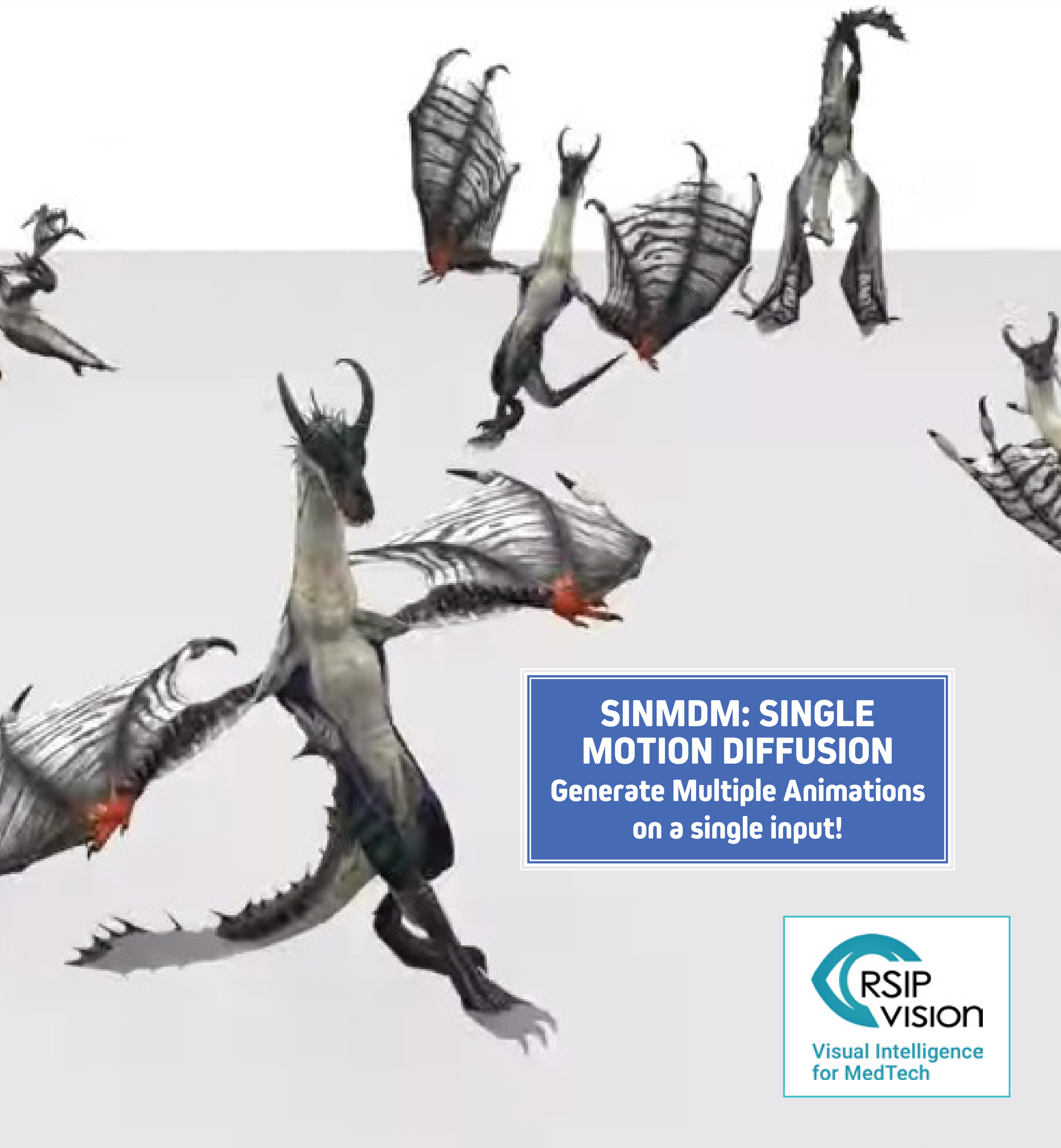
APRIL 2023

Computer Vision News

& Medical Imaging News

The Magazine of the Algorithm Community



SINMDM: SINGLE MOTION DIFFUSION

Generate Multiple Animations on a single input!

This photo was taken in peaceful, lovely and brave Odessa, Ukraine.

Computer Vision News

Editor:
Ralph Anzarouth

Engineering Editors:
Marica Muffoletto
Ioannis Valasakis

Publisher:
RSIP Vision

Contact us

Free subscription

Read Previous Magazines

Follow Us

Dear reader,

Welcome to the April edition of **Computer Vision News**. We have an action-packed issue for you this month, so let's get straight to it!

It's said the eyes provide a glimpse into the soul, but did you know your eyes can be a window into your overall health? **Miguel Bernabeu** and **Ylenia Giarratano** from the **University of Edinburgh** take us on a deep dive into their cutting-edge retinal imaging research and their wider work developing **AI systems for medical applications**.

Master's students **Rodrigo** and **Hanna** from the **University of Calgary** discuss their **Magnetic Resonance Spectroscopy (MRS) challenge** for the upcoming **ISBI 2023 conference**. Can MRS join MRI and break through the clinical barrier?

We chat with **Hilla** and **Coralie**, co-founders of **MeNow**, an Israeli startup shaking up the cosmetics industry **using the power of AI to curate personalized skincare solutions for consumers**.

If you're a fan of animation, you won't want to miss our report on **SinMDM**. **Sigal** and **Inbal** from **Tel Aviv University** tell us about their new lightweight algorithm that can **generate multiple animations from a single input**. Why settle for one breakdancing dragon when you could have an army of them?

We also celebrate the achievements of two exceptional scholars, **Camila González from TU Darmstadt** and **Alberto Testoni from the University of Trento**, in our regular feature: **Congrats, Doctor!**

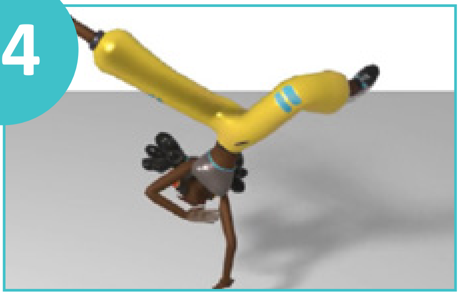
This month, our **Woman in Computer Vision** is Iranian-born **Seyran Khademi** from the **Delft University of Technology**. As well as filling us in on her fascinating work at the intersection of architecture and computer vision, Seyran speaks about the brave women in her homeland fighting for their freedom, making a heartfelt plea: "**Stand with the right side of history!**"

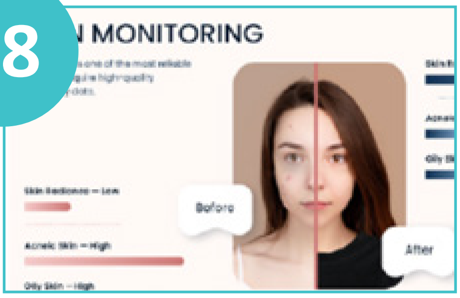
Finally, mark your calendars because, on April 19, **Associate Professor of Urology Michael Gorin** from the **Icahn School of Medicine at Mount Sinai in New York** will host a free webinar: **AI-Enabled Urology – A Clinician's Perspective**. Don't miss his valuable insights into recent advancements in this exciting field. **Sign up now!**

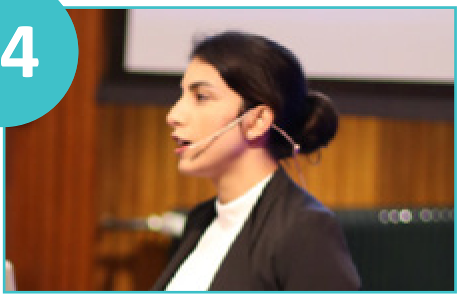
Sit back, relax, enjoy the read, and please, spread the word!

Ralph Anzarouth,
Editor, **Computer Vision News,**
Marketing Manager, **RSIP Vision**

Computer Vision News

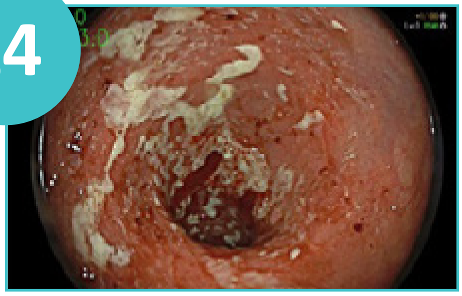
04


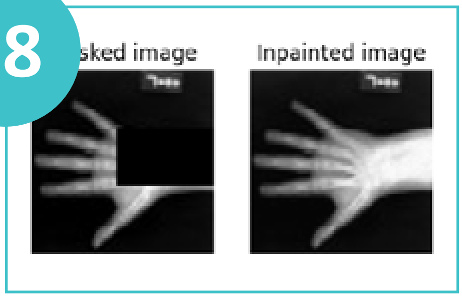
08


14


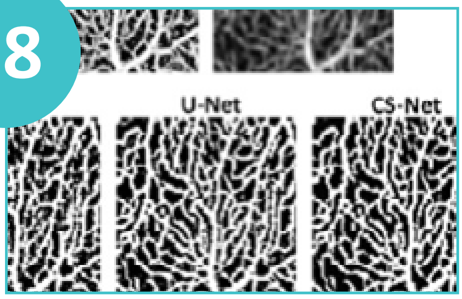
20


Medical Imaging News

24


28


37


38


SINMDM: SINGLE MOTION DIFFUSION

Sigal Raab is a PhD candidate, and Inbal Leibovitch is a master's student at Tel Aviv University. They speak to us about their paper proposing a novel technique for synthesizing realistic animations.

Sigal Raab

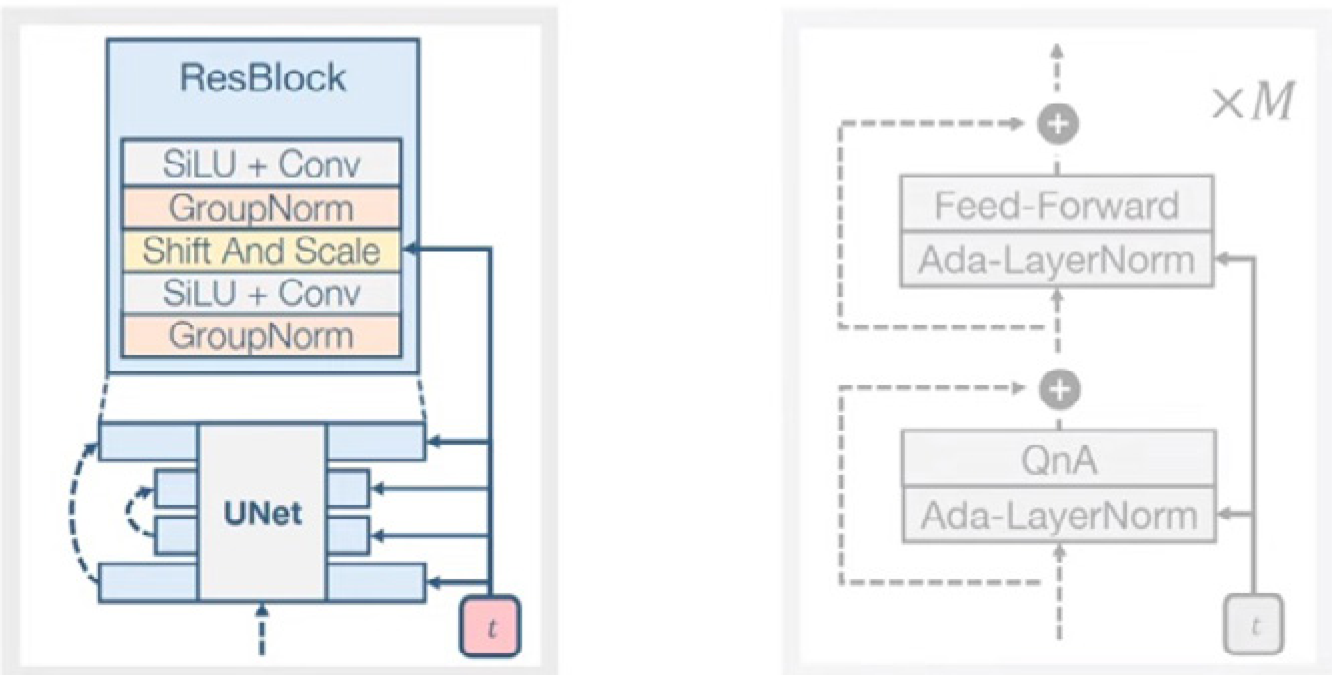
Inbal Leibovitch

Creating animations for films can be a time-consuming and costly manual or semi-manual process. Even if similar animations exist, animators must start from scratch to make any necessary variations.

Sigal and Inbal present **a new lightweight algorithm that can generate multiple different animations based on a single input**. This breakthrough could revolutionize the animation industry, making the process of creating animations faster, more efficient, and more cost-effective.

"In this work, our goal is to take one input

UNET Architecture



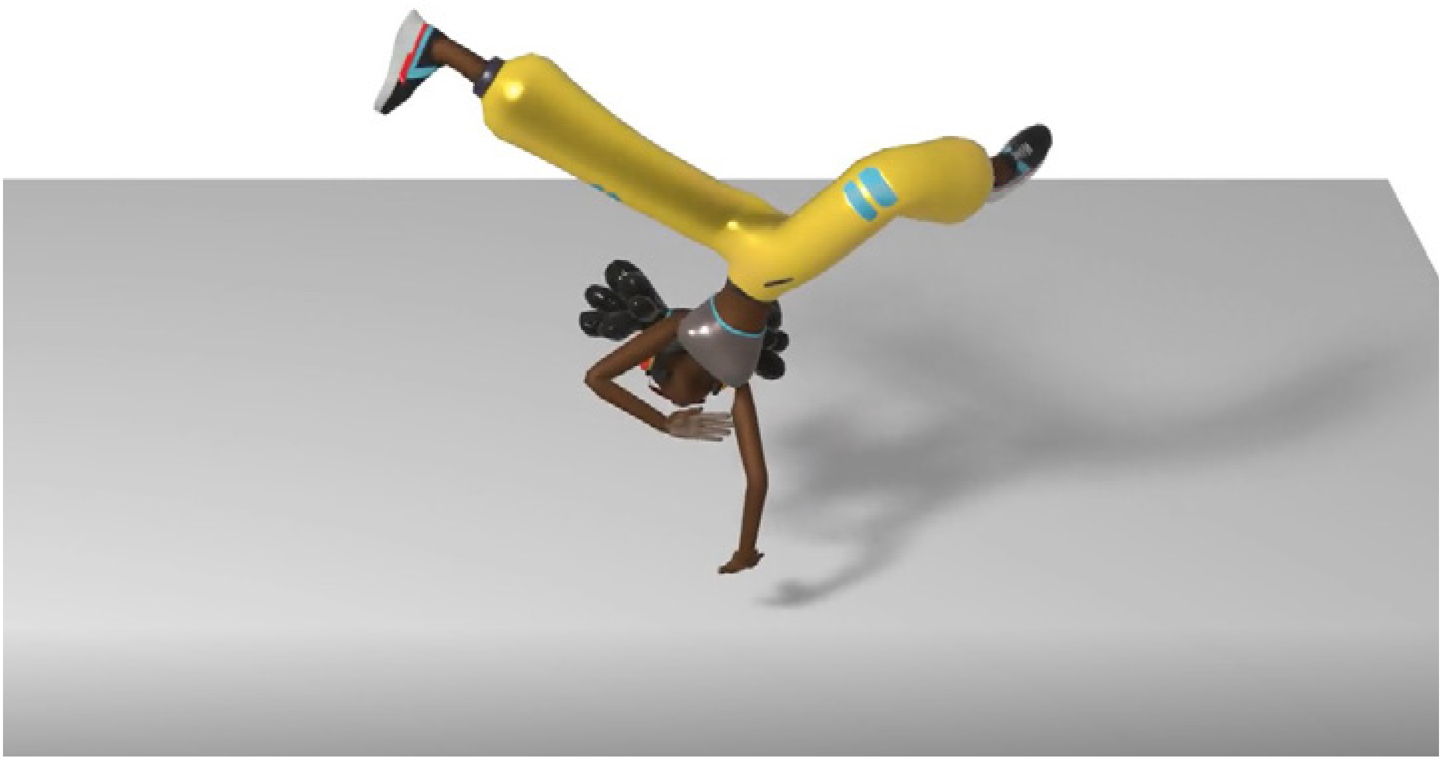
animation – for example, a breakdancing dragon – and create as many different animations as we want," Sigal tells us. *"All these different animations are faithful to the motion motifs of the original, but no motion is the same as another. **The dragons dance together but not in the same order. Only the semantics are the same!**"*

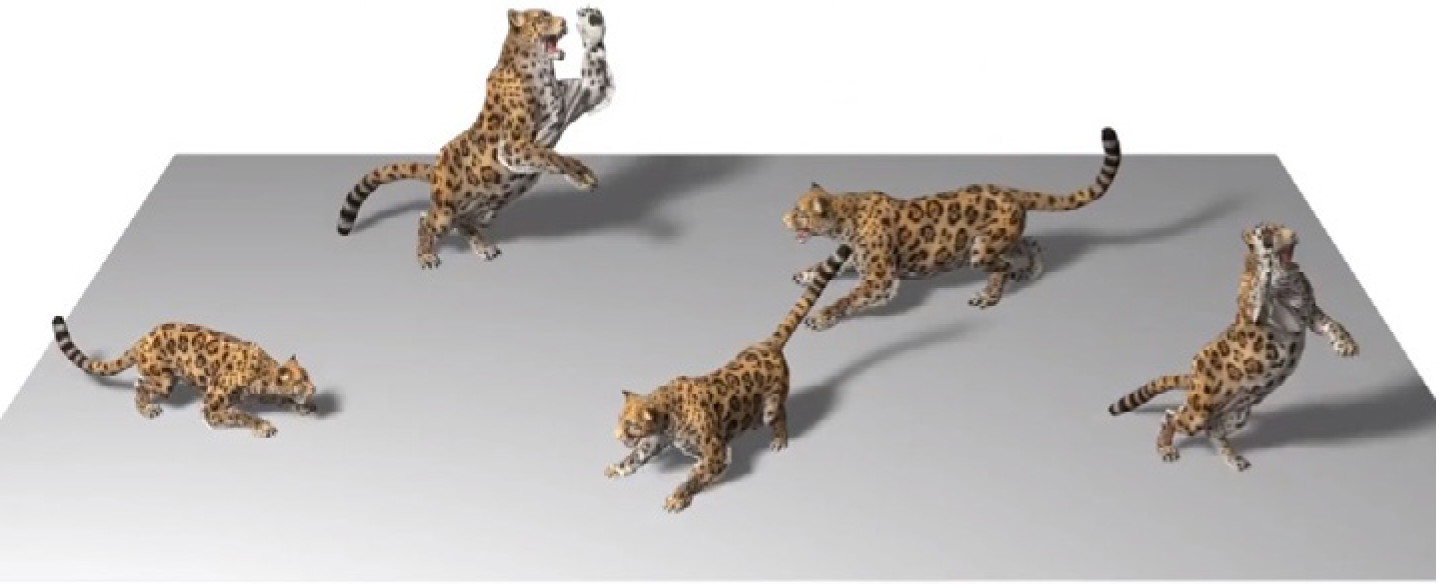
The new algorithm has significantly improved the speed and memory requirements for generating animations by utilizing the power of diffusion models and a denoising network. Traditionally, diffusion models were believed to require extensive datasets, but **this model can run on just one input**. The resulting animations are superior to previous work that used GANs, leading to a more complicated, slow, and memory-intensive algorithm.

