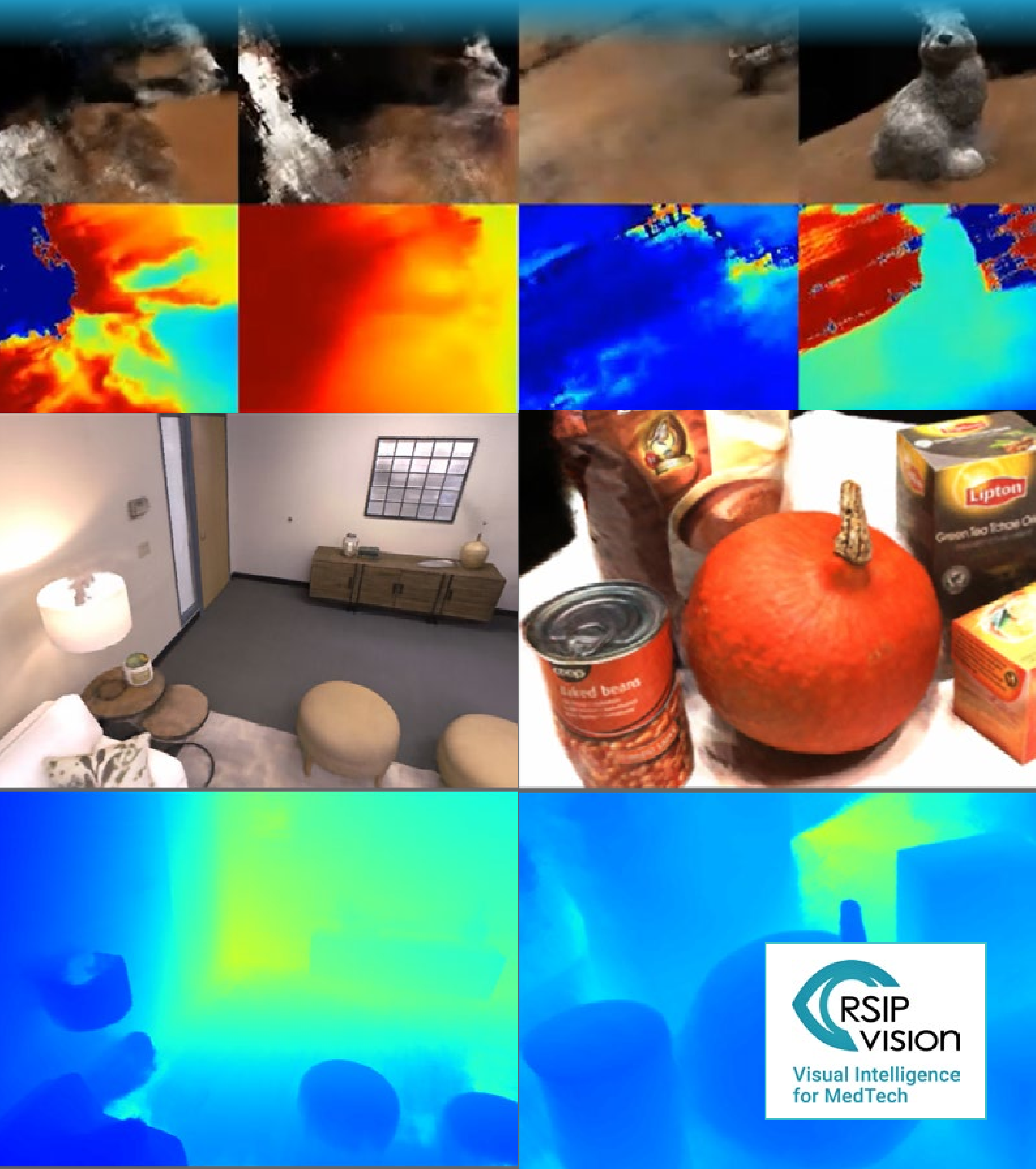
MAY 2023

Computer Vision News

& Medical Imaging News

The Magazine of the Algorithm Community



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

Copyright: RSIP Vision
All rights reserved
Unauthorized reproduction
is strictly forbidden.

RSIP Vision and Computer Vision News claim no copyright on any of the presented works, but only on our editorial work.

Contact us

Free subscription

Read Previous Magazines

Follow Us

Dear reader,

Welcome to the May issue of Computer Vision News!

This month, we keep up our tradition of reviewing the **Best Papers** at the world's leading conferences, with two excellent but very different award winners at **ISBI 2023**: the **Best Paper** presenting a novel solution to dataset similarity and transfer learning in medical imaging and the **Best Oral Presentation** proposing improved diagnosis for a common but hidden medical condition affecting millions of women globally.

The premier international computer vision event, **CVPR 2023**, is just around the corner! Register now to join thousands of delegates in **Vancouver from 18-22 June**. In the meantime, we've reviewed a promising **CVPR highlights paper** proposing a new joint pose-NeRF training strategy designed to be more robust in the real world.

As CVPR prepares to take Canada by storm, the medical image analysis community is gearing up for its next big thing – **IPMI 2023 in Argentina!** Early bird registration ends on 15 May, and spots are still available for this fascinating **MICCAI-endorsed event**.

Get ready to dive into another compelling **application of deep learning** with Ioannis Valasakis as we explore the use of computer vision for blood cell classification. This medical application has broad implications, and we have all the details (**including code!**) for you.

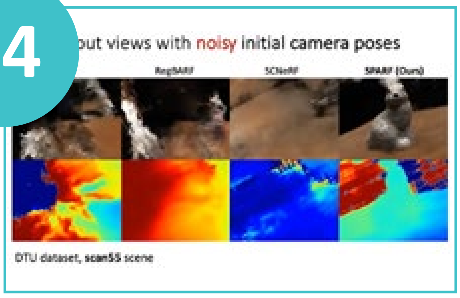
Two friends of Computer Vision News, **Sebastien Ourselin** and **Tom Vercauteren**, are involved in an exciting startup allowing surgeons to **go beyond the boundaries of human vision**. Read on to find out more about **Hypervision Surgical** and its revolutionary plans.

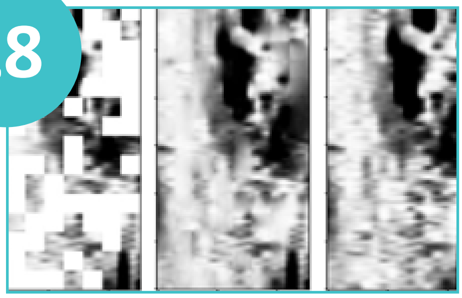
We hope you enjoy this month's feature-rich content! Remember to share our link with others and encourage them to subscribe for free.

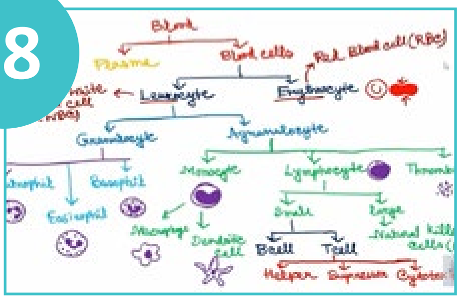
Ralph Anzarouth,
Editor, **Computer Vision News,**
Marketing Manager, **RSIP Vision**

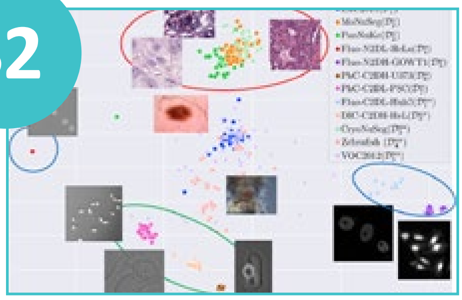
Computer Vision News

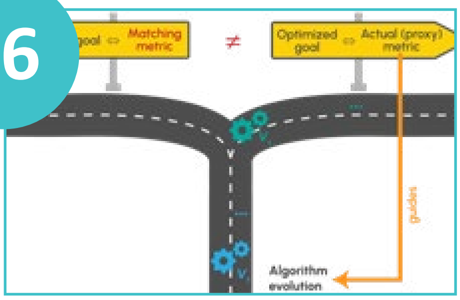
Medical Imaging News

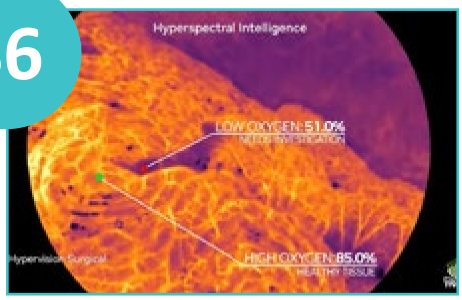
04


28


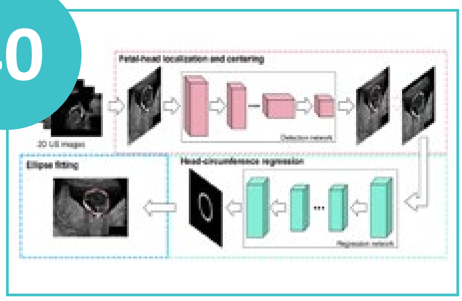
08


32


16


36


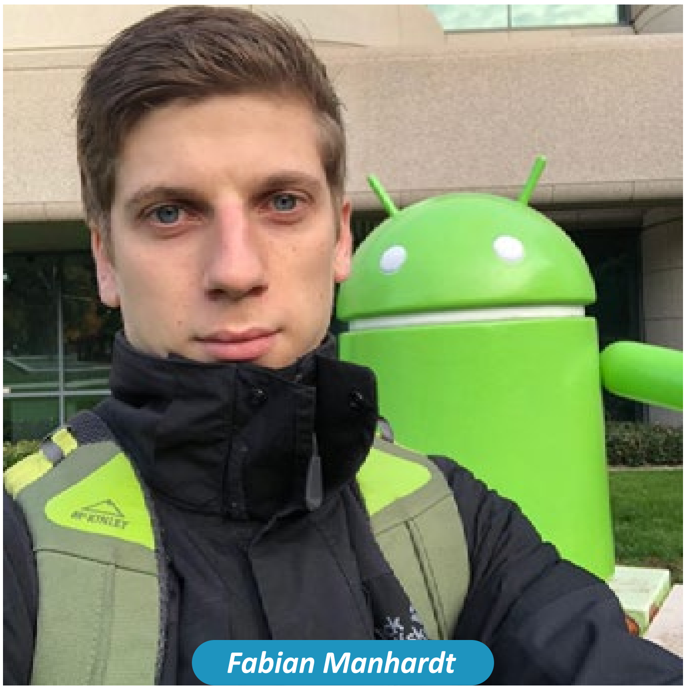
18


40


SPARF: NEURAL RADIANCE FIELDS FROM SPARSE AND NOISY POSES

Prune Truong is a fourth-year PhD student in the Computer Vision Lab at ETH Zurich and an intern in Federico Tombari's team at Google Zurich, where Fabian Manhardt is a Research Scientist. Prune and Fabian speak to us about their paper proposing a new joint pose-NeRF training strategy designed to be more robust in the real world, which has been accepted as a highlight at CVPR 2023.

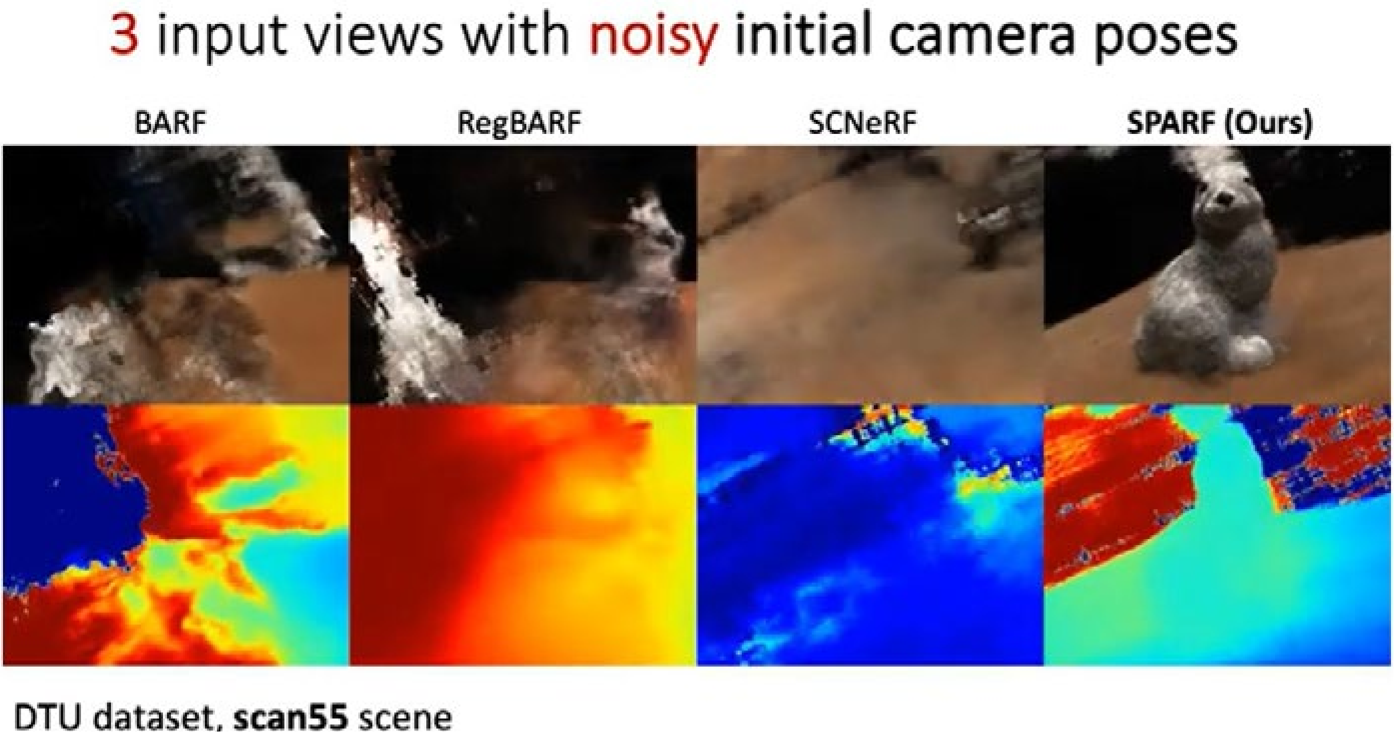

Prune Truong


Fabian Manhardt

NeRF (Neural Radiance Fields) is a cutting-edge technology with remarkable potential in generating 3D reconstructions and rendering novel views. NeRF has been shown to work best under two conditions: dense coverage of the 3D space and highly accurate camera poses. This scenario limits its application in the real world, where input views are often sparse, and poses are noisy.

"When you train a NeRF model on only a few images, it will instantly overfit," Prune explains. *"You'll have very nice training renderings, RGB will be good, and the photometric loss will be low, but when you look at the depth renderings, you'll see that the model doesn't learn any meaningful geometry. Therefore, it will be really bad if you try to render a novel view."*

The standard process to estimate per-scene poses is to use a **Structure-from-Motion approach, such as COLMAP**, which works well with many input views. However, with fewer views or an increased baseline between the images, it becomes much more challenging, and the pose estimation results are degraded.

