MARCH 2023

Computer Vision News

& Medical Imaging News

The Magazine of the Algorithm Community

nerf2nerf

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

Dear reader,

Welcome to the March issue of **Computer Vision News**, full of all the latest news and updates from our community!

This month, we hear about **nerf2nerf**, a fascinating paper by Lily Goli and Andrea Tagliasacchi accepted at the upcoming robotics conference **ICRA 2023**. It proposes a novel technique for registering neural radiance fields, which could enable many innovative applications in 3D vision.

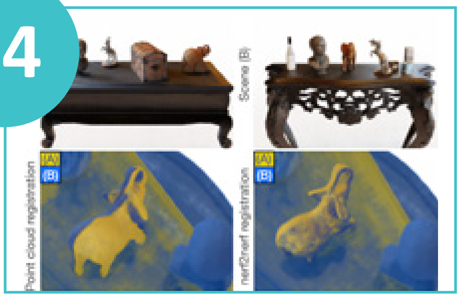
“We’re currently at an intersection where advancements in AI are leading to a new revolution in intelligent imaging.” **Paul Yi** and **Vishwa Parekh** speak to us about bringing their distinct clinical and technical skills together in their new lab at the **University of Maryland**, aiming to improve human health by advancing AI in medical imaging from bench to bedside.

Meanwhile, the team here at **RSIP Vision** explore one of AI’s latest surgical breakthroughs – the **real-time tracking of soft tissue during brain surgery**. How does AI provide a better solution to this important challenge?

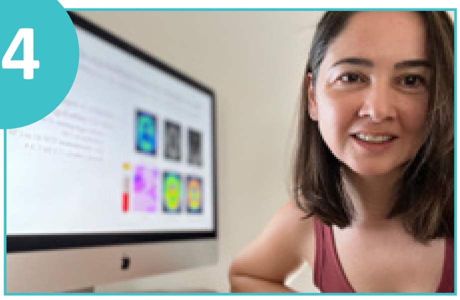
I’ve only given you a taster of the exciting content we’ve prepared for you this month, so I won’t keep you any longer. Happy browsing, and remember to spread the word about our free subscription to your friends and colleagues!

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

Computer Vision News

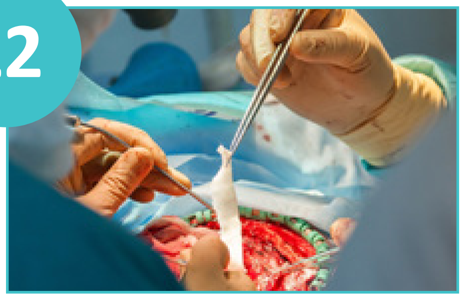
04


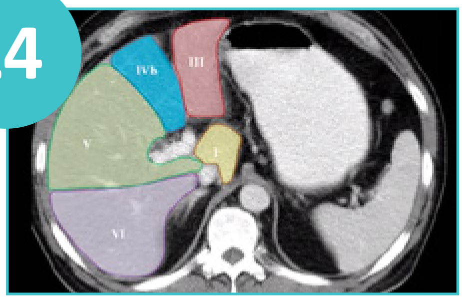
10

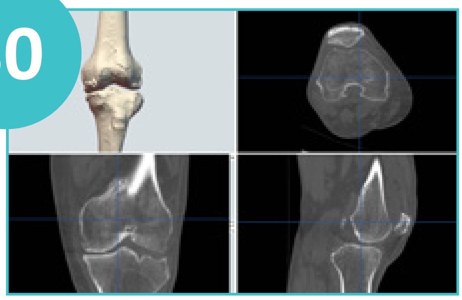

14


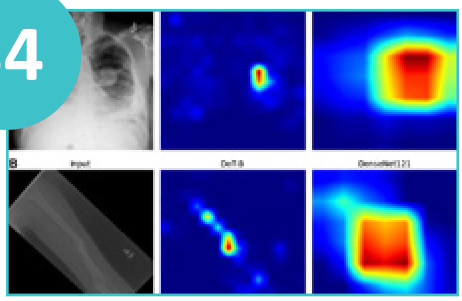
20


Medical Imaging News

22


24


30


34


NERF2NERF: PAIRWISE REGISTRATION OF NEURAL RADIANCE FIELDS

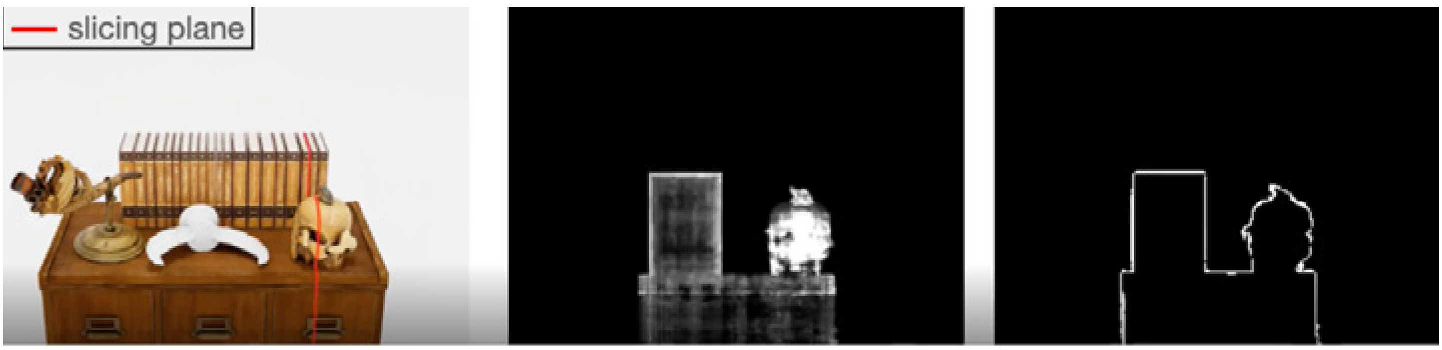
Lily Goli is a second-year PhD student at the University of Toronto, co-supervised by Andrea Tagliasacchi and Alec Jacobson.

Her paper proposing a new technique for registering neural radiance fields has been accepted at ICRA 2023, an upcoming robotics conference in London. Lily is here to tell us what it is all about.

With the growing popularity of **neural radiance fields (NeRFs)** as 3D scene representations, there are an increasing number of datasets and databases of NeRFs from different scenes which need to be processed, as is already the case for 2D images.

Registration is a technique widely used in image processing, particularly in medical applications for detecting longitudinal changes in images, such as MRIs. It has also been utilized in robotics for bundle adjustment, city planning, and coordination between multiple robots.

In 2D, you can easily perform a basic operation of **rigid registration between two images**, making the same objects in both images fit together by aligning the two with some rotation and translation. This work aims to achieve the same functionality on NeRFs. Where there are two 3D scenes with different lighting and backgrounds but



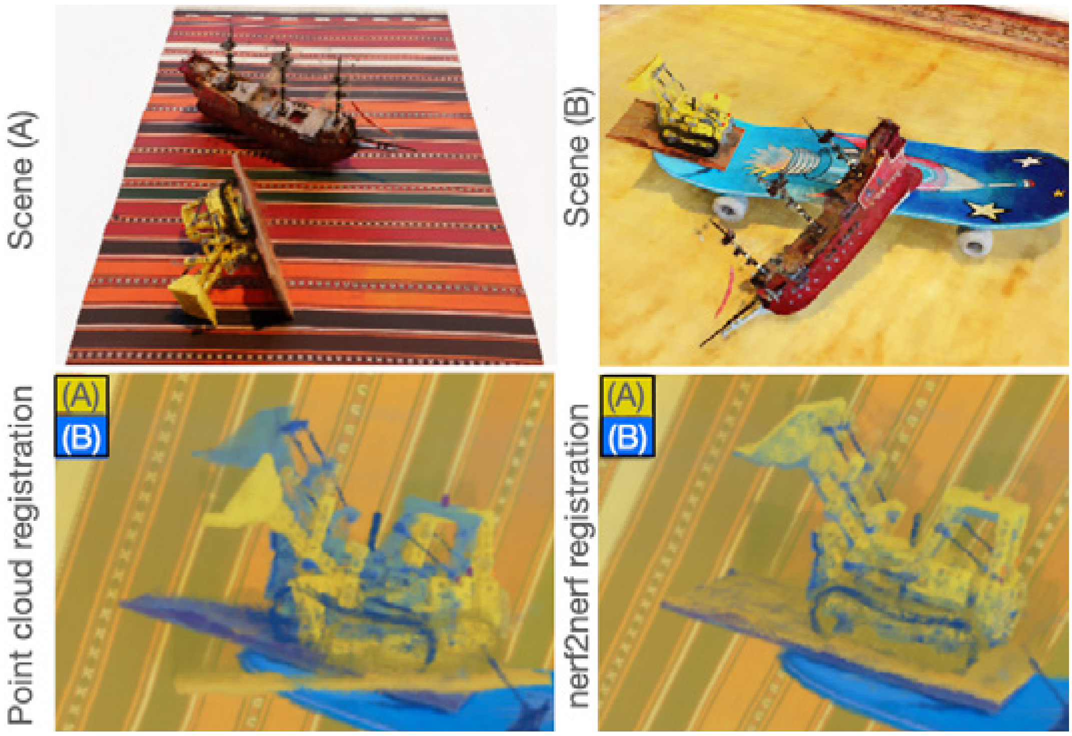
NeRF Scene A 2D Slice in Density Field A 2D Slice in Surface Field

the possibility of the same instance of an object, **the goal is to find the perfect rigid transformation to align these objects**.

However, this is more challenging as **NeRFs are implicit representations** without explicit access to the radiance, color, or other geometry used to compare images. Therefore, Lily and the team had to devise creative methods for performing this registration.

"The main challenge here is that NeRFs bake in the appearance of the objects inside their geometry, so ***there is no clear decoupling of appearance and geometry,*** *"* she explains. *"We wanted this to work between any two NeRFs that have the same object in them, so with different lighting or at a different time of the day, the appearance of the object, the color of it, might change. A cup, for example, in daylight might be very white, but at night, it might be gray, so we can't use the radiance colors outputted from NeRF directly to compare two points on the two NeRFs to see if they match."*

Instead, Lily had to **extract a representation that showed the object's geometry to compare it**. Geometry is reliable because it does not change with different lighting and is not view-dependent, whereas color is completely view-dependent. She extracted a new field called a surface field from NeRF, which can be performed as **a complete post-processing step without altering the NeRF pipeline**. Surface fields give the probability that each point in a space is located on the surface of an object. This approach offers a clean domain with zero or one values describing the scene's geometry.



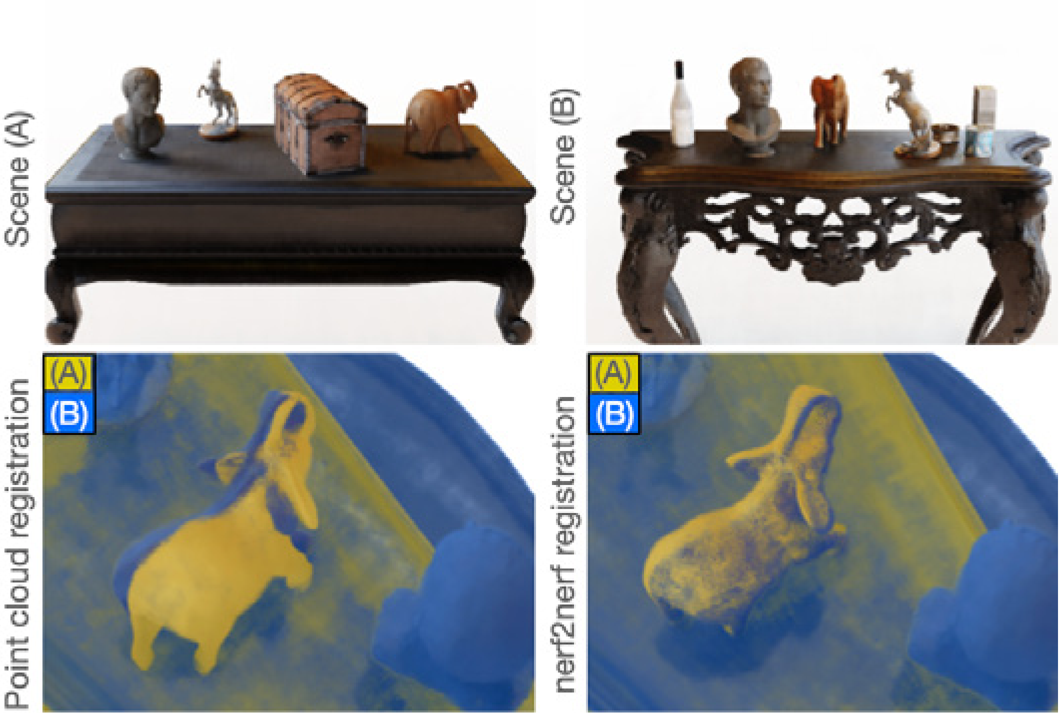
"Now that we had the domain to compare the two NeRFs against, we could sample points in the two NeRFs near the object we wanted to register and then compare the pointwise surface values," Lily explains. *"To perform the classic optimization loop for rigid transformation, very similar to ICP, **we compare the surface value of the sample pairs and optimize their rotation and translation to make these match iteratively.**"*

nerf2nerf is a high-level tool that could be used in different applications – bundle adjustment, SLAM, or structure from motion, which are hot topics in robotics right now. As 3D vision is likely to replace 2D vision in robotics, people are finding novel ways to incorporate NeRFs or other neural field domains into **3D vision for robotics**. Lily hopes to see registration being used as a first step to more complicated robotics algorithms. She submitted the work to **ICRA 2023**, a robotics conference, with these goals in mind.

*"Something I want to achieve for my thesis and PhD is to **collect assets from the real world**,"* she tells us. *"If we have a NeRF of a cup on a table, I will see the complete 360 degrees of the cup, but not the underneath. If I make a model from that one NeRF, it will always miss the bottom part. Although there are ways to find this bottom part, through inpainting and generative models, they're not 100% accurate."*

To solve this, Lily came up with the idea of **making another NeRF of the cup from a different angle and registering it with the original NeRF** to obtain complete information about the object. Having completed the registration step, she is now working on **combining the two NeRFs to extract the final model**.

Thinking about a medical scenario, is it possible to envision a scenario where



medical teams perform an ultrasound during an intervention and use NeRF technology to register the real-time ultrasound with a pre-intervention CT or MRI, allowing them to make decisions during the operation?

"It could potentially exist," Lily muses. *"NeRFs are very fast now and can be trained with 3D or sparse 3D information, and then you can register them. However, I've only performed rigid transformations, which is a good baseline for medical applications, but deformable registration is usually needed. Some works are pointing towards that, so it should be the next step in NeRF registration for medical purposes."*

Before arriving in Canada only a couple of years ago, Lily was a bachelor's student in Iran. She graduated with high honors from the **Sharif University of Technology** and then took a direct PhD.

"This was my first project as a PhD student, so the adjustment was hard for me," she recalls. *"**Andrea [Tagliasacchi]** was very helpful. He suggested a few problems to see which one I liked, and I thought this was the coolest one. He guided me through every step, gave me many resources to start my knowledge on NeRFs, and did a personal reading course for another student and me. He's been very supportive through all of this. He also connected me with the second author of this work, **Daniel Rebain**, a brilliant student and a senior in doing NeRF stuff, so he helped collect the data sets."*

As part of her bachelor's, Lily did two internships, working with some of the most eminent researchers in the field. While at the **Technical University of Munich**, supervised by Nassir Navab, her work was accepted at **MICCAI 2021**. She also worked on medical imaging at the **University of British Columbia** with Purang Abolmaesumi.

