JUNE 2023

Computer Vision News

& Medical Imaging News

The Magazine of the Algorithm Community

Katrin Binner

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

Computer Vision News

Editor:
Ralph Anzarouth

Publisher:
RSIP Vision

Contact us

Free subscription

Read Previous Magazines

Follow Us

Dear reader,

It's a significant month for the computer vision community as CVPR returns in the eclectic city of Vancouver, Canada. RSIP Vision and Computer Vision News will be there, and we're so excited to meet all of you! We'll be publishing three new CVPR Dailies across the week. Receive them in your inbox daily from 20 June by subscribing to our traditional initiative: **Feel at CVPR as if you were at CVPR!**

With CVPR in mind, we love bringing you the BEST of the BEST. Turn to page 4 for a preview of a fascinating paper by the **Robotics and Perception Group at the University of Zurich** that is candidate to win a significant award at CVPR later this month: **Data-driven Feature Tracking for Event Cameras**, featuring the talent of **Nico Messikommer**, **Mathias Gehrig** and **Carter Fang**. Best of luck for the award!

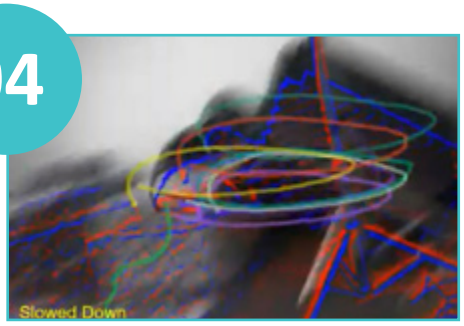
We have a treat for **robotics** aficionados this month: a charming interview with a freshly named professor, the awesome **Georgia Chalvatzaki**. You can watch the interview on video, it's much better!

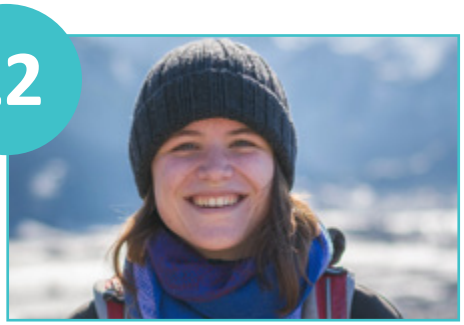
We review a captivating paper that will be presented at MIDL2023 next month by **Nandita Bhaskhar and Rogier van der Sluijs**: brilliant Nandita tells us (in video!) about the need to apply large self-supervised models to medical images, focusing specifically on chest X-rays. We will learn about her aim to train these models **without explicit labels, using self-supervised learning techniques such as pseudo-learning and contrastive frameworks.**

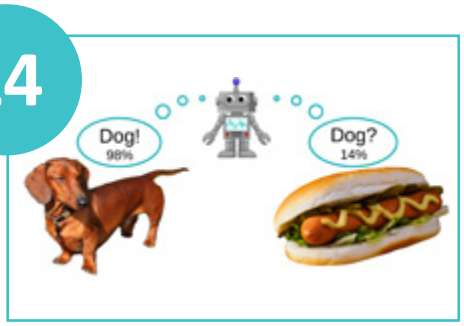
Enjoy reading all this rich content and more in June's edition of Computer Vision News and our supplement Medical Imaging News. Next month, we'll have an extraordinary **BEST OF CVPR** section. You won't want to miss it!

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

Computer Vision News







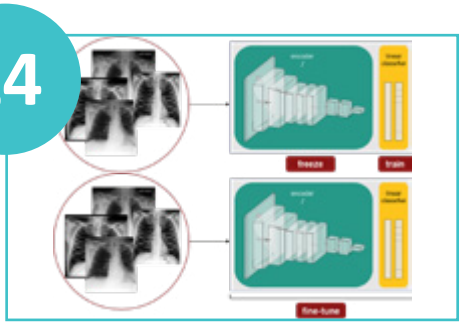
04 **Data-driven Feature Tracking for Event Cameras**
CVPR Candidate Award Paper

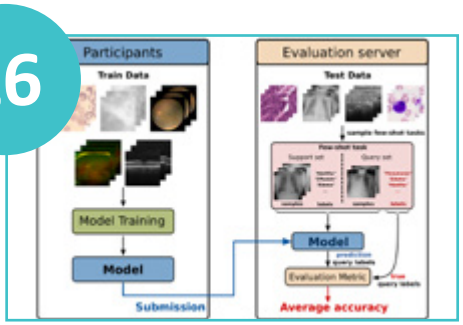
12 **Measuring and Modeling Mountain Glaciers**
AI Newcomer Award

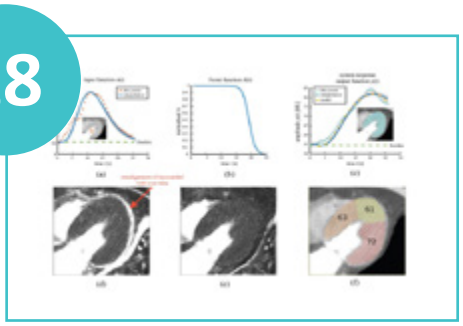
14 **Alexander Meinke**
Congrats, Doctor Alexander!

16 **Georgia Chalvatzaki**
Women in Computer Vision

Medical Imaging News







24 **Exploring Image Augmentations for Siamese ...**
Oral Paper at MIDL 2023

26 **Learn2Learn - Cross-Domain Few-shot Learning**
MICCAI 2023 Challenge

28 **Andrea Lara**
Congrats, Doctor Andrea!

36 **Trustworthy Machine Learning for Healthcare**
ICLR 2023 Workshop

DATA-DRIVEN FEATURE TRACKING FOR EVENT CAMERAS

Nico Messikommer and Mathias Gehrig are PhD students in the lab of Davide Scaramuzza in the Robotics and Perception Group at the University of Zurich. Carter Fang was a master's student at ETH and is now a Research Engineer at Waabi working on autonomous driving for trucks. All three are co-authors of a fantastic paper that is a candidate to win a significant award at CVPR later this year.

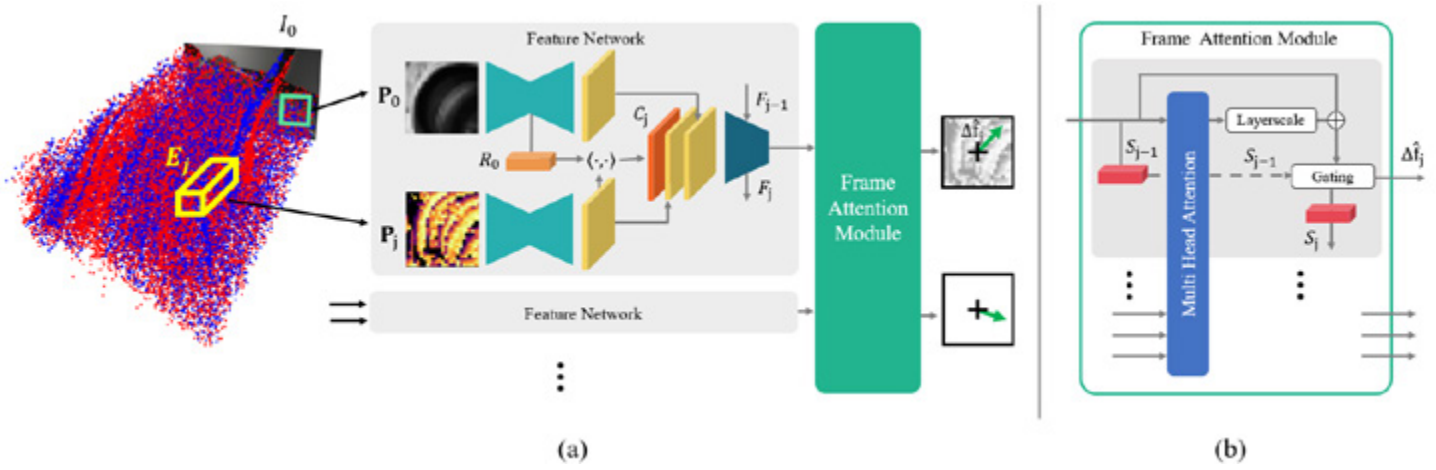
Nico Messikommer

Mathias Gehrig

Carter Fang

Event cameras differ from standard cameras in how they measure light at the image sensor. Instead of capturing frames continuously, **event cameras only output a binary signal when there is a significant change in intensity above a predetermined threshold**. Consequently, no information is obtained when an event camera is pointed at a static scene since there are no intensity changes. However, if an object moves within the scene, the pixel locations experiencing intensity changes will generate output events, primarily representing the edges surrounding the object.

Event cameras draw inspiration from **biological receptors in the human eye**, which also focus on detecting changes in the scene. Our brain often disregards motion,

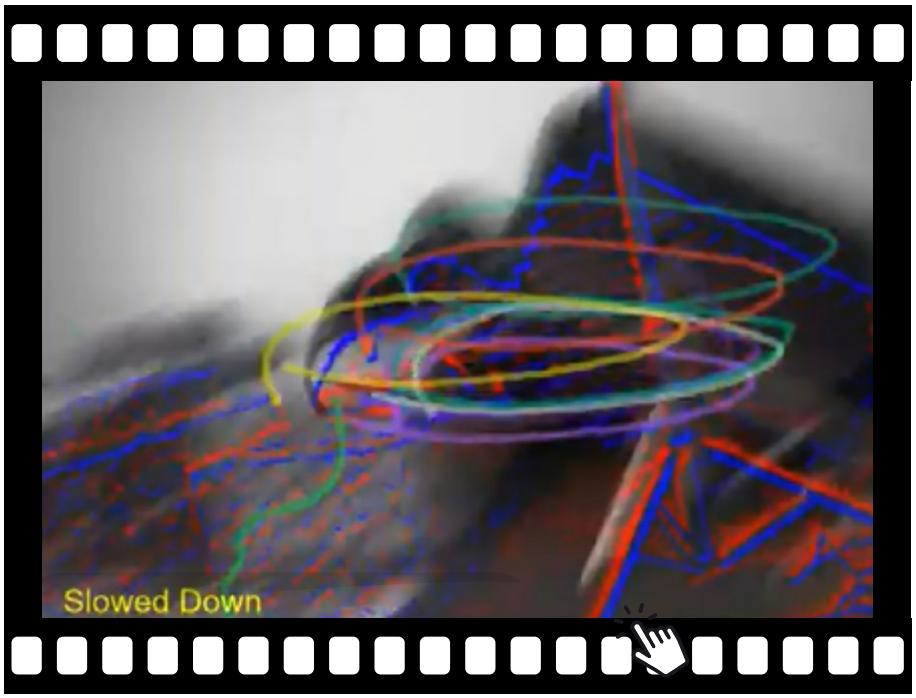


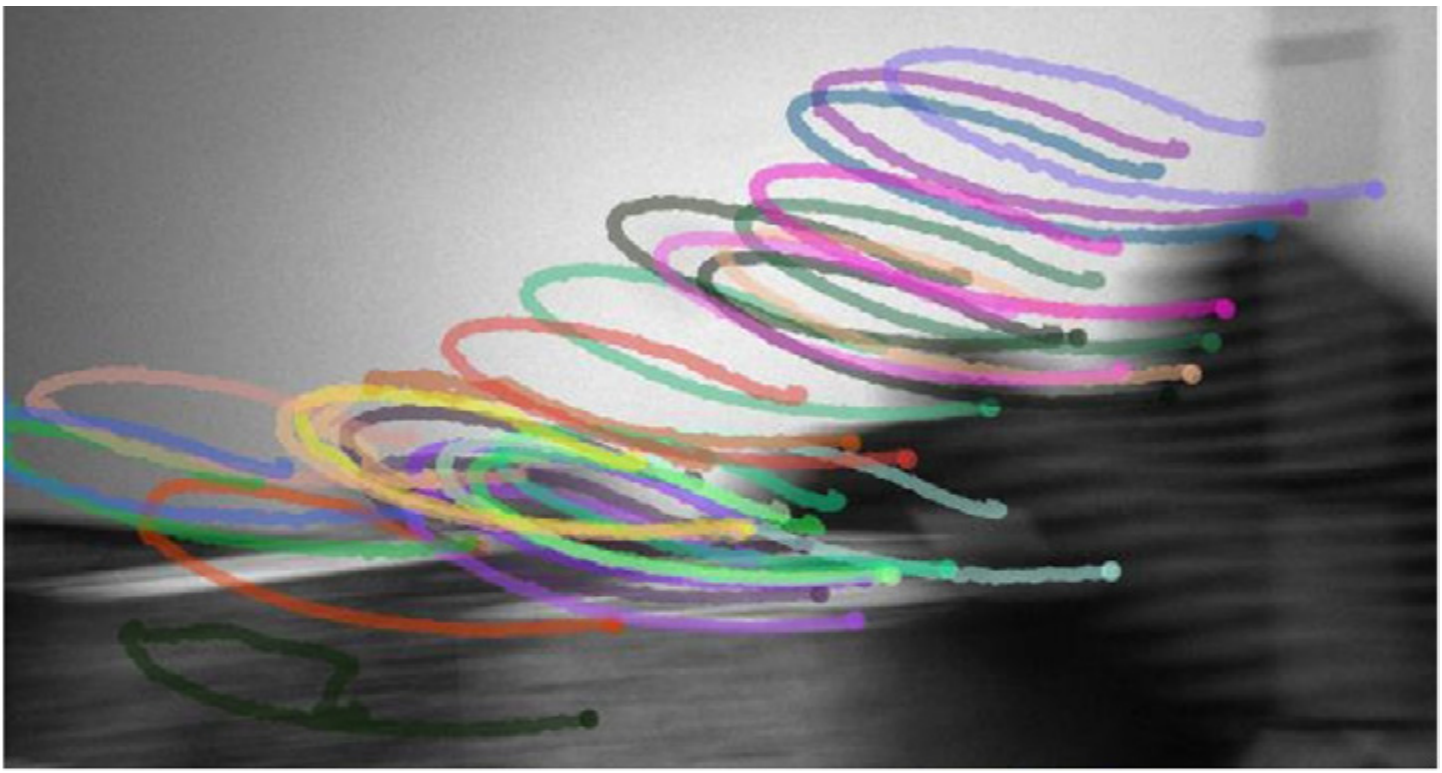
but if it deems it relevant or significant, it uses that information downstream. Similarly, event cameras exclusively process changes and ignore static scenes to avoid processing redundant information.