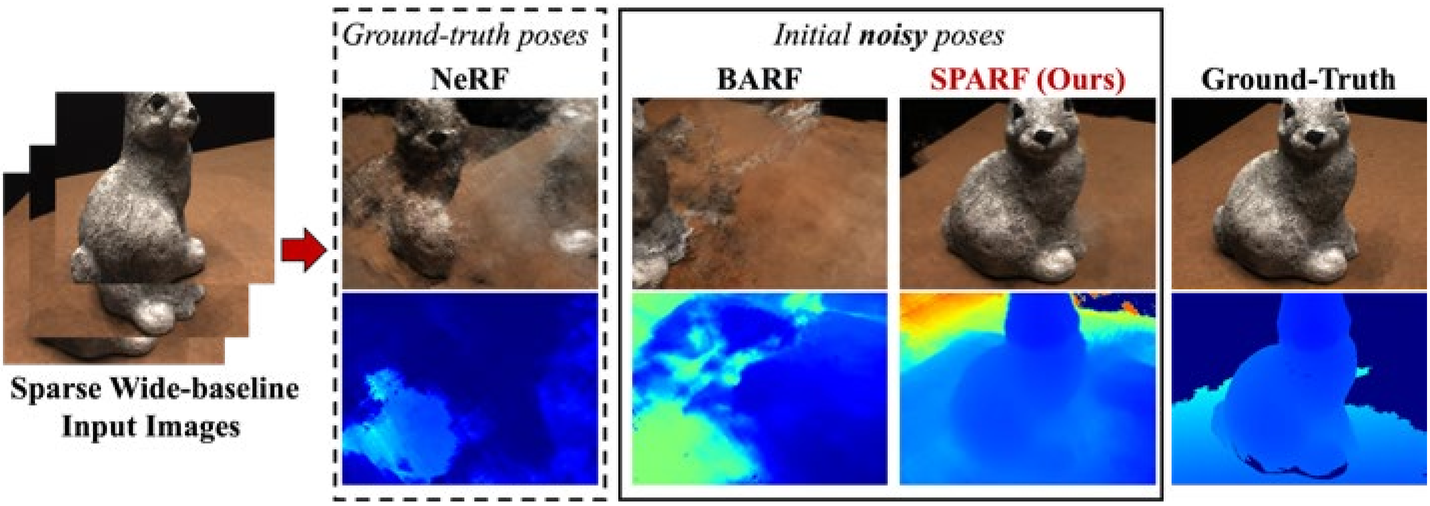


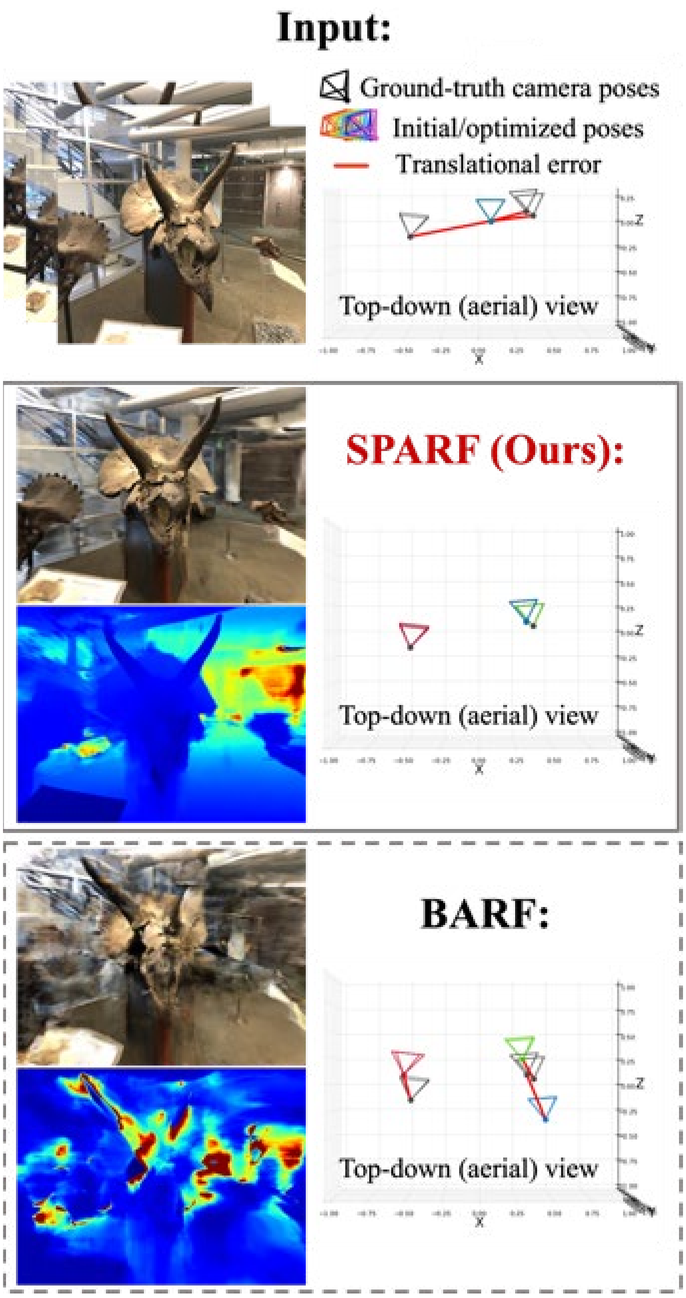
This paper turns the **best-case scenario on its head**, aiming to address the challenge of novel-view synthesis based on a neural field representation using as few as two to three views and noisy poses.

The **potential applications** of this technology are significant. In **robotics**, it could capture 3D reconstructions from a few images, saving time and resources. It could also be used in **AR or VR applications**, such as remodeling apartments, with users feeding images of furniture or other objects into the model to generate novel-view renderings and visualize how it would look in their living room.

After completing a literature review, Prune discovered that several works had already tackled sparse input views and found solutions to help with the overfitting problem. However, these works typically used perfect ground truth poses. There had also been works on a joint pose-NeRF refinement, but they assumed dense images.

"All these NeRF papers assume there are





hundreds of images, but you don't want to have to take hundreds of images of an object," Fabian points out. *"You take a few, and then this problem becomes severe because the poses you get are usually bad, and you have both problems simultaneously. That's why* ***this is an important research direction for real applications!****"*

Prune agrees: *"For us, it was still interesting. Even though this joint pose-NeRF refinement didn't work for the sparse scenario, I started from this and worked my way up by adding new constraints."*

For the constraints, Prune took inspiration from **multi-view problems and bundle adjustments**, which are well explored in **classical computer vision and geometry**, and set about integrating them into the NeRF framework.

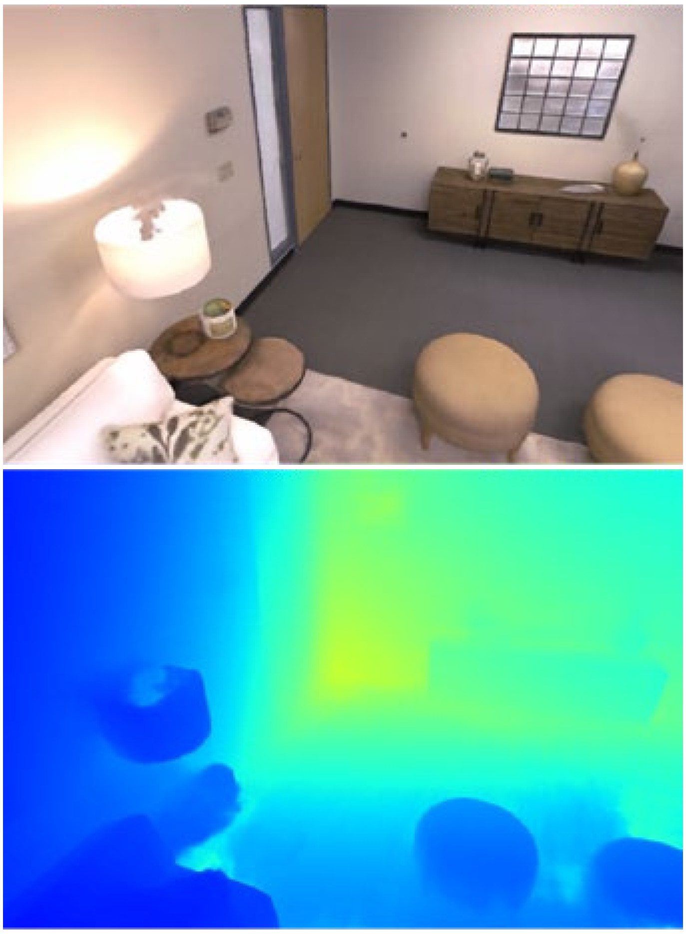
SPARF can train in a much shorter window, but the output is still far from the rendering quality that could be achieved with dense views. Prune is keen to point out that although they have made great strides, there is still a road ahead.

"There's still a lot of work towards getting perfect rendering from only two or three views," she tells us. *"We can do this joint pose-NeRF refinement much better than before, but we rely on point correspondences between the different views. The problem is getting these point correspondences is itself a research problem.* ***It would be great if there were a way to train or refine everything together.*** *For example, if you're training the NeRF and refining the poses based on the correspondences, could you also update the correspondences based on the NeRF and the poses? Like a system where everything can get better all at once. We haven't reached that point yet!"*

Fabian continues: *"It's a chicken and egg problem. COLMAP uses correspondences to get the poses, but you need the poses to get the correspondences.* ***We need a tool that does it all at the same time.****"*

Could this be the seed for their next paper or perhaps open the door for other researchers?

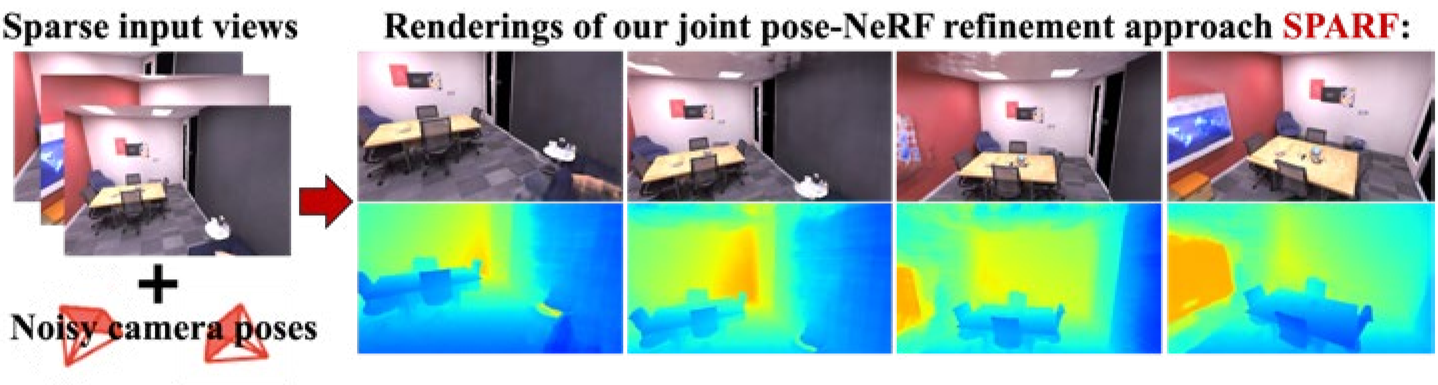
"That would be pretty nice!" Prune responds. *"We did a couple of experiments in that direction, and it's turning out more challenging than we hoped, but it would be a very nice output and future direction."*



As the first author of the work, Prune will present at **CVPR 2023 in June**. She worked on the code during her internship at **Google**, where in the beginning, she was a newbie to the NeRF field and can remember being surprised at the results when the state-of-the-art methods were applied to sparse views.

"I was like, how can it be so bad?" she recalls. *"I'm most proud of the difference in outputs we can now get. That's because I saw how it was initially, and I thought, how are we ever going to be able to get something reasonable? Now, we get something that looks realistic."*

Fabian adds: *"The most important thing is that rather than having to record for hours,* ***you can just take a bunch of images or crawl the internet.*** *That opens a lot of doors and makes it more accessible for everybody!"*



IOANNIS VALASAKIS, KING'S COLLEGE LONDON

@WIZOFE

Hi everyone! Today, we're diving into another fascinating application of deep learning in medicine: blood cell classification using computer vision. This has far-reaching implications, from diagnosing blood disorders to analyzing infection responses. It’s a medical application indeed, but I’m actually using a super-nice version of the net we developed in the lab and it has great use in the broad field of computer vision applications! So, let's get started!

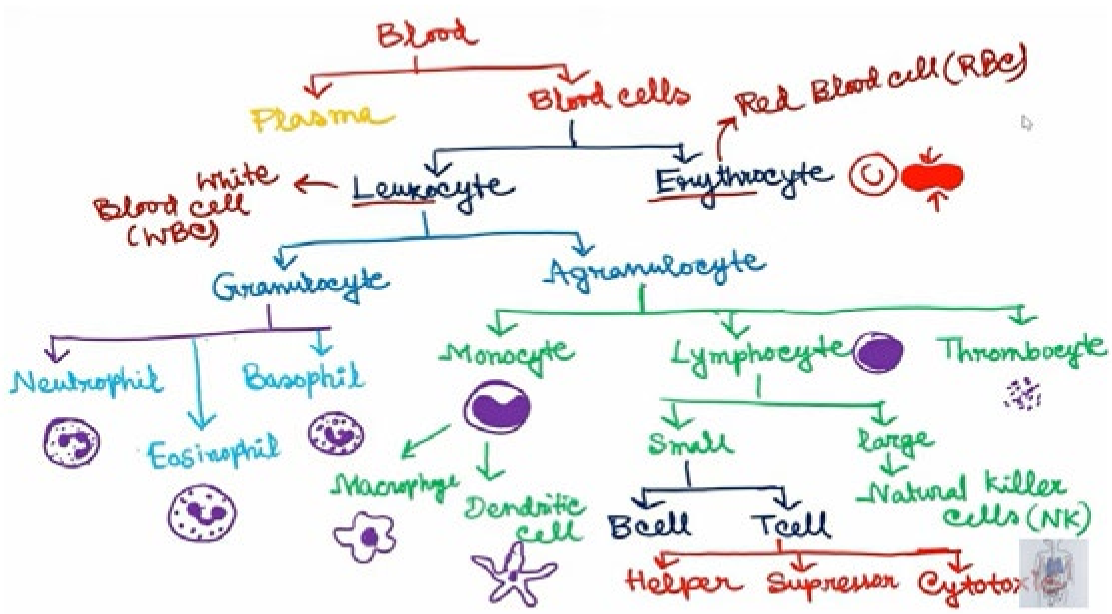
Introduction

Blood cell classification is crucial in diagnosing various medical conditions, such as anemia, leukemia, and infections. Traditional classification methods often involve manual examination of blood samples under a microscope, which can be time-consuming and prone to human error. Deep learning models, particularly convolutional neural networks (CNNs), have shown great potential for automating this process with high accuracy.

One popular CNN architecture for image classification tasks is the ResNet, which employs residual connections to facilitate the flow of information between layers. In this article, we will explore how ResNet can be adapted for blood cell classification.

Blood cell classification can be performed on different types of cells, such as red blood cells, white blood cells, and platelets. Here, we will focus on classifying white blood cells, which play a vital role in our immune system. You can find real blood cell images online, such as this dataset on Kaggle: [White Blood Cell Classification Dataset](#).

Here is a very useful video for understanding blood cells, their types and why we attempt the classification approach:



Here's another one from NVIDIA:

Approaches

This is an image showing the different types of blood cells which we usually need to know before attempting any classification:

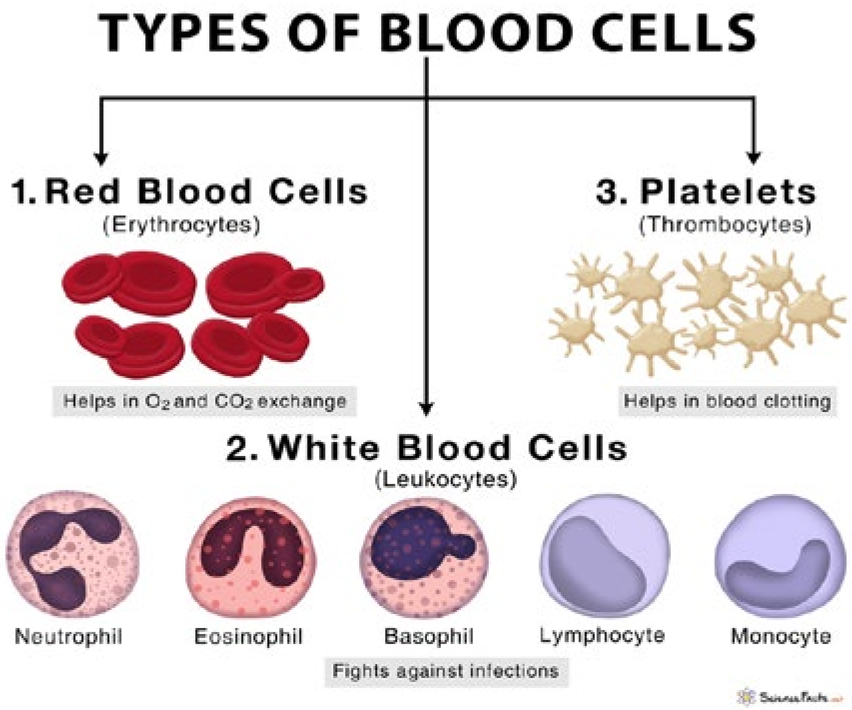


Image 1: Blood cell types, adapted from Sciense Facts

A common approach to blood cell classification is to use transfer learning with a pre-trained ResNet model. Let's see how to implement this using TensorFlow and Keras:

```
import tensorflow as tf
from tensorflow.keras import layers, models
from tensorflow.keras.applications import ResNet50
# Load pre-trained ResNet model
resnet = ResNet50(weights='imagenet', include_top=False, input_shape=(224, 224, 3))
```

```
# Add classification head
x = layers.GlobalAveragePooling2D()(resnet.output)
x = layers.Dense(256, activation='relu')(x)
x = layers.Dropout(0.5)(x)
x = layers.Dense(4, activation='softmax')(x)

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

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

# Load blood cell images and train the model
# See the reference Kaggle
```

In this example, we load a pre-trained ResNet model and add a classification head to the top of the network. The classification head consists of a global average pooling layer, followed by two fully connected layers with dropout and a softmax activation function for multi-class classification. During training, we can use a categorical cross-entropy loss function to optimize the model.