"Purang is a very kind-hearted professor," Lily reveals. *"He takes his lab to lunch every week, which makes the whole team much closer. Having experienced PhD life now, I can see that's rare. He has a very friendly group and creates that atmosphere himself."*

Register now for ICRA 2023 at ExCeL London from 29 May to learn more about Lily's work and many more.

TRAINING WITH MALICIOUS TEACHERS: POISONING ATTACKS AGAINST MACHINE LEARNING

Antonio Emanuele Cinà has recently finished his PhD at Ca' Foscari University of Venice. His research interests include ML security, system reliability, and interpretability. Recently, he has focused on studying ML vulnerabilities during training to categorize these risks and define guidelines for developing secure models. Antonio is now a Postdoctoral Researcher at CISPA – Helmholtz Center for Information Security, Saarbrücken. Congrats, Doctor Antonio!

Machine Learning (ML) models are nowadays becoming the de facto standard for diverse and sensitive tasks, such as cancer detection, malware detection, and road sign recognition, as vital tools for data analysis and autonomic decision-making.

The key strength of learning models is their ability to infer patterns from data that can be used for future predictions.

Nevertheless, ML has traditionally been developed under the assumption that the environment is benign during training and usage of the model. These assumptions have helped design efficacious ML models but do not cover cases where malicious users try to alter this condition to reach their goal. Therefore, the increasing pervasiveness of ML in critical applications poses an issue about their robustness in the presence of malicious manipulations. For example, in 2005, Nelson et al. **[Nelson08SF]** showed that malicious users could evade ML-based spam email filters by appending words indicative of legitimate email. This scenario is an instance of a ML security threat called **data poisoning**. Under this setting, malicious users may cause failures in ML systems (e.g., spam filters) by tampering with their training data, thereby posing real concerns about their trustworthiness.

Fig.1 depicts the influence of a poisoning attack on a ML model decision boundary. Specifically, compared to the pristine model, the poisoned one has a different decision boundary that now meets the attacker's goal.

Antonio's work aimed at shedding light on existing types of poisoning attacks, categorizing them with respect to their assumptions and attack methodologies. He then investigated four more aspects of poisoning attacks, namely: (i) their scalability issues, (ii) the factors affecting the vulnerability of ML models, and (iii) their effectiveness when limiting the attacker's knowledge.

He also paved the way toward a novel kind of security violation caused by poisoning attacks, i.e., energy-latency attacks. The resulting attack, i.e., **sponge poisoning**, increases energy consumption and latency of the victim model at inference time **[Cina2022SP]**.

Finally, Antonio identified the relevant open challenges limiting the advancement of the poisoning literature, i.e., missing scalable and effective attacks and lack of attacks for realistic threat models, together with reasonable research directions that can tackle them **[Cina2022SV]**.

In conclusion, Antonio's contribution clarifies what threats an ML system may encounter when malicious users influence part of the training pipeline and establishes guidelines for developing more reliable models against this threat.

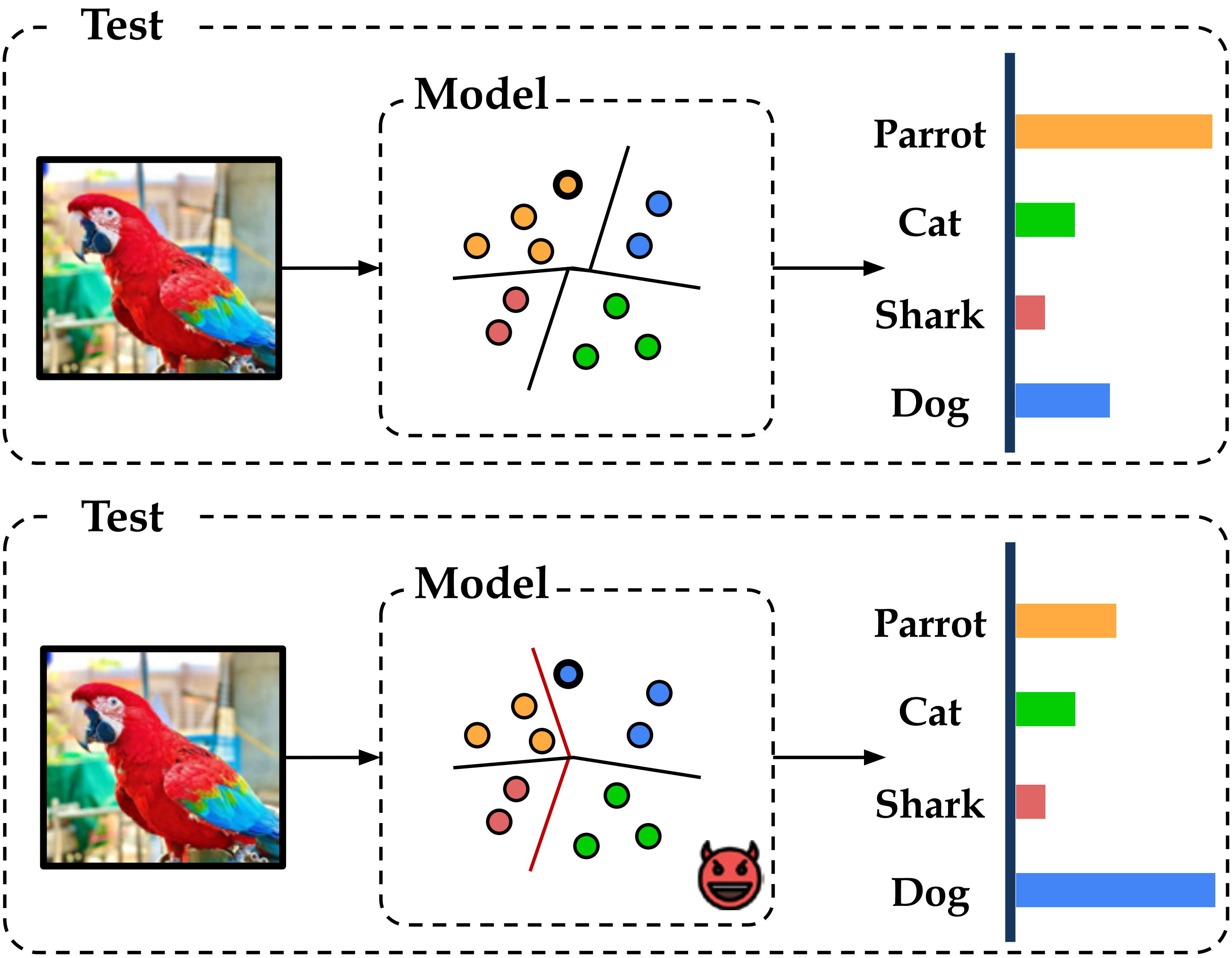


Figure 1. Example of a poisoning attack on the animal classifier. On the left, a clean model that behaves as expected, i.e., it correctly classifies the input image. On the right, the poisoned ML model misclassifies the pristine input as desired by the attacker. The red decision boundary indicates how the model has changed after being targeted by poisoning.

AI ACCELERATOR INSTITUTE

Tim Mitchell is the Community and Events Manager at AI Accelerator Institute, an alliance of leading innovators in the AI ecosystem dedicated to developing next-gen machine intelligence. He is here to tell us more about its journey and upcoming series of in-person summits.

Tim Mitchell

AI Accelerator Institute (AIAI) provides a global platform for AI practitioners to connect and find the resources and events they need to advance in industrial AI and deploy and scale AI-driven business solutions. It is hosting a worldwide series of in-person summits starting **in Tel Aviv on 23 March**, then **San Jose in April**, and **Berlin**, **Boston**, and **London** later in the year.

AIAI had previously run in-person events in limited US locations but developed a thriving online presence during the pandemic, encouraging it to expand its reach internationally.

"The pandemic allowed us to focus on building our online community and growing some aspects we hadn't considered before," Tim tells us. *"We've now got virtual events, podcasts, e-books, courses, and many more strings to our bow. Bringing back in-person events after Covid means we've got a global base in many areas to expand this series of summits."*

The summits attract a mix of technical teams and business leaders, including AI/ML/CV engineers, data scientists, and decision-makers, such as heads of R&D, innovation, and emerging technologies.

"We're not the type of event people will have attended before," Tim declares. *"There are hundreds of AI events, but they're either academic-focused or filled with the wrong people to hold business conversations. We're a very focused group of **business and technical leaders** with an almost exclusively end-user-led line-up. We don't sell many speaking slots, so it's not a day of sales pitches. That's not what we're about. We want people to be there for knowledge and strategy-sharing purposes. It's a great place to network and make connections."*

The summits showcase leading applications and cutting-edge solutions, demonstrating how organizations already use them. They include small expo areas where attendees can learn more and meet the people behind the scenes.

Having received positive feedback from

the three events AIAI held in 2022, Tim feels confident in the series' success this year.

"We've been doing this for several years now, but to be honest, we were nervous coming back from the pandemic," he reveals. *"It was two years since we'd run an in-person event, and there are always a few nerves, but all three shows went brilliantly. We were pleased with the turnout, the quality, and the feedback, and we wouldn't be expanding the series if we weren't confident."*

Regarding quality, Tim tells us it comes down to a focus on actionable outcomes, ensuring participants leave armed with something they can immediately apply to their work.

"One of my favorite bits of feedback we got last year was from a senior director at Shutterstock who said there was a clear focus on outcomes," he recalls. *"That's what quality means. You're left with something actionable in your business that you didn't know or have access to at the start of the day."*

The choice of Tel Aviv as the starting point of this year's series is driven by the exciting **computer vision landscape in the region** and its burgeoning **start-up scene**.

"The Israeli landscape is growing hugely and will continue for the next few years, so we wanted to share what we can bring to the audience there," Tim says. *"Hopefully, people are going to enjoy what we put together."*

The events cover **autonomous vehicles, healthcare, banking, manufacturing, and retail**, although these fields vary by location. Tel Aviv has a start-up and innovation focus, while San Jose has a more established market and landscape, so the agenda is slightly different.

"I tend to give key themes I want the day to explore," Tim clarifies. *"It's about the state of the industry, the technologies being used, and applications being facilitated. That's a nice story to walk through and leaves people feeling up to date with the state-of-the-art. They'll be surprised at how effectively industries use these technologies in ways they won't have considered."*

The first event opens with a keynote from **Arne Stoschek**, Project Executive at **Acubed by Airbus**, speaking about **the autonomous landing of its A380 airliner**, and features other companies that may be unexpected in this domain, such as **Lego** and **Nike**.

"It's going to open people's eyes to solutions," Tim continues. *"They'll be asking, how can I use that in my business? How can I use that on this project? They'll be able to find those methods and solutions."*

As for the future, can we expect more shows and new destinations in 2024?

"Absolutely!" Tim responds. *"We've got a very exciting Indian location on the horizon, so if things go well, we'll be in Bangalore next year. There are plenty more places for us to explore."*

Duygu Tosun-Turgut is currently an Associate Professor of Radiology and Biomedical Imaging at the University of California, San Francisco. Duygu is also the current chair of Alliance of Women Alzheimer's Researchers (AWARE), part of Alzheimer's Association.

Duygu, can you tell us about your current position?

I actually got promoted to a full professorship, but it's not going to be active until a couple more months. I'm also a Director of Medical Imaging Informatics and Artificial Intelligence in our department.

Congratulations on the promotion! What work did you do to deserve it?

I'm trained as an electrical engineer, and my expertise is in medical imaging using multimodal, different medical imaging modalities, extracting information - using these modalities to better understand diseases, better diagnose diseases, and better monitor diseases. I mostly specialize in Alzheimer's disease and psychiatric disorders like depression. It's hard to tell what I did to deserve this, but I think it's years of work accumulating, and my research has been dedicated to imaging biomarker discovery. One of the things I really focus on is making biomarkers accessible and scalable for everyone, especially in Alzheimer's disease. Not like having state-of-the-art biomarkers, which require special equipment or expensive imaging technology, so it can only be accessed by certain countries or by certain specialized centers. Biomarkers have been playing a huge role in us defining what is really Alzheimer's disease and what kind of pathological changes are happening. In the past ten years, my research has been dedicated to developing computational approaches to discover imaging biomarkers that can be used in clinical trials to identify participants and also to track how they're responding to treatment.

Does this mean that we can detect Alzheimer's disease and chronical depression in medical images or in blood test?

Both things are happening right now. We do have very special molecular imaging approaches. These are positron emission tomography (PET) approaches where we have specially engineered or designed radiotracers that specifically bind pathologies or protein changes in the brain. We've been using that for Alzheimer's disease as well as depression. The way we conceptualize Alzheimer's disease is as a proteinopathy, meaning there are abnormal protein accumulations. And these molecular imaging modalities allow us to detect and see how these proteins accumulate in individuals while they're alive before autopsy. We can identify these individuals very early, even before they start having a cognitive impairment or cognitive decline. Similarly, depression, for example, is conceptualized as a synaptic dysfunction in particular brain regions. We've been using particularly glucose metabolism PET imaging and functional MR imaging to identify the underlying pathophysiological changes that might be related to their symptoms. We're actually slowly seeing it come to clinical practice and clinical trials, not just seeing research ideas or research findings. But some of my work is actually on alternatives to PET imaging, which is extremely expensive. It costs around $5,000 to $7,000 for imaging modality, even in resource settings. It's difficult to find those imaging modalities in every clinical setting. So, we've been focusing on some other imaging approaches: not molecularly specific, but still might serve as a surrogate marker of these functional and structural changes in the brain, like

... science is definitely more accessible, which is great!

MRI imaging, and developing what I call computational biomarkers. It's not a true pathological or molecular biomarker, but it captures how the brain has been affected over the years or how the brain is affected by the insult of these pathologies or syndromes. We interpret, or we

impute from MRI imaging measurements what are the true pathological changes. Again, these are really relying on imaging modalities that are sometimes not easy to access, especially when you think about underrepresented populations or low-income or low-socioeconomic-status regions. The new thing that's very exciting, especially in the Alzheimer's domain, is the plasma biomarkers, the blood biomarkers. Things are really connected: your brain is not in isolation and there is constant interaction with the blood. Even though other things contribute to the composition of the blood, we still see a contribution from the nervous system to the blood. And one really exciting development happening in our domain is the development of specific blood biomarkers for the Alzheimer's pathology. Again, it's in the very early development at the early stages, but it's very promising. If it works, people can get a blood assessment like we get every year for cholesterol, vitamin levels, or glucose level. A similar screening can be used very easily in every clinical setting to detect Alzheimer's pathology related changes.

You told me where we stand today, but you still have decades of research in front of you. Where do you think science will stand when you retire? What will you know, and what will you not know by then?

I think my retirement will be a little far away [*laughs*].

Tell me then about the next ten years.

Technology is developing really fast, compared to the years I did my PhD or my early faculty years. Even today, we are seeing a really fast pace of development in biomarkers and technology. That is a combination of techniques being more advanced, but we're also seeing the effect of big data. There are more efforts where before siloed data sets are now shared! So, creating larger and diverse data sets coming together, allows us to see a bigger picture of how diseases affect individuals. And at the same time, machine learning

It's not our data, it's people's data, and we owe it to science to make it available!

and AI are progressing at a really fast-pace. All these new techniques take advantage of bigger data sets. What I would hope to see when I'm closer to retirement is a better understanding of the disease. Even right now, when we define Alzheimer's, we define it based on amyloid and tau pathologies. There are all these other earlier changes happening in the brain as well as in the entire body, particularly focusing on neuroinflammation or peripheral inflammation, or microglial activation, again related to inflammation. We are still learning a lot about what contributes to Alzheimer's disease. One of the interesting things or one of the avenues we still just started exploring, is how the rest of the body is actually affecting the brain. It's not in isolation: everything you do for the health of your body actually affects the brain. That includes nutrition. That includes exercise or lifestyle or how your body develops at younger ages, or chronic diseases like diabetes or hypertension. All of those have an impact on the brain, how your brain is functioning, and also how the brain reacts to these insults of protein pathology. We are just learning about that. I think that over the years, we will have a bigger understanding of these multi-system interactions. It's not going to be an isolated look at what an amyloid pathology is doing to the brain, or what tau pathology is doing to the brain and try to fix that one at a time. It's going to be a full picture and a whole system approach, a multi-domain approach to understand the disease as well as to treat the disease at the same time. It's happening, and one of the key steps to get there is really bringing data from multiple sources, from multiple cohorts, where we see diversity. Not just in terms of race and ethnicity but also in lifestyle diversity or genetic diversity. That truly gives us a more complete picture of the disease. We really need that collaboration and big data science approaches.