"Diffusion models for human motion modeling is a new region," Sigal, who recently took the relatively unusual path of returning to academia after many years working in the industry, points out. *"I was part of one of the pioneering works, which showed that, unlike image diffusion models that use the famous **U-Net architecture**, it's better to use **transformers** for human motion diffusion. However, standard transformers have a receptive field that contains the whole motion. If we run a network with that, we get an overfitting problem. The output will be the same motion, again and again, just replicating the input, when we want diversity."*

To run a transformer that uses an attention mechanism on motion, you can employ local attention in non-overlapping windows, similar to **ViT - Vision Transformers**. However, using non-interleaving windows tends to constrain the cross-window interaction, adversely affecting model performance. Smaller windows are necessary to reduce the size of the receptive field. **The QnA or query and attend algorithm provides a solution to this problem**.

"Existing algorithms can use interleaving

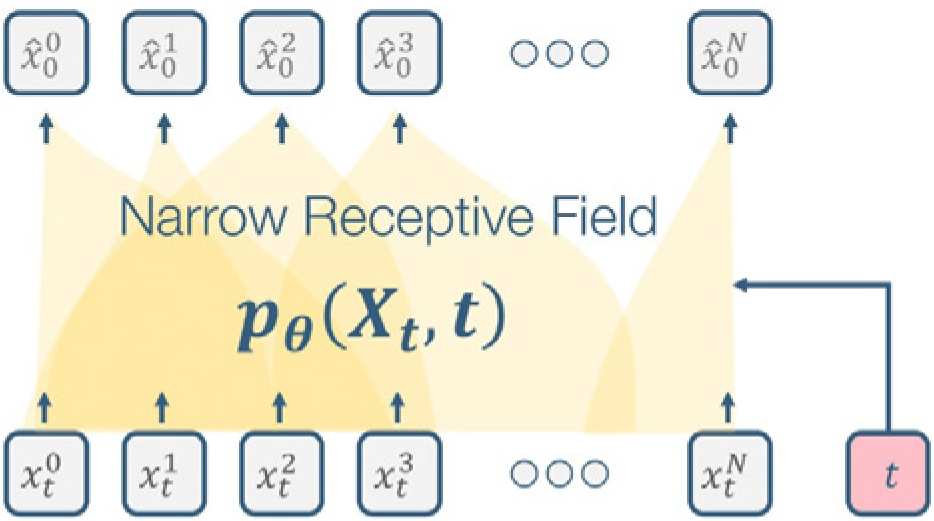




windows, like partially overlapping windows, but will take hours, days, or weeks to run them," Sigal explains. *"QnA solves this, using partially overlapping windows without a performance problem. The key idea behind QnA is that all those windows learn and share the queries, unlike the traditional attention mechanism. QnA is very successful for the imaging domain, and we're the first to use it for the motion domain, which is how we get the temporal receptive fields to be smaller."*

Relative positional encoding is also essential to achieving optimal results. In traditional transformers, positional encoding is global, meaning that when relating to a timeframe, you always relate to its global index in the time series. This approach fixes each frame to a specific point in time, hindering the mixing of motions across different timeframes. However, relative positional encoding can shift sub-motions from the end of the sequence to the beginning and vice versa, enabling greater flexibility and creativity in the animation process.

The use of diffusion models in **SinMDM** can also facilitate various applications at inference time without further training for specific tasks.



"The first and most basic application would be ***crowd animation****, where you want a few animations that look like they're doing the same thing, but not exactly the same,"* Inbal explains. *"**Synthesising motions with our network would**

give you that. The network can also generate long motions. *Let's say we train on a two-second dance sequence, we can then generate a full minute of the same dance, and it would look continuous rather than repetitive. Another application is style transfer, where we train the model on a style motion, like someone walking happily, and then transfer that style to a new motion using a harmonization technique."*

Other applications include **temporal composition**, where given the prefix or center of a motion, the network can synthesize the rest of it, introducing diverse outputs, and spatial composition, where the network uses certain body parts as input and generates others. For example, the motion of the arms is based on a reference motion, and the model uses learned motion motifs to synthesize the motion of the legs or vice versa.

Moving forward, the team anticipates that other researchers could build upon their work by leveraging the key concepts that made the diffusion model effective in learning from a single animation and exploring new applications beyond those already introduced.

Apart from contributing to this project, Inbal has been actively pursuing her master's degree in computer science and works at Microsoft.

"It was a big honor for me to work with such smart people on this paper," she reveals. *"I've learned a lot from them, including how they work, how to approach things, and how to get the best paper done in time. It's been great fun!"*

MENOW

WINNING OVER THE COSMETICS INDUSTRY WITH THE POWER OF AI

Hilla Ben-Hamo Arbel and Coralie Ebert are the co-founders of MeNow, an Israeli startup winning over the cosmetics industry with the power of AI. They are here to tell us all about their revolutionary software.

Cosmetics is an industry worth $570 billion globally and predominantly targeted toward women. In the US alone, **women apply an average of 168 chemicals to their skin daily**, with younger women exposing themselves to even more. While cosmetics encompass a wide range of decorative products, MeNow focuses more on the medical side, specifically skincare and haircare.

CEO Hilla, who has a PhD in biotechnology engineering, tells us that most individuals are unaware of which products are right for them: *"We apply a lot of ingredients on our skin, and we don't know how it will react to them. There's no way to check it. It's not like drugs, which go through many experiments. Ingredients must be approved, but they mix them in the lab, approve it, and sell it. Many people suffer from side effects because of that. Sometimes they don't even know why because no one did the research."*

There are over 50,000 ingredients in the European cosmetics databases and many plants you can use, but MeNow CTO Coralie tells us the **regulation is lenient**: *"I can decide to put a plant in a cosmetic product and build a story behind it.* ***There's no one checking that!***"

Every person is different, and many factors can affect how our skin reacts, including gender, age, and ethnicity. The average person buys products to improve, protect, or feel better about their skin but does

not know if they are worsening the situation.

"Go into Facebook groups, and there's a whole host of people suffering from atopic dermatitis, acne, or pigmentation," Hilla points out. *"Some talk about how a product was amazing for them and changed their lives, but some say it made their lives worse. There are many contradictions."*

Only AI!

Personalized solutions matching people to the right ingredients for their skin type do not yet exist in the market. With **skin diseases affecting almost one-third of the world's population**, around 85% of young people experiencing acne at some point, and up to 50% of babies suffering from diaper dermatitis, MeNow has identified a massive need. Is AI one of the ways to tackle this?

"**Only AI!**" she clarifies. "I don't see any other way it's going to work. There's not a person that can solve this issue, and people are trying."

Coralie agrees: *"Doctors can read clinical studies and scientific publications and get some understanding about what a drug is good for, but usually drugs have one active ingredient, and with cosmetics, there are often seven or eight. Some are plants, and you don't have scientific papers on their effects. Ingredients can enhance or negate each other. Sometimes they do the same thing and compete.* ***It's impossible to make any prediction without a computer!****"*

It sounds like magic, but it's not!

MeNow's software-as-a-service platform is the solution. Their patent-pending AI can simulate the effects of a particular combination of ingredients without requiring a human to sift through all the potential interactions manually.

"There are hundreds of molecules in a plant," Coralie notes. *"You can't sit and read about each molecule and what it does and compute the interaction in your head. For most, no one knows what they do, and what is healthy for one person can be bad for another."*

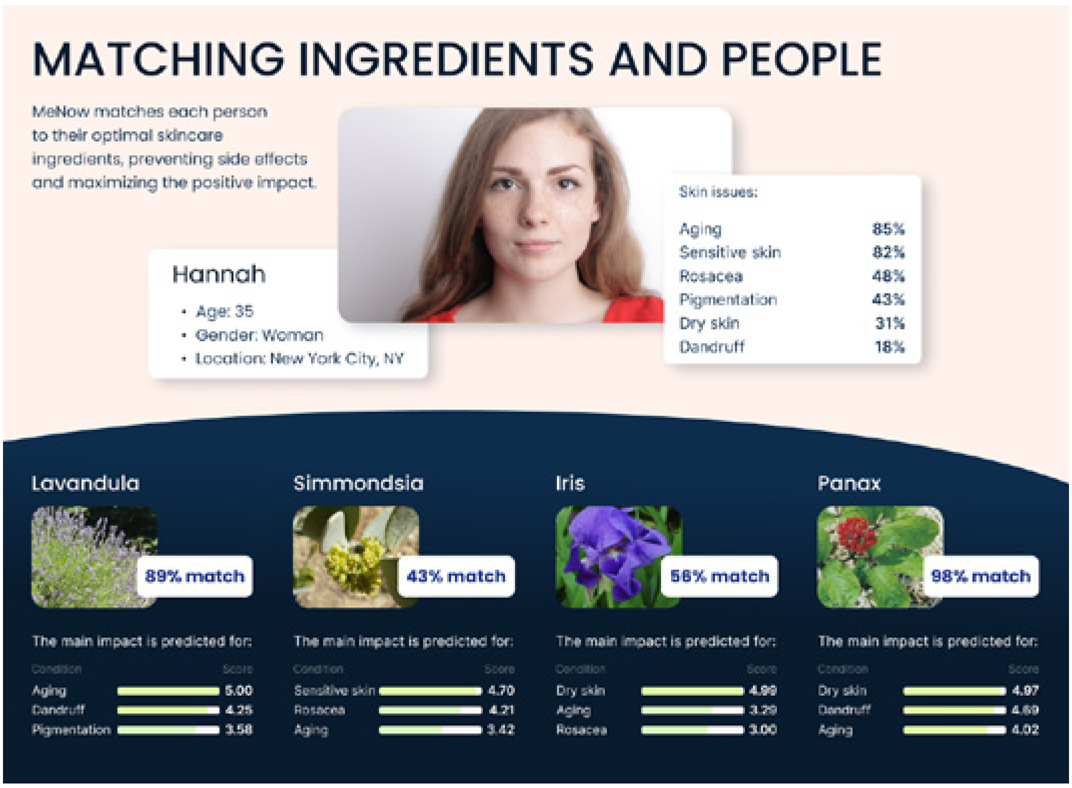
Chemicals with different names can be identical, while the same molecule in different contexts can have a different formula. Therefore, much effort goes into standardization, reconstructing the picture, and integrating information for analysis.

"We have an algorithm that can take the structure of the molecules and, based on these structures and these properties, predict what will happen," Coralie continues. *"**It sounds like magic, but it's not!** In biology, there are things you know the molecule is going to bind with, and we use this to predict the effects."*

Understanding the person is also crucial. It is not just about whether an ingredient is good or bad but whether it is good or bad for specific people under specific conditions. It is not enough to categorize skincare products as being for dry, oily, or normal skin. Ethnicity and skin tone also plays a role in determining the appropriate skincare routine. To identify these factors and cater to each person's unique needs, **MeNow uses a combination of survey questions and imaging**.

"People use the app to take pictures, and we use the images to decrease the number of questions we need to ask," Coralie reveals. *"We have an algorithm that profiles the user using a mix of computer vision and their answers. We detect their face, do a 3D reconstruction, and have a database of features and specific conditions we're detecting and segmenting. At the end of the process, we recommend a person with specific products. It's also about monitoring. **If you return to me and say, 'This cream didn't work,' we can see what went wrong.**"*

This monitoring aspect can be especially important for younger women with cyclic hormonal variations that can impact the skin. The AI considers this, using **monitoring and landmarks in the images to track changes over time**. However, Hilla and Coralie are keen to point out that the software provides valuable insights but should not be seen as diagnostic, and they have no intention for it to be classified as a



medical device.

It's not a black box!

Coralie, who has a PhD in computational neuroscience from the **Weizmann Institute in Israel**, explains that the system uses **Bayesian networks rather than deep neural networks** more familiar to the field: *"People are starting to understand that* ***the brain works like a Bayesian machine****. I love this model because Bayes' rule is so simple. You use less computational power and have predictions you can understand.* ***It's not a black box! These networks are probabilistic, and we live in a probabilistic world.*** *Having a framework that can take probability straight as input and give back probability in return is convenient and easy to use and interpret. It allows us to integrate a lot of information and compute anything missing."*

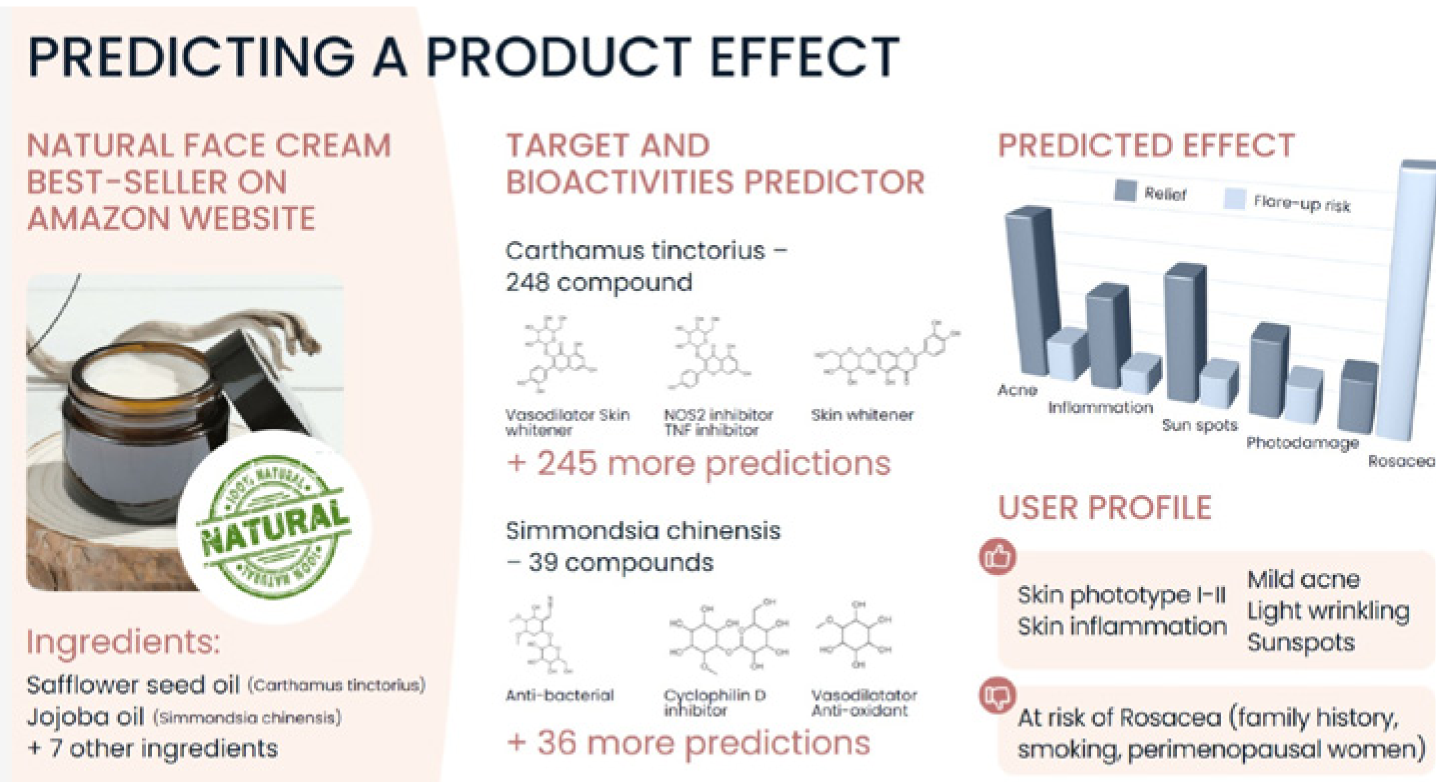
Behind the scenes, MeNow is a B2B company that has a diverse team of experts in many fields, ranging from chemoinformatics, bioinformatics, and plant biology to computer vision, machine learning, and AI.

It's in their interest to give you something good for you!

The software is already generating a buzz as **companies recognize personalization as the direction the whole field is going**. Consumer behavior has changed from the days when people stuck rigidly to their chosen brand.

"Cosmetics is so competitive," Coralie affirms. *"It's not a field where people can be scared of change. You have to go with it. Most companies want to do good. We can suggest modifications to their products or different populations for them to target.* ***They're enthusiastic because they see it as a way to gain more information and improve things!****"*

Hilla adds: *"The new generation is just looking for the next best thing. Companies won't survive if they don't supply the right products for their customers.* ***It's in their interest to give you something good for you!****"*



Congrats, Doctor Alberto!