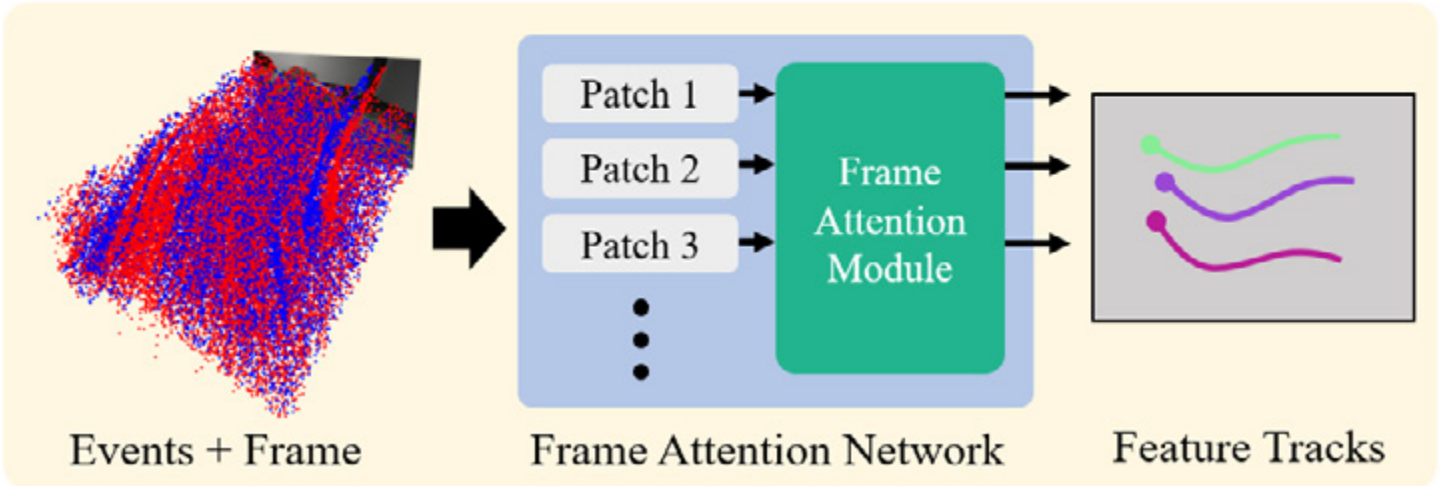
In this paper, Nico and Mathias explore using event cameras in combination with frame cameras to achieve robust feature tracking in sequential images. Feature tracking is crucial for various applications. The **Robotics and Perception Group** focuses on **SLAM algorithms and pose estimation using cameras placed on robots**. The researchers also work with **drones**, involving high-speed motion and maneuvers, with robust feature tracks to compute the pose using SLAM and VO backends.

"Robust feature tracking is often called the front end of visual odometry pipelines," Mathias tells us. *"**These VO pipelines are the foundation of mobile robotics** because they're required for control algorithms in the robots to tell the robot where it needs to go or what it needs to do. If you don't have access to these visual feature tracks, it's almost impossible to tell the robot what it needs to do from its own perception."*

Previous work in this domain has predominantly relied on standard image cameras, such as video cameras, which have some drawbacks. Existing trackers based on standard camera images are affected by issues such as motion blur in high-speed scenarios, resulting in a loss of scene structure. Additionally, the frame rate of standard image cameras is typically limited to around 20 fps. This work proposes the inclusion of an event camera alongside a traditional camera, which offers higher temporal resolution and enhances the







overall performance of feature tracking.

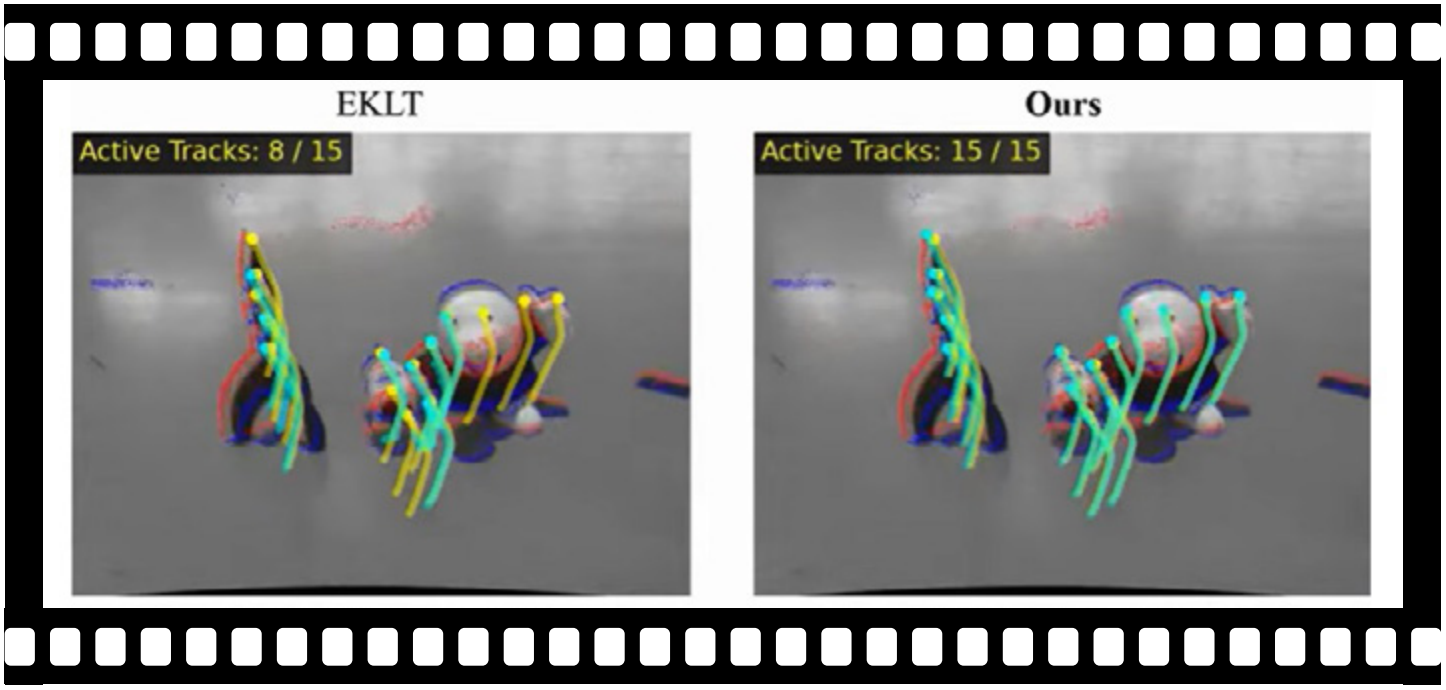
*"What is new about this paper is that we have an **end-to-end solution to feature tracking that combines both frame-based and event-based domains**," Mathias explains. "One of the big challenges, if you use both modalities, is how can you associate information from one modality to the other while not losing the advantages of each modality?"*

Existing event-based trackers use model assumptions to solve the problem. In this work, the team uses a novel data-driven approach so that the network implicitly learns the association between a grayscale patch and event sequences.

*"Another big challenge was the overfitting to training data," Nico tells us. "We needed to come up with different augmentation strategies to augment the synthetic data. We proposed **a novel supervision method that can be used directly on real data and relies only on the camera poses**. With that, we can fine-tune our network on the real data."*

Carter adds: "***Our model predicts the***

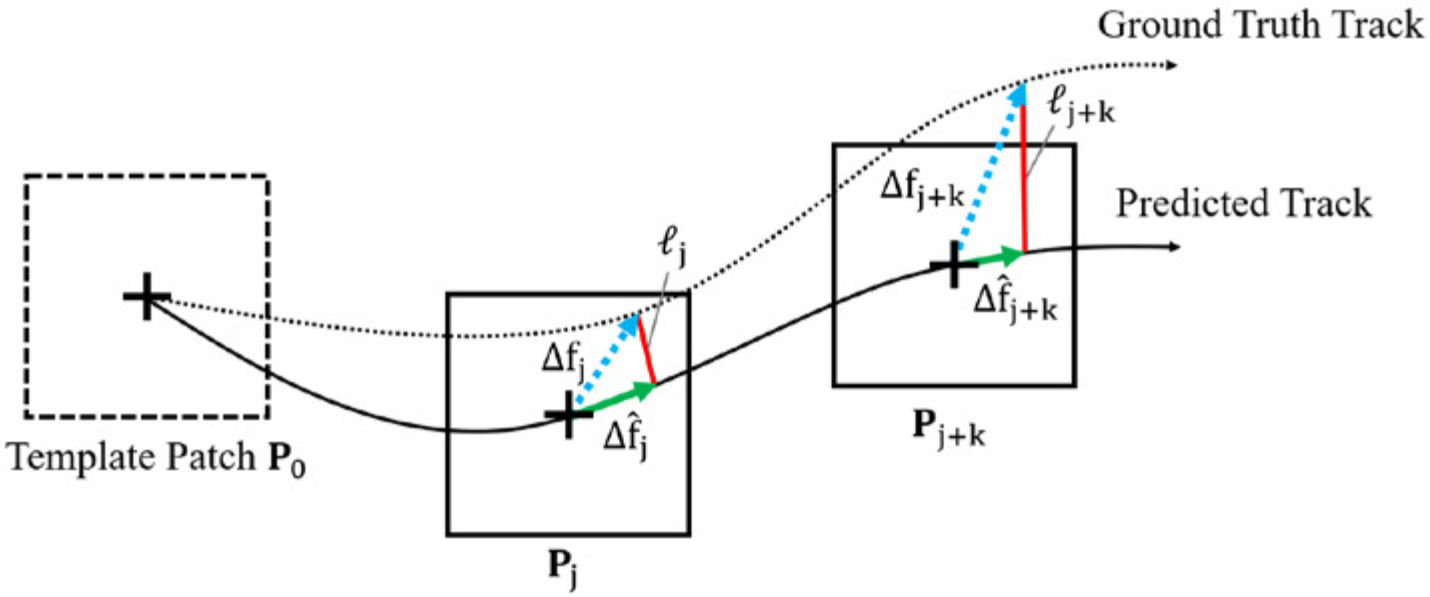

HAPPINESS IS...
HOLDING TICKETS
THAT SAY
VANCOUVER
See you at CVPR 2023 in Vancouver.
Come Over. In-Person. It will be fun!

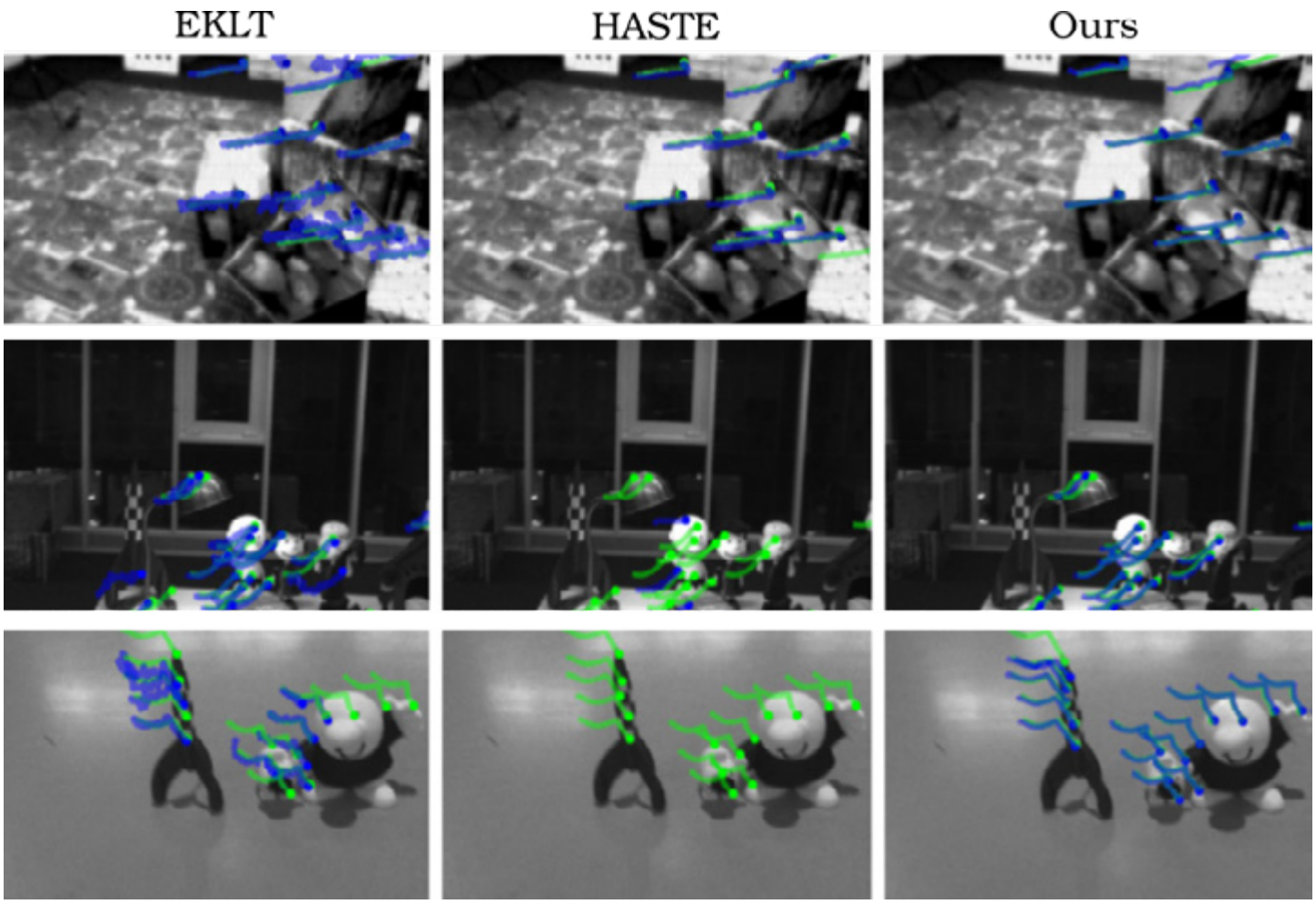


incremental flow at every step. *Many feature trackers currently in the wild are based on optical flow. To make the model more robust to the different sorts of geometric changes that a patch might undergo at each step, we take the patch and apply similarity transform – rotating it, scaling it, and then forcing the network to predict the augmented flow. We found this strategy to be effective in closing the sim-to-real gap."*

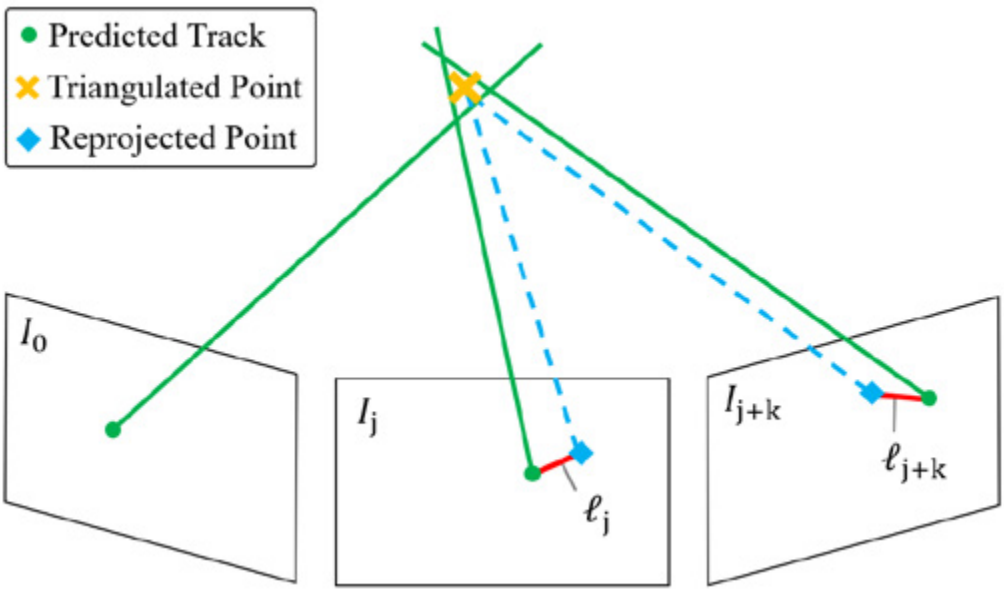
The model begins using a feature network to process each detected feature track separately. **This network consists of existing common, fully convolutional layers, a correlation layer, and a recurrent layer**. The correlation layer allows for correlating a grayscale image patch with event sequences, while the recurrent layer utilizes temporal information in the events, which inherently contain motion.