It is the integrity of science that is the most important thing in our careers!

Remind me to check that in a few decades...

Okay, let's see.

This month has been very traumatic for Turkey, shattered by a terrible earthquake. I am also quite sensitive to it, because my roots are at the border between Syria and Turkey. What message would you like to send now?

This is something my husband and I have been discussing. One big challenge right now is actually sheltering. They lost their homes. It's the middle of winter and finding food or other resources has been challenging. There's lots of effort happening there. But providing shelter to these families, these individuals who lost their homes, that's challenging. We're looking into if there are mobile containers or really protective tents that can be deployed and distributed to them. It's temporary. It's mobile, but it gives them some shelter. I think that would have a huge impact if we can figure out how to deploy those to them.

With all the things that you have learned in science, can you teach us one lesson that particularly impressed you?

That is a tough one. You're right, I have had different mentors over the years. I still have my mentors, because you definitely need mentors at every stage of your career. One thing that stuck with me, and one thing that I try to teach my students, goes back to my PhD years. It wasn't like open, dictated learning. It's observational learning. Showing me that it is the integrity of science that is the most important thing in our careers! It's not the number of publications or the number of presentations you're giving. It's really the rigor and the integrity of your work that's going to carry you forward. That's what everyone is going to remember and appreciate, and that's going to make a difference. So, I always keep telling my students that, rather than rushing, it is really important to understand all the details. But, also ask yourself: why are you doing it? It's not just for another paper in your CV. It needs to contribute to others' work as well as to your overarching goal. In my case, it's clinical research, clinical trials. I do enjoy seeing the impact of the approaches and techniques we've been developing in clinical trials. So, I'm just stepping back and questioning why we're doing this.

… And how do you rate the integrity of science now?

[*laughs*] It's more accessible. Compared to my PhD years, science is definitely more accessible, which is great! So, if you're interested in the topic, you can always find something to read or learn how it is done. You can watch a YouTube video and everything. Also, data is more accessible because we really advocate for open data sharing: it's not our data, it's people's data, and we owe it to science to make it available! There is also more production happening both in terms of papers and new projects and at a very fast pace. And that pressure in science, especially in academia! At the end of the day, your performance is still mostly assessed based on your product activity in terms of your papers or your grant-funded grants, with everything being more accessible, more people working on similar ideas. It's a bigger competition, and I think that puts lots of pressure and leads to some unfinished ideas being published. Nowadays, it's even more important to be more critical when we are reading things or when we are assessing what's happening in the literature. It's good to see that we are moving forward. But at the same time, the pace is so fast, and people are constantly trying to catch up with everything. Sometimes, things stay on the surface rather than really getting deeper.

It is the integrity of science that is the most important thing in our careers!

If you had the magic wand, what would be one specific thing that you would change in science?

Maybe just figure out a way to reduce funding pressure and produce science, just keeping those a little bit more separate. Right now, though, that goes really hand in hand. Sometimes it just turns into more like an industry rather than a really idealistic scholarly environment. Not every position, not every institute is what we call hard money, right? They don't have a tenure position, so they really rely on grant funding. If we could find a balance on that, rather than pressuring people to constantly producing ideas and selling them…

COMPUTER VISION EVENTS

ACM/IEEE Int. Conf. on Human-Roabot Interaction Stockholm, Sweden 13-16 March	Digital Health Rewired London, UK 14-15 March	Emerging Medtech Summit Dana Point, CA 20-23 March
Computer Vision Summit Tel Aviv, Israel 23 March	Bern Interpretable AI Symposium: Medical Image Analysis Bern, Switzerland 24 March	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		

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!

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!

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

MARCH 2023

SOFT TISSUES TRACKING DURING BRAIN SURGERIES

Most of the tissue handled by the surgeon (or by a robotic hand) during brain surgeries is flexible: With the obvious exception of bones, which are rigid, all the structures touched during the surgery change shape and place. For instance, by changing the force applied to the brain, like when the skull is opened, the pressure on the brain changes and, as a consequence, its shape changes too.

Where did the tumor move to?

It is key for the surgeon to know where each point has moved to: if the tumor that needs to be treated was in one place, where is it now? This is a very important challenge that needs to be solved by the surgical team in real time.

How did surgeons do until now?

Surgeons made assumptions: for instance, they assumed that motion is rigid. This assumption is obviously incorrect, but it was the best available hint to what was going on in the treated region. Of course, the most expert surgeons could do excellent work even with some amount of guessing. The better the human expertise involved, the better the guess - and, generally, the outcome as well.

Unfortunately, the opposite case could also be true: surgeons at the beginning of their career might lack the educated guess of their seniors and, as a consequence, may bring about less positive outcomes.

Can AI provide a better solution?

Very encouraging advances have seen the light in recent years, in particular in the field of AI-assisted image analysis; new technologies now give access to better solutions, providing a precious assistance to the surgical team. Some of these breakthroughs are adding a whole set of possibilities to previously existing classics, like segmentation and feature tracking. In particular, advanced deep learning algorithms can process calculations much faster and offer much more reliable results; they can do that in real time, during the surgery, when supporting information is most needed.

How does it work?

For instance, if the surgical team wants to focus on a specific region, they could start by running a segmentation of that region as a first step, and decide to track what they deem important to be tracked, choosing at will what soft tissues need to be followed. They can even decide to ignore surgical tools and remove them from the field of view as if they weren't obstructing the eye of the surgeon. Key landmarks correspondence is done much better, making it easier to take crucial decisions regarding the physical soft tissues being treated.

These new solutions, thanks to advanced AI algorithms, are much more timely, accurate and reliable than ever before. They are more stable and, as such, more convenient, since they need fewer human corrections, many of which had some level of guessing. When old algorithms did their calculations, these were valid only for some time until the next important change in the region, after which the whole calculating process needed to be started over again. Now it is much easier to keep track even of the soft tissues in the operated region.

OK, what's the next step?

RSIP Vision's engineers are fully trained to develop AI-based soft tissue tracking solutions that are crucial during brain surgeries. **Contact RSIP Vision** to discuss your project.

IOANNIS VALASAKIS, KING'S COLLEGE LONDON

@WIZOFE

Hi everyone! I mentioned last time ChatGPT, didn't I? It takes already the scientific world and not only! Everybody is watching in awe!

Impressive improvements, the usage of search agents, the big fight with Google's Bard... Let us see what the future will bring. Overly exciting times. In other news, in our news, in fact, this month I am going to talk about **a new age in medical image segmentation**. Are you excited? I was myself and while writing this code exploration I discovered some new things, especially in relation to dense layers and how useful they prove!

Let us get into that :)

Introduction

Medical imaging is a crucial tool in modern medicine that helps physicians diagnose and treat a wide range of illnesses. However, interpreting these images can be a challenging task due to the complexity and variability of the human body. One way to assist physicians in analyzing these images is through the use of image segmentation, which involves dividing an image into meaningful regions. In recent years, deep learning models have shown remarkable success in medical image segmentation. In particular, the **DenseNet architecture** has emerged as a powerful tool for image segmentation in the medical field.

The DenseNet architecture is a type of **neural network** that facilitates the flow of information between layers and has demonstrated impressive results in **image classification tasks**. It consists of densely connected layers that allow the network to learn more complex features and improve the efficiency of the training process. This architecture is particularly effective for medical image segmentation, as it can accurately identify regions of interest in complex images.

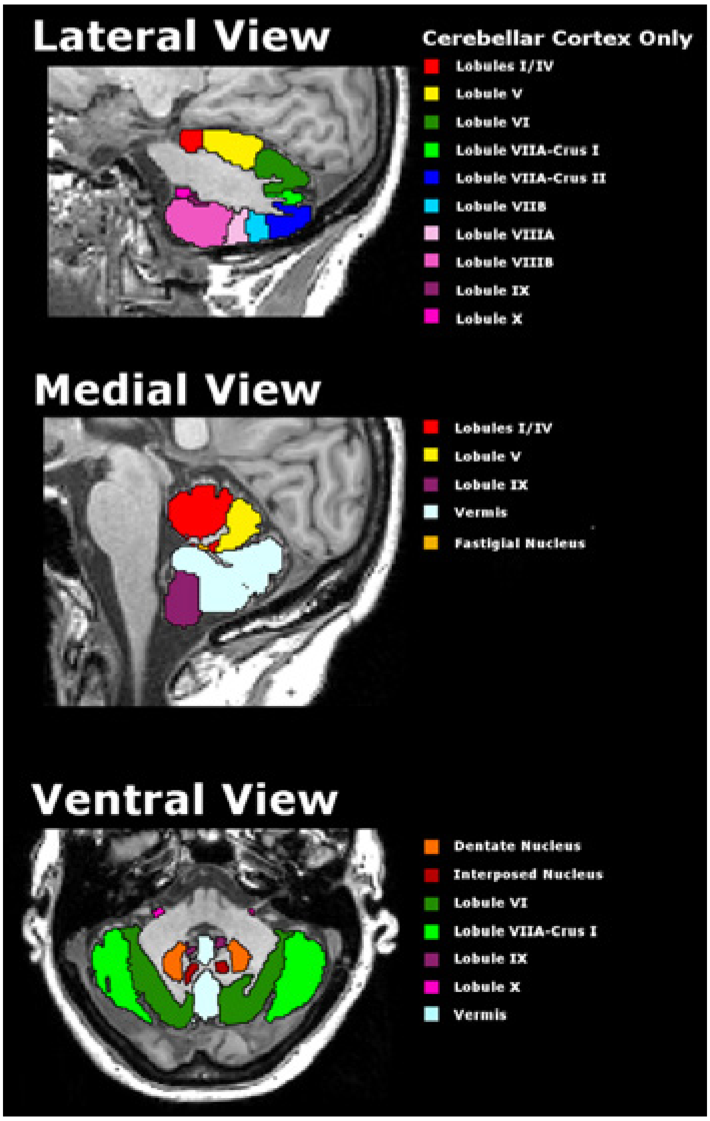
Medical image segmentation using DenseNet has many applications, including tumor detection, identifying and measuring organ sizes, and tracking disease progression. For example, in cancer diagnosis, DenseNet-based image segmentation can help identify the exact location and size of a tumor, enabling physicians to develop targeted treatment plans, as you can see in a detailed segmentation example from the Wikimedia entry in the figure next page.

In conclusion, medical image segmentation using DenseNet has the potential to revolutionize diagnostic precision in the medical field. By automating the segmentation process, physicians can spend less time analyzing images and more time developing treatment plans, resulting in improved patient outcomes. As deep learning models continue to evolve, we can expect even more breakthroughs in medical image analysis in the future.

Approaches

One common approach to using DenseNet for image segmentation is to use a fully convolutional network (FCN). As we talked previously, an FCN takes an image as input and produces a segmentation map, which assigns a class label to each pixel in the image.

Now I will show you an example of how a FCN can be implemented using DenseNet in PyTorch:




```
import torch
import torch.nn as nn
import torch.nn.functional as F

from torchvision.models import densenet121

class DensenetFCN(nn.Module):
  def __init__(self, n_classes):
    super(DensenetFCN, self).__init__()

    # Load pretrained Densenet model
    self.densenet = densenet121(pretrained=True)

    # Replace final classification layer with 1x1 convolution
    n_features = self.densenet.classifier.in_features
    self.densenet.classifier = nn.Conv2d(n_features, n_classes, kernel_size=1)

    # Upsampling layer
    self.upsample = nn.Upsample(scale_factor=32, mode='bilinear', align_corners=False)

  def forward(self, x):
    features = self.densenet.features(x)
    out = self.densenet.classifier(features)
    out = self.upsample(out)
    return out
```

In this example, we load a pre-trained DenseNet model and replace its final classification layer with a 1x1 convolution. We then add an upsampling layer to resize the output to match the input image size. During training, we can use a loss function such as binary cross-entropy to compare the output segmentation map to the ground truth labels.

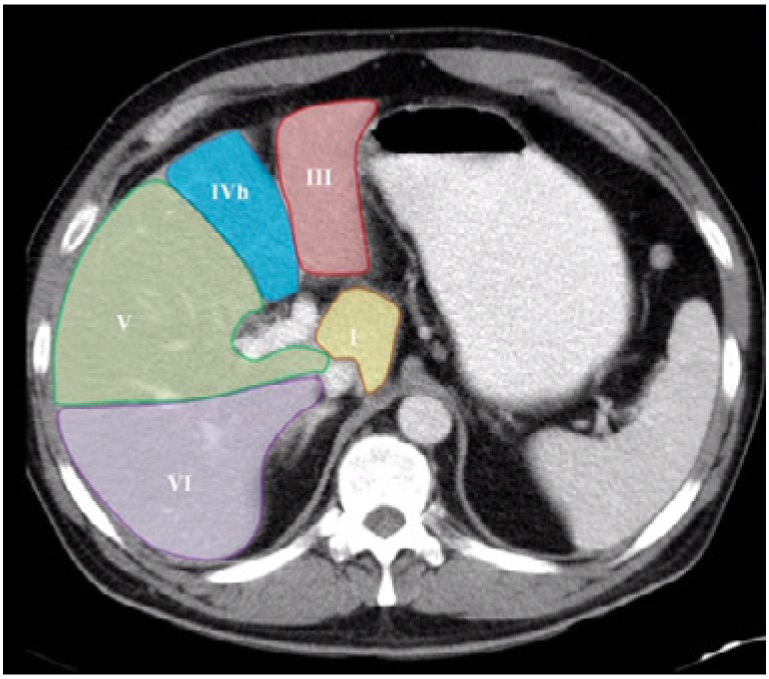
Another approach to using DenseNet for medical image segmentation is to use a 3D convolutional network. This is particularly useful for segmentation tasks that involve volumetric data such as MRI scans. Now just for the sake of inclusivity and comparison, I want to show you an example of a 3D convolutional network implemented using DenseNet in Keras:

```
from keras.models import Model
from keras.layers import Input, Conv3D, Dense, Dropout, GlobalAveragePooling3D
from keras.applications.densenet import DenseNet121


def build_3d_densenet(input_shape, n_classes):
    inputs = Input(shape=input_shape)

    # Load pretrained Densenet model
    base_model = DenseNet121(input_shape=input_shape, include_top=False)

    # Add 3D convolutional layers
    x = Conv3D(64, (3, 3, 3), padding='same')(inputs)
    x = base_model(x)
    x = Conv3D(64, (3, 3, 3), padding='same')(x)
    x = GlobalAveragePooling3D()(x)
    x = Dropout(0.5)(x)
    x = Dense(256, activation='relu')(x)
    x = Dropout(0.5)(x)
    outputs = Dense(n_classes, activation='softmax')(x)

    # Build the model
    model = Model(inputs, outputs)
    return model
```

In this example, we load a pre-trained DenseNet model and add 3D convolutional layers to the beginning and end of the network. We then train the model on volumetric data and use a Softmax activation function to produce a segmentation map with multiple classes.