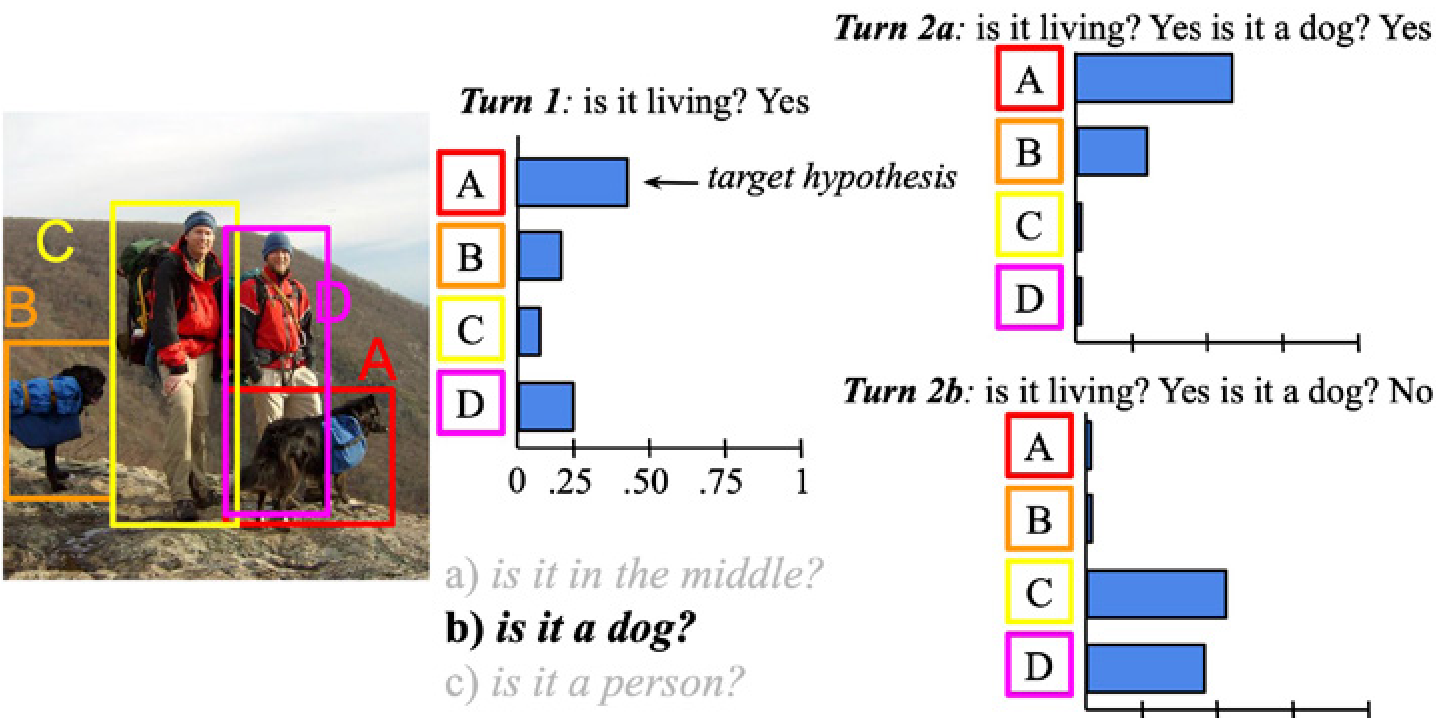
Alberto Testoni has recently completed his PhD at the University of Trento, supervised by Raffaella Bernardi (with him in the photo). His research focused on investigating Natural Language Generation techniques in Visual Dialogue tasks. After demonstrating that existing metrics do not represent a good proxy of dialogue quality and visual grounding skills, he proposed to evaluate dialogues on a deeper level by taking into account the strategy across dialogue turns. He explored how cognitive studies on question-asking and information-seeking strategies can inspire the development of better decoding strategies. He is now a Postdoctoral Researcher at the University of Amsterdam. Congrats, Doctor Alberto!

Chatbots and conversational agents have recently become part of our daily lives. They are often conceived as passive agents that reply to users' inquiries, but it is crucial to evaluate their ability to actively ask questions to solve problems. Asking questions is essential to resolve ambiguities, gather information about the world, and identify objects mentioned in the conversation, especially when chatbots can process visual inputs about the surroundings. From Cognitive Science, it is well known that asking informative questions requires a plethora of high-level cognitive skills, and modeling this feature in computational models represents an important challenge.

In his work, Alberto takes referential guessing dialogue games as a testbed to evaluate the question-asking skills of visually-grounded Natural Language Generation models. Using existing surface-level metrics, he shows that machine-generated dialogues are often repetitive, and improving over these metrics does not help to achieve a higher task success

Alberto proposes to evaluate machine-generated dialogues on a deeper level by capturing the interplay between the Encoder and Decoder components of neural architectures. He considers entity hallucinations (generation of words that are not coherent with the image upon which the conversation occurs) as a case study. Hallucinations are shown to create a detrimental cascade effect on consecutive dialogue turns. By adapting Transformer-based models to these tasks, he finds that more sophisticated visual processing plays a crucial role in reducing hallucinations.

The progressive advance towards even deeper evaluation criteria led Alberto to study the effectiveness of the question-asking strategy in humans and machines. Inspired by cognitive studies on children and adults, Alberto proposes Confirm-it (Figure 1), a model based on a beam search re-ranking technique that implements a confirmation-driven strategy. Confirm-it outperforms different decoding strategies against both surface-level and more fine-grained metrics, as well as generating dialogues that are most informative also for humans playing the same task. Finally, Alberto broadens the horizons on what is still missing from achieving human-like dialogue systems by presenting a large-scale study of human conversations used to train computational models to unveil the pragmatic phenomena that make human communication successful in Visual Dialogue tasks. For more information, see [Alberto's website](#).



In the "GuessWhat?!" game, an Oracle is assigned a target object in an image, and a Questioner has to ask questions to identify it. The charts show the probability distributions over different objects after dialogue exchanges. Among the possible follow-up questions(a-c), Confirm-it selects the one that tests the target intermediate hypothesis (b).

Seyran Khademi is an assistant professor at Delft University of Technology in the Netherlands at the Faculty of Architecture and the Built Environment.

Seyran, what do you do at Delft?

I came to Delft in 2011 to do my PhD in signal processing. Signal processing is between mathematics and data science. It's the science behind our digital life, basically. I did my PhD for four to five years at the Department of Signal Processing in the signal processing lab. Then I moved to computer vision as a postdoc. In between, I did a bit of speech processing as well. So it's been a journey for me to come from low-level signals to really complex signals like images, videos, and any kind of visual data. Right now, I am working on AI for visual data processing at the Faculty of Architecture and the Built Environment. As architecture is by nature a visual discipline, designers and architects communicate their design ideas very much with visual data. And with the whole era of big data, we need intelligent models to be able to process all this kind of data and to make sense of them. What I'm doing right now is developing and working on AI methodologies for architects to better explore their visual domain.

Our readers know already one lady who works at the intersection between architecture and computer vision, Iro Armeni. Maybe you have heard about her?

Yeah, I know her! I'm actually kind of collaborating with her right now. [*laughs*]

It's a small world after all!

[*laughs*] Yes, indeed. In this intersection, it is. I think the difference is that she actually came from the world of architecture to computer science. And for me, it's the other way around. I started as a computer scientist and moved to the domain of architecture.

Do you have a preference between one world and the other?

Well, my comfort zone is actually computer science because I've been there almost forever now. So that's a comfort zone. It's not a preference per se, but I think architecture is also fascinating, especially the people working in this domain: they are kind of mediators between us, nature, and our built environment. They have to develop lots of skills in lots of directions, from art to engineering, to social science. And that makes a very interesting profile. For me, being in this context, it's a very valuable experience to learn from different domains. I think, in general, we are kind of of moving away from a single discipline. Our problem as humanity has become so complex that it really needs to be solved at the intersection of many disciplines.

When you become very famous, we will see your name on a building. Or maybe there will be a computer program with your name. We don't know yet.

[*laughs*] Very interesting! Projecting to the future, I would not know, actually. But I think both are fascinating.

Tell me Seyran, if you were not working in this field, what would you do? What would your career have been?

Very interesting question! You're taking me to the parallel world of me being something else [*laughs*] This is very unimaginable, but I think I would still be teaching. I think the teaching part, apart from the research, is very rewarding for me. I would probably still work in some kind of academic education institute, and I would be teaching something on science still. I cannot imagine working outside of science. For me, it's again a safe zone, a haven that can help me understand

this complex word.

I know that you were not born in Holland. You come from another place.

I'm from Iran originally, from a western part of Iran called Kurdistan, or a Kurdish area in Iran. It's a very mountainous area, so I'm having a bit of a hard time in the Netherlands, which is so flat. [*laughs*]

Were you born in a village?

Not a village but a small town. I was born there, and I was there until I was 18 years old. And then, I started my university in another city, which is a bit north of Iran, in the province of Azerbaijan in Iran. I studied for my bachelor's there in electrical engineering and communication. Telecommunication engineering, back in time. Then I moved to Sweden for my master's.

Did you study at Tabriz University?

Yes, I did my bachelor's at the University of Tabriz. I have strong roots there. I think it's appropriate to bring it up here because of this woman-life-freedom movement that is happening in Iran. I'm not sure if you're aware of it. I think it's like the biggest feminist movement happening in the world now. I think it's worth mentioning here because we are talking about women in science, and this is something that we are proud to advocate for diversity.

We will talk about that in a minute. But before, you were talking about Sweden. Did you go to Sweden to do your master's?

Yes, I went to Sweden to continue my education and to get a master's degree in 2008. That was the point I left Iran. I stayed there for two years before moving to the Netherlands for my PhD.

Do you visit Iran often?

Yes, my family still lives there, so my emotional source was still there.

Many of us have Iranian colleagues and friends. And we don't know exactly what to say because on the one hand, we are friends of Iran. On the other hand, some people in the country are protesting against their government. Which side should we be on? Can you give us your take on this?

It's a very interesting question. Stand with the right side of history. There is enough evidence now with all the news that is coming from Iran. What is happening is not really a political matter. It's a more humanitarian issue. As I said, it's a feminist movement. It started with women trying to gain back their rights. I even wrote this as a tribute in my PhD thesis, to all the women in my country who were fighting for freedom. Because I think that's the battle, an unfair battle. It doesn't matter if you are a man or woman, but especially for women. They have been suppressed. They have not been treated absolutely equally, and they have fought for their rights now for many years. They are very brave. They are in the front line. I think the right side of this story is to support them to be their voice. It's also my mission to use any kind of platform, including this one, to be their voice. Even though I'm not there, I feel I have this responsibility.

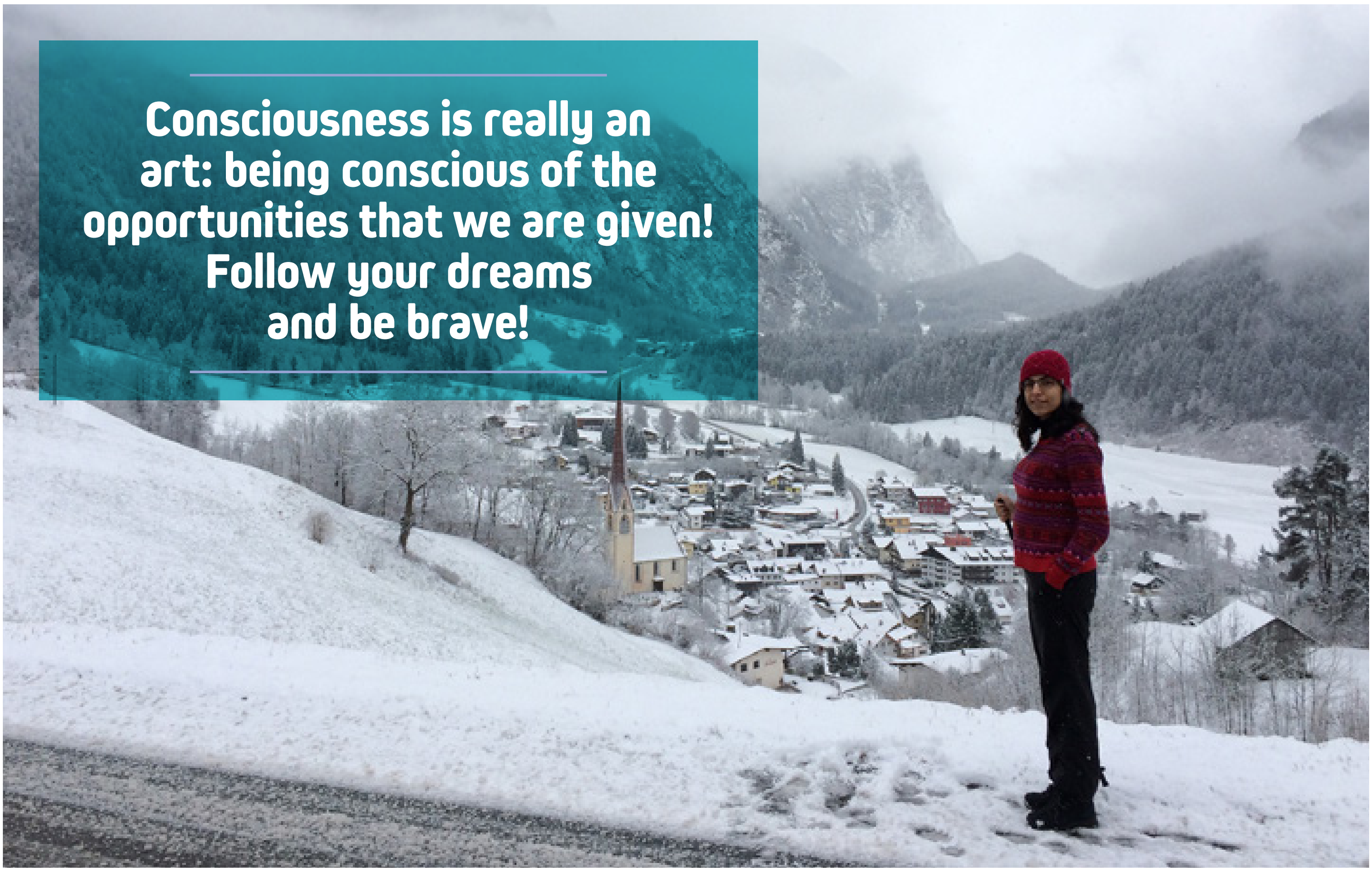
Let's talk now about Iranians outside Iran. How is it to be Iranian thousands of miles from your homeland? What is the feeling?

There are many people actually who share the same experience as expats, especially in science, because you are seeking communities interested in the topic that you're working on. You're willing to travel

anywhere to basically do what you love, right? I think this is very common for people who are working in science to be away from home.

Let's go back to science. We spoke about the past. We spoke about the present, and now let's speak about the future. What does your future hold?

I hope that I will still be doing what I'm doing right now because I'm really loving it. That's basically what I hope for. And I also think the topic that I'm working on, which is AI, is very fascinating and really the topic of the day. I would say this is something that is disrupting both ways, good ways and bad ways, with our life as humans. This is really fascinating for me to understand what would be the future of humans with the presence of this kind of technology. I would like to work on the same topic in the future.



Okay, tell me do you feel more like a researcher or like a teacher?

[*laughs*] Well, definitely both. Teaching is something that I really love. I've always been doing this, since I was a high school student. I love to teach people. I have to explain things. I have to make things simple for others to understand. But I'm also fascinated by technology and how technology is changing our world. And explaining it again to people so that they can do amazing things with it. I think it's really like an equal contribution. They are both very rewarding. Research is long-term rewarding. Teaching is really instantaneous. You have your class, and then you get the energy back from the students.

It has been almost 20 years since you started your academic engagement. Maybe you want to tell me, what is the greatest fruit of your work in these 20 years? What has been the thing that you are most proud of?

If you asked this question to directors or movie makers, they would say, I love all my works! I think when you work on a topic, you really put your whole self into the project. So I have a little bit of something from all the things that I've done. But I think moving from pure science in computer vision and signal processing to more application architecture is very rewarding and interesting. Also, personal development, in a sense, because, as I said, I came out of my comfort zone to explore something new. And then, within this intersection of architecture and computer vision, I have fascinating projects where I've learned a lot. I think this has been the highlight of my career. It's not something that many people are working on. It's a small community, and I'm trying to be part of the community but also, at the same time, build the community around this topic.

Maybe now we can play a game. Let's set an objective for you ten years from now, and we'll check back in ten years to see if you achieved it. What would you like to achieve in the next ten years?

[*laughs*] This is a difficult question. But I think if I can just imagine myself, I would have a group of scientists working on really challenging problems that we are facing right now as humanity. For example, sustainability and climate change. These are the things that really touch our life. Using AI within this context to really do something meaningful. I don't mean to really move to the hype but to solve things fundamentally. Recognizing the challenges that might change people's lives and start working on them. So I think for me, having this community of people with the same vision and me being part of it and maybe leading it. That would be my mission for the next ten years.

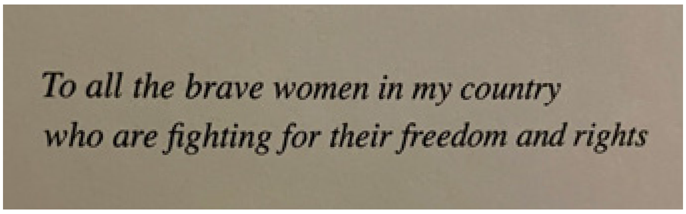
Your last message?

Maybe it sounds cheesy, but follow your dreams and be brave!

Really? Do you know that some people break their teeth by following their dreams?

[*laughs*] Well, we are talking about wisdom here, right? So we are building these dreams based on the knowledge and the wisdom of using that knowledge within our life. So when I'm saying follow your dreams and be brave, it's not that you should ignore the laws of physics.

[*laughs*] I am ignoring the laws of physics, actually. Because I want to be good at roller skating. That's my dream. But I have



broken every possible bone in my body doing that. This is why I do not recommend to follow one's dreams!

[*laughs*] Was it worth it for you? Only you know! I can't be you. It's a trade-off. You can do nothing and be in your safe haven. It's very safe, but then you're missing out on experiencing life! At the same time, you could try things outside your comfort zone. But I think another thing that I could add is consciousness. In this world of AI and technology disruption, consciousness is really an art: being conscious of the opportunities that we are given!

Over 100 inspiring interviews with successful Women in Science!

