Images trompeuses et modèles d'intelligence artificielle

Misleading images and artificial intelligence models
Study day and workshops

Journée d'étude et ateliers
12-13 Oct. 2023 au Fresnoy
Studio national des arts contemporains
Tourcoing, France
Organisation Gaëtan Robillard et Renée Bourassa
Groupe de recherche international Arcanes CRILCQ
Université Laval (Canada)

CRSH SSHRC ARCANES UNIVERSITÉ LAVAL Fonds de recherche Société et culture Québec CRILCQ CENTRE DE RECHERCHE INTERUNIVERSITAIRE SUR LA LITTÉRATURE ET LA CULTURE AU QUÉBEC LE FRESNOY STUDIO NATIONAL DES ARTS CONTEMPORAINS Tourcoing umr sciences et technologies de la musique et du son ircam Centre Pompidou Culture cnrs SORBONNE UNIVERSITÉ

Images trompeuses et modèles d'intelligence artificielle

Misleading images and artificial intelligence models

Journée d'étude et ateliers les 12 et 13 octobre
Au Fresnoy - Studio national des arts contemporains. Tourcoing, France

Intervenants

Constantine Boussalis
Terence Broad
Olga Kisseleva
Alain Lioret
Nicolas Obin
Frank Madlener
Kazushi Mukaiyama
Vincent Nozick
Fabien Richert

Organisation

Gaëtan Robillard
Renée Bourassa
Université Laval et CRILCQ
Avec l'aide d'Éric Prigent
Le Fresnoy

Images trompeuses et
modèles d'intelligence artificielle

Misleading images and artificial intelligence models

12 octobre. Journée d'étude

8h30 - 9h00	Accueil
9h00 - 9h15	Introduction, Gaëtan Robillard et Renée Bourassa.
9h15 - 10h30	My lines drawn by AI (en ligne), Kazushi Mukaiyama. Artful intrusions—A hacker's guide to generative AI, Terence Broad.
10h30 - 11h00	Pause café
11h00 - 12h15	Mesonet : Détecter la falsification de vidéos deepfakes, Vincent Nozick. Deepfakes ou synthèse vocale, Nicolas Obin.
12h15 - 14h00	Pause déjeuner
14h00 - 15h15	Studying images with the help of computer vision, Constantine Boussalis. Agents conversationnels (en ligne), Fabien Richert.
15h15 - 15h45	Pause café
15h45 - 17h00	De l'intelligence en essaim à l'intelligence artificielle, et vice versa, Olga Kisseleva. Après l'intelligence artificielle : l'éthique, les tics et l'esthétique, Frank Madlener.
17h00 - 17h45	Conclusion et débat, Gaëtan Robillard et Renée Bourassa.

Argumentaire

Avec les récents développements de l'apprentissage profond (deep learning), force est de constater que les médias génératifs se hissent au rang des médias de masse. DALL·E, Midjourney et Stable Diffusion sont trois exemples de modèles qui permettent au profane de générer des images synthétiques à partir d'instructions écrites, des prompts, plus ou moins sommaires. Cette massification des médias génératifs va de pair avec une ère de l'information en crise – une ère « post-vérité » (Shepard, 2022) dans laquelle il devient difficile de distinguer le vrai du faux, aussi bien que de départager l'image authentique de l'image falsifiée (avec par exemple les images issues des technologies appelées deepfake). Dans ce contexte, des problèmes importants apparaissent : les régimes traditionnels de l'authenticité sont en crise ; il est nécessaire d'en comprendre les raisons.

Pour une part, l'entraînement des Intelligences Artificielles (IA) génératives requiert un très grand nombre de données dont l'origine ou les biais sont habituellement peu mis en avant, et dont les pratiques de travail « soulèvent de sérieuses inquiétudes d'ordre éthique, méthodologique et épistémologique » (Crawford, 2021). D'autre part, en s'opposant aux approches symboliques des algorithmes (Cardon et al., 2018), les réseaux de neurones entraînés, notamment par leur échelle, rendent particulièrement difficile leur étude critique et technique. Il faut aussi évoquer le rôle des modèles d'apprentissage profond dans le domaine de la vision-machine. Lié au pouvoir ou au contrôle, et impliquant de nouvelles stratégies de contournement comme l'obfuscation (Fourmentraux, 2023), le voir des machines entraînerait une culture visuelle dans laquelle la perception humaine ne serait plus nécessairement au centre (Paglen, 2016), ce qui n'est pas sans poser de questions.

Dès lors, comment envisager l'historicisation ou l'explicabilité des images artificielles ? Comment caractériser les nouveaux régimes d'authenticité des médias génératifs et synthétiques ? Quelles façons de faire peuvent paraître ? Même si des auteurs comme Lev Manovich, Emmanuele Arielli, ou Yuk Hui ont proposé un certain nombre d'approches, soulignons que du point de vue des images, une philosophie technique articulée à une compréhension esthétique de ces objets semble faire défaut. Ainsi, la journée d'étude entend rassembler diverses approches théoriques des modèles d'IA générative dans l'art et les médias : réseau génératif adversaire (GAN), auto encodeur variationnel (VAE), ou encore modèle de diffusion, – afin d'en dégager des formes d'authenticité. (Nous comprenons ici le terme d'authenticité à la lumière des écrits de Walter Benjamin.) Il s'agit ainsi de développer une meilleure compréhension des mécanismes activant les « puissances du faux » (Deleuze, 1985), en particulier du point de vue de l'image algorithmique, de sa production et de sa réception.

