

JANUARY 2023

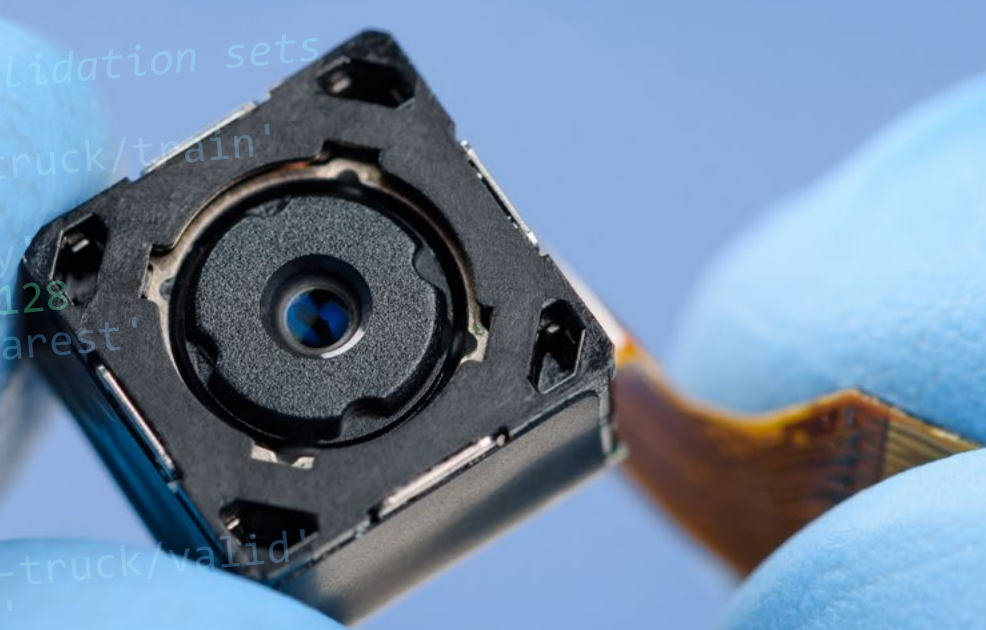
Computer Vision News & Medical Imaging News

The Magazine of the Algorithm Community

```
'PYTHONHASHSEED' = '1'  
'TF_DETERMINISTIC_OPS' = '1'
```

```
Matplotlib defaults  
'figure' = True  
'axes' = 'bold' = 'large'  
'image' = 'bold' = 18 = 10  
= 'magma' = 18 = 10  
"ignore" # to clean up output cells
```

```
and training and validation sets  
=   
'../input/car-or-truck/train'  
= 'inferred'  
= 'binary'  
= 128 128  
= 'nearest'  
= 64  
= True  
=   
'../input/car-or-truck/valid'  
= 'inferred'  
= 'binary'  
= 128 128  
= 'nearest'  
= 64  
= False
```





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.

Dear reader,

Happy New Year! On behalf of **RSIP Vision**, we wish all our readers and the community a successful 2023. This year, we will continue our pioneering work in Medical Image Analysis and AI and have many surprises in store, so stay tuned!

This month, we have an overview of not one but **TWO award-winning Best Papers**, continuing our tradition of bringing you the best of the best work out there.

We hear about **SMILE lab**, where Ruogu Fang and her team are doing revolutionary work on AI and machine learning for neurodegenerative diseases.

Lucky students of December's **Hamlyn Winter School on Surgical Imaging and Vision** will begin the year fully up to date on the state of the art in surgical robotics. Matina Giannarou fills us in on the week-long event.

Also, Ioannis Valasakis reviews feature extraction layers in ConvNets, a fabulous computer vision tool.

Enjoy all that and so much more in January's issue of Computer Vision News. Remember to tell all your friends and colleagues about us so they can subscribe for free too!

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

A big thank you to Ralph Anzarouth and the team at RSIP Vision for their dedication to serving the community with such an amazing magazine!

Fabio Pizzati,
Postdoc at University of Oxford

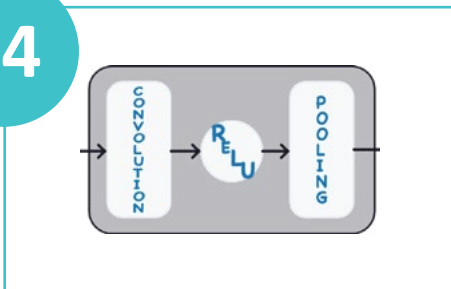
Follow Us



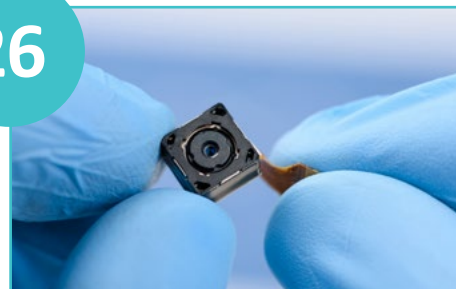
Computer Vision News

Medical Imaging News

04



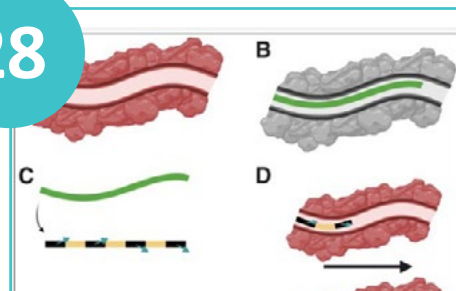
26



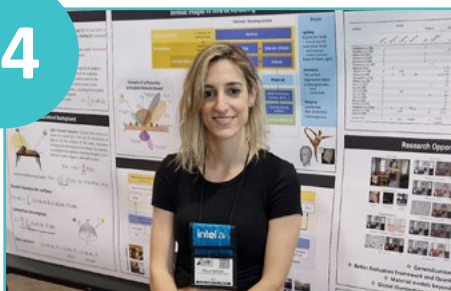
10



28



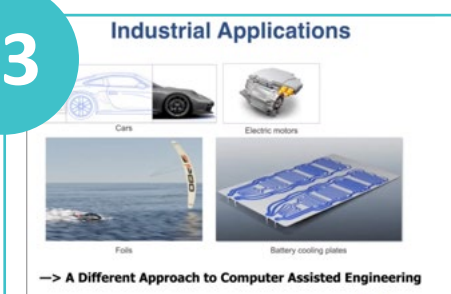
14



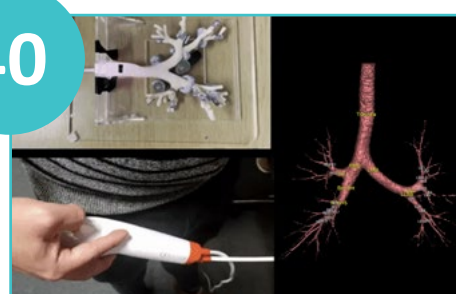
32



23



40



04 Feature Extraction Layers in ConvNets
Computer Vision Tools

26 Soft Tissue Tracking with HR Cameras
Medical Imaging with RSIP Vision

10 Trans2k with Jiří Matas and team
Best Paper BMVC



28 Patient-Specific
Magnetic Catheters for...
Best Paper Leeds University



14 Elena Garcés
Women in Computer Vision

32 The SMILE Lab with Ruogu Fang
Brain/Bio-inspired AI

23 Pascal Fua - EPFL
Keynote Speech at BMVC

40 The Hamlyn School
with Matina Giannarou
Robotic Surgery



IOANNIS VALASAKIS, KING'S COLLEGE LONDON



Happy New Year for all the people who changed one, and Season's Greetings for everyone! This is an exciting month. New beginnings. And one of the most exciting news lately is the ChatGPT. I am sure you read, you saw tweets, everybody got excited, their systems were overloaded.

It can do many things, with impressive accuracy and determination. OK, I am so excited myself, that I will give space for the next paragraph.