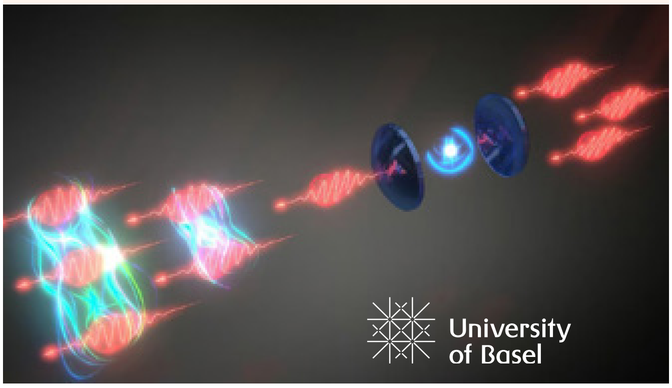
Computer Vision News has found great new stories, written somewhere else by somebody else. We share them with you, adding a short comment. **Enjoy!**

Everything You Know About Computer Vision May Soon Be Wrong

This MIT spinoff called Ubicept believes that computer vision would be a lot faster and better if we skip the concept of still frames and instead directly analyze the data stream from a camera. There are "45 billion cameras in the world, and most of them are creating images and video that aren't really being looked at by a human", so the guys ask why bother with the concept of frames, when this is meaningful for humans and not from machines. Instead, they suggest to measure the individual photons that hit an imaging sensor directly. Their goal is that in 10 years, 50% of cameras will have their technology. Read More

"Quantum light" Manipulation a Step Closer

We never speak about quantum computing and its potential in fields like medical imaging, biology, advanced manufacturing and more. So today we do it. Physicists at the Australia's University of Sydney and Switzerland's University of Basel are among the first team to manipulate photons - particles of light - which are interacting with each other. The scientists like this experiment, "not only because it validates a fundamental effect – stimulated emission – at its ultimate limit, but it also represents a huge technological step towards advanced applications." Read More



French Surveillance System for Olympics Moves Forward, Despite Civil Rights Campaign

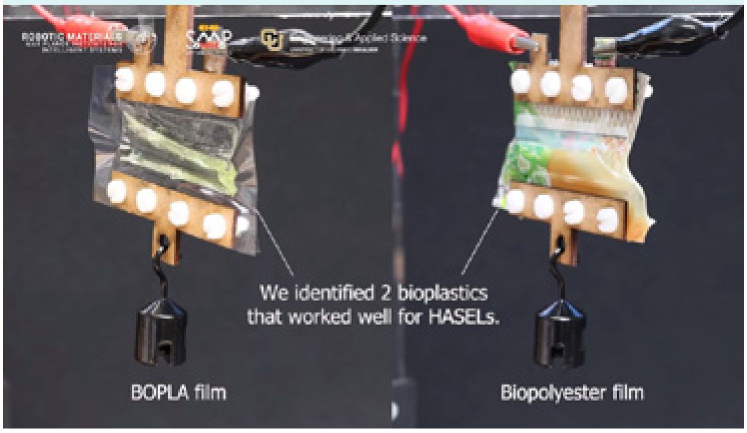
Do you plan to go to the **Paris Olympics** next year? Someone will be watching you. A controversial video surveillance system was just approved by the **French parliament**: it is expected to use automated (with Artificial Intelligence) behavioral surveillance of public spaces during the **2024 Paris Summer Olympics**, ignoring objections from the opposition, **Amnesty International** and more. Opponents of the intrusive plan argue that it will infringe upon privacy rights. Promoters argue instead that the algorithm will help spot suspicious behavior, including unsupervised luggage and alarming crowd movements like stampedes. Read More

The Importance of Reducing Food Waste in Commercial Aviation

How do you like airline food? OK, not much. The story here says that much of it (tasty or not) goes sadly wasted. IATA estimates that about 25% of all cabin waste is untouched food and beverages. Reducing food waste levels in the sky is therefore a goal for carriers. Steps to do this can be very different: besides identifying waste at the production source, one of the methods uses computer vision and machine learning to reduce inflight waste. This technology, via identifiers on meal packaging, can track and differentiate discarded food items and create more relevant and efficient catering plans. Read More

Biodegradable Artificial Muscles: Going Green in The Field of Soft Robotics

Let's end this month's AI Spotlight News with a positive note: journal Science Advances has published the work of an international team of researchers from the Max Planck Institute for Intelligent Systems (MPI-IS), the Johannes Kepler University (JKU, and the University of Colorado (CU Boulder). The scientists collaborated to design a fully biodegradable, high performance artificial muscle – based on gelatin, oil, and bioplastics. They show the potential of this biodegradable technology by using it to animate a robotic gripper, which could be especially useful in single-use deployments such as for waste collection. **Watch the Video**



AeroCloud Launches Computer Vision Passenger Tracking Technology

Stay on board, please! Airport management platform **AeroCloud** has launched **AeroCloud Optic**, a product that uses **computer vision** to track passengers intelligently, anonymously and accurately as they move through an airport. The real-time monitoring of passenger flow means that alerts are triggered in response to bottlenecks in operations, such as extended wait times at check-in or security, allowing these to be immediately addressed with additional resources. **AI** and **ML** are used to enable airport staff to identify trends, predict future scenarios take accurate decisions in both short and long term. Read More

COMPUTER VISION EVENTS

Deepimaging 2023 Summer school Lyon, France 17-21 April	ISBI Cartagena de Indias, Colombia 18-21 April	Innovation Summit Dublin, Ireland 25-27 April
Computer Vision Summit San Jose, CA 26-27 April	ICLR Kigali, Rwanda 1-5 May	WSCG - Computer Graphics Pilsen, Czech Republic 15-19 May
ICDIP - Digital Image Processing Nanjing, Chin 19-22 May	Embedded Vision Summit Sta Clara, CA 22-25 May	
Ital-IA Pisa, Italy 29-31 May		
ICRA - Robotics and Automation London, UK 29 May - 2 June		

FREE SUBSCRIPTION
(click here, its free)

Did you enjoy reading Computer Vision News? Would you like to receive it every month?

Fill the Subscription Form
it takes less than 1 minute!

SUBSCRIBE!

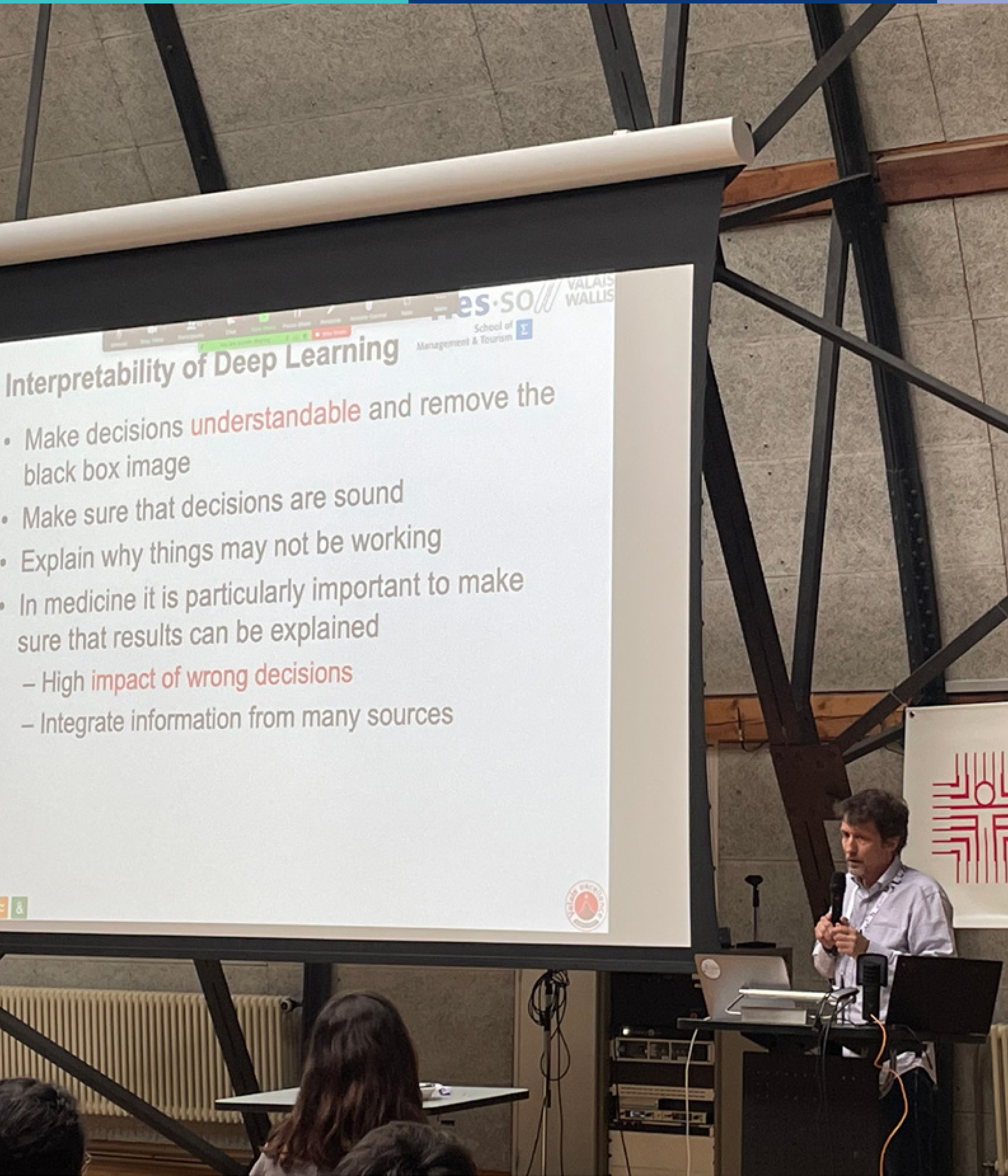
Join thousands of AI professionals who receive Computer Vision News as soon as we publish it. You can also visit our archive to find new and old issues as well.

We hate SPAM and promise to keep your email address safe, always!

Due to the pandemic situation, most shows are considering going virtual or to be held at another date. Please check the latest information on their website before making any plans!

MEDICAL IMAGING NEWS

APRIL 2023



IBD SCORING – CLARIO, GI REVIEWERS AND RSIP VISION TEAM UP TO PRESENT A NEW AI SOLUTION TO ADVANCE CLINICAL TRIALS FOR INFLAMMATORY BOWEL DISEASES

Innovative, human-level AI technology will improve efficiency and consistency of Inflammatory Bowel Disease (IBD) scoring, advancing clinical trials of novel treatments for these debilitating ailments.

Clario, a leading healthcare research and technology company that generates the richest clinical evidence for the clinical trials industry and **GI Reviewers**, LLC Gastroenterology consultants and central readers, together with **RSIP Vision**, an experienced developer of groundbreaking AI technologies for medical imaging, today announced their new AI-based scoring system for inflammatory bowel diseases. The solution is expected to improve the reproducibility of colonoscopy video scoring, optimize workflow by shortening the time required for a human expert reader to score, and ultimately improve recruitment and reduce costs associated with clinical trials. Peer-reviewed results demonstrating the human-level performance of this innovative system were presented to the pharmaceutical and clinical trials industry community at the prestigious European Crohn's and Colitis (ECCO) conference this week.

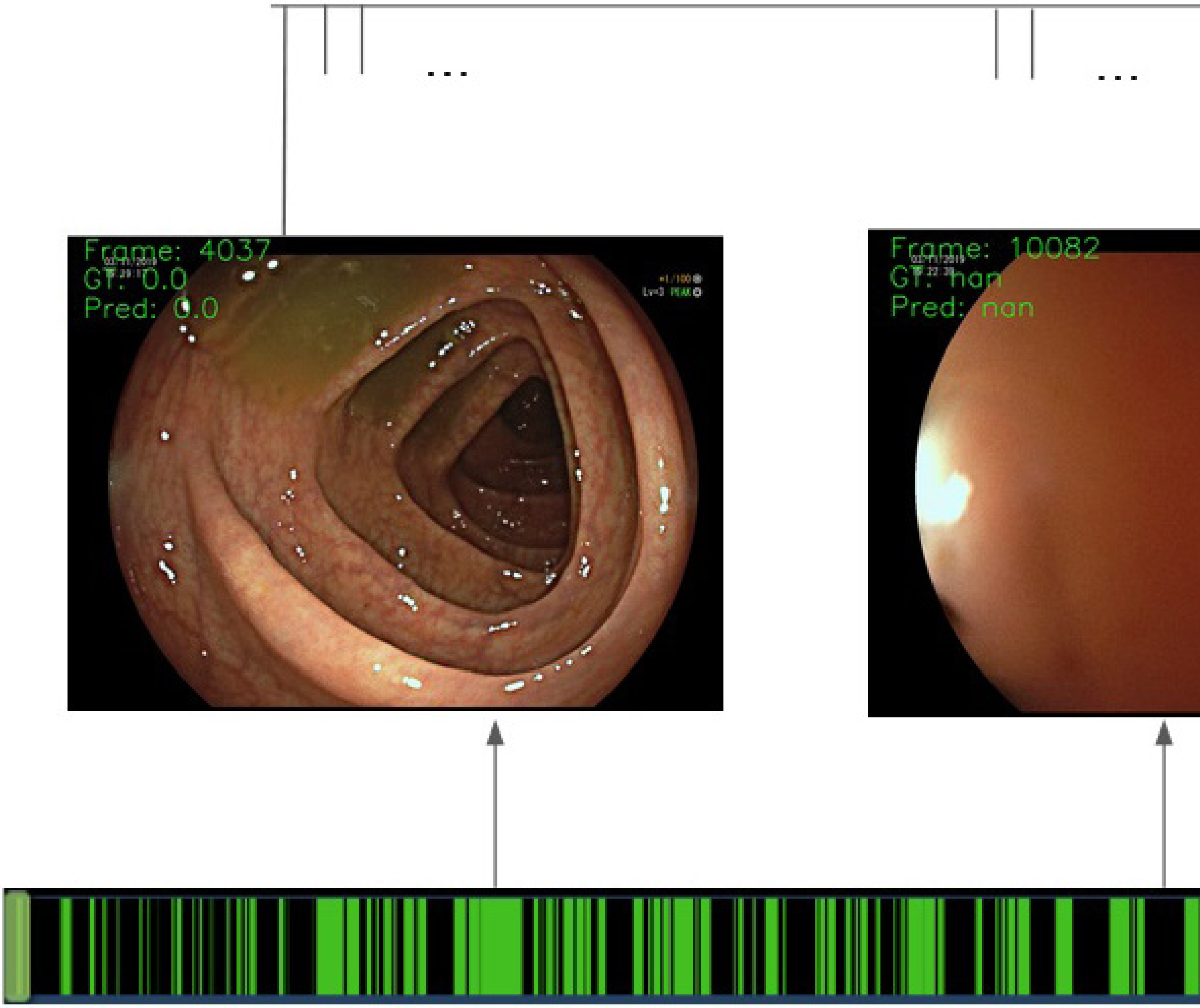
Inflammatory bowel diseases (IBD) such as Ulcerative Colitis (UC) and Crohn's disease are mostly diagnosed by performing a colonoscopy. The level of disease is assessed by visual findings such as ulcers, erosions, and erythema, which are graded using visual scoring approaches. However, IBD scoring is a complex and time-consuming task, and reader assessments are often subjective, based on their training, experience and individual interpretation of the colonoscopy. The new AI-based system, developed at RSIP Vision under the scientific leadership of Clario and the clinical guidance from GI Reviewers, automatically computes the Mayo Endoscopic Score (MES) from colonoscopy videos in Ulcerative Colitis patients. The system was trained on hundreds of videos from multiple sites and devices, containing >100,000 images, and all scored under expert gastroenterologist supervision. Algorithm performance, as assessed by the Kappa metric, was found to be comparable to the performance of human reviewers. Building on the successes in UC, the system is being extended to support Crohn's Disease scoring. This technology will allow generation of rich and detailed data from IBD videos, and increase the automation of clinical trial enrollment procedures.

"Central reading has become common in gastroenterology clinical trials due to the need for an unbiased, independent review of images ensuring the most reliable outcome data for assessing treatment efficacy," **said Dr. Marcela Vieira, Medical Director of Gastroenterology at Clario.** *"Our new AI-based solution will reduce known limitations of central reading, such as fatigue and inter-reader variability, and improve the reproducibility and speed of the readings. Using AI to support our reviewers is improving our capabilities across a range of clinical trial use cases, and will continue to be a game changer for medical research."*

Endoscopic videos such as colonoscopies present steep challenges for AI-based processing. Implementing this successful solution requires merging state-of-the-art neural network architectures with strong, robust, and ongoing validation processes, integrating the work of the clinical team as an integral part of the algorithm development.

