

# Quand les artistes s'emparent des réseaux de neurones

Par Olivier Perriquet

Artiste

En avril 2016, une quinzaine de robots Pepper entonnait publiquement l'*Hymne à la joie*. La presse titrait : « *Japon : une chorale de robots massacre du Beethoven* ». Elle nous rassurait ensuite en nous proposant une interprétation *humaine*, nettement plus poignante, sous la baguette de Daniel Barenboïm. Les fantasmes que suscite la procréation artistique accordée aux « machines » se mêlent en effet à une crainte, que l'on pourrait nommer le *mythe de Frankenstein* dont on retrouve le thème central (un être artificiel s'émancipant dangereusement de son créateur humain) à différentes époques et sous des formes variées (Victor, Pygmalion, le Golem, ...). Les journalistes concédaient toutefois avoir été lors de cette représentation « *émus comme des parents qui vont assister à un spectacle de leurs enfants* ». Si nous en sommes en robotique et en intelligence artificielle à un stade qui pourrait évoquer celui de l'enfance, la comparaison s'arrête là, car nulle conscience, nulle intention, n'habite ces artefacts, aussi sophistiqués qu'ils puissent être. Pas pour l'instant, en tous cas. L'intelligence artificielle suscite pourtant un intérêt notoire de la part des artistes tout comme de celle des institutions, en témoignent la tenue d'expositions comme « *AI: More than Human* » au Centre Barbican à Londres (2019) et, simultanément, « *Entangled reality – Living with Artificial Intelligence* » à la HeK (Maison des Arts Electroniques) de Bâle, ou bien encore « *Neurones : les intelligences simulées* » à Beaubourg au printemps dernier, et bien d'autres. Un tel intérêt se comprend au regard des avancées scientifiques récentes en intelligence artificielle. Les cinq dernières années ont en effet été le théâtre de progrès significatifs en apprentissage automatique par réseaux de neurones artificiels : ils ont surpris les experts du domaine eux-mêmes. Or selon l'expression de Kurt Vonnegut, certains artistes sont dans ce contexte comme les cellules sensorielles à la surface du corps social, tirant la sonnette d'alarme lorsqu'une société est en grand danger.

## Un art réticulo-neuronal ?

À l'été 2015, les réseaux sociaux ont vu proliférer des images psychédéliques aux reflets irisés et aux contours caractéristiques, conglomérats de chiens en convolution aux yeux proéminents, rappelant les célèbres compositions d'Arcimboldo. Ces images « virales » (un nouveau type de memes<sup>1</sup>) présentaient des scènes reconnaissables passées au filtre de *Deep Dream*, un système initialement développé par les scientifiques de Google à des fins d'analyse des réseaux de neurones convolutifs (*Deep learning*) et détourné par des artistes pour sa capacité à produire ces images inédites. Si des chiens ou d'autres animaux apparaissaient, c'est que le réseau

avait été entraîné par la firme américaine à reconnaître ce type d'images et que la plupart des internautes n'avaient pas eu à cœur de le ré-entraîner – entreprise qui leur aurait demandé plusieurs mois. Le résultat est tautologique : si on lui apprend à reconnaître des chiens, le réseau de neurones « verra » des chiens. Le phénomène n'est pas bien différent de ce qu'on nomme chez les humains la paréidolie, c'est à dire la propension à reconnaître des formes significatives dans ce qu'ils voient, même si celles-ci sont fortuites : le visage du Christ sur un toast grillé, la forme d'un éléphant dans un rocher (la toponymie reflète parfois ce que de nombreux visiteurs ont communément éprouvé), ou l'esquisse d'un dromadaire ou d'un lapin dans les nuages, ou bien simplement un sourire dans la juxtaposition typographique « :) ». *Deep Dream* pousse ce principe à l'extrême et fait dégorger les images de tous les chiens qu'elles peuvent « évoquer ». Mais là aussi, il faut se garder de tout anthropomorphisme : outre le caractère purement mécanique et dénué d'intention du traitement de l'information par un réseau de neurones artificiels, celui-ci n'est neuro-mimétique que jusqu'à un certain point, et l'oeil n'est pas non plus un capteur matriciel. Ces images ne sont pas plus mystérieuses que celles obtenues grâce à un filtre appliqué avec un logiciel de traitement graphique.

En quelques années, les techniques d'apprentissage par réseaux de neurones se sont diversifiées, les codes sont devenus accessibles, un ensemble d'artistes a commencé à en exploiter les possibilités. La presse anglophone s'est emparée de ce phénomène, parlant d'un *AI art* et le dénigrant, parfois pour son hypermédiation (*AI hype*), ou notant que, du point de vue de l'histoire de l'art, un certain nombre de questions prétendument soulevées par les *AI artists* ont déjà largement été explorées depuis l'art génératif apparu à la fin des années 1950, au moment où se concevaient les premiers ordinateurs. En particulier, celle de l'auteur (doit-on attribuer cette qualité à la machine ou au concepteur de la machine ?) ou de la possibilité d'une génération infinie d'images. Le jugement esthétique, comme l'explique Frieder Nake, un des pionniers en *computer art*, doit s'exercer non sur le résultat (l'artefact) mais sur le procédé mis en œuvre. On comprend alors que cette réponse convoque encore d'autres volets de l'histoire de l'art, qui n'ont plus nécessairement à voir avec l'usage de la technologie (les instructions données par Sol Lewitt pour la réalisation de ses *Wall drawings*, par exemple, s'apparentent à – ou devrait-on dire : sont – un algorithme).