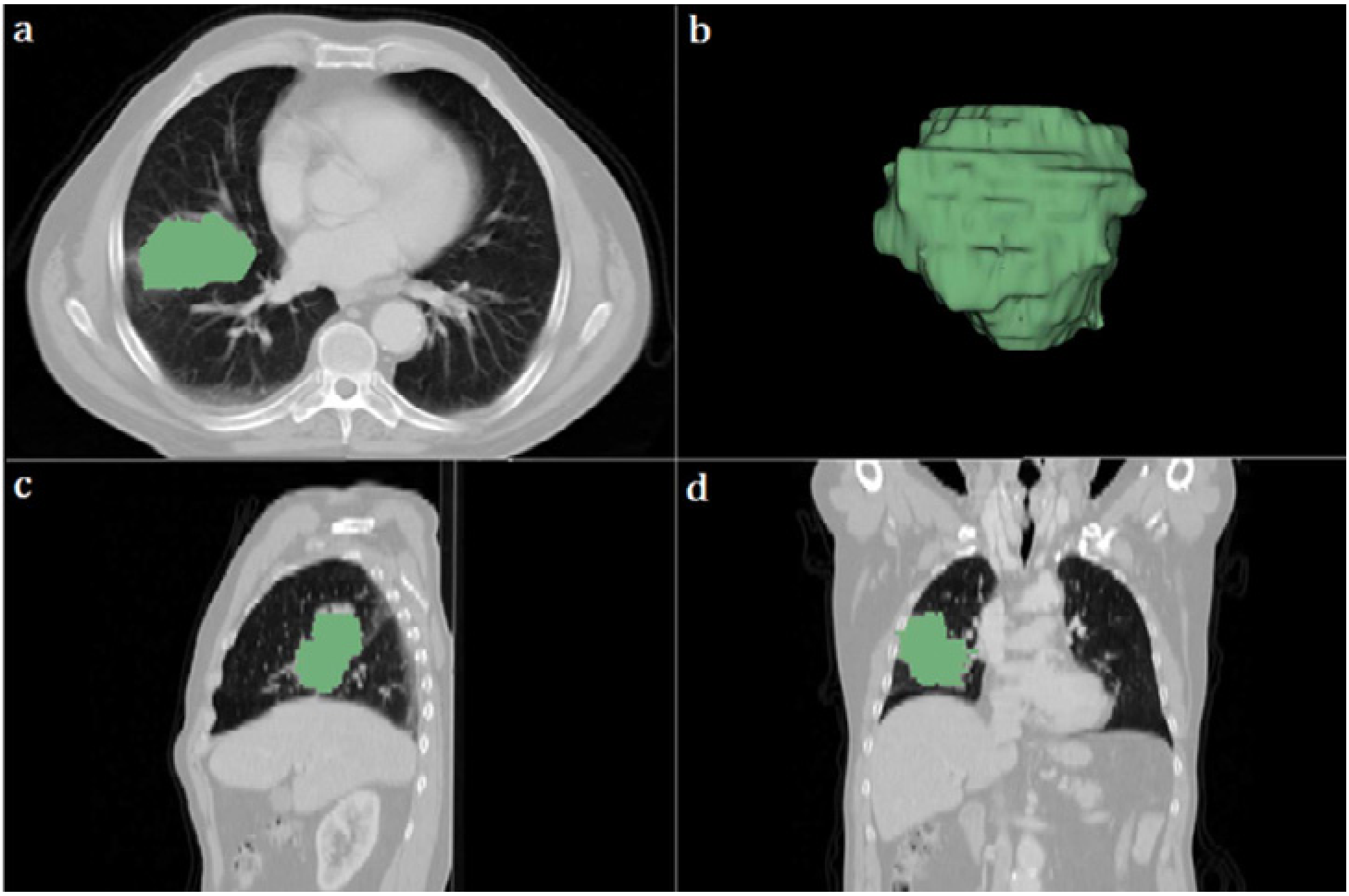
But how about applying segmentation with a different *data modality?* Medical image segmentation is important and can be used with images from different modalities such as MRIs, CT scans, and X-rays. One common application of medical image segmentation is the segmentation of tumors, which is an essential step in the diagnosis and treatment of cancer.

Traditionally, medical image segmentation has been a labor-intensive process that involves manually delineating the boundaries of a tumor in each image slice. However, with the development of deep learning models like DenseNet, we can now automate this process to a considerable extent.



So, let us have a look at doing this for a CT scan! Now for a change, I will use TensorFlow and Keras and I will let you implement the final two steps. Feel free to send your results and/or questions to my email or social media! Happy to hear from you:

```
import tensorflow as tf
from tensorflow.keras import layers
from tensorflow.keras.models import Model
from tensorflow.keras.applications import DenseNet121

# Load pretrained Densenet model
densenet = DenseNet121(weights='imagenet', include_top=False, input_shape=(224, 224, 3))

# Add segmentation head
x = layers.GlobalAveragePooling2D()(densenet.output)
x = layers.Dense(256, activation='relu')(x)
x = layers.Dropout(0.5)(x)
x = layers.Dense(1, activation='sigmoid')(x)

# Build the model
model = Model(inputs=densenet.input, outputs=x)

# Compile the model
model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])

# Load CT scan data and train the model
# exercise for you!

# Use the model for tumor segmentation
# Same here, please email me your results!
```



In this example, we load a pre-trained DenseNet model and add a segmentation head to the top of the network. The segmentation head consists of a global average pooling layer, followed by two fully connected layers with dropout and a sigmoid activation function to produce a binary segmentation mask. During training, we can use a binary cross-entropy loss function to compare the output mask to the ground truth segmentation.

Another approach to using DenseNet for tumor segmentation is to use a 3D convolutional network. Here is an example of a 3D convolutional network implemented using DenseNet in PyTorch. In the previous code we built a 3D DenseNet. Now let's create a class which implements a 3D segmentation model:

```
import torch
import torch.nn as nn
import torch.nn.functional as F

from torchvision.models import densenet121

class Densenet3DSegmentation(nn.Module):
    def __init__(self, n_classes):
        super(Densenet3DSegmentation, self).__init__()
```

```
        # Load pretrained Densenet model
        self.densenet = densenet121(pretrained=True)

        # Add 3D convolutional layers
        self.conv1 = nn.Conv3d(1024, 256, kernel_size=3, stride=1, padding=1)
        self.conv2 = nn.Conv3d(256, 128, kernel_size=3, stride=1, padding=1)
        self.conv3 = nn.Conv3d(128, n_classes, kernel_size=1, stride=1)

    def forward(self, x):
        features = self.densenet.features(x)
        out = F.relu(self.conv1(features))
        out = F.relu(self.conv2(out))
        out = self.conv3(out)
        return out
```

Pretty neat, right?!

In this example, we load a pre-trained DenseNet model and add 3D convolutional layers to the end of the network. We then use the output of these layers to produce a segmentation mask with multiple classes. During training, we can use a loss function such as cross-entropy to compare the output mask to the ground truth segmentation.

Conclusion

I hope you got a great flavor of different implementations for segmentation networks. I wanted to show you how different frameworks can be applied and how different they are when we implement a solution. At the end of the day, it is up to you to decide which one to use, depending on your familiarity, or the rest of the code (if there is some existing code that you are working on).

The most important thing is to get what is really happening behind the scenes; not just memorize or copy the code, but rather be able to create each step and building element!

Let me know if you have any questions and if not... then let us meet the month after the next!

Next month

Again I'll see you in a couple of months. Meanwhile stay tuned and as always you can enjoy the amazing work done by the magazine and my amazing colleague Marica!

Have a wonderful time and always stay curious!

Ganymed Robotics is a robotics start-up founded in 2018 aiming to improve the surgical practice in orthopedics by bringing computer vision algorithms and robotics to the operating room. Marion Decrouez, Computer Vision Team Lead at Ganymed, speaks to us about its work developing a robot for total knee arthroplasty.

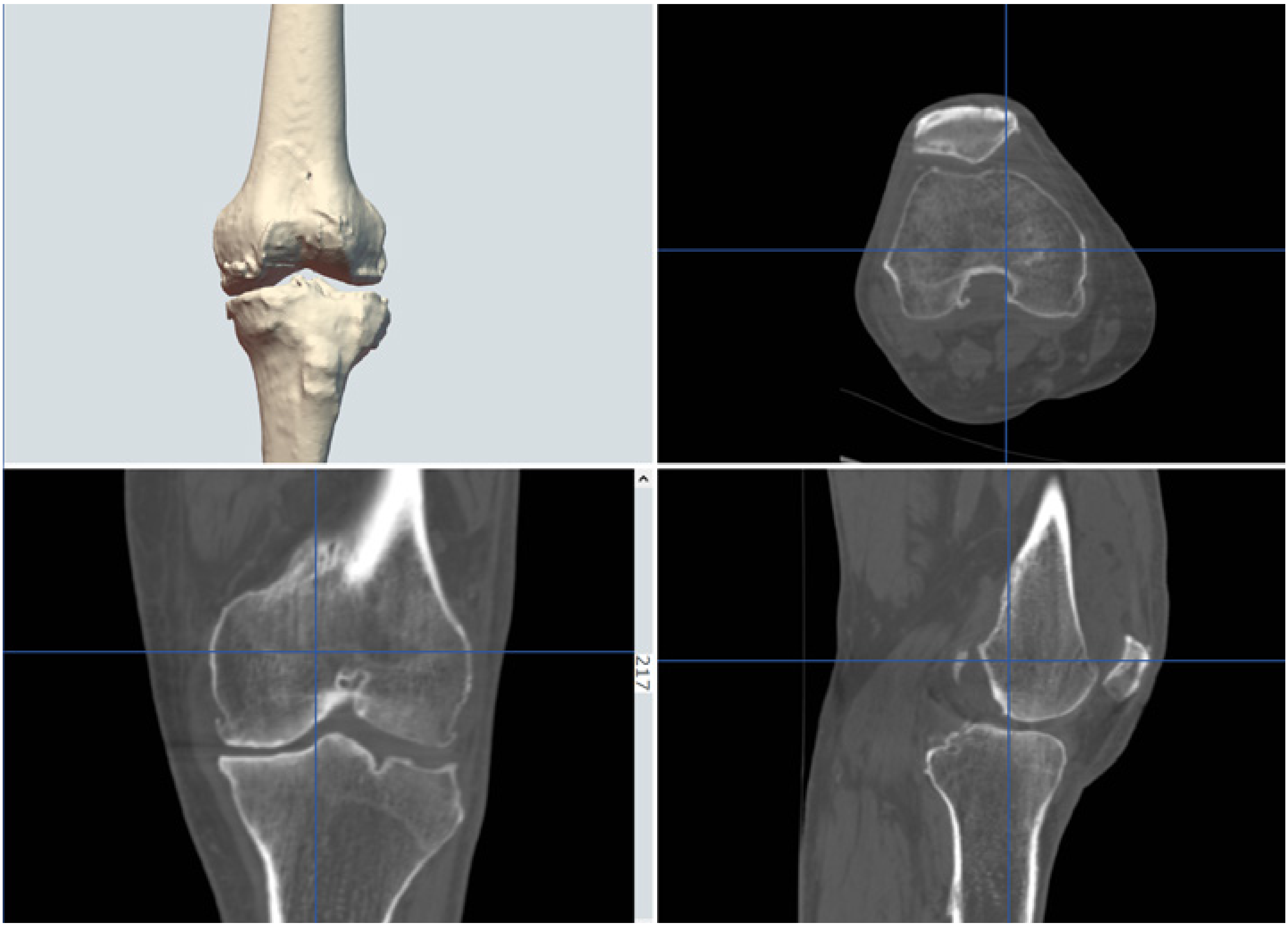
Total knee arthroplasty is a widespread and rapidly growing procedure. As a standard manual operation, it is invasive and associated with unsatisfactory and inconsistent patient outcomes. Even technological solutions offered by other MedTech companies are less than ideal.

"Surgeons tell us it's like woodworking!" Marion laughs. *"They use a hammer, a screwdriver, and mechanical tools to localize the patient during the procedure. Our competitors screw percutaneous 3D trackers on the patients. They're like stars with balls at the end of their branches, tracked by a localizer. That's a big robot with a big footprint on the OR. You have to hurt the patient more than in the standard procedure."*

As part of the procedure, the femur and tibia must have several plane cuts. These are related to the good placement of the implant and better patient outcomes, so they must be exact. In Ganymed's solution, a computer vision algorithm and a vision sensor embedded in a robot help localize the patient intraoperatively to precisely guide the surgeon's gesture. It is **minimally invasive**, avoiding attaching extra trackers to the patient's bone or making unnecessary cuts.

"We want a really intuitive and compact robot," Marion tells us. *"My team wants robust algorithms that work for different operating room conditions and patient morphologies.* ***This procedure has to work for all patients!****"*

The core innovation of the vision part of the system involves a novel approach to intraoperatively locating a patient's bone using the embedded vision sensor, which collects data during the operation.

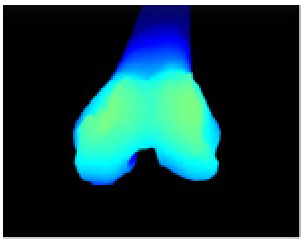


Preoperative data comes from a 3D model of the bone extracted from a CT scan.

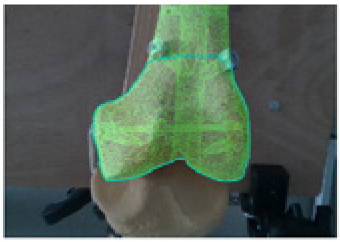
"We begin by detecting the bone in the images and merge the preoperative and intraoperative data to locate it," Marion explains. *"Once we know the orientation and position of the bone, we know it won't move with respect to the robot. Then the robotic part of the system leads on tracking and guides the surgeon with precision and accuracy."*

The team has used different methods to verify and validate the efficacy of their algorithm across a range of data sources, including **simulated and phantom data using synthetic plastic bones and rigorous testing on cadavers under laboratory conditions**. The cadaver testing used trackers or landmarks to obtain the ground truth as a benchmark for evaluating the algorithm's performance.

