is, it resembles or imitates the object, according to a quality that is shared by both and present in both. In fact, in this case the situation is more complex. Peirce differentiated the iconic signs into three classes: (a) the image, which is a quality of the object that it imitates or resembles with its own quality; (b) the diagram, whose internal relations represent by analogy the relations from something to which they refer; and (c) the metaphor, which represents something in parallel with something else, and it is not necessarily visual but it can be constructed from signs like words. As for me, I would add the new and widespread species of images obtained by means similar to photography, technical images, icons that use other resources for recordings, such as ultrasound, radar, Mars images, etc. In these cases, certain forms are related or attributed in some way to certain aspects that they denote, in some kind of relationships that are similar with cause - effect indicial relationship. In the case of the diagram, which can be especially geometric, arithmetic or algebraic, we have to deal with similarities of much stronger relationships, which work according to certain rules or norms.

So, in a much more gradual transition from index signs to symbolic signs, I would differentially place the coded technical images towards the index signs, and the diagrams, towards the symbolic ones (words, physical, mathematical formulas). It would remain in the midst of iconic signs, images and metaphors, which are based only on resemblance or similarity, that is, only on pure analogy, supported only by human perception. In order to

understand the metaphor - as in the case of the image - we must resort to the starting point of the bodily experience, he supports his older experientialist theory. (Lakoff & Johnson, 2003) (In a Review of this book Peter Norvig - from UC Berkeley - notes the difficulties of metaphor for AI.) (Norvig, 2019). These are the signs that, either in the identification in non-emblematic situations of the images or in the automatic translation of the sentences, the AI often misses them. (There remains the possibility that I prefer not to analyze it here, because I did it elsewhere, that the image should be more than a sign that is a speech.) (Codoban, 2011) Therefore, the great problem of recognizing images - either visual or poetic (and verbal) - by AI raises the analogy in a strict sense, that is, one that cannot be clearly operationalized. Nothing guarantees it because it is neither the necessary effect of a cause, nor the obligatory consequence of a rule. Therefore we have no criteria to appreciate the truth of the signs of the image type built on a pure likeness. About the truth and correctness of the indexes and "symbols", which are verifiable after all, we can at least discuss. Hence the suspicion of Plato, of the Jewish prophets, of the iconoclasm and Islam of the images.

However, the analogy is deeply embedded - even fundamental - in the human condition in the form of the presence of another, of otherness. What the analogy tells us about an iconic sign of the strict type of image - or metaphor - is that it is and is not at the same time identical with what it means - it just resembles with what it means. The same thing happens with the presence of another: he is a self - like me - but another self than myself, the only real one for me, because I can experience it as such. Our self can appear only because of the presence of another and under the gaze of another. The relationships I maintain with the Other are constitutive and decisive: they mediate my relationship with the world. If the world exists, it is to the extent that we are, to the extent that we are not alone, that is, to the extent that we are in relation to another / others. But in the case of the presence of the other, of another, we take an excessive step in the belief in the external reality, that is, besides the faith, which is not logically demonstrable, the existence of something outside ourselves, we consciously credit this presence of the other. However, this faith is one of the hardest to prove: the other just looks like me.

For the purely objective world of modern scientific knowledge, of operational rationality, other human beings can only be objects among other objects: humans can be perfectly assimilated to animals or machines, with small specific differences. The scientific method excludes from its field the presence of another as such: the behavior of another is only apparent and

momentarily more complex, less rigorous and less predictable, than the trajectory of a stone, the algorithm of a computer or the behavior of an animal. (In this sense Marvin Minsky, 1927-2016, mathematician and AI pioneer, can say about intelligence that it disappears as soon as we discover it.) On the other hand, communication cannot accept neither total human identity, nor absolute difference. If the identity is total, the communication, in the happiest case, becomes, at most, the transmission of signals, if it is not really useless and denied by its redundancy. Conversely, if the difference is total, communication is impossible, given the absence of a common code. In the first case communication would be useless, in the second it would be impossible. Otherness thus comes as a condition of the possibility of communication: to be able to communicate, my partner must be, like me, an I, but another one than myself! We could say that the otherness that calls into question the idea of communication as a mere transmission of signals, forces us to consider the higher level of the signs. More precisely, because of the analogy involved, Alterity presents the same complications and difficulties and is as difficult to understand as the iconic signs, and among them, especially as the image, the sign most dependent on the pure analogy. Otherness suggests that, as in the case of the image, analogy implies and demands interpretation. Only in the case of otherness the analogy relates to the existing (*Dasein*) as the image. To understand the other means to interpret it, to be able to tell a story about it. After all, you can support the analogy between yourself and the other's self because you can tell a story about the other as about yourself. That is, you understand in a narrative sense what is going on with him.

You can understand the behavior, activities and actions of the other, even what he communicates as messages, because you interpret them as if it were about yourself. To interpret means, very simply saying, to put yourself in someone else's place. But to put yourself in the place of someone else, with your life. Of course, the limits of your interpretation are the limits of your life. But people can say - and at the same time are somewhat forced to say - stories only because they exist, that is, not a Being endowed with an essence, but existence that gains its essence over time, existence that has a history. Probably, the subjectivity is related to the ability to tell a story and the self-consciousness is the protagonist self-perception of a personal story that his memory tells.

The AI can algorithmically identify the identity of a static object, which appears as an entity - such as the school bus that emblematically displays itself as a school bus - based on relationships between its parts or on the basis of the pixels that make up its image. But it can no longer

identify it in cases where it does not represent itself fully enough - overturned on one side, perpendicular to the road, or overturned with wheels up in the ditch - that is, in different positions where it no longer appears as emblematic - respectively coded in pixels - for its significance. The truth of the meaning of an image also contains the road to it, the story of how we reach that situation. The meaning of the image is, as in the case of the other's behavior, meaning plus the road to it, the subjective story of the road to it, the story of how meaning can be taken.