"Then we add a frame attention module on



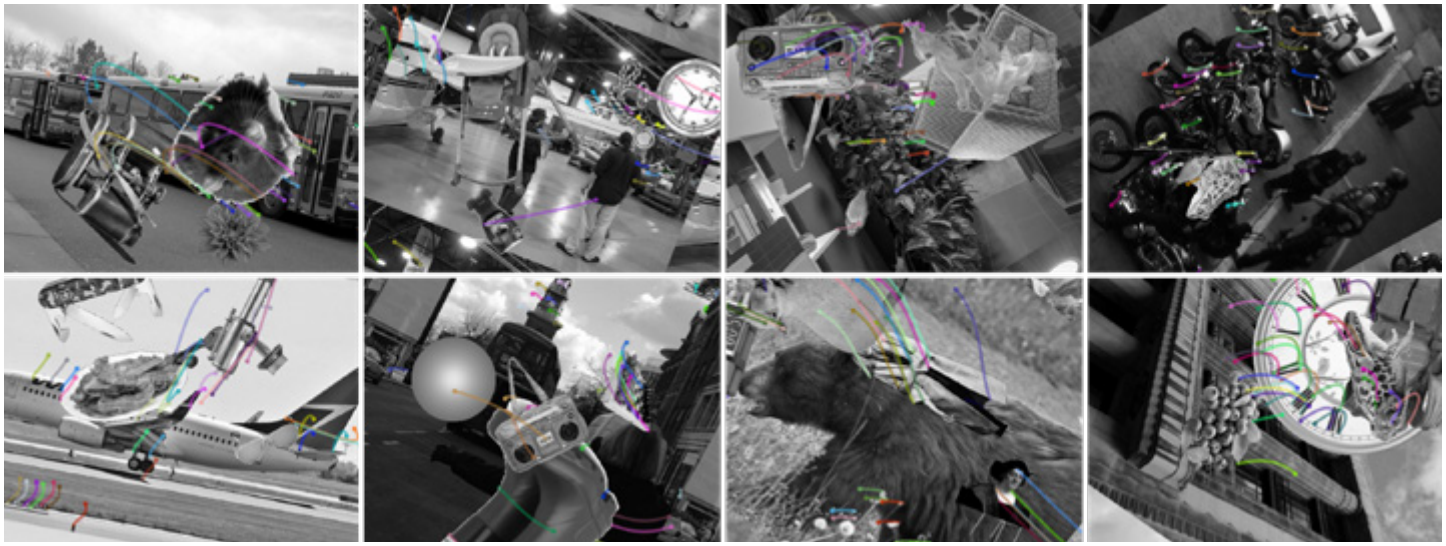


top of the feature network," Nico explains. "*We have feature networks for each track in one frame, and then we want to share the global information inside one frame, so you have multiple tracks in one frame. The frame attention module relies on a multi-head attention layer to combine the information from the different tracks in one frame and then output the final displacement vector.*"



Reflecting on being recognized among **the top 20 papers at CVPR**, we want to know if Nico, Mathias, and Carter have contemplated the reasons behind the judge's decision. How has the paper transcended being a mere poster or highlight material to deserve an award?

"***Event-based feature tracking was often regarded as very***



***promising for the application of event cameras**, but it's never been robustly shown to work to such a degree that people think you could apply it directly and gain huge benefits," Mathias points out. "One issue was that modeling event cameras is very difficult. Even for our lab, which is quite expert in this, it's not straightforward. In our solution, we overcome the longstanding challenge of event-based feature tracking in a data-driven fashion, letting the benefits of event cameras shine. This is reflected also in the experimental results, where we show **over 100% gain over the baseline**. It's not very common in computer vision to have such strong results."*

Nico says the future application of the method has likely played a role in the paper's success. It demonstrates the possibility of combining traditional image approaches with events, reducing bandwidth in certain applications. While not explicitly shown in the paper, **this reduced bandwidth is significant for virtual and augmented reality scenarios**. Tracking images with events makes it possible to make real-time decisions about feature track quality and request new images when necessary. This approach can result in lower frame rates, bandwidth requirements, and power consumption.

*"Event cameras have been promising, especially for applications like **feature tracking, where the benefits of high speed and high dynamic range have been promised**, and our work unlocks that in a data-driven fashion,"* Carter ponders. *"**At a higher level, our proposed algorithm is quite simple!** It's a straightforward idea that works well, so that's also something in our favor."*

Thinking about the next steps for this work, the team plan to revisit the initial motivation to use the feature tracker for pose estimation. They want to integrate it with a VO pipeline to improve the overall estimation of the camera pose. Additionally, an open question remains regarding which features to track.

"In this paper, we predetermined which features we wanted to track," Mathias reveals. *"You can imagine that certain features are easier to track in certain scenarios than others, but at the moment, this is completely handcrafted. Eventually, the process could be optimized end to end with **a visual odometry or visual-inertial odometry pipeline**."*

ARE YOU GOING TO MISS CVPR 2023? YOU CAN BE ALMOST THERE ANYWAY!

Here's the trick to follow CVPR day by day: click here and subscribe for free to CVPR Daily (20-21-22 June) with awesome highlights from CVPR!

KI 2023 CAMP
AI Newcomer 2023

Nora Gourmelon is a PhD candidate at the Pattern Recognition Lab at the Friedrich-Alexander University in Germany and the recent winner of the AI Newcomer Award in the field of natural and life sciences at KI-Camp 2023, an event for young AI researchers organized by the German Association of Computer Science and the German Federal Ministry for Education and Research.

Nora tells us about several projects she has been involved in, including her current PhD research on the recognition of **calving fronts**, which is the position where a glacier ends, and the ocean starts. She uses **deep learning techniques for segmentation** and has transitioned from an **optimized U-Net** to a **transformer approach**. This research has been ongoing for two years and has resulted in the publication of a benchmark data set and baselines.

In another project, Nora is collaborating with the **German Federal Agency for Nature Conservation** to automate the process of bird counting in aerial videos. In the first study, she employed a **neural network to recognize birds automatically in the videos**, which indicated positive results.

Also, Nora was recognized for her research during her master's thesis, which focused on smart water meters and time series analysis. She concluded that **classification requires labeled data** as clustering methods did not yield satisfactory results using only the time series information.

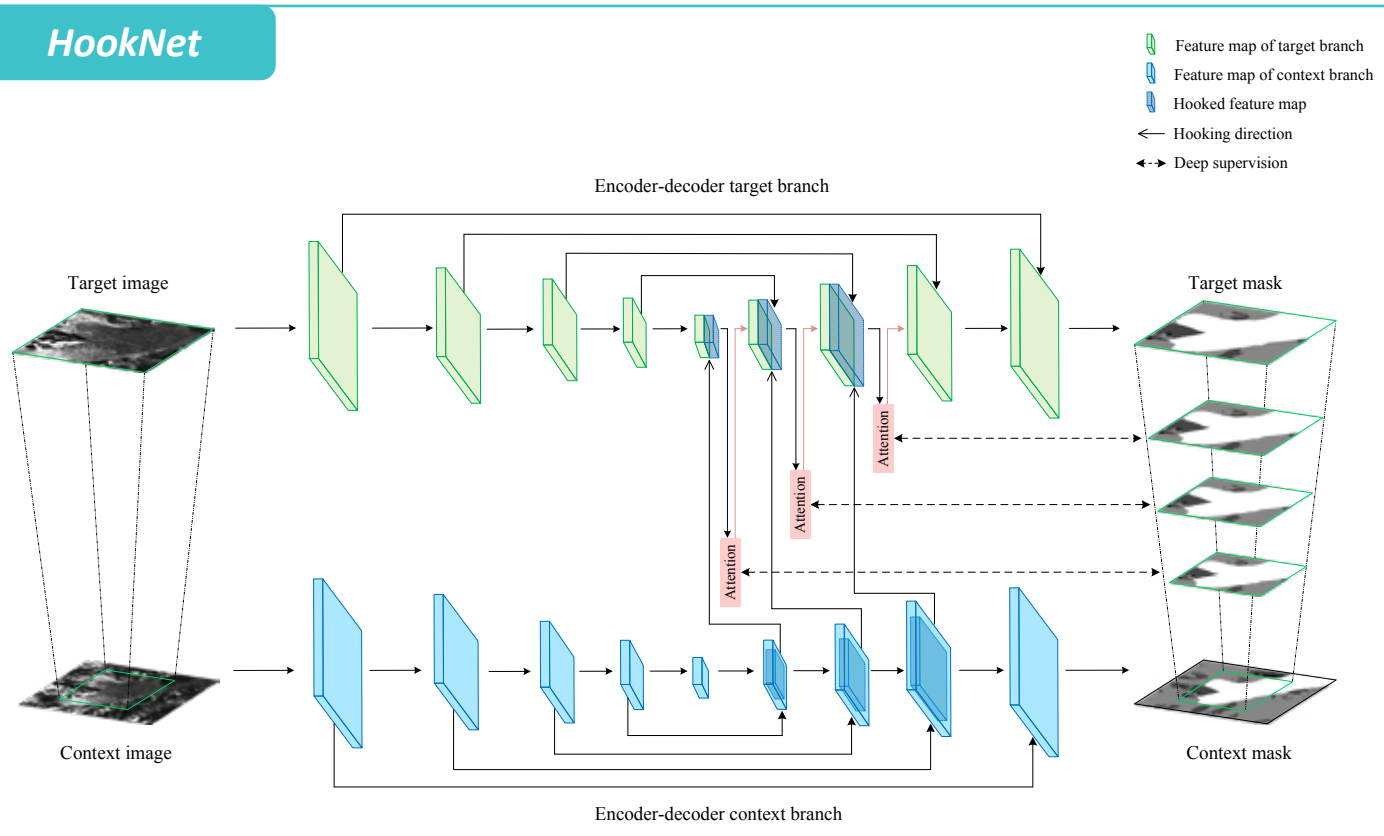
Nora is modest about what specifically impressed the jury and scooped the **AI Newcomer Award** but emphasizes her dedication to her PhD research.

"I'm most proud of the work I'm doing right

now with the calving fronts, so with the glaciers," Nora tells us. *"I have put, I think, the most effort in those because I've been working on it now for two years. We published a* ***benchmark data set*** *there, which is really nice, and some* ***baselines****. We now have very many new ideas and want to go in the direction of* ***uncertainty****!"*

Watch our video interview now to hear Nora speak in more detail about the technical and scientific parts of her work, including the specific models involved in the award.

HookNet



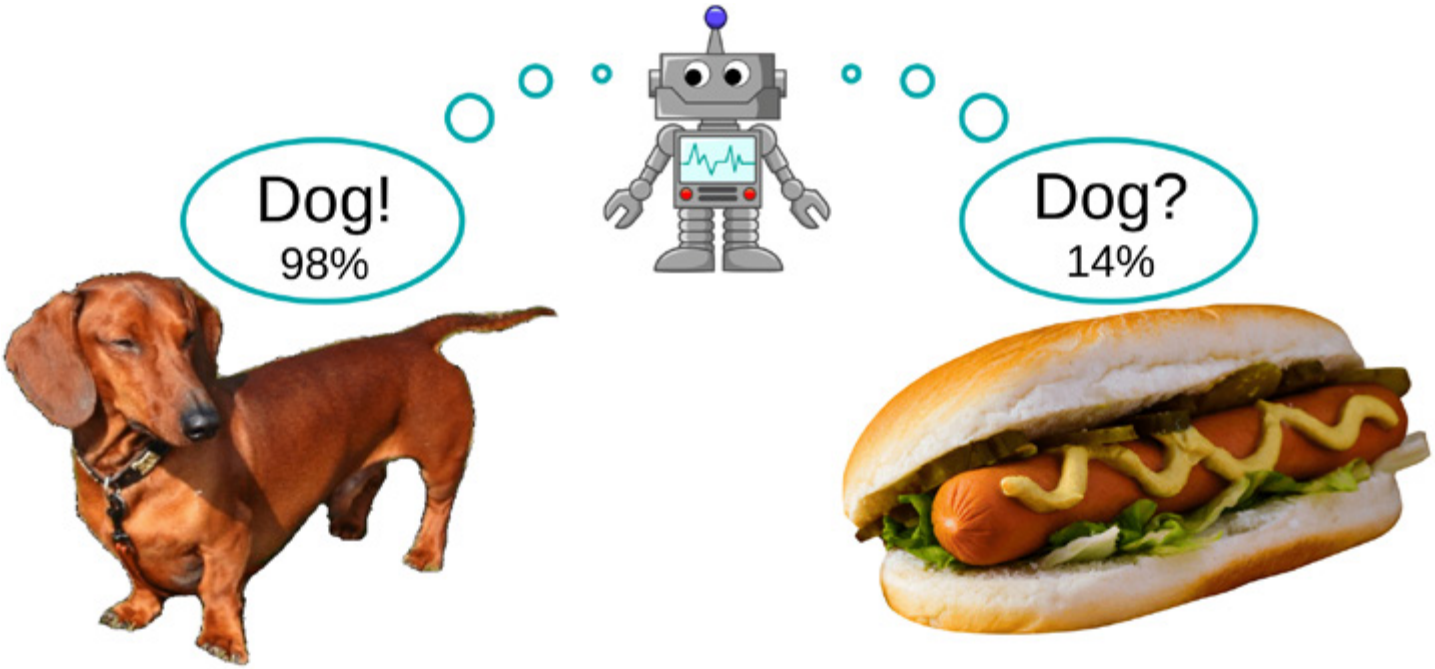
Congrats, Doctor Alexander!