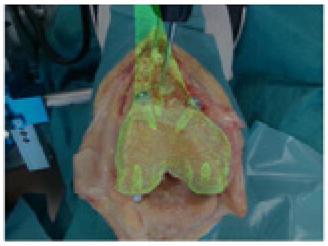
Working with hard structures like bones



Synthetic data

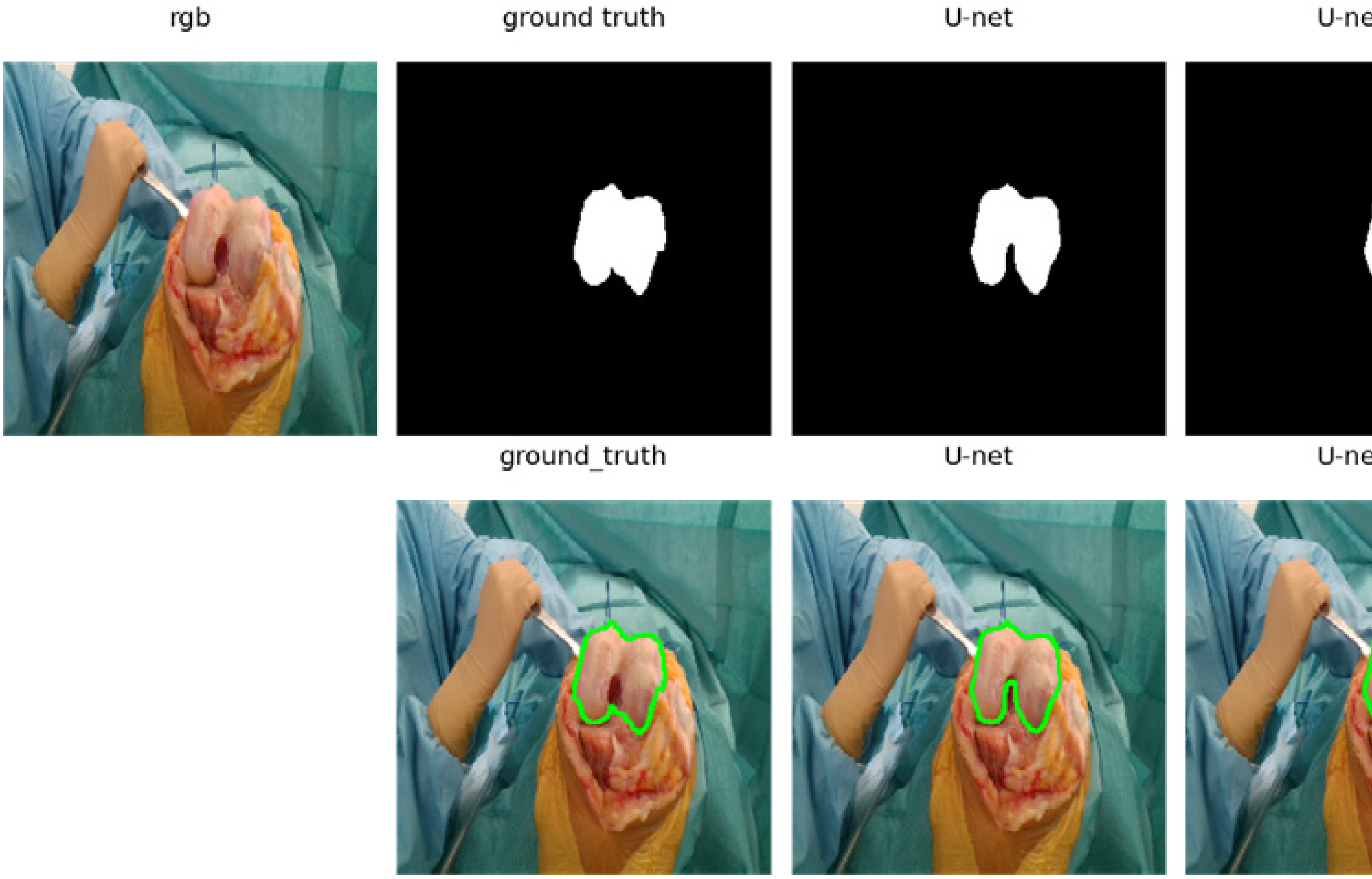


Sawbone Data



Cadaver Data

Patient Data



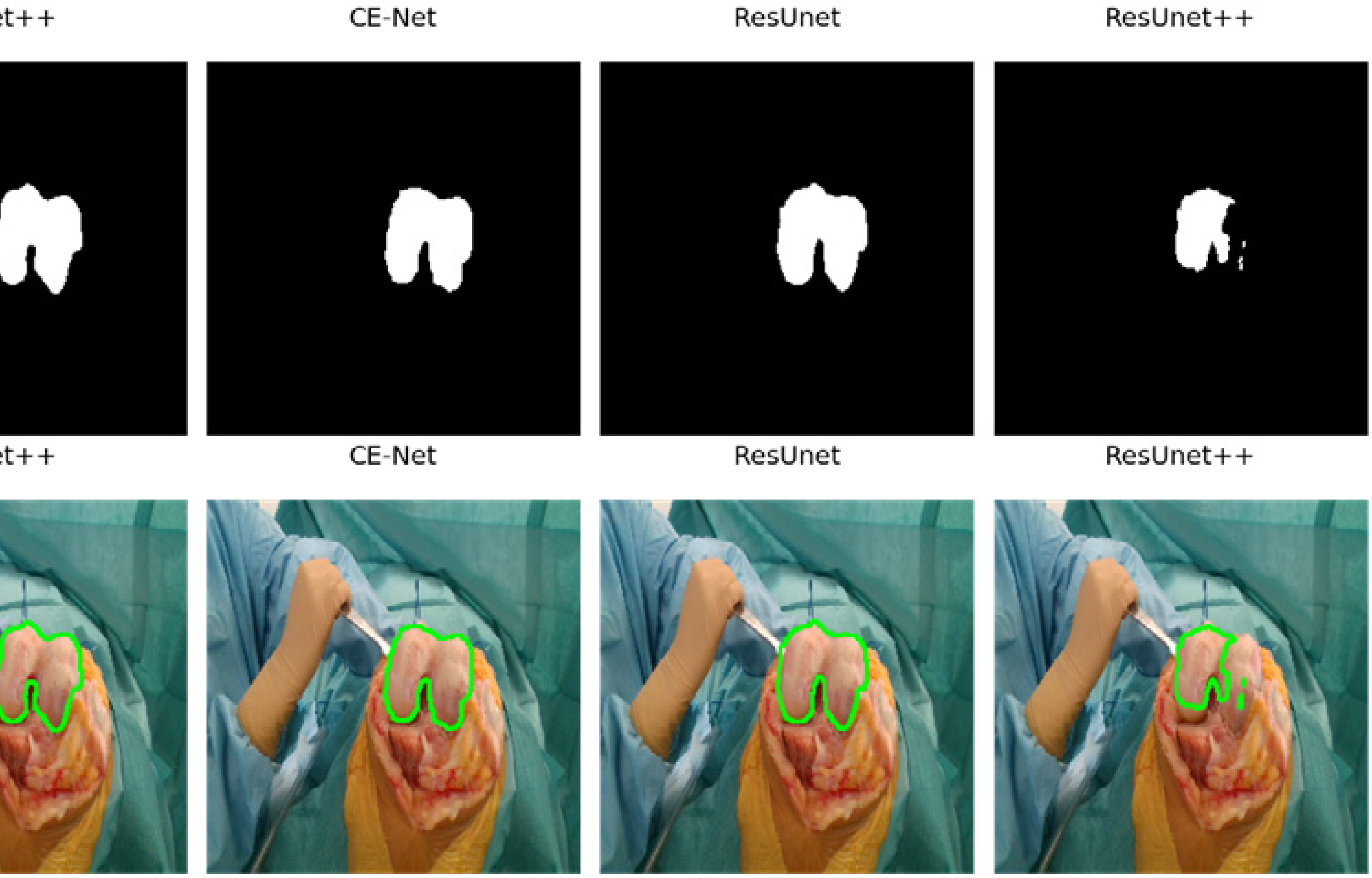
differs from the surrounding soft tissue, like muscles and tendons. The algorithms have been tested on simulated and cadaver data with and without soft tissue to evaluate their robustness. The system provides information to help preserve the soft tissue surrounding the bones' rigid structures.

Ganymed collected data from **100 patients** through a semi-quantitative evaluation to assess whether their algorithm performed optimally under operating room conditions, considering factors such as **bone exposure, lighting, and draping**, and ensured accurate bone detection within the images in all cases.

"When I joined Ganymed, my initial focus was on making something that worked qualitatively," Marion recalls. *"As soon as the bone registration seemed to work, I wanted to know how well it worked. A big part of my work was to develop a means to get this ground truth evaluation to compare with our competitors. We've used the trackers our competitors use, and we've developed their algorithms. We don't know if they're exactly the same, but we have the same data. We can tell the orientation and position errors by comparing the algorithm registration with the ground truth registration."*

Ganymed Robotics has a prototype that works, but when developing a medical device, the road to market can be long and winding, incorporating regulatory approval and industrialization of the hardware and software. It currently anticipates completing all steps by the end of 2024.

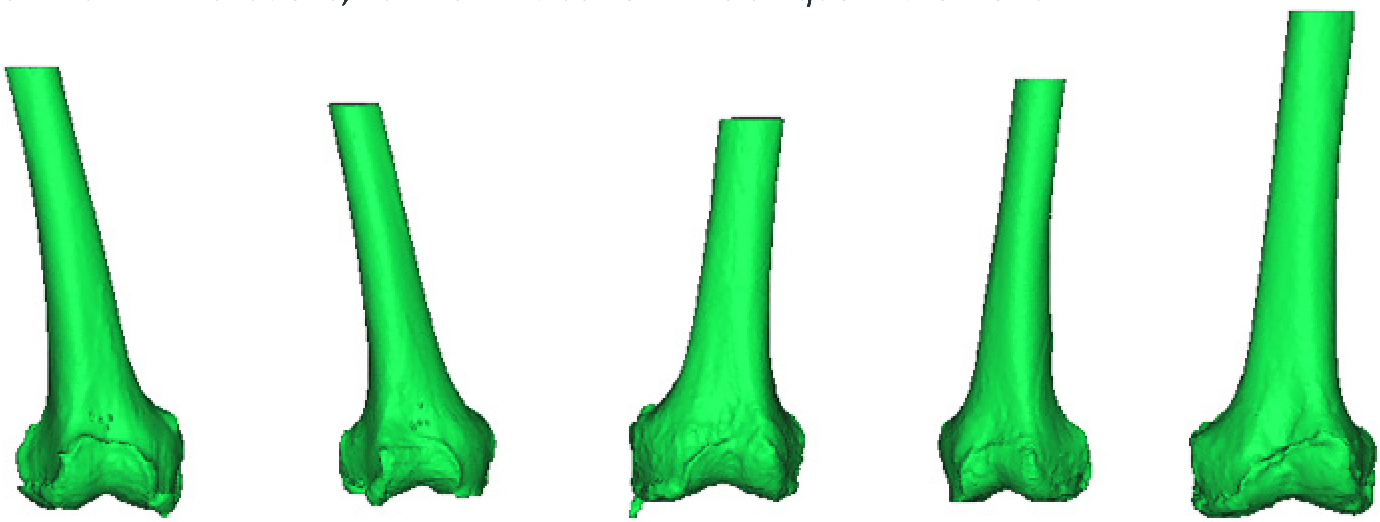
"We've had four patents accepted for the vision part," Marion tells us. *"We've started the process with the FDA, but it's*



a very long process. In the meantime, we have two papers in the ***Computer Assisted Orthopaedic Surgery or CAOS 2023 conference****. This year we'll also launch a PhD thesis in collaboration with a French laboratory, an expert in medical image processing."*

In summary, the Ganymed solution offers two main innovations, a non-intrusive computer vision approach to locate bones and a robot guiding precise surgeon gestures.

"Our contactless and minimally invasive process is new," Marion points out. *"Contrary to our competitors, we don't put extra trackers on the patient's bone; we just need vision data to locate it. This approach is unique in the world!"*



UNIVERSITY OF MARYLAND MEDICAL INTELLIGENT IMAGING CENTER (UM2II)

Paul Yi

Vishwa Parekh

Paul Yi and Vishwa Parekh are the Director and Technical Director at the University of Maryland Medical Intelligent Imaging Center and Assistant Professors of Diagnostic Radiology and Nuclear Medicine at the University of Maryland School of Medicine. They are with us today to tell us more about their exciting work.

The **University of Maryland Medical Intelligent Imaging Center**, also known as **UM2ii**, aims to improve human health by advancing AI in medical imaging from bench to bedside. It benefits from the distinct skill sets that Paul and Vishwa bring to the table. Paul, a radiologist by training, brings a wealth of **clinical knowledge**. On the other hand, Vishwa, a computer scientist, offers valuable **technical expertise**.

"Much of what I do focuses on clinical translation, thinking about how we can use AI to answer clinically important questions, as well as how we solve problems that

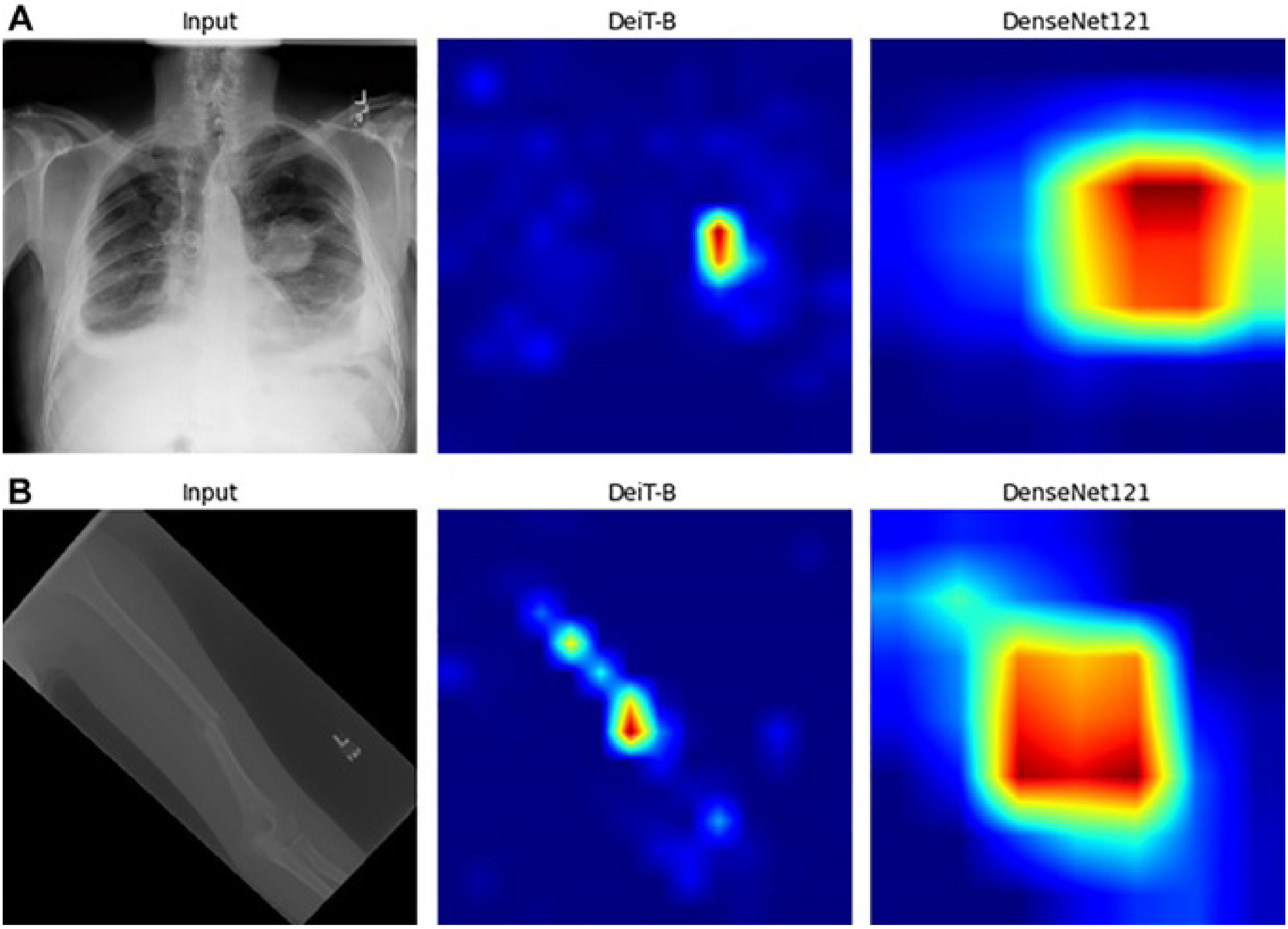


Figure showing our recent work in Vision Transformers (ViT) compared to Convolutional Neural Networks (CNN) for radiograph diagnosis. Heatmaps for ViT (DeiT-B, middle) show more precise localization to radiographic abnormalities than CNNs (DenseNet121, right).

prevent the AI being translated," Paul tells us. *"Problems like fairness, trustworthiness, and explainability of AI are things that I care very much about as a practicing radiologist. Vishwa brings a robust set of skills on the technical side. How do we translate things from computer science into the medical world?"*

