Workshop on Tangible xAI

Leonardo Angelini [1,2], Nadine Couture [3], Mira El Kamali [2], Quentin Meteier [2], and Elena Mugellini [2]

[1] *School of Management, HES-SO//Fribourg, Chemin du Musée 4, Fribourg, 1700, Switzerland*
[2] *HumanTech Institute, HES-SO//Fribourg, Boulevard de Pérolles 80, Fribourg, 1700, Switzerland*
[3] *Univ. Bordeaux, ESTIA Institute of Technology, Bidart, 64 210, France*

Abstract

The goal of this workshop is to discuss the potential benefits and the open challenges of giving a physical form to Artificial Intelligence (AI), towards the definition of tangible, or graspable AI. In the workshop we will focus on the use-case of a convolutional Neural Network for image analysis and we will carry out a hands-on paper prototyping activity to imagine tangible interactive AI-powered systems where critical parameters of the Neural Network are physicalized. Such systems have the potential to lower the barrier for AI education and for making AI systems more trustable and explainable.

Keywords

Tangible AI, Graspable AI, xAI, CNN, education

1. Background

Artificial intelligence systems have a critical flaw, they are taught and designed by the developers that build them. The one using an AI system (teacher and final user) is often not an AI expert. Thus, the AI teaching part can easily be very biased because the teacher is a computer scientist highly knowledgeable in machine learning, who will eventually optimize the learning and target their own goals without a broader view of the problems. A new research domain emerged from the current limitations of human-AI interactions, the Human-Centered AI (HCAI) [1]. This field combines research on AI algorithms with user experience design methods to shape technologies that amplify, augment, empower, and enhance human performance. Researchers and developers for HCAI systems value meaningful human control, putting people first by serving human needs, values, and goals. There is a sub domain of HCAI, called "interactive machine learning" [2]. Interactive Machine Learning (IML) is the design and implementation of algorithms and intelligent user interface (IUI) frameworks that facilitate machine learning (ML) with the help of human interaction, i.e., is an interaction paradigm in which a user or a user group refines a mathematical model to describe a concept through iterative cycles of input and review. Dudley and Kristensson [3] note that while HCAI has gained traction and interest in research and actual applications, there is a real lack of knowledge on interfaces' design in such applications. They thus synthesized observations done in the field of IML with relevant HCI theories and propose a list of principles for interface design for IML:

- Make task goals and constraints explicit
- Support user understanding of model uncertainty and confidence
- Capture intent rather than input
- Provide effective data representations
- Exploit interactivity and promote rich interactions
- Engage the user

[1]Proceedings of ETIS 2022, November 7–10, 2022, Toulouse, France
EMAIL: leonardo.angelini@hes-so.ch (L. Angelini); n.couture@estia.fr (N. Couture); mira.elkamali@hes-so.ch (M. El Kamali); quentin.meteier@hes-so.ch (Q. Meteier); elena.mugellini@hes-so.ch (E. Mugellini)
ORCID: 0000-0002-8802-5282 (L. Angelini); 0000-0001-7959-5227 (N. Couture); 0000-0003-2895-6867 (M. El Kamali); 0000-0003-2568-9898 (Q. Meteier); 0000-0002-0775-0862 (E. Mugellini)

Amershi et al. [14] proposed 18 guidelines for Human-AI interaction, suggesting, among others, that the AI should provide global controls to the user, supporting efficient invocation, dismissal and correction.

Tangible User Interfaces are good candidates to address these challenges. Indeed, Tangible interaction proposes to use the affordances of physical objects to interact with data, exploiting metaphors and interaction concepts borrowed from the physical world. Using physical objects for explaining and interacting with AI is a very recent research trend. To the best of our knowledge, there have been only a couple of workshops on this subject [4,5,6], identifying pathways for future research, but no original research has been produced so far.

In our workshop, we would like to explore further such topics, opening new research pathways in collaboration with the workshop participants. The central question that would be tackled in the workshop is:

- Can we use physical representations and controls to make AI more explainable and trustable?

Therefore, the main goals of the workshop will be to:

- Explore how the principles of tangible interaction could increase trust and understandability of deep learning algorithms
- Reflect on how such principles could be included in the design of future XAI systems

In the workshop, we will focus on the use case of image analysis with Convolutional Neural Networks (CNNs). CNNs are a popular network topology for image classification, constituted by several layers of convolutional filters and subsampling filters (e.g., max pooling). A typical representation of a CNN is shown in Figure 1.