Alexander Meinke recently defended his PhD at the University of Tübingen. In his research he focused on the robust quantification of uncertainty in neural networks. He is now working on AI safety and existential risks associated with AGI. Congrats, Doctor Alexander!

Do neural nets know when they don't know?

In general, the answer is no, how could they? That is to say there are two problems: 1) A **neural classifier** will give highly confident predictions on samples that don't even contain any of the classes that they were trained on and 2) you can actually mathematically prove that neural classifiers become **more and more confident** the further you move away from their training data, instead of less and less confident. Let's take these issues one by one. For the first problem we are saying, what if we are trying to classify cats and dogs and suddenly somebody shows a toaster? Then the model better assign low confidence. We can achieve this behavior by simply handing the model a whole bunch of samples that do not contain the classes during training and telling the model to have uniform predictions there. For example, in vision you could scrape a huge set of random images off the internet and use this as your training out-distribution. Turns out that in many cases the neural net will learn to generalize its low confidence for unseen out-distributions as well.

That all sounds well and good but of course there are some issues with this. As you might know, neural networks are generally not robust to **adversarial perturbations**. In this context this means that even if our neural net correctly assigns low confidence to an unknown class, **we can adversarially manipulate this out-distribution sample so that it looks almost identical but suddenly receives very high confidence from the model**. Of course, there is lots of research on adversarial robustness and the most successful method from



that literature, known as adversarial training, can also be applied here. Basically, you run adversarial attacks at train time and teach the model to retain low confidences even on the perturbed out-distribution samples. **But verifying that your model really is robust is a computationally intractable problem.** So even if you have a perfectly robust model, you can never know that for sure. This is where provable methods come to the rescue. **In my work I managed to come up with such a method that is robust, scalable, simple to train and computationally cheap**.

And the best part? You get a second guarantee for the price of one. Earlier I said that one can mathematically show that standard neural nets become asymptotically overconfident on far-away data. **Our method actually provably fixes this so that as you move far away from the training data your model has decreasing confidence. You could say, the model knows, that it doesn't know!**


Georgia Chalvatzaki is a Professor for Interactive Robot Perception and Learning at TU Darmstadt in Germany.
Over 100 inspiring interviews with successful Women in Computer Vision in our archive!
Katrin Binner

You are a new professor – congratulations!

Yes, I am a new professor! [*Georgia laughs*] Thank you!

You have chosen a very special field for your research. Do you want to tell us about it?

I research in the general field of intelligent robotic assistants. In order to enable robotic assistants to become useful to our society, meaning to become service robots in our homes or nurses at hospitals or in elderly houses, we need to endow them with specific capabilities. For example, we want our robots to be able to close this very difficult perception-action loop. We want this to happen not only for low-level actions, so how the robot would move and manipulate objects, how it can effectively grasp objects, and so on, but also for how it can use its perception in order to do reasoning. We are looking into the interplay between perception and action and reasoning to enable these robots to become of good service to us.

I am sure you have tried to convince robots to do many things; some have succeeded, and some have not. What can we convince a robot to do, and what are they unable to do yet?

What we have succeeded in is that we are now able to have a mobile manipulator robot, which means a robot that has two arms, a head, and a mobile base, so it is like a humanoid mobile robot. We have managed to make it effectively understand from a scene the object that it needs to grasp and then to be able to do sequential pick and place autonomously without us needing to program something on the robot. But when it comes to more challenging tasks, like needing these kinds of robots to open drawers, fridges, doors, put things in cupboards, take things out, or more fine manipulation skills like preparing a coffee for you, that is beyond just picking and placing stuff, we are very far away still from making this robot understand how to operate with all their body in an environment. This is still very challenging.

What is missing for us to be able to do it?

We need many things. First of all, we need to be able to perceive well the components of the scene, such that we know which parts hold enough for them, which can be graspable, and for which task they can be used. Then we need this higher level of reasoning, so the robot will be able to think of the sub-tasks that you have to do to do these more complicated, easy-for-us things to do. The other part is that you need better coordination in the body of the robot. For example, for us, it is very intuitive to open the door of the fridge, and while we are opening it, we are also adapting our body while this door is being manipulated, but this is very challenging to do with a robot without the robot hitting the door of the fridge, hitting itself, colliding, and so on. We have a lot of progress to do there. For this, we have looked into neural representations, so geometric representations. We are trying to understand distances by querying batches of thousands of points and trying to adapt the way the robot would move so that it does not collide and effectively completes a task. This combines ideas both from geometric deep learning and reinforcement learning, so that the robot can learn based on what it perceives how to adapt.

Do you ever get frustrated with the robot and say, I am done with that? I do not want to work with robots anymore!

[*Georgia laughs*] Yes, this is an everyday experience for us! Basically, you start working with the robot, you set up everything, and then suddenly, when you have everything very nicely done in simulation, you try to do it on the real robot, and actually, it does not work. Other things that happen are that the robot may overheat and then stop working, or you program things in a very good situation, and then suddenly your camera needs to be calibrated again, and then you have to start all over until you find out what is going on. You can look at it and you can see that it was working, but what is happening? Then you say, oh no, I have to calibrate this camera again, and hours have passed. So, yes, frustration is a part of our job! [*She laughs*]

But I am sure that on the other side, there are many things that you have taught the robot to do that we cannot do.

There are things definitely that, by reinforcement learning, for example, we can see robots finding solutions that could be beyond what I could have taught the robot by kinesthetic teaching. I could have taken the arm of the robot and tried to put it in a specific location, and then if I just allow the robot to explore its own capabilities in some sense because robots and humans do not have exactly the same structure in our bodies, and every robot is completely different, through reinforcement learning, we can see robots discovering their abilities into learning how to do things so that they can complete a task without doing exactly what I would have done. As long as they can deliver the thing. For me, it is okay if they do it in a completely different fashion than I would do it, but of course, they have to be successful. How you can motivate the robot to learn is actually one complexity of reinforcement learning. How do you define the reward? How do you define what is good behavior and what is bad behavior so that the robot is able to experience things and learn through this experience?

What was your most impressive eureka moment?

Ah, eureka! Yes! [*Georgia laughs*] This is a very, very interesting question. I would say that my first experience with reinforcement learning was the one, trying to understand the interplay between thinking ahead and being very reactive. That is more the part of reinforcement learning and how the robot can figure things out that go beyond what I would have expected. For example, one of the first works that I did here in Darmstadt when I arrived was on reinforcement learning. We had a robotic hand that we wanted to learn on the robot. We were not doing simulation at all; we were training the robotic hand itself. We wanted it to learn to open valves. We had a hand with four fingers, and the robot found out that it can do this very effectively using only two out of the four fingers. It did not lose energy to move the other two fingers because it figured out that "*I can do it very well by just using my two fingers!*" This is connected to what I said before. This amazed me. Yeah, it was very, very interesting for me to see.

You have chosen to work and have a career outside of your own country. I ask the question because I visited Greece


Claus Voelker

only a few days ago, so it is a very recent memory for me.

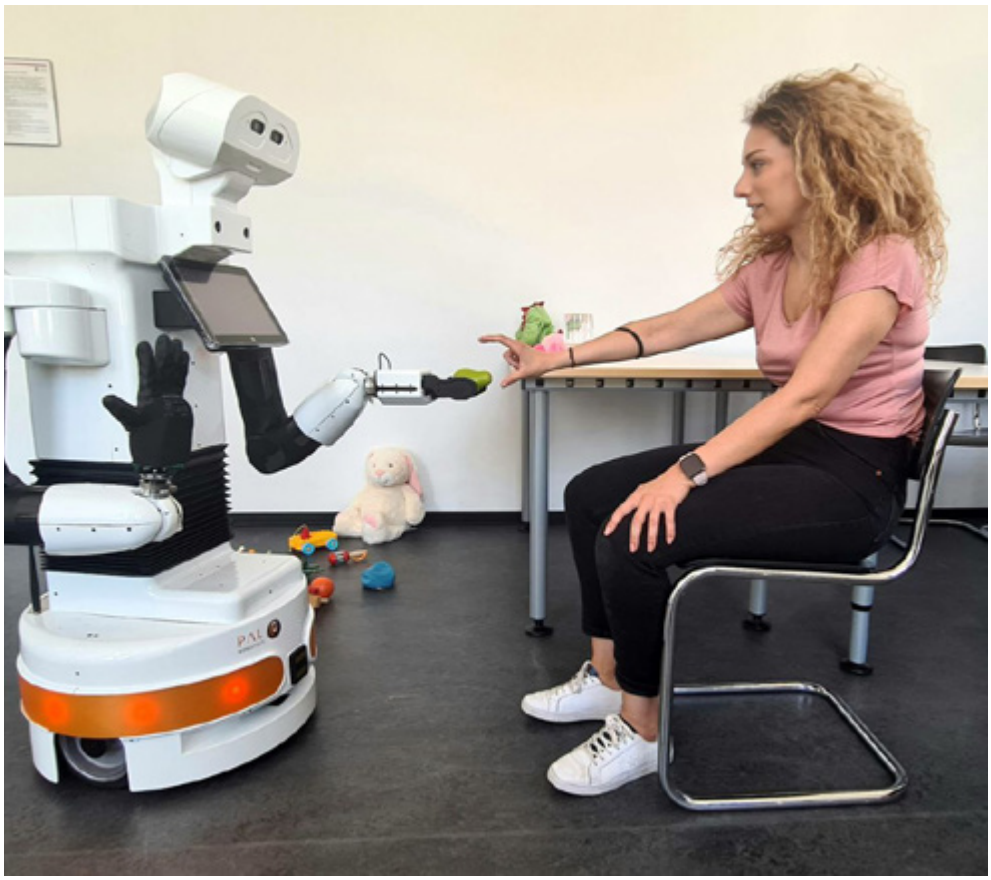
Oh, nice. I was also in Greece a few days ago! [*she laughs*] I visited my parents. I was in Rhodes.

I adore Rhodes. How can you find your place so far away from home?

This was a difficult part because I lived most of my life in Greece. I am from Rhodes, and most of my studies were in Athens. It was a bit difficult because the climate is different compared to the climate of Greece, and especially the climate of the island that I come from, so you understand that was a very difficult process. On top of that, there was Covid. I came to Darmstadt in November 2019, and in two or three months, there was Covid quarantine, so for two years, I was basically at home.

What I find here, what really makes me feel complete, one part is the work. I can do here things that I did not have the ability to do when I was in Greece because we did not have the resources, we did not have the robots, and we did not even have GPUs until the very end of my PhD. We took a lot of time to make a small cluster to be able to do deep learning. I am very grateful that I have this opportunity here, and I am surrounded by people that come from so many different nationalities. This is very interesting for me, and I think it makes me grow as a person.

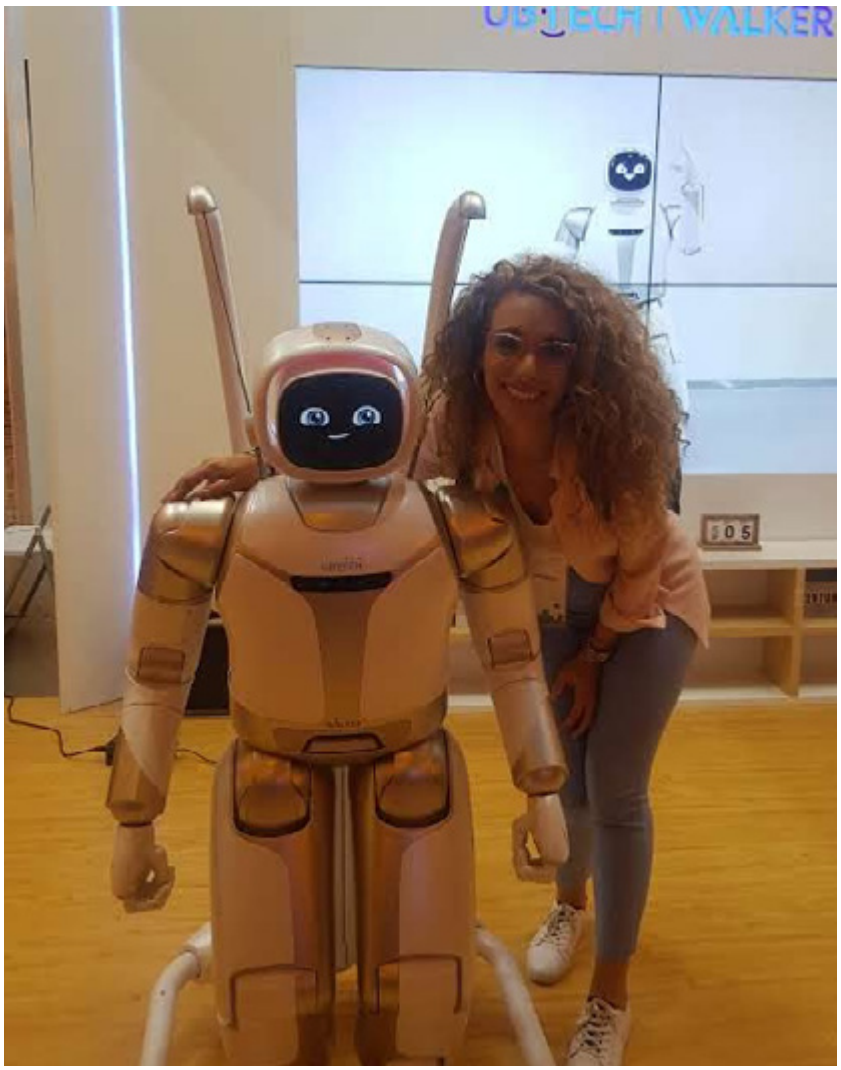
The other part that I am grateful for is, of course, my personal life here. I have friends already and my partner here, so it makes me feel well on a personal level. Now that there is no Covid, what I try to do is to take small trips to Greece whenever I can so that I can visit my family. Eventually, my mother visited us last October. She had not visited us all the time due to Covid. Now, given that we are very close to Frankfurt Airport, it gives us the opportunity to travel easily. Especially when the summertime starts there are direct flights from Frankfurt to Rhodes. I try to balance this, but I have to do it only when there is no teaching in the university. This is what I try to do, but to be honest, I become homesick, especially during summertime!



You did not have the resources for your research in Greece. Can you give Greece some advice?

The advice to Greece is to

invest some money in education and into making the universities more self-sustaining. The universities do a very good job in providing us with good foundations, and I think we can see many people that stem from Greece have very good careers outside of Greece. Unfortunately, Greece does not make any effort to go one step further to create facilities so that they can keep people there and advance research.



I can imagine, on the one hand, that the main income of Greece is tourism, so everything would go into this infrastructure. They might think that deep learning and robotics are not relevant, but this is very short-sighted because, for example, right now on my island, the hotels have a shortage of people because during Covid many people found other jobs or acquired skills online. They do not want to do the job of cleaning rooms in a hotel, so who is going to do this in the future? You need robots. Also, you need AI systems to become more effective to be able to process bookings and all these things. If we think about it, automation is inevitable, and the countries that do not invest in the direction of science in general – I would put not only AI – are not going to be able to keep up. Unfortunately, this goes for not only my country but also other countries, especially in the south of Europe.