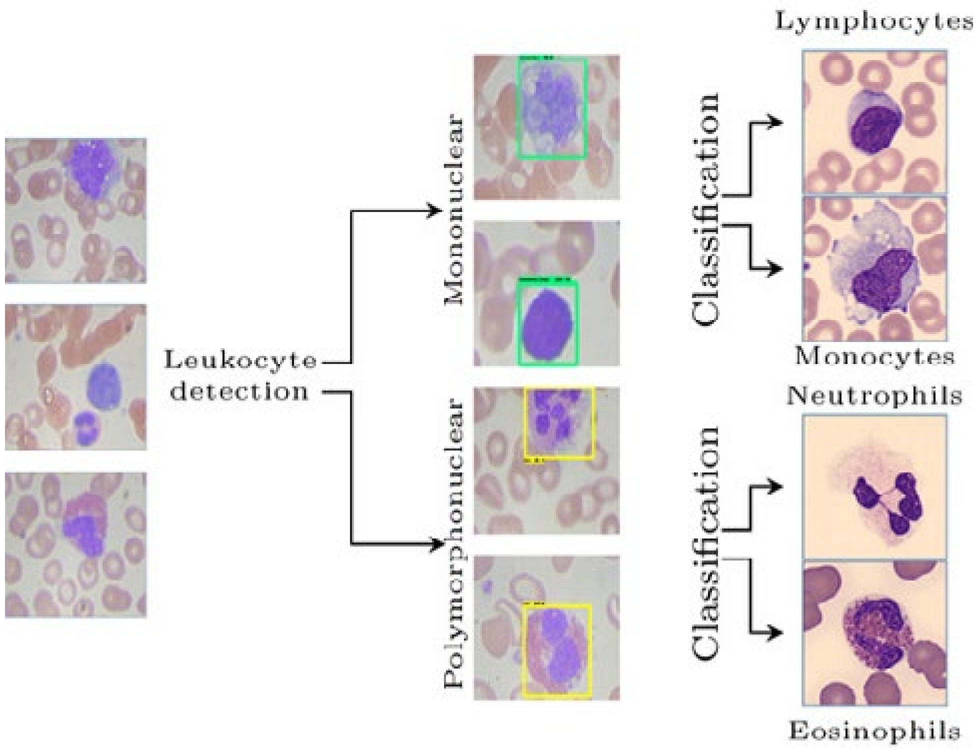


Image 2: Leykocyte detection: An application of Mono- vs Polymorpho- nuclear from Cheuque, C.; Querales, M.; León, R.; Salas, R.; Torres, R. An Efficient Multi-Level Convolutional Neural Network Approach for White Blood Cells Classification. Diagnostics 2022*, 12, 248.* https://doi.org/10.3390/diagnostics12020248

Another approach to blood cell classification is to use data augmentation techniques. This involves applying random transformations to the input images, such as rotation, scaling, and flipping, to increase the size and diversity of the training dataset. Let's see how to implement data augmentation using the ImageDataGenerator class in Keras:

```
from tensorflow.keras.preprocessing.image import ImageDataGenerator
# Create an ImageDataGenerator with data augmentation
data_gen = ImageDataGenerator(rotation_range=20,
                              width_shift_range=0.1,
                              height_shift_range=0.1,
                              horizontal_flip=True,
                              vertical_flip=True)
# Load blood cell images and apply data augmentation
# exercise for you!
```

In this example, we create an ImageDataGenerator with various data augmentation options. During training, the ImageDataGenerator will randomly apply these transformations to the input images, effectively increasing the size and diversity of the dataset.

We will now expand upon this example by adding data preprocessing, training, and evaluation steps. Additionally, we will provide two more coding examples: one for data augmentation using TensorFlow and another for fine-tuning a pre-trained ResNet model for blood cell classification.

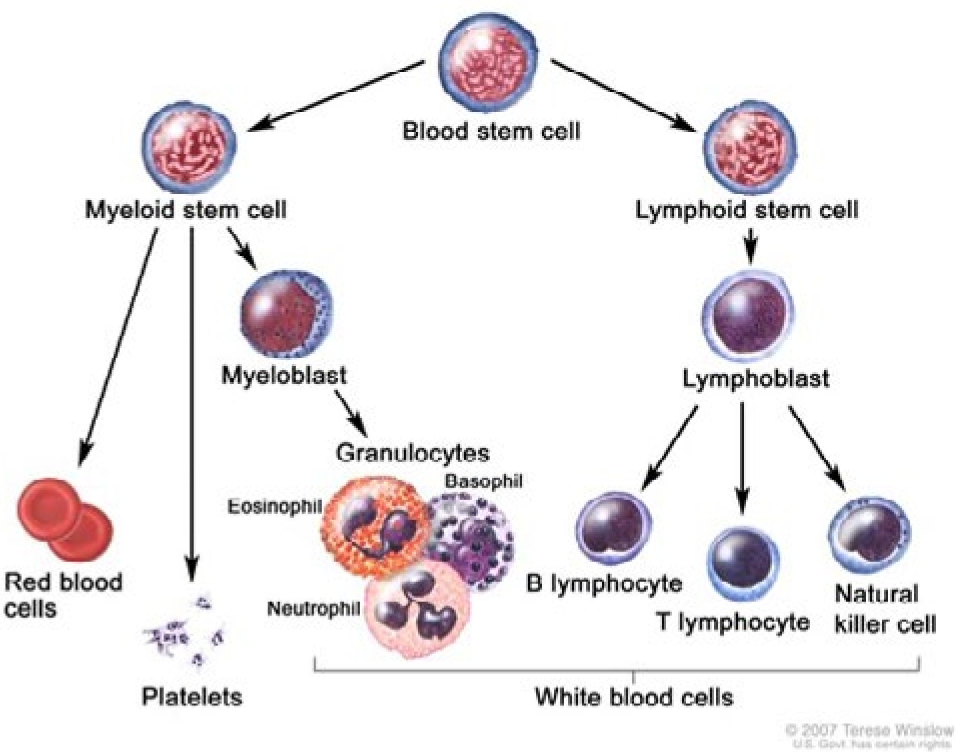


Image 3: Definition of white cells © Terese Winslow

Expanded 3D Densenet example:

```
import numpy as np
from keras.models import Model
from keras.layers import Input, Conv3D, Dense, Dropout, GlobalAveragePooling3D
from keras.applications.densenet import DenseNet121
from keras.preprocessing.image import ImageDataGenerator
from keras.optimizers import Adam
from sklearn.model_selection import train_test_split

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

# Load your dataset (X and y)
# X: (num_samples, depth, height, width, channels)
# y: (num_samples, n_classes)

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Create the 3D Densenet model
input_shape = (224, 224, 224, 3)  # Replace with your input shape
n_classes = 2  # Replace with your number of classes
model = build_3d_densenet(input_shape, n_classes)

# Compile the model
```

```
optimizer = Adam(lr=1e-4)
model.compile(optimizer=optimizer, loss='categorical_crossentropy', metrics=['accuracy'])

# Train the model
batch_size = 8
epochs = 20
model.fit(X_train, y_train, batch_size=batch_size, epochs=epochs, validation_split=0.1)

# Evaluate the model
scores = model.evaluate(X_test, y_test)
print("Test loss:", scores[0])
print("Test accuracy:", scores[1])
```

And for a final treat, what if we finetune our network?? Back to Tensorflow ResNet and I'll leave the comparison to you as an exercise:

```
import tensorflow as tf
from tensorflow.keras import layers
from tensorflow.keras.models import Model
from tensorflow.keras.applications import ResNet50
from tensorflow.keras.optimizers import Adam
from tensorflow.keras.preprocessing.image import ImageDataGenerator

def build_resnet(input_shape, n_classes):
   inputs = tf.keras.Input(shape=input_shape)

   # Load pretrained ResNet model
   base_model = ResNet50(include_top=False, weights='imagenet', input_shape=input_shape)

   # Add custom layers
   x = base_model(inputs, training=False)
   x = layers.GlobalAveragePooling2D()(x)
   x = layers.Dense(256, activation='relu')(x)
   x = layers.Dropout(0.5)(x)
   outputs = layers.Dense(n_classes, activation='softmax')(x)

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

# Load your dataset (X and y)
```

```
# X: (num_samples, height, width, channels)
# y: (num_samples, n_classes)

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Create the ResNet model
input_shape = (224, 224, 3)  # Replace with your input shape
n_classes = 4  # Replace with your number of classes (e.g., 4 for red blood cells, white blood
cells, platelets, and background)
model = build_resnet(input_shape, n_classes)

# Compile the model
optimizer = Adam(lr=1e-4)
model.compile(optimizer=optimizer, loss='categorical_crossentropy', metrics=['accuracy'])

# Data augmentation
train_datagen = ImageDataGenerator(
   rescale=1./255,
   rotation_range=20,
   width_shift_range=0.2,
   height_shift_range=0.2,
   shear_range=0.2,
   zoom_range=0.2,
   horizontal_flip=True,
   fill_mode='nearest')

test_datagen = ImageDataGenerator(rescale=1./255)

# Train the model
batch_size = 32
epochs = 20
model.fit(train_datagen.flow(X_train, y_train, batch_size=batch_size),
        steps_per_epoch=len(X_train) // batch_size,
        epochs=epochs,
        validation_data=test_datagen.flow(X_test, y_test, batch_size=batch_size),
        validation_steps=len(X_test) // batch_size)

# Evaluate the model
scores = model.evaluate(test_datagen.flow(X_test, y_test, batch_size=batch_size))
print("Test loss:", scores[0])
print("Test accuracy:", scores[1])
```

Conclusion

In conclusion, the application of deep learning and computer vision techniques for blood cell classification has the potential to revolutionize diagnostic accuracy and efficiency in the medical field. By automating the identification and categorization of different blood cell types, physicians can spend less time analyzing blood samples and more time focusing on developing targeted treatment plans, ultimately leading to improved patient outcomes.

Furthermore, as deep learning models continue to evolve and improve, we can expect even more breakthroughs in medical diagnostics and other healthcare applications. The use of advanced models, such as ResNet and DenseNet, has already demonstrated remarkable success in image classification tasks, and their application to blood cell classification is just one of many potential areas for exploration.

In this article, we have introduced the concept of blood cell classification, provided examples of how to implement deep learning models for this task using TensorFlow and Keras, and discussed the potential benefits of these techniques in medical diagnostics. We have also highlighted the importance of data augmentation in improving the performance of our models and the use of pre-trained models to accelerate the development process.

As we move forward, it is essential for researchers and practitioners alike to continue exploring novel deep learning approaches and techniques, pushing the boundaries of what is possible in medical diagnostics and other healthcare applications. The integration of deep learning into medicine is still in its early stages, and there is much to be discovered and developed, as we strive to create more effective and efficient healthcare solutions for patients worldwide.

Next month

Stay tuned for more fascinating insights into the world of deep learning and its transformative potential in healthcare!

As always, feel free to reach out with any questions or suggestions. We love hearing from our readers! In the meantime, keep exploring, stay curious, and enjoy the incredible possibilities that deep learning has to offer!

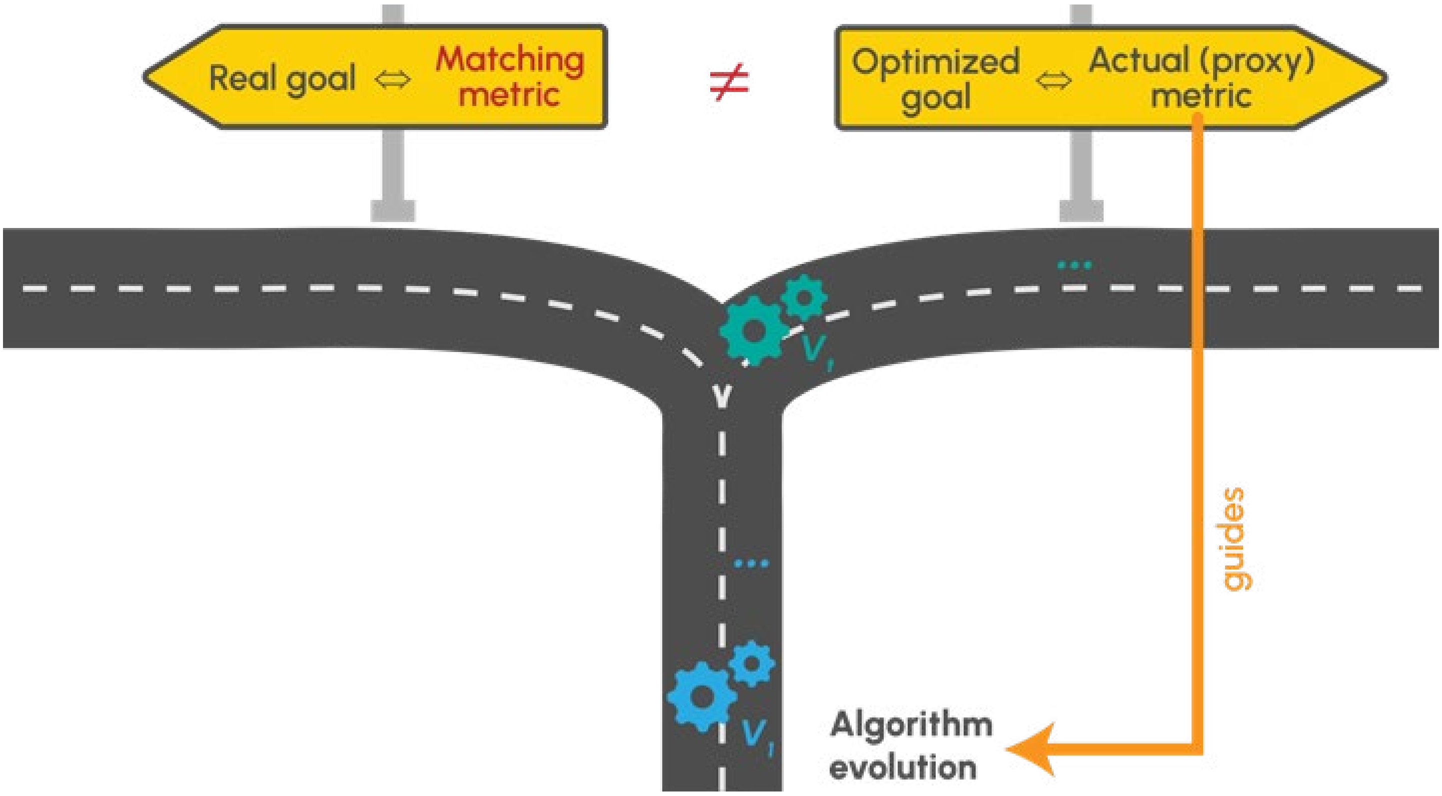
Annika Reinke recently defended her PhD at the German Cancer Research Center (DKFZ). Her research focuses on the validation of biomedical image analysis algorithms - making sure that our algorithms do what they are supposed to do. After completing her PhD, she will continue working on the topic as a postdoc at the DKFZ. Congrats, Doctor Annika!

Why do we need to care about algorithm validation?

In fact, proper validation of AI algorithms is frequently neglected in favor of a strong focus on the development and exploration of new models. This research practice can, however, be really risky as it may propagate poorly validated algorithms that could cause adverse outcomes for patients! Thus, a thorough and high-quality validation is crucial for any algorithm to potentially be used in clinical practice. This particularly holds true for image analysis competitions (challenges), which have emerged as the standard technique for comparative assessment of AI algorithms and determining which is the most effective in solving a certain research question. Given the tremendous importance of challenges for the research, it is surprising that hardly any attention has so far been given to quality control.

Validation pitfalls

During her PhD, Annika and team raised awareness of severe flaws in challenges and algorithm validation. She evidenced how effortlessly both challenge participants and organizers could, in theory, manipulate challenges by taking advantage of security holes in the challenge design. She also found that researchers typically favor common performance metrics without being aware of numerous pitfalls pertaining to their use. She even put together a comprehensive list



of metric pitfalls. When comparing algorithms against each other, she demonstrated that rankings are typically unstable, meaning that an algorithm could be the best simply due to the nature of a ranking calculation and not due to actually being the best fit for solving a particular research task. Finally, she found that challenges and submitted algorithms are typically not reproducible.

Improving validation

Uncovering problems is good, solving them is even better! To overcome the described issues, Annika and her team proposed several improvements: a structured challenge submission system now collects comprehensive information about challenge designs, which can be critically reviewed by independent referees. To promote the selection of validation metrics based on their suitability to the underlying research problem rather than popularity, she proposed a problem-driven metric recommendation framework that empowers researchers to make educated decisions while being made aware of the pitfalls to avoid. To enable uncertainty-based ranking analysis, she presented an open-source toolkit including several advanced visualization techniques for benchmarking experiments. To facilitate and enhance challenge transparency, she presented a guideline for challenge reporting and introduced challenge registration, i.e. publishing the complete challenge design before execution. Finally, she showed that challenge results can be used for a dedicated strength-weakness analysis of participating algorithms, from which future algorithm development could heavily benefit in addressing unsolved issues.

We hope that Annika's work paves the way for high-quality and thorough algorithm validation, which is crucial to avoiding translating inefficient or clinically useless algorithms into clinical practice.

Abby Stylianou is an assistant professor of computer science at St. Louis University. She is also a fellow of the Taylor Geospatial Institute.

Over 100 inspiring interviews with successful Women in Computer Vision in our archive!

Wow! So you are busy with two roles! Which one will we start with?

[*laughs*] They go together. It's an appointment that I have as a faculty here at St. Louis University. I got appointed this past year as the inaugural fellow of a new eight institution geospatial institute that is really looking to push the limits on geospatial technology at large.

Can you tell us more about it?

Absolutely! I work broadly in computer vision and machine learning. But my specific research has focused on global-scale image retrieval or image search. And I've focused particularly on one application, which is combating child sexual abuse by recognizing where victims of child sexual abuse are photographed. I work with the National Center for Missing and Exploited Children here in the United States. We have an investigative platform for when they are investigating cases of child sexual abuse that, it turns out, are often taken in hotel rooms. They can recognize what hotel it was taken in.