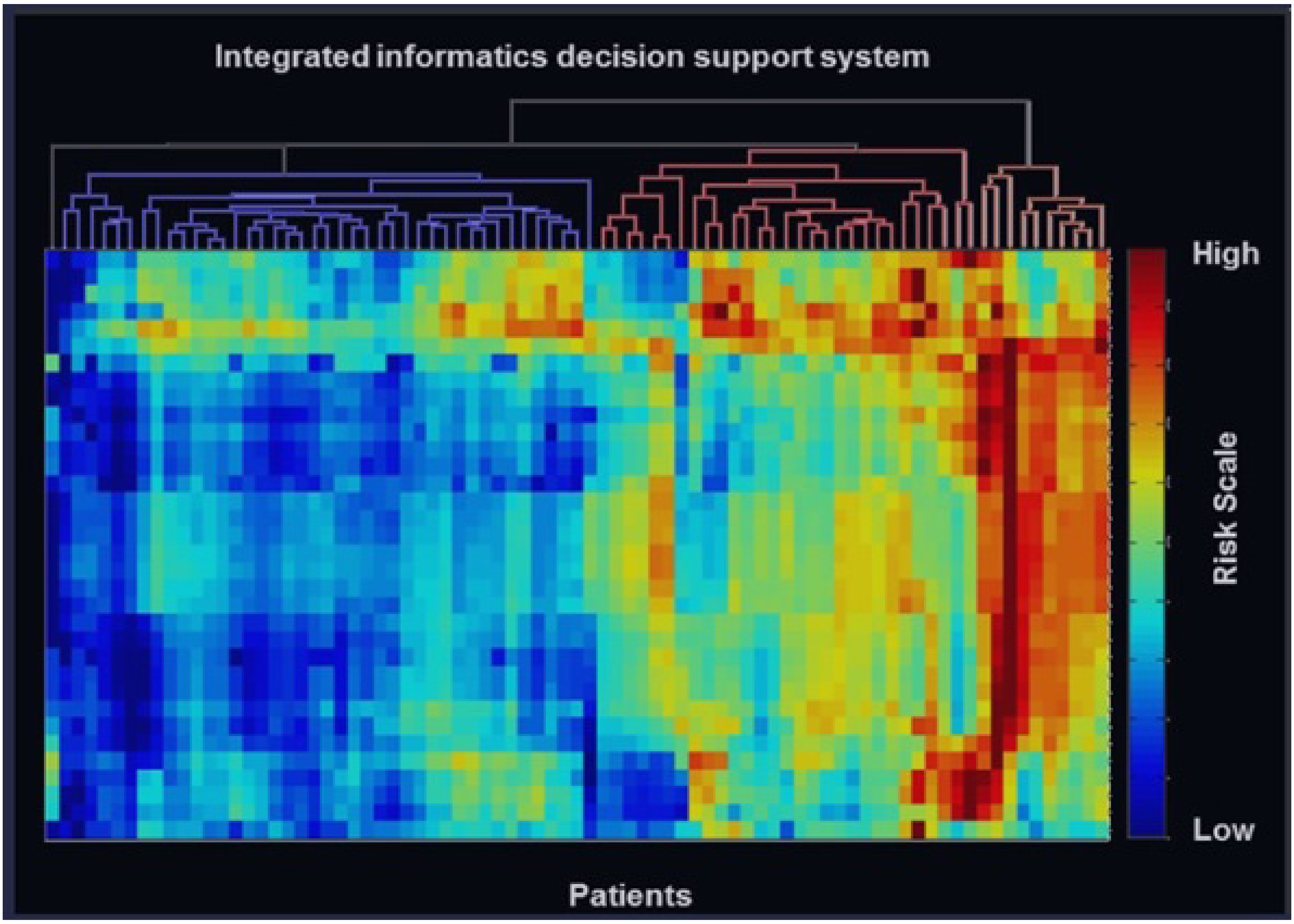
By design, the pair are building a lab that rejects the previous silo working in these two domains. This collaboration allows them to bounce ideas off each other and discover the problems that really require attention.

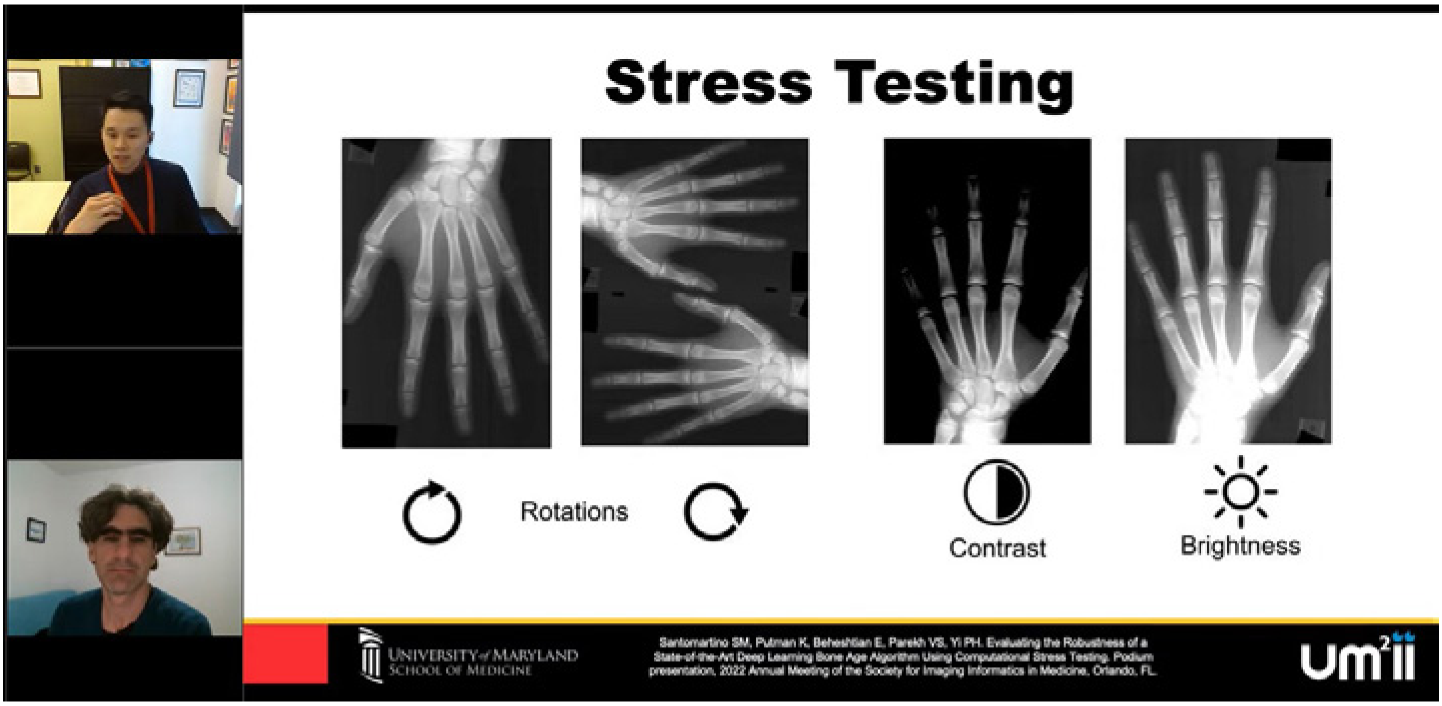
"I went through four years of medical school, five years of residency, and a one-year fellowship, where most of my time was spent learning how to do clinical work," Paul points out. *"Vishwa spent a similar amount of time on the computer science side. We have different perspectives and vantage points and can cover each other's weaknesses."*

Vishwa agrees: *"As a computer scientist, I would go off and work on problems that wouldn't even be useful from a clinical standpoint. What we do here is work on problems that can be translated and* *will actually be useful in clinical practice. Imaging has completely revolutionized modern medicine. We are currently at an intersection where* ***advancements in AI are leading to a new revolution in intelligent imaging!****"*

He explains that the current method of **biopsying** only a small part of a tumor for cancer diagnosis and treatment planning, for example, is suboptimal and can lead to ineffective treatment. However, a virtual biopsy utilizing AI and imaging together can provide a more comprehensive understanding of the tumor. **Collaboration between computer scientists, radiologists, and clinicians is necessary to develop models for predicting tumor behavior and designing the most effective treatment plan**.

"We see eye to eye about our goal, which is bringing together our expertise to solve the problems of AI and radiology," Paul tells us. *"We know how to speak each other's language. I'm not a computer scientist, and I don't know as much about the technical detail as Vishwa does, not even close, but I've spent enough time working with engineers and learning the concepts that I'm able to translate these things. Vishwa, having worked with physicians for several years, can come to me and talk to me about things in a language that I understand. This translation piece is often the most important thing."*





The group recently published a paper about **vision transformers** – a new type of deep learning algorithm different from traditional convolutional neural networks. Having seen vision transformers garnering attention in other computer vision and AI communities, Paul wondered if they could have a similar impact in medical imaging.

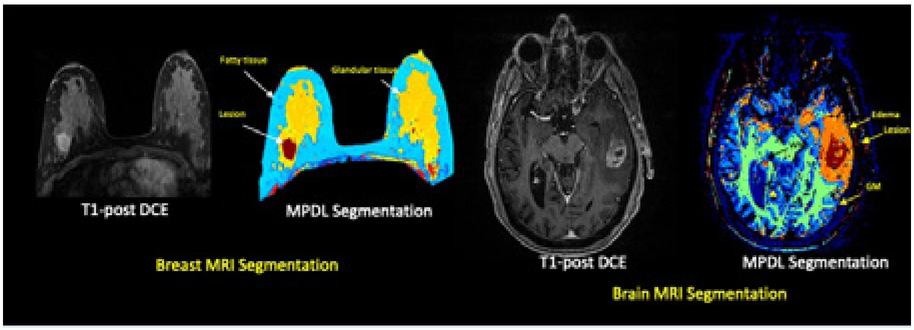
"Working with one of my collaborators at Johns Hopkins, we took this new technology that Facebook and Google had developed at the time and found it had some interesting advantages," he explains. *"If I hadn't had that kind of collaboration, if I hadn't been reading and aware of the literature on the computer science side, I never would have figured it out. Vision transformers will be very exciting to evaluate because they're just starting out. Convolutional neural networks changed the game, and I think vision transformers will have big implications."*

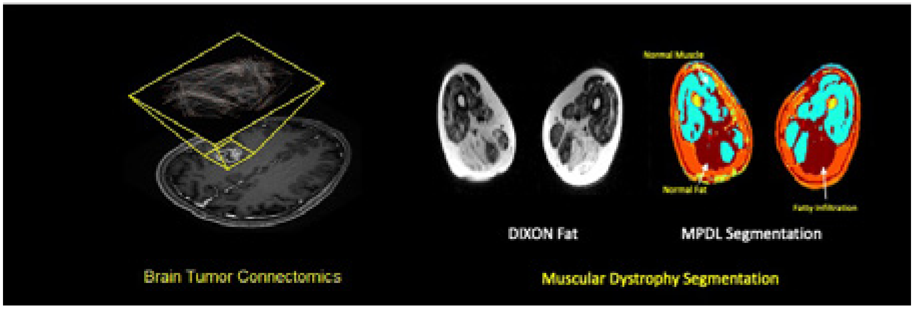
UM2ii is only just approaching its first birthday. Paul, also an Adjunct Assistant Professor of Bioengineering at the University of Maryland, College Park, and Adjunct Assistant Research Scientist at the Johns Hopkins Malone Center for Engineering in Healthcare, was recruited by the University of Maryland to build it, and it launched when Vishwa joined last year.

Joining Paul and Vishwa on the leadership team are **Clinical Director and neuroradiologist Peter Kamel** and **Innovation Director Florence Doo**, who is finishing her Body Imaging Fellowship at Stanford. The team has three full-time staff, two software engineers and one program coordinator, and around 10 students, including graduates, undergraduates, and medical students.

If you think Paul seems familiar, that may be because he spoke brilliantly at our webinar last year, **Can we trust AI?**

"I'm interested in the fairness and bias of algorithms," he reveals. *"Does an algorithm*





perform better on males vs. females, or on one ethnic group compared to another? Few of us on the MD side evaluate these pitfalls."

The team recently published a paper in **RSNA's Radiology journal** showing biases in a leading bone age prediction model, indicating inconsistent performance across diverse populations. While they work to solve this problem, one may wonder why they have prioritized such intangible concepts over the numerous physical issues that still need to be addressed globally.

"It's important because if algorithms are biased, they can perpetuate health inequalities in our country in other places," Paul responds. *"We want to improve human health, but pieces like fairness and explainability are just as important as diagnosing disease because if these things are not solved, then we can't achieve the ultimate goal. If an AI model says you have cancer, but we can't confirm that, it will be hard to get anyone to trust it.* ***If they don't trust it, they won't use it!****"*

The team's work interests a broad audience, including journals such as RSNA's Radiology and **Radiology: Artificial Intelligence**, and conferences and meetings, including the annual meeting of the Society for Imaging Informatics in Medicine (**SIIM**), the Conference on Machine Intelligence in Medical Imaging (**CMIMI**), the Medical Imaging with Deep Learning (**MIDL**) conference, and of course the Conference on Medical Image Computing and Computer Assisted Intervention (**MICCAI**).

With these events still a few months away, what might we expect to see Paul and Vishwa presenting this year?

"We're developing a ***federated***

The "Spirit of Charm City" by artist Sean Garin

This is the mural in our lab painted by our program coordinator.

Description from Sean:

"Water has been a staple in the thematic characterization of the UM2ii Lab, as collaboration is guided by the fluidity of thought and the passion to "make waves" in the medical imaging bioinformatics space. With a local landmark being so geographically close to the University of Maryland School of Medicine, it felt like a natural scene to depict, welcoming visitors to the lab and giving a hint of the spirit of its members."

***learning framework** for homogeneous and heterogeneous datasets,"* Vishwa reveals. *"If one group works on liver tumors while another works on pancreas tumors, each can benefit from the other's work. We don't need to curate a completely new dataset, but can develop a more generalizable collaborative framework that works on everything."*

The team will also present lifelong learning technology, which is essential because many models are static and do not adapt well to real-life situations with new diseases, scanners, or protocols. This technology enables AI systems to continually learn and integrate new information, including developments like Covid-19.

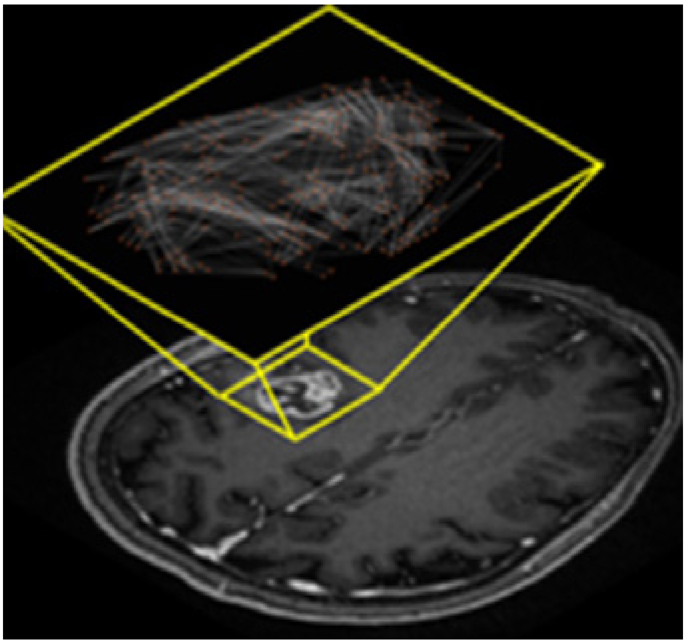
Looking to the future, the team plans to scale up in the coming months and years. Its lab already has a reputation as somewhere synergy is happening, and MDs and PhDs work together in ways that will move the needle.

"How do we move together rather than pulling or pushing each other?" Paul poses. *"We're trying to make this multidisciplinary by design. If we can achieve that, many amazing things will happen, and our work will be more impactful."*

And in five years, where will UM2ii be?