You gave your advice to Greece. Now, I would like you to share some advice for a younger student from Greece who wants to succeed.

My advice would be to shoot for the stars regardless of the difficulties. If you really want to do something, it does not have to be robotics or AI, whatever it is, you should feel an inner motivation and not get disappointed and lose your courage. You have to be persistent, and you have to be resilient. People that come from Greece, especially after the things that have passed (and also other areas that I could imagine), we have learned how to be resilient, and these are qualities that you can use to your benefit. Keep trying. This is the point. Do not give up, and keep trying!

Over 100 inspiring interviews with successful Women in Science!

COMPUTER VISION EVENTS

CVPR
MEET US THERE
Vancouver, Canada
18-22 June

Deep Learning World
Las Vegas, NV
18-22 June

FIMH Functional Imaging and Modeling of the Heart
Lyon, France
19-22 June

IPMI
San Carlos De Bariloche, Argentina
19-23 June

CARS
Munich, Germany
20-23 June

IPCAI
Munich, Germany
20-21 June

Hamlyn Symposium on Medical Robotics
London, UK
26-29 June

BVM Workshop
Braunschweig, Germany
2-4 July

MIDL- Medical Imaging with Deep Learning
Nashville, TN
10-12 July

MIUA - Medical Image Understanding and Analysis
Aberdeen, Scotland
19-21 July

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

JUNE 2023

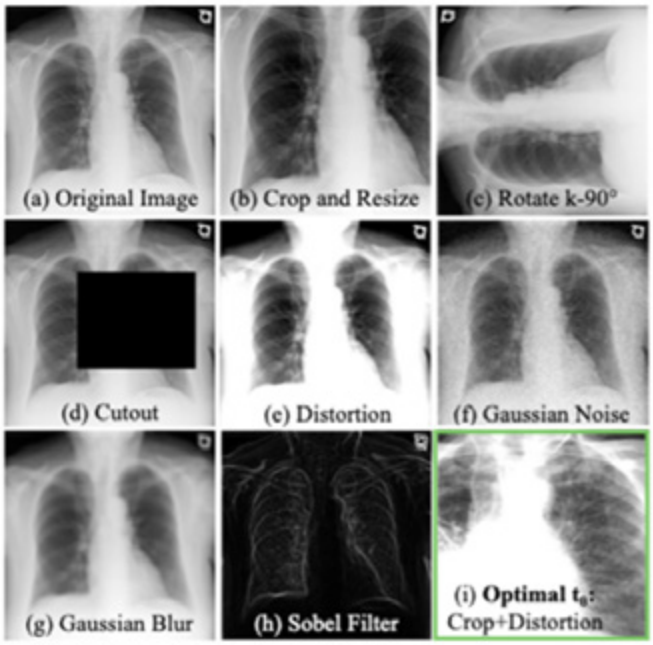
EXPLORING IMAGE AUGMENTATIONS FOR SIAMESE REPRESENTATION LEARNING WITH CHEST X-RAYS

Nandita Bhaskhar has just completed her PhD in the Department of Electrical Engineering at Stanford University.

With the success of large language models like ChatGPT, there is an avid interest in doing the same thing but for **medical images**. However, medical images are a completely different beast. They have domain-specific properties that differentiate them from natural images, such as those of cats and dogs.

In this upcoming MIDL paper, co-written with **Rogier van der Sluijs**, Nandita addresses the need to apply large self-supervised models to medical images, focusing specifically on chest X-rays. She aims to train these models **without explicit labels, using self-supervised learning techniques such as pseudo-learning and contrastive frameworks.**

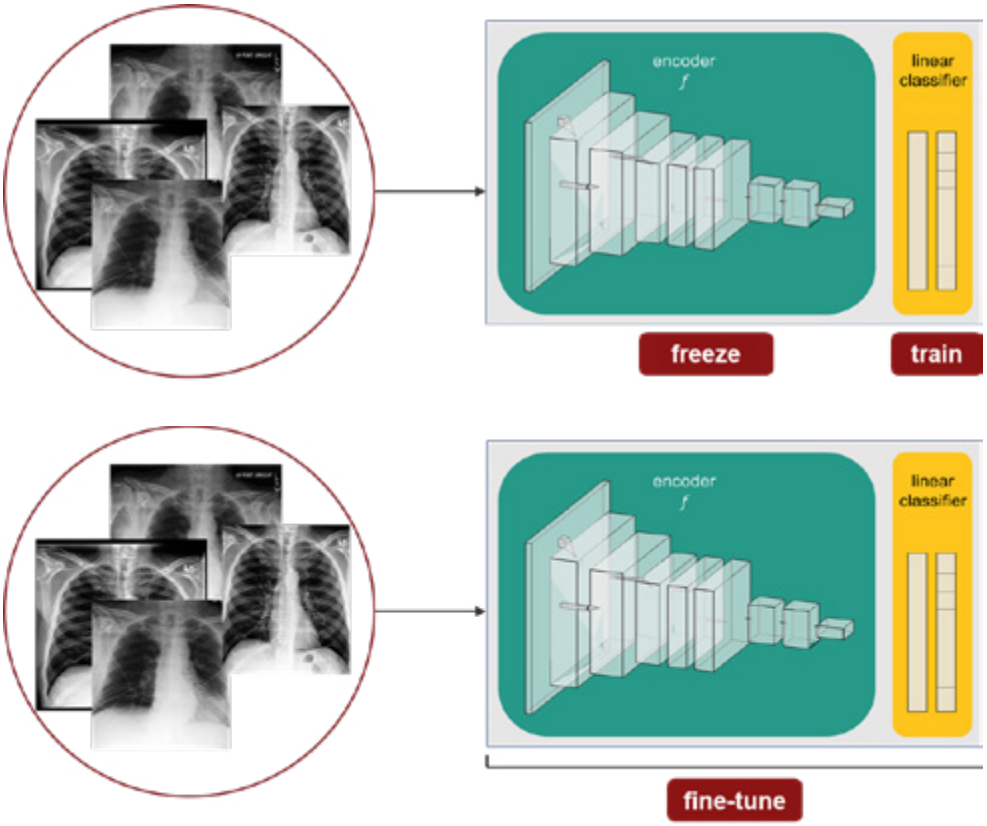
The key element that drives the success of self-supervised learning is the use of **augmentations**. Augmentations involve applying various computer vision transformations to an image, such as **Gaussian blurring or random resized cropping**. In the self-supervised framework, the image is transformed into two different augmentations, and their similarity or dissimilarity is assessed, which is the



property by which these methods learn. The paper explores **how to leverage the domain-specific properties of medical images to determine suitable augmentations that enhance self-supervision.**

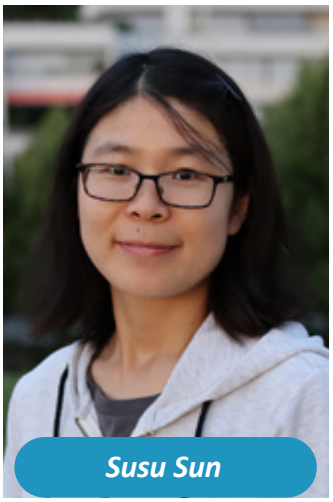
"There are a few things that make our paper unique," Nandita explains. "There are so many different types of augmentations that can be done for medical images. ***This is the first paper that did a systematic evaluation of various pairs of augmentations over a large scale with multiple data sets****. We also focus specifically on* ***domain specificity****. We introduced a set of augmentations with* ***optimal hyperparameters****. It's a recipe for practitioners to start using these techniques for their models. In our evaluations, we showed that not only are these well-performing augmentations, but they generalise well to different hospitals and domains. In our paper, we show that our models can even generalise to* ***unseen diseases****. For example, tuberculosis was never seen before by our model, but now it's able to understand something is new, and something is tuberculosis."*

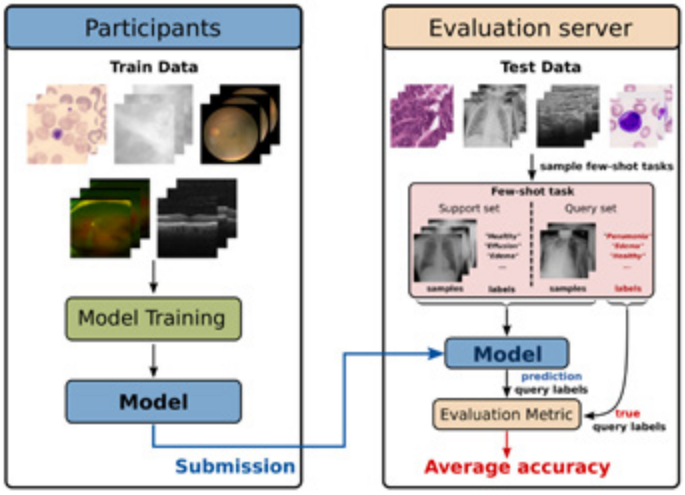
To learn more about this work, including the team's challenges and Nandita's work outside of this paper, please watch our video interview. All images are used with permission from: Nandita Bhaskhar - Robust, Data-efficient, and Trustworthy Medical AI - 2023.



THE MICCAI LEARN2LEARN CHALLENGE

Stefano Woerner and Susu Sun are PhD students in the Machine Learning for Medical Image Analysis Group in Christian Baumgartner's lab at the University of Tübingen. They are co-organizers of the MICCAI Learn2Learn Challenge, running now, with results announced at the conference in October.

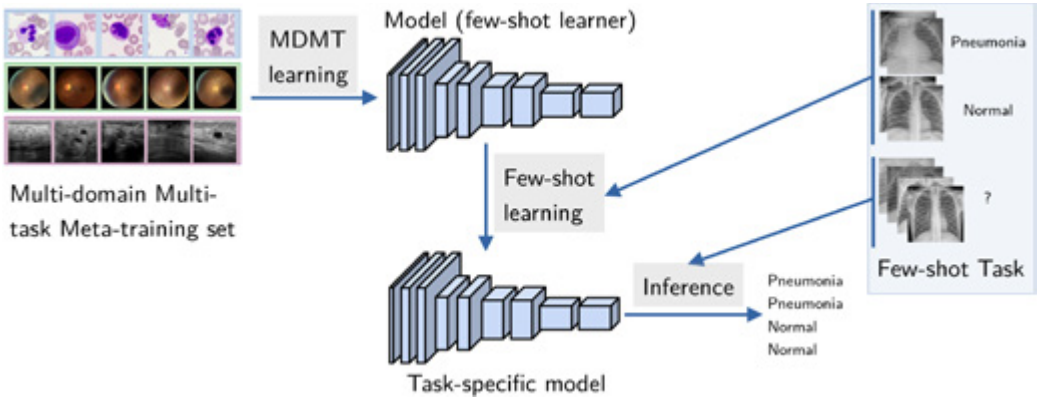


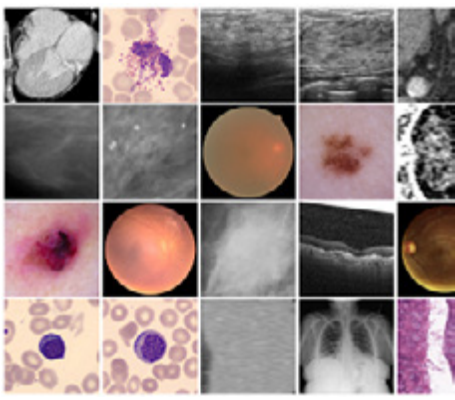


The **Learn2Learn Challenge** focuses on **cross-domain few-shot learning and meta-learning in the field of medicine**. These are relatively unexplored areas, particularly in realistic scenarios. Existing meta-learning research often revolves around toy problems using similar tasks from the same dataset, which fails to reflect the complexity of real-world clinical tasks and diverse domains. To address this, the organizers have created a benchmark dataset and initiated a challenge for other researchers to test their algorithms.

The challenge aims **to develop an algorithm capable of learning from multiple medical domains**. The researchers provide data from 17 different datasets, comprising a total of 28 tasks. The algorithm should possess **the ability to adapt to a new domain in a few-shot learning scenario where the specific target domain remains undisclosed**.

The datasets are presented in a unified format and accessible through a **PyTorch interface**. They consist of resized images, facilitating straightforward input to algorithms. While specific domain knowledge is not necessary, it might prove beneficial.

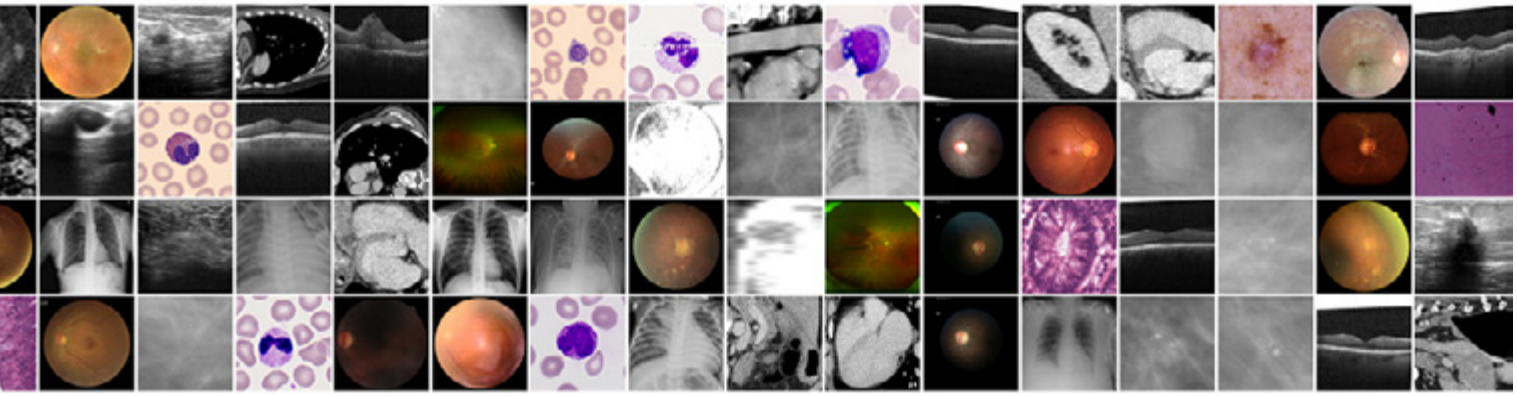




Of course, this is a medical challenge, so the **medical imaging community** is invited to join. However, Stefano and Susu also expect participation from scholars in **non-medical machine learning fields**, as few-shot learning and meta-learning are currently popular topics in machine learning research. The challenge has been designed to be easily accessible to machine learning researchers from any domain.