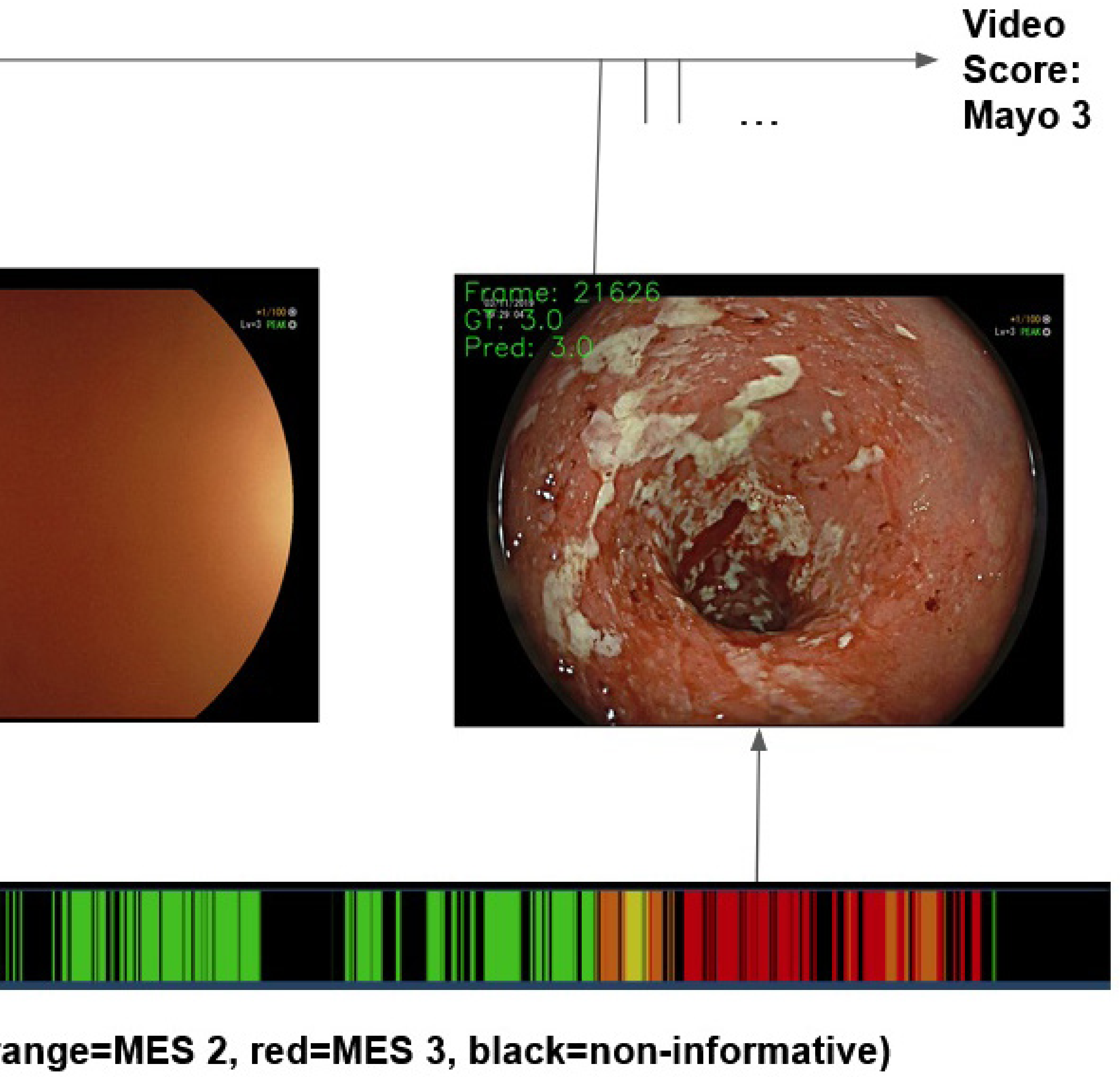
"In addition to the Mayo score, our team trained the AI algorithm to detect which parts of the video lack medical interest and indicate this to the user, saving time and increasing efficiency," **said Moshe Safran, U.S. CEO at RSIP Vision.** *"The accuracy of the algorithm has been dramatically improved over the past year, and the performance of the AI has reached a level similar to that of a human expert,"* Safran said.

"Colonoscopy scoring is a complex task, requiring years of training and experience to have the confidence, especially in more complex cases," **said Dr. Daniel Mishkin, CEO & Medical Director at GI Reviewers.** *"It has been a fascinating journey to see how this*



Per-frame scores (green=healthy, yellow=MES 1, or *knowledge can be embedded in an automated algorithm that will help bring benefit to the patients of the future,"* Mishkin said.

Clario, GI Reviewers, and RSIP Vision continue to collaborate to bring the benefits of AI to all aspects of IBD clinical trials. Upcoming plans include the expansion of the study to additional sites, AI-enabled extraction of rich and detailed video-based data to support deeper analytics by Pharma companies, expansion to Crohn's disease scoring, AI-based automation of clinical trial enrollment, and ultimately the potential development of novel, quantitative, AI-enabled IBD scoring systems.



Walter Pinaya

MONAI+ GENERATIVE MODELS

By Marica Muffoletto (twitter)

Welcome to a new article fully dedicated to tools for medical imaging!

Today we will be talking about the new extension of **MONAI**, an extraordinary platform that we already reviewed more than 2 years ago and that is continuing to grow, thanks to a shared and collaborative effort and to a research interest lead by facilitating use of deep learning methods in medical practice.

The new extension, available **here**, is a prototyping repository for generative models, and since the writer is a big fan of those, I jumped at the opportunity to try it out and talk about it here. In a previous article about MONAI, our tutorial was using **GANs**, but the Generative Models extension introduces new state-of-the-art architectures such as Diffusion Model, Autoencoder-KL, VQ-VAE, Autoregressive transformers, (Multi-scale) Patch-GAN discriminator. Moreover, it offers a wide choice of losses: Adversarial, Spectral, and Perceptual losses, and different Diffusion Model Noise Schedulers: DDPM, DDIM, and PNDM. You can use it for applications such as Anomaly Detection, Inpainting and Super-resolution challenges.

Before writing this article, we spoke to Walter Pinaya, one of the main contributors of this distribution, and a great colleague at KCL. He highlighted that the strength of **MONAI Generative Models** lies in the implementation and support for diffusion models in 3D medical imaging, a special feature from the team. Inspired from him, we chose to show 3 applications, two of which use the denoising diffusion probabilistic model to generate 3D synthetic images from the **Brain tumour Decathlon dataset** and for inpainting on the **2D MedNIST dataset**. The other application uses **VQVAE for 3D reconstruction**. Enjoy!

Google Colab setup and imports

```
from google.colab import drive
drive.mount('/content/gdrive')
!python -c "import monai" || pip install -q "monai-weekly[tqdm]"
!python -c "import matplotlib" || pip install -q matplotlib
%matplotlib inline
!git clone https://github.com/Project-MONAI/GenerativeModels.git
cd GenerativeModels

!python setup.py install

import os
import shutil
import tempfile
import time

import matplotlib.pyplot as plt
import numpy as np
import torch
import torch.nn.functional as F
from monai import transforms
from monai.config import print_config
from monai.data import CacheDataset, DataLoader
from monai.utils import first, set_determinism
from torch.cuda.amp import GradScaler, autocast
from tqdm import tqdm

directory = os.environ.get("MONAI_DATA_DIRECTORY")
root_dir = tempfile.mkdtemp() if directory is None else directory
print(root_dir)
set_determinism(42)

device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
print(f"Using {device}")
```

Download Decathlon dataset (Brain Tumour)

```
from monai.apps import DecathlonDataset
    [
        transforms.LoadImaged(keys=["image"]),
        transforms.Lambdad(keys="image", func=lambda x: x[:, :, :, 1]),
        transforms.AddChanneld(keys=["image"]),
        transforms.ScaleIntensityd(keys=["image"]),
        transforms.CenterSpatialCropd(keys=["image"], roi_size=[176, 224, 
155]),
        transforms.Resized(keys=["image"], spatial_size=(32, 48, 32)),
    ]
)

train_ds = DecathlonDataset(
    root_dir=root_dir, task="Task01_BrainTumour", transform=data_trans-
form, section="training", download=True
)
```

```
train_loader = DataLoader(train_ds, batch_size=16, shuffle=True, num_work-
ers=8)

val_ds = DecathlonDataset(
    root_dir=root_dir, task="Task01_BrainTumour", transform= data_trans-
form, section="validation", download=True
)
val_loader = DataLoader(val_ds, batch_size=16, shuffle=False, num_work-
ers=8, persistent_workers=True)
```

Download MedNIST Validation only dataset

```
from monai.apps import MedNISTDataset

val_data = MedNISTDataset(root_dir=root_dir, section="validation", down-
load=True, progress=False, seed=0)
val_datalist = [{"image": item["image"]} for item in val_data.data if
item["class_name"] == "Hand"]
```

GENERATION OF SYNTHETIC IMAGES WITH DENOISING DIFFUSION PROBABILISTIC MODEL (DDPM)

```
Note: You will probably need to lower the batch size before running this
experiment!

from generative.inferers import DiffusionInferer
from generative.networks.nets import DiffusionModelUNet
from generative.networks.schedulers import DDPMScheduler
```

Model: Diffusion Model U-Net (3D)

```
model = DiffusionModelUNet(
    spatial_dims=3,
    in_channels=1,
    out_channels=1,
    num_channels=[256, 256, 512],
    attention_levels=[False, False, True],
    num_head_channels=[0, 0, 512],
    num_res_blocks=2,
)
model.to(device)
```

Parameters: Adam Optimiser and DDPM Scheduler

```
scheduler = DDPMScheduler(num_train_timesteps=1000, beta_schedule="scaled_
linear", beta_start=0.0005, beta_end=0.0195)

inferer = DiffusionInferer(scheduler)
optimizer = torch.optim.Adam(params=model.parameters(), lr=5e-5)
```

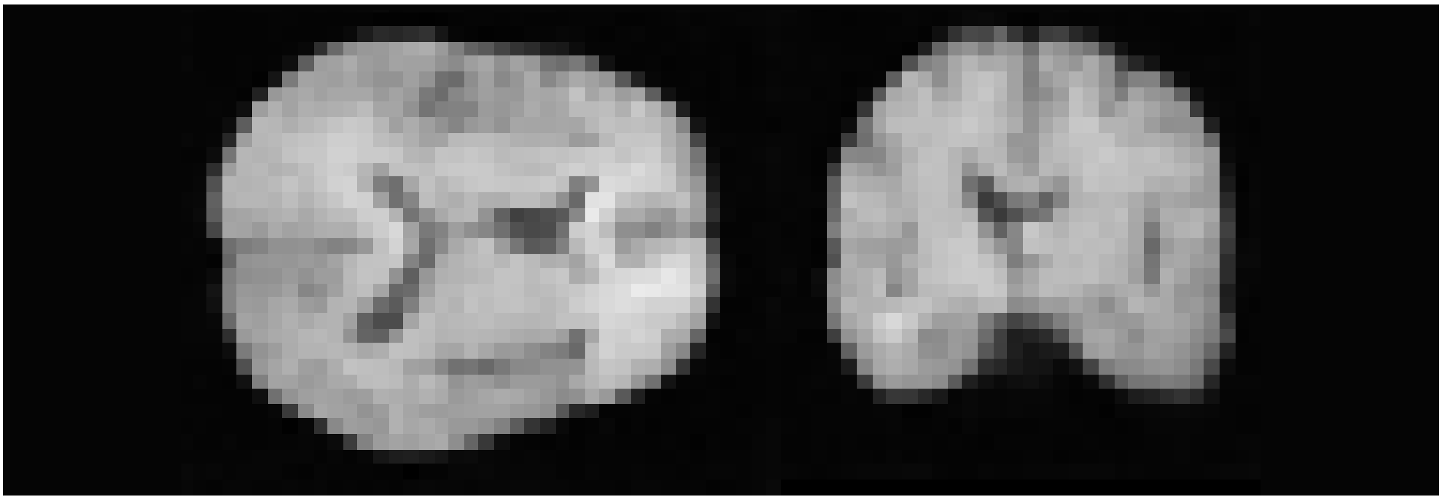
Training

```
n_epochs = 150
val_interval = 25
```

```
epoch_loss_list = []
val_epoch_loss_list = []

scaler = GradScaler()
total_start = time.time()
for epoch in range(n_epochs):
    model.train()
    epoch_loss = 0
    progress_bar = tqdm(enumerate(train_loader), total=len(train_loader),
ncols=70)
    progress_bar.set_description(f"Epoch {epoch}")
    for step, batch in progress_bar:
        images = batch["image"].to(device)
        optimizer.zero_grad(set_to_none=True)
        with autocast(enabled=True):
            # Generate random noise
            noise = torch.randn_like(images).to(device)

            # Create timesteps
            timesteps = torch.randint(
                0, inferer.scheduler.num_train_timesteps, (images.
shape[0],), device=images.device
            ).long()

            # Get model prediction
            noise_pred = inferer(inputs=images, diffusion_model=model,
noise=noise, timesteps=timesteps)

            loss = F.mse_loss(noise_pred.float(), noise.float())

        scaler.scale(loss).backward()
        scaler.step(optimizer)
        scaler.update()

        epoch_loss += loss.item()

        progress_bar.set_postfix({"loss": epoch_loss / (step + 1)})
    epoch_loss_list.append(epoch_loss / (step + 1))

    if (epoch + 1) % val_interval == 0:
        model.eval()
        val_epoch_loss = 0
        for step, batch in enumerate(val_loader):
            images = batch["image"].to(device)
            noise = torch.randn_like(images).to(device)
            with torch.no_grad():
                with autocast(enabled=True):
                    timesteps = torch.randint(
                        0, inferer.scheduler.num_train_timesteps, (images.
shape[0],), device=images.device
                    ).long()

                    # Get model prediction
                    noise_pred = inferer(inputs=images, diffusion_model=-
model, noise=noise, timesteps=timesteps)
```

```
                    val_loss = F.mse_loss(noise_pred.float(), noise.float())

                val_epoch_loss += val_loss.item()
                progress_bar.set_postfix({"val_loss": val_epoch_loss / (step +
1)})
            val_epoch_loss_list.append(val_epoch_loss / (step + 1))
```

After training, we tried to sample brain images using as starting point an image containing just noise and this is the result:



INPAINTING (USING PRE-TRAINED MODEL)

```
from generative.inferers import DiffusionInferer
from generative.networks.nets import DiffusionModelUNet
from generative.networks.schedulers import DDPMScheduler
```

Loading MEDNIST Dataset

```
val_transforms = transforms.Compose(
    [
        transforms.LoadImaged(keys=["image"]),
        transforms.EnsureChannelFirstd(keys=["image"]),
        transforms.ScaleIntensityRanged(keys=["image"], a_min=0.0, a_
max=255.0, b_min=0.0, b_max=1.0, clip=True),
    ]
)
val_ds = CacheDataset(data=val_datalist, transform=val_transforms)
val_loader = DataLoader(val_ds, batch_size=128, shuffle=False, num_work-
ers=4, persistent_workers=True)
```

Model: Diffusion Model U-Net (2D)

```
model = DiffusionModelUNet(
    spatial_dims=2,
    in_channels=1,
    out_channels=1,
    num_channels=(128, 256, 256),
    attention_levels=(False, True, True),
```

```
    num_res_blocks=1,
    num_head_channels=256,
)
model.to(device)
```

Parameters: DDPMScheduler and Adam Optimiser

```
scheduler = DDPMScheduler(num_train_timesteps=1000)
optimizer = torch.optim.Adam(params=model.parameters(), lr=2.5e-5)
inferer = DiffusionInferer(scheduler)
```

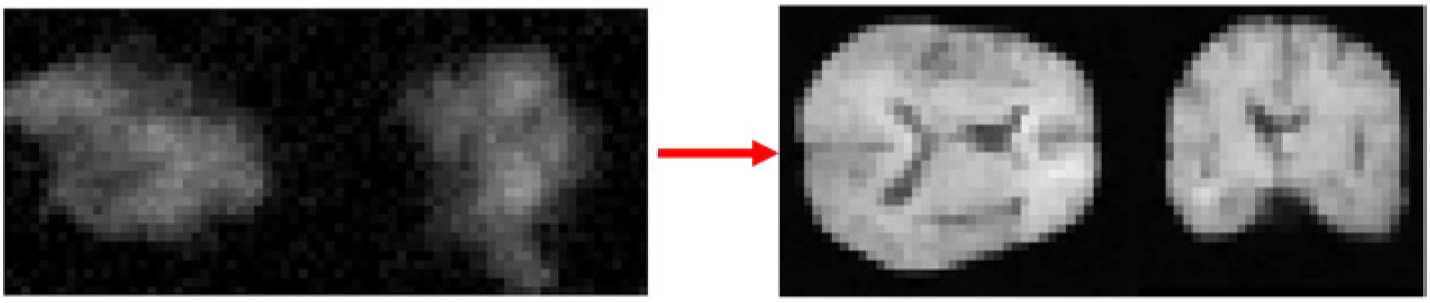
Loading pre-trained model

```
model = torch.hub.load("marksgraham/pretrained_generative_models:v0.2",
model="ddpm_2d", verbose=True).to(device)

model.eval()
mask = mask.to(device)
val_image_masked = val_image_masked.to(device)
#timesteps = torch.Tensor((999,)).to(noise.device).long()
val_image_inpainted = torch.randn((1, 1, 64, 64)).to(device)

scheduler.set_timesteps(num_inference_steps=1000)
progress_bar = tqdm(scheduler.timesteps)
```

Inferencing on validation cases

```
num_resample_steps = 4
with torch.no_grad():
    with autocast(enabled=True):
        for t in progress_bar:
            for u in range(num_resample_steps):
                # get the known portion at t-1
                if t > 0:
                    noise = torch.randn((1, 1, 64, 64)).to(device)
                    timesteps_prev = torch.Tensor((t - 1,)).to(noise.de-
vice).long()
                    val_image_inpainted_prev_known = scheduler.add_noise(
                        original_samples=val_image_masked, noise=noise,
timesteps=timesteps_prev
                    )
                else:
                    val_image_inpainted_prev_known = val_image_masked

                # perform a denoising step to get the unknown portion at
t-1
                if t > 0:
                    timesteps = torch.Tensor((t,)).to(noise.device).long()
                    model_output = model(val_image_inpainted, timesteps=-
timesteps)
                    val_image_inpainted_prev_unknown, _ = scheduler.step(-
model_output, t, val_image_inpainted)
```

```
                # combine known and unknown using the mask
                val_image_inpainted = torch.where(
                    mask == 1, val_image_inpainted_prev_known, val_image_
inpainted_prev_unknown
                )
                # perform resampling
                if t > 0 and u < (num_resample_steps - 1):
                    # sample x_t from x_t-1
                    noise = torch.randn((1, 1, 64, 64)).to(device)
                    val_image_inpainted = (
                        torch.sqrt(1 - scheduler.betas[t - 1]) * val_im-
age_inpainted
                        + torch.sqrt(scheduler.betas[t - 1]) * noise
```



3D RECONSTRUCTION WITH VQVAE

```
from generative.networks.nets import VQVAE
from torch.nn import L1Loss
from monai.apps import DecathlonDataset
```

Model: VQVAE

```
model = VQVAE(
    spatial_dims=3,
    in_channels=1,
    out_channels=1,
    num_channels=(256, 256),
    num_res_channels=256,
    num_res_layers=2,
    downsample_parameters=((2, 4, 1, 1), (2, 4, 1, 1)),
    upsample_parameters=((2, 4, 1, 1, 0), (2, 4, 1, 1, 0)),
    num_embeddings=256,
    embedding_dim=32,
)
model.to(device)
```