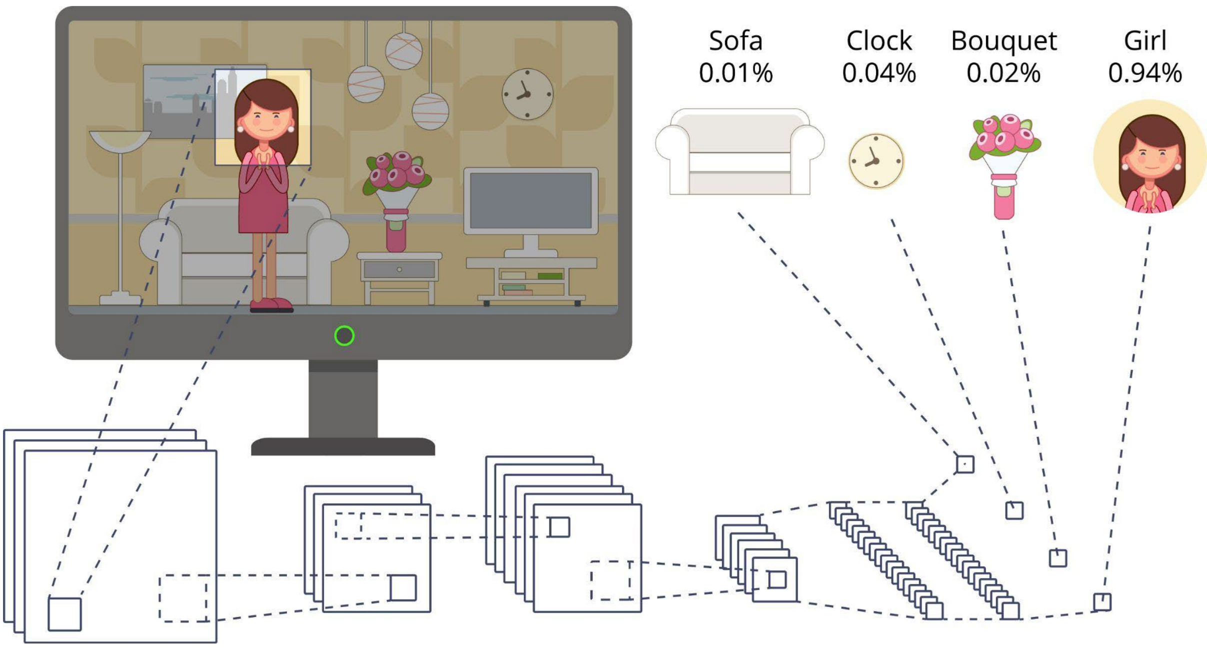

Figure 1. Example of CNN for image recognition (Adobe Stock).

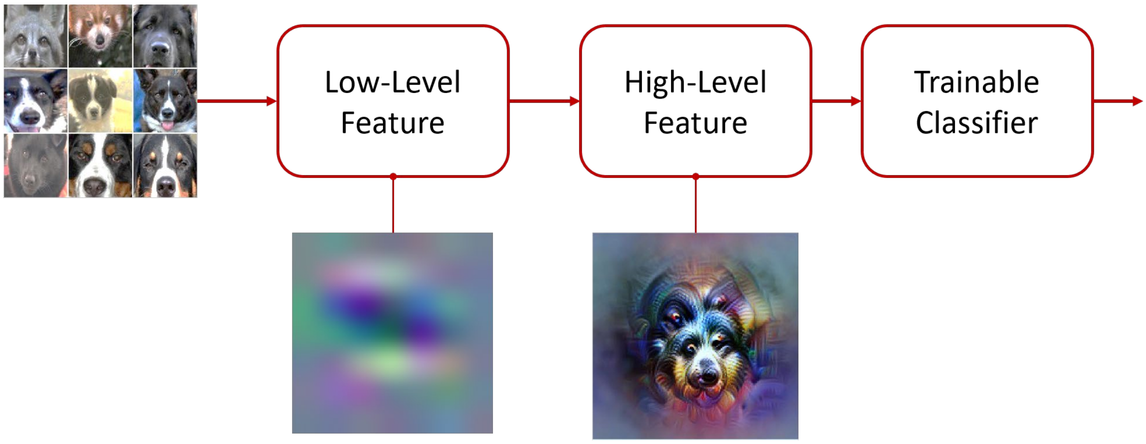

Figure 2. Example of feature visualization for the different layers of a CNN, from [7]

To better understand how CNNs work, Zeiler and Fergus [8] provided a visual representation of what each layer of the architecture would learn, showing that the first layers would focus on low-level features, such as edges in the image, while the last layers would be activated with the most complex visual features, such as part of the objects or the whole object that should be recognized (Figure 2).

Typically, topologies differ by the size of the windows on which the convolution or the subsampling is applied, the depth of each layer and the shift distance applied to each window. We believe that having a 3D representation of the network and allowing to play physically with such parameters to alter the behavior of the algorithm could help students learn how CNNs work. Eventually, it could also help researchers to get a better understanding of what is happening inside the network and to find new network topologies that could improve the performance, or the explainability of the algorithms. Designing explainable and physically controllable AI algorithms should help increase trust in AI and spread the adoption of Interactive Machine Learning among a larger user base.
The workshop will build on previous research carried out by the organizers in the field of the Internet of Tangible Things [9], in particular to reflect on tangible properties and AI system properties that could be physicalized, and on the previous research topics of growth, unpredictability and intentionality highlighted by Ghajargar et al. [6].

2. Workshop Structure

The duration of the workshop was half day (4 hours).
The workshop was organized as follow:

- 15 minutes: Participants' introduction with few fords on their previous experience on Machine Learning and Tangible Interaction (previous knowledge is not considered as a prerequisite, but was useful to form heterogeneous groups)
- 20 minutes: Short presentation by the organizers of the main important concepts (xAI, deep learning and CNN) and recent work on tangible xAI
- 50 minutes: Group formation (4 to 6 people) and 45 minutes discussion on potential use cases and benefits of Tangible XAI. The main research questions investigated are the following:
    - How can a CNN be materialized?
    - How can we physically explore and control the functioning (feature learning and classification) of a CNN algorithm?
- 15 minutes: Discussion of the ideas, selection of a use-case application for each group and instructions for the next phase
- 15 minutes break
- 60 minutes: Paper prototyping of the selected use-cases
- 20 minutes: Presentation of the ideas, user testing

- 30 minutes: Analysis of the main aspects that should be physicalized and the expected impact on the users. Using IoTT cards as starting point of the reflection [9]
- 10 minutes: Conclusion and next steps

3. Workshop results

About 15 people attended our workshop. We formed three groups, ensuring the presence of at least one knowledgeable person in the domain of machine learning and in particular of CNNs. To get a better understanding of how CNNs work we used four online tools for the visualization of CNNs networks [10-13]. These tools allowed the participants to familiarize with the main concepts and hyperparameters of CNNs, such as the convolution operation, the number of layers, the kernel size, the stride, padding, and the features that are typically learned at each layer. The three groups decided to work on two main topics: tangible interfaces for learning the functioning of CNNs and interactive machine learning with tangible interfaces in healthcare.

3.1. Tangible interfaces for learning CNNs

Two groups explored the potential of tangible interfaces for representing the CNNs structure, enabling the possibility of manipulating physical artifacts to manipulate the architecture of the network.

The first group explored the metaphor of a transparent box (in contrast to the black box representation usually perceived with deep learning algorithms) to represent a layer of the network. In the box a transparent filter would overlap with an image in order to give the different output of the layer (Figure 3, left). Stacking more boxes would be equivalent to adding more layers to the network (Figure 3, center).

The second group used instead cardboards of different colors to represent different operations (i.e., convolution and max pooling). Stacking cardboards on a table was perceived as an intuitive interaction for building the different layers of a CNN (Figure 3, right).

Figure 3. Workshop prototypes from group 1 (left and center) and from group 2 (right)

Pulling a thread from a wool yarn was also identified as another tangible interaction metaphor for exploring and visualizing the features learned by a CNN.

3.2. Interactive machine learning with tangible interfaces in healthcare

The third group worked on an application of tangible XAI for the healthcare domain. In the context of CNN for breast cancer diagnosis, doctors may struggle in understanding, and trusting, the model predictions. In order to make informed decisions, workshop participants proposed an interactive machine learning system where the doctor can use playdough to cover parts of the x-ray images and see in real-time what would be the effect of ignoring that part on the algorithm prediction (Figure 4, left). One of the participants of group 3 presented also a system developed few years before (unpublished work) to explore the functioning of CNNs through a tangible tabletop installation (Figure 4, right).

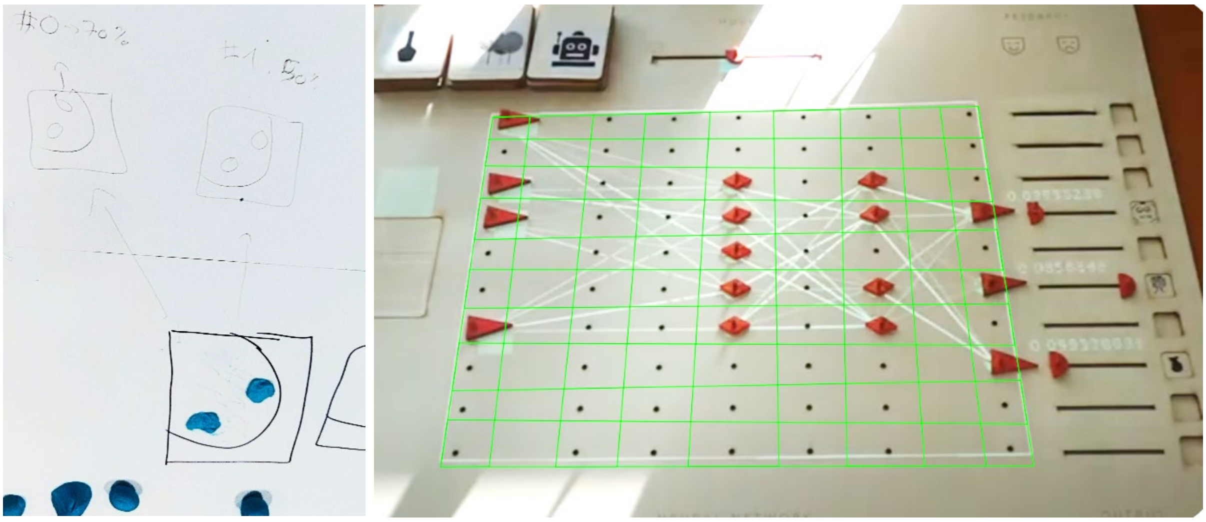

Figure 4. Workshop prototype from group 3 (left) and tangible tabletop system for learning CNNs (right) presented by Jérémy Laviole. It was realized in the context of the Hackathon Hack'1 Cerveau, in 2017 at Cap Sciences (Bordeaux, France)

3.3. Workshop discussion and conclusion

Based on the ideas and paper prototypes presented by each group, we identified several insights that could inspire designers of tangible interfaces for explainable AI:

- Tangible interaction could help understanding the structure and functioning of deep learning algorithms, facilitating group discussion and giving multiple controls for shared interaction.
- Most representations of CNNs layers were done using paper or cardboard sheets. This may be due to the fact that the workshop focused on the example of image recognition. Nevertheless, participants found interacting with paper sheets intuitive and easy.
- Tangible interaction encourages exploration, which could help to understand the functioning of deep learning algorithms

4. Organizers

Leonardo Angelini is Assistant Professor at the School of Management of Fribourg (HES-SO) and affiliated to the HumanTech Institute (HES-SO) as postdoctoral researcher. Leonardo has a PhD in Computer Science and his thesis investigated gesture interaction with tangible objects in the context of smart environments. Leonardo has several years of experience in the field of Human-Computer Interaction and Artificial Intelligence, with a particular focus on application for seniors' well-being, eHealth and automotive interfaces.

Nadine Couture is Professor in Computer Science at ESTIA Institute of Technology in Biarritz (France). Her research is conducted at ESTIA-Recherche, which she leads. Nadine is interested in Tangible Interaction, from the physical embodiment of data to interaction with the whole body, applied to Affective Computing and Mixed Reality. She develops research-enterprise links and, as such, she co-founded the PEPSS platform about Human Factors for Interactive Technology, she is involved in the Aerospace-Valley Cluster on data economy and artificial intelligence, and vice-president of the Aquitaine Centre for Information Technologies and Electronics.

Elena Mugellini is Professor at the University of Applied Sciences of Western Switzerland in Fribourg (HES-SO). She holds a Diploma (Bsc and Msc) in Telecommunications Engineering and a Ph.D. in Computer Sciences from University of Florence, Italy. Elena is the leader of HumanTech Institute (Technology for Human well-being, humantech.eia-fr.ch/). The institute aims at improving the quality of life and well-being of human beings thanks to the ingenious use of new technologies, in order to strengthen their abilities as individuals, as well as members of an increasingly dynamic, nomadic and globalized society. Her research expertise lies in Human Computer Interaction (tangible, multimodal and natural interaction, conversational interface and empathic interaction) and Intelligent Data Analysis (machine learning, artificial intelligence, human analytics).

Mira El Kamali is a phD candidate at the University of Fribourg. She is also a scientific collaborator at the University of Applied Sciences of Western Switzerland in Fribourg (HES-SO). Her main research interests are in human computer interaction, multimodal systems, multimodal conversational agents. She worked in a European project “Nestore” with 15 partners where she designed a virtual coach for older adults’ wellbeing. She is also currently working in an Innosuisse project that aims at building recommender systems for smart homes. She holds a master of Science in Computer Engineering where her master thesis was about activity recognition using machine learning.

Quentin Meteier is a PhD candidate at University of Fribourg. He is also a scientific collaborator at the University of Applied Sciences of Western Switzerland in Fribourg (HES-SO). He is working on conditionally automated driving and focuses his research on evaluating the physiological state of the driver using machine learning techniques.

The organizers have collaborated in the past to organize workshops in the field of Tangible Interaction, and in particular in the field of tangible interaction with the Internet of Things.

References

[1] Ben Shneiderman, Human-Centered AI: Reliable, Safe and Trustworthy, April 13, 2021, ACM IUI’21.

[2] Tegen, Agnes, Paul Davidsson, and Jan A. Persson. "Activity recognition through interactive machine learning in a dynamic sensor setting." Personal and Ubiquitous Computing (2020): 1-14.

[3] Dudley, John J., and Per Ola Kristensson. "A review of user interface design for interactive machine learning." ACM Transactions on Interactive Intelligent Systems (TiiS) 8.2 (2018): 1-37.

[4] Ghajargar, M., Bardzell, J., Smith-Renner, A. M., Höök, K., & Krogh, P. G. (2022, February). Graspable AI: Physical Forms as Explanation Modality for Explainable AI. In Sixteenth International Conference on Tangible, Embedded, and Embodied Interaction (pp. 1-4).

[5] Ghajargar, M., Bardzell, J., Renner, A. S., Krogh, P. G., Höök, K., Cuartielles, D., ... & Wiberg, M. (2021, February). From” explainable ai” to” graspable ai”. In Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction (pp. 1-4).

[6] Ghajargar, M., Bardzell, J., Smith-Renner, A., Krogh, P. G., & Höök, K. (2021). Tangible XAI. ACM Interactions.

[7] Olah, et al., "Feature Visualization", Distill, 2017.

[8] Zeiler, M. D., & Fergus, R. (2014, September). Visualizing and understanding convolutional networks. In European conference on computer vision (pp. 818-833). Springer, Cham.

[9] Angelini, L., Mugellini, E., Abou Khaled, O., & Couture, N. (2018, January). Internet of Tangible Things (IoTT): Challenges and opportunities for tangible interaction with IoT. In Informatics (Vol. 5, No. 1, p. 7). MDPI.

[10] Harley, A. W. (2015, December). An interactive node-link visualization of convolutional neural networks. In International Symposium on Visual Computing (pp. 867-877). Springer, Cham. Available at: https://adamharley.com/nn_vis/

[11] Edward Z. Yang, Convolution Visualizer, Accessible at https://ezyang.github.io/convolution-visualizer/. Last accessed: November 2022

[12] Wang, Z. J., Turko, R., Shaikh, O., Park, H., Das, N., Hohman, F., ... & Chau, D. H. P. (2020). CNN explainer: learning convolutional neural networks with interactive visualization. IEEE Transactions on Visualization and Computer Graphics, 27(2), 1396-1406. Available at https://poloclub.github.io/cnn-explainer/

[13] Cloudera Fast Forward Labs, ConvNet Playground, Available at https://convnetplayground.fastforwardlabs.com/. Last accessed: November 2022.

[14] Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., ... & Horvitz, E. (2019, May). Guidelines for human-AI interaction. In Proceedings of the 2019 chi conference on human factors in computing systems (pp. 1-13).