*"One thing we were trying to do ourselves is **to find out the best method to build an algorithm that can learn from the plethora of data we have** and then be adapted with very few examples on a new domain,"* Stefano tells us. *"That's what we want to know through the challenge. Through many people trying out their methods and ideas, we hope to get a better benchmark than just us in our group thinking about this and benchmarking a few methods."*

To learn more about cross-domain few-shot learning and the background to this challenge, watch excerpts from our video interview with Stefano and Susu.

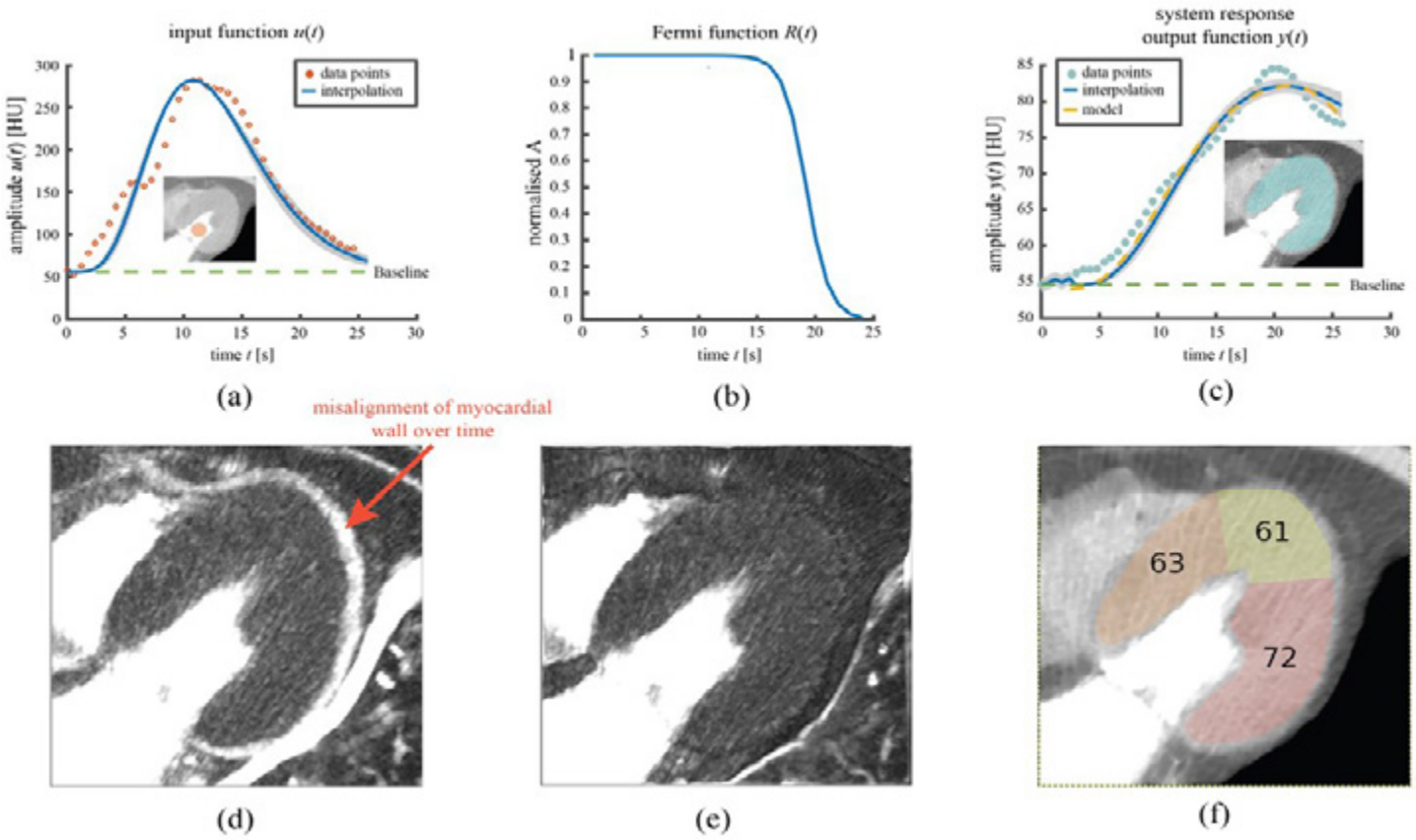


Congrats, Doctor Andrea!

Andrea Lara has recently completed her PhD at the Institute of Health Care Engineering with European Testing Center of Medical Devices (iHCE) at Graz University of Technology under the supervision of Prof. Christian Baumgartner. Her research focused on deep learning-based tools for medical spatiotemporal image processing and analysis. Currently, Andrea is the Director of the biomedical research laboratory (BiomedLab) at Universidad Galileo in Guatemala. Congrats, Doctor Andrea!

Recent advances in medical imaging have made possible to improve the acquisition and interpretation of spatiotemporal data. In many medical application areas, temporal information has high diagnostic value, as it provides additional information about the dynamics of physiological mechanisms. However, the analysis of this spatiotemporal information can be a complex and time consuming task. Deep learning techniques have proven to be particularly useful in processing medical image data. While most approaches focus on spatial image information only, innovative image processing techniques based on spatiotemporal imaging are still limited.

In her PhD thesis, Andrea introduces deep learning-based methods for the analysis and interpretation of dynamic imaging. Specifically, cardiac and musculoskeletal imaging where temporal information provides insight into the dynamics and physiology that can lead to a more accurate assessment of various conditions. First, she presents a systematic review of the current state of deep learning methods in cardiac imaging using spatiotemporal data and introduces an evaluation scheme to assess their clinical usability [1].



Result of clinical example.: (a) CT values time curves obtained from a ROI in the LV cavity (input function u(t)) (b) Fermi-function for deconvolution (c) measured and estimated CT values-time curves in the segmented LV myocardial wall (output function y(t)). (d) difference in HU values over time for the unregistered images (misalignment of the myocardium over the sequence). (e) difference in HU values over time for the registered images (aligned LV myocardium after LCV registration). (f) calculated regional myocardial perfusion in ml/100g/min for the apical (yellow), septal (orange) and lateral wall region (red).

Motivated by the above, she introduces the *first* deep learning-based registration method for dynamic myocardial perfusion CT studies [2]. The proposed approach is based on a novel loss function and a recursive cascade configuration that performs 2D registration without affecting the changing contrast agent concentration, which is essential to perform perfusion measurement. Moreover, this method is tested in a clinical example to quantify myocardial perfusion using an unseen patient with known aortic valve insufficiency as shown in **Fig 1.**

Finally, she investigates other applications where temporal information is essential in the understanding of the dynamics and physiology, such as the detection and tracking the muscle tendon junction [3]. In this work, she contributed by studying different deep learning architectures that exploit the spatial and temporal information present in musculoskeletal US sequences.

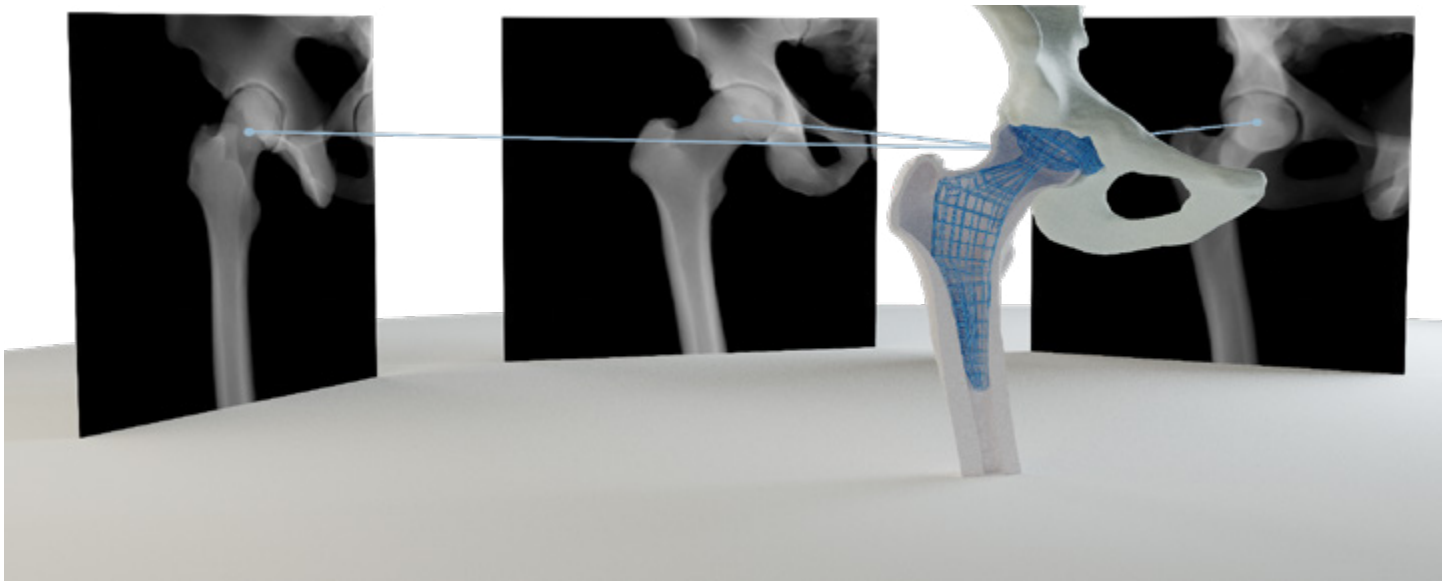
In the future, Andrea plans to extend her research to other dynamic imaging modalities and focus on applications that address current clinical needs in low-middle income countries (LMICs). Also, she wants to continue working to strengthen and empowering researchers from LMICs through initiatives like **RISE-MICCAI** and **SIPAIM**.

DEEP-LEARNING-BASED 2D-TO-3D JOINT RECONSTRUCTION FROM X-RAY IMAGES

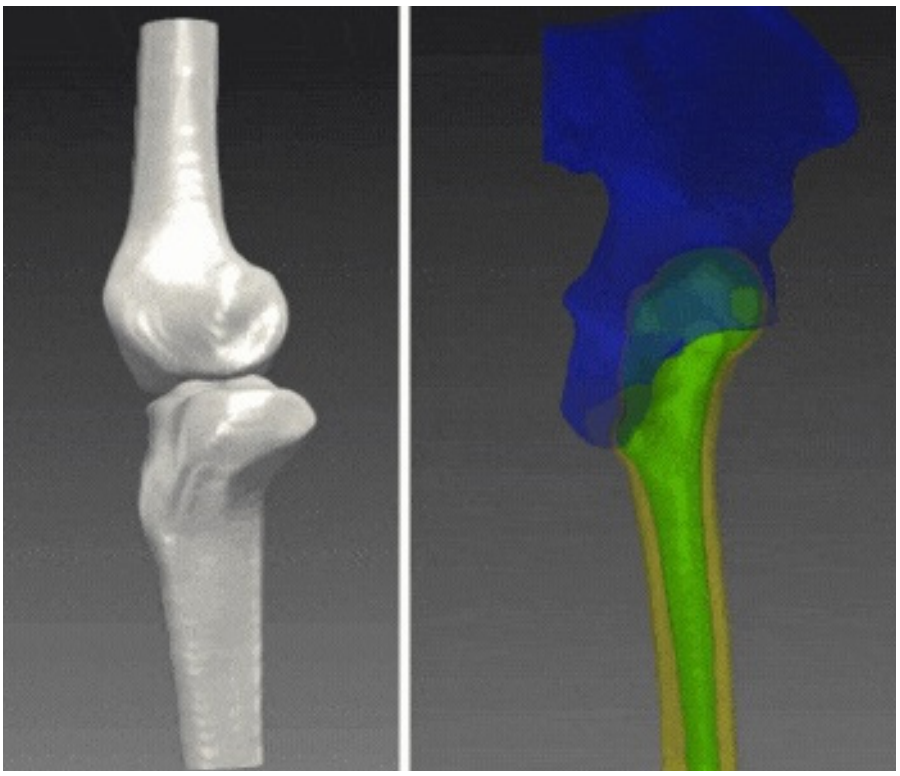
This article was first published here

3D models of joints and bones are widely used for a large variety of orthopedic clinical purposes including preoperative planning, patient specific implant design, jig printing, intraoperative navigation and other clinical purposes for which precise and detailed anatomy is required. An example of such a procedure for which a 3D model of joints is required is Arthroplasty. **Arthroplasty**, one of the most common procedures in the US with more than **1 million procedures performed annually**, is an orthopedic surgical procedure that aims to restore functionality and relieve pain by a full or partial replacement or remodeling of joints. This procedure goes through high-complexity steps, starting from planning to execution, all requiring accurate understanding of the patient anatomy, oftenly represented in 3D models.

Arthroplasty outcomes are highly dependent on the **technique accuracy**, during both planning and operation. Different types of **Patient Specific Instrumentation** (PSI) were developed for the purpose of **elevating accuracy** and **improving fit** of the implant to individual patients' anatomy. Fine adjustments and precise implant fit during pre-operational planning using technological solutions producing accurate and detailed 3D models



of the joints significantly improve procedure outcomes. A multitude of additional **new innovative technologies** aiming to improve procedures' outcomes are being developed - starting from **automated planning**, through **robotic surgeries** and usage of **augmented reality**. All of these require **precise anatomical 3D information**.



The standard 3D reconstruction is based on the 3D information captured in **CT scans**. Along with the benefit of having the information required for the generation of such models, CT scans have significant operative barriers including **low accessibility**, **high costs** and **reimbursement difficulties**. From the clinical perspective, for gaining high precision 3D information, CT scanners expose patients to high doses of **ionizing radiation**.

RSIP Vision's 2D to 3D reconstruction technology enables a clinical-grade 3D reconstruction of the bone from two or more 2D **standard X-ray images** without the need for CT scan. Using **advanced artificial intelligence algorithms** and **deep-learning based models** that allow capturing statistical information and features represented in complex data, this technology enables an end-to-end interpolation of the 2D information into 3D and reconstruction of **a highly accurate 3D model**. The resulting 3D model provides the information required for the variety of clinical purposes while replacing the need for a full CT scan. The alternative usage of x-ray images allows reduction in radiation, lower-costs with higher reimbursement rates, and higher accessibility, significantly improving the existing pipelines. These benefits potentially influence all parties involved, including patients, clinicians and healthcare providers.

Clinical evaluation of RSIP Vision's **2D to 3D Reconstruction** technology indicates **sub-mm accuracy** for the knee bones 3D-reconstructed from x-ray images. The technology is now extended for additional joints including hip, ankles and shoulders, and can potentially be extended to any other bony structure.

RSIP Vision is a global leader in **artificial intelligence** and **deep learning** for **medical imaging** with proven industry experience developing **clinical-grade** breakthrough AI-technologies. They will be happy to introduce you to their extensive records and use their experts to find the right solution for boosting up your product.