This is a project that started while I was actually still a PhD student. It was with Robert Pless, who was my PhD advisor at Washington University in St. Louis: he was amazingly supportive and kind to let me have free reign to do what I wanted to do. I appreciate that he let me take this project on. He's now at George Washington University. And then another collaborator, actually an academic sibling of mine, Richard Souvenir, who's at Temple University. He and I built this starting in 2016, and we've been working on it since.

I know Robert. We were introduced at CVPR 2016, in Las Vegas.

I was there. I was pregnant with my first child.

How does a computer vision academic get into this, rather than counting trees in a forest or that kinds of things?

My story is a strange one. My undergraduate degree is not in computer science. I have an undergraduate degree in environmental studies and geoscience. It's a geology degree. Then I discovered that I thought rocks were really boring, and I was trying to figure out what I wanted to do with my career. I had taken a remote sensing class that I really liked. So I was looking for jobs straight out of undergrad, and I was lucky enough to land in Robert's lab as a research assistant. At the time, he was funded by IARPA, which is the Intelligence Advanced Research Projects Activity. They were

working on the Finder project, which is this idea that I have a picture and nothing else but just the pixels in the picture. How can I figure out where in the world it was taken? So I was working with Robert on tools for the really fine-scaled refinement of camera geolocalization. Which is, if I sort of have a rough estimate of where I am, how can I really precisely define the parameters of the camera that took that picture? That was in 2012 when I started working for him. And in 2013, there was a story in our local paper about a 1980s murder victim who had been buried in a cemetery in North St Louis. The case had run cold. They didn't know who she was. And they wanted to exhume her to do modern forensic analysis, but they couldn't find her gravesite. They had pictures of the burial, though, from the 1980s. It is a testament to Robert that I've been able to have the career that I have. He gave me full support to take the things we were developing for the intelligence agencies and to instead go and help on this cold case to figure out where the 1983 picture was taken. So we went out and tried to find things like the headstones and the billboards, the things that were in the cemetery in the 1980s. If we could find where those were today, then it's a fairly straightforward optimization problem to solve for the parameters of the camera that could have taken the picture. And one of those pictures was taken right on the corner of the grave.

Was this a eureka moment for you?

It was, absolutely! This changed the trajectory of my career. We were able to identify her grave. As I said, the camera geometry is fairly straightforward. This was a point at which I really recognized that I was able to do something good with

the skills that I learned. And so, I took that as a direction for my career. Then I got introduced to this group of women here in St. Louis called the Exchange Initiative. They had this idea to build a database of hotel room images. Their idea was just to build the database to make it easier for the investigators to scroll through instead of having to go and search on Google Photos or something. I approached them about making this sort of AI-based approach or a machine learning-based approach rather. They were really excited. But it all stems from having helped on this cold case, and then taking that and saying, okay, let's keep that as a focus.

It’s not the only thing that I do. In St. Louis, we have a lot of plant science expertise. There's the Donald Danforth Plant Science Center. So I also work really significantly on agricultural applications. I have two paths in my career. One is image retrieval for societally beneficial combating child sexual abuse. The other is thinking about agricultural applications. And where those intersect to me is that it is like thinking about humanity. In agricultural applications, what we're really thinking about is making better crops so that we can feed people better or more healthily or grow things in a climate that is changing very rapidly. Coming up with computer vision-based approaches to understanding the health or the vitality of plants that are in the field at a large scale. I consider myself extremely lucky to have this emphasis on doing something good while at the same time having really interesting computer vision questions to think about. It's also really challenging from an image retrieval perspective. So there have been opportunities for algorithmic development.

I'm so happy for you! Tell me something

about the system that you have put in place. Is there one thing that you would like your computer to tell you, and it does not tell you yet?

[*laughs*] That's a great question. I would love to have much better models of uncertainty. This is something we've been thinking about in recent work that is not published yet but hopefully will be soon. In our image search tool for investigators, the way we present results is we say you gave us this image. Here are the thousand images that we may think are most similar to the hotels that they came from, but that have no kind of notion of uncertainty. And I think it would be a much more useful system for investigators if we could say, I actually don't know what image is most similar, but I really think it's this particular hotel chain, right? I don't have a good image retrieval match, but I have a reasonable estimate of some classes. Or, I think that it's from this geographic region. The hierarchical notion of uncertainty is something that I've really been thinking a lot about lately. It's not super well integrated into a lot of the image retrieval approaches that exist. But I think it would make our tool much more powerful for investigators rather than just always returning an image search result. Here's a similar image that's not actually that similar. If we could say I don't actually know what a similar image is, but I sure think that that is from New York City. Then we would explain to an investigator why.

We have work that I've published at WACV, going back a few years now, trying to explain image similarity results. I get two images that I say are similar. How can I explain to an investigator why the model anthropomorphizes. I think we've made heat maps as the approach so far. It says

like, here's the thing in the query image that we focused on, and here's the thing in the resulting image. Much of computer vision when we're trying to do explainability, we make heat maps. They're kind of a horrible way of explaining anything. It doesn't really help, especially for somebody who's not an expert, which is what we're really dealing with, what the model thinks is important. So I think at having better natural language explanations, for example. We've really had this very recent explosion of impressive language models. I would love to be able to tell an investigator the reason we think this is similar. It is the red couch or something like that. And to give them the ability to correct that and say like, no, I don't care about the red couch. Focus on this. You're hearing all of the things that I'm writing into my next grant!

Isn't the scientific community making a lot of progress dealing with uncertainty?

They are. This is clearly something that, broadly, there's a lot of interest in. I think it's something that I've not seen as much in the image retrieval community. I will admit that it just may be that I have missed important work. There's so much coming out of our field, right? It is entirely possible there is important work that I have not seen yet. But no, I think as a field, we do generally realize we need much better notions of uncertainty and confidence. But it's something I really want to build into our models because I think it's really important for investigators.

You've told us a lot about your current work. And I see that you have several decades of career still in front of you to do. Do you see yourself on this path forever?

[*laughs*] It's funny when I look at my

career so far, it has been very driven by happenstance and sort of coming across the right thing at the right time. So I think I have general ideas of things that I care to continue working on, which is thinking about how computer vision can be used to help people that need help. I'm really interested in that very broadly. But you're right that working to combat child sexual abuse by recognizing hotels is one very specific thing. So I think my plan for the future is generally to keep thinking about how I can take the things I've already done and keep doing a little bit of good in the world. Here in St. Louis, there is a real opportunity to focus on the plant science side, and I have some really amazing collaborators here. One of the things I'm really interested in there is how we can use low-cost sensors to build better models of fields. This ties into thinking about the variety of types of people in the world and the technology that they have access to. If we think about large-scale production farming from big companies, they have the budget to have crazy technology out in their field.

St. Louis University is an interesting place for me. We didn't have a PhD program until a year ago. So in some ways, there have been challenges in being an effective researcher here. I haven't had the sort of support systems or research programs that my peers necessarily have had. What I have had is support to work on the problems that I think are important. I think it's important actually in the field for people to realize not every university is your top, such as Stanford, CMU, and ETH Zurich. We don't all have hundreds and hundreds of PhD students, and we're still doing really good and impactful work at these universities. So I sometimes struggle with this. I look at my peers, and I'm like, oh, I wish I could have ten CVPR papers a year! That's not the reality of where I'm at. But I get to work on work that I think is really important, and my university supports me to do that. That's something that's really special.

And that's what you wanted, or it's a goal that came together with time?

I've mentioned I have two kids. I have a six-year-old, and I have a one-year-old. I also have my entire family in St. Louis. So I live in the house behind my parents. We bought the house behind my parents. I have two brothers that also live in St.

Louis with young children. I wasn't leaving St. Louis. I wanted to be with my family. So I invited the chair of the Computer Science Department at St. Louis University at the time to my dissertation defense at Washington University. And I said, I don't think you have a position right now, but I think I do really cool work, and I'd be interested in joining your faculty if there were ever an opportunity.

And the rest is history?

Yes! He made a position for me. He got the research institute at St. Louis University and said, "Let's create a position!" And I've been here since 2019.

I would encourage PhD students: if there is something that you're passionate about, find ways to integrate that into the work you're doing. Push to have a paper about that. And I would encourage PhD advisors, research advisors, or research mentors to let people do that. I think that's something we all fight for. We need the CVPR publication. We need the next grant, and they're not necessarily aligned, but we all need to make sure we're still doing something impactful in the world.

Over 100 inspiring interviews with successful Women in Science!

COMPUTER VISION EVENTS

WSCG - Computer Graphics
Pilsen, Czech Republic
15-19 May

ICDIP - Digital Image Processing
Nanjing, Chin
19-22 May

Embedded Vision Summit
Sta Clara, CA
22-25 May

Ital-IA
Pisa, Italy
29-31 May

ICRA - Robotics and Automation
London, UK
29 May - 2 June

IEEE Intelligent Vehicles Symposium
Anchorage, AK
4-7 June

IEEE Conference on AI
Sta Clara, CA
5-6 June

Robots and Vision
Montreal, Canada
6-8 June

AIWorld Congress
London, UK
7-8 June

CAOS
Pattaya, Thailand
7-10 June

FREE SUBSCRIPTION
(click here, its free)

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

Fill the Subscription Form
it takes less than 1 minute!

SUBSCRIBE!

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

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

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

MEDICAL IMAGING NEWS

MAY 2023

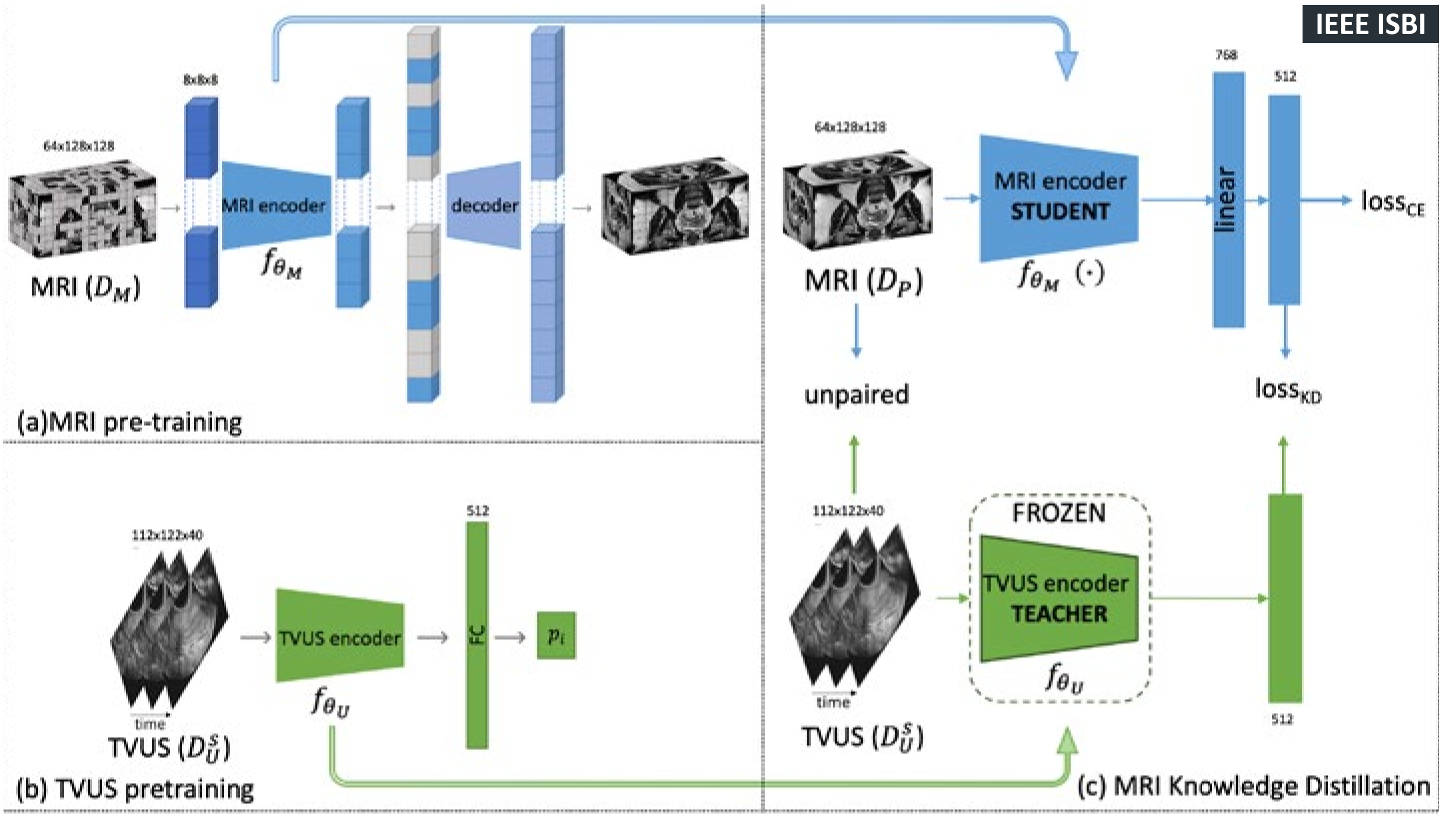
DISTILLING MISSING MODALITY KNOWLEDGE FROM ULTRASOUND FOR ENDOMETRIOSIS DIAGNOSIS WITH MAGNETIC RESONANCE IMAGES

Yuan Zhang is a second-year PhD candidate from the University of Adelaide in Australia. She joins us fresh from winning the Best Oral Presentation award at last month's International Symposium on Biomedical Imaging (IEEE ISBI 2023) in Colombia to tell us more about her important work on endometriosis diagnosis.

Endometriosis is a common gynecological disease affecting around one in nine women worldwide and can cause chronic pain, infertility, and other health problems for those who are affected. The condition occurs when endometrial-like tissue grows outside the uterus, and its diagnosis can often be challenging. **The pouch of Douglas (POD) obliteration**, one of its characteristic features, is typically detected using **transvaginal gynecological ultrasound (TVUS) scans and MRI**.

"These two modalities provide some distinct and complementary markers, so doctors get a better diagnosis using both, but patients are not usually able to access that," Yuan tells us. "Even in Australia, where investment is at the top level in the world, it's really hard to find doctors who can diagnose endometriosis using TVUS or MRI alone, let alone together."

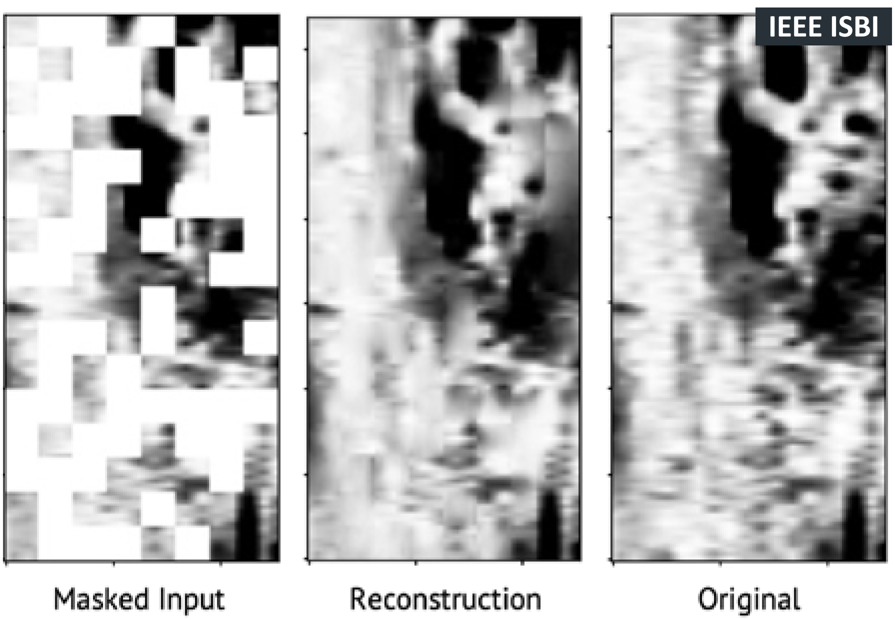
For TVUS, doctors can diagnose POD obliteration using the 'sliding sign' by mobilizing the patient's uterus during the