Parameters: Adam Optimiser and L1 Loss

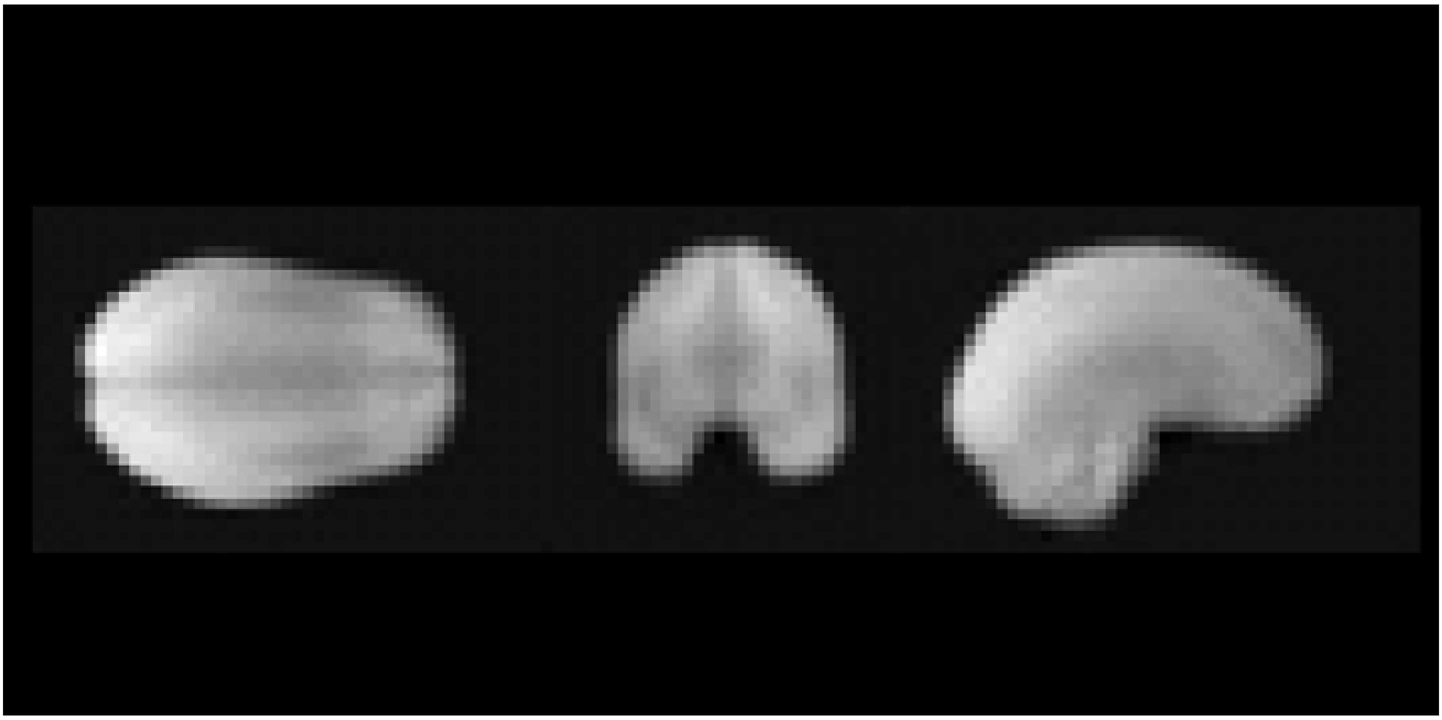
```
optimizer = torch.optim.Adam(params=model.parameters(), lr=1e-4)
l1_loss = L1Loss()
```

Training

```
n_epochs = 100
val_interval = 10
epoch_recon_loss_list = []
epoch_quant_loss_list = []
val_recon_epoch_loss_list = []
intermediary_images = []
n_example_images = 4

total_start = time.time()
for epoch in range(n_epochs):
    model.train()
    epoch_loss = 0
    progress_bar = tqdm(enumerate(train_loader), total=len(train_loader),
ncols=110)
    progress_bar.set_description(f"Epoch {epoch}")
    for step, batch in progress_bar:
        images = batch["image"].to(device)
        optimizer.zero_grad(set_to_none=True)

        # model outputs reconstruction and the quantization error
        reconstruction, quantization_loss = model(images=images)

        recons_loss = l1_loss(reconstruction.float(), images.float())

        loss = recons_loss + quantization_loss

        loss.backward()
        optimizer.step()

        epoch_loss += recons_loss.item()

        progress_bar.set_postfix(
            {"recons_loss": epoch_loss / (step + 1), "quantization_loss":
quantization_loss.item() / (step + 1)}
        )
    epoch_recon_loss_list.append(epoch_loss / (step + 1))
    epoch_quant_loss_list.append(quantization_loss.item() / (step + 1))

    if (epoch + 1) % val_interval == 0:
        model.eval()
        val_loss = 0
        with torch.no_grad():
            for val_step, batch in enumerate(val_loader, start=1):
                images = batch["image"].to(device)

                reconstruction, quantization_loss = model(images=images)

                # get the first sample from the first validation batch for
                # visualizing how the training evolves
                if val_step == 1:
                    intermediary_images.append(reconstruction[:n_example_
images, 0])
```

```
                recons_loss = l1_loss(reconstruction.float(), images.
float())

                val_loss += recons_loss.item()

        val_loss /= val_step
        val_recon_epoch_loss_list.append(val_loss)

total_time = time.time() - total_start
print(f"train completed, total time: {total_time}.")
```

To check the training, you can plot the reconstructed images at every x validation steps (here set to 10), e.g.



To write this article, we used the tutorials found on **Github**, an easily approachable way to get into MONAI Generative Models. For the second experiment on Inpainting, we used a pre-trained model offered by the creators and available **here**. This is an additional asset of this tool, which will soon offer models (in the MONAI model zoo platform) pre-trained on big datasets ready for fine-tuning. For now, **at this link**, you will find the weights from two pre-trained models for brain images and chest X-Ray generator.

In our previous article on MONAI, we also generated synthetic images of hands, using GANs. **Check it out** to understand what amazing progress the MONAI team has achieved in just a bit more than 2 years!

Miguel Bernabeu is the Deputy Director and Director of Research at The Bayes Centre, the University of Edinburgh's innovation hub for data science and artificial intelligence. He is also Professor of Computational Medicine at the Usher Institute within the Edinburgh Medical School, where he established his first research group in the Centre for Medical Informatics. He speaks to us about its work in the field of computer vision for medical imaging.

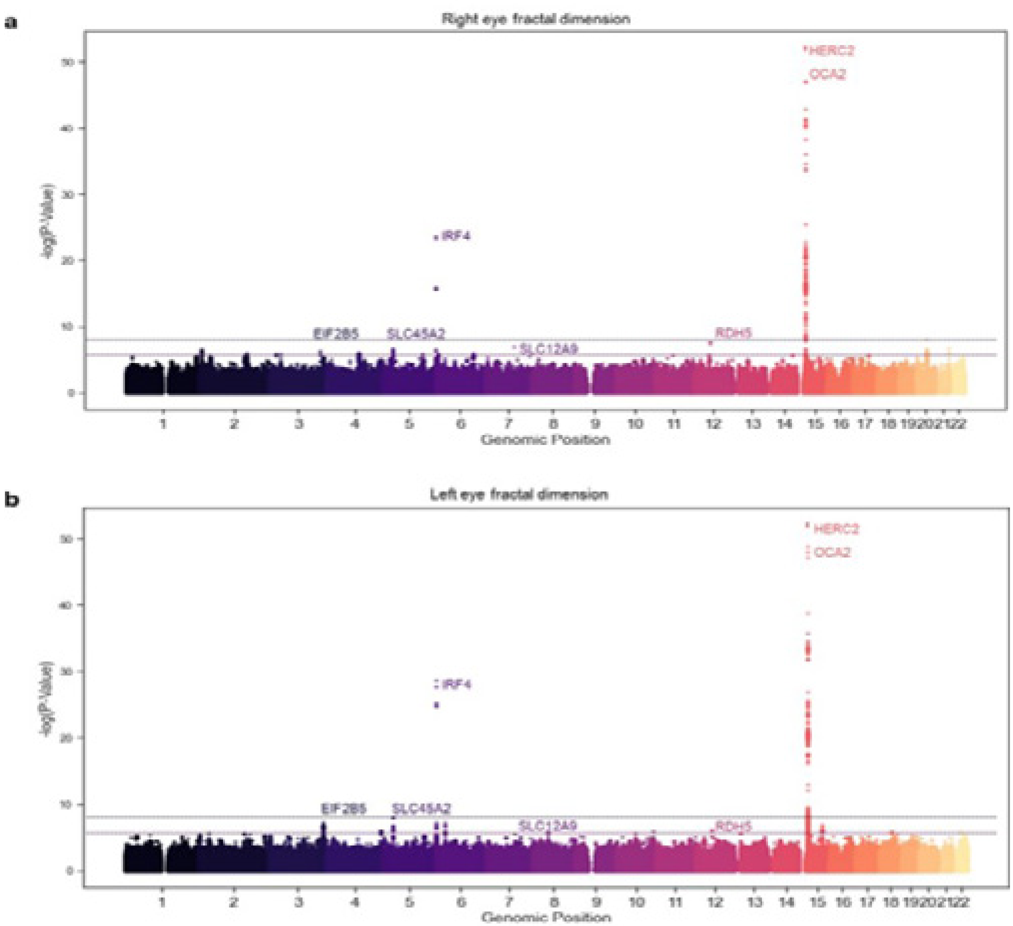
Artificial intelligence has made significant advancements in recent years, and its potential to transform healthcare has become increasingly apparent. With a mix of largely PhD students and postdocs, the **Computational Medicine Group at the University of Edinburgh** works across every aspect of developing **AI systems for medical applications**, from early-stage exploratory work to real-world evaluation.

"We work on interesting projects that go all the way from gathering data and creating models to implementing them in practice and evaluating them," Miguel tells us. *"I don't think many labs cover this whole pipeline of activity."*

The group's research includes using retinal and brain imaging to identify biomarkers for disease diagnosis and prediction, generating novel insights and approaches that will hopefully be translated into clinical practice further down the line.

"To support that, we do a lot of work on gathering large datasets," he continues. *"Before any model can be trained, you need data. One of the big challenges in computer vision for medical imaging is that we don't have the humongous datasets that people use to train models in other areas of computer vision. From very early on, we've paid special attention to that."*

There are ongoing efforts to release open-source datasets, including at **MICCAI**, where challenges regularly provide data for people to work on. However, Miguel says the way forward is to develop more advanced methods of accessing highly sensitive information and handling it with the utmost care.



"*[Health Data Research UK](#) allows researchers to access data that was not necessarily consented for research,*" he points out. "*No one has the legal right to release this data out in the open, but they provide sophisticated IT infrastructures for you to do your work in a way that ensures it remains confidential and is treated according to the regulations. I think the future will be in creating these infrastructures for us to deal with as much patient data as is legally allowed.*"

As well as discovery science, the group has activities closer to implementation in practice. Currently, it is working on deploying **computer vision algorithms for interpreting electrocardiograms (ECGs)** in hospitals, supporting cardiologists to identify features in the ECG that indicate someone is presenting with a heart attack.

"*When someone walks through the door at A&E with chest pain, it might be a heart attack or something else,*" Miguel explains. "*ECGs are one tool that gets used for diagnosis, and there tends to be quite a lot of variability between how different cardiologists interpret an ECG. We're using computer vision to extract information from them, and with the help of these algorithms, we can begin to standardize the process.*"

ECG interpretation requires experience, but an expert may not always be available in every hospital at any given time. These algorithms can bridge that gap; furthermore, automated tools can help streamline the work of cardiologists and enable them to see more patients. The group is collaborating with the NHS as part of this project to gain insight into the requirements for integrating its algorithms into clinical trial systems.

The Computational Medicine Group is also working with **Aidence**, a Dutch AI company, on a joint evaluation project called [INPACT](#). Aidence has a commercial computer vision system that detects **lung nodules on chest CTs**, which has already been rolled out in several hospitals across the UK and other countries.

"This is the final piece of the puzzle – you develop the models, put them into practice, and then demonstrate that they're ready," Miguel asserts. *"We're doing a clinical study to understand whether, with the help of these tools, radiologists make more consistent decisions in line with the guidelines they receive."*

Evaluating computer vision systems in the real world is a multifaceted process. One part of the INPACT study is the clinical evaluation, which checks how accurately radiologists interpret an image against the ground truth with or without AI. Another aspect is health economics, demonstrating that the solution can be cost-effective for the health system. Qualitative research is a further aspect, using interviews with hospital clinicians and IT managers to understand the enablers and barriers to adoption and with patients to understand their perspectives.

"This type of work is pioneering because we carry out a holistic interpretation of the problem," Miguel shares. *"It's not only the computer vision task; it's everything around it. Providing a paper reporting on a computer vision task's accuracy is very different from demonstrating that it's improving patients' lives."*

Finally, with such a comprehensive portfolio of work, is there one thing Miguel is most proud of?

*"The project we're getting off the ground on **retinal image data curation** at scale is really innovative,"* he responds. *"It's going to make a real impact on the field!"*

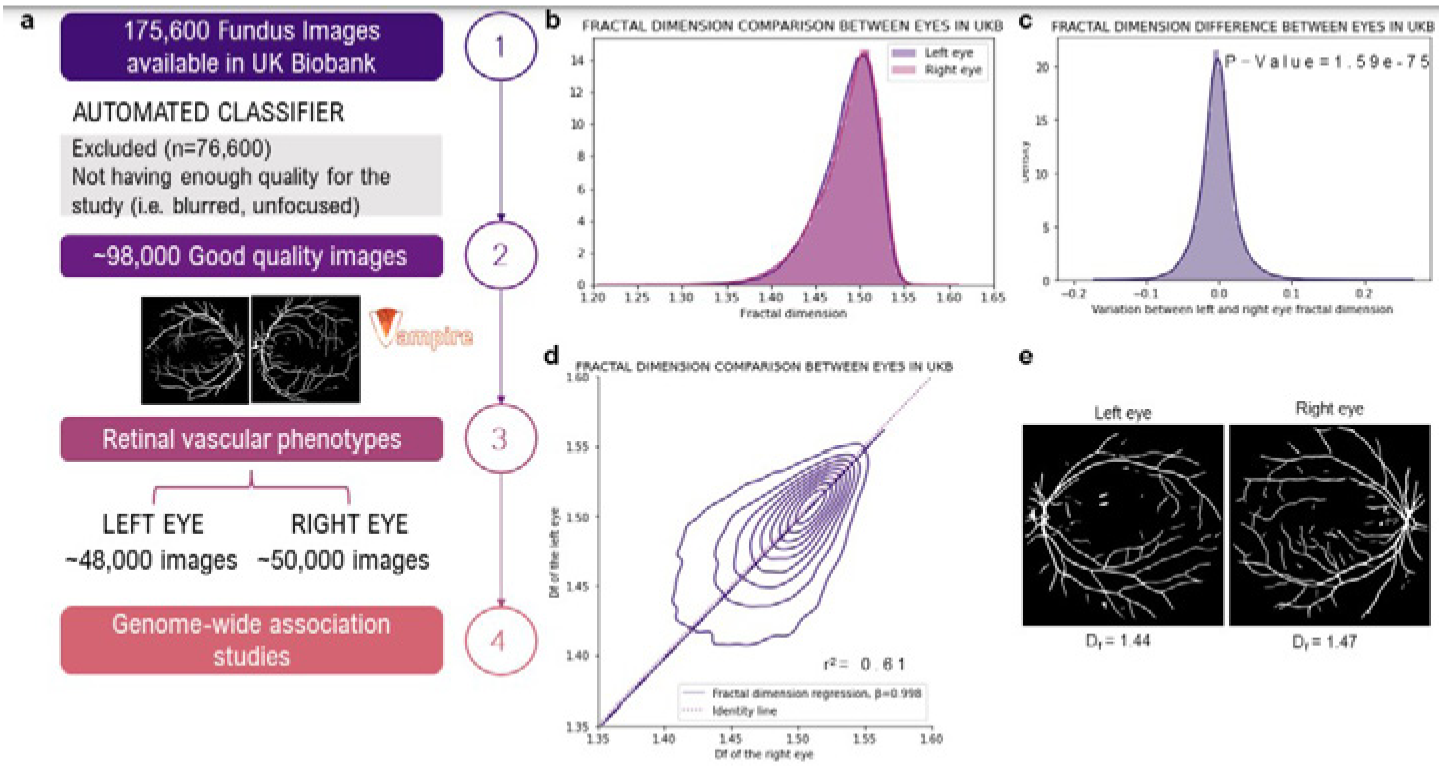


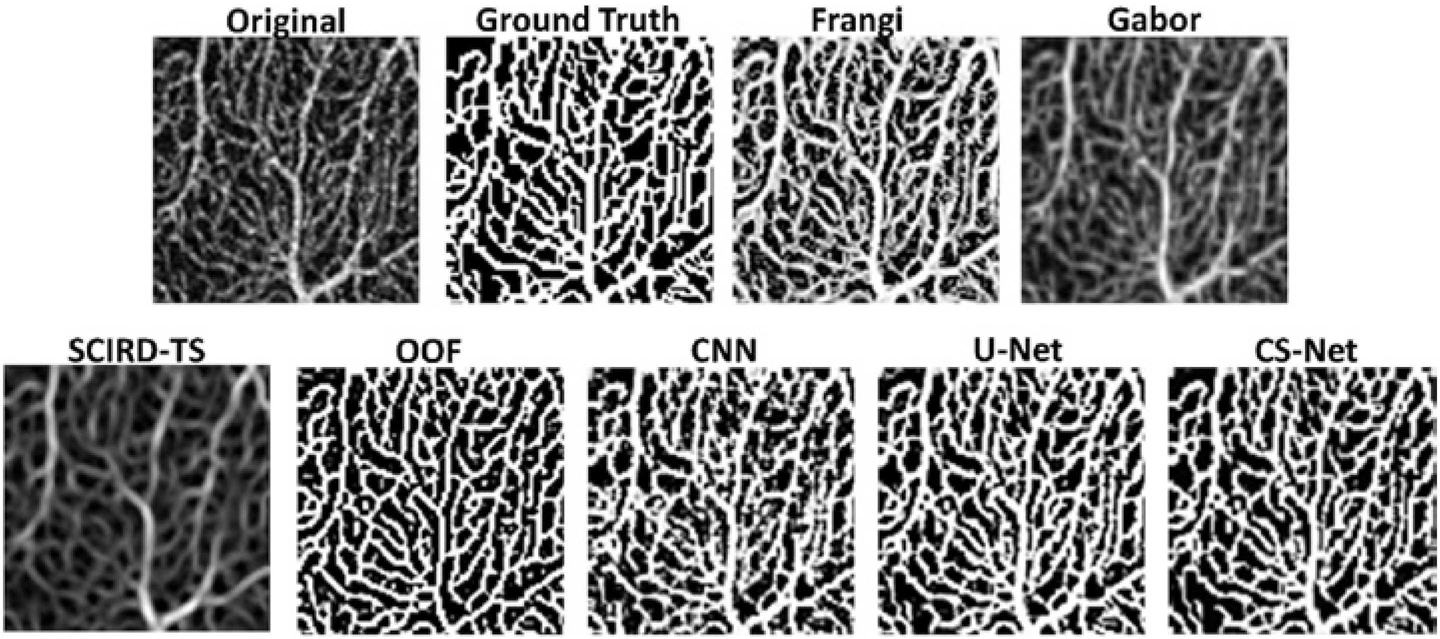
Figure 1. Pipeline and D_f characteristics. a Study design's diagram describing the stepwise development of this project. **b** Left and right D_f histogram. **c** Individual variation distribution between left and right D_f. **d** Overlapping left and right D_f histograms including the regression line. **e** Example of individual interocular asymmetry in UKBB fundus images.