*"I'd love to see **some technology from this lab get deployed into clinical practice**,"* Vishwa ponders hopefully. *"It's a very ambitious five-year goal, but even if we are in a position where we are piloting at multiple places, I think that would be a success."*

Paul adds, finally: *"I hope that in five years we have funding and we're publishing papers, but the most important thing is that I wake up every day and I'm having fun, and the people in the lab are having fun too because life is short. The moment this stops being fun, I'll go and find something else to do."*



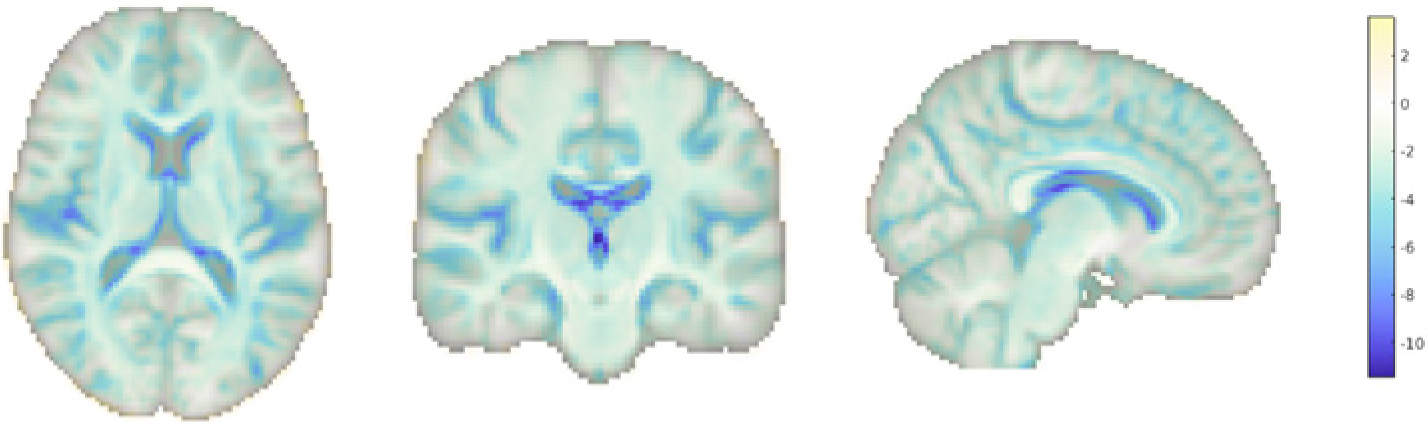
Congrats, Doctor Chiara!

Chiara Mauri recently finished her PhD with the Computational Neuroimaging group at Technical University of Denmark. Her research aimed to develop a novel way of making accurate and interpretable predictions based on brain images. She is soon starting a postdoc at Harvard Medical School, where she will continue her research on computational imaging methods. Congrats, Doctor Chiara!

Magnetic Resonance Imaging (MRI) is a technique for image acquisition which is extensively used in neuro-clinical practice, to support clinicians in making diagnosis and planning treatments. In the last decades, researchers started to develop computational methods to make automatic predictions of variables of interest, such as a subject's diagnosis or prognosis, based on brain MRI scans. Since MRI is able to detect very subtle changes in brain anatomy, these methods have a huge potential in many clinical applications, such as providing early diagnosis, better therapy planning and monitoring, paving the way to personalized treatments.

Current image-based prediction methods mainly focus on deep discriminative learning techniques. While these methods are able to produce accurate results, they have been proven to be difficult to interpret. This may be problematic, since, in many neuroimaging tasks, it is important not only to predict well, but also to interpret morphological changes underlying predictions.

Given the limitations of existing techniques, during my PhD I developed a novel method for image-based predictions [**springer**]. This method is based on a lightweight generative model,



which not only yields accurate predictions, but is also inherently interpretable, and simple and fast to use.

I first demonstrated the effectiveness of the method by using age and gender prediction tasks: I showed that the method achieves prediction performances that are competitive with state-of-the-art techniques, especially when the sample size is at most of a few thousand subjects, which is the typical scenario in many neuroimaging applications. I then employed the method to perform automatic diagnosis of multiple sclerosis patients.

A huge advantage of this method is that, in addition to being accurate, it is also inherently interpretable. In fact, it automatically provides maps displaying the morphological effect of the variable of interest on the brain, which are straightforward to interpret. Figure 1 shows an example of these maps obtained for age prediction: they mostly highlight areas around the ventricles and in the cortex, expressing therefore ventricles enlargement and cortical thinning, which are known morphological effects of brain aging.

This prediction method also allows to answer counterfactual questions, such as "*How will my brain look like when I am 80 years old?*". Figure 2 shows the real brain image of a 47 years old person, together with the same person's brain, artificially aged to 80 years old. We can see that the model is able to encode typical aging patterns, such as ventricles enlargement.

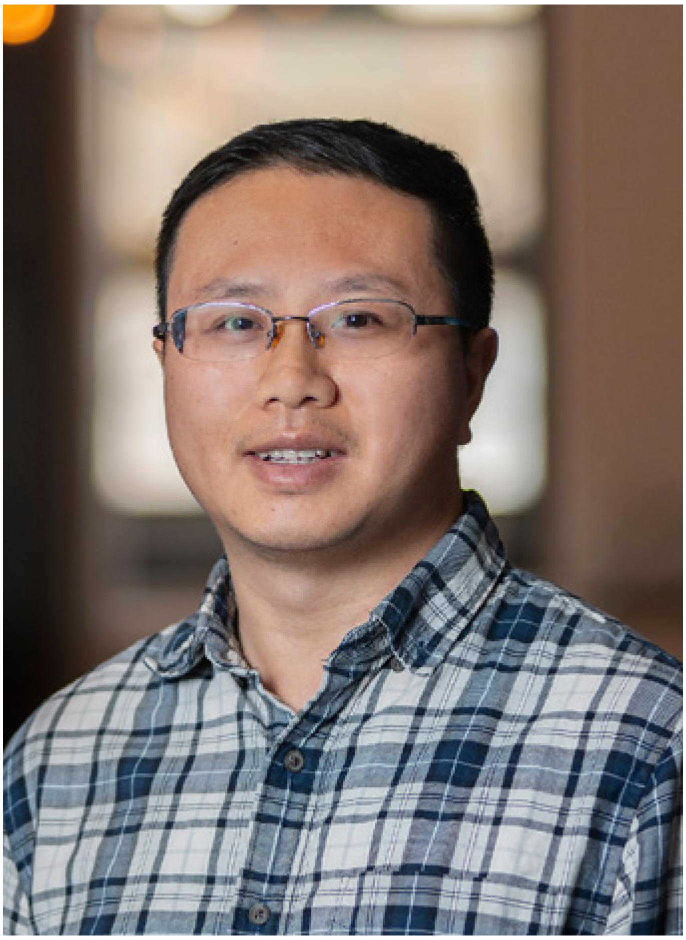
Therefore, this prediction method that I developed has the advantage of being both accurate and interpretable, but let's not forget about its simplicity and speed: training the method on a few hundred of subjects for example takes only a few minutes (on a CPU)!

47 years (real)

80 years (counterfactual)

WASPSYN CHALLENGE AT ISBI 2023

Jingpeng Wu is an Associate Research Scientist at the Flatiron Institute in New York. He is organizing an interesting ISBI challenge exploring the potential of out-of-domain generalization methods for large-scale connectomics applications. He speaks to us about what is involved.



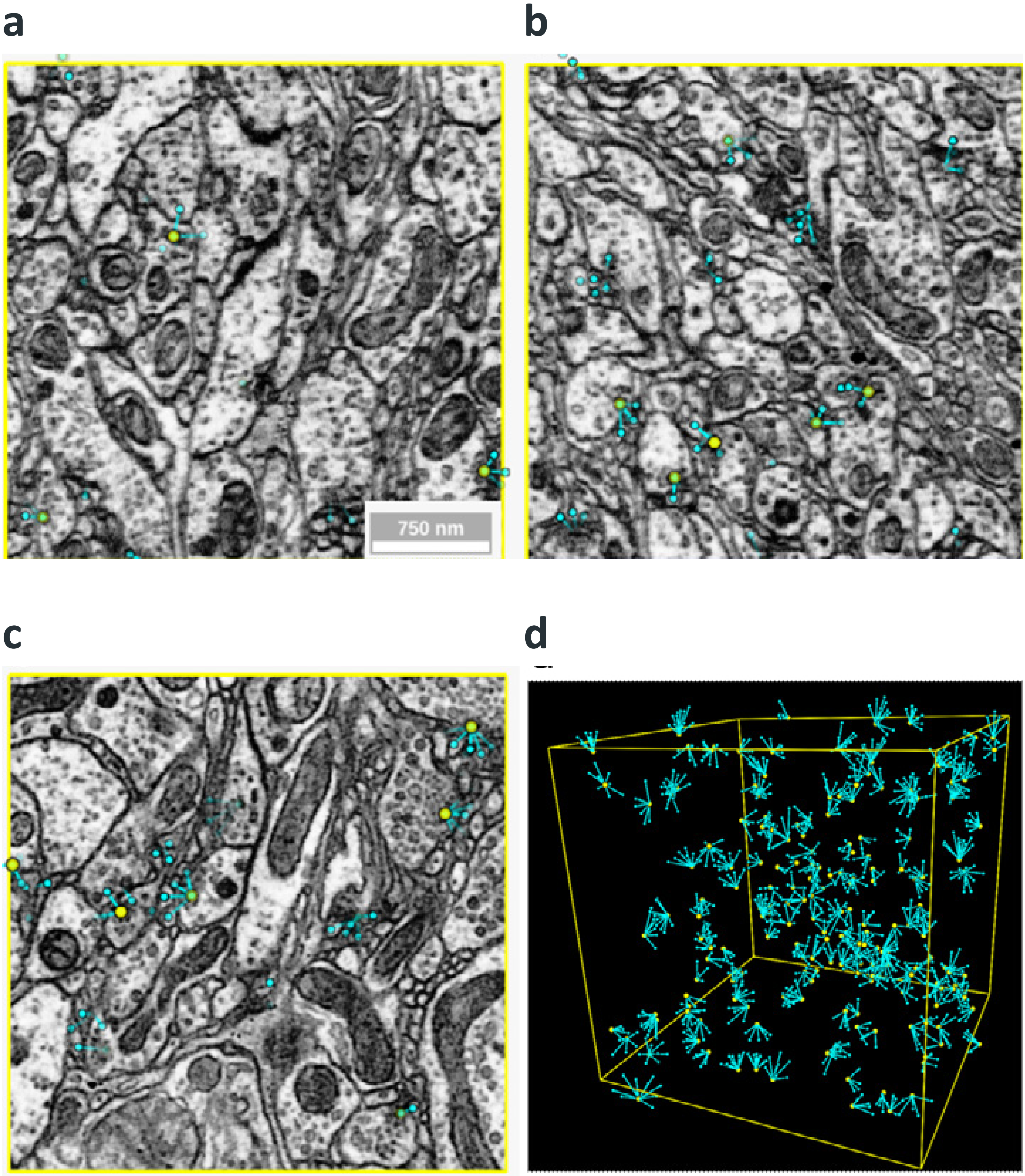
The objective of the **WASPSYN Challenge**, part of the ISBI 2023 conference in April, is **to detect synapses within electron microscope images**. It is an out-of-domain challenge, introducing three brain samples, with participants expected to train on one and test on the other two.

The task represents a significant challenge in **connectomics** to understand the intricate network of neuron connections within the brain. The brain is sliced into thin sections – in this case, a mere 8 nanometers (nm). The resulting voxel size is 8 x 8 x 8 nm, providing extremely high-resolution 3D imagery.

"The goal is to understand the complete biological network in the brain," Jingpeng tells us. *"We normally talk about artificial neural networks, artificially generated or manually designed networks, but we're studying real-world biological neural networks."*

The brain is represented as an extensive graph, with **neurons as nodes connected by synapses**. Every synapse is an edge in this directed graph, and there can be multiple edges in these two nodes because there can be multiple synapses connecting two neurons.

"The neurons split out the neurotransmitters



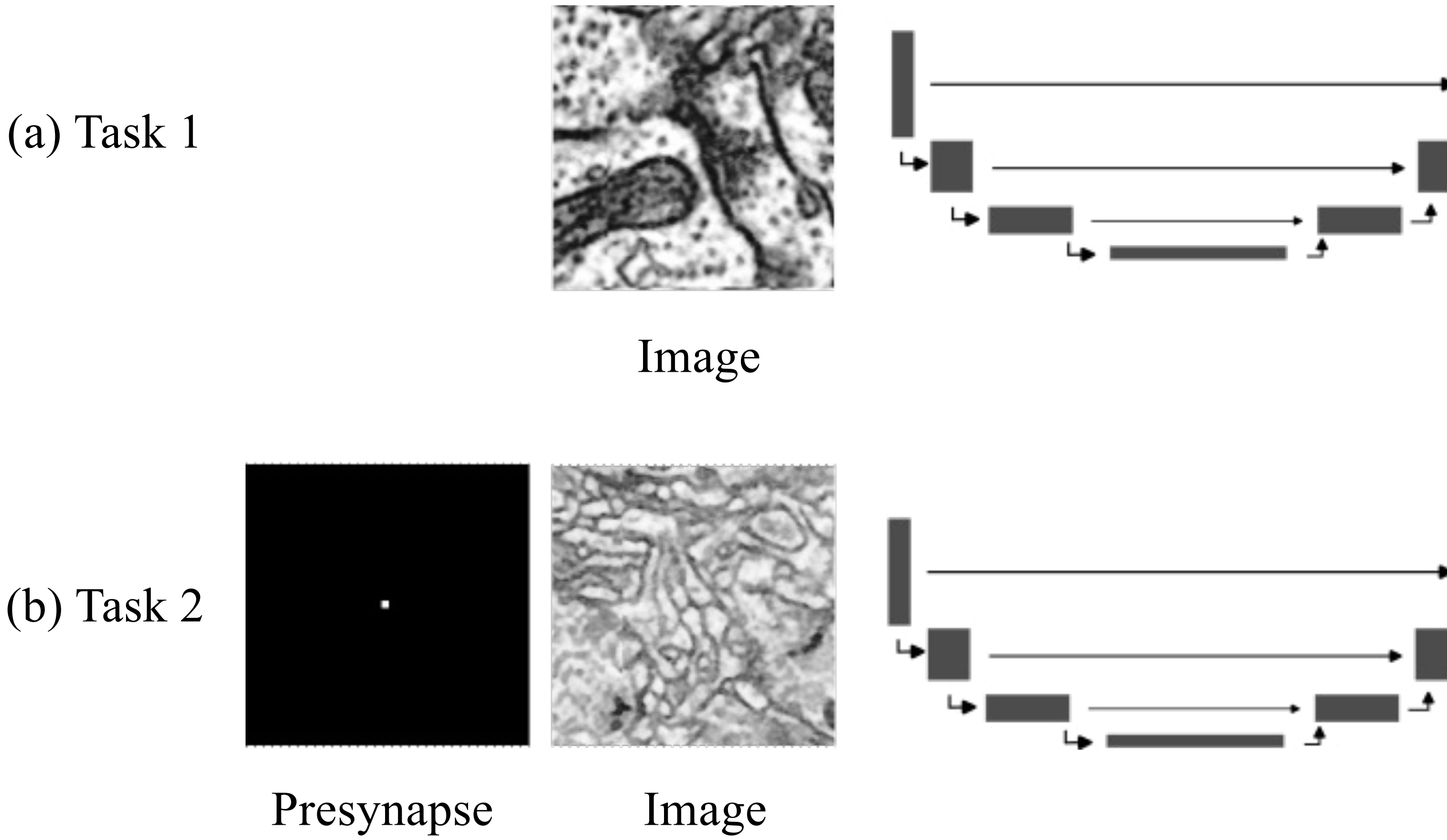
Synapse detection from 3D electron microscopy (EM) image vol-ume.
(a,b,c) The ZY, XZ, and XY planes of the 3D volume with manually annotated synapses;
(d) 3D point cloud visualization of the annotated synapses: presynapses represented as yellow dots and postsynapses as cyan dots and edges connected to the corresponding presynapses.

to another neuron from the pre-synapse to post-synapse," he explains. *"It has a direction from neuron A to neuron B. In this challenge, we want to detect the synapses automatically and accurately."*

WASPSYN is distinct from an existing **circuit reconstruction challenge called CREMI**. The primary difference lies in the type of electron microscopy used for imaging. In CREMI, physical sectioning is employed, resulting in a section thickness of 40 nm, compared to 8 nm here, and an XY lateral resolution of 4 nm. Also, the training and test volumes are very close. The image texture is similar, which Jingpeng says **does not reflect the diverse neuron morphology in different brain regions**.