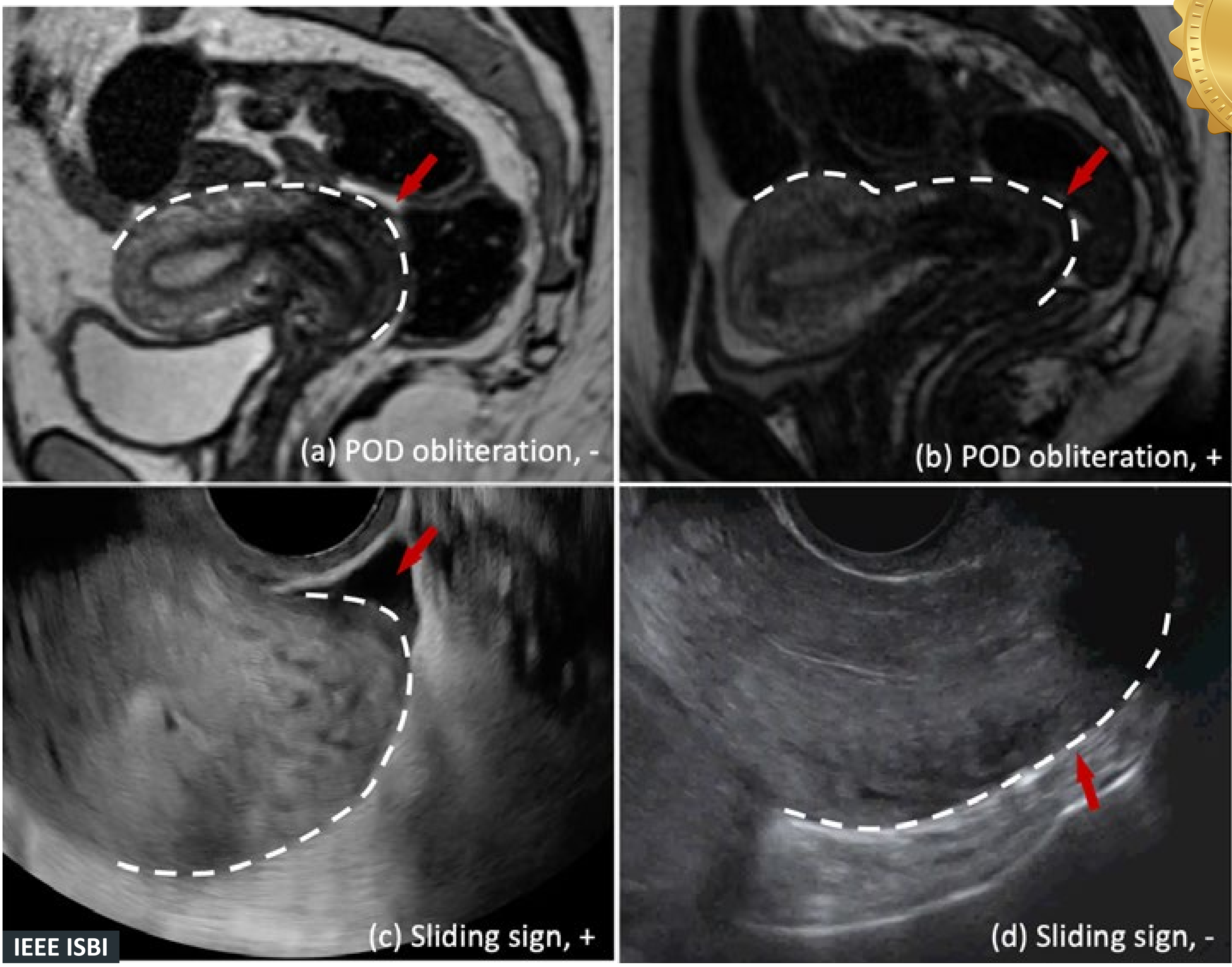


ultrasound to check if their organs freely glide over each other. However, **this is highly dependent on the operators' skills**. MRI provides more information, but again, **diagnosis requires an experienced clinician**, and it is generally more challenging to detect POD obliteration from MRI than TVUS.

To address this problem, Yuan has developed an **AI model to improve endometriosis diagnosis from MRI, using knowledge distilled from unpaired TVUS data**. To achieve this, she recognized the need for multidisciplinary expertise. The work is backed by a team of researchers and clinicians working together on a large project called IMAGENDO.

"I'm a computer science student, but this is a huge medical issue, and even on the medical side, it's still very challenging," she tells us. *"I needed to use some cross-field knowledge. Doctors and surgeons have helped me and explained everything. They're super busy but can always spare me 30 minutes of their time. They know how to diagnose this disease from MRI, so we can mimic what they do in our algorithm."*





Yuan used **data augmentation techniques** to create more samples from a small MRI dataset of only 89 cases. Preparing the data, she used the **3D AHE** technique to enhance the definition of edges, while a pre-training method called **3D MAE** helped the student encoder learn better about MRIs and achieve a promising reconstruction.

"The pre-training and fine-tuning datasets used in the model are very different, but we expect the pre-training to learn some patterns for normal or common MRIs," she explains. *"**With this pre-training, we see a big improvement in our model's performance!**"*

The team's focus on POD obliteration is the first step in their project. They plan to focus on other symptoms in the future, with the ultimate goal being that one day there will be a functioning system able to diagnose endometriosis for any sign based on any modality.

Winning an award at a prestigious international event is an outstanding achievement for Yuan in more ways than one. Born in Southwest China ("*Where the panda comes from!*"), not only is she working in a non-native language, but she is also an oral presentation winner among

people whose language is English, which is even more impressive. Did we also mention that it was her **first-ever conference**? We can see why this paper piqued the judges' interests, but what does she think was the secret of its success?

"To be honest, I always have impostor syndrome!" she reveals. *"I'm not very confident about my work, so when I got this award, I was really surprised. But I think it's a good algorithm. Plus, we don't just want to publish a paper;* ***we're trying to solve a problem for women that really exists and needs to be solved!****"*

Yuan would like to express her gratitude to the IMAGENDO project team, particularly her mentors and co-authors, **Gustavo Carneiro** and **Louise Hull**. She would also like to thank the **Australian Institute for Machine Learning (AIML)** and **Robinson Research Institute (RRI)** for their support.

"Endometriosis is such a common issue," she adds. *"I'm surprised it hasn't been solved yet because so many women are suffering, but we are just at the beginning of this research."*

Computer Vision News wishes to thank the ISBI Awards Co-Chair **Diana Mateus** for her very kind help.

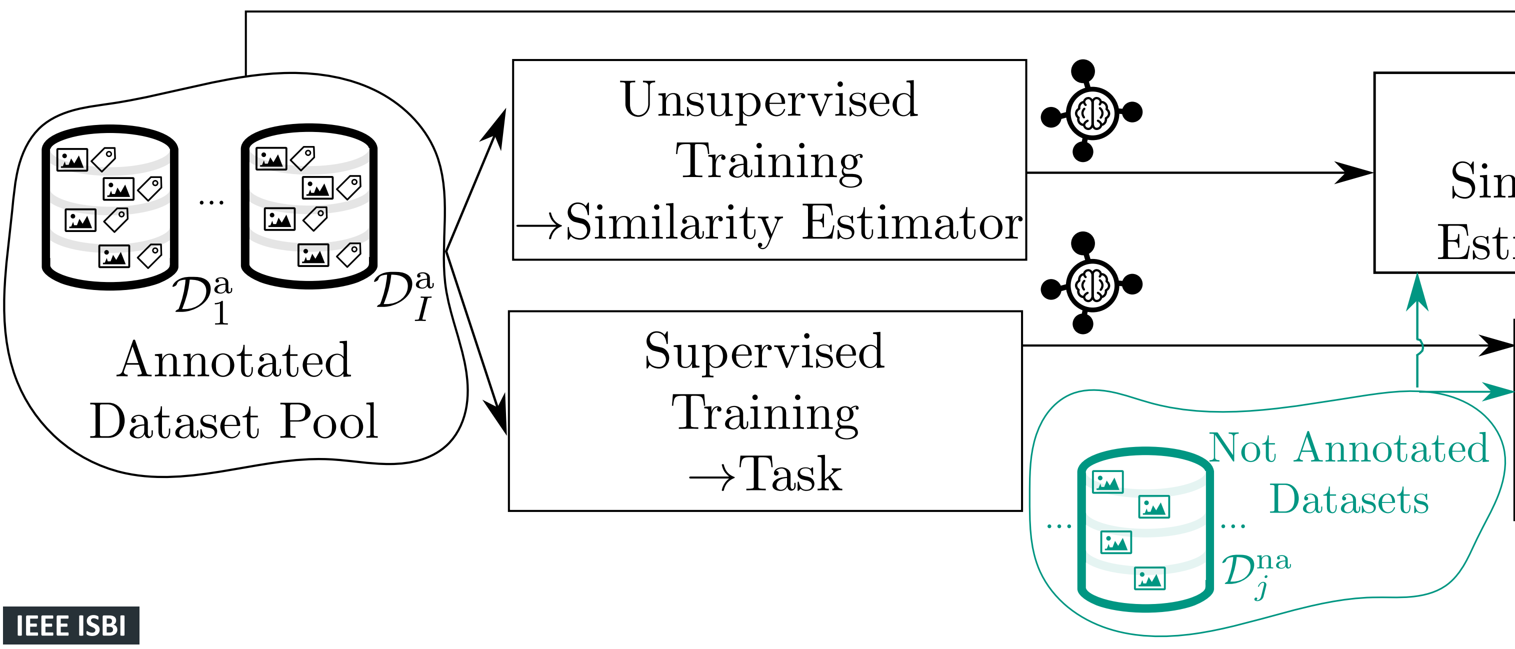
AUTOMATED STYLE-AWARE SELECTION OF ANNOTATED PRE-TRAINING DATABASES IN BIOMEDICAL IMAGING

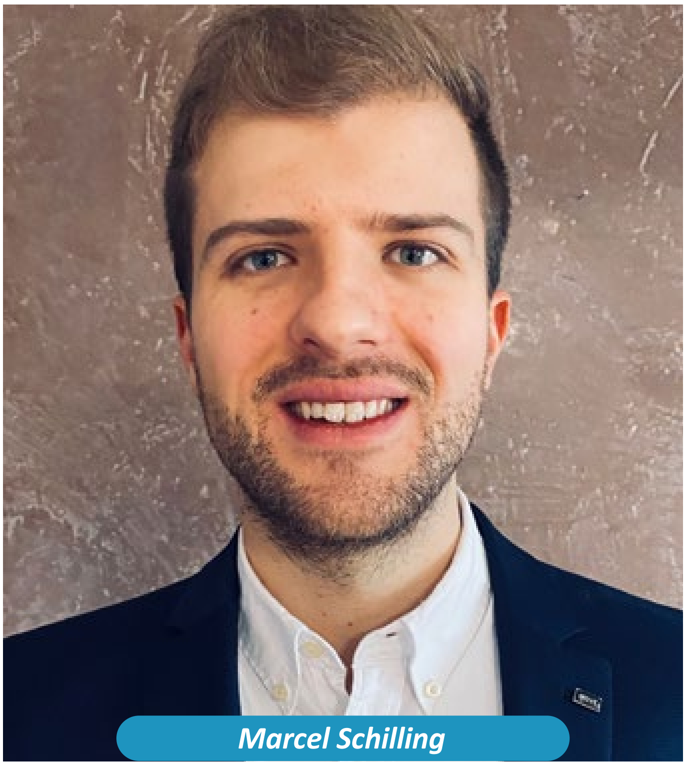
BEST PAPER ISBI 2023

Marcel Schilling is a doctoral student at the Institute for Automation and Applied Informatics at Karlsruhe Institute of Technology (KIT) in Germany. Miguel Molina-Moreno is a PhD student in the Multimedia Processing Group at the Charles III University of Madrid. They speak to us as the winners of the Best Paper Award at IEEE ISBI 2023 for their novel approach to tackling the problem of dataset similarity.

Deep learning relies heavily on big datasets being created, annotated, and made available for different tasks. However, these datasets often remain largely untapped due to the difficulties in performing transfer learning between them. **Transfer learning in medical imaging** is challenging due to the diversity of image modalities, organs, and cells and the scarcity of annotated data. Selecting an appropriate database for a given problem is still a significant hurdle.

"In the last few years, there has been a lot of focus on developing fancy new architectures to solve a specific task 0.5% better," Marcel tells us. *"However, this always assumes the **availability of annotated data**, which is not always practical in real-world applications. That's*

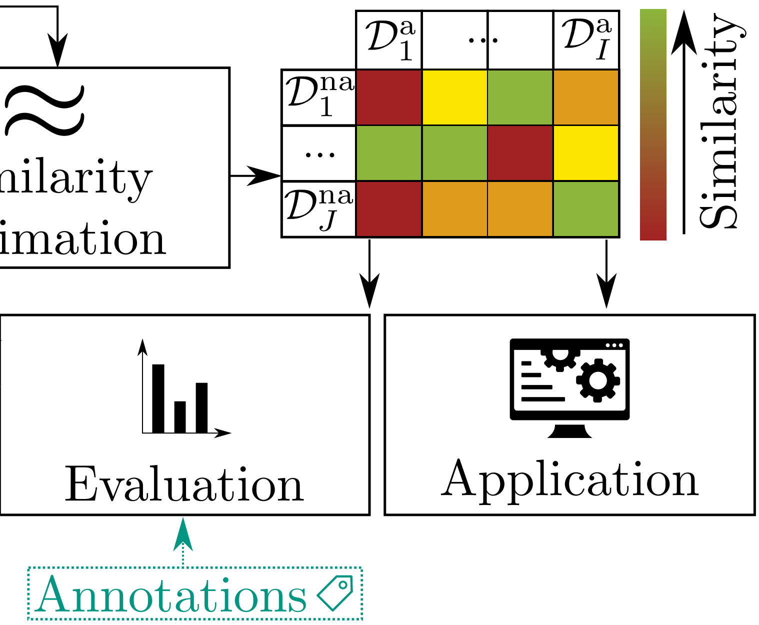



Marcel Schilling


Miguel Molina-Moreno

why we're working on these data-related problems and can see developments toward a more ***data-centric AI****. We believe* ***focusing more on the data and less on the architectures is important.****"*

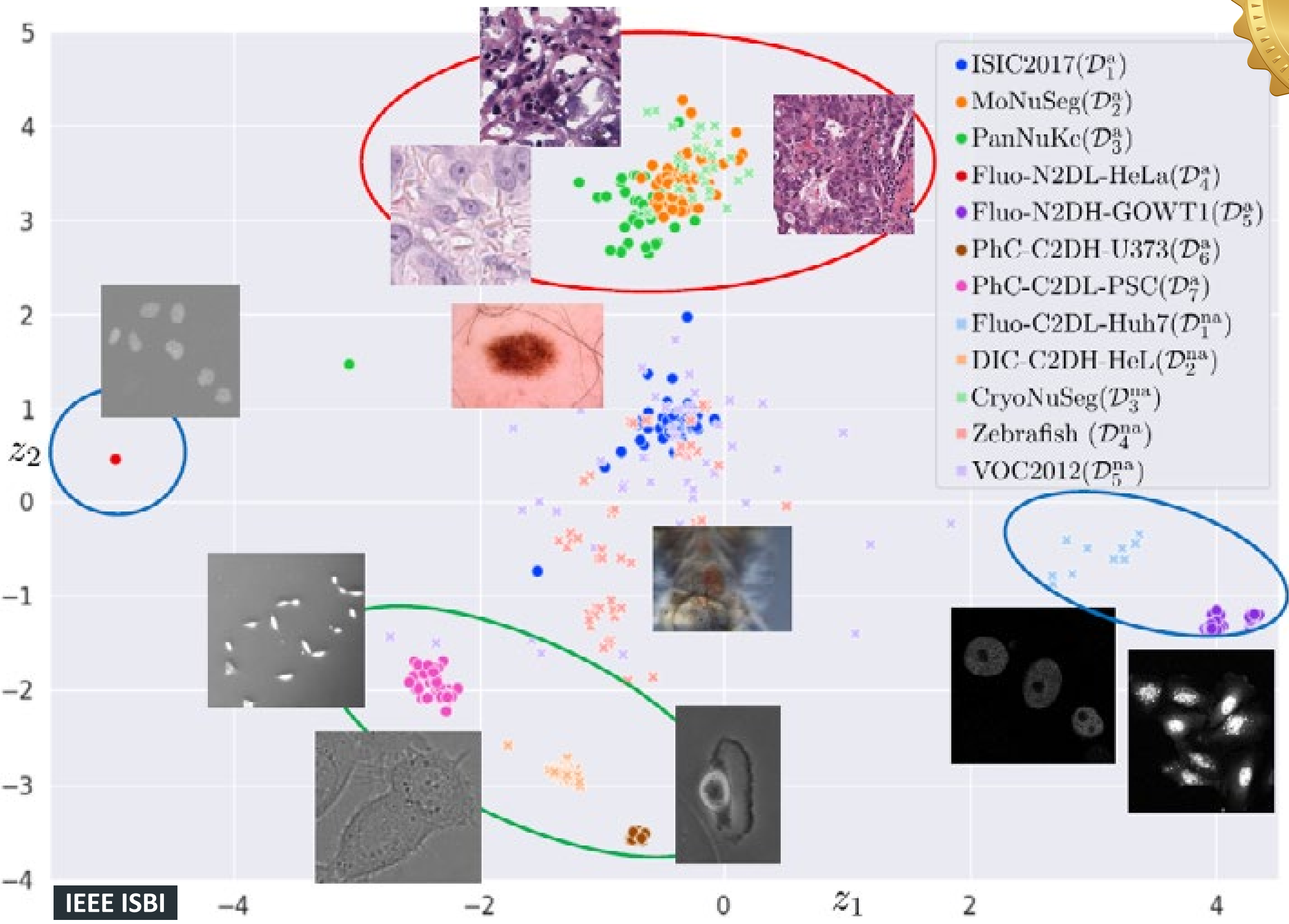
The idea behind the approach is to **automate the selection of suitable databases for a given problem by identifying the most similar database that is already annotated**. To achieve this, the team has developed a similarity estimator consisting of an autoencoder with a specific loss function and arrangement.



Miguel believes their method is more objective than relying solely on human assessment. It considers various factors essential for **determining dataset similarity** – not only the modality of the data but also the style.