Ylenia Giarratano is a Research Fellow at the University of Edinburgh. She speaks to us about the innovative work Miguel's lab is doing using retinal imaging as a window into the health of other organs.

Retinal imaging is a fascinating field because the eyes are a window into the human body's inner workings, allowing us to observe the vascular system in vivo. This imaging technique is not restricted to studying ocular diseases but can also be employed **to investigate other medical conditions with a vascular component**.

"In my work, I use one of the most recent retinal imaging technologies: optical coherence tomography angiography (OCTA)," Ylenia tells us. *"This device allows us to visualize the tiniest vessels in our eyes, so it provides images at a microscopic level, going well beyond the common standard in retinal imaging tools."*

Ylenia has been developing a new methodology for analyzing the vasculature captured in OCTA images. She is building a pipeline in which **machine learning tools pre-process the image, including a U-Net architecture for segmenting the blood vessels**, some concepts borrowed from graph theory to model the vasculature, and then implementing a broad spectrum



of structural and functional features to characterize the entire retinal landscape.

"The vasculature already has this network structure, so we're using the graph theory concept of edges and nodes to reconstruct it," she explains. *"Using this structure, we can then compute features such as bifurcation points, and there could be nodes of degree three, for example, or we can measure the length of one edge, which would be the segment of one vessel between two bifurcation points."*

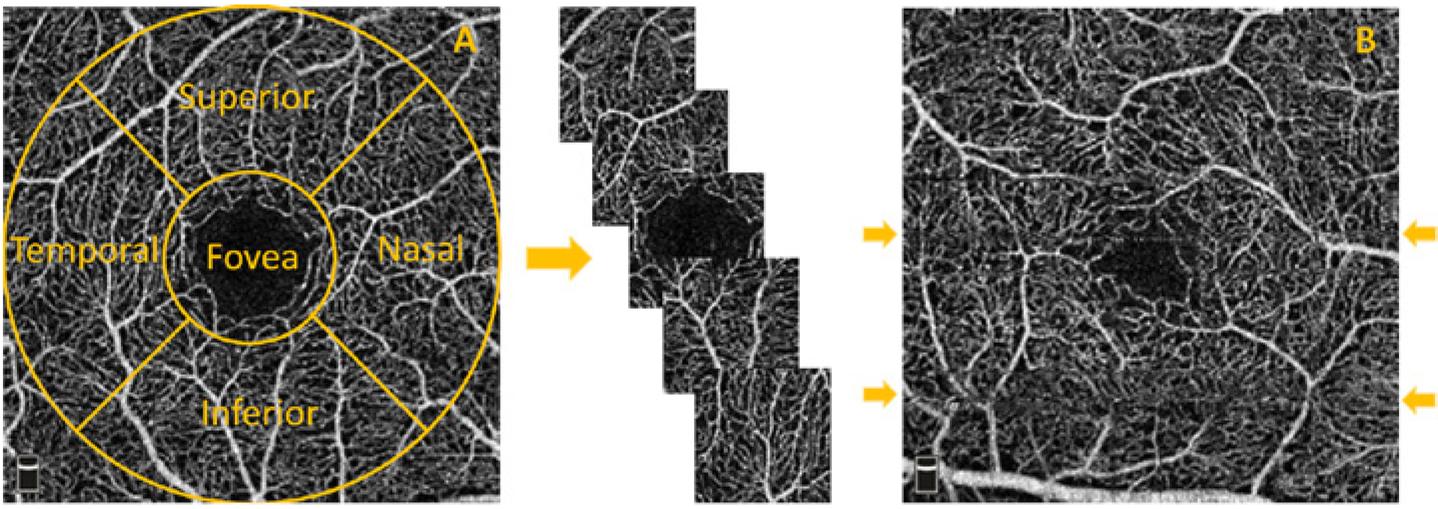
Ylenia tried different methodologies for **segmenting the vasculature**, surveying vessel enhancement methods and neural network architectures, and choosing the **U-Net architecture** because it worked best with her specific device.

"There are different OCTA devices commercially available, so there's not one method that fits them all," she points out. *"Every year at MICCAI, more and more people are interested in the segmentation of the vasculature coming from these OCTA images, so other neural networks are trying to get better at it."*

The device can compute features such as vessel density or the area free from vessels, known as the vascular area in the center of these images. However, changes in the vasculature can be observed in multiple diseases, and these basic features are insufficient for distinguishing between different conditions.

"We're working on developing and implementing more and more metrics characterizing vessel morphology and structure, including tortuosity and caliber, and the size and shape of the intercapillary spaces," Ylenia reveals. *"Also, looking at the topology matrix and the repeated patterns in the network, specifically exploring sub-graphs in the vasculature to see how often we see a similar structure, for example."*

The idea is to use those measurements to discriminate a group of participants and establish a baseline control group for observing changes in specific diseases. The Computational Medicine Group has



been studying diabetes and diabetic retinopathy, chronic kidney disease, and the potential of using retinal biomarkers to assess the suitability of individuals for kidney donation. Recently, Ylenia has been working on a project concerning small vessel disease in the brain.

"The eye is actually part of the central nervous system," she explains. *"It's an outgrowth of the developing brain, and they're physically linked. Changes in the vasculature of the eye may mirror changes in the microvasculature of the brain, so we can exploit retinal imaging to get information on what's happening at the vascular level in the brain."*

The group is also launching a longitudinal pilot study that is acquiring retinal images of pregnant individuals at higher risk of stillbirth to demonstrate the potential utility of retinal biomarkers for the early detection of pregnancy complications.

Having moved to Scotland for a six-month internship with an Erasmus+ scholarship after her master's, Ylenia is still there six years later. We wonder, does it feel like home now, and how is she coping with the British weather?

"Scotland is lovely, and with all this climate change, the weather is getting better!" she laughs. *"I like living here in Edinburgh, but I don't know what the future holds just yet."*

Of all her work, she tells us she is most proud of her paper describing the pipeline for analyzing OCTA images, which was published at **MICCAI** in the **OMIA workshop**, and hopes to submit an extended version soon describing each feature in more detail.

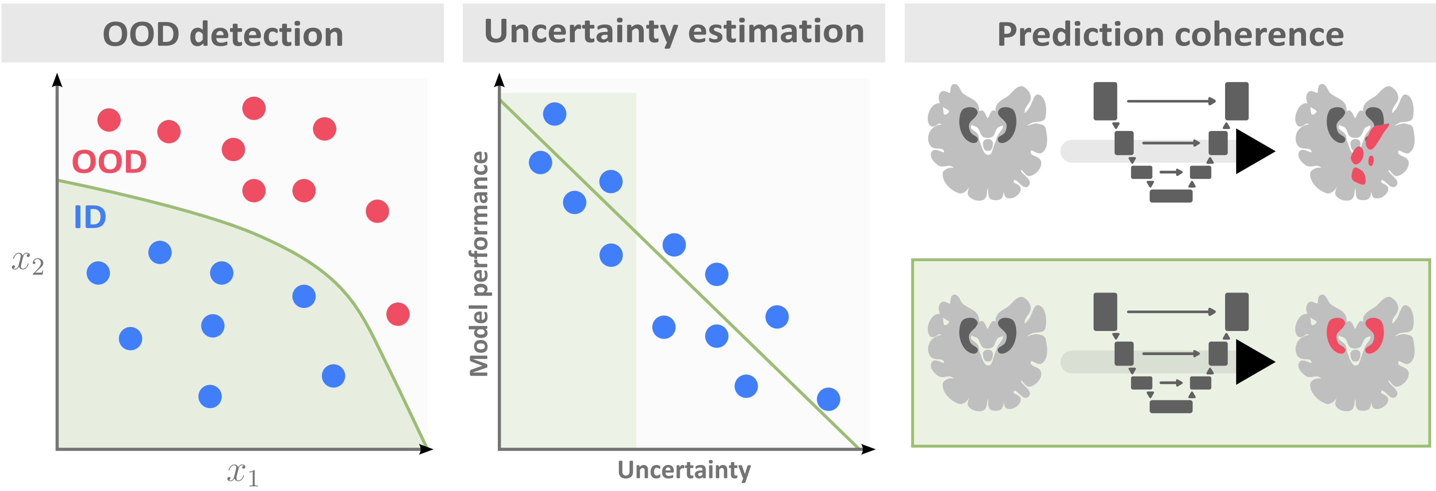
"I think retinal imaging can be a little overlooked in the computer vision world," Ylenia ponders thoughtfully. *"There is so much potential we're not exploiting. We're acquiring OCTA images with these tiny vessels and can do so much with them. With all this new technology, we've got a huge opportunity with retinal imaging!"*

Camila González recently completed her PhD at the Technical University of Darmstadt, Germany. Her research mainly focused on adapting deep learning methods for dynamic clinical environments. She plans to continue her research as a postdoctoral scholar at Stanford University, in the Computational Neuroimage Science lab. Congrats, Doctor Camila!

Despite mounting evidence that data drift causes deep learning models to deteriorate over time, the majority of medical imaging research is developed for - and evaluated on - static close-world environments. There have been exciting advances in the automatic detection and segmentation of diagnostically-relevant findings. Yet the few studies that validate their performance in real clinics are met with disappointing results and little utility as perceived by healthcare professionals. This is largely due to the many factors that introduce shifts in the data distribution, from changes in the acquisition practices to naturally occurring variations in the patient population. If we truly wish to leverage deep learning technologies in the healthcare sector, we must move away from close-world assumptions and start designing systems for the *dynamic clinical world*.

Detecting a deterioration in model performance

This entails, first, the establishment of reliable quality assurance mechanisms with methods from the fields of uncertainty estimation, out-of-distribution detection, and domain-aware prediction appraisal. We propose two approaches that identify outliers by monitoring a self-supervised objective or by quantifying the distance to training samples in a low-dimensional latent space. We also explore how to maximize the diversity among members of a deep ensemble for improved calibration and robustness; and present a lightweight method to detect low-quality masks using domain knowledge.



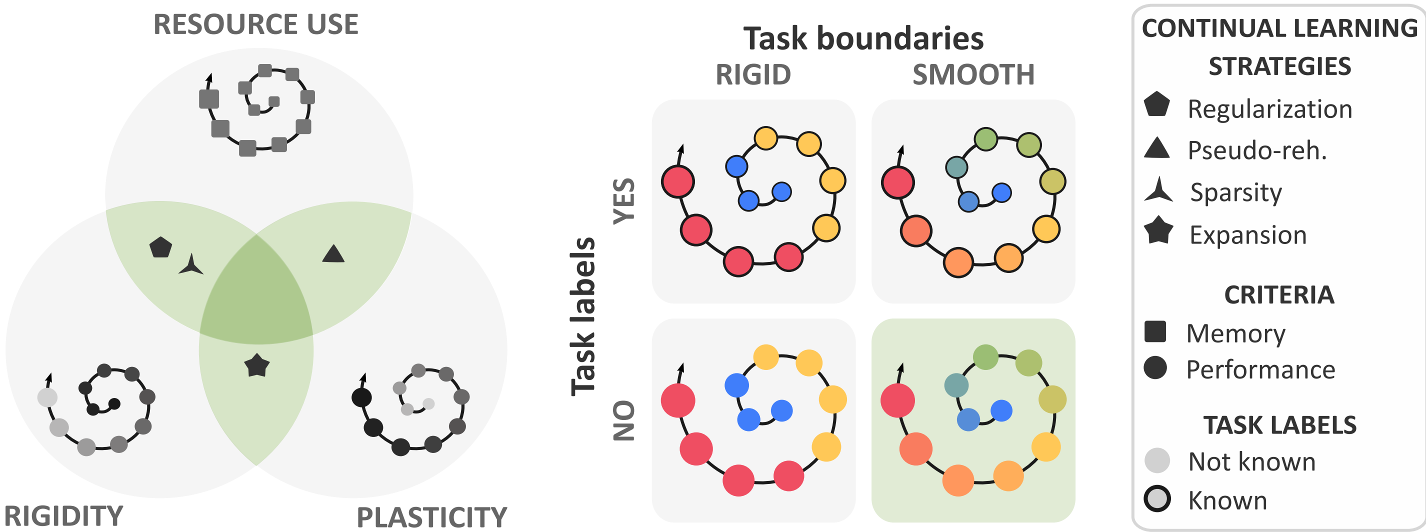
Moving beyond catastrophic forgetting to adapt systems over time

Of course, detecting failures is only the first step. We seek to train models that are reliable for a large portion of the data. Domain generalization and adaptation only increase robustness to a certain extent. As time goes on, models can only maintain acceptable performance if they learn from newly acquired cases. The goal of *continual learning* is to adapt to changes in the environment without forgetting previous knowledge. One practical strategy to approach this relies on *expansion*, whereby multiple parametrizations of the model are maintained and the most appropriate one is selected during inference. We present two expansion-based methods that do not rely on information regarding when the data distribution changes.

Overcoming practical hurdles

Even when appropriate mechanisms are in place to fail safely and accumulate knowledge over time, this will only translate to clinical usage insofar as the regulatory framework allows it. Current regulations in the USA and European Union only authorize *locked* systems that do not learn post-deployment. Fortunately, regulatory bodies are noting the need for a modern *lifecycle regulatory approach*. We review these efforts, along with other practical aspects of developing systems that learn through their lifecycle.

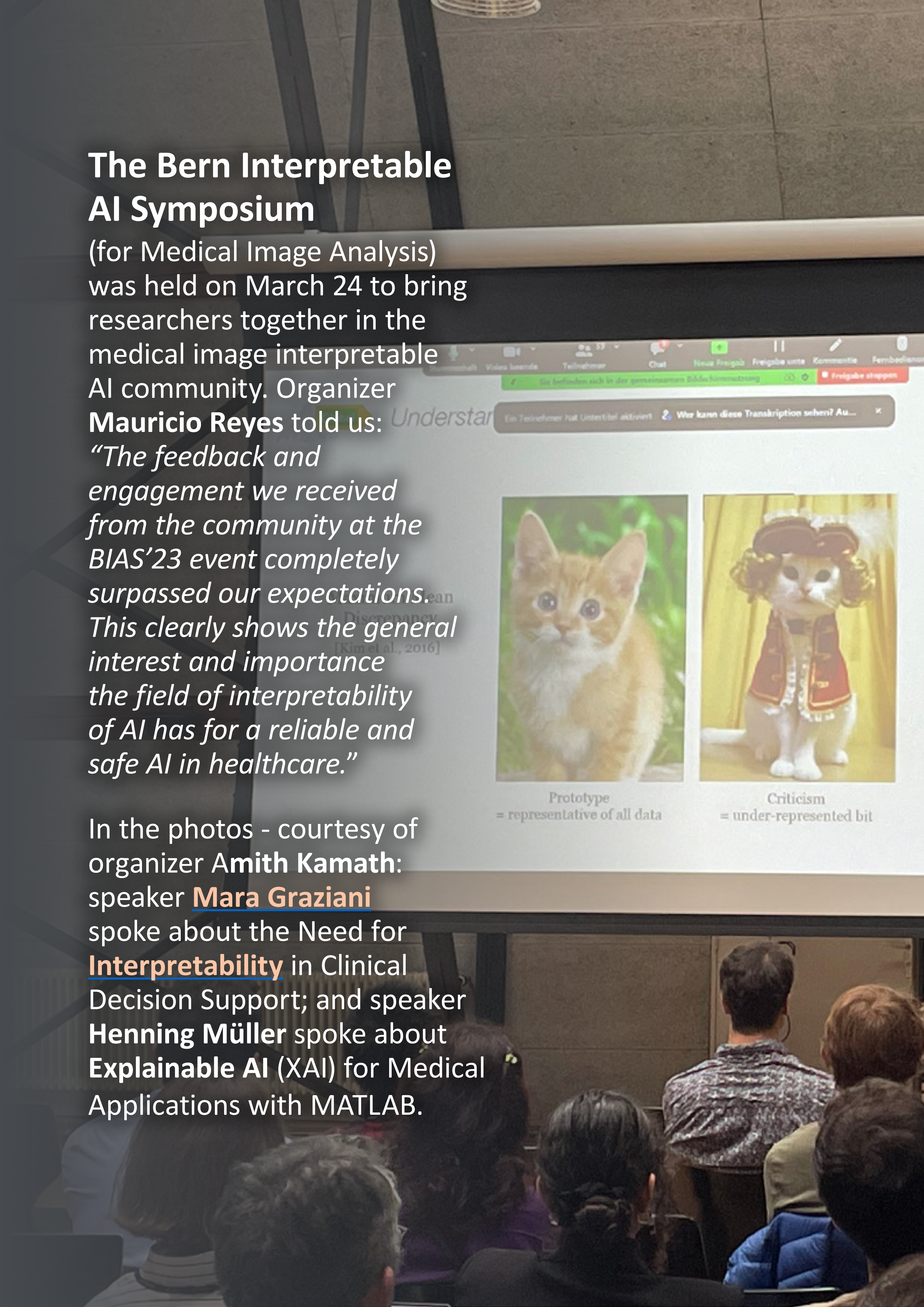
We are finally at a stage where healthcare professionals and regulators embrace deep learning. The number of commercially available diagnostic radiology systems is skyrocketing. This opens up our chance - and responsibility - to show that these systems can be safe and effective throughout their lifespan.