<sup>1</sup> Un meme, anglicisme du mot *meme*, formé à partir du terme grec *mimesis*, est un concept, un texte, une image, une vidéo, massivement repris, décliné et détourné sur Internet. Le terme a été forgé par Richard Dawkins dans les années 1970 et popularisé dans *The Selfish gene*.

## Des artistes antagonistes

En octobre 2018, le collectif français Obvious, composé d'un informaticien et de deux étudiants en école de commerce, tous trois âgés de 25 ans, vend aux enchères chez Christie's à New York pour la somme de 432.500 \$ un tableau calculé par un réseau de neurones et imprimé sur toile : *le portrait d'Edmond de Belamy*, un bourgeois fictif du XIX<sup>e</sup> siècle. Les images ont été créées à l'aide d'une architecture particulière appelée GAN (Generative Adversarial Networks, ou Réseaux Antagonistes Génératifs) qui consiste en la mise en relation dialectique de deux réseaux de neurones, l'un jouant le rôle de « faussaire » en produisant des images « à la manière de » à partir d'un corpus qui l'a nourri (ici 15.000 portraits classiques allant du XV<sup>e</sup> au XX<sup>e</sup> siècle) et l'autre jouant le rôle de « l'expert », dont la réponse permet au faussaire de s'améliorer. Si *Belamy* est une évocation de Ian Goodfellow, le chercheur en IA inventeur des GANs, dont il semble être une traduction, le journaliste Laurent Carpentier note, dans *Le Monde*, que ce nom convoque également le personnage Bel-Ami de Maupassant... figure de l'arriviste. L'exploit ressemble en effet plus à une opération de marketing – dont le trio maîtrise bien la langue – qu'à un réel acte artistique. Il donne surtout la mesure de la fascination que peut exercer sur le public ce type de techniques.

Des artistes de premier plan, qui ne revendiquent pas l'appartenance au mouvement susmentionné, se sont récemment emparés des GANs pour produire des œuvres au contenu manifestement plus trouble. Hito Steyerl a ainsi présenté à la Biennale de Venise en 2019 une installation monumentale intitulée *This is the future*, mettant en scène, avec une certaine dose d'ironie, l'intelligence artificielle comme un oracle. On y reconnaît son style caractéristique, une narration poétique, mélange d'images de la culture pop, de séquences documentaires et d'animations par ordinateur, dans une œuvre qui présente une vision de notre monde à la fois humoristique et effrayante, mêlant la satire et le sérieux, tout en abordant une série de sujets allant de la mondialisation à l'effondrement économique. Grégory Chatonsky présentait quant à lui au Palais de Tokyo au même moment l'exposition *Terre seconde*, où une intelligence artificielle génère une version alternative de la Terre, comme un lot de consolation, ou comme la promesse d'une résurrection future, après que l'humanité ait disparu. Créée en 2017 par Trevor Paglen, *Adversarially Evolved Hallucinations*, est une série d'images générées également par des GANs qui ont été entraînés sur des corpus spécifiques. Le réseau de la série

*Monsters of Capitalism* a ainsi été entraîné à voir des monstres ayant été historiquement utilisés comme des allégories du capitalisme : le *Frankenstein* de Mary Shelley, la figure du Vampire (Marx évoquait des morts-vivants se nourrissant du sang de la classe ouvrière), des zombies haïtiens, des poulpes, etc. C'est un environnement non moins inquiétant qu'a mis en scène Pierre Huyghe à la galerie Serpentine à Londres en 2018, où cinq écrans présentent ce qu'il nomme des « images mentales », produites dans le laboratoire du Dr Kamitani à Kyoto. L'artiste a présenté à des individus un ensemble d'images pendant que leurs ondes cérébrales étaient scannées ; un GAN recomposait ensuite cette imagerie mentale en puisant dans une banque d'images préexistantes. Défilant à grande vitesse, les images produites sont livrées telles quelles, brutes comme des esquisses qui n'ont pas encore trouvé leur forme définitive, créant une atmosphère qui rappelle celle de Lynch ou de Cronenberg. Les sons qu'on entend ont tout à la fois de la machine, de l'humain et de l'animal. Il en va de même pour l'odeur étrange diffusée dans l'air. La poussière s'accroche aux chaussures des spectateurs et se répand dans les jardins, les rues, le métro. Parfois, les images des panneaux se figent, en fonction des conditions de l'espace de la galerie : la température, l'humidité et la lumière, le nombre de visiteurs, le temps qu'il fait dehors, le rythme des naissances des mouches qui peuplent également la galerie...

La manifestation Transmediale (Berlin 2020), intitulée *Adversarial Hacking in the Age of AI*, a précisément choisi son thème en référence aux GANs, tout en donnant à « adversarial » la signification d'un engagement et d'un rapport de force : il s'agit moins d'appliquer le style de Van Gogh sur une photo de vacances ou de gagner le gros lot au marché de l'art que d'affirmer le rôle politique qu'ont les artistes vis à vis de ces technologies. Comme le souligne le chercheur Fabian Offert, l'avenir de *l'LA art* ne réside pas tant dans son utilisation pour la création d'images que dans son potentiel critique face à une utilisation de plus en plus industrialisée de l'intelligence artificielle. Reste en premier lieu à prendre conscience de l'envoutement que peuvent procurer ces images qui sont sans doute moins mystérieuses qu'on veut nous le faire croire. Il faut par exemple se souvenir qu'au XVIII<sup>e</sup> siècle, croyant réellement assister à l'apparition de spectres et de revenants, le public s'enfuyait en hurlant lors des *Fantasmagories* de Robertson. Mais c'est précisément ce qui fait la magie de tout spectacle, pourvu qu'on comprenne bien où l'on met les pieds.