"The style means the aspect, color, type of annotation and types of organs or cells represented in the image, and it considers all these things as a whole to match the datasets," he explains. *"Humans can see by sight if a database is similar, but they don't know if it's correlated with the pre-training results. We've demonstrated that* ***our similarity estimator corresponds to the pre-training results*** *in our semantic segmentation task."*

The new method is designed to work with **any 2D image dataset**, making it applicable to a wide range of medical and non-



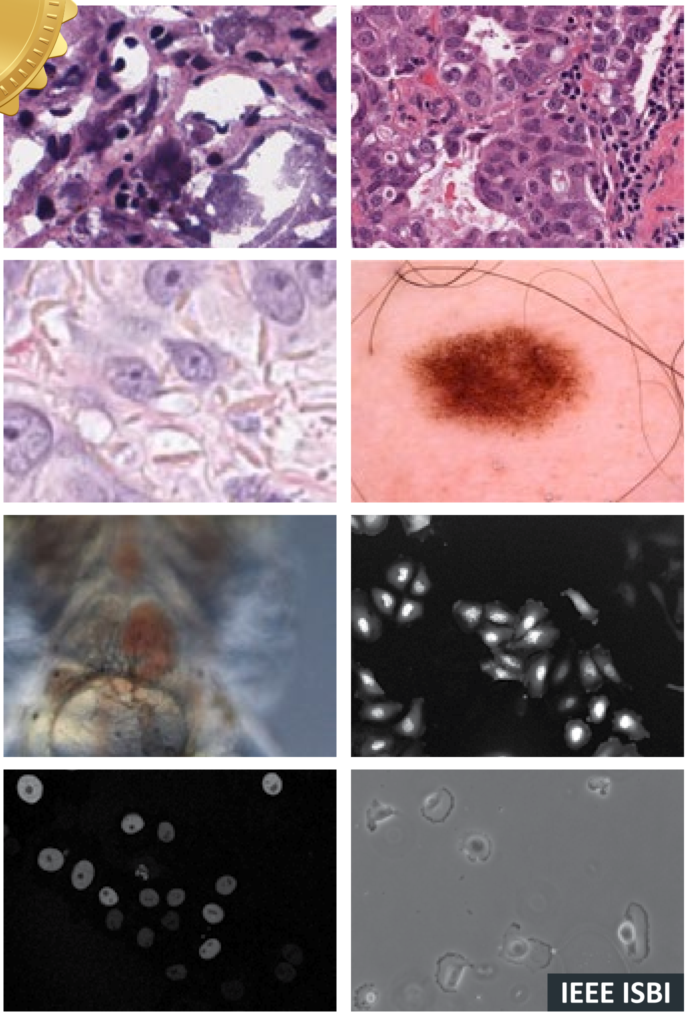
medical problems, and it has been tested on multiple datasets, including **histology**, **dermatology**, and **real-world images**. Marcel and Miguel intend to build a latent space that includes all the standard medical imaging modalities, including ultrasound, CT, and MRI. They will add more datasets and improve the architecture to solve the problem more robustly.

"*We'll continue working on this topic, and when this task of selecting the best available database is solved, it will create huge potential in the application of deep learning,*" Marcel says. "*Once you overcome the issue of finding annotated datasets, there will be plenty of future work in this direction.*"

Miguel agrees: "*Yes, particularly in the biomedical image domain, where you have small datasets that require expert annotation, an automated measurement like ours can open the doors for people with few resources to use other datasets and build up their projects.*"

Away from this paper, Marcel is working at **KIT** on an interdisciplinary project about **personalized oncology**. Miguel has worked on tasks such as license plate and gas leakage detection, but his research focuses on medical and biomedical problems. In that regard, his PhD thesis is devoted to **behavioral cell analysis**.

Taking home the **Best Paper Award at ISBI 2023** last month is still sinking in for both, so is it too soon to ask what factors they think influenced the judges' decision to



choose their paper as the top contender among all submissions?

"That's a tough question to answer," Marcel responds. *"I'm struggling because so much great work was presented at ISBI this year."*

Miguel, who has been thinking carefully, reflects: *"I believe the main advantage of our work is that it is transversal. You could use our method with every imaging modality at ISBI.* ***We're trying to solve a real problem that people in the medical and biomedical field face daily.****"*

As joint first authors of the work, they count their different perspectives, ongoing discussions, and spirit of collaboration and cooperation as key to their success.

"Miguel and I discussed a lot during his research stay here in Germany," Marcel reveals. *"He helped from a technical perspective because he had already worked with* ***contrastive learning****, which was very important for our embedding, and knew about* ***style awareness****. We were a match on the technical and personal side, and it was great attending the conference and presenting together too."*

Miguel agrees and has a tip for other researchers: *"****Be as collaborative as possible!*** *Marcel and I are friends and get along very well. He knows a lot more about the biomedical image domain than me, and he was the one that formulated the problem in the first place. He told me about it, and we found a solution together. We're now developing the next stage of the project!"*

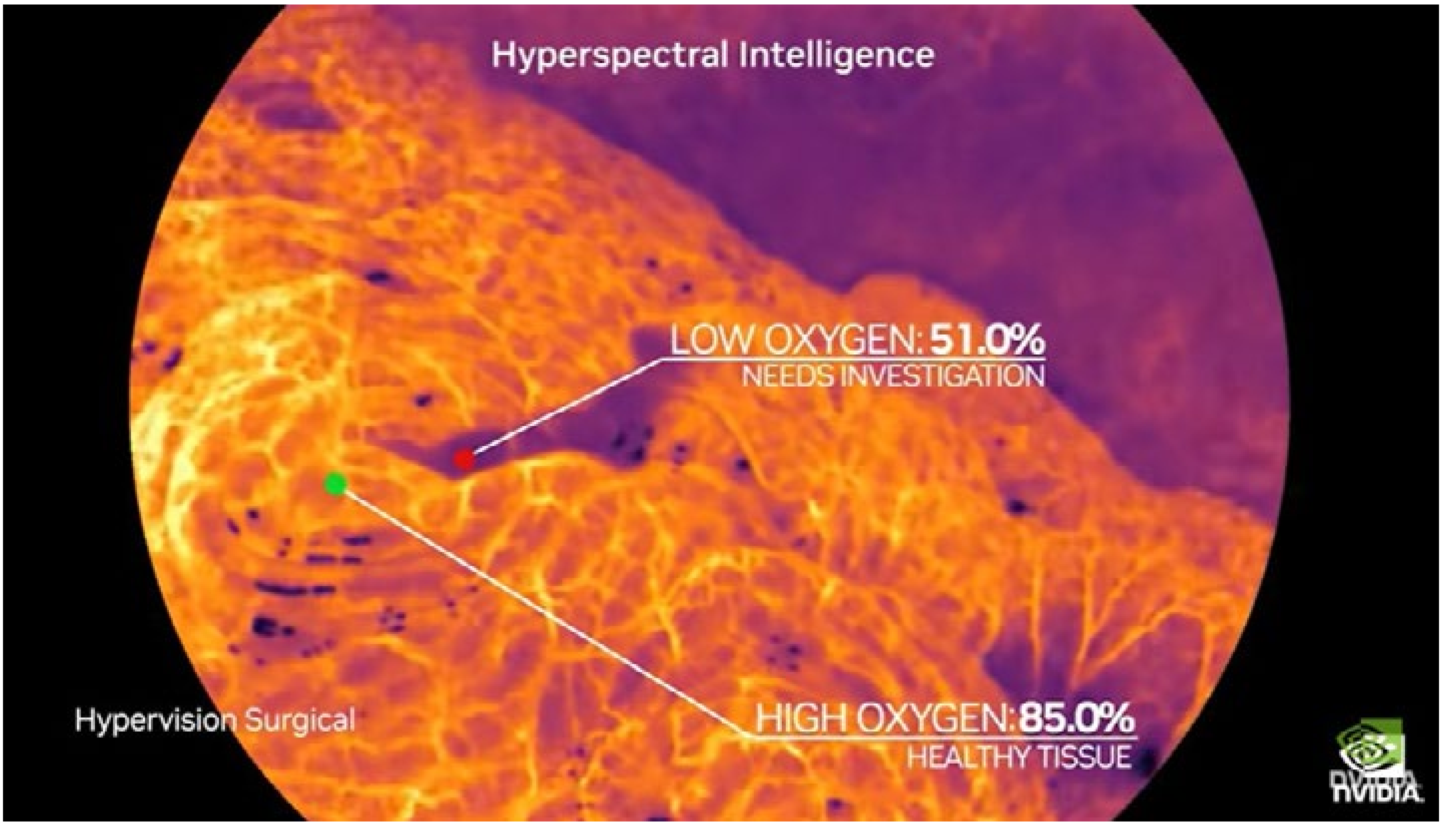
HYPERVISION SURGICAL

Michael Ebner and Tom Vercauteren are co-founders of King's College London spin-out Hypervision Surgical. Michael is the CEO, and Tom is the Chief Scientific Officer and a Professor of Interventional Image Computing at KCL. They speak to us about its emerging technology which offers a promising solution to the limitations of human vision in surgery.

Precision and accuracy are essential factors in ensuring successful surgical outcomes for patients. Every day, surgeons rely on their naked eye to make critical decisions, such as identifying tumor margins, detecting nerves, and assessing tissue perfusion. However, what if they could go beyond human vision and gain access to objective analytical information that could help them make better decisions?

The need for **better imaging technology** comes as surgical complication rates range from 5-20% and are anything from minor to severe. The impact on patients can be devastating.

"Anastomotic leakage is known to occur in around 15% of colorectal surgery patients in the US alone," Michael tells us. *"That's 15% of 600,000 people and costs around $30bn each year. Up to a third of those patients may not survive. That's just one indication,*



but it holds across the board. Whether in colorectal surgery, urology, or neurosurgery, there's a need to help surgeons see better and understand better."

Hypervision Surgical aims to augment the vision of surgeons and provide them with tools that are easy to use and offer real-time guidance. Its technology is based on **hyperspectral imaging**, a technique long used in biomedical imaging, but its surgical use has been restricted due to acquisition and reconstruction complications.

"Recently, changes on the ***machine vision*** *side have led to new sensors being able to* ***capture low-resolution hyperspectral data in real-time****,"* Tom reveals. *"We took advantage of those developments, and together with our computational expertise, we could bridge the gap to deliver* ***high-resolution real-time hyperspectral imaging****. A technology that fits the surgical need!"*

AI is at the heart of everything Hypervision Surgical does. The low-resolution data it captures has many inherent artifacts, but its AI tools reconstruct pristine images for surgeons. It also employs AI to analyze the captured data and extract relevant information displayed to the surgeon to enhance their **intraoperative decision-making abilities**.

"Instead of dividing the entire spectrum of light into three bands – the standard red, green, and blue – we split it into multiple spectral bands," Tom explains. *"That's not something we can visualize as humans, as our eyes are limited to three color bands. We capture 16 and use AI to* ***extract imaging biomarkers of interest in real-time*** *and feed that back to the surgeon in a visually plausible manner."*

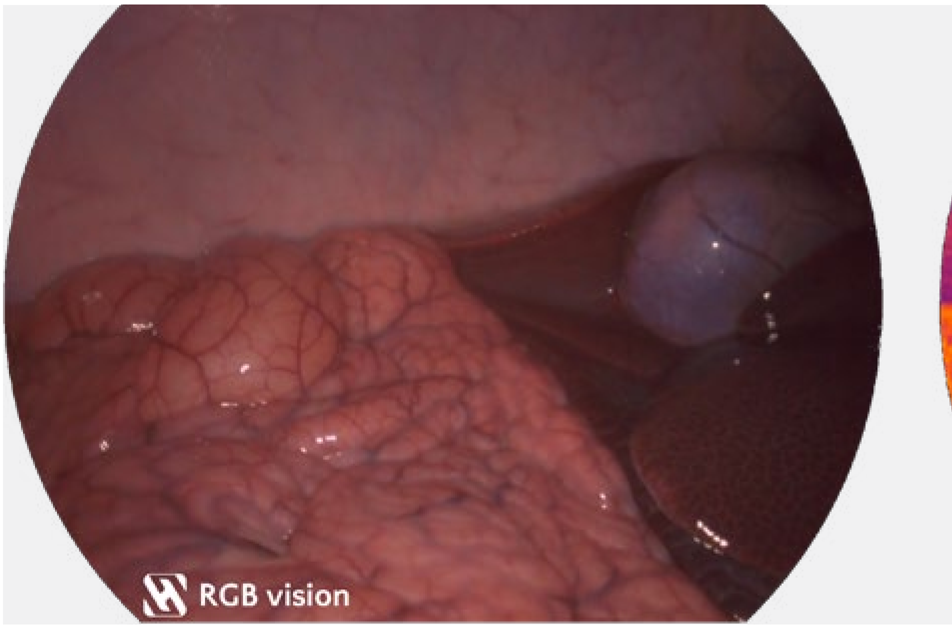
It takes roughly 20,000 hours to train as a surgeon. However much you believe in your technology, convincing surgeons to change their regular practice can be a difficult task,

but luckily, Michael and Tom had factored this in from the outset.

"The link to surgeons is core to what we do," Michael reports. *"We've taken on surgeon input from the very start. That's why we've designed a* ***seemingly conventional laparoscopic system****. It has the same form factor, usability, and color vision, but under the hood, we capture so much more hyperspectral data in a sparse manner. It's all about the* ***simplicity of integration****. To the user, nothing changes in how they use the camera or perform surgery, but at the flick of a button, they can switch from the RGB vision that a surgeon already sees and is comfortable with to an overlay they don't see yet but need to."*

Hypervision Surgical faced a not-unexpected challenge that comes with developing new technology with a new sensor: the **scarcity of data**. It had to devise self-supervised learning approaches to make the most out of the available data and set up controlled environments to capture more information to validate its algorithms.

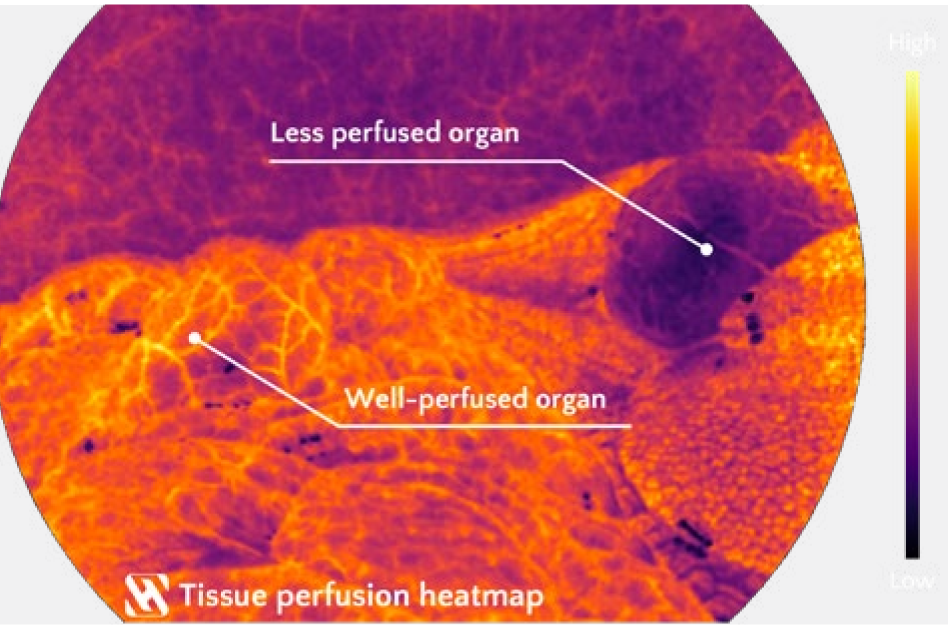
Another risk the company was keen to mitigate was the **real-time AI hallucinating features**, driving the surgery in an unintended direction. The regulatory



burden of bringing a medical device to surgery becomes advantageous here because it sets robust quality control protocols and safeguards that must be in place to ensure the models are appropriate for their end use.

The company is preparing to close its seed round, and in the coming months, it will work towards achieving ISO certification for medical clearance, UK CE marking, and FDA clearance in quick steps, as well as running clinical studies to test its technology and gather valuable data.

"It's a very exciting period for us," Michael smiles. *"It means we'll have all the dedicated resources needed in an early-stage startup. As well as working more on the technology, we'll be able to get through the regulatory requirements, which will allow us to bring this technology closer to patients in a safe but fast-paced manner."*



If you want to join a dynamic and innovative startup, you will be pleased to hear that Hypervision Surgical is hiring. It is actively seeking to broaden the talents and diversity of its team, so do not miss your chance to be a part of its success story. Who knows, in a few years, you might be the one sharing your journey with us. Take a chance and apply now!

Congrats, Doctor Maria Chiara (cum laude)!