Pour résumer, l'événement se concentrera principalement sur deux aspects. Tout d'abord, nous nous intéresserons à la manière dont les pratiques artistiques contemporaines et la culture visuelle sont impactées par l'accroissement des modèles génératifs et des médias synthétiques dans l'espace informationnel, et à la manière dont ces pratiques artistiques ouvrent de nouvelles perspectives. Deuxièmement, nous nous intéresserons aux implications techniques et sociopolitiques comprises par l'utilisation de modèles d'apprentissage automatique (ou IA) et par la diffusion inhérente de contenus trompeurs dans les médias ou la sphère publique. Il s'agira notamment d'étudier les relations entre désinformation, médias génératifs et modèles informatiques.

Abstract

How to think of the historicization or the explainability of artificial images? How to characterize the new authenticity regimes of generative and synthetic media? How do artists tackle this issues? To answer these questions, the seminar brings together various theoretical approaches to generative AI models in art and media: generative adversarial network (GAN), variational autoencoder (VAE), or diffusion model, – in order to identify forms of authenticity. The event will focus on mainly two aspects. Firstly, we are interested in how visual culture and artistic practices are impacted in regards to the surge of generative models. We will look into how those artistic practices can redefine themselves in this context, leading to new perspectives. Secondly, we will focus on the technical and socio-political implications of the use of machine learning (or AI) models, in regards to the dissemination of misleading content in the media or the public sphere. This will include studying the relationships between disinformation, generative media and computer models.

13 octobre. Ateliers

La journée d'étude donnera suite à une série d'ateliers. L'objectif de cette seconde journée est de rassembler diverses approches pratiques des images algorithmiques et génératives, à la fois d'un point de vue critique et d'un point de vue expérimental.

9h-10h30. Machine Unlearning, Gaëtan Robillard

Nous nous intéresserons ici au modèle MobileNET, un réseau profond entraîné pour la classification d'images. Il s'agira d'interagir avec ce modèle à travers un programme JavaScript de quelques lignes (librairies p5.js et ml5.js). Grâce à des input graphiques et aléatoires, nous chercherons à déjouer la reconnaissance d'image pour ensuite explorer les biais d'entraînement du modèle.

11h-13h. GANs et usage de faux, Gaëtan Robillard

Nous nous appuierons sur un environnement de travail spécialement conçu pour la recherche sur les réseaux adversaires génératifs (GAN). Cet environnement est pensé pour quiconque souhaiterait découvrir, façonner ou critiquer ce type de modèle, en particulier dans un contexte de recherche et d'expérimentation visuelle. Nous explorerons des jeux de données et nous verrons comment produire des images en mouvement. L'atelier ouvrira sur la présentation de prompts pour la génération vidéo.

14h-16h. Art Génératif Quantique, Alain Lioret

On verra comment utiliser l'informatique quantique pour la générativité artistique et comment les techniques avancées d'intelligence artificielle ouvrent-elles de nouvelles perspectives avec les algorithmes quantiques. On découvrira comment les principes de superposition, d'intrication, de téléportation quantiques peuvent apporter des outils nouveaux pour les artistes.

Bibliographie

Sofian Audry, *Art in the Age of Machine Learning*, Cambridge, MA, MIT Press, 2021.

Walter Benjamin, *L'Œuvre d'art à l'époque de sa reproductibilité technique*, Allia, 2003.

Anthony Glyn Burton et Wendy Hui Kyong Chun, *Algorithmic Authenticity : An Overview*, Lüneburg, Meson Press, 2023.

Cristian S. Calude et Giuseppe Longo, « The Deluge of Spurious Correlations in Big Data », *Foundation of Science*, 22 (3), 2017.

Dominique Cardon, Jean-Philippe Cointet et Antoine Mazieres, « La revanche des neurones : L'invention des machines inductives et la controverse de l'intelligence artificielle », *Réseaux*, La Découverte, 2018.

Kate Crawford, *Contre-Atlas de l'intelligence artificielle*, Paris, Éditions Zulma, 2021.

Gilles Deleuze, *L'image-temps. Cinéma 2*, Paris, Les Éditions de minuit, 1985.

David Foster, *Generative Deep Learning : Teaching Machines to Paint, Write, Compose and Play*, Sebastopol, CA, O'Reilly Media, 2019.

Jean-Paul Fourmentraux, *antiDATA*, Dijon, Les presses du réel, 2020.

Sophia Kourkoulakou, Jean-François Jégo et Gaëtan Robillard, « GAN et Arts (Atelier) », Université Paris 8, 2022.

Lev Manovich et Emanuele Arielli, « Even an AI could do that », in *Artificial Aesthetics: A critical guide to AI, media and design*, Manovich and Arielli, 2021.

Fabian Offert et Peter Bell, « Perceptual bias and technical metapictures: critical machine vision as a humanities challenge », *AI & SOCIETY*, 36 (4), 2021.

Trevor Paglen, « Invisible Images (Your Pictures Are Looking at You) », *The New Inquiry*, 2016.

Matteo Pasquinelli, *The Eye of the Master. A Social History of Artificial Intelligence*, Brooklyn, NY, Verso Books, 2023.

Philipp Schmitt, « Introduction to Blueprints for Intelligence », 2021.

Mark Shepard, *There are no Facts: Attentive Algorithms, Extractive Data Practices and the Quantification of Everyday Life*, Cambridge, MA, MIT Press, 2022.

George Stiny et James Gips, *Algorithmic aesthetics : computer models for criticism and design in the arts*, Berkeley, University of California Press, 1978.