Find more about AI for orthopedics

IPMI 2023 – INFORMATION PROCESSING IN MEDICAL IMAGING

Alejandro (Alex) Frangi is a Professor of Computational Medicine at the University of Leeds, with joint appointments at the School of Computing and School of Medicine. Well-known to readers of this magazine, he speaks to us as General co-Chair of next month's IPMI 2023 conference in Argentina. We also hear from Program co-Chair Demian Wassermann and Local co-Chair Enzo Ferrante. Demian is Director of Research at INRIA in Saclay, France and Enzo is Faculty Researcher at CONICET, the Argentinian National Research Council.

Established in 1969, the **Information Processing in Medical Imaging (IPMI) conference** is a biennial event exploring the mathematical and algorithmic building blocks of **medical image analysis and processing**. Alex tells us its focus on supporting these methodological foundations helps to ensure sufficient underpinning and backing of other events in the field.

"IPMI has a core computing flavor, but it's not that we don't value translation

Alex Frangi

in the conference," he explains. *"In fact, many participants also attend **MICCAI** and **ISBI**. However, this is about the core methodological science that, as a community, we want to nurture and maintain."*

IPMI has one of the most distinctive formats of all the medical imaging conferences. With up to 140 attendees, it intentionally keeps numbers low **to support a highly collaborative environment**. Kicking off with a Sunday evening reception and finishing up on Friday lunchtime, the **presentations are longer than usual, with 20 minutes for orals** and, in theory, unlimited time for

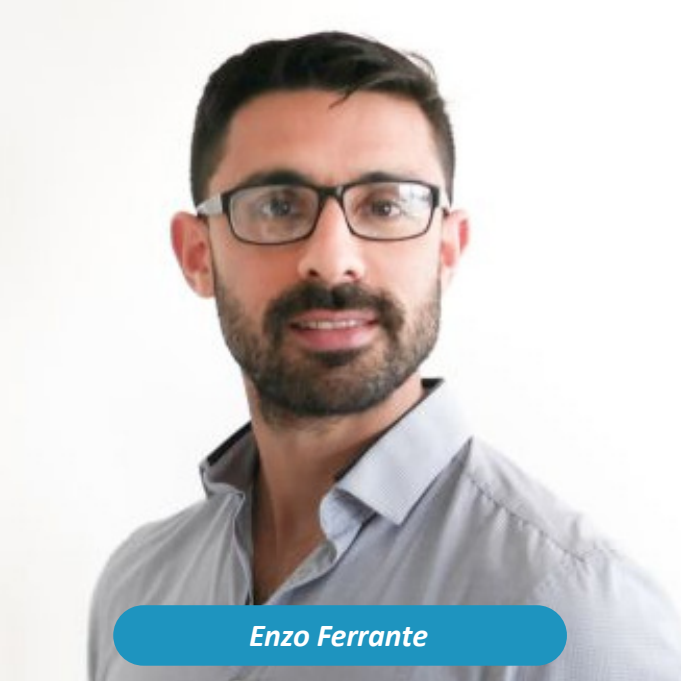
discussion and questions on each paper.

"*We allocate about 45 minutes to every paper, but the schedule we propose is tentative, and sometimes we need to be flexible,*" Alex points out. "*We do a thorough assessment of the work to benefit the authors and the rest of the community. That's why* ***we need to keep IPMI small so that it's properly interactive!***"

Demian adds that this format gives the community a great setting for exchanges and conversations: students enjoy a rare occasion to get much quality time with many scholars in their field. The event is **single-track**. "*There are also two main features*", adds Demian. "*one is that each oral presentation is introduced by a mixed team of about 10 people (professors, students and maybe industry) who studied that work beforehand. The other is that there is no fixed time for questions.*" Indeed, a unique feature of the conference is those **reading groups**. This process allows for more informed discussions during the presentations and adds to the supportive and collaborative atmosphere on site. "IPMI has always been for the conference with the most interesting exchanges and the most fun!" he adds.

The organizers set aside a few minutes at the beginning of each Q&A session to ensure that **early-career academics can participate and ask questions ahead of longer-standing professors**. This concept dates back to the Jewish Sanhedrin, 2,000 years ago, when in cases of capital law young scholars were encouraged to speak first so that the opinions of their elders did not unduly influence them, and they were able to gain confidence and form independent views.

"*That's superb,*" Alex responds. "*I wouldn't be surprised if there is Jewish influence in some of the regional members that set up IPMI. There may have been people who had that historical background.*"



Enzo Ferrante

This year's event will host three keynote speakers: **Jong Chul Ye** from **Korea Advanced Institute of Science and Technology (KAIST)** will speak about diffusion models for inverse problems in medical imaging; **Petar Veličković** from **DeepMind** will deliver an exciting and methodologically-focused talk on geometric deep learning; and **Gitta Kutyniok** from **Ludwig Maximilian University of Munich** will be discussing the foundations of deep learning. Enzo points out that "*Bringing internationally reknown scientists to our country is very motivating. Some may even give seminars in Buenos Aires…*"

What about the local scene? Enzo knows all about it: "*IPMI will be held in LatAm for the first time. It was very exciting to hear Alex and Demian propose to hold it in Argentina! This will happen in* ***Patagonia****!*" Bariloche is an idyllic city in the foothills of the Andes, the heart of Argentina's

Demian Wassermann

ski resorts, as the seasons change from autumn to winter. Although it won't be as cold as it can get, there will likely be snow. Staging it in a remote location encourages participants **to interact and build a sense of community**. There will be traditional social activities, including a choir and a football match, and visitors will have the chance to explore a beautiful city surrounded by picturesque forests, lakes, and mountains.

"*There's a huge amount of culture and nature to admire,*" Alex adds. "*It's an area where, for obvious reasons of looking like the Alps, there has historically been a big community of Swiss people. There are many regional delicatessens and a specific type of chocolate is made in the area.*"

Born in Argentina, Alex lived there for 18 years. He advises visitors to make the most of their trip by spending a day or two in Buenos Aires on the way to Bariloche to get a taste of the capital. With 10 million people, it's one of the largest cities in Latin America.

"***It was a dream for the three of us to make this happen in our country,***" he reveals. "*I wanted people to experience and admire the country in which I was born and to which I'm highly in debt for so many things, but I also wanted to bring the conference closer to the people in Argentina, Chile, Brazil, and Latin America in general, who have more challenges attending events outside Argentina.*"

Does Alex ever worry that the size of the event may mean people who want to attend will be left out? Organizers considered holding a hybrid meeting but ultimately decided it would detract from the personal experience and retreat-like atmosphere.

"*Of course, at one level, with a small attendance, you leave people out,*" he responds. "*But what reassures us is that we don't choose who comes to the conference. The secured ticket to get in is submitting a paper and having it accepted. We register all the people with accepted papers first, then open general registration. There is a risk that IPMI isn't as visible as other conferences, but we spend plenty of time on preparation and publicity to maximize*

its exposure."

Alex is joined as General Chair by Marleen de Bruijne. Taking on such a prestigious role at the helm of an international conference is no easy task, especially when you are already a busy professor and proactive community leader. We must ask: What entices him to give more of himself to the community?

"*I studied at Utrecht University with **Max Viergever**, **Wiro Niessen**, and many other colleagues, like **Erik Meijering**, **Bram van Ginneken**, **Josien Pluim**, and Marleen,*" he recalls. "*We've received a huge amount of benefit from the community. With 20 years of distance, you realize just how lucky we were to be doing our PhDs together in big laboratories at a time when **medical image computing** was starting to come out from being seen as purely computer vision and having a corpus of its own. It focused on translation but without removing the core methodological components. That wasn't the situation for everyone. Many people worked in laboratories with a different approach. I feel we're responsible for maximizing the impact of those benefits for others in any way we can.*"

There is still time to register for IPMI 2023. For anyone unable to attend, conference proceedings will be published in Lecture Notes in Computer Science.

TML4H: ICLR WORKSHOP ON TRUSTWORTHY MACHINE LEARNING FOR HEALTHCARE

Hao Chen is an Assistant Professor at the Department of Computer Science and Engineering, the Hong Kong University of Science and Technology. He leads the Smart Lab focusing on AI in healthcare and serves as an Associate Director in Center of Medical Imaging and Analysis, HKUST. He has also organized the very first ICLR 2023 workshop on Trustworthy Machine Learning for Healthcare (TML4H) this year. Luyang Luo, the co-organizer of TML4H and also a postdoctoral fellow from the Smart lab, is here to share the recent news on the workshop and what to expect afterwards.

Hao Chen

Luyang Luo

Machine learning (ML) has achieved or even exceeded human performance in many healthcare tasks, owing to the fast development of ML techniques and the growing scale of medical data. However, ML techniques are still far from being widely applied in practice. Real-world scenarios are far more complex, and ML is often faced with challenges in its trustworthiness such as lack of explainability, generalization, fairness, privacy, etc. **Improving the credibility of machine learning is hence of great importance to enhance the trust and confidence of doctors and patients in using the related techniques.**

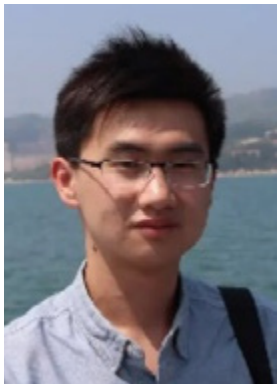
"*This is the key reason why we organize this workshop*," Luyang said, "*We aim to bring together researchers from interdisciplinary fields, **including but not limited to machine learning, clinical research, and medical imaging, etc.**, to provide different perspectives on how to develop trustworthy ML algorithms to accelerate the landing of ML in healthcare. This can also be seen by the diversity of our program committee: we have Yuyin Zhou, Jing Qin, Yueming Jin, Xi Wang, Pheng-Ann Heng, Danny Z. Chen, and Tim Kwang Ting Cheng from the Computer Science field; Daguang Xu and Le Lu from the industry; Vince Varut Vardhanabhuti and Marius George Linguraru from the clinical institute; as well as Jiguang Wang and Xin Wang from the life science field.*"

The workshop has received in total 31 valid submissions and finally accepted 18 papers, with 10 short orals and 8 long orals. "*Holding the very first TML4H workshop, we are surprised and also excited to receive such a large number of submissions. It revealed how much attention the research community is paying on this topic.*"

The workshop began with the keynote entitled "**Trustworthy Machine Learning in Medical Imaging**" by **Lena Maier-Hein** from DKFZ in Heidelberg (Germany). Apart


Lena Maier-Hein

from developing trustworthy machine learning algorithms, Lena also raised the question of a lack of standardization in the field of medical image analysis, **which impeded successful adoption of modern ML research into clinical use**. She further discussed pervasive shortcomings in


Minghao Cheng


Huazhu Fu


Shandong Wu

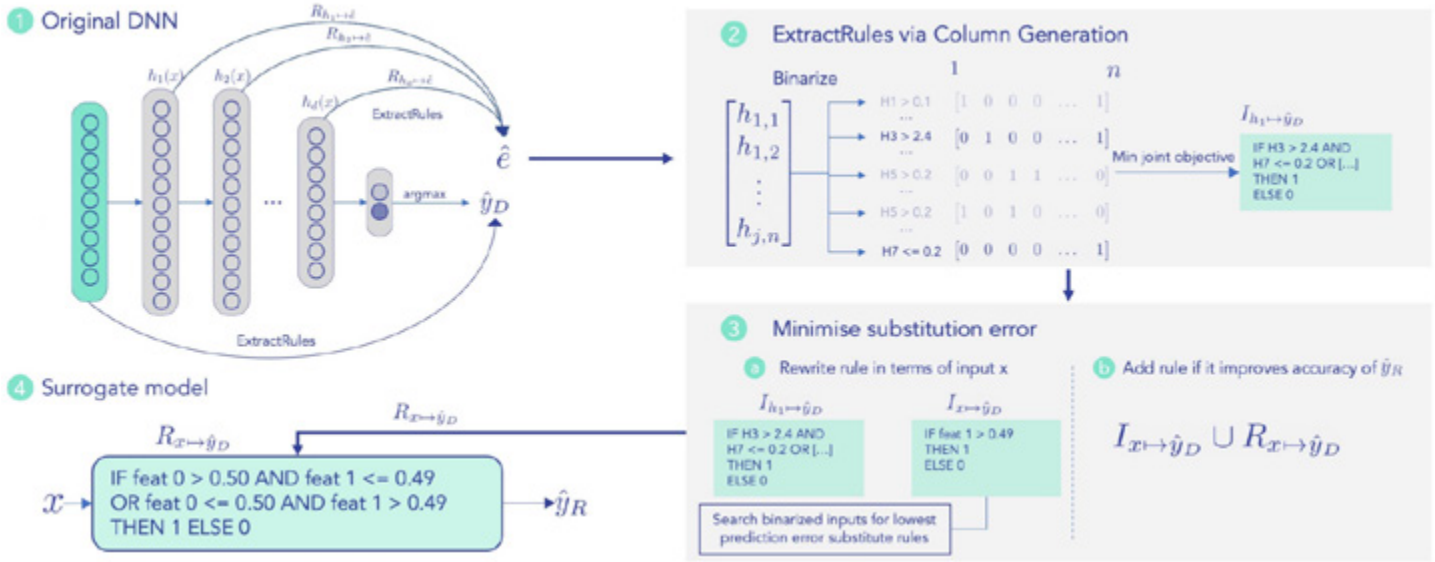

Georgios Kaissis


Xiaoxiao Li

current medical imaging procedures and presented possible solutions.

"*Standardized and transparent evaluation is one of the keys to successfully deploying any healthcare solutions into practice, to make sure that an AI product is robust, fair, and trustworthy for clinical use.*" Luyang told us.

The workshop also has five other invited talks. **Minghao Cheng** shared their effort on visualizing a global explanation in the input space for every class learned in the training procedure by his talk "Generating Class-wise Visual Explanations for Deep Neural Networks". **Huazhu Fu** introduced how to develop reliable and robust machine learning algorithms with uncertainty evaluation in the talk "**Safely Utilizing AI Model in Open Clinical Environment**". **Shandong Wu** discussed the trustworthiness of AI technology from a more clinical perspective in the talk "**Trustworthy Medical AI in the Loop of Algorithm and Clinic**", where he emphasized that the AI algorithms and clinic needs mutually fostering each other. **Georgios Kaissis** shared their works on protecting data privacy by "**Unlocking the Potential of Differential Privacy in Medical Imaging: Enabling Data Analysis while Protecting Patient Privacy**". Meanwhile, **Xiaoxiao Li**, who has also been working on privacy-preserving machine learning, discussed their effort on federated learning with the talk "**Overcoming Data Heterogeneity**

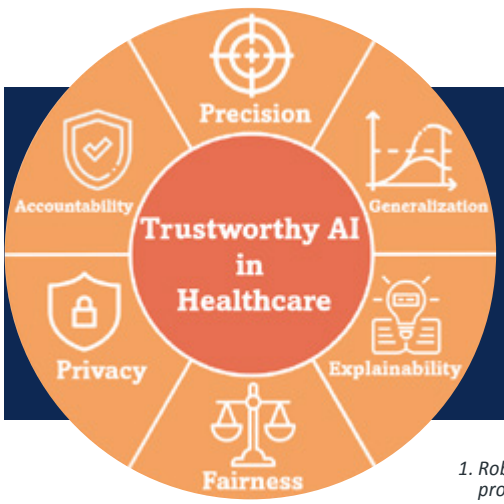


Challenges in Federated Learning".

