

June 2024

Computer Vision News & Medical Imaging News

The Magazine of the Algorithm Community



ICLR

THE INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS

Vienne, Austria | May 7-11, 2024



ICLR

ICLR

ICLR



**SPECIAL
ICLR**

**Special ICLR Edition with full reviews
of award-winning papers! Pages 2 to 33**

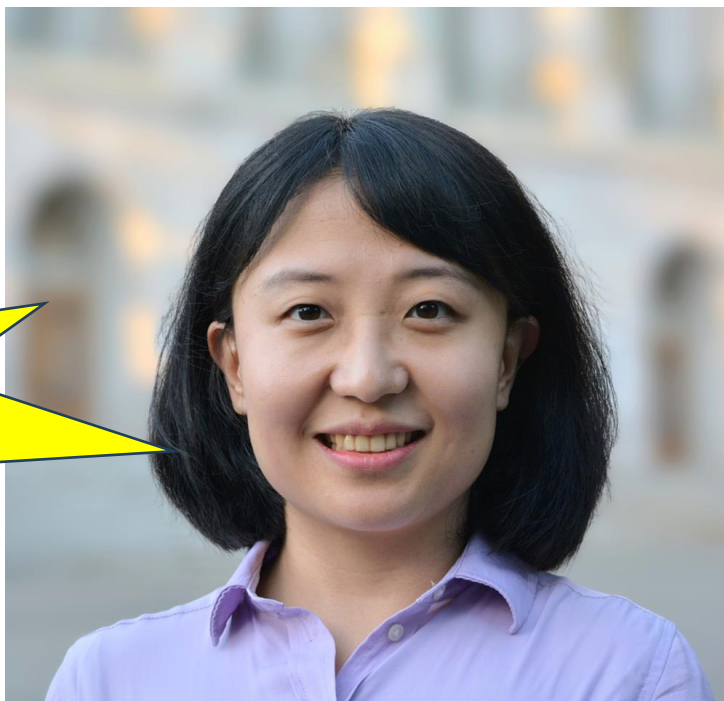
A publication by



Learning Interactive Real-World Simulators



**ICLR
Outstanding
Paper Winner!**



Sherry Yang is a Research Scientist at Google DeepMind and recently graduated from UC Berkeley. She talks to us about her work on learning a realistic interactive simulator through video generation, which earned a super selective oral presentation slot at ICLR 2024 and scooped a coveted Outstanding Paper Award.



A huge amount of video data is available on the internet showing humans performing activities, from cooking to assembling furniture. Typically, this content is consumed passively, but Sherry proposes to **train a model that can absorb these videos and allow users to control the action with language instructions.**

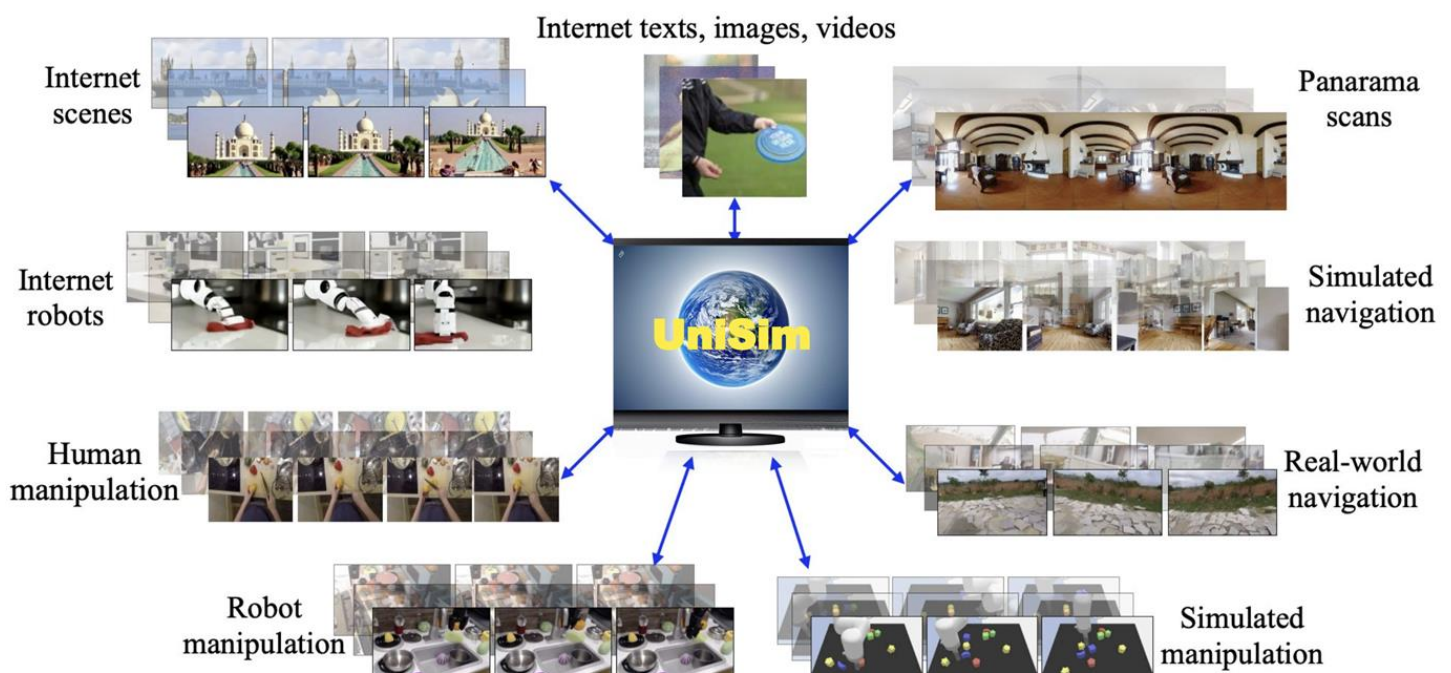
“Starting from a particular frame – for example, me facing my laptop with my hands in the air – I can give an instruction saying, ‘Touch the screen of the laptop,’ and then generate the video of my hand reaching out and touching the screen of the laptop,” she explains. “This is simulated because a user is not

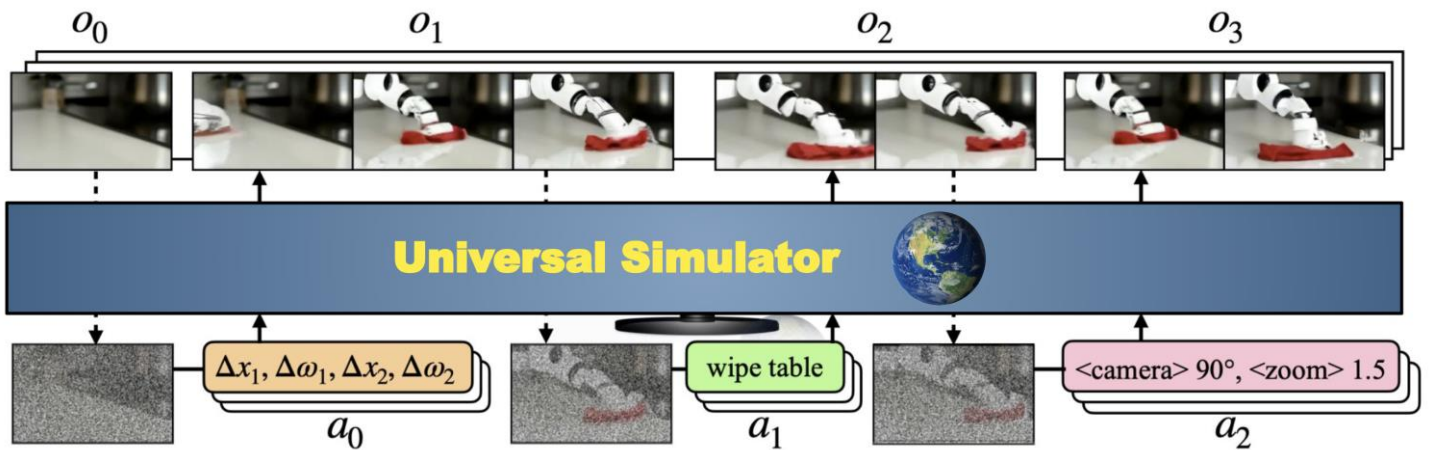
physically interacting with the laptop screen; it's generated in the video by the user giving this language instruction."

There are many applications for such a realistic simulator. While game or film production immediately comes to mind, **augmented reality** is another possibility, where users can interact with an imagined world by issuing commands. **This research primarily focuses on embodied AI, leveraging these simulated experiences to control robots.** "We start with an image of a robot facing a table with some objects on the table, and then give a language instruction to the simulator conditioned on this first frame," Sherry describes. "We say something like 'Grasp the banana' or 'Open the drawer and put the fruits in the drawer.' The simulator will then generate a video of the robot

executing this action – moving its arm closer to the object, picking it up, and putting it into the drawer."

The model can simulate this because it has been trained on videos of other robots or humans performing tasks, enabling it **to interpolate and predict actions based on what it has learned.** Translating the generated videos into real-world robot actions involves training an **inverse dynamics model.** That model takes the video of the robot executing the task and predicts the low-level control actions between two frames, such as joint movements, required to perform it. "After we have this inverse dynamics model, we convert the generated video into low-level robot controls and execute these on the actual robot," Sherry reveals. "In the paper, we've shown situations where the simulated execution looks similar to the real execution of the robot."





Another significant application of this technology is in **reinforcement learning**, drawing parallels with the success of **AlphaGo**, Google DeepMind's program that mastered the game of Go. AlphaGo defeated professional player Lee Sedol in 2016 by learning optimal strategies through self-play simulations, ultimately creating a player superior to humans. It has a perfect simulator, with the game's rules hardcoded for training and a **reinforcement learning algorithm that encourages winning behaviors and discourages losing ones**.

However, replicating this success in real-world scenarios is challenging due to a lack of perfect simulators. *"There's no simulator for the real world,"* Sherry tells us. *"There's no way we can have an agent or computer automatically interacting with a simulator and learning from making mistakes because the simulators people use are often some kind of toy simulator that looks very different from the real world. People could directly have*

real-world interactions and learn from making real mistakes, but it's expensive and unsafe. We can't just have robots breaking things to learn!"

Here, the paper's idea of **UniSim**, using diverse real-world data to create **realistic simulations**, comes in. It can visualize the effects of executing various language instructions, even unsafe ones, **without real-world risk**. By combining this advanced simulation with reinforcement learning algorithms, the model can **train agents to achieve superhuman performance** in a range of tasks, not just games like Go.

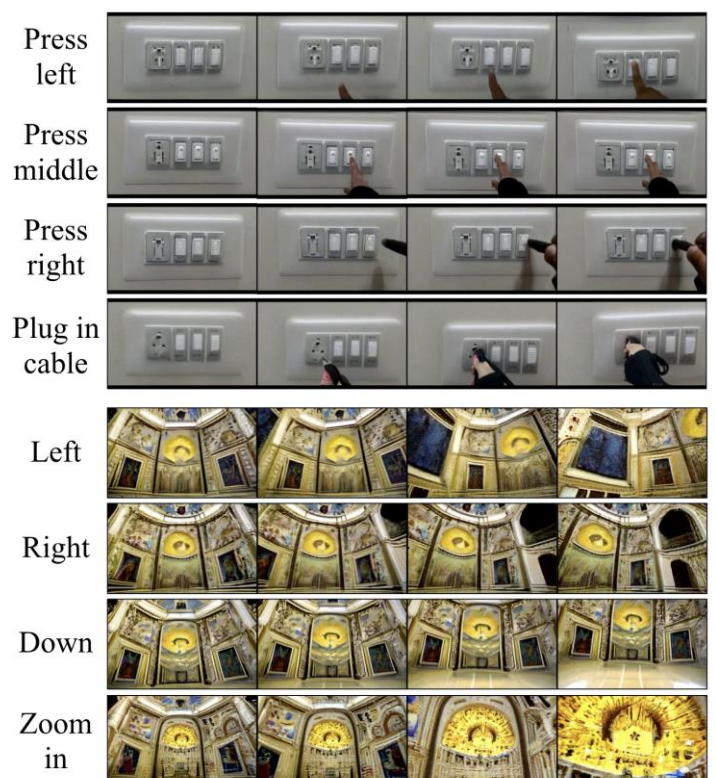
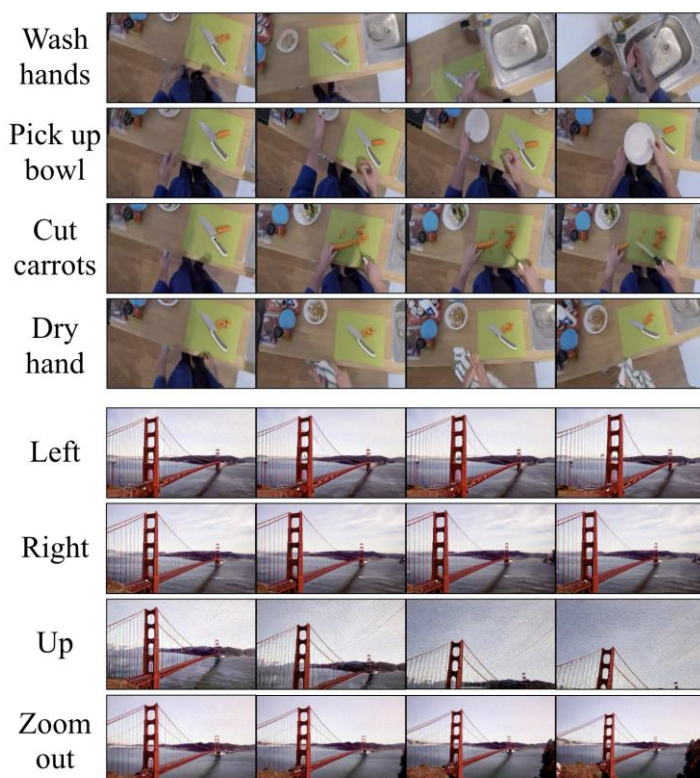
Data curation was one of the project's biggest challenges. Most of the videos on platforms like YouTube have speech transcripts but no detailed action annotations. Robotics datasets inherently document low-level control actions but use formats like Δx , Δy , or endpoint movements and forces. *"We have to convert those*

continuous values into a text description through tokenization,” Sherry explains. “Having a unified language to describe actions is difficult. Once we have that, it’s just about combining datasets and merging information in a single model.”

Another technical challenge was ensuring **the model maintained consistency and memory over time**. For example, if the model places an apple in a drawer and needs to open it later, it must remember the initial action. To address this, it uses history conditioning, aggregating past frames and conditioning on that to generate future video segments. However, the model is currently limited to a fixed number of past frames. *“If something happened days ago, how can I ensure the*

*model remembers that?” she ponders. “This is not addressed in this work, but there are other works in **Google Gemini** or long-context learning where **people can fit millions of tokens into these large models**. These will be considered for future work to empower generative simulators to incorporate history.”*

The field of generative models is motivated by **computer vision**. Classical tasks like segmentation, tracking, and depth estimation play a role in this work, but it connects them to embodied AI, taking a broader, end-to-end approach to simulating the effects of executing actions. Rather than those intermediate tasks, it focuses on image-to-video prediction with some control in the middle, blending **robotics, computer vision, control, and reinforcement learning**.