Me: Can you tell me a phrase about yourself, describing your technology to a very knowledgeable audience and write an introductory paragraph for my computer vision tutorial article in the new year's issue of RSIP Vision's magazine?

ChatGPT: Certainly! Here is a phrase that describes my technology: "Assistant is a large language model trained by OpenAI, using state-of-the-art machine *learning techniques to generate human-like text.*"

As an introduction to your computer vision tutorial article in the new year's issue of RSVIP magazine, you might consider mentioning the increasing importance and widespread applications of computer vision in various fields, from self-driving cars to medical image analysis. You could also mention the potential for computer vision to transform industries and improve quality of life and introduce the concept of using machine learning techniques to teach computers to interpret and analyze visual data. You might also provide some context on the current state of the field and mention any exciting recent developments or advances.

Wow! So, with this introduction, I'll start my tutorial which is a computer vision analysis using computer vision models with convnets. Let's dive in!

Introduction

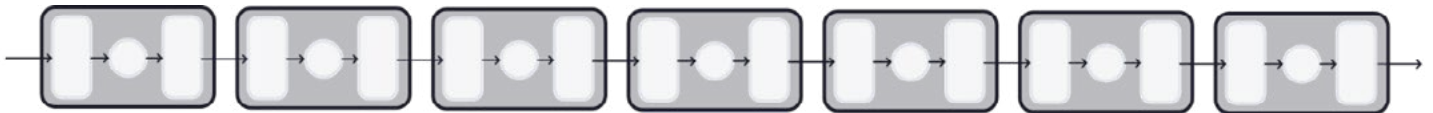
The idea is that, as we saw on the last time, we use different layers of a convnet to extract features, this time we'll use those feature extraction layers and put them together to build a network of our own 😊

We have also seen in my previous tutorials, that convolutional networks perform **feature extraction** through three operations: **filter**, **detect**, and **condense**. A single round of feature extraction can only extract relatively simple features from an image, things like simple lines or contrasts. These are too simple to solve most classification problems. Instead, convnets will repeat this extraction over and over, so that the features become more complex and refined as they travel deeper into the network.

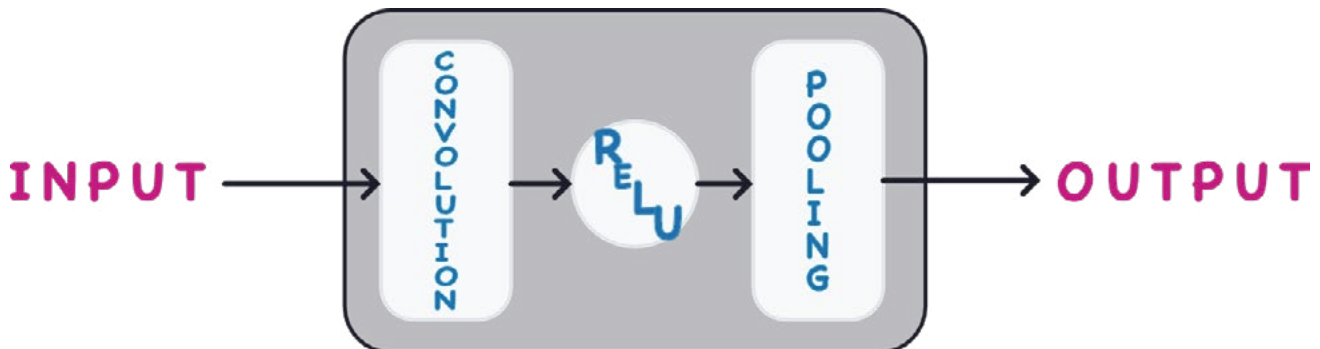


Convolutional Blocks

It does this by passing them through long chains of **convolutional blocks** which perform this extraction.



These convolutional blocks are stacks of Conv2D and MaxPool2D layers, whose role in feature extraction we learned about in the last few lessons.



Each block represents a round of extraction, and by composing these blocks the convnet can combine and recombine the features produced, growing them and shaping them to better fit the problem at hand. The deep structure of modern convnets is what allows this sophisticated feature engineering and has been largely responsible for their superior performance.

Example - Design a Convnet

Let's see how to define a deep convolutional network capable of engineering complex features. In this example, we'll create a Keras Sequence model and then train it on our Cars dataset.

Step 1 - Load Data

This hidden cell loads the data.

```
# Imports
import os, warnings
import matplotlib.pyplot as plt
from matplotlib import gridspec
```

```
import numpy as np
import tensorflow as tf
from tensorflow.keras.preprocessing import image_dataset_from_directory
```

Reproducibility

```
def set_seed(seed=31415):
    np.random.seed(seed)
    tf.random.set_seed(seed)
    os.environ['PYTHONHASHSEED'] = str(seed)
    os.environ['TF_DETERMINISTIC_OPS'] = '1'
set_seed()
```

Set Matplotlib defaults

```
plt.rc('figure', autolayout=True)
plt.rc('axes', labelweight='bold', labelsize='large',
       titleweight='bold', titlesize=18, titlepad=10)
plt.rc('image', cmap='magma')
warnings.filterwarnings("ignore") # to clean up output cells
```

Load training and validation sets

```
ds_train_ = image_dataset_from_directory(
    '../input/car-or-truck/train',
    labels='inferred',
    label_mode='binary',
    image_size=[128, 128],
    interpolation='nearest',
    batch_size=64,
    shuffle=True,
)
ds_valid_ = image_dataset_from_directory(
    '../input/car-or-truck/valid',
    labels='inferred',
    label_mode='binary',
    image_size=[128, 128],
    interpolation='nearest',
    batch_size=64,
    shuffle=False,
)
```

Data Pipeline

```
def convert_to_float(image, label):
    image = tf.image.convert_image_dtype(image, dtype=tf.float32)
    return image, label
```

```
AUTOTUNE = tf.data.experimental.AUTOTUNE
```

```
ds_train = (
    ds_train_
    .map(convert_to_float)
    .cache()
    .prefetch(buffer_size=AUTOTUNE)
)
ds_valid = (
    ds_valid_
    .map(convert_to_float)
    .cache()
)
```

```
.prefetch(buffer_size=AUTOTUNE)
)
```

Found 5117 files belonging to 2 classes.
Found 5051 files belonging to 2 classes.

Step 2 - Define Model

Here is a diagram of the model we'll use:

Now we'll define the model. See how our model consists of three blocks of Conv2D and MaxPool2D layers (the base) followed by a head of Dense layers. We can translate this diagram more or less directly into a Keras Sequential model just by filling in the appropriate parameters.

```
import tensorflow.keras as keras
import tensorflow.keras.layers as layers

model = keras.Sequential([

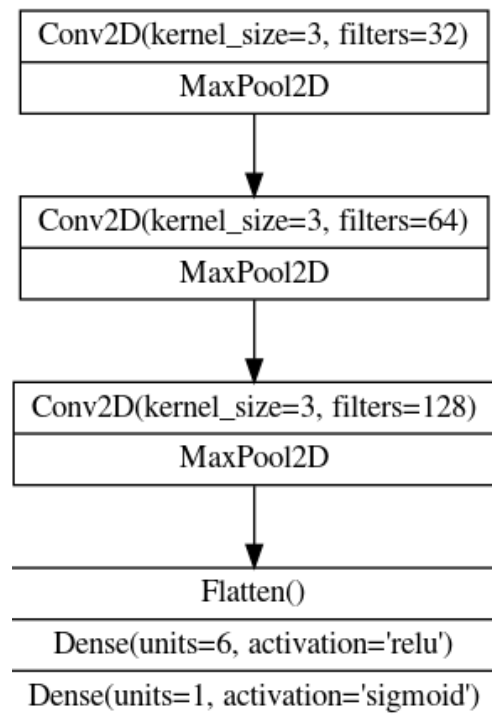
    # First Convolutional Block
    layers.Conv2D(filters=32,
kernel_size=5, activation="relu",
padding='same',
                # give the input
dimensions in the first layer
                # [height, width,
color channels(RGB)]
                input_shape=[128,
128, 3]),
    layers.MaxPool2D(),

    # Second Convolutional Block
    layers.Conv2D(filters=64, kernel_size=3, activation="relu", padding='same'),
    layers.MaxPool2D(),

    # Third Convolutional Block
    layers.Conv2D(filters=128, kernel_size=3, activation="relu", padding='same'),
    layers.MaxPool2D(),

    # Classifier Head
    layers.Flatten(),
    layers.Dense(units=6, activation="relu"),
    layers.Dense(units=1, activation="sigmoid"),
])
model.summary()

Model: "sequential"
```



Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 128, 128, 32)	2432
max_pooling2d (MaxPooling2D)	(None, 64, 64, 32)	0
conv2d_1 (Conv2D)	(None, 64, 64, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 32, 32, 64)	0
conv2d_2 (Conv2D)	(None, 32, 32, 128)	73856
max_pooling2d_2 (MaxPooling2D)	(None, 16, 16, 128)	0
flatten (Flatten)	(None, 32768)	0
dense (Dense)	(None, 6)	196614
dense_1 (Dense)	(None, 1)	7
Total params: 291,405		
Trainable params: 291,405		
Non-trainable params: 0		

Notice in this definition is how the number of filters doubled block-by-block: 64, 128, 256. This is a common pattern. Since the `MaxPool2D` layer is reducing the *size* of the feature maps, we can afford to increase the *quantity* we create.

Step 3 - Train

We can train this model just like the model from Lesson 1: compile it with an optimizer along with a loss and metric appropriate for binary classification.

```
model.compile(
    optimizer=tf.keras.optimizers.Adam(epsilon=0.01),
    loss='binary_crossentropy',
    metrics=['binary_accuracy']
)

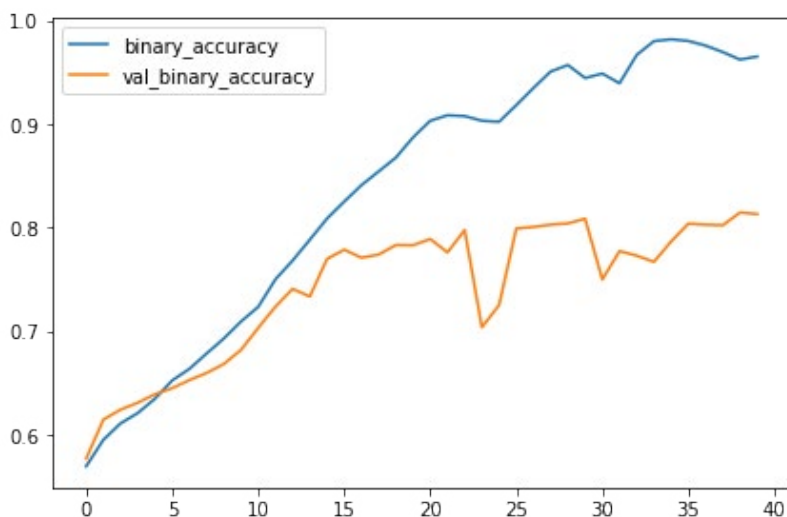
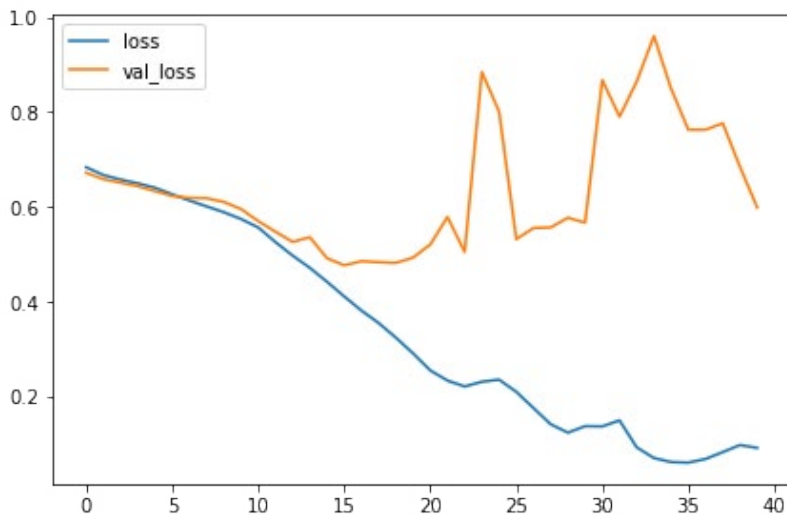
history = model.fit(
    ds_train,
    validation_data=ds_valid,
    epochs=40,
)

Epoch 1/40
80/80 [=====] - 19s 239ms/step - loss: 0.6830 - binary_accuracy:
0.5697 - val_loss: 0.6714 - val_binary_accuracy: 0.5771
[. . .]
Epoch 40/40
80/80 [=====] - 3s 40ms/step - loss: 0.0905 - binary_accuracy:
0.9650 - val_loss: 0.5981 - val_binary_accuracy: 0.8131

import pandas as pd
```



```
history_frame = pd.DataFrame(history.history)
history_frame.loc[:, ['loss', 'val_loss']].plot()
history_frame.loc[:, ['binary_accuracy', 'val_binary_accuracy']].plot();
```



This model is much smaller than the VGG16 model from Lesson 1 -- only 3 convolutional layers versus the 16 of VGG16. It was nevertheless able to fit this dataset fairly well. We might still be able to improve this simple model by adding more convolutional layers, hoping to create features which are better adapted to the dataset. This is what we'll try in the exercises.

Conclusion

In this tutorial, you saw how to build a custom convnet composed of many **convolutional blocks** and capable of complex feature engineering.

Next month

I've received two emails asking for articles but (if you missed my last review), I am writing a coding tutorial every second month. Don't worry though; meanwhile you can enjoy the amazing work done by the magazine and my amazing colleague Marica!

Take care and as always have a great time and always be curious! 😊

TRANS2K: UNLOCKING THE POWER OF DEEP MODELS FOR TRANSPARENT OBJECT TRACKING



Alan Lukezic is a postdoc in the Faculty of Computer and Information Science at the University of Ljubljana, where data scientist Žiga Trojer has just completed his master's degree, and Matej Kristan is a professor. They speak to us fresh from winning the Best Paper Award at the British Machine Vision Conference (BMVC) 2022 for their work on a new training dataset for tracking transparent objects.

In **visual object tracking**, the task is to locate an object in all video frames, given only its location in the first frame. Matej, who organizes annual **Visual Object Tracking (VOT) Challenges** with Alan, co-author Jiří Matas from the Czech Technical University in Prague, and other colleagues, tells us the field has seen many changes over the last few years.

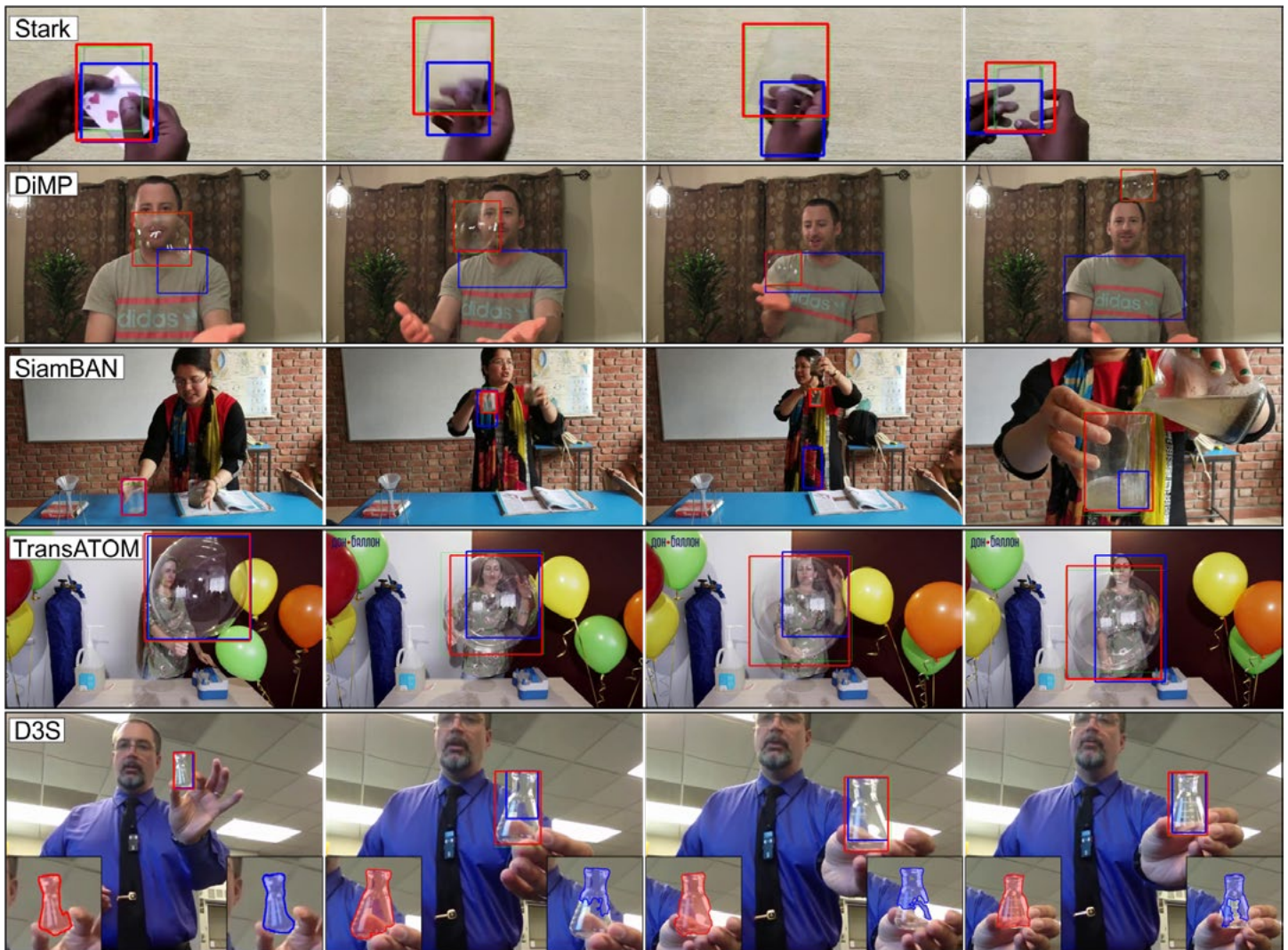
"We've been running these challenges for the last decade and have seen how significantly the field has developed," he points out. *"The methodology has more or less completely shifted every three years. Most research has been on opaque objects,*



Alan Lukezic



Žiga Trojer



which is already a challenging problem.”

In this problem, **the tracker is initialized on an object in a single frame**, the only training example it gets. It does not know what it will track beforehand. Through **self-supervised learning**, the algorithm has to learn the appearance of what it is supposed to track so that it can adapt to robustly localize the target in all remaining frames, even though it may change its appearance or be occluded.

The development of this field depends on the **datasets available for evaluation and training**. Recently, a dataset for evaluation on transparent objects emerged. Transparent objects do not just change appearance when they change their position in space; as they are see-through,

their appearance also varies due to the background. That presents an even more challenging problem.

“We saw that this problem was not solved in the community,” Alan tells us. *“Maybe a year ago, the authors presented this benchmark for **evaluating tracking algorithms on transparent objects**. We saw that there was a possibility to evaluate the algorithms, but there was no data to train these algorithms. That’s how we came to this idea to solve the problem.”*

The team proposes a **new training dataset for tracking transparent objects** and hopes to foster the development of **new algorithms** in that direction. The paper’s contribution is a dataset rather than a tracking method.



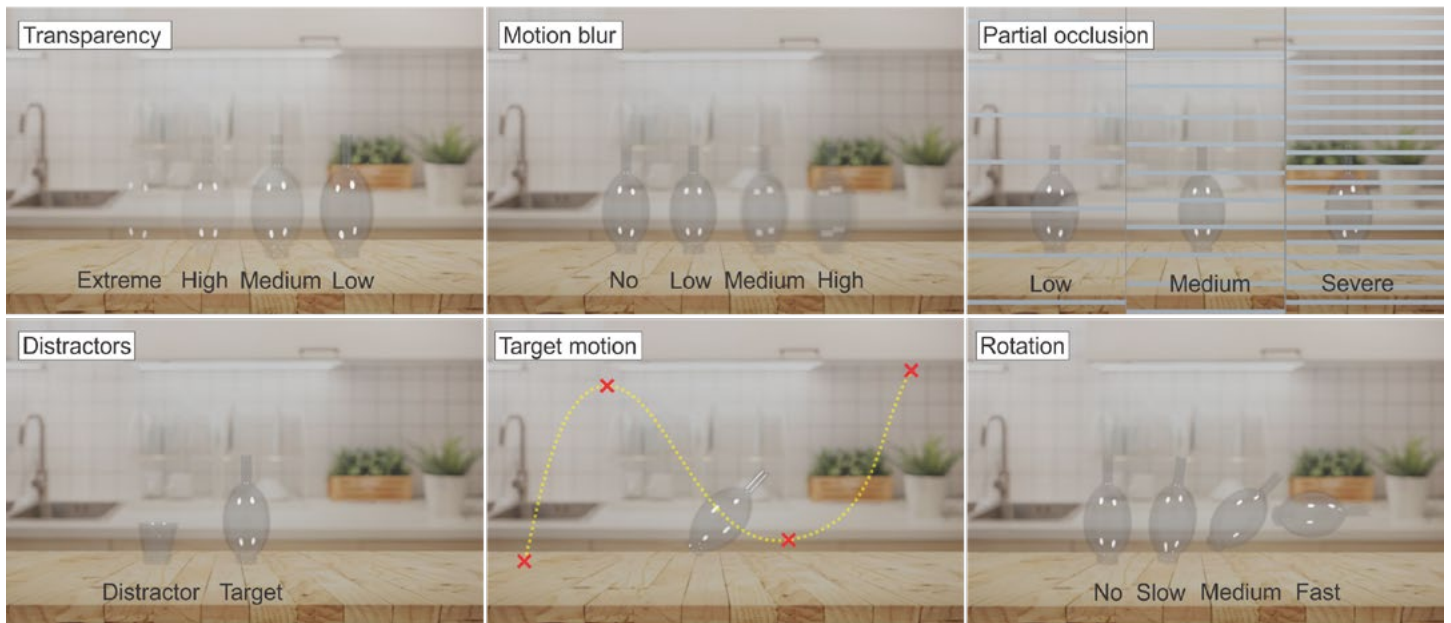
“When we started the challenges, the subfields of computer vision were separated methodologically,” Matej recalls. “You could draw from other fields, but not directly. We’re always drawing from object detection, feature matching, and geometry when we’re designing tracking algorithms, but what changed was **deep learning**. Deep learning came along, and the computer vision field started to converge methodologically. That allowed us to disseminate results faster across the subfields. We’re now drawing more strongly from subfields such as **detection, optical flow, and motion estimation** and integrating those into **new algorithms based on deep learning.**”

The team must be doing something right because scooping the **Best Paper Award** at an international conference is no mean feat. Can Alan, Žiga, and Matej offer any insights into what they think captured the judges’ attention?

“Maybe they realized that the paper is useful for a wider audience,” Alan ponders. “It could be useful in the industry – there are many people in the glass industry, for example. Also, it’s presented for tracking transparent objects, but it could be applied to other fields, such as object detection. I think the paper shows **how to extend the training dataset for other problems.**”

Matej agrees:





“When you present a method, the method may change the direction of development of methodology in certain fields, but with a training dataset in this context, we enabled the development of the whole field of transparent object tracking regardless of the method. The methods will be based on deep learning, for sure, allowing the development of any method and research and exploration in that. It might be impactful, but we will see that only after several years.”

Žiga adds:

“We showed that you can generate synthetic training data, and if you use synthetic data, the boost is still possible, which is good because the cost is low, and it’s not time-consuming to gather.”

In **opaque tracking datasets**, there are objects you are supposed to track, and there may be other objects in the scene, but they will not typically be very similar to the tracked object. In **transparent tracking**, this is different. Where you have a number of transparent objects in a scene, you select one object, but there will be many distractors – other objects that can

confuse your tracker and cause it to fail by jumping to them. The team is setting out to address this challenge in their next paper.

“The data is just one part of the research, and the other part is the algorithm,” Alan explains. “Under an existing method, your algorithm will improve with this dataset, but **now we’re looking at changing the algorithm to make tracking even better when multiple transparent objects are present in the scene.** That’s the topic we’re looking at right now.”

Away from their day jobs, the team has a taste for some more adventurous pursuits. Alan’s hobbies include climbing, cycling, and woodworking. Žiga enjoys sports and reading. Matej is a climber, hiker, and scuba diver who likes traveling to exotic places.

“I’m at an age where I’m not making illusions, and I can’t avoid the computer!” Matej laughs. “But hopefully, getting back to normal life after the crisis we’ve all been through, we’ll be able to continue with these adventures.”

Elena Garcés is Senior researcher at Universidad Rey Juan Carlos (Madrid) and Director at SEDDI (Madrid), a startup which she co-founded where she leads a technology team with expertise in render, optical capture, and AI.

A Survey on Intrinsic Images: Delving Deep Into Lambert and Beyond
International Journal in Computer Vision 130, 836-868 (2022)
 Elena Garcés^{1,2}, Carlos Rodríguez-Pardo^{1,3}, Dan Casas², Jorge Lopez-Moreno^{1,2}
 SEDDI¹, Universidad Rey Juan Carlos² Universidad Carlos III³





Elena, let's start from the beginning. You're not from Madrid.

No, I'm originally from Zaragoza, a city nearby.

Our readers already know an awesome female scientist from Zaragoza, [Clara Fernandez](#). What was it like to grow up in Zaragoza as a young, talented future scientist?

I always loved math. In my high school, I had a math teacher who was with us for six years. This man was also very much into programming, so I learned to code very young, and I really enjoyed it. Then it was clear to me that I wanted to study computer science. But this offered many possibilities: biocomputing, consulting, software engineering, or NLP, for example. Computer science is so broad that you can choose any of these. But then, when I started the degree, I was very much into photography and into visual effects. I realized that I wanted to work on tools that allow artists to improve their performance, to allow them to create better content, and create cooler things. I wanted to work on

this kind of tooling for artists so that they can develop the best of their art. I found that computer graphics and computer vision were the way to go. I have always been on the frontier between these two areas and I decided to do my thesis on this with Adobe.

You said two things. One was about the process of studying these kinds of things. And one was about the outcome: you want the end product, the output text or image to be good. Which is more important? The process of studying computer science or the result of computer science?

I really want to think about the application. I always have the application in mind because I think that is what is powerful. I love to do research, and I love to advance science, but you always need to have some clear use case. I particularly love that my job can be used by somebody. It has a clear purpose. I guess that's why I'm into this company, because I really like to do things that *work*. That, to me, is the main difference between pure research in academia and research in the industry. In the industry, you really



need to do something that has some value.

Elena, if I understand correctly, you prefer to eat the cake than to prepare it?

To eat the cake? *[laughs]* Well, I also enjoy preparing it. Of course, I enjoy preparing as it's fun. And then if you don't spend time preparing the cake, maybe when you eat it, it won't taste well.

Was it as linear as you tell it? "I discovered that I like to do this, then I decided that I would do it. And then I did it!"



I don't think it was very hard for me to make the decision. I consider myself a lucky person. I took the right opportunities when they came to me. My family has no scientists. They don't have strong degrees or anything. I had to fight for that when I was a kid. Then, after overcoming that part with my family, I'm in a better place.

Why was it difficult? Didn't your family want a scientist in it?

My family, at some point, thought that computer engineering, like engineering degrees, was too hard.

I think that you showed them!

I think so! *[laughs]* But yeah, it's fine with my family.

What do they think now?

They are super proud!

We are also super proud of having you, Elena. How was it to be a student in Zaragoza?

It was fun! I enjoyed it. My professors were very good in general. The research group where I did my PhD was very strong and it is still very strong today. I could have gone to

bigger universities, but I do not regret that at all. There are opportunities everywhere!

What brought you to Madrid?

That's a different story. I was doing a postdoc in France, at Technicolor. At some point, I had to decide what to do next, whether to stay at Technicolor or find a postdoc elsewhere, find a job, or whatever. I was contacted by some of my colleagues from Zaragoza, about to build this startup, Seddi. I absolutely loved the idea! I have been working here for four years. We do digital fashion and sell it. As you know, the fashion industry is super polluting. It's one of the most polluting industries in the world. It contaminates water. There are many carbon emissions, and so on. It's not sustainable. Everything is fabricated. You need to fabricate clothes to be able to predict how it's going to fit on your body. Our mission is to build digital clothes that can be tried on in 3D in the virtual world. Then you don't need to fabricate them to see how they are going to be. The idea is to reduce the processing cycle, the fabrication

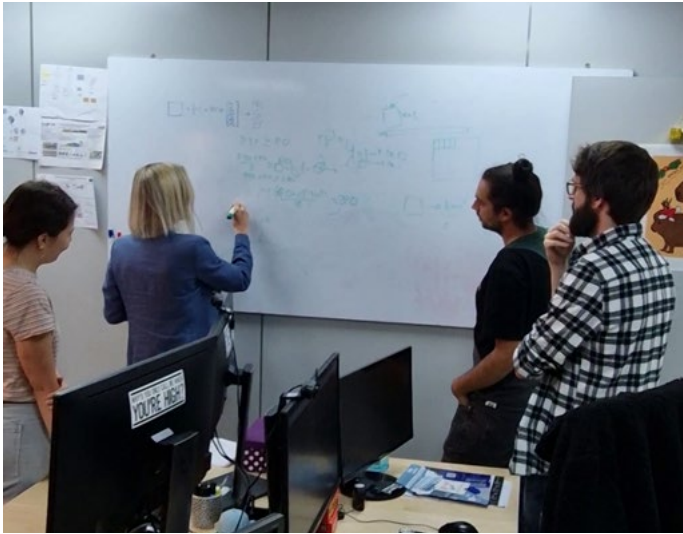
cycle, to reduce oil and to reduce pollution, and the like. I really loved that idea and I really love fashion, so to me, working on that is the perfect combination of my expertise, and it's a field that I love. I could not be in a better place right now!

Who is the client for this kind of application? Who will buy it from you, and who gets value from it?

The brands and the designer, because right now they create the pattern, like building the physical patterns of the garments they cut up, and place them and sell them. With our tools, they can introduce the digital workflow in 3D. They can communicate the style online without the need to fabricate it. That is very powerful. And this is for each particular body, not just a general mannequin. It's for brands and consumers. For example, you or I can go shopping, and we can have our avatar and test these particular clothes. We can see how it's going to fit us, not the model!

So instead of looking at myself in the





mirror, I will look at myself on the computer?

Exactly.

Does it make it a longer process until it gets in the hands of the consumer?

In theory it doesn't, because you will pick your exact size, the size that fits you. There won't be any returns. The clothes will be built according to my size, after I order it. Whereas when I am in a shop, I take it and bring it home.

It sounds fascinating. Maybe you want to give us a taste of what is your day-to-day work on these tasks.

My days are full of meetings. I have three PhD students, so I meet with them to track their research. I also have meetings with my engineering team. We plan and review their work to check that everything is correct. Then we have other kinds of meetings to decide what to do next. That's one part of my day, and then the other part, I do research myself. I read papers or write articles. I also teach a little bit.

Do you also code?

I code very little. Like, almost nothing.

I thought that you liked coding!



I used to like coding, but when the team starts to grow that much, and there are many more management activities, I cannot code anymore. I have to keep track of my student's research which is super time-consuming. I also teach a little, so I also have these obligations. Then I also have Masters students. When you have that many management responsibilities, there's little time to focus on coding.

You teach and work in Madrid.

In Madrid, yeah.

What does Madrid have that Zaragoza does not have?

It's a much more vibrant city! It's much more fun. It's big, and it's one of the capitals of the world. You can find people in the streets at 4 AM in the morning!

I was there one day this summer on my way back from CVPR, but it was the wrong day because there was the G7 conference. The city was completely blocked by police! I chose the wrong day. Or maybe the G7 chose the wrong day... [we laugh]

Everything was locked by the police.

You went abroad to intern here and there, but you seem to spend most of your young career in Spain. Is there a reason?



No particular reason. I found a very nice research group that I was very comfortable working with. I also have a partner; he is also a doctor from the same university. Maybe that shaped my decision a little. But I didn't feel like I was missing something by staying in Spain. I felt that I had the same opportunities.

We spoke about the past and the present. Let's speak about the future. Where are you going?

Where am I going? Hopefully, to change the world with digital fashion. As for myself, I have to say; I don't know. I see myself probably as a professor at the university in ten years or so. I would like to be there. But I am very excited about what I do now.

Let's play a game! Can you give yourself a target goal in ten years to achieve? Then we will speak again ten years from now and see whether you achieved it or not.

In ten years, I would like to be a professor at a university with a large student body and a group. That's where I will be.

Building a sustainable model in this market and passing on your legacy to some successful students seems to be good enough for you. Is this the career



that you want?

Yeah, this is exactly what I want.

Do you have a message for humankind that you'd like to share?

That's a huge responsibility. I only have a message for women out there.

Fine with me...

[laughs] I don't know why, but sometimes we are more insecure than men. My recommendation for them is that they trust themselves. They should trust their instincts and believe in themselves. I don't think we have emphasized that enough. We have been afraid of everything and we have had all these fears and insecurities; I don't know why, but women have more!

I will test your confidence, Elena. How many chances do you give yourself to achieve the dream career that we mentioned one minute ago?

I'm going to make it!

A percentage. 100%?

Yeah!

So you are already secure enough. That's a good thing. Ladies, learn confidence from Elena. She's got it! [we laugh]

Mohamed Ali Souibgui has recently completed his PhD in Universitat Autònoma de Barcelona. He has been working as a researcher in the Computer Vision Center of Barcelona. His research mainly focuses on the enhancement and recognition of historical document images, especially the low-resource manuscripts, i.e, the scripts with few available datasets like the ciphered manuscripts. The thesis was done under the European project [DECRIPT](#). Congrats, Doctor Mohamed!



Optical and Handwritten Text Recognition (OCR/HTR) systems are now achieving good results using recent **deep learning** tools. However, those systems can fail in some scenarios. For instance, when the image quality is degraded or the addressed script is with low-resource (lack of labeled training data). To address this, novel deep learning models and training strategies are proposed within the thesis. Those contributions can be categorized into two main lines of research: **document image enhancement** and HTR in low-resource scenarios.

For the document image enhancement, the problem was treated as an **image-to-image translation task** and several deep learning tools were employed as solutions. First, a conditional **Generative Adversarial Network (cGAN)**, composed of a generator and a discriminator, was used to generate clean images by conditioning on the degraded ones. During training, the generator is producing an enhanced image and passing it to the discriminator to decide whether it is real (originally clean) or fake (cleaned by the generator). This training was done using an adversarial loss with a min-max game. Second, an evolved approach was introduced by adding another component during training that read the text of the enhanced images and forces the generator to produce more readable text while enhancing the image by a CTC loss. This was improving the results to produce images that are as *clean* and *readable* as possible. Finally, a model was proposed basing on **vision transformers** instead of the convolutional layers, which further improves the results. An example of enhancing degraded images by our models is presented in Fig. 1.

For text recognition in low-resource, the goal was to propose generic approaches to recognize different types of historical ciphred texts, as the ones shown in Fig. 2. First, A **few-shot learning** approach was proposed to recognize the historical ciphred manuscripts requiring only a few examples (usually 5) from any new script, even not seen during training. Second, a **one-shot** character generation method was used to generate handwritten text line images for training. The generation was requiring a single example of each character. After that, a **self-supervised learning** approach was also used to learn rich representations from unlabeled data. In the thesis, it was shown that these representations serve as a good start to train models, especially when the labeled training data is limited. Thus, it is a good solution for the low resource scenario when there is a big amount of data, but most of it is unlabeled.

In the future, Mohamed plans to extend his work by exploring **diffusion models** for image enhancement and **continual learning** for HTR in low-resource. He will also focus on model **robustness** and **explainability**.

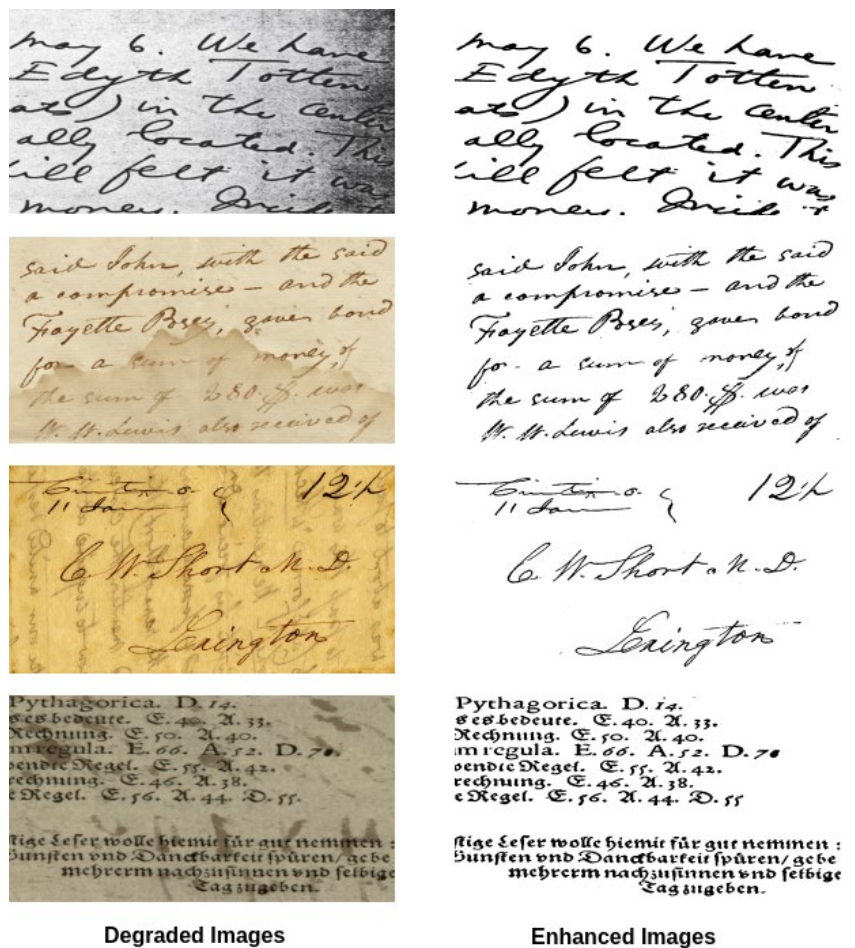


Fig. 1: Qualitative results of enhancing the quality of the degraded images by our models.

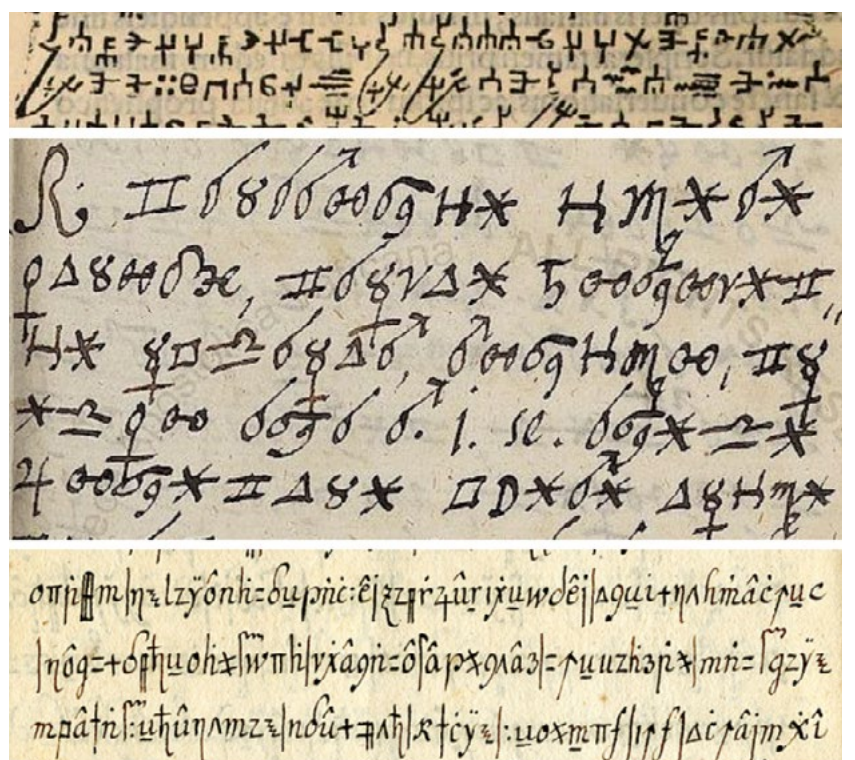


Fig. 2: Examples of handwritten ciphers dated from the 16th to the 18th century. | Top: Devil cipher Middle: Borg Cipher. Bottom: Copiale cipher.



wellcome
EPSRC centre
interventional+
surgical sciences

Robotic systems like those, which
assist in minimally invasive, minimally
invasive and even to encourage
medical techniques.

Modern medicine is a highly dynamic
technology-driven field and it
continues to evolve rapidly.

It is exciting to see the
evolution of these systems and
how they are being used to
improve patient care.

Please address all enquiries to:



SIT

UCI
Division of Surgery
and Interventional Science

G.R01

SHE IS A
DOCTOR NOW

She's a doctor now!

Read about her exciting
work on page 38

9:41



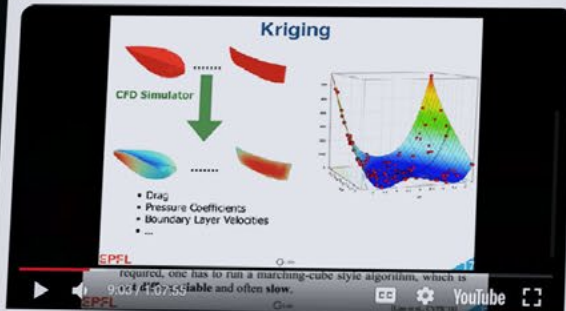
Search Twitter



Guang Yang
@gyangMedIA



A fantastic [#keynote](#) was given by Prof Pascal Fua [@FuaPv](#) on [#Deep 3D](#) [#Surface](#) [#Meshes](#) at [#BMVC2022](#), not only [#research](#) but also lots of [#commercialised](#) [#innovations](#). Now you can watch the full [#video](#) at:



youtube.com
BMVC PF Keynote

12:56 AM · Nov 29, 2022

5 Retweets 11 Likes



COMPUTER VISION EVENTS

Northern Lights Deep Learning Conference

Tromsø, Norway
10-12 January

World Artificial Intelligence Cannes Festival

9-11 February
Cannes, France & online

RE•WORK AI Summit West

San Francisco, CA
15-16 February

SPIE Medical Imaging

San Diego, California, USA
19-23 February

ICCTech Int. Conf. on Computer Technologies
Kuantan, Malaysia
23-25 February

CMVIT Machine Vision and Information Technology
Xiamen, China
24-26 February

ACM/IEEE Int. Conf. on Human-Robot Interaction
Stockholm, Sweden
13-16 March

Digital Health Rewired

London, UK
14-15 March

Emerging Medtech Summit

Dana Point, CA
20-23 March

ISBI

Cartagena de Indias, Colombia
18-21 April

FREE SUBSCRIPTION

(click here, its free)

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

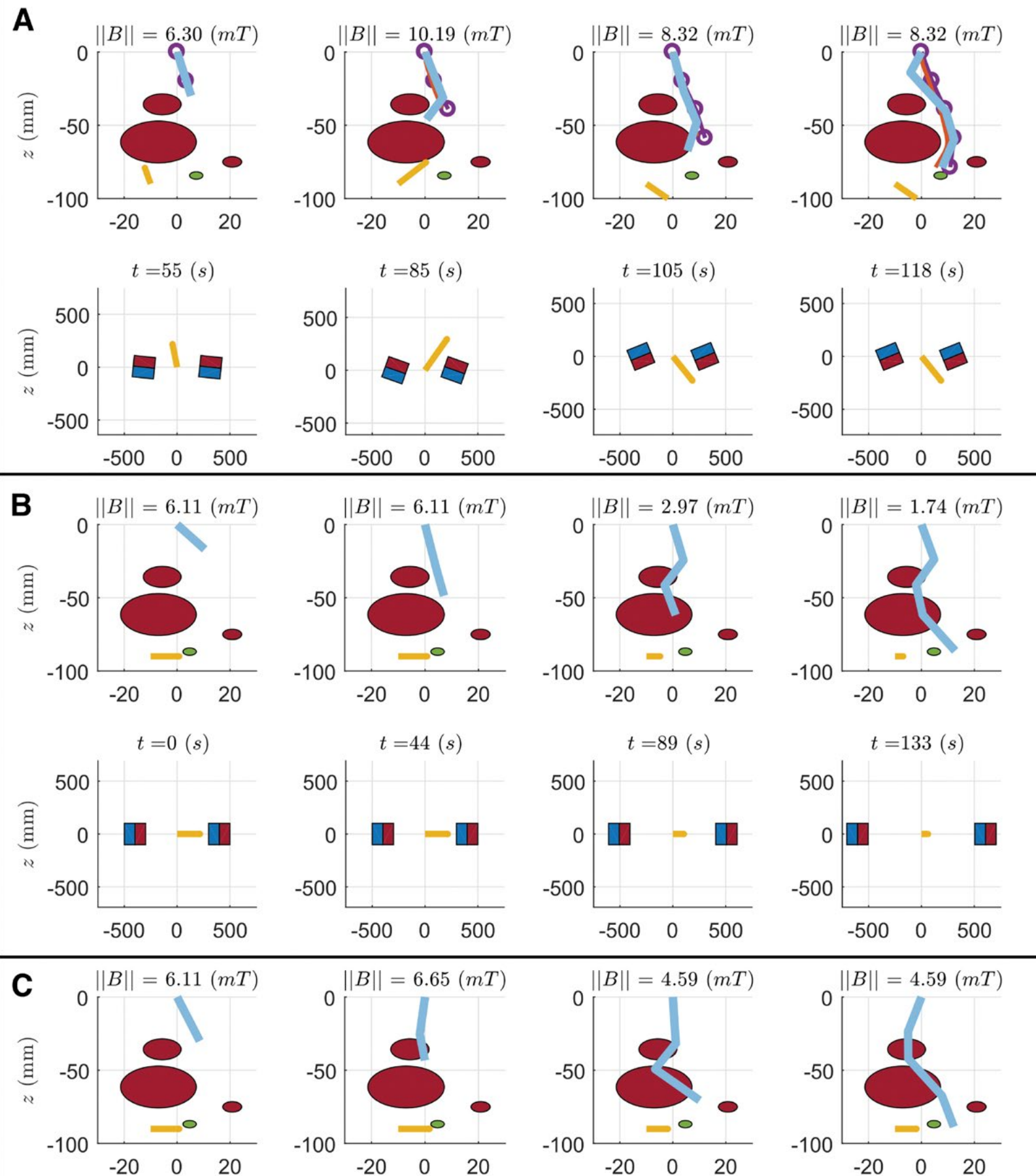
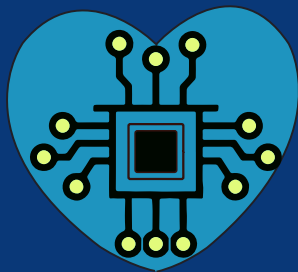
Fill the Subscription Form
it takes less than 1 minute!

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!

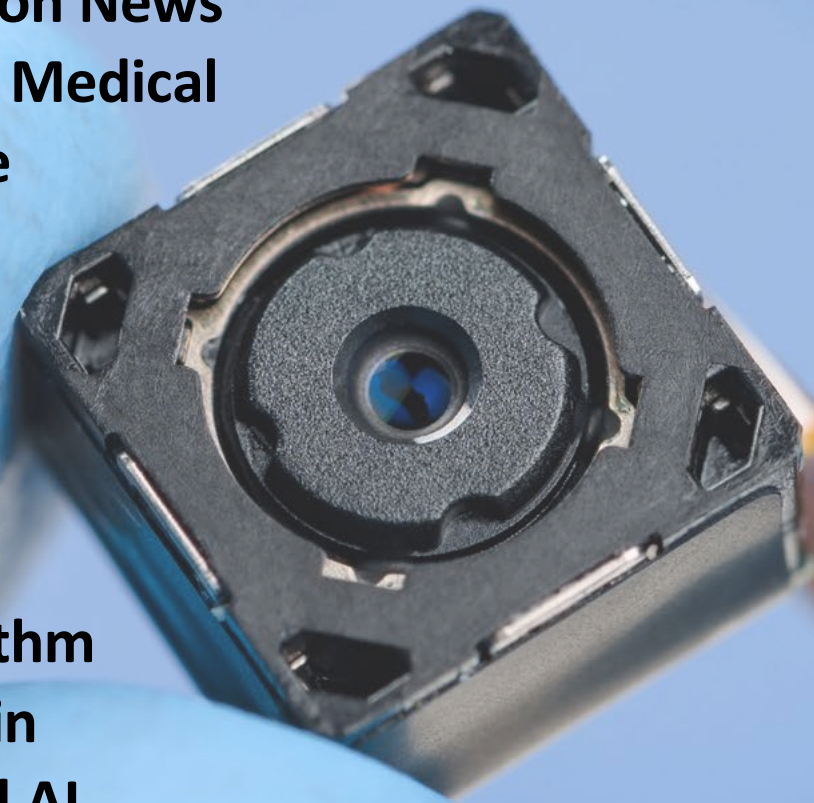
Due to the pandemic situation, most shows are considering going virtual or to be held at another date. Please check the latest information on their website before making any plans!



SOFT TISSUE TRACKING WITH HIGH-RESOLUTION CAMERAS

As our readers well know, we share regularly on Computer Vision News (and on its MedTech insert Medical Imaging News) some of the pioneering work done by [RSIP Vision](#) in the field.

For new subscribers who might not know about it, RSIP Vision is a service company for algorithm development, specializing in medical image analysis and AI. This month we are going to tell you about how high-resolution cameras enable soft tissue tracking, by providing AI with more data to perform stable tracking.



RSIP Vision is transforming the way that we track and analyze soft tissue in medical imaging with the use of **high-resolution cameras**. In fact, high-resolution cameras are becoming increasingly important in the field of medical image analysis and AI, particularly when it comes to **tracking soft tissues**. These cameras have a higher pixel count and better input/output quality than traditional cameras, allowing for greater detail and accuracy in the captured images. This increased resolution is particularly useful when it comes to tracking soft tissue, as it allows for the capture of more data points that can be used by **artificial intelligence (AI) algorithms** to perform stable tracking.



Yehiel Shilo

RSIP Vision's R&D manager **Yehiel Shilo** confirms that the new cameras enable much more visual intelligence, allowing for more accurate tracking and analysis of soft tissue: *"This is particularly beneficial in the field of surgery"*, he adds, *"where precise measurements are often critical for successful outcomes. Moreover, high-resolution cameras are being used in other areas of medicine to track the movement and behavior of soft tissue, such as monitoring the healing process of injuries."* This is essential for a variety of modern medical applications.

Yehiel insists on soft tissues tracking, as these structures are often difficult to distinguish in lower resolution images. In medical applications it is crucial to have a clear visualization, which leads to a **sharper understanding of the location and movements of these tissues**. Innovative AI algorithms are the key to making this possible.

What are the MedTech fields which most benefit from the breakthroughs allowed by high resolution cameras? *"We have been able to achieve significant advances in a variety of medical applications"*, Yehiel clarifies. *"This includes work in **surgery, navigation, rehabilitation, and monitoring of patient health.**"*

This is done through a combination of classic algorithms like **SIFT** (Scale-Invariant Feature Transform) and **SLAM** (Simultaneous Localization and Mapping) with AI neural networks trained through unsupervised deep learning. These algorithmic techniques are used in a very similar way to solve very similar problems in other fields (like autonomous driving) where the algorithms need to choose which features are relevant to the task at hand.

"The use of high-resolution cameras in medical imaging", Yehiel concludes, *"enables us to spot features which are generally undetected, giving us a stable view of soft tissue - like blood vessels. **When analyzing anatomy, the more you improve the resolution the more you see! And this leads to better patient outcomes and more successful medical procedures.**"* Ask your [RSIP Vision contact](#) how they can do it for your project too.

PATIENT-SPECIFIC MAGNETIC CATHETERS FOR ATRAUMATIC AUTONOMOUS ENDOSCOPY



Giovanni Pittiglio has recently completed his PhD and is a Research Fellow at Boston Children's Hospital. He speaks to us about his work designing catheters for electromagnetic navigational bronchoscopy, which has just won the Carter Prize for Best Paper from the University of Leeds.



Lung cancer is one of the deadliest forms of cancer, and bronchoscopy is currently the state-of-the-art clinical practice for its diagnosis and treatment. Once a surgeon has identified a mass inside the lung, they attempt to reach the target in a minimally invasive way, using CT scans of the lungs to extract a preoperative map of the bronchial anatomy to aid navigation.

The surgeon uses a **bronchoscope with a camera** to reach a certain level inside the lung before a **smaller, semi-rigid catheter** inserted into its tool channel takes over and navigates to the target. However, the rigidity of these tools creates several problems, including causing the anatomy to move and the clinician to lose track of where they are.

"We have used a preoperative image to design a magnetically actuated catheter

that navigates autonomously between two given points," Giovanni tells us. *"It's made of silicone, which is as soft as rubber, if not softer. As it navigates, it won't cause pain or create disruption, and it doesn't deflect the anatomy, which is generally more rigid than the catheter. We can navigate in a smoother, less invasive, and more accurate way because we can trust the preoperative image more than if we were using a rigid tool."*

Using magnetics in bronchoscopy is not a new concept. Five years ago, when Giovanni started his PhD, magnetic catheters were predominantly made of a soft tube with a magnet at the tip. Similar platforms are used for colonoscopy, in which Giovanni's PI for this research, **Pietro Vadastrì**, does impressive work. **You can use an external magnet to pull the endoscope from the tip** instead of pushing from the rear as you