The concept of AI is based on the idea of building machines capable of thinking, acting and learning as humans. Artificial Intelligence is a broader concept, but learning is the most common application of AI. To be more accurate with our objection, we must say that the difference here is between the way AI learns and the way humans live, between the way an artificial organism learns and the man lives as a biological being. AI uses metadata sets (big computer data as dimensions) to "learn", create models and then be able to recognize what's unknown.

3. AI exists already knowing, while man knows existing in time

That's why for now - and most likely, I think, forever - AI operates with predefined significances rather than meanings. Because AI does not have the personal history of the road to significances, which it receives as such, readily elaborated by programming. In the case of man, the existence, the individual life - which, let us not forget, is that of an organic entity - is the one that allows the road to the truth of significances, to the meaning. We could say that AI exists already knowing, while man knows existing in time. AI has a very different relation with time: it remains imperturbable and constantly functional, capable of awake attention, which does not get tired and which retains every second. While man, as a biological entity, gets tired with the passage of time and is very oscillating in his attention, as time went on he may be hungry or thirsty, or he may have unpredictable desires. The relation of man with time is essential, therefore defining, for him. Man as a being has a premature existence: he lives, not knowing how to live better, neither strictly according to nature, nor strictly according to culture. In this sense, man, as Heidegger says, is a being of adventure.

It can be argued that history means having experiences over time, which may resemble the learning process. But the process of computer learning is not comparable with the existence in time, with the life of a biological entity, of an adventure being who is forced to live by choosing, before being able to be sure of the outcome of her choices. The problem of

learning is the key problem even if it concerns only part of the AI. AI has multiple sensors and huge memory, to detect physical data from the real world, is able to learn from its mistakes and can adapt to the new environment. The difference appears here, but it is of a different nature. The man, as a biological being, has the human experience of reality that usually is structured on three levels: first, the perception of sensitive objects, respectively the knowledge of things, the second, the knowledge about knowledge, and the third is the knowledge about second-order knowledge, respectively the understanding of their meaning in relation to his existence. This third order knowledge provides meaningful premises for one's life. Man has great flexibility and adaptability as long as the third order premises about his existence and the meaning of the world in which he lives remain intact. It is very difficult to say exactly what these premises are in human experience. An old American movie from 1972 - 'The Effect of Gamma Rays on Man-in-the-Moon Marigolds' - is one of the best metaphors for what I would like to say.

What separates man from AI is the existence of the level of consciousness of his experiences that is more than knowledge about second-order knowledge. Deep learning is a surrogate of existence, it is the artificial form of existence. It is, after all, an *a priori* deduction, without novelty and especially without the adventure and risk of Kantian synthetic judgment. Conversely, man bases his existence on adventurous and risky induction and intuition, without certainty. Different from man, in AI the essence precedes existence. Even if the algorithms of the analogy process will be infinitely perfected, that analogy will always lack the interpretation that comes from the life of the existing one. That's why AI knows digital, but only man understands analog. The AI knows, the man, the existent, interprets; AI understands from knowledge, in man, after Heidegger it is the opposite: he knows from and after understanding. Both can interpret images and otherness (other humans and animals), the first through learning and algorithms, the second through the story of his own existence, through empathy and the imaginary experience of art. The difference between AI and man in relation to the iconic signs that are the images is between reading the signs and interpreting the signs.

Perhaps man has reached the limit of his abstracting abilities, as it results from the psychiatric implications of the biographies of the great mathematicians, but his capacity for understanding nevertheless has wider limits than the ability to abstract. From the perspective of the existence of man, AI, Big-Bang and God are singularities outside the categories of space and time that can be more properly understood in terms of negative

theology. But this perspective is reversible and we can ask ourselves whether or not from a future AI - or from divinity - will man also be a singularity, something that, even if a future absolute AI can know, cannot understand?!

References

- Ayzenberg, V., & Lourenco, S. F. (2019, June 27). *Skeletal descriptions of shape provide unique perceptual information for object recognition*. <https://www.nature.com/articles/s41598-019-45268-y>
- Bergman, M., & Paavola, S. (2019). *The Commens dictionary. Digital companion to C. S. Peirce*. Commens. <http://www.commens.org/dictionary>
- Codoban, A. (2011). *Imperiul comunicării. Corp, imagine și relaționare* [The empire of communication. Body, image and relationships]. Idea.
- Feldman, J. (2019, September 19). What you see that the machine doesn't. <https://mindmatters.ai/2019/09/what-you-see-that-the-machine-doesnt/>
- Lakoff, G., & Johnson, M. (2003). *Metaphors we live by*. The University of Chicago Press.
- Loclair, C. M. (2019). Narciss. <https://christianmioloclair.com/narciss>
- Norvig, P. (2019). Review of metaphors we live by. <https://norvig.com/mwlb.html>
- Ray, T. (2018, November 30). Google's image recognition AI fooled by new tricks. <https://www.zdnet.com/article/googles-best-image-recognition-system-flummoxed-by-fakes/>
- Searle, J. (1980). Minds, brains and programs. *The Behavioural and Brain Sciences*, 3(3), 417-457. <https://doi.org/10.1017/s0140525x00005756>
- Watzlawick, P., Bavelas, J. B., & Jackson, D. D. (1967). *Pragmatics of human communication: A study of interactional patterns, pathologies, and paradoxes*. Norton.