"The discussion in this workshop was versatile." Luyang said, *"Apart from the keynote and the invited talks, the oral presentations covered various aspects of trustworthiness, such as **explainability, privacy, cross-domain robustness**, etc. Many papers are related to explainability, and the best paper selected by the program committee is also on this topic. Other papers also impressed us a lot, from which we selected two best paper honorable mentions."*

*"What distinguishes our workshop from others is that we are trying to draw the attention of the research community on the important topic of trustworthy ML. We are indeed facing such a challenge. For example, a paper published at the Nature Machine Intelligence on 2021 reviewed 62 studies developing AI models for COVID-19 analysis and found that **none of them are of clinical usage due to methodological flaws and/or underlying biases**[1]. This strongly motivated this workshop to share, discuss, and explore trustworthy machine learning algorithms for a more practical clinic use in the future."*

Also, as part of the post-workshop activities, Hao and Luyang are co-organizing a special issue on the **IEEE Journal of Biomedical and Health Informatics (J-BHI)** entitled "**Trustworthy Machine Learning for Health Informatics**"."*The 1st TML4H workshop is just a beginning, and we look forward to more fruitful studies on the special issue as well as the next workshop!"*

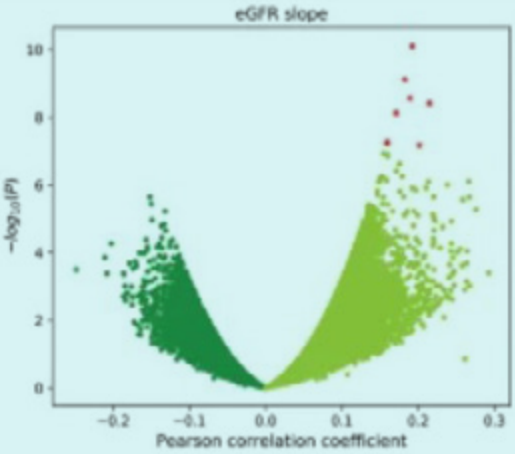


1. Roberts, Michael, et al. "Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans." Nature Machine Intelligence 3.3 (2021): 199-217.

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

New Algorithm can Predict Diabetic Kidney Disease

Researchers from Sanford Burnham Prebys and the Chinese University of Hong Kong have developed a computational approach **to predict whether a person with type 2 diabetes will develop kidney disease**, a frequent and dangerous complication of diabetes. They published their work in **Nature Communications**. This is expected to help clinicians determine who is at greatest risk of developing diabetic kidney disease and optimize the treatment of type 2 diabetes to prevent kidney disease. The study used a precious dataset of 1,200 patients in the **Hong Kong Diabetes Register**, which are followed up for many years. Read More



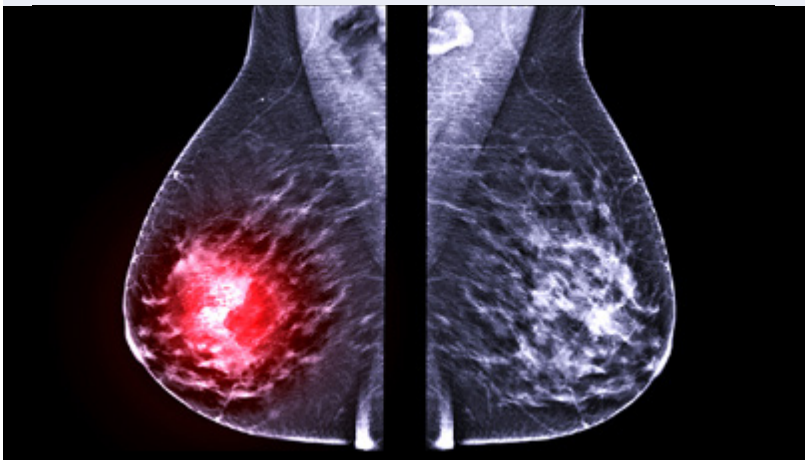
ESOC 2023
THE VOICE OF STROKE IN EUROPE

AI Tool Outperforms Human Emergency Call Handlers in Identifying Stroke

We know this story and the previous one thanks to the folks at **Digital Health News**. Many times, **emergency call handlers** do a fantastic job and save human lives. At other times, alas, the response is not as good. It is therefore very useful to help these guys take the best decision, for instance in identifying **stroke cases**. Researchers from Denmark did just that, working on a dataset of over 1.5 million calls made to the **Copenhagen Emergency Medical Services**. They have developed **a new AI framework** to find out many strokes that go unrecognized by human call handlers and have just presented it at ESOC. Read More

Standalone AI for Breast Cancer Detection at Screening Digital Mammography

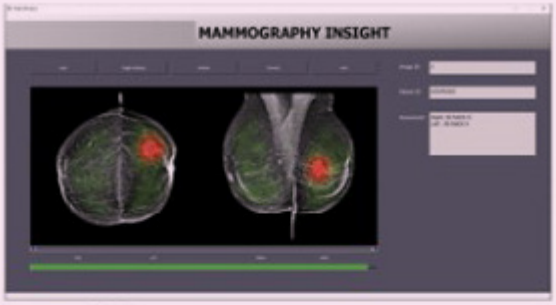
Many studies investigate the potential use of Artificial Intelligence systems in **mammographic screening**. However, what do we know about the performance of AI? Can it become a modality used for **independent mammographic interpretation**? A new study published on **Radiology** only a few days ago found out that standalone AI for screening digital mammography performed as well as or better than radiologists. The research was conducted on databases of studies published from January 2017 to June 2022, that included a total of almost 500,000 women. Read More

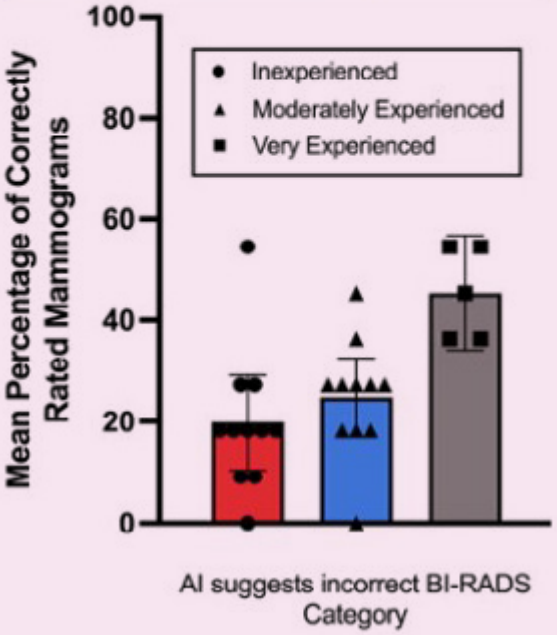


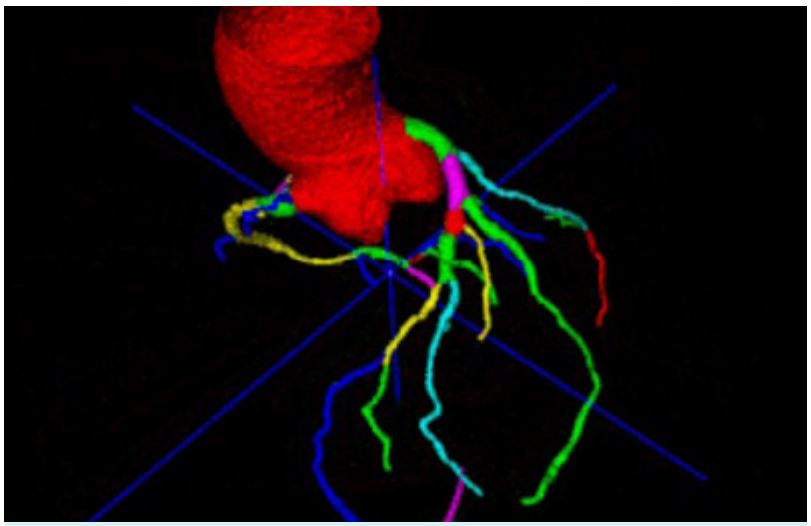
Automation Bias in Breast AI

Another interesting story from **Radiology**. It says, you guessed it, that automation can be a game-changer in **breast imaging**, both in quantity (volume of cases examined) and quality (providing an efficient diagnosis). What is the **automation bias** then? A group of researchers from Germany and the Netherlands found that inexperienced radiologists tend to rely more heavily on the automated system versus moderately and highly experienced radiologists. The latter were more disposed to **challenge / verify the AI automated response**. The result is that the automation bias may lead to missed breast cancer diagnosis and unnecessary recalls or biopsies.

Read More



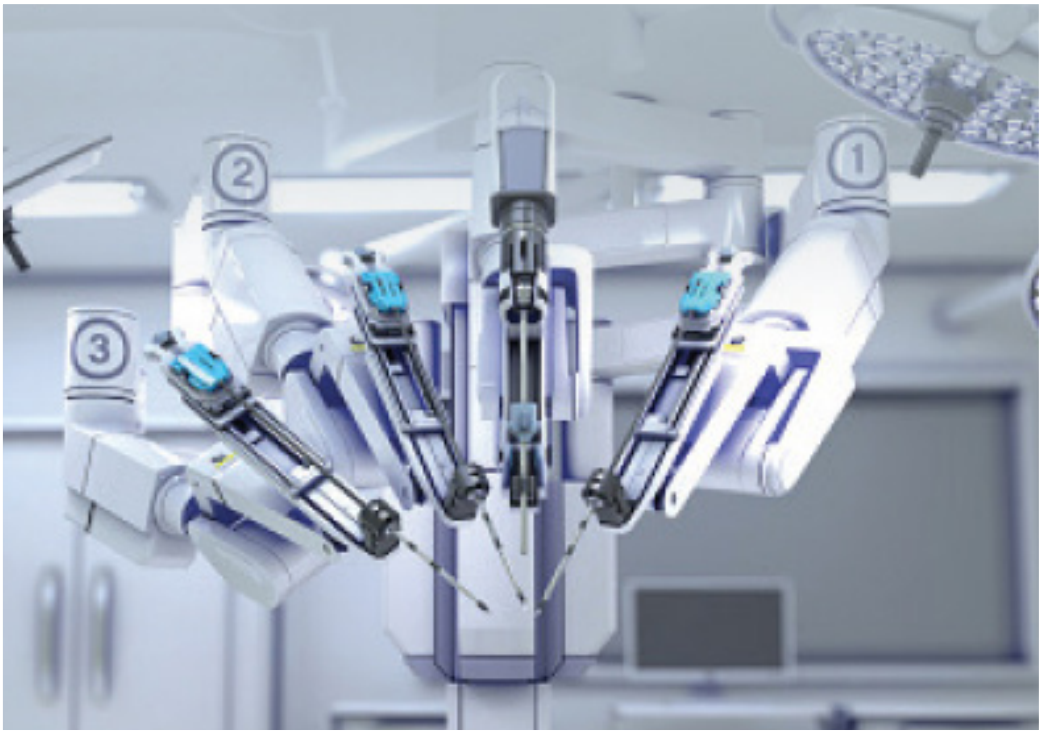




Cardiology Has Embraced AI More than Most Other Specialties

This fine article informs us that **Artificial intelligence algorithms** are being used more and more by **cardiologists and other cardiovascular professionals**. How do they know? Apparently, of the more than 500 clinical AI algorithms cleared between January and March 2023 by the **U.S. FDA**, cardiology has received 58 and only one other specialty has more: guess which one that is? Yes, radiology has almost 400! **RSIP Vision does exciting work in R&D of AI for cardiology, you can check that here**. One of the experts interviewed in the article suggests the inspiring term of "**collaborative intelligence**" for AI in the medical field. Read More

COME AND MEET US AT MICCAI 2023 IN VANCOUVER, CANADA!



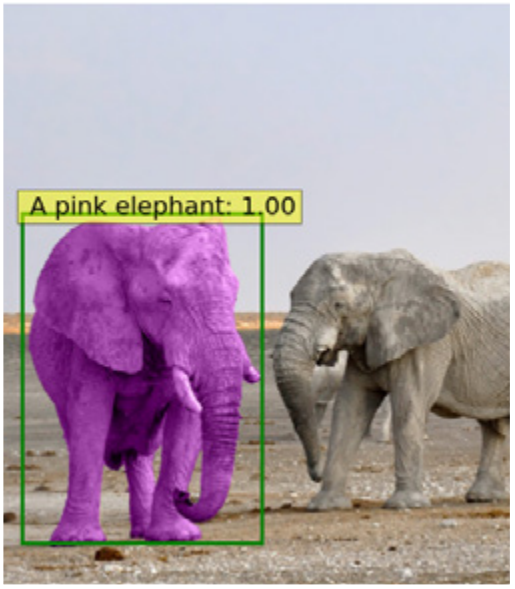
IMPROVE YOUR VISION WITH

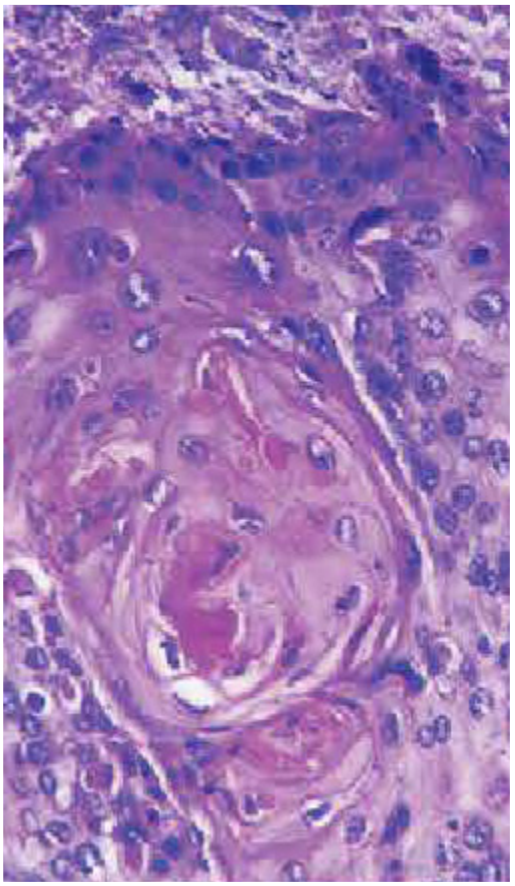
Computer Vision News

SUBSCRIBE

to the magazine of the algorithm community

and get also the new supplement Medical Imaging News!





RSIP VISION