"We want to apply our model to the whole brain or even another brain, so we require the trained model to be generalizable to many brain regions and datasets," he points out. *"Then, these methods can be applied in real-world connectomics. The CREMI methods don't have this guarantee."*

Solving this problem proves challenging due to the **extensive variability of the neuron morphology in brain regions**. Furthermore, there is a lack of consistency even when examining different brain samples with the same microscope. Illumination and sample preparation can

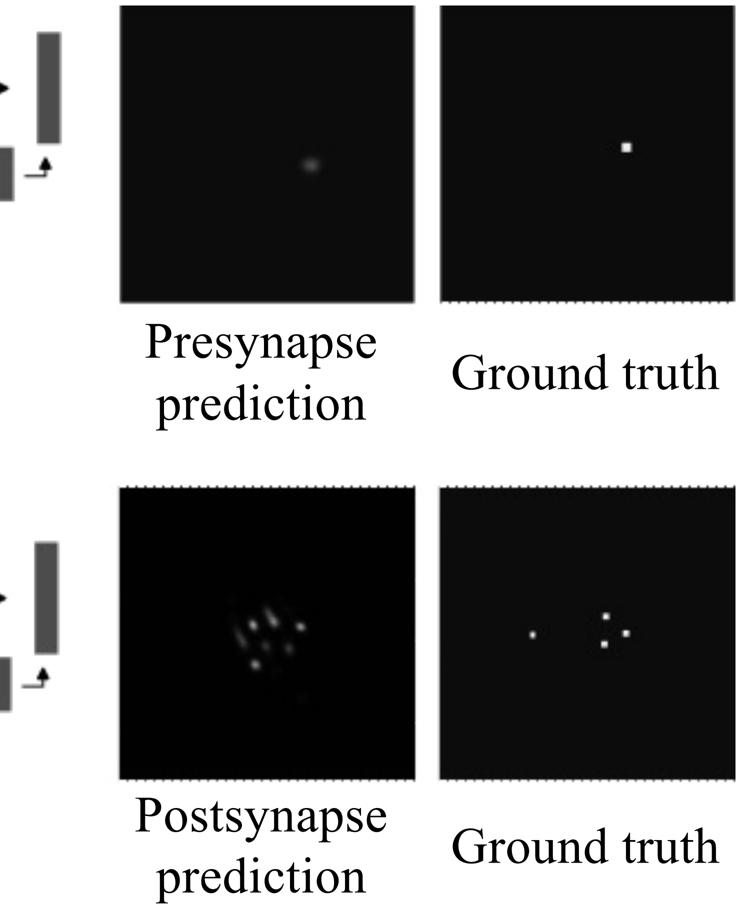


Baseline method using 3D U-Net. (a) Training of T-bar detection network. The patches are randomly sampled
(b) Training of post-synapse detection network. The T-bar is in the center of each patch and a fixed patch with a ce
the illustration is 2D while both the patches and network are 3D and of

cause images of the same species, sex, and age to appear noticeably different.

"***We want to build an atlas of the brain***," Jingpeng reveals. "*This requires very high accuracy. We'll approach humans as proofreaders to correct errors, but there's a balance between how much human work is needed and how accurate the model must be. Improving these models will gradually reduce the work required by humans.*"

He hopes the challenge attracts computer vision community members interested in out-of-domain object detection, often experts at building models with good generalizability.



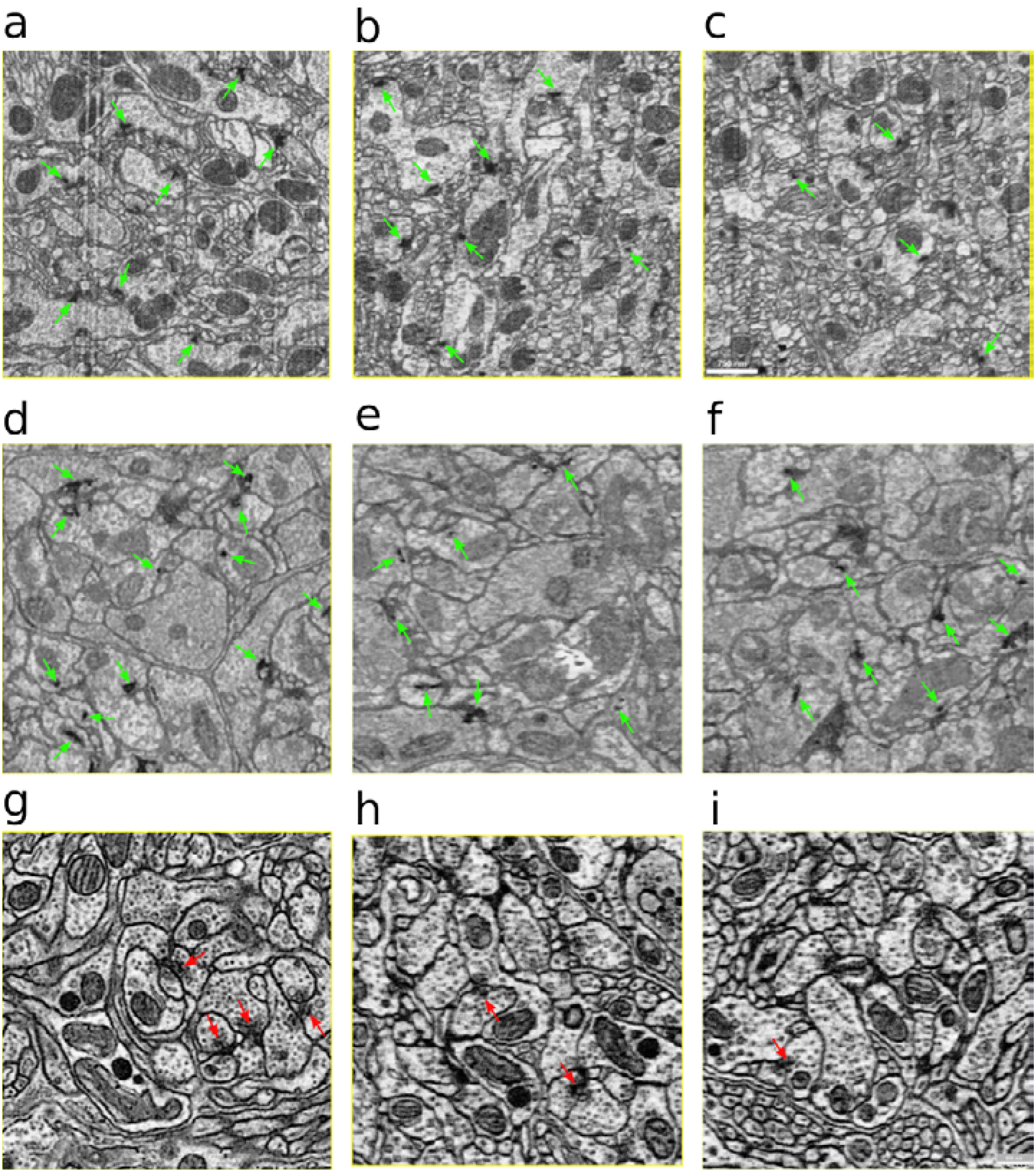
so that some patches might not contain any T-bar. ntral point is used as a channel of input patch. Note that isotropic size.

"*I'm not a specialist in this field, so I hope they will try their best to solve our problem,*" he appeals. "*This is a real problem I encountered in my research. I'm working on **a microwasp connectomics project to map several microwasp brains**. It's a rare study area, and to the best of my knowledge, we're the only organization doing it. We have three data sets covering the whole brain, and I work mostly on one of them. I've built a baseline method, but it doesn't generalize well across brain regions and samples.*"

The wider connectomics community is experiencing rapid growth, and Jingpeng's work extends beyond microwasps to include **flies, mice, monkeys, and even humans**. However, the methods used in mammalian brains cannot be directly applied to microwasps. One of the most significant differences is in **the post-synapse density or synaptic cleft**.

"*In a mammalian brain, you can see a very thick membrane **when two neurons meet with a synapse**,*" he explains. "*The contacting face of the synapse is very dense, dark, and thick and is called Post Synapse Density. In microwasps, the synapses are much smaller, making it harder to see this post-synapse density. Also, our imaging method has a lower lateral resolution than transmission electron microscopy, making this problem more difficult to solve than in a mammalian brain.*"

You can learn more about this competition on the challenge website, and Jingpeng's full proposal is available on arXiv.



Images from three specimens.
The columns of images are XY, XZ, and YZ planes from left to right.
The images are from specimens one to three from top to bottom.
The arrows indicate T-bars identified in the section.
The red arrows indicate T-bars from mushroom bodies specifically

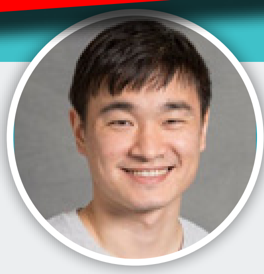
RSIP VISION Visual Intelligence for MedTech

UCL

RSIP VISION MEETUP

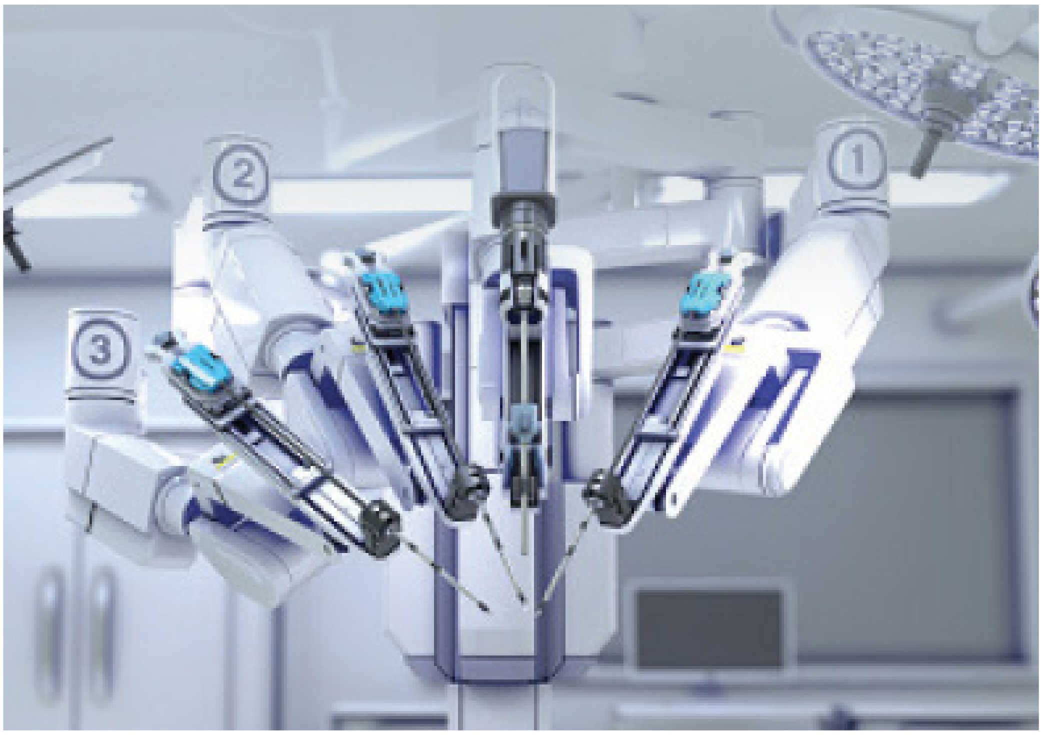
ML & Clinical Use in Ultrasound-guided Surgery

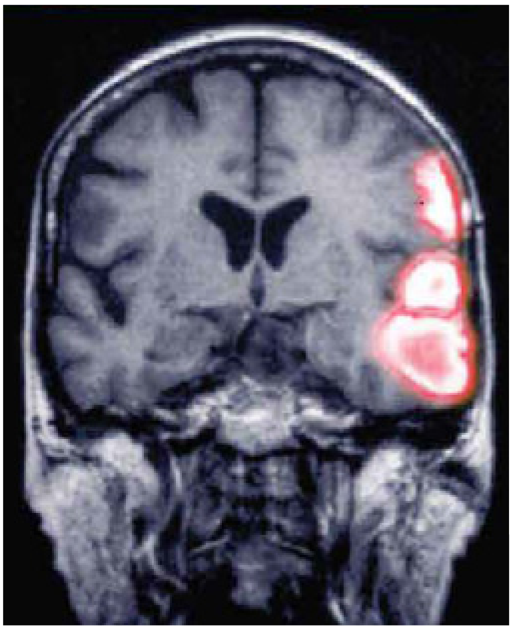
If you missed the webinar, don't miss the video!

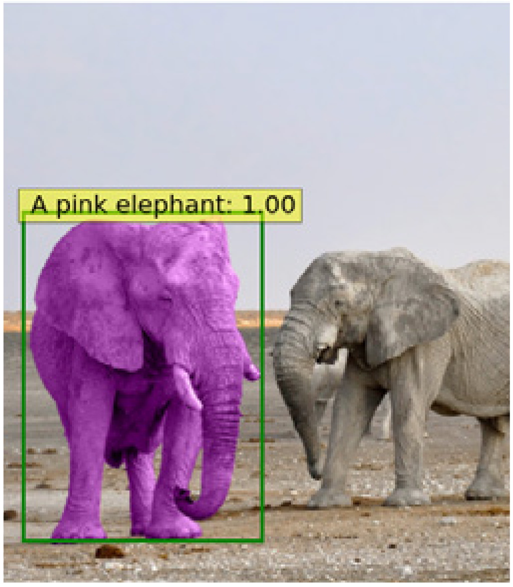


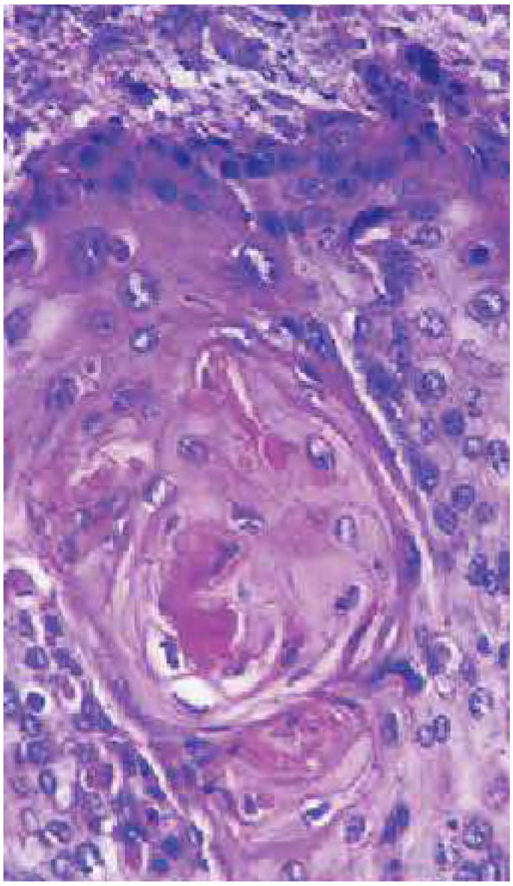
Hosted by
Moshe Safran,
CEO RSIP Vision USA

Speaker
Yipeng Hu,
Associate Professor at UCL
& Affiliated Researcher at WEISS









RSIP VISION