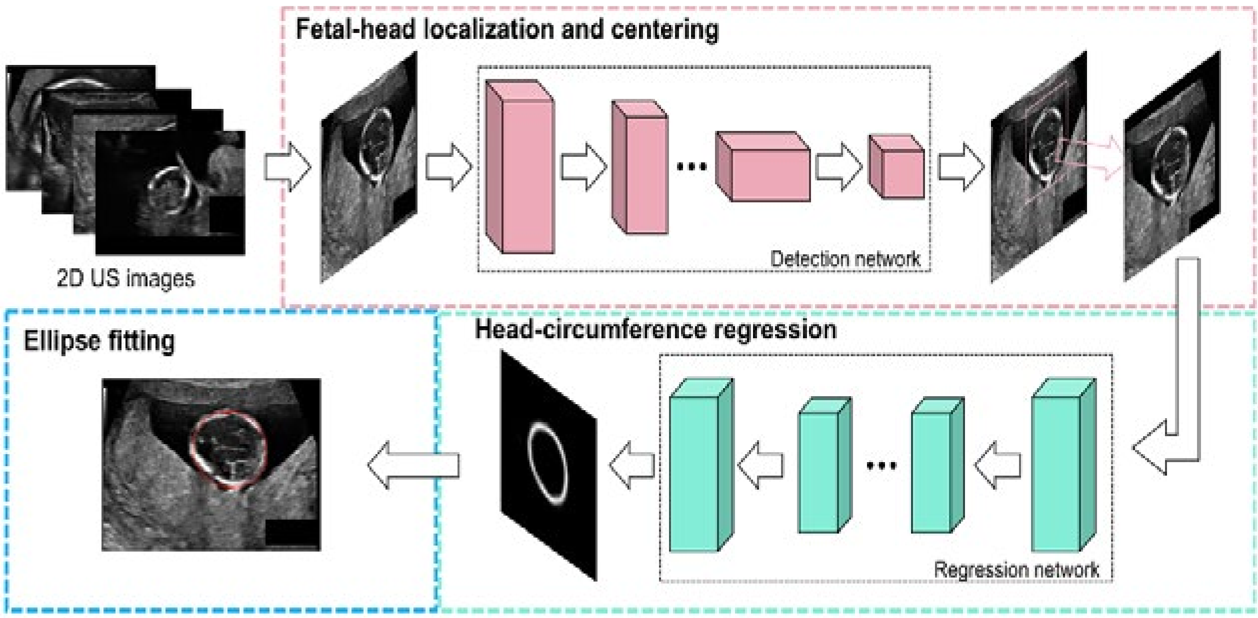
Maria Chiara Fiorentino recently finished her PhD with the Vision Robotic and Artificial Intelligence (VRAI) group at Università Politecnica delle Marche. The aim of her thesis is to exploit Deep Learning (DL) algorithm for Ultrasound (US) image analysis. Congrats, Doctor Maria Chiara!

Due to its real-time nature, **Ultrasound** imaging stands out among many imaging diagnostic options as a **highly cost-effective modality** that provides the clinician with an unrivaled and invaluable level of interaction along with other advantages (e.g. absence of ionizing radiation, portability, accessibility and cost effectiveness). Despite the well-recognized clinical utility especially for soft-tissue examination, Ultrasound presents **unique challenges**, such as presence of **speckle and acoustic noise**, which hinder image interpretation, **high dependence on diagnostic operator experience** and **high inter- and intra-observer variability** across different institutes and manufacturers' Ultrasound systems.

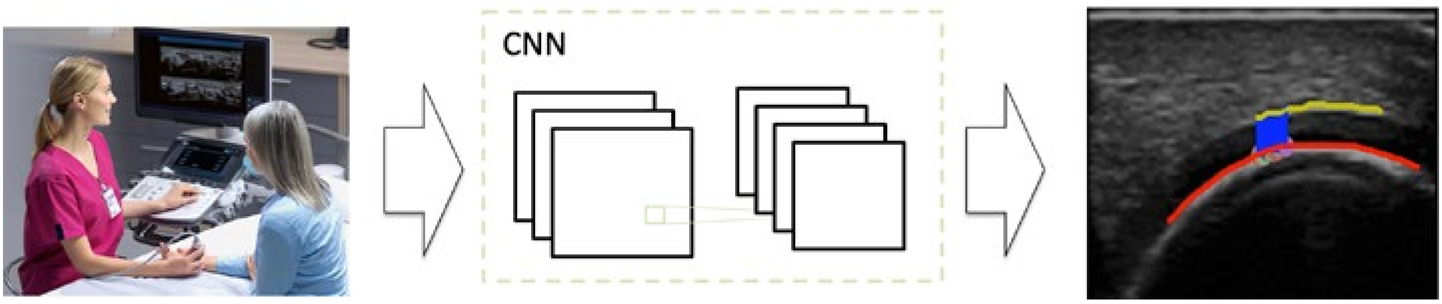
Deep Learning has undergone an increasing role in Ultrasound image analysis to offer decision support to clinicians. However, despite its huge potential in the Ultrasound field, there is still room for improvement.

My thesis aims at moving the Deep Learning state of the art forward for already investigated Ultrasound fields, such as **gynecology**, a well-known and studied domain of applications; but it also aims at investigating the potentiality of Deep Learning approaches for fields, such as **rheumatology**, in which it has not been fully exploited.

Regarding Ultrasound in gynecology, I have proposed a novel distance field network approach to automatically measure the fetal head circumference from fetal 2D Ultrasound images. **Automated measurement of fetal head circumference** is a crucial clinical task to assess fetus growth and its assessment is strongly dependent from clinicians expertise posing issues of both intra- and inter-operator variability and time consuming. Experimental results have demonstrated that distance field produces precise delineations of Head Circumference (HC) boundaries invariably from shapes, heterogeneous appearances and gestational age of fetus head (Figure 1).



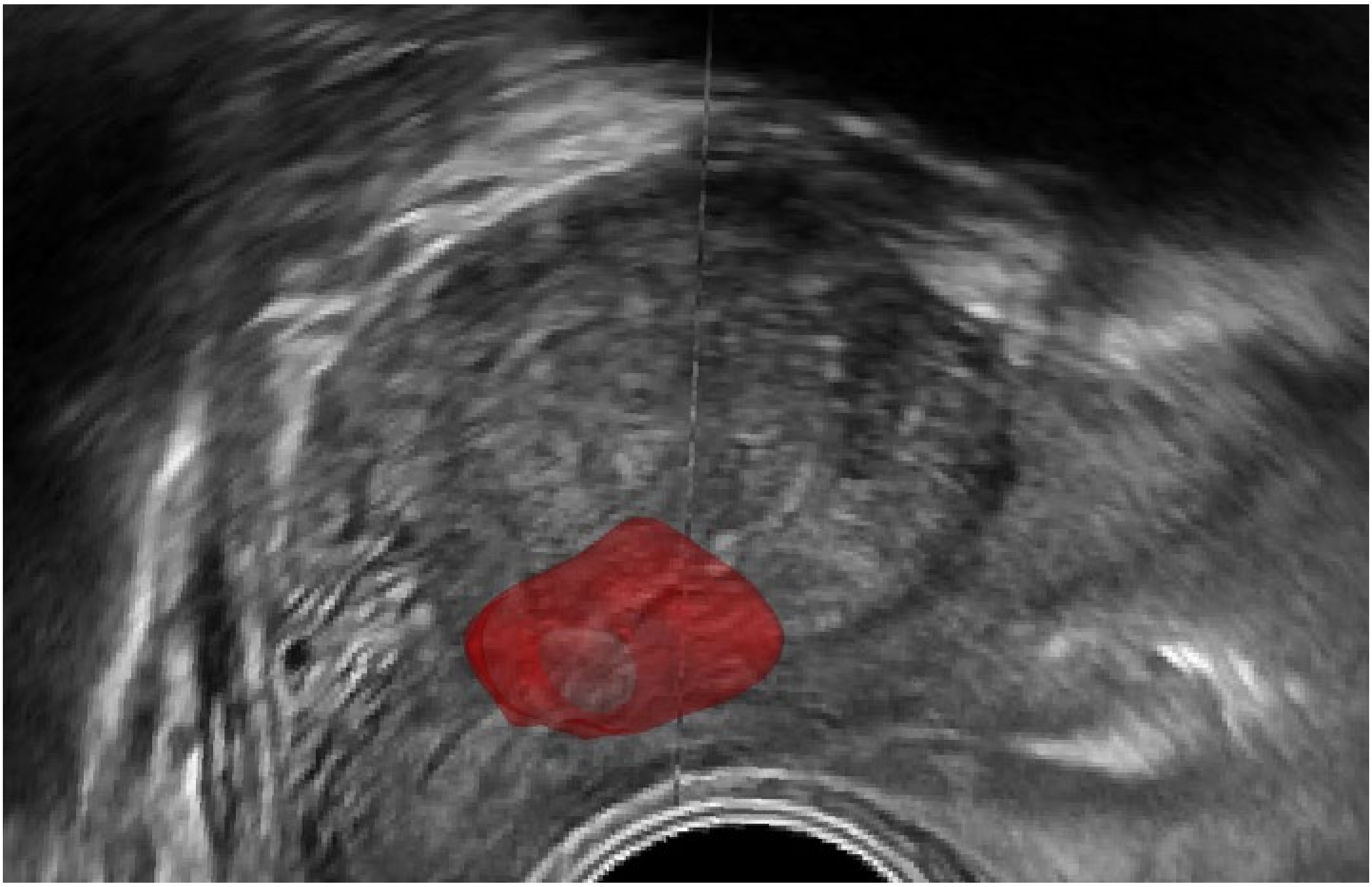
Concerning rheumatology domain, Deep Learning in **musculoskeletal Ultrasound** has been receiving significant attention and it is expected to be pivotal in the development of next-generation cutting-edge Ultrasound imaging systems. In this regard, I have developed a framework able to automatically estimate cartilage-thickness from Ultrasound images, paving the way for future research in the field. **Cartilage damage** is one of the most relevant determinants of physical disability in patients who suffer from Rheumatoid Arthritis (RA), a chronic disease characterized by erosive symmetrical polyarthritis. From a clinical perspective, the proposed framework proved to be a valuable tool to support RA clinical examination (Figure 2).



My thesis had also **technology transfer implications**. In fact, I had the opportunity to join **Esaote R&D**, one of the world's leading producers of medical diagnostic systems, and **Vicomtech** eHealth and Medical Devices department, a technological center set up as a private non-profit Foundation. Within these periods, I had the opportunity to focus on implementing and validate AI-based solutions for Ultrasound and MRI.

INTRA-OP PROSTATE GUIDANCE

Prostate guidance with accurate lesion localization and prediction of tissue distortion.



Approximately one million prostate biopsies are performed each year in the United States to **diagnose and monitor prostate cancer**. Doctors commonly use transrectal ultrasound (TRUS) to generate **real-time images of the prostate to help guide the biopsy**. To mitigate the risk of missing a lesion, physicians usually carry out multiple collections from the gland. Nonetheless, there is a risk that some tumors may go undetected. Therefore, it is common practice for the patient to have a **pre-op MRI scan**, which generates a more detailed image of the prostate.

This provides the physician with a more accurate depiction of the organ and lesions.

The pre-op MRI scan must be **registered with the live ultrasound image**, to enable the physician to accurately locate lesions in real time. Registration of non-rigid tissue is challenging, as **tissue moves during the procedure**, for example when the patient breathes or when the US probe is inserted and applies pressure on the surrounding area. In addition, pre-op MRI images may be acquired some time before the procedure, such that the anatomical representation captured in the pre-op scan has altered before the ultrasound is performed. Tissues may deform, distorting the live image compared to the pre-op image, **making it difficult to superimpose them accurately**.

Physicians often use **naive registration** to merge live ultrasound images to pre-op MRI scans. In this method, tissue motion is assumed to be rigid, which allows for simpler mathematics and algorithms. However, the assumption of rigid motion is coarse and may lead to mistakes in positioning tumors correctly. Furthermore, naive registration relies heavily on the surgeon's knowledge and expertise to compensate for any differences between the live ultrasound and pre-op MRI images that may occur as a result of tissue movement. **A limited number of physicians are qualified to perform this technique, which is time-consuming and susceptible to human subjectivity**.

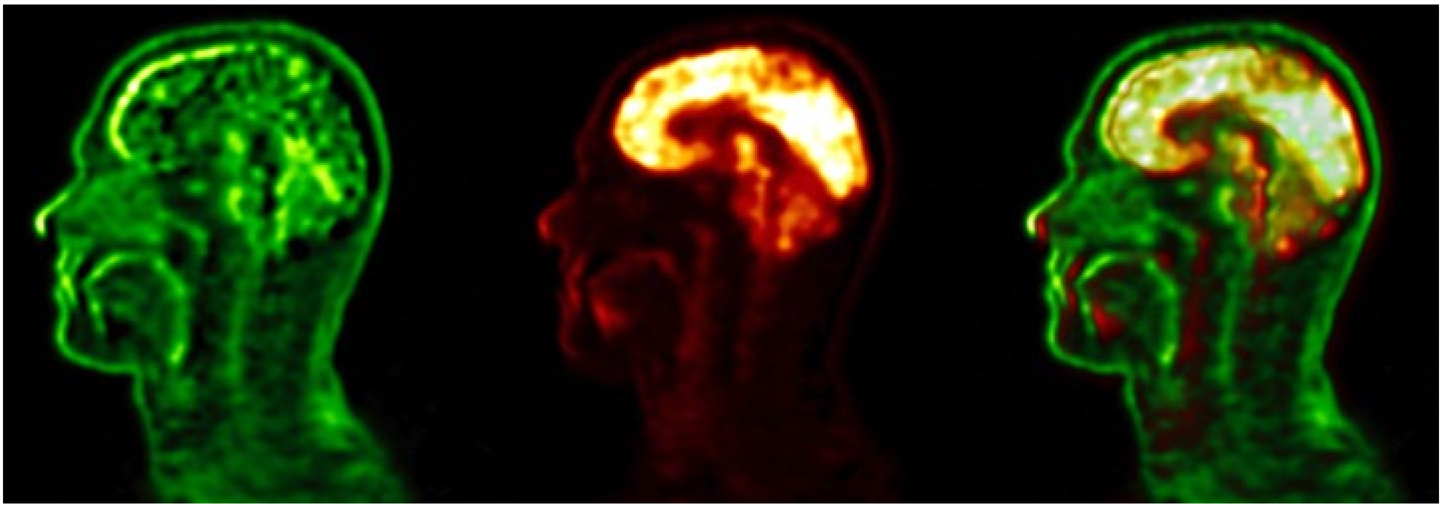
RSIP Vision is developing sophisticated technology to perform **non-rigid registration to account for soft tissue motion during procedures. AI algorithms** are trained to predict tissue distortion, allowing for **live ultrasound images to be accurately merged with pre-op MRI scans**. This provides the surgeon with a detailed MRI image that is overlaid on a real-time ultrasound, allowing for quick and accurate detection of tumors.

In addition, RSIP Vision uses **probe tracking** in a wide variety of **navigation systems**. These systems provide position and orientation data using external hardware. This provides better inputs for the algorithms generating more detailed information for physicians. Tracking devices allow for precise localization of every pixel in the ultrasound image, leading to **more accurate lesion localization**.

At RSIP Vision we are utilizing **AI and computer vision** in combination with advanced tracking hardware to **improve the speed and accuracy of imaging-based navigation systems**. These methods will enhance detection and evaluation of tumors to ultimately improve the diagnosis and treatment of prostate cancer. Contact RSIP Vision today to learn how you can implement these techniques for improved clinical results.

This article was first published on RSIP Vision's website

MACHINE LEARNING IN CLINICAL NEUROIMAGING (MLCN)



Thomas Wolfers is the Group Leader of the Laboratory for Mental Health Mapping at the University of Tübingen in Germany. Nicha Dvornek is an Assistant Professor at Yale University in the Department of Radiology & Biomedical Imaging and the Department of Biomedical Engineering. Vinod Kumar is a Postdoc at the Max Planck Institute for Biological Cybernetics Tübingen. They speak to us today as co-organizers of a fascinating workshop later this year at MICCAI 2023 in Vancouver.

The Machine Learning in Clinical Neuroimaging (MLCN) workshop was established in 2018 and remains as relevant as ever. Over the last five editions, it has seen a range of exciting discoveries and engaging talks. One standout keynote was given by a friend of this magazine, **Jorge Cardoso**, in 2020, when MONAI was in active development. The workshop aims to bridge the gap between machine learning and clinically applied research. With that aim in mind, it is divided into two tracks.