The Bern Interpretable AI Symposium

(for Medical Image Analysis) was held on March 24 to bring researchers together in the medical image interpretable AI community. Organizer **Mauricio Reyes** told us: *“The feedback and engagement we received from the community at the BIAS’23 event completely surpassed our expectations. This clearly shows the general interest and importance the field of interpretability of AI has for a reliable and safe AI in healthcare.”*

In the photos - courtesy of organizer A**mith Kamath**: speaker **Mara Graziani** spoke about the Need for **Interpretability** in Clinical Decision Support; and speaker **Henning Müller** spoke about **Explainable AI** (XAI) for Medical Applications with MATLAB.




Mehr
hes
CAIM
Interpretability of Deep Learning
Make decisions understandable and remove the black box image
Make sure that decisions are sound
Explain why things may not be working
In medicine it is particularly important to make sure that results can be explained
– High impact of wrong decisions
– Integrate information from many sources

EDITED-MRS RECONSTRUCTION CHALLENGE

Rodrigo Berto and Hanna Bugler are master's students in Biomedical Engineering at the University of Calgary. They speak to us as co-organizers of an intriguing Magnetic Resonance Spectroscopy challenge at the upcoming ISBI 2023 conference.

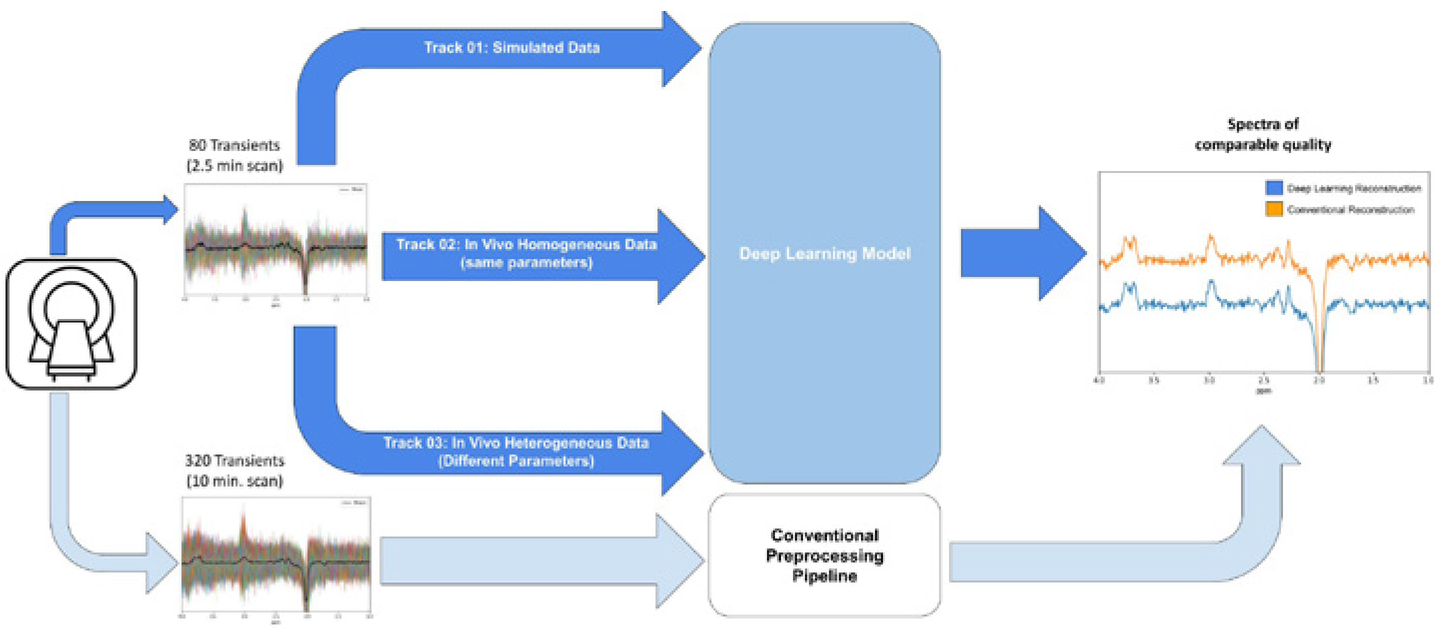
Magnetic resonance spectroscopy (MRS) is an imaging modality similar to magnetic resonance imaging (MRI), but instead of looking at the anatomic structure of the brain, for example, it provides information on **the chemical concentrations of neurochemicals**. However, MRS has yet to reach its full potential due to issues such as a lengthy scan time that can lead to decreased spectrum quality if the subject moves.

"You want the scanner to be available to more people, but if a scan takes a long time, then doctors might not ask for it, and even researchers are going to have a hard time with it," Rodrigo tells us. *"There are many reasons why making it faster is beneficial for both research and clinical applications, as well as the quality of the scan itself."*

Several competitions have been designed to **accelerate MRI**, including one by Rodrigo and Hanna's supervisor, **Roberto Souza**, but until now, there has not been a similar effort to speed up MRS. That is where Rodrigo and Hanna's new challenge comes in. It explores **the edited MRS modality** and aims to solve the problem of **hidden neurochemicals**, which are difficult to study because of a **low signal-to-noise ratio (SNR)**.

"For edited MRS, we measure two spectra and subtract them," Rodrigo explains. *"Each spectrum by itself has a lot of noise compared to the signal, and then when you subtract them,* ***you double the noise and decrease the signal even more!*** *There simply isn't enough signal to work with in the first place."*

Hanna adds: *"There's this paradox: if we make the scan longer, we're getting more data, and we're able to reduce the SNR, but we're increasing the chance of motion artifact. If we reduce it, we're not collecting the signal we need. If we can accelerate it, we're eliminating the higher probability of getting motion on the scan, which would be much better."*



If a patient was tracked, could that diminish the impact of their movement?

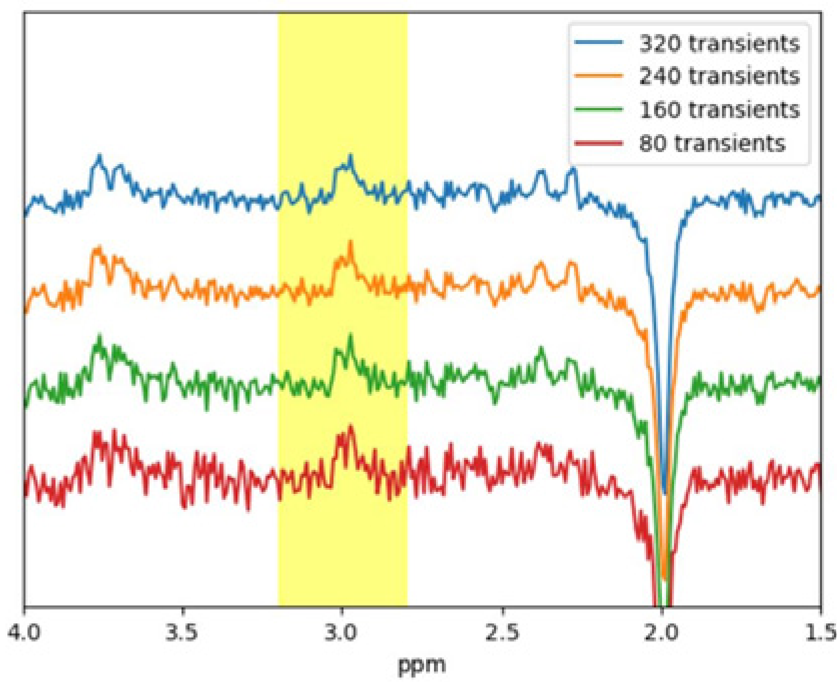
"Prospective motion correction is still in the works," Hanna responds. *"As far as I know, it doesn't exist for edited MRS. Retrospective motion correction algorithms do exist. They're not as good for MRI as for MRS, but even for MRS, they're not great."*

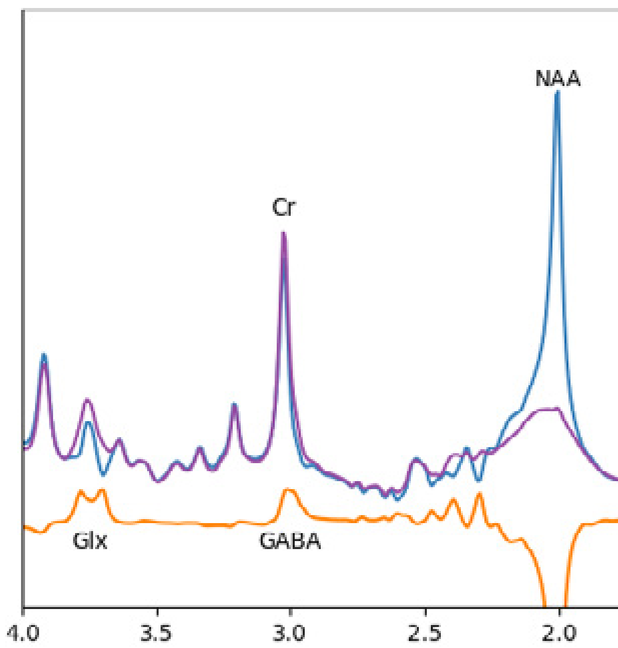
Competitors are asked to develop ways to reconstruct different spectra from data with fewer transients than regular scans. Usually, there would be 320 transients per scan, but they are given 80. The challenge has three tracks; people can participate in one, two, or all three. **The goal is to go from a specific to a more generalized model.**

"The first track is with ***simulated data*** *because there isn't enough MRS data,"* Rodrigo points out. *"A few researchers are working on it, but it's not enough to train machine learning models as we see with MRI, where there's just a lot of data. We provide* ***in vivo data*** *for the second track, but it comes from the same vendor and has similar parameters. Then on the third track, we provide* ***data from three different vendors with different parameters****. The idea is that each track of the challenge is a little bit harder than the previous one."*

The challenge is open to everybody and could be a great entry point for machine learning engineers and data scientists looking to explore medical imaging or people already familiar with MRS but wanting to enter the machine learning field. The team has a fascinating **website** with tutorials and guides to help you get started, and they are happy to answer any questions.

The challenge is Rodrigo and Hanna's first, but they have their supervisors – **Ashley Harris**, who works at the **Alberta Children's Hospital**, and **Roberto Souza**, who co-conceived the **Calgary-Campinas**





public MRI dataset – on hand to advise and guide them through it.

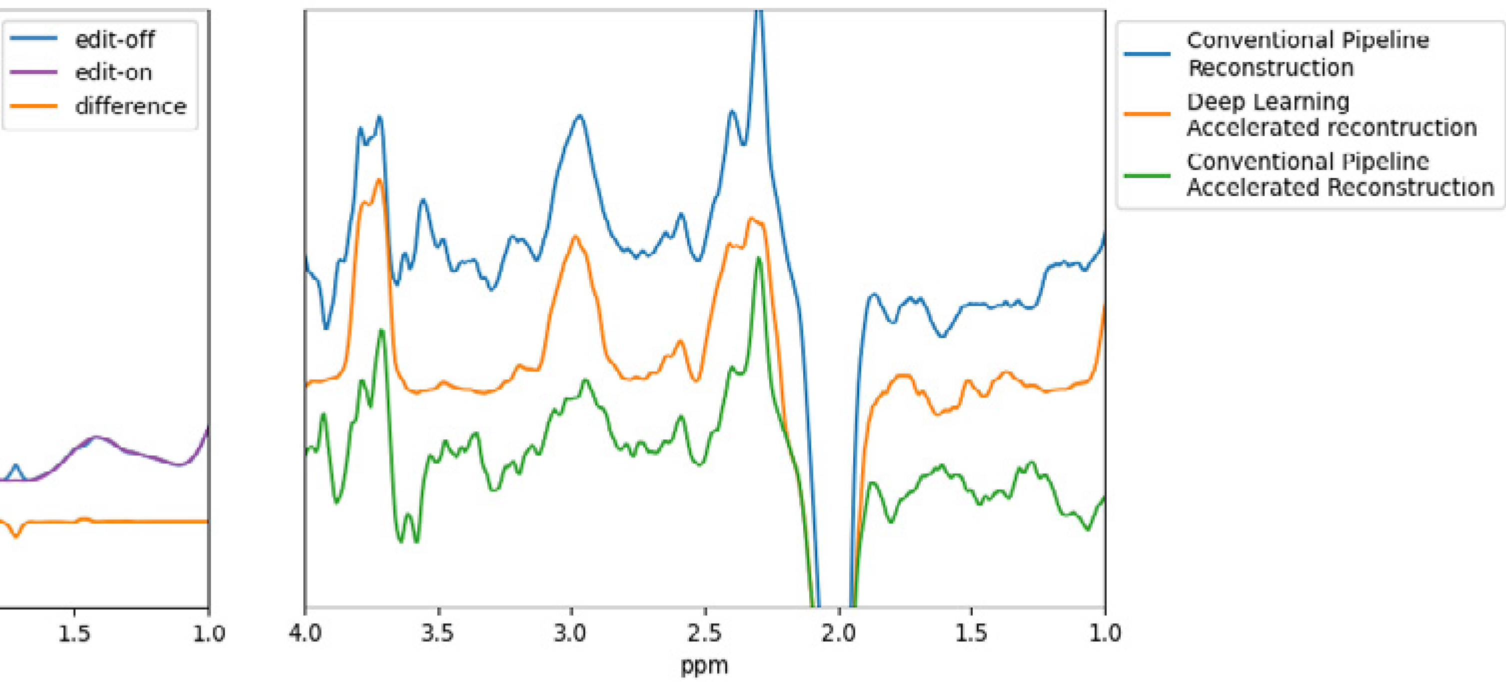
“The University of Calgary is pretty good at MRI and MRS – specifically, their combination,” Rodrigo reveals. *“We have many colleagues who work on projects related to MRI, MRS, and machine learning. It’s a big field here.”*

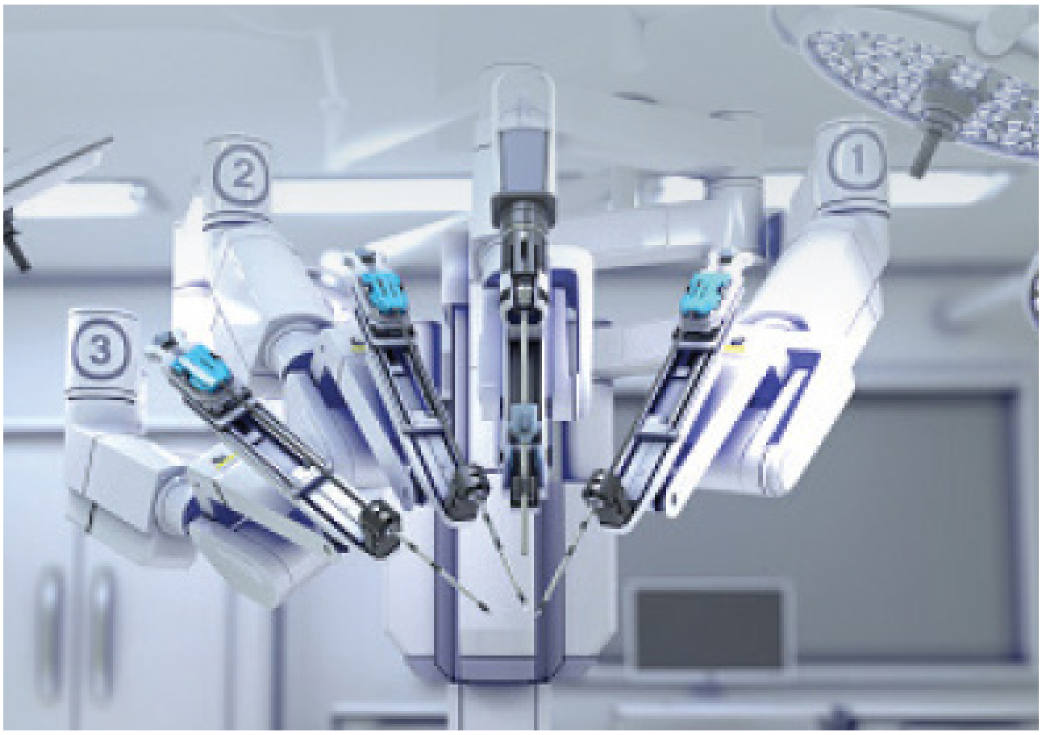
Hanna, currently transferring to a PhD at the university, agrees: *“It’s a **pretty big imaging hub** here, and our supervisor came to us with this great opportunity to propose a challenge at **ISBI**. Many articles on MRS and deep learning have emerged recently, but I was amazed that **I could only find one MRS challenge in the last 10 years**. If you can’t find a challenge, why not start one and drive that innovation forward? I think now is the right time to push that boundary.”*

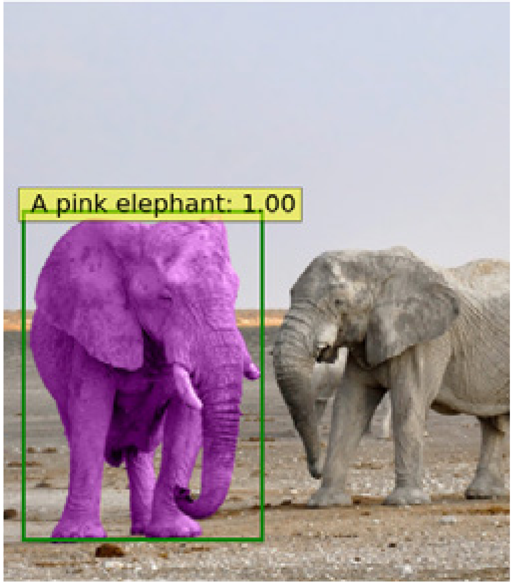
So far, there has been a greater focus on MRI because it has a **more clinical application than edited MRS, which is still stuck in the research setting**. Rodrigo is exploring this problem as a part of his thesis, aiming to break down that clinical barrier.

“I hope this challenge attracts more people from the medical imaging community and brings greater awareness to edited MRS and how it can be improved with deep learning,” he tells us. *“I’d like to establish an initial benchmark on how that improvement could be performed specifically for accelerating MRS. Hopefully, we will find it reaching that clinical and research barrier in the next few years.”*

Hanna concludes: ***“I hope the challenge increases the confidence between science and machine learning**. There’s still a question: Can machine learning improve, and can we trust it? Doing more of these challenges and seeing the outcomes will help improve that confidence that **machine learning is a way to go for certain things!**”*






A pink elephant: 1.00