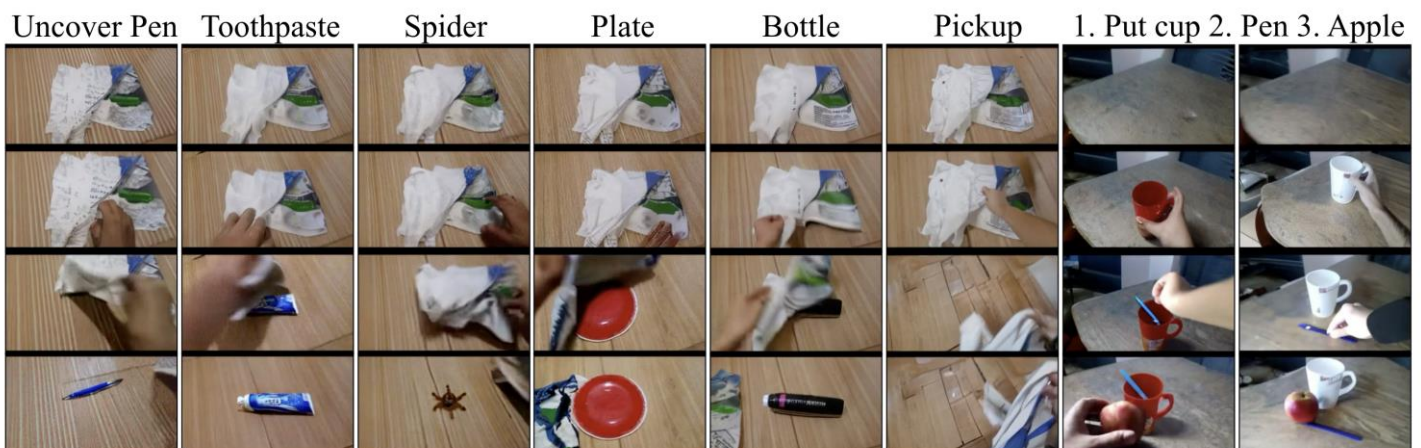


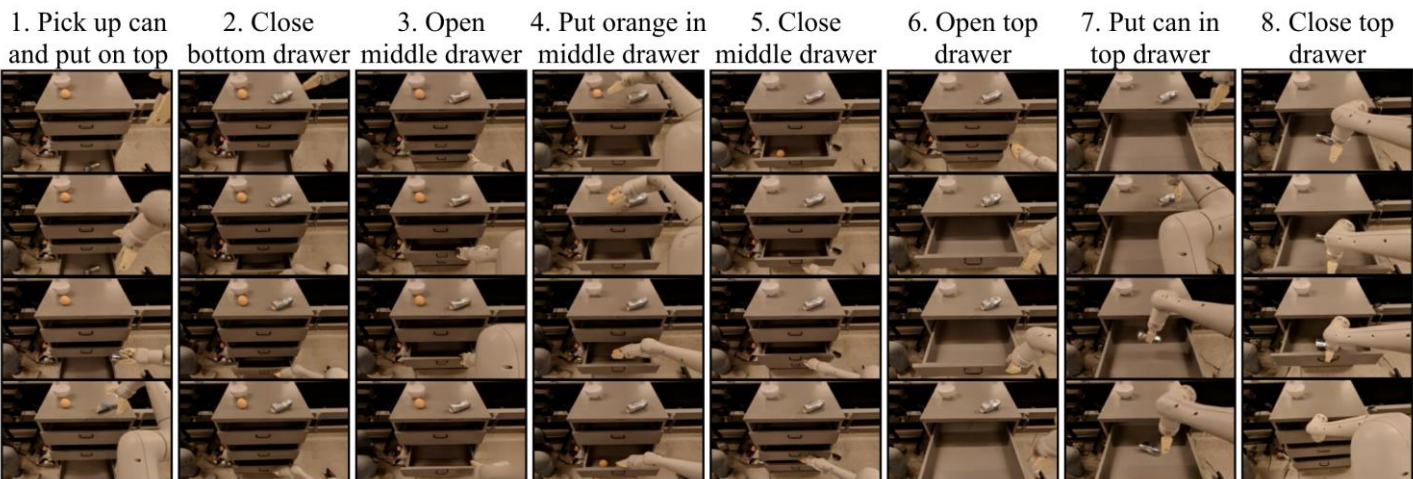
Sherry says the text-to-video generation community has been focused on entertainment applications, favoring creating videos of cute animals in unusual situations over real-world scenarios. This trend followed text-to-image generation, where the focus was on generating pictures that did not exist. *"I'm not saying people went down the wrong route,"* she adds. *"I'm saying we were being narrow by only thinking about creative media. When we think about videos, one of the most interesting applications is modeling the real world because physics is hard. Interaction with soft objects, fluid dynamics, and cloud movement is hard to model using mathematical equations. **Learning a generative model of videos using a data-driven approach, with millions of video clips to learn this dynamics model, is a more natural approach with these large models.**"*

Reflecting on why her work was chosen as an **Outstanding Paper**, Sherry points to the significance of treating video generation as a

simulator of the real world. This shift in perspective is significant, opening new avenues for generating robot and human videos. *"This is the novelty of the idea,"* she reiterates. *"What does it mean if we have **a perfect simulator**? The work demonstrates a few examples. You can use it to train agents. You can use it to generate additional experiences to train video-captioning models. There's great potential for what people can use it for in the future."*

In February, an **OpenAI blog introduced Sora**, its text-to-video AI model capable of generating realistic videos from text instructions. Mirroring Sherry's research, which had been completed months before, it discussed the concept of video generation models as world simulators. *"We put out this idea from a scientific perspective much earlier,"* she points out. *"People from academia and industry are thinking about applications of video generation along similar lines at different times, but I don't see us*





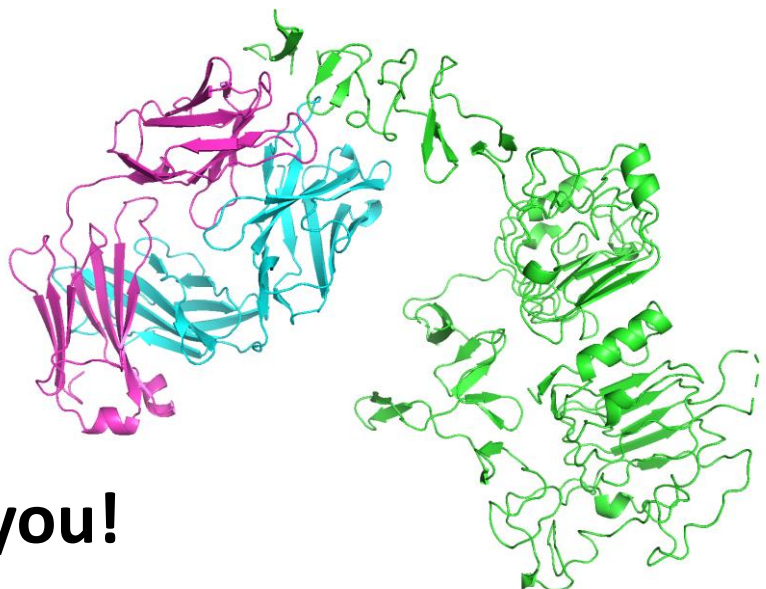
competing with Sora because the focus and scalability are different. I see Sora as reinforcing the idea that generative simulators are simulators of reality, further justifying this approach."

As generative models evolve, Sherry envisions a future where **many internet videos are generated rather than real**, similar to how large language models already create much online content. While concerns around authenticity persist, she thinks advancements in **watermarking technologies** may offer a solution, distinguishing

between real and generated content.

When she is not writing award-winning papers, Sherry is immersed in cutting-edge work at **Google DeepMind**. *"One of our missions is to discover the next research breakthroughs that bring innovations, such as this paper,"* she smiles. *"Other things we're exploring are, for example, **how to use generative models to discover new materials**. That's one I'm focusing on. The goal is just to develop groundbreaking research ideas."*

What is this?
Turn the page:
another ICLR
Outstanding
Paper Award
winner awaits for you!



Protein Discovery with Discrete Walk-Jump Sampling

**ICLR
Outstanding
Paper Winner!**

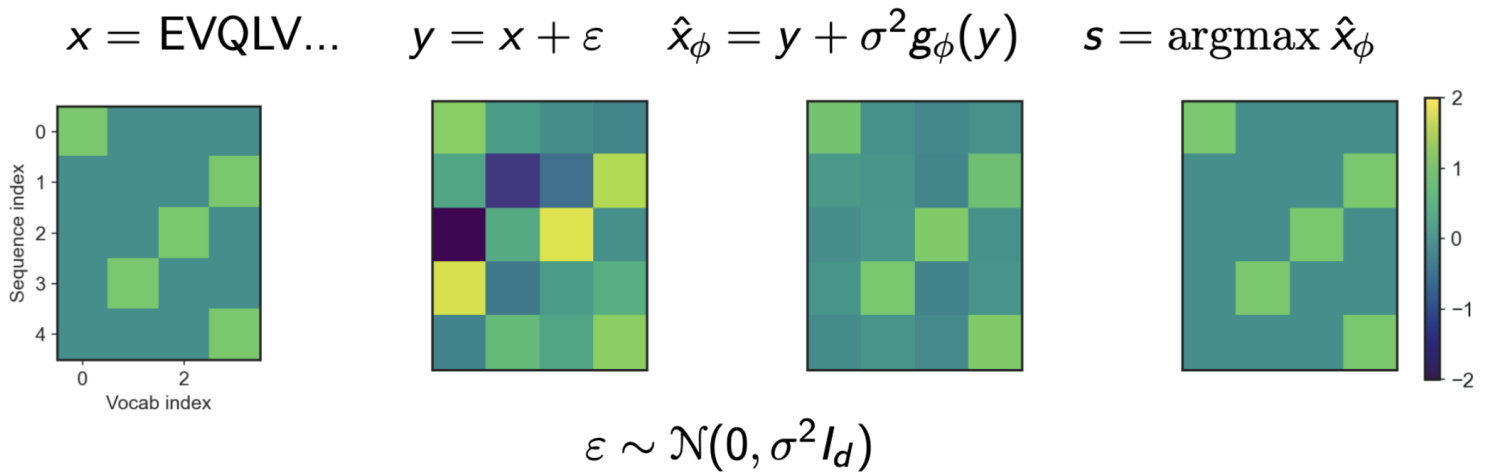
Nathan Frey is a Principal Machine Learning Scientist and Group Leader at Prescient Design, Genentech. He speaks to us fresh from winning an Outstanding Paper Award at ICLR 2024 for his pioneering work on generative modeling for drug discovery.

Nathan Frey and the team at **Prescient Design in Genentech** are pushing the boundaries of drug discovery by applying generative modeling techniques, commonly associated with **image generation**, to create surprising new proteins and discover antibodies and other molecules to cure disease.

“We have the major advantage that everything we design can be tested in the laboratory!”

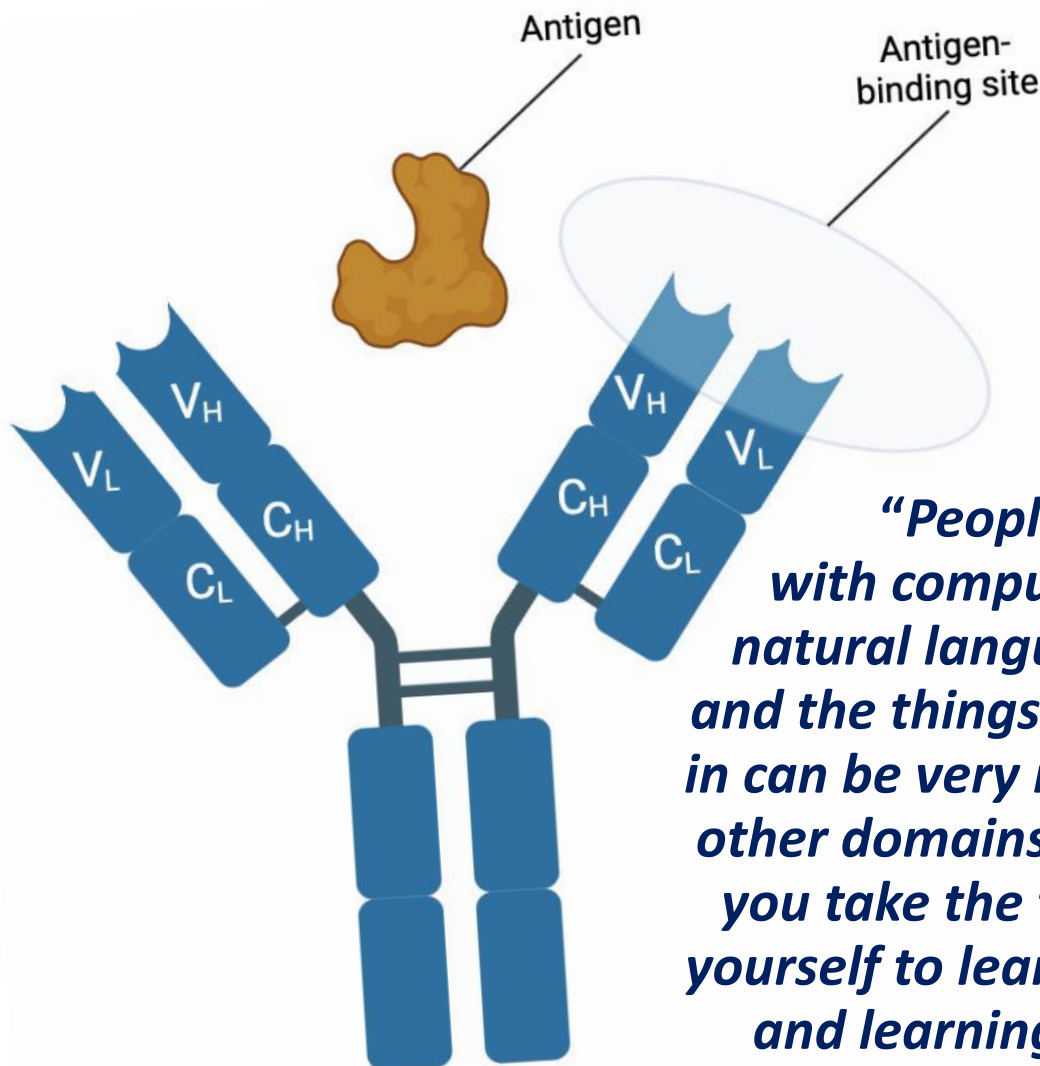
However, as these are complex

molecules critical for various biological functions, **how can the validity of these protein designs be verified?** *“We have the major advantage that everything we design can be tested in the laboratory,”* Nathan explains. *“We work with our experimental colleagues and look at sequences together. They tell us if we’ve done something catastrophically wrong, if there’s something we haven’t thought of or been aware of, or if the human body or immune system would never do this. They teach us those kinds of things, and then we can teach the model and test them.”*



There are many families of generative models and approaches for generative modeling. The motivation behind this project was to resolve the problems previously seen in the protein space when using energy-based and diffusion models.

Protein design is an instance of the **discrete sequence generation problem**, where amino acids are the building blocks of protein, and each amino acid sequence is a string of characters with only 20 possible choices at each position. Precision is crucial in determining the composition



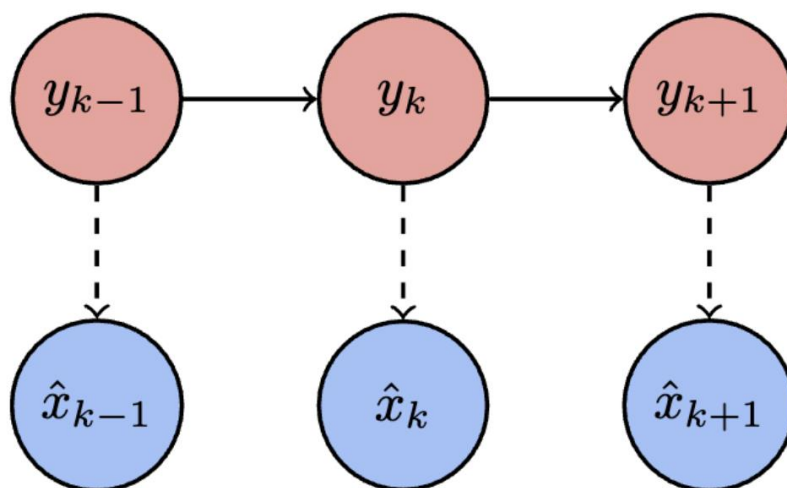
“People familiar with computer vision and natural language, your skills and the things you’re interested in can be very impactful in these other domains in the sciences if you take the time to dedicate yourself to learning from people and learning those areas!”

of each position and modifying existing proteins. “Basically, a lot of generative models are very bad at doing that,” Nathan points out. “Coming up with **a good, robust, sample-efficient method** for generating these discrete sequences was the problem we were interested in.”

The paper represents these sequences as discrete characters that look like language but are one-hot encoded to make an image representation, where the pixels tell you what character is in each position. The team then applies a noising and denoising process. This approach is already used to **create new molecules, aid in sequence diversification, and hit expansion in the real world, which are critical phases in drug development.** “What we’re trying to do is take some starting protein sequences and use the generative model to explore around them to find other real proteins that are reasonable and interesting to look at,” Nathan tells us. “That’s part one of a much bigger process called **lab-in-the-loop for protein design.**”

Still coming down from his **Outstanding Paper Award** win, Nathan says he is excited about the **growing acceptance of machine learning in biology.** “As far as I can tell, this is the first bio ML paper that’s won an outstanding paper award at ICLR – maybe at any major machine learning conference,” he reflects. It is a fantastic achievement with **only five Outstanding Paper winners out of 7,262 submitted and 2,260 accepted research papers.** What does he think convinced the judges that it deserved such an accolade? “I would guess that, if anything, **it’s the real-world impact!** What we did is not in the mainstream of generative modeling right now. There’s an aspect of going against the grain, developing something new, and actually showing that we needed to develop something new for real-world impact.”

This work has been a collaboration between Nathan and his co-authors, **Dan Berenberg** and **Saeed Saremi**, as well as **many other scientists and engineers across Prescient and Genentech.** “One of many things



that brought me to Prescient Design in Genentech was the hope of building things that are actually used in the laboratory to make real things,” Nathan recalls. “I think that’s a sentiment shared by many of our scientists and engineers. For this process to work, you need all the machine learning and biology domain knowledge and great engineering. You really need all of those things together.”

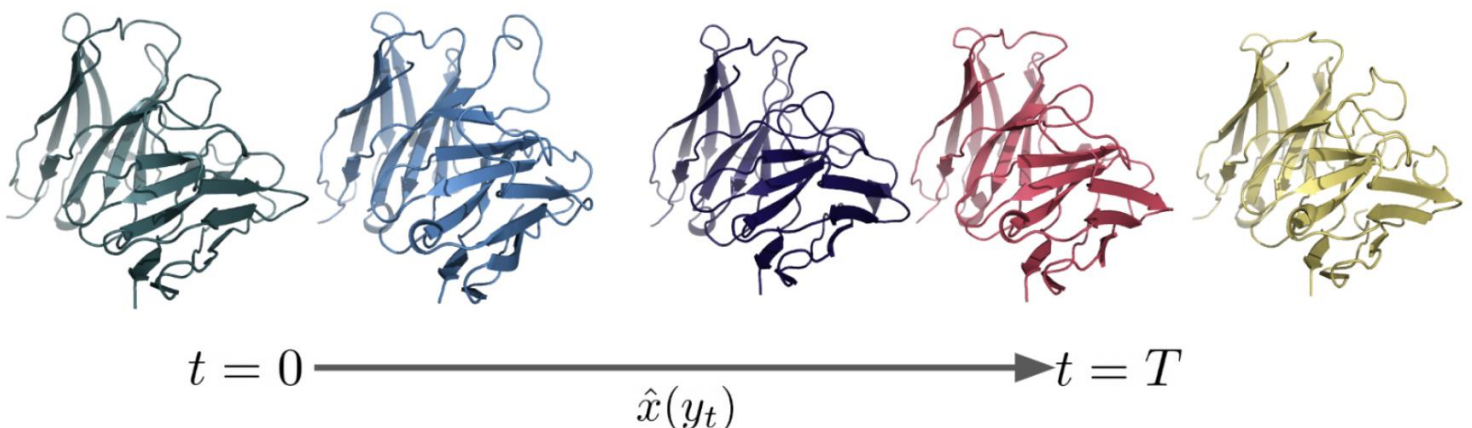
“When we solve real problems and have something interesting to say to the community, that’s when we write something!”

At his lab, Nathan tells us he does not spend much time working on research papers. Instead, he is learning about and contributing to drug discovery. “Writing papers is a byproduct of solving problems,” he points out. “When we solve real problems and have something interesting to say to the community, that’s when we write something.”

Looking ahead, he hints at ongoing collaborations and upcoming

projects within and outside Prescient building on this work, including the **Generative and Experimental Perspectives for Biomolecular Design (GEM) workshop** that just took place at ICLR. “Stay tuned for much, much more work!” he reveals. “People familiar with computer vision and natural language, your skills and the things you’re interested in can be very impactful in these other domains in the sciences if you take the time to dedicate yourself to learning from people and learning those areas.”

If, after reading this, you think you have what it takes to join Genentech, you will be pleased to hear that it is hiring. You, too, could be advancing the frontiers of science very soon! Genentech sees itself as **the first biotechnology company going back many decades** and is now trying to be **the first machine learning-based drug discovery company!** “It’s a reinvention of a company that has had massive success in drug discovery,” Nathan adds. “We want to continue leading on that front over the next many decades.”



ValUES: A Framework for Systematic Validation of Uncertainty Estimation in Semantic Segmentation



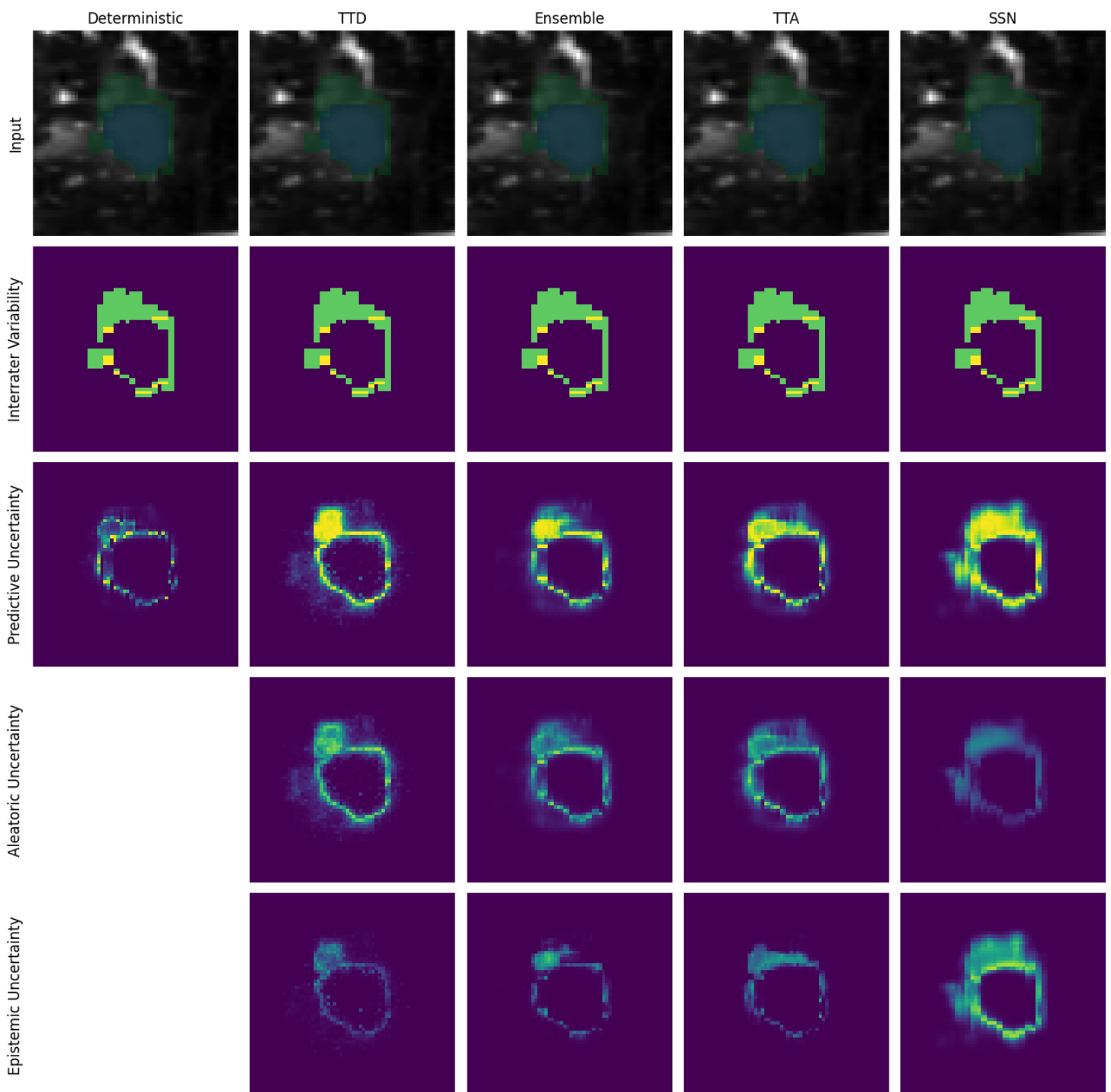
Carsten Lüth and Kim-Celine Kahl are PhD students in the Interactive Machine Learning Group at the German Cancer Research Centre DKFZ under the supervision of Paul Jaeger. Their work on uncertainty estimation in semantic segmentation was selected for a coveted oral presentation at ICLR 2024 - no mean feat when only 1% of submissions made the cut - and they are here to tell us more about its innovative framework and implications for real-world medical applications.

In this paper, Carsten and Kim propose **ValUES**, a framework for systematically analyzing and evaluating uncertainty methods for image segmentation tasks. By defining the essential components

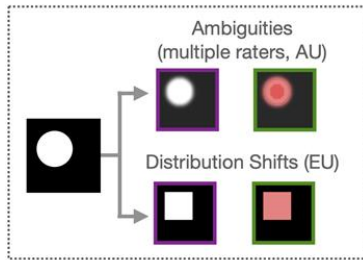
of these theoretical methods and establishing benchmarks for comparison, ValUES empowers practitioners to make informed decisions about using them in their downstream tasks.

In the medical field, certainty is critical, but understanding uncertainty is equally important. *"It's crucial we can detect cases where the model is uncertain,"* Kim tells us. *"If we always assume our model to be 100% certain, it would be impractical for clinical relevance because a doctor would need to check these cases."*

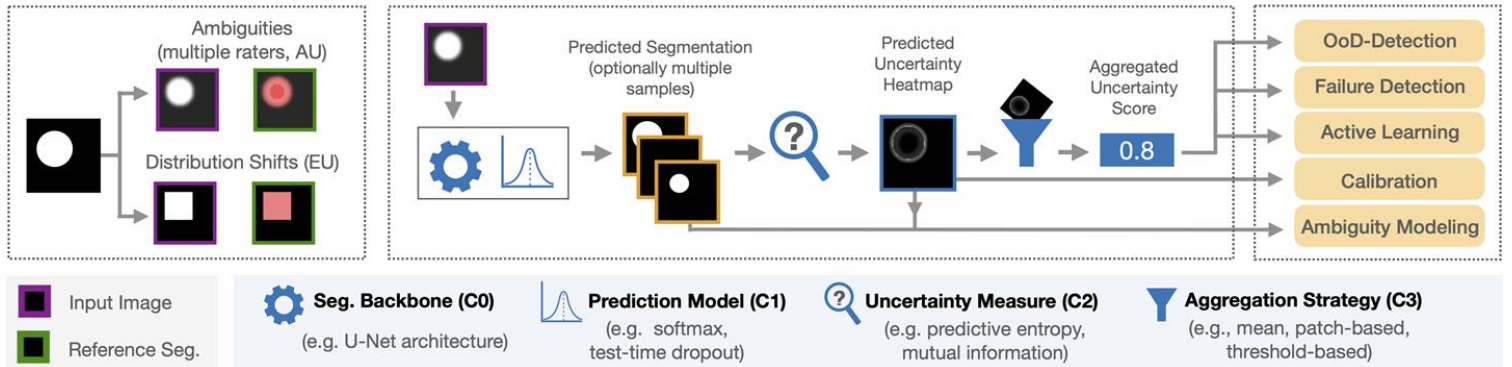
For example, when segmenting lung lesions, it is essential to identify cases where the model fails to detect the tumor outline accurately. **Implementing a failure detection mechanism in the segmentation model enables automatic flagging of potentially inaccurate segmentations** so that they can be reviewed by a human annotator or



Explicit Uncertainty References (R1)



Uncertainty Method (R2)



Downstream Tasks (R3)

- OoD-Detection
- Failure Detection
- Active Learning
- Calibration
- Ambiguity Modeling

physician for correction, ensuring the reliability of the final estimates.

Selecting datasets with specific properties conducive to the study's objectives was a key challenge. Carsten recalls working with a medical dataset focused on lung nodules. *"Before we could even start, we looked at different types of uncertainties because you don't only have one type of uncertainty,"* he explains. *"You have an uncertainty about the border regions, for example, where everyone agrees it has to be diagnosed, but where they would draw the boundaries differs. Or you have cases distinct from one that the model has already seen, and that's a different type of uncertainty. How do you create an environment where you can measure both?"*

Supervisor **Paul Jaeger** says the question of whether the separation of uncertainty types has any real-world effects has bothered him since he started his PhD in 2016: *"I don't remember being so genuinely curious about the outcome of an analysis before. I am happy the community appreciates our insights*

and hope they will help to streamline research in the field."

The team used **U-Net** and **HRNet** architectures as the computer vision backbone. Within this framework, they define a prediction model, employing techniques such as **test-time dropout**, incorporating dropout not only during model training but also at test time for uncertainty estimation, and ensembling multiple U-Nets to generate uncertainty estimates over the ensembles of the predictions.

Delving into more specific uncertainty measures, the team focused on **calculating the uncertainty score at the pixel level**. Previous studies typically addressed image classification tasks, where aggregating uncertainty into a single score per image was unnecessary since the analysis didn't operate at the pixel level. Consequently, these studies lacked pixel-level uncertainty heatmaps and inherently produced uncertainty scores at the image level. *"We use predictive entropy, mutual information, and expected entropy between our multiple predictions,"* Kim explains. *"We have*

Ambiguities for inducing AU

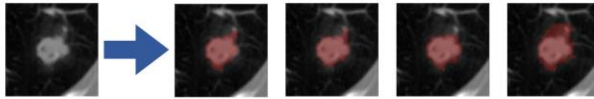
Toy Dataset

Settings: Toy 1 (AU), Toy 2 (EU),
Toy 3a&b (AU&EU)

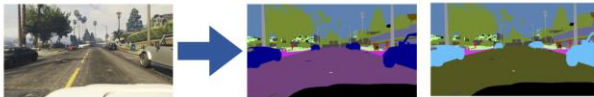


LIDC-IDRI

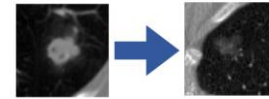
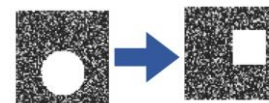
Settings: Shift Solid→Non-solid
(LIDC Tex), Shift Non-
malignant→malignant (LIDC Mal)



GTA5 / Cityscapes



Distribution Shifts for inducing EU



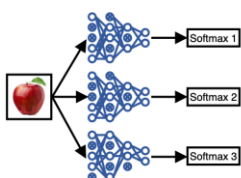
uncertainty heat maps at the pixel level, and then we aggregate them into one score so that we have an uncertainty for the whole image. For failure detection, where we want to have whole images being regarded as a failure, we need to have an automatic system to say, at the image level, this is a failure."

The question arises: **what significance does a glowing pixel of uncertainty hold when determining the uncertainty for a lesion?** Here, the aggregation process becomes vital to facilitate human interaction with the model's output. *"It's a similar thing for active learning, where you select data points to be annotated,"* Carsten points out. *"Just one glowing pixel does not have an inherent meaning. It has to be aggregated. **Aggregation is a crucial part of uncertainty methods**, which we found in our study also has a very large factor for the final performance."*

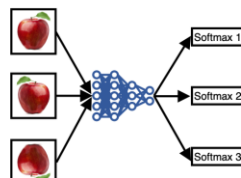
Looking ahead, Kim highlights ValUES's potential to serve as a **foundational framework** rather than an endpoint, opening avenues for the wider community to benchmark and refine new uncertainty methods. She says: *"We argue that our benchmark should be used for newly developed methods to inform practitioners how they can use these methods."*

Carsten agrees and sees the work as bridging the gap between theory and practice. *"There are a lot of great developments in theory that never make it into practical applications because they're very complex, and practitioners are unsure about the benefits for their downstream tasks,"* he tells us. *"Developments can be benchmarked in our setting, and then practitioners can look at the results and make an informed decision to use some newly developed method or not."*

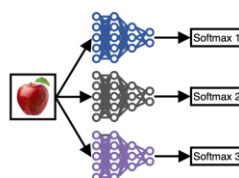
Test-time dropout



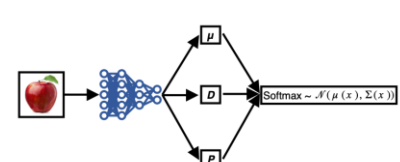
Test-time augmentations



Ensembles



Stochastic Segmentation Networks



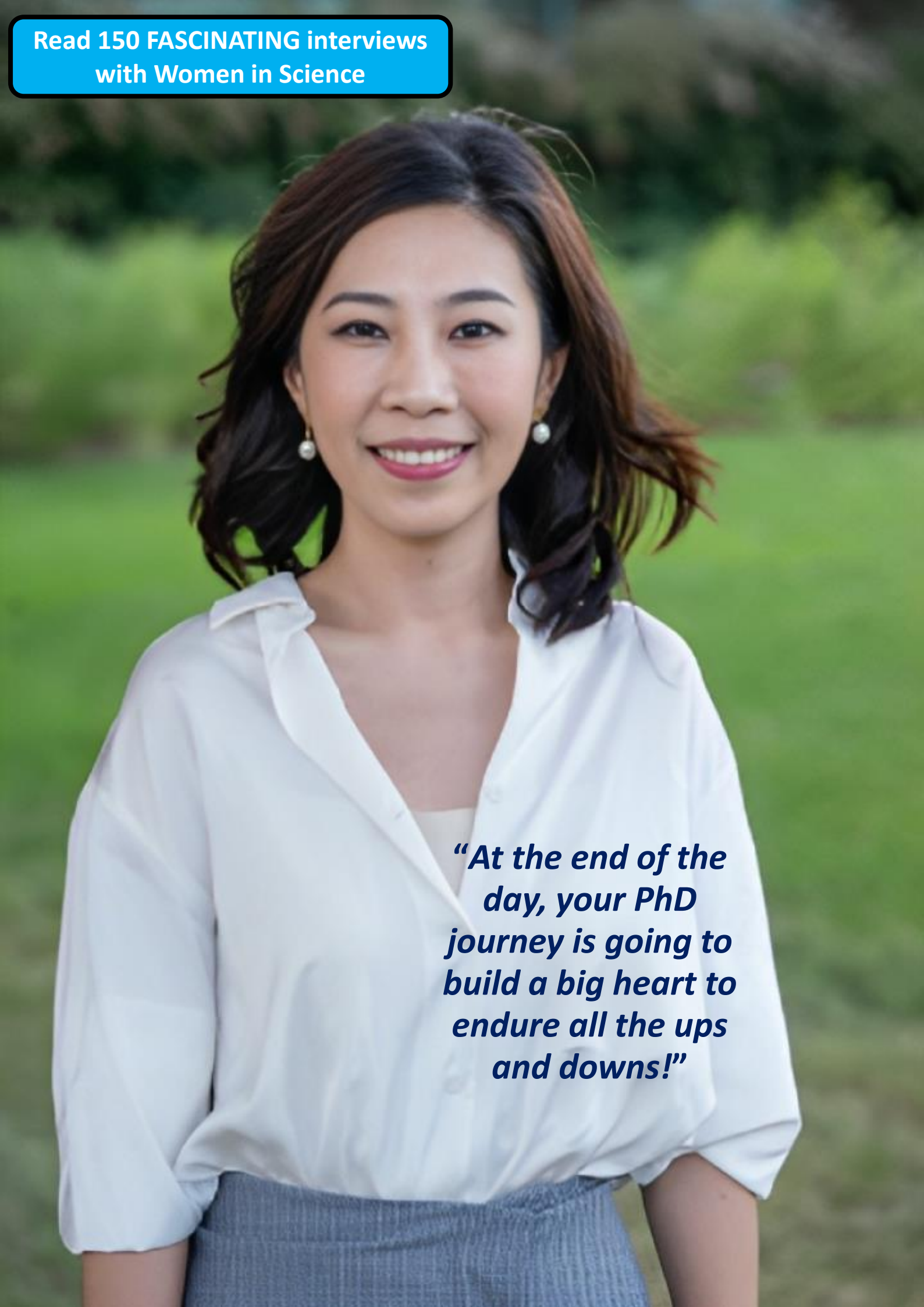


The Ukrainian community connected at ICLR 2024 for a moment of solidarity and friendship. Standing from left to right: Ruslan Partsey, Dmytru Kotsur, Nina Narodytska, Veronika Solopova, Andrii Zadaianchuk, Kate Lytvynets. Thank you to Ruslan for making this happen. I was able to find a spot in the small photo...

For those who don't know (yet), Yann LeCun is also cofounder and chair of the ICLR Foundation. Hence his repeated visits to the conference and - obviously - his presence on the Meta booth. 6 months after our last interview, I had the chance of a brief exchange, during which Yann shared with me his optimistic outlook about the ICLR community and the progress of Artificial Intelligence in general. Of course, he was exquisite and kind as usual. If you missed that interview, [read it here!](#)



**Read 150 FASCINATING interviews
with Women in Science**



***“At the end of the
day, your PhD
journey is going to
build a big heart to
endure all the ups
and downs!”***

Furong Huang is an Assistant Professor at the University of Maryland, College Park. She was at ICLR 2024 to present two spotlights and eight posters, as well for the mentorship program and to interview with Computer Vision News. She was everywhere!

Furong tells us about her career so far and her work on trustworthy AI and machine learning.

Thank you for accepting my invitation, Furong. Can you tell us what your work is about?

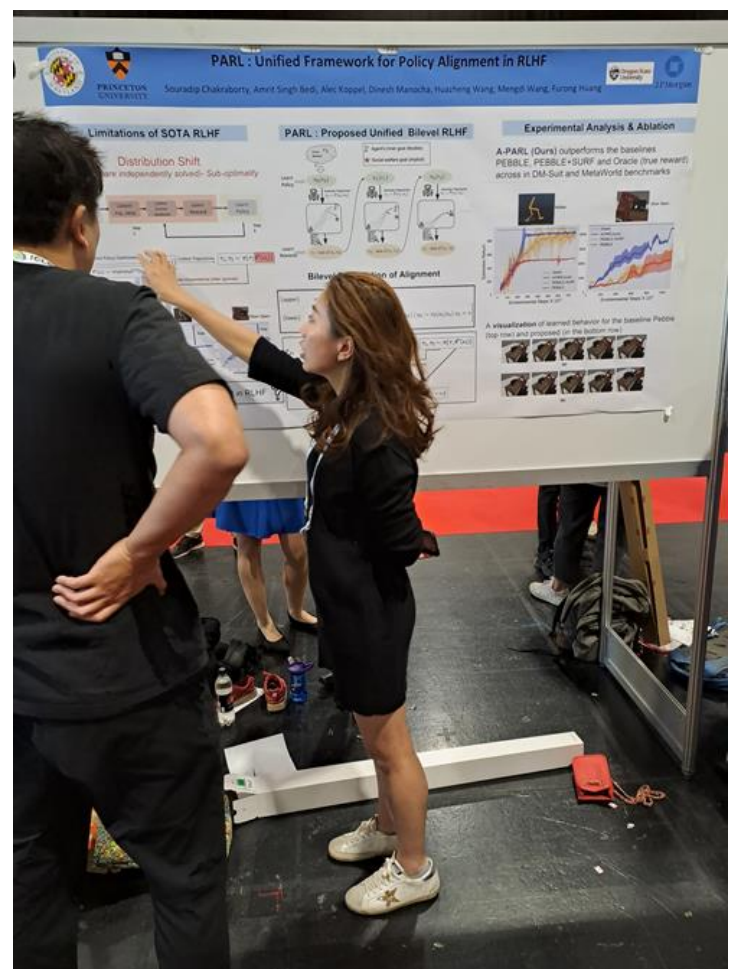
I work on trustworthy AI and ML. Specifically, I want to understand how to align AI with human and social values, the risks and ethical issues in AI deployment, and make sure that AI is always in service of humans in this highly dynamic world that always changes over time.

When we identify things that we should or should not do, how do we enforce ethical protocols and ensure that both the good and bad guys follow them?

[she laughs] That's a great question. Actually, my recent research on AI security is to understand the concept called jailbreak in large language models (LLMs). You might have heard that these LLMs, before they go out for deployment, usually have to go through some security safeguards in the sense that you want to make sure they don't answer illegal or inappropriate questions. Especially

those questions that could harm your business. For example, chatbots should not always say yes when they're asked for a refund if you don't think these are legitimate questions for a refund. But that's more from the business level. There's also the level of security in general in the sense that you shouldn't allow LLMs to give you very detailed guidance on how to hack into a government database, for example. You shouldn't allow those kinds of instructions to happen.

Of course, your question is how to understand the ethical issues or security issues for AI when there are good and bad social actors. This jailbreak problem is for the bad social





For example, AI alignment is a very hot topic that I'm also working on right now to understand how to align these LLMs or even generative AI in general with social and human values. The bad social actor is one thing, but there's also the good social actor problems. When it comes to asking generative AI or LLMs to do things that you need them to do, which are very legitimate requests, you want to make sure these LLMs do not make mistakes or that they actually follow your rules. These LLMs may hallucinate. They may come up with something that doesn't exist, or they may just give you a wrong answer in a very convincing, persuasive way. All these things are concerns we have right now. My research is addressing those issues.

actors that are trying to break the law. They're trying to do malicious things. You want to make sure that LLMs are not their tool. There is an active research area trying to address these questions by first understanding how to do the red teaming of these LLMs to understand where the vulnerabilities are. That's the research we do when it comes to bad social actors trying to take advantage of or exploit LLMs to do bad things.

What direction do you think this is going?

A lot of the research is, like I said, about understanding the bad things that the LLMs can do, and you want to make sure that you patch those.

How do you convince the world that your recommendations are right?

That's my job to make sure my research is being seen by the research community. I'm doing a lot of work making sure that I have my voice heard, for example, on Twitter. *[she laughs]* I have a Twitter account where I try to tell people what my research is about, but also, the research community, in general, is very good at keeping up with literature. If you publish your results, your voice will be heard.

I guess your question is more relevant to, let's say, your research has been accepted by your peers in academia, but what about outside of

academia? What about industry folks? What about government? What about policymakers and so on? Yeah, I don't really have control over it. I don't know for sure outside of the US, but at least in the US, where I work, the government is very open-minded in terms of soliciting ideas and so-called expert opinions from folks in academia and the industry. I've been reached out to by government entities quite a few times in terms of asking my opinions about this kind of rapid growth in AI and what the securities and concerns are. I believe my colleagues have been going through the same. That being said, I think there needs to be some time for policies to be in place and laws in effect, but I'm optimistic that they will eventually catch up.

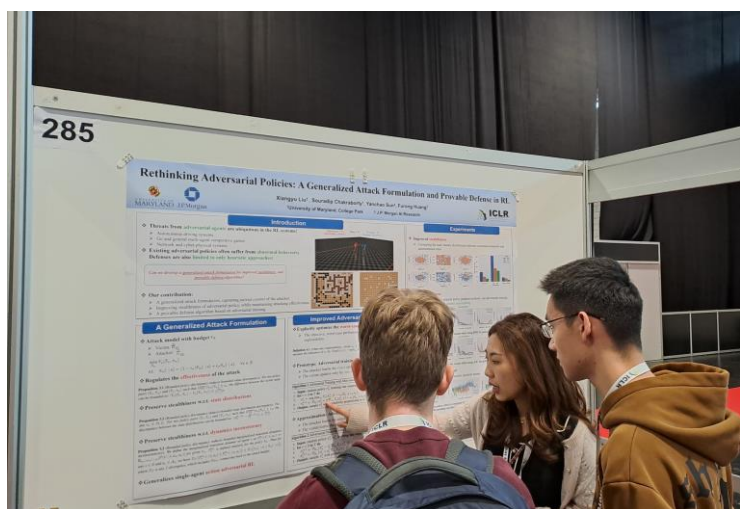
What about the 200 other governments? Are they as open as the US administration?

I don't know a lot in terms of all these countries, but I know that in Europe, the EU is definitely concerned and they're very proactive in terms of making sure they're prepared for this risk. That's a good sign. I know some Asian countries are also very proactive.

One way of your ways to be heard is by having eight presentations here at ICLR! You present, conduct research, and even mentor people. How do you do it all?

[Furong laughs] That's part of my job! I'm very proud to say that I have a very productive research group that I work with. We have two spotlight





presentations and eight poster presentations, all because of their hard work. It's our opportunity to go out and present our research product and make sure that our voices are heard. Unfortunately, many of them have visa issues. They couldn't come in person, so I'm responsible for a lot of the presentations because I could come in person! *[she laughs]* It's a great opportunity for our research to be acknowledged by our colleagues and peers.

Thank you, for finding the time to take this interview during such a busy conference.

It's my pleasure.

Earlier, you told me you felt a bit embarrassed running a mentoring session. You're successful, have a successful team, and have been asked to do it because the community greatly respects you and your work. You were in front of a cheering assembly of young students who would like to be in your place one day. What more would you need to feel at ease?

[she laughs] I would say I'm flattered that I got invited to do this. I'm very happy I could make an influence or maybe help some of the younger researchers to find, hopefully, their career path if they're struggling right now. But to be honest, there are so many great minds and people out there in academia, so I just feel very humble to be able to have this opportunity to just sit there and talk to people.

Do you think that young female students now have the same imposter syndrome that you have had over the years, or do you think that has less power now?

I can't really say on behalf of the female researchers. Yeah, I think you're probably hitting it right with the point that I should be more confident.

What keeps you motivated now that you're getting closer to the middle of your career? Are you still as motivated as you were when you were a PhD student?

That's a great question. What has always kept me going is my passion for how I can make a more tangible impact on the real world. As someone who's passionate about research and passionate about doing something that was not possible previously, I hope I can really make an impact or make a difference. For example, my research is about trustworthy AI. As a human being, I'm concerned that I have to coexist with very intelligent machines. I want to make sure there is harmony. I want

to make sure we can trust the environment we'll have to coexist in. I want to make sure I can contribute to that. I wanted to do my research so that I could help make these machines a little bit more trustworthy. They're going to come. The thing is already out of the cage, whatever it is. I just want to make sure that the AI or the machine learning algorithms we develop are in service of humans, not in competition with humans. We'll make sure that people have security when they're actually living with them. Maybe they don't even realize they're already there. They may be invisible. Like autonomous agents. They don't have to be real robots. They could just be invisible. They could just be a piece of code that can help you do a lot of things in an autonomous way. Like the chatbot is invisible, right? We have to have people doing this kind of research to make sure that we have security.

Were you born in America?

No, I was born in China.

You are spending the significant adult years of your life in America, but what do you feel your identity is?

That's a hard question. I've never really given it very serious thought. I think it's very natural I came to the US to do a graduate program, to do my PhD, and I got a world-class education with great help from my mentors, my colleagues, and my

peer students, but I still feel like I'm not confined by a specific identity. Like, I'm now living in the US, or I was born in China. I feel like the effort to do AI doesn't have boundaries between countries. It's more of an international collaborative effort. So many researchers from all around the world come to Vienna for this great ICLR event. We all have a passion for AI and ML, and we're here, so I don't think the identity of your country is very important.

Did you know America before you arrived there to do your PhD?

My knowledge about the US is more from the movies! *[Furong laughs]* I don't really have any social ties there.

How brave did you need to be across your career to get to where you are today?

When you're young, you're very brave. You don't really think too much! *[she laughs]*

I will not ask you to make the same move in another 20 years! *[she laughs]* Do you have a message for any young scholars reading this?



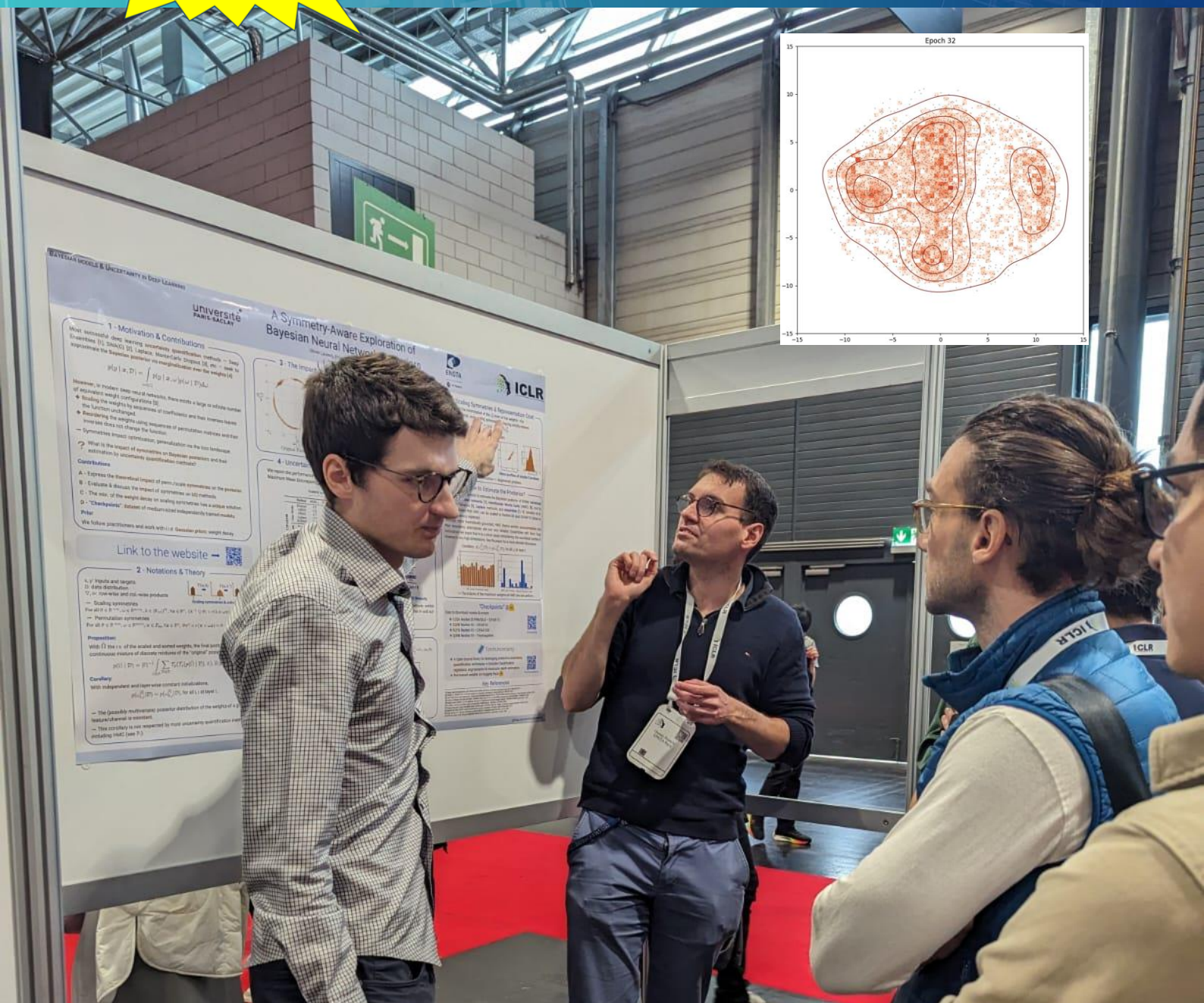
For young researchers, maybe I could just talk to myself, like 14 years ago. When I started this mentoring session, someone was asking me how to be a good PhD student. I don't think I have a very concrete answer about it because a good PhD student can come in many different forms, but probably the most important is your passion and your motivation about what you're going to do. It's going to be a long journey, a very lonely sort of journey. There will be ups and downs for sure. Your papers will likely get rejected. I don't think anybody has a 100% success rate. You will have to go through a lot of rejections, either fellowship applications or competitions, but at the end of the day, I think the most important thing is that you learn during the process. First of all, you

should understand the real problems and challenges out there, but also develop very good communication skills by working with your advisor, working with your peer students, presenting your work, and going out to make elevator pitches.

At the end of the day, your PhD journey is going to build a big heart to endure all the ups and downs, but also to be well connected to build your own community of your research. Who are going to be your mentors? What's your social network? What's your support system? Then, finally, learn to communicate to the world!

**Read 150 FASCINATING interviews
with Women in Computer Vision!**





Olivier Laurent (left) is currently pursuing his PhD at Institut Polytechnique de Paris and Paris-Saclay University under Gianni Franchi's (center) guidance. He presented his paper *A Symmetry-aware Exploration of Bayesian Neural Network Posteriors* at ICLR 2024.

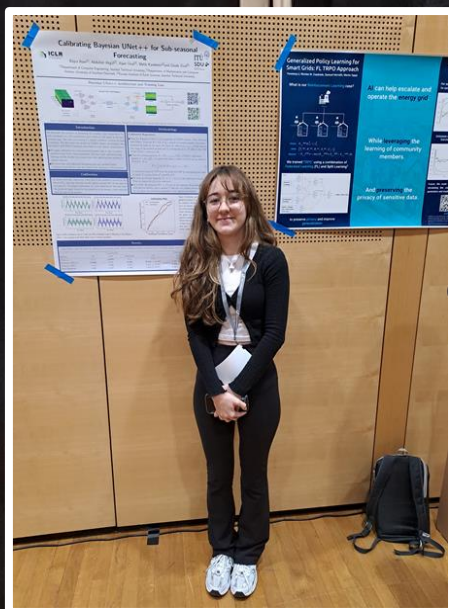
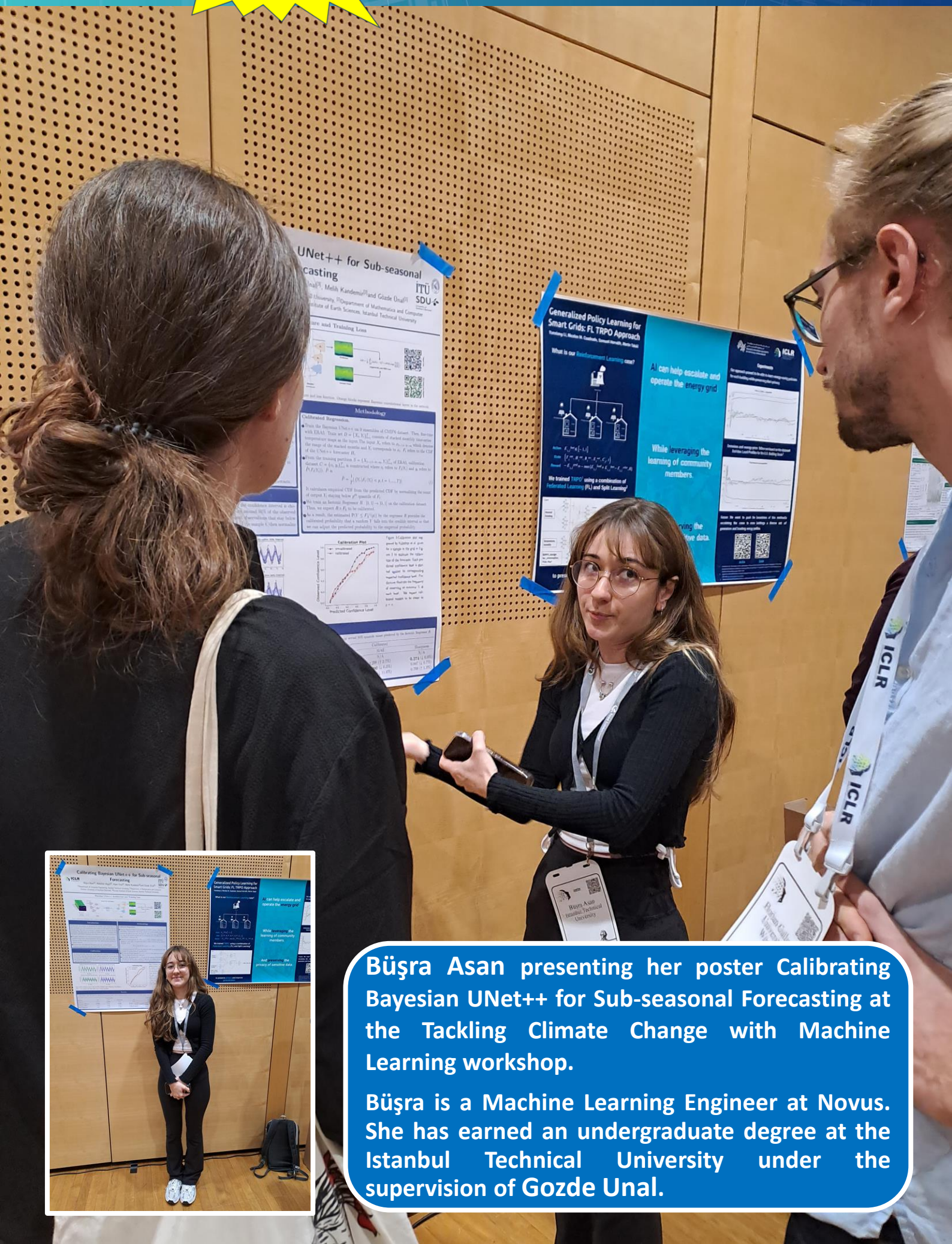
Gianni, an assistant professor at ENSTA, told us that the video above illustrates the progression of the posterior of a basic DNN trained on CIFAR-10, following dimension reduction for clearer visualization. The image depicts the histogram of the joint distribution of a neural network consisting of only 4 neurons. "We specifically plot 2 neurons," said Gianni, "to observe the impact of symmetries."

"This paper holds interest due to its capacity to visualize the behavior of DNN distribution throughout DNN training," he also declared. "It marks the inaugural dataset of DNN checkpointing. Additionally, it enables the acquisition of ground truth regarding the posterior, facilitating the exploration of the relationship between symmetries and Bayesian Neural Networks."



Full house for the traditional Women in Machine Learning meeting. *"The message that most stuck with me from the panel discussion is that your decisions themselves do not define your career path, it is about how you deal with them,"* told us Lisa Weijler, one of the WiML Social @ICLR2024 organizers. *"Rather than focusing on planning every step ahead, your focus is better drawn to exploring what excites you and developing your personal touch on the work you do!"*

Our readers certainly recognize awesome Devi Parikh, second panelist from left. Devi also received enthusiastic reactions to her Invited Talk: Stories from my life.



Büşra Asan presenting her poster Calibrating Bayesian UNet++ for Sub-seasonal Forecasting at the Tackling Climate Change with Machine Learning workshop.

Büşra is a Machine Learning Engineer at Novus. She has earned an undergraduate degree at the Istanbul Technical University under the supervision of Gozde Unal.

CONTRASTIVE POINCARÉ MAPS FOR SINGLE-CELL DATA ANALYSIS

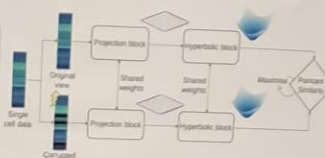
Nithya Bhasker¹, Hattie Chung², Louis Boucherie³, Vladislav Kim¹, Stefanie Speidel¹, and Melanie Weber¹

¹Department of Translational Surgical Oncology, National Center for Tumor Diseases (NCT) Dresden, Dresden, Germany; ²German Cancer Research Center (DKFZ), Heidelberg, Germany; ³Faculty of Medicine and University Hospital Carl Gustav Carus, T20 Dresden University of Technology, Dresden, Germany; ⁴Technical University of Denmark; ⁵Max Planck Institute for Informatics, Saarland University

Contrastive Poincaré Maps: a contrastive learning-based approach for representing single-cell data in hyperbolic space

Motivation

- Single-cell data encode rich hierarchical information from complex biological processes
- Widely used single-cell data analysis methods represent them in a low-dimensional Euclidean space, distorting inherent hierarchies
- Hyperbolic representation methods are known to achieve high accuracy and less distortion for complex hierarchies
- However, existing hyperbolic representation method (Kimovskaya et al., 2020) preserves local structures at the cost of global representation accuracy, and memory for deeper, more complex hierarchies



Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Key features

- Higher global representation accuracy with less distortion
- Memory efficient
- Does not require additional feature processing steps
- Can be efficiently updated as new data becomes available

Method	$Q_{\text{test}}(1)$	$Q_{\text{test}}(1)$	MAP (1)	$D_{\text{test}}(1)$
CPM without PCA	0.46	0.81	0.10	5.83
CPM with PCA	0.43	0.84	0.05	5.98
PM without PCA	0.31	0.74	0.11	12.30
PM with PCA	0.32	0.73	0.13	6.75

Data: Chicken heart development (Martini et al., 2021)

Mantri et al., 2021 - 22,315 cells

Expression in chicken embryos

Individual lineages and relative

expression in the non-cardiac erythrocyte.

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells

Method: PM, CPM

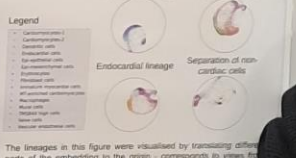
Time (min)

Number of cells

Method: PM, CPM

Time (min)

Number of cells



The lineages in this figure were visualized by translating different parts of the embedding in the origin - corresponds to views from different lineages

Paper link

References

1. Mantri et al., 2021 - 22,315 cells

2. Mantri et al., 2021 - 22,315 cells

3. Mantri et al., 2021 - 22,315 cells

4. Mantri et al., 2021 - 22,315 cells

5. Mantri et al., 2021 - 22,315 cells

6. Mantri et al., 2021 - 22,315 cells

7. Mantri et al., 2021 - 22,315 cells

8. Mantri et al., 2021 - 22,315 cells

9. Mantri et al., 2021 - 22,315 cells

10. Mantri et al., 2021 - 22,315 cells

11. Mantri et al., 2021 - 22,315 cells

12. Mantri et al., 2021 - 22,315 cells

13. Mantri et al., 2021 - 22,315 cells

14. Mantri et al., 2021 - 22,315 cells

15. Mantri et al., 2021 - 22,315 cells

16. Mantri et al., 2021 - 22,315 cells

17. Mantri et al., 2021 - 22,315 cells

18. Mantri et al., 2021 - 22,315 cells

19. Mantri et al., 2021 - 22,315 cells

20. Mantri et al., 2021 - 22,315 cells

21. Mantri et al., 2021 - 22,315 cells

22. Mantri et al., 2021 - 22,315 cells

23. Mantri et al., 2021 - 22,315 cells

24. Mantri et al., 2021 - 22,315 cells

25. Mantri et al., 2021 - 22,315 cells

26. Mantri et al., 2021 - 22,315 cells

27. Mantri et al., 2021 - 22,315 cells

28. Mantri et al., 2021 - 22,315 cells

29. Mantri et al., 2021 - 22,315 cells

30. Mantri et al., 2021 - 22,315 cells

31. Mantri et al., 2021 - 22,315 cells

32. Mantri et al., 2021 - 22,315 cells

33. Mantri et al., 2021 - 22,315 cells

34. Mantri et al., 2021 - 22,315 cells

35. Mantri et al., 2021 - 22,315 cells

36. Mantri et al., 2021 - 22,315 cells

37. Mantri et al., 2021 - 22,315 cells

38. Mantri et al., 2021 - 22,315 cells

39. Mantri et al., 2021 - 22,315 cells

40. Mantri et al., 2021 - 22,315 cells

41. Mantri et al., 2021 - 22,315 cells

42. Mantri et al., 2021 - 22,315 cells

Nithya Bhasker presented her poster at the MLGenX workshop: Contrastive Poincaré Maps for Single-cell Data Analysis.

Nithya is currently pursuing her PhD at the Department of Translational Surgical Oncology NCT/UCC Dresden, with Stefanie Speidel as PI.



Full house also for the fascinating MLGenX Workshop, aiming at bridging the gap between machine learning and functional genomics.

Panel was moderated by awesome Aïcha Bentaieb (standing right) and Alma Andersson (standing left) both from Genentech.

"My main takeaway is how complex this field is and neither biologists, nor computer scientists or machine learning experts can expect to solve the problems on their own," Alma declared. "It requires a joint effort and a willingness to work with people outside of yourself own domain!"





Hello everyone!

I am Christina - I do research in deep learning applied to ophthalmology and soon also remote sensing. I am happy to give a quick summary of the ICLR workshop day, which I attended together with Ralph from RSIP Vision in my lovely hometown, Vienna!

by Christina Bornberg

I mainly attended three of the workshops, since they match my research interest: climate change, remote sensing and representational alignment.

So, let's get started with my top 3 presentations!

I want to start with **"SNAP"** which was presented by **SueYeon Chung**. SNAP stands for **Spectral theory of Neural Prediction and Alignment**. They show that regression-based neural predictivity can be analytically decomposed into a set of contributing factors such as spectral bias and task-model alignment.

Another talk from the representation alignment session was given by **David Lipshutz**. His work focuses on quantifying (dis)similarity of

neural representations. Based on their statement that representations are stochastic and dynamic, they use stochastic shape distances (SSD) as metrics to disentangle noisy dynamic systems with different recurrent interactions.

Finally, I really enjoyed **Emily Shuckburgh's** keynote in the **Tackling Climate Change with Machine Learning** workshop. She was speaking about different tasks that can be helped by decision-support systems, including parameterisations, predictions, classification, systems analysis, evidence synthesis and ethical design.

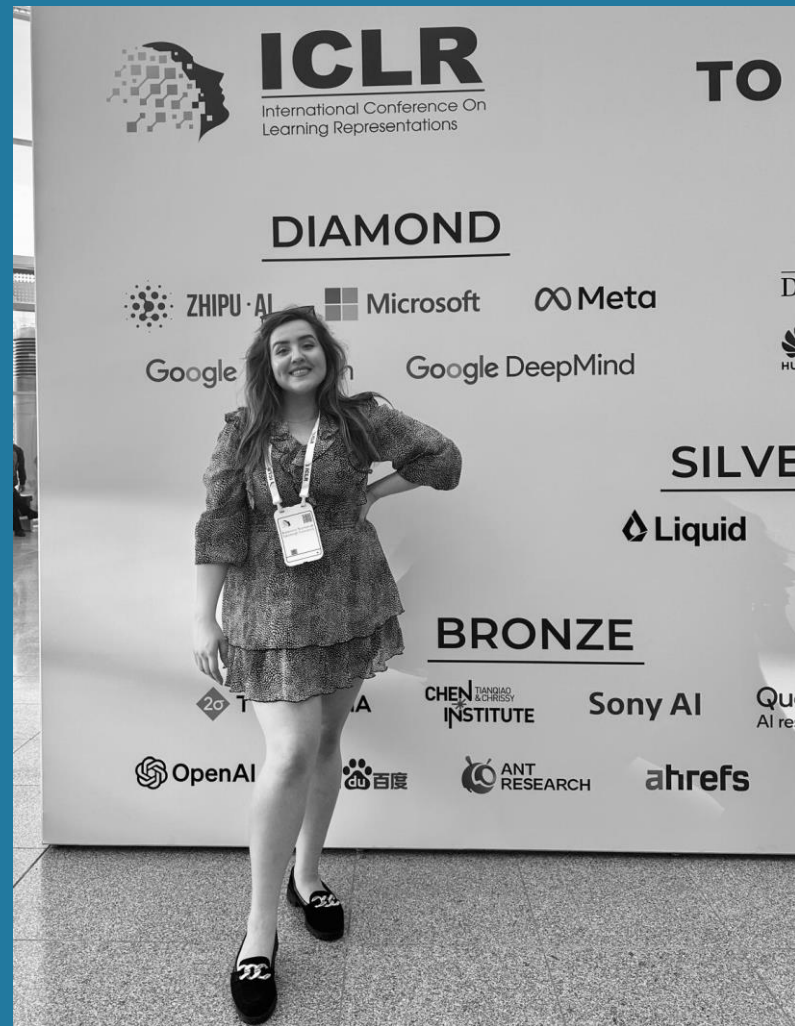
Meet Christina (almost) every month on Computer Vision News with her regular column datascEYence.

Top right: Katarzyna Szymaniak, a PhD student at the University of Edinburgh, UK.

"I'm a researcher dedicated to enhancing human-computer interaction through time-series data analysis," Kasia told us. "With the increasing prevalence of wearables and biosignals, my work focuses on decoding these signals to create more intuitive HCI."

Kasia interned at Meta Reality Labs, where she used advanced deep learning techniques and signal processing to bridge the gap between HCI and practical applications. *"By incorporating inductive biases from ASR architectures, I improved my research in my own field,"* she added. *"In my PhD, I am concentrating on myoelectric control for upper-limb prosthetics. I am particularly interested in using active learning to enhance the long-term adaptation of EMG biosignals, as well as developing compute-efficient methods to address distribution shifts caused by intrinsic and extrinsic factors in the data."*

Bottom right: the undisputed star of ICLR 2024.





ICLR

International Conference On
Learning Representations

DIAMOND



ZHIPU · AI



Micro



Meta

Google Research



DeepMind

SIL



Liqui

ONZE



TWO SIGMA

Sony AI



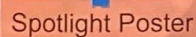
OpenAI

Baidu 百度

T

SEARCH

ahrefs



The second paper **Bounding Box Stability against Feature Dropout Reflects Detector Generalization across Environments**, nominated as a spotlight, explores a brand-new problem of estimate detection accuracy in terms of mAP without test ground truths.





Mathis Petrovich has recently completed his PhD in the IMAGINE team at the École des Ponts ParisTech and in collaboration with the Perceiving Systems department of the Max Planck Institute for Intelligent Systems. He was supervised by [Gül Varol](#) and [Michael J. Black](#) and his research focuses on Natural Language Control for 3D Human Motion Synthesis.

Mathis is going to move into industry, but hasn't decided which one yet.

He's a catch! Congrats, Doctor Mathis!

3D human motions are at the core of many applications such as the film industry, healthcare, augmented reality, virtual reality and video games. However, these applications often rely on **expensive and time-consuming motion capture data**. The aim of Mathis' thesis is to explore **generative models as an alternative route to obtain 3D human motions**. More specifically, the goal is to allow a natural language interface as a means to control the generation process. To this end, Mathis develops a series of models that synthesize realistic and diverse motions following the semantic inputs.

In the first work on his thesis, he addresses the challenge of **generating human motion sequences conditioned on specific action categories**. He introduces **ACTOR**, a conditional variational autoencoder (VAE) that learns an

action-aware latent representation for human motions. He shows significant gains over existing methods thanks to the new Transformer-based VAE formulation, encoding and decoding SMPL pose sequences through a single motion-level embedding.

In his second work, he goes beyond categorical actions, and dives into the task of synthesizing diverse 3D human motions from textual descriptions allowing a larger vocabulary and potentially more fine-grained control. This work stands out from previous research by not deterministically generating a single motion sequence, but by synthesizing multiple, varied sequences from a given text. He proposes **TEMOS**, building on his VAE-based ACTOR architecture, but this time **integrating a pretrained text encoder to handle large-vocabulary natural language inputs**.

A person is walking in a circle clockwise and then sitting down while raising the right hand.

(a) Traditional text input

Walking in a circle clockwise and then sitting down
Raising the right hand

(c) Spatial composition

Walking in a circle clockwise while raising the right hand
Sitting down while raising the right hand

(b) Temporal composition

Walking in a circle clockwise
Sitting down
Raising the right hand

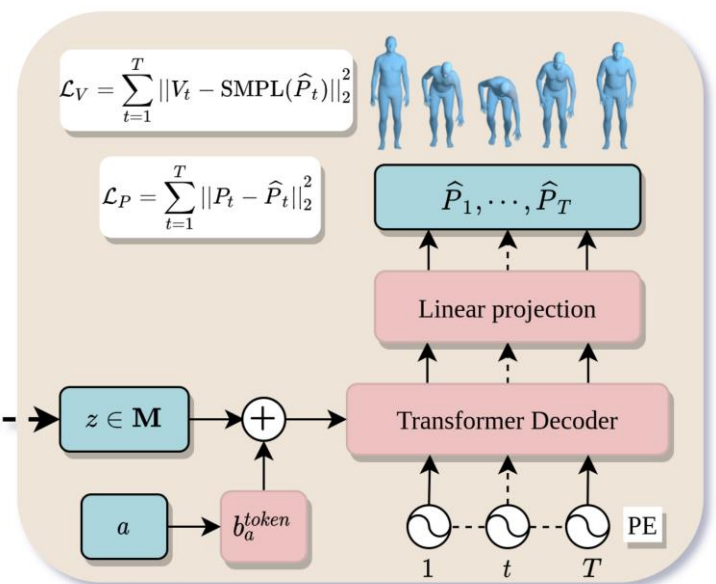
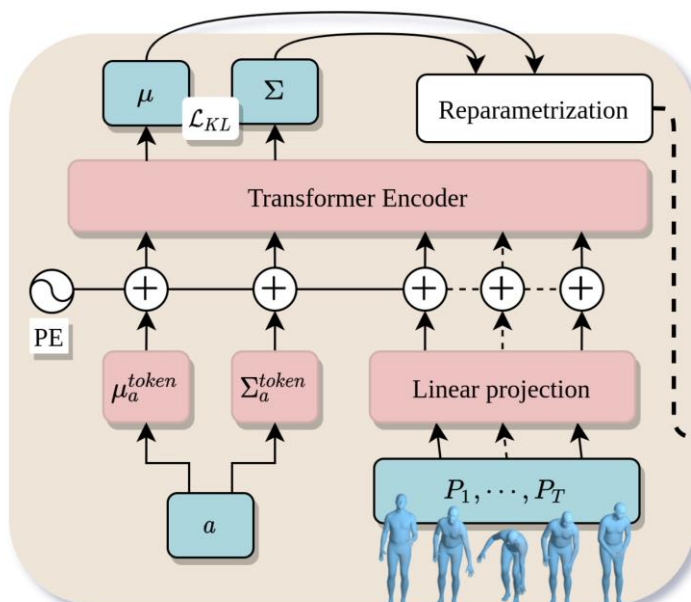
(d) Multi-track timeline control

In his third work, he addresses the adjacent task of **text-to-3D human motion retrieval**, where the goal is **to search in a motion collection by querying via text**. He introduces a simple yet effective approach, named **TMR**, building on his earlier model TEMOS, by **integrating a contrastive loss to enhance the structure of the cross-modal latent space**. His findings emphasize the importance of retaining the motion generation loss in conjunction with contrastive training for improved results. He establishes a new evaluation benchmark and conduct analyses on several protocols.

In his fourth work, he introduces a new problem termed as “**multi-track**

timeline control” for text-driven 3D human motion synthesis. Instead of a single textual prompt, users can organize multiple prompts in temporal intervals that may overlap. He introduces **STMC**, a test-time denoising method that can be integrated with any pre-trained motion diffusion model. His evaluations demonstrate that his method generates motions that closely match the semantic and temporal aspects of the input timelines.

Mathis has also played a major role in other projects: generating human motions with spatial compositions or temporal compositions. For more information, see [his website](#).



NRDF: Neural Riemannian Distance Fields for Learning Articulated Pose Priors



Yannan He is a PhD student with the Real Virtual Humans Group within the Department of Computer Science at the University of Tübingen, supervised by [Gerard Pons-Moll](#).

His recent paper, which has been accepted as a highlight at CVPR, follows on from [second author Garvita Tiwari's award-winning work, Pose-NDF](#).

He will present at the first poster session, on Wednesday June 19, poster 145.

The seeds for this project were sown when Yannan and the team behind **Pose-NDF**, a continuous model for plausible human poses based on **neural distance fields**, began to investigate the failure cases of Pose-NDF. They discovered an underlying issue in its **training data distribution**, which should show decreasing samples as you move away from the pose manifold. In this new paper, NRDF, Yannan aims to set that right.

"We're still modeling it as a distance

field," he explains. *"During inference, it could help if you draw more samples nearby the manifold because, during the projection, you're moving closer and closer to the manifold. If there are more negative samples near the manifold, the network prediction will be more accurate, and it'll help you achieve more stable and accurate projection results."*

The goal is to model distance field-based pose priors in the space of plausible articulations. The learned

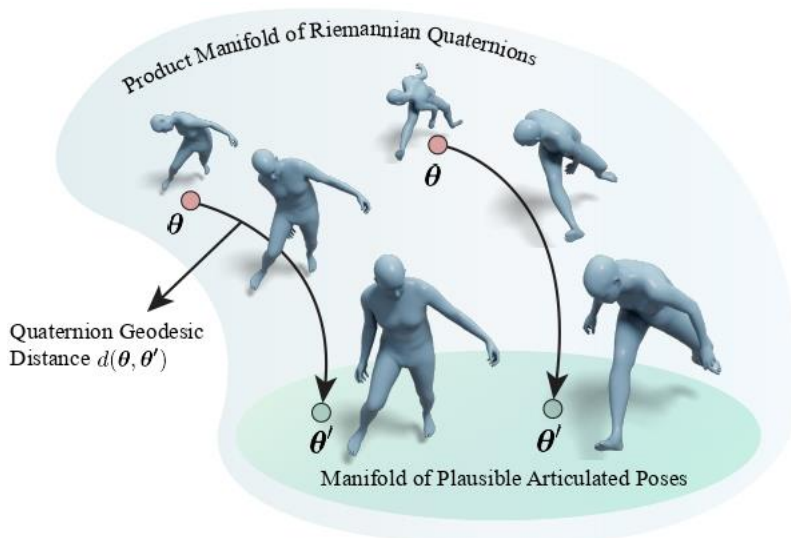
pose priors are versatile and can be applied to various downstream tasks such as pose denoising, 3D pose estimation from single images, and solving inverse kinematics from sparse observations. “Sometimes existing methods return 3D poses where the image overlay is good, but if you view it from another point of view in 3D, it’s implausible,” Yannan points out. “The 3D pose itself may have self-occlusions, interpenetrations, and also some implausible pose patterns, like a knee bending outwards.”

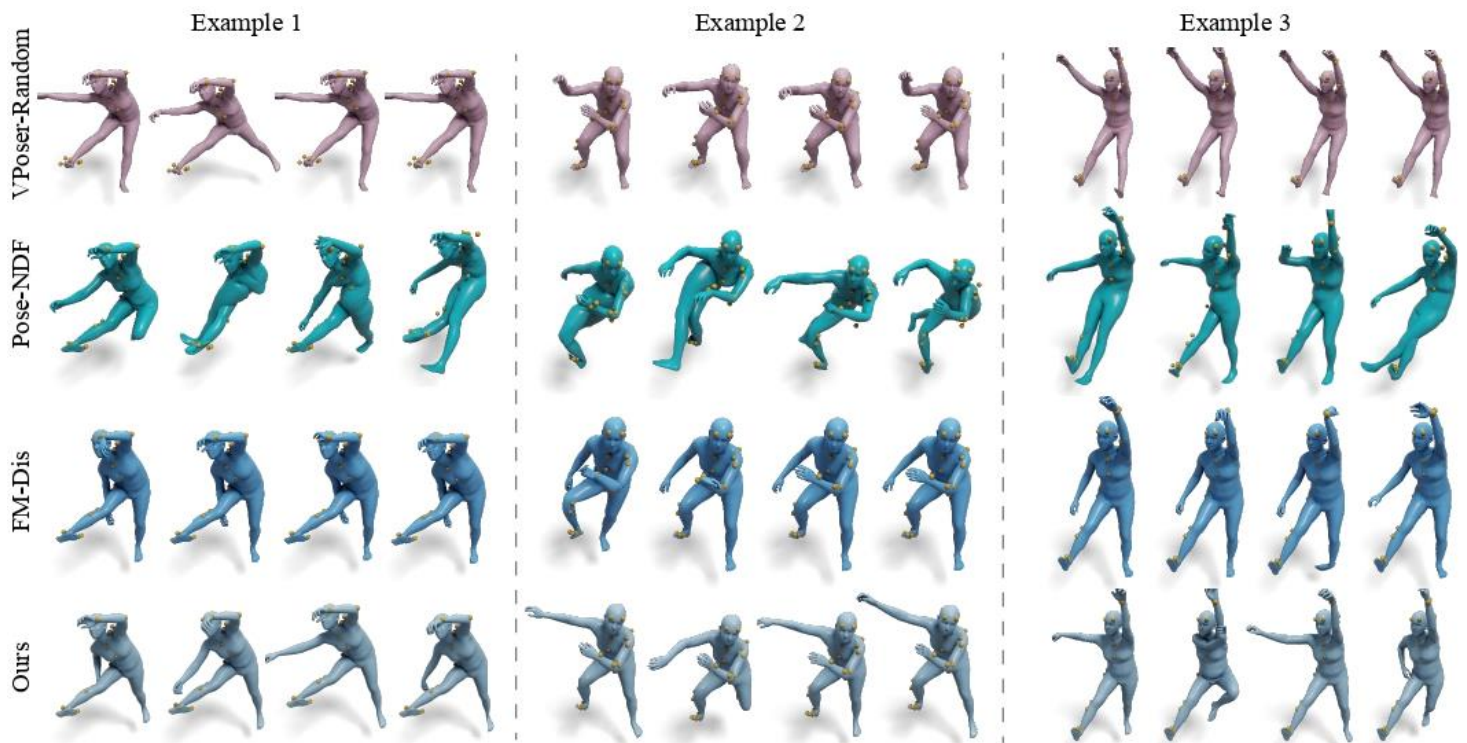
Besides humans, NRDF can be extended to any articulated shapes, such as hand and animal poses. It can return plausible and valid results with only wrists and ankles as surface markers.

Yannan outlines two critical ideas presented in NRDF. Firstly, a **novel sampling approach to address inaccurate distance prediction**, where there are no training samples near the manifold. “During the

training data preparation, we aim to draw more samples near the surface, with a gradual decrease as we move to the faraway regions,” he explains. “In the distribution of Pose-NDF, there is a huge gap between the zero-level-set and the mean, which lies in the center with a big distance value. After we propose the sampling algorithm, we could obtain a distribution shape like a half-Gaussian distribution. Actually, it could be any distribution that the user specifies. It could be an exponential distribution or a uniform distribution.”

Most distance fields work with 3D point clouds, so they are learning 3D shapes in Euclidean space, but this work features **high-dimensional articulated poses represented by K quaternions in the product space of $SO(3)$** . “Sampling points in Euclidean space is totally different from directly sampling rotations,” Yannan tells us. “The special thing about our work is we propose an easy way to



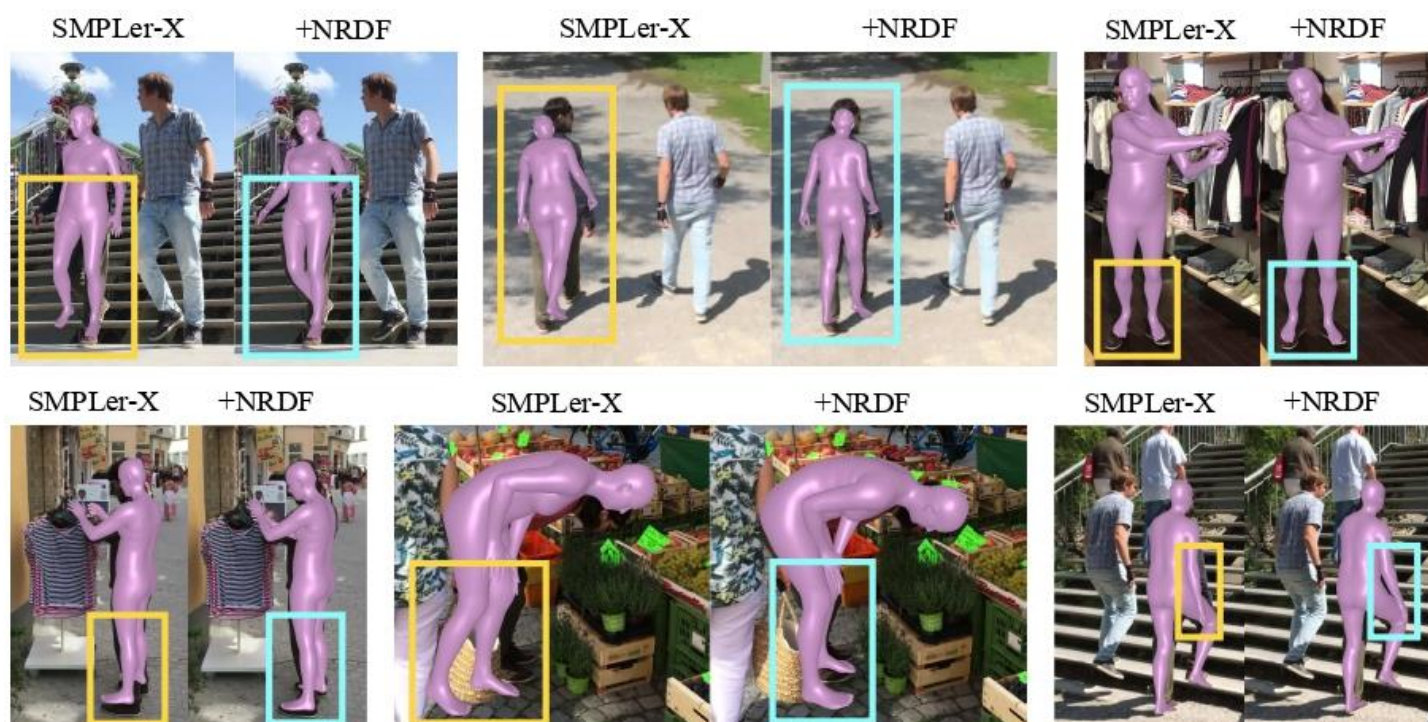


*directly sample articulated rotations in the articulated $SO(3)$ space, which we call **a product manifold of Riemannian quaternions**.*

Additionally, this work introduces **RDFGrad**, an innovative technique that streamlines the gradient descent process during inference. For Pose-NDF, after a normal gradient descent, you have to reproject the resulting pose onto the quaternion space because the quaternion should also be uniform, which slows down the projection process. In contrast, NRDF extends the original gradient descent procedure onto the Riemannian manifold during inference time projection. *“Given a noisy pose, we obtain the gradient direction*

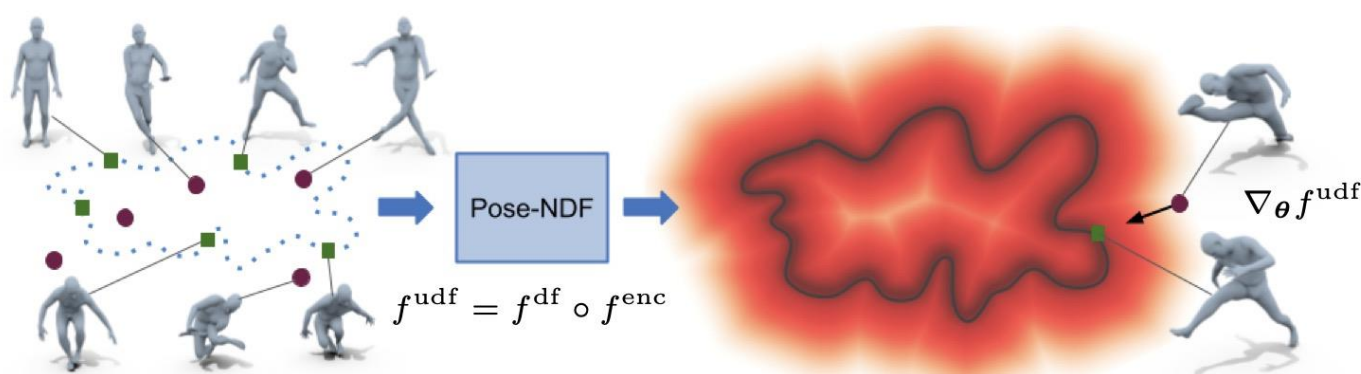
returned by the network propagation, which is the Euclidean gradient, and we iteratively project it onto the tangent space of a given pose and directly work along the geodesic” he explains. *“This is crucial and makes the process faster.”*

Yannan tells us the breakthrough came with the realization that **optimizing the training data distribution is crucial for distance fields**. Finding this, out of all possible explanations for Pose-NDF limitations, was one of the biggest challenges and took several months of intense investigation. **For NRDF, no network architecture was modified compared with Pose-NDF.**



Collaborating with [Gerard Pons-Moll](#) and a talented team of researchers was a rewarding experience for Yannan. “It’s a lot of fun working with Gerard because he’s a really nice person. Even away from research, I can learn a lot from him – he’s a big fan of music, and I also enjoy making music!” he laughs.

“Also, our other collaborators, [Jan Eric Lenssen](#), [Tolga Birdal](#), and [Garvita Tiwari](#). We’re diving deeper into the mathematics behind the problem and exploring the essential reasons behind the phenomenon. It’s really cool, and I learned a lot from them.”



Garvita Tiwari’s work Pose-NDF that won a Best Paper Honorable Mention award at ECCV 2022. [Read our full review.](#)

Are We Building Biases in AI-Driven Fetal Diagnostics? Uncovering Ethical Issues in Public Ultrasound Datasets



Mariachiara Di Cosmo is a PhD candidate at Università Politecnica delle Marche, focusing on the application of deep learning techniques in medical image analysis.

She accepted to sum up the key insights from her contribution at the AI Fairness Cluster Inaugural Conference, presented within the workshop: AI Bias - Measurements, Mitigation, Explanation Strategies.

by Mariachiara Di Cosmo

In the dynamic field of **medical imaging**, AI is gaining prominence in prenatal diagnostics. **Fetal ultrasound (US) is crucial in prenatal care**, providing visualization and monitoring of fetal development within the womb. AI holds promise in expanding US diagnostic capabilities, increasing accuracy and efficiency, while addressing challenges associated with this complex operator-dependent technique.

As evidence of this, a burgeoning body of literature emerged recently also thanks to the availability of benchmark international datasets open to research community.

However, for a researcher in medical imaging analysis and deep learning developer, critical questions arise:

- *Will prenatal diagnostics truly take advantage from AI tools?*
- *Can we ensure AI models benefit all populations and prenatal contexts equally?*
- *Are we designing “fair” diagnostic support systems, considering minorities, unusual fetal anatomies and real clinical US setup?*

Looking for answers and guidelines, we examine public fetal US datasets, used for algorithm development. **The integrity of AI models outcomes heavily depends on training datasets.**

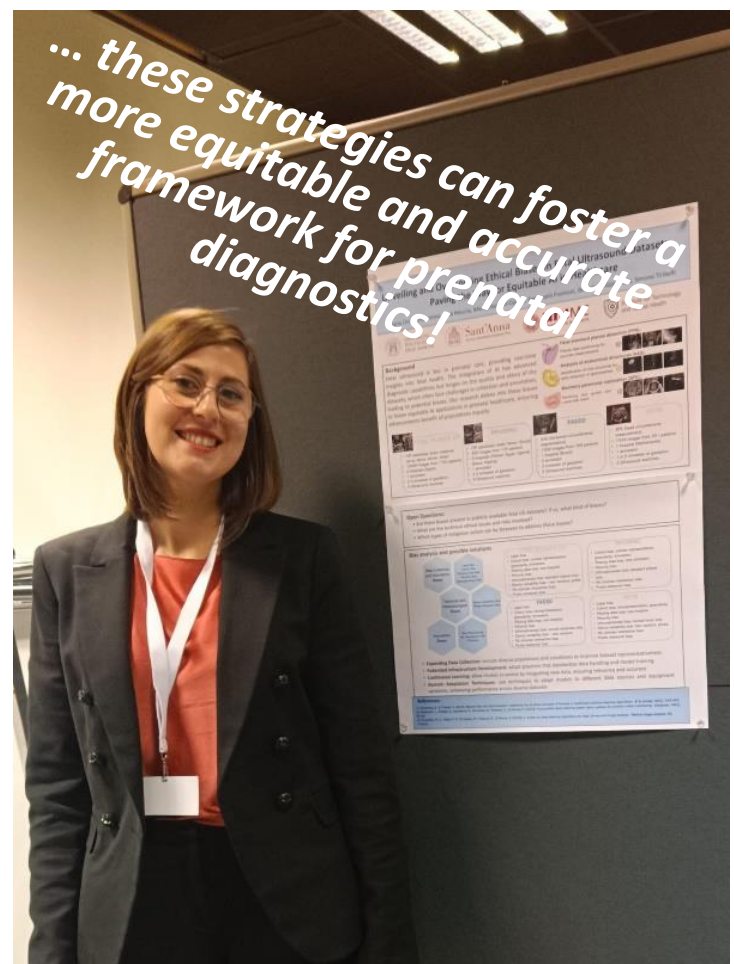
Common biases include a lack of demographic representativeness, leading to models performing inequitably across different populations or hospitals.

Additionally, the **datasets lack diversity in fetal conditions**, skewing algorithm development towards the typical fetus anatomy and neglecting pathological or abnormal fetal growth and multiple gestations. Variability in US technology, both in vending machines and device setup (probe type and working frequency), also leads to discrepancies in image quality and AI outcomes. **Such biases can lead to AI models that perform well within represented cohort, but fail in others, potentially leading to misdiagnoses or inadequate medical advice.**

This underscores the need for careful dataset evaluation and technical strategies to mitigate these issues. Diversifying data collection to include a broader range of demographics and conditions ensures more representative datasets. **Incorporating annotations from diverse medical experts** addresses inter-observer variability. **Federated learning**, by pooling data from various sources, and continuous learning, by updating models with new data, help standardizing and updating diagnostic models. **Domain adaptation techniques** ensure model consistency across different US technologies, while **generative AI** enhances dataset variety by creating synthetic examples of rare conditions. Together, **these strategies can foster a more equitable and accurate framework for prenatal diagnostics.**

Public fetal US datasets are invaluable resources, but may carry inherent biases, potentially skewing research and clinical applications. By promoting awareness of dataset biases and advocating for ethical AI practices, we can pave the way for accurate, fair, and clinically reliable diagnostic tools in prenatal care.

"As engineers and AI ethicists that work on designing trustworthy deep-learning algorithms for medical image analysis, fostering equitable AI is now a priority for us," says Mariachiara's PhD Tutor **Sara Moccia**. *"Thanks, Mariachiara, for your work!"*





Shmulik Shapiro (left), RSIP Vision's EVP of Business Development & Strategy, talks to us about his visit with Arik Rond (right), Head of AI Solutions for Urology and R&D Team Leader, to this year's Annual Meeting of the American Urologists Association (AUA).

Participants from healthcare, academia, and industry came together last month for the Annual Meeting of the American Urologists Association in picturesque San

Antonio. The event was a hive of innovation and enthusiasm for the new technologies on display and the future potential of AI in this field.

"I was happy to see the whole ecosystem again," Shmulik tells us. "Seeing all the participants, physicians, companies, and new technologies was exciting. San Antonio is a beautiful city. I didn't even know how beautiful it was, with its unique River Walk and water system. That was a great surprise. The conference was larger than last year. That's very positive to see, very encouraging."

Shmulik spotted two key trends: **a new line of disposable solutions getting out to the market and AI in urology becoming more of a reality.** *“Four or five years ago, it was only a dream and a buzz talking about AI,” he recalls. “Today, we see it getting more and more into the solutions and the next generation of products, in different aspects of the market. That’s a great blessing because there’s no doubt that these tools can assist in the workflow and diagnostic issues of urology all across the board.”*

RSIP Vision is a regular visitor to the AUA Annual Meeting, bringing the most advanced technologies it has developed, including innovations like **MRI and ultrasound fusion solutions, automated segmentation of organs and tumors, and AI engines for lesion detection and classification.** Shmulik noted a growing acceptance and interest in these technologies, with ongoing projects aimed at integrating AI into next-generation medical devices.

Accompanied by **Arik Rond**, RSIP Vision’s very own expert in

urological AI, the visit was also an opportunity to learn and network. *“It was great being with Arik again,” he smiles. “Arik is a great team member and a real expert in urological AI. I think one of the leading in the world. He’s very modest and would never say that, but I will! Also, he has a lot of respect for the technical teams of the large manufacturers. They already know him, so having these meetings together is great.”*

The pair observed significant progress in **robotic and endoscopic systems**, promising to bring advanced surgical capabilities not there in the market today. Endoscopic systems are harnessing the power of AI and becoming smarter, while the hardware itself is constantly improving, particularly in image quality. Better images mean the algorithms can work better, leading to enhanced diagnostic accuracy and improved treatment options for patients.

Comparing the progress of urology to that of cardiology, where AI is already widespread, Shmulik points



**Visual Intelligence
for MedTech**



event in Las Vegas is already underway. Shmulik tells us there is already a strong sense of anticipation in the air. *“By next year, we’ll have more solutions to show, and the existing solutions we’re working on will be much more advanced,”* he reveals. ***“I want to see these AI tools in the market and the clinics! We want to see them in the booths being sold and the physicians appreciating the assistance they get from them!”***

out that urology is playing catch up, but it is getting there. *“That’s very exciting, and everybody is talking about it,”* he says. Laughing, he adds: *“I met a colleague, and they had a joke during the AUA that until the age of 50-55, the best thing to have is a cardiologist friend, but from that age going on, you should have a good urologist as well!”*

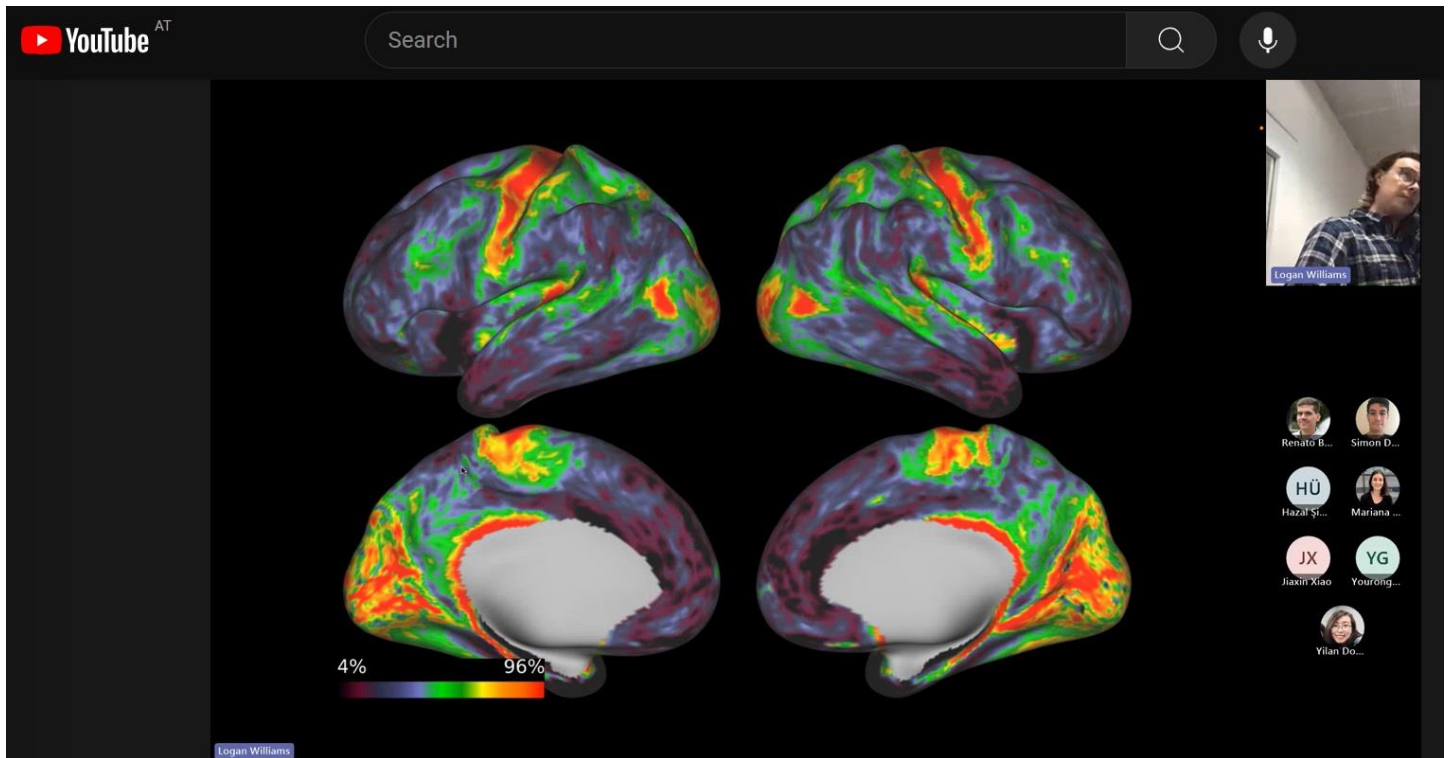
Is that a hint that RSIP Vision has more innovation in the works? *“Yeah, definitely,”* he teases. *“We always have something great cooking!”*

If you think we could help with your project, [contact RSIP Vision](#) today for an informal discussion about your work!

One of the most memorable and engaging experiences at the event was a booth designed around a Back to the Future theme, symbolizing **older treatment concepts being used in new ways to improve results today**. Its creative setup featured a DeLorean car, complete with the iconic flux capacitor and a hoverboard, reminiscent of the famous movies. *“Everybody went around it and took pictures,”* he beams. *“It was a lot of fun, and from a marketing perspective, it was great because everybody came and heard their spiel. Hopefully, they had a lot of orders as well.”*

No sooner has AUA2024 drawn to a close than planning for next year’s

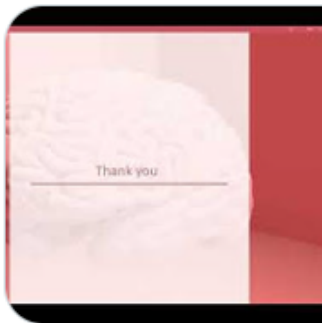




Emma R - @metricsemma.bsky.social

@emrobSci

Hi everyone, our group run tutorials alternate Fridays on cortical surface processing, MRI pre-processing, neuroscience and AI methods (soon to be expanded to other organs) - I've started to upload our talks to YouTube. I hope they are useful



youtube.com

MeTrICs group meetings

Recordings from our group meetings - which offer tutorials on a range of topics from cortical surface modelling, to ...

10:07 PM · May 14, 2024 · 7,760 Views



1



40



93



60



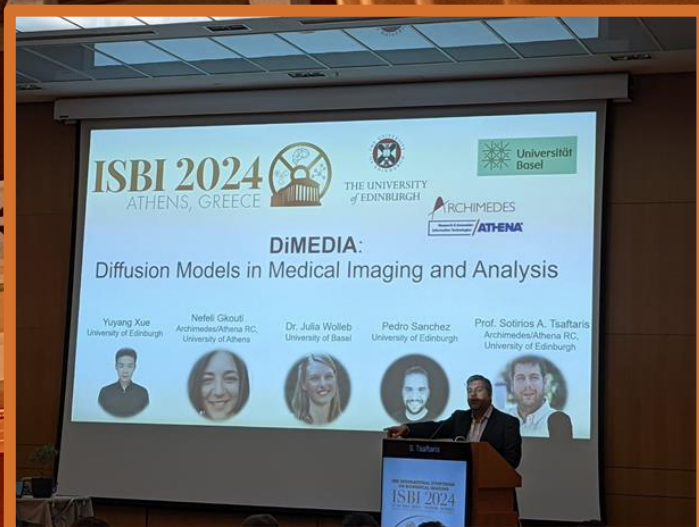
You can subscribe to awesome Emma Robinson's YouTube channel and watch a new tutorial every two weeks 😊



ISBI 2024 was held in Athens, GR.

All photos courtesy of Sidra Aleem. Sidra is a PhD student at the School of Electronics, Dublin City University. At ISBI she presented her work titled "ConvLoRA and AdaBN based Domain Adaptation via Self-Training".

The proposed adapter "ConvLoRA" is a parameter efficient adapter specifically designed for application in CNNs, presenting a novel approach to address domain adaptation challenges in the context of image data. She's standing (left) at her poster near Claudia Mazo, an Assistant Professor at DCU."





Stéphanie Allassonniere • 1st

Professor and Vice-President Valorisation, Uni...

3w • 🌐

I am so glad to see the tremendous achievements of [Sonio](#) now joining [Samsung Healthcare](#)!

This adventure started in 2016 with the PhD of [Remi Besson](#) that I had the chance to co-advise with [Erwan LE PENNEC](#) and with the continuous inputs of [Yves Ville](#), [Julien Stirnemann](#) and [Emmanuel Spaggiari](#).

Then [Cecile Brosset Dubois](#) joined the team and made our dreams a reality.

I am full of gratitude to the Sonio team for this incredible journey.



Samsung Healthcare

57,098 followers

3w • Edited • 🌐

+ Follow

We are very pleased to announce that we signed an agreement to acquire 100% of the shares of [Sonio SAS](#), a fetal ultrasound AI software company.

Stay tuned for more updates on this groundbreaking collaboration! 🙌



Collaboration

samsunghealthcare.com

Congrats,
Stéphanie Allassonnière
and Sonio team!

HELMHOLTZ
MUNICH

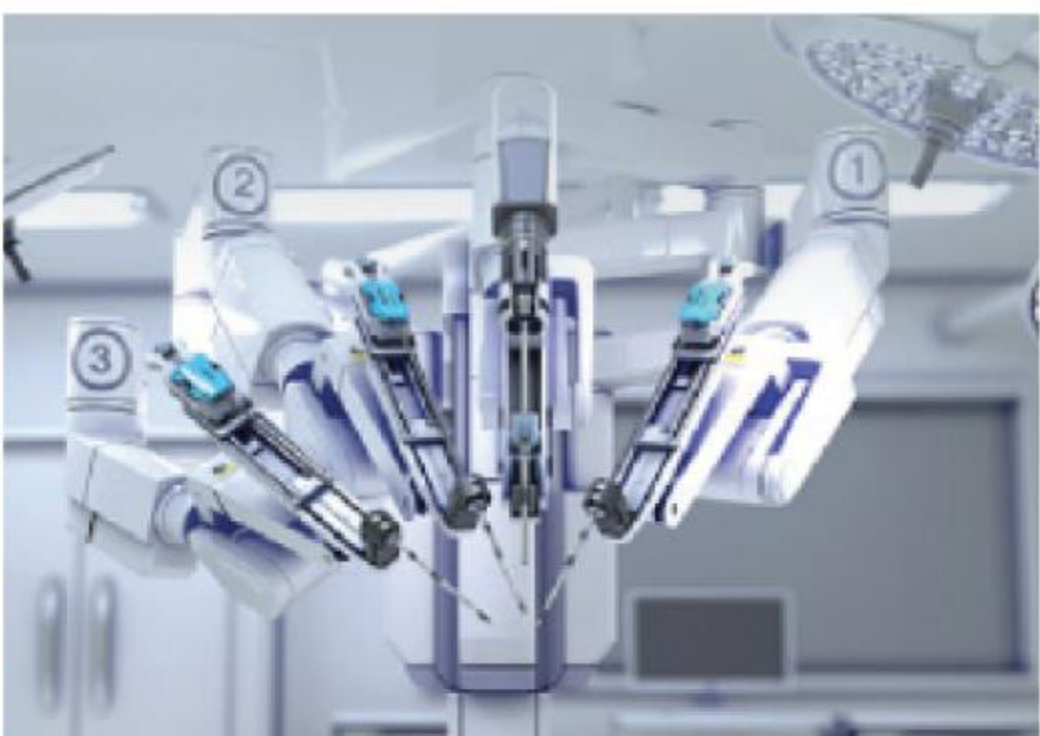
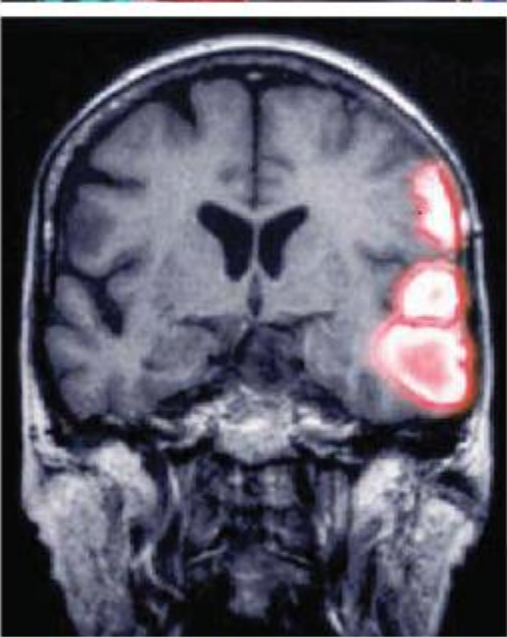
News



Alfred-Breit-Preis der
Deutschen Röntgengesellschaft

Prof. Julia Schnabel

Congrats to awesome Julia Schnabel for receiving the Alfred Breit Prize of the German Radiological Society. Julia is Director at the Institute for Machine Learning in Biomedical Imaging at Helmholtz Munich and she has been honored for her outstanding achievements in the field of radiological research, which have made a significant contribution to advances in the treatment of both benign and malignant diseases. For example, by applying machine learning, her work enables precise and non-invasive measurements of placental properties such as blood flow, size, and tissue composition. These findings help to detect pregnancy complications at an early stage and improve prenatal care.



IMPROVE YOUR VISION WITH Computer Vision News

SUBSCRIBE

to the magazine of the
algorithm community
and get also the
new supplement
Medical Imaging News!