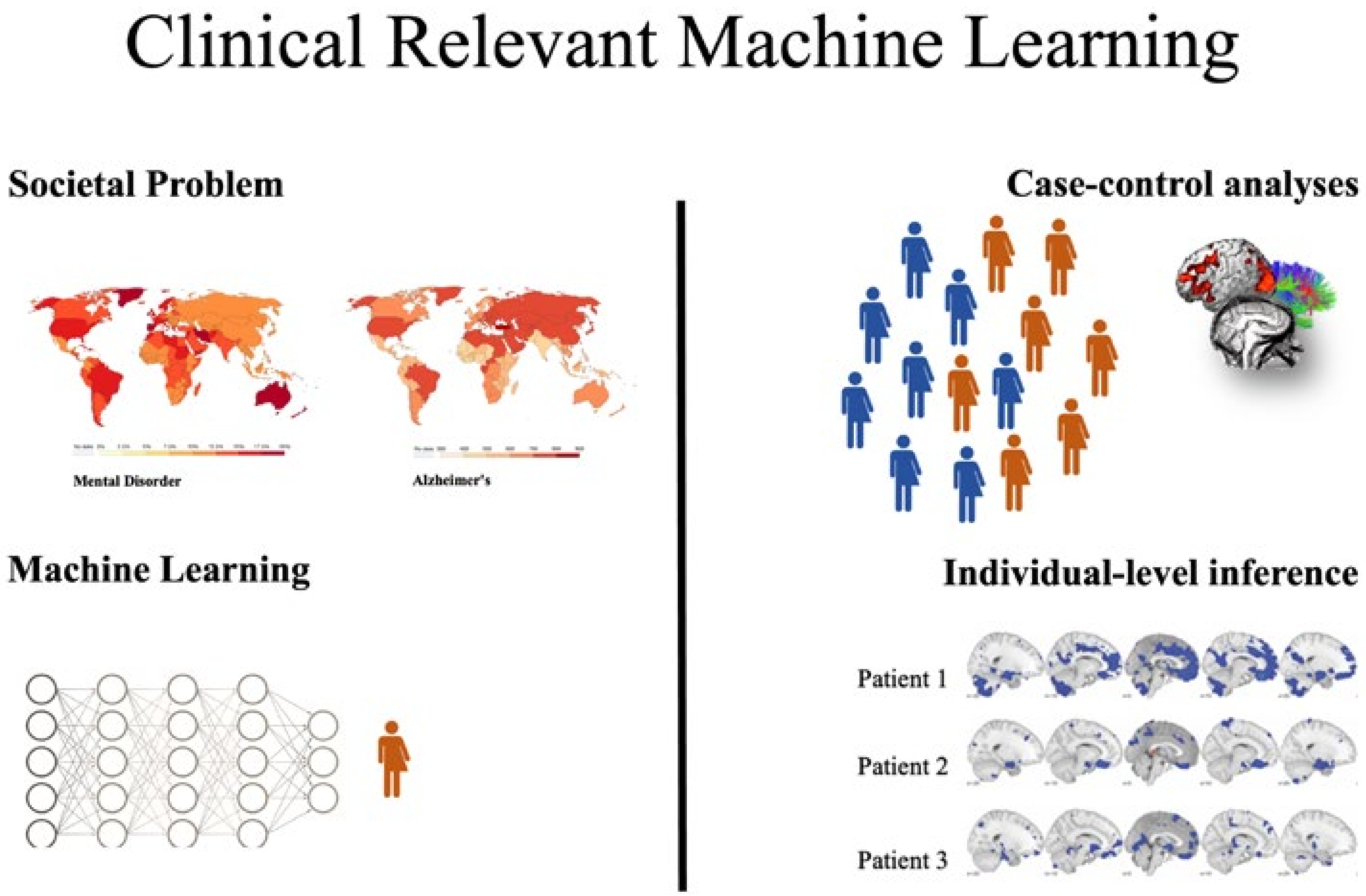
"The first track is about the ***development of novel machine learning technology****, with an eye on clinical utility in the context of* ***brain imaging and neuroscientific research****,"* Thomas tells us. *"The second track is more applied. We want to find applications that have, for instance, found* ***an interesting parcellation of the brain, sub-stratification of a disease, or made some clinically relevant conceptual changes****."*

We are yet to understand our brains fully, and there is still much to learn and discover in **neuroimaging**. Vinod tells us he has been investigating the brain using a combination of high-resolution MRI and machine learning. He is fascinated by how it works. *"There are many clinical applications that are in dire need of new methods to understand the brain, different*

diseases, and biomarkers for neurological and psychiatric disorders," he points out. *"Also, new methods to segment the brain and to identify where there are lesions or increased or reduced sickness in a certain brain area."*

MICCAI participants are often spoilt for choice regarding workshops, but the organizers are confident that MLCN will continue to stand out from the crowd.

"We don't just focus on the methodological development, like the machine learning side, but also the translation side," Nicha emphasizes. *"It's a nice forum to bring together the machine learning scientists and the neuroscience folks. How can we*



The scope of the MLCN workshop, including the societal and methodological problems it addresses, the approaches that are relevant, and the ambition to find individual-level decision points for complex brain disorders.

get these new techniques adopted by clinicians to improve both analysis and clinical workflow?"

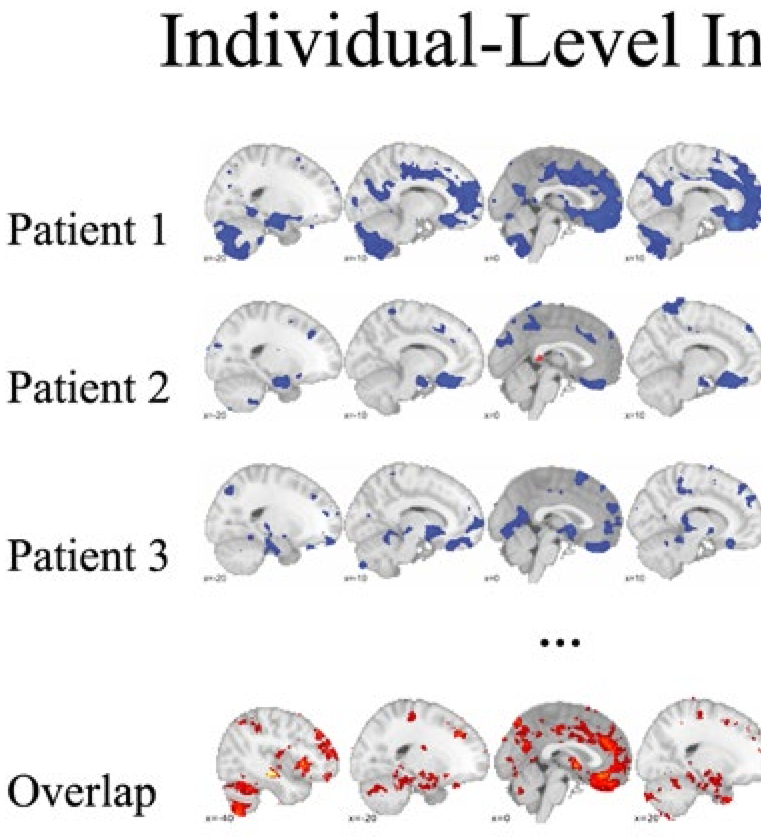
The team hopes the workshop will ultimately lead to **robust machine learning algorithms** that can handle messy clinical practice data with the same ease as large-scale off-the-shelf or consortia datasets with well-resourced acquisition protocols. Mental and neurological disorders are often heterogeneous, and identifying individual differences is crucial for finding possible illness markers and stratifying patients with similar symptoms.

"We've seen large datasets that estimated a norm across a population and then placed individuals across a range of biomarkers in reference to this population," Thomas explains. *"This eventually allows us to find possible markers of illness and stratify differences between patients that supposedly have the same illness. In the case of schizophrenia or other neurological diseases,* ***symptomatic representation doesn't map well onto the brain on the group level****."*

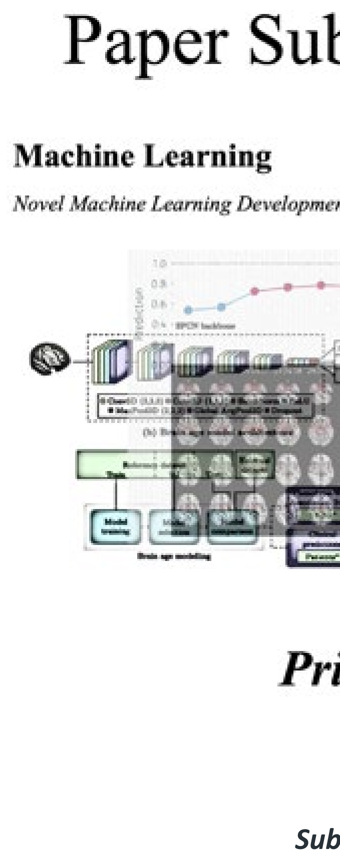
Nicha adds: *"It sounds abstract sometimes talking about finding these different biomarkers to help stratify patients, but that's important to impact the treatment. In some of my work, we work with children with autism, and no one treatment works for each child. It's trial and error. If we could somehow, perhaps through baseline imaging, predict that this specific treatment will be effective for this specific child, that kind of personalized or* ***precision medicine approach*** *would be the dream."*

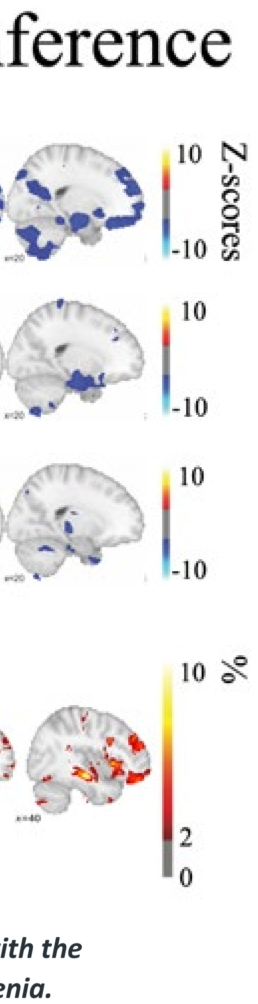
Thomas, Vinod, and Nicha each have personal reasons for focusing on the brain at this stage in their career. All agree that it is one of the most fascinating organs of the human body, and largely still a mystery, even though it realizes many of the functions that make us human.

"We operate our consciousness and experience emotions through the brain," Vinod muses. *"It's why we're enjoying this beautiful conversation. You're curious, and we're also curious to talk*



The degree of variability among individuals w
same brain disorder, in this case, schizophre

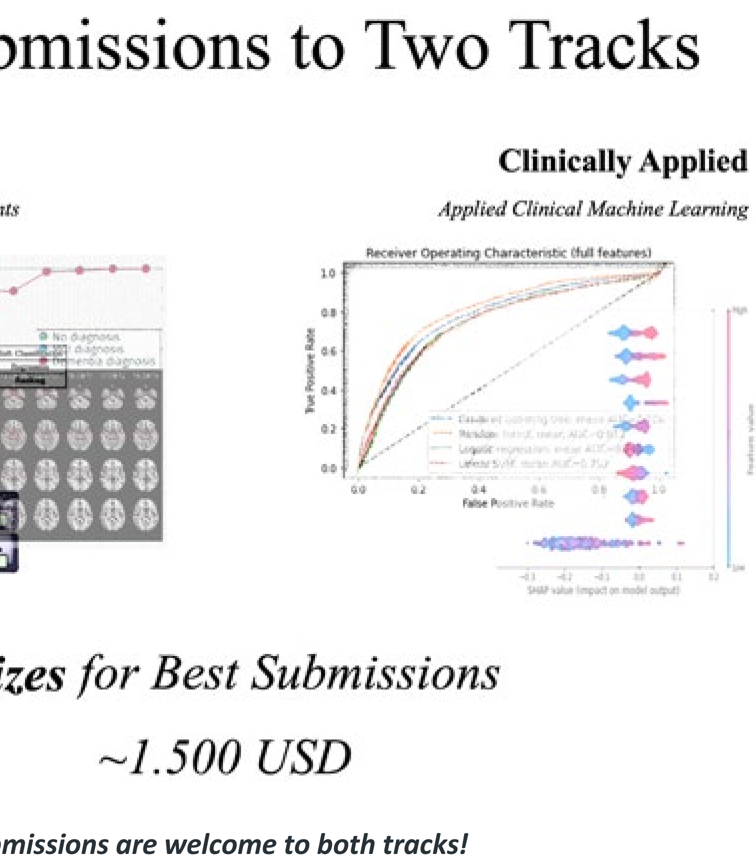




about it. We're experiencing all this through the brain. ***That's magical!***"

Thomas adds: "*Many of the illnesses supposedly realized in the brain are ill-understood. There's a huge potential for research in that area. Some of the most devastating conditions, like mental disorders and neurological diseases, are linked to brain states.*"

For Nicha, landing here was more of a happy accident: "*When I started my PhD, I wanted to go into* ***cardiac image analysis****. I was really interested in the heart and circulation. Then I fell into a brain imaging project, and it was one of my best accidental turns!*"

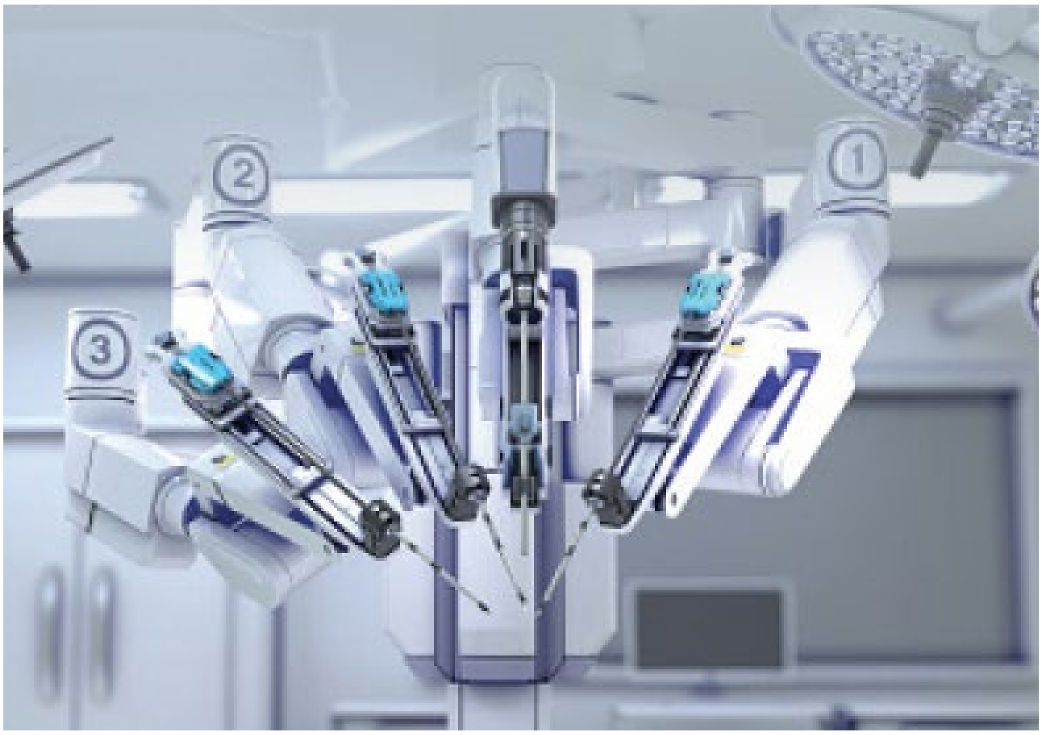


We ask the group what they think is the most significant contribution this workshop could make to the field. However, Vinod emphasizes the importance of incremental impact, building on current knowledge, and **impacting clinical studies**, specifically regarding certification and awareness of variance in the population.

Thomas has a more specific hope: "*It would be great if we could find a replicable, clearly discernible, treatment-relevant subcategory for schizophrenia that is distinct from most other individuals with schizophrenia and can be shown across at least three different samples.*"

Nicha agrees with Thomas that **reproducibility is a big problem**, and she points to the importance of the panel discussion at the end of the workshop, which is often a space to identify the most critical questions and prioritize what to work towards. "*Many times, we get stuck on: How can I apply this method to this problem? But really, are we solving the right problems?*" she ponders rhetorically.

Vinod highlights last year's helpful conversation on the impact of demographics, race, and gender on machine learning and the need to understand more about generalizability beyond a particular population. "*If you train a model in America, it may not be translatable in China or India, for example,*" he explains. "*It's not generalized enough, and we must know the limitations. This subject opens new dimensions for research. There are more things to learn.*"



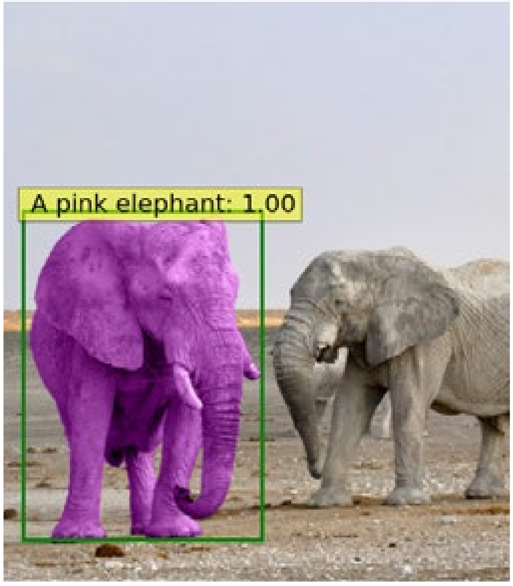
IMPROVE YOUR VISION WITH

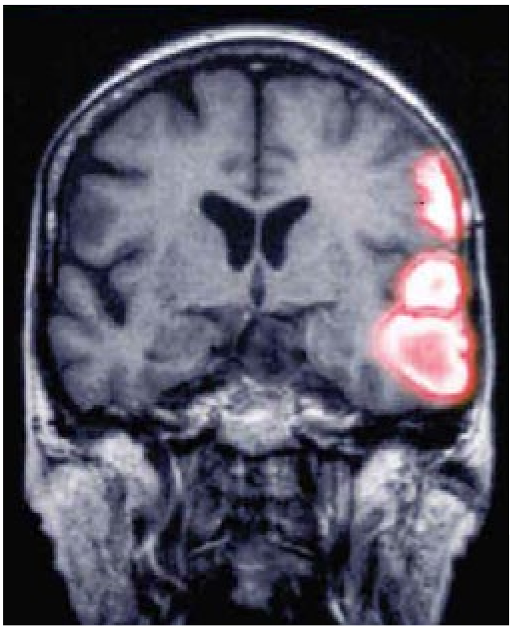
Computer Vision News

SUBSCRIBE

to the magazine of the algorithm community

and get also the new supplement Medical Imaging News!





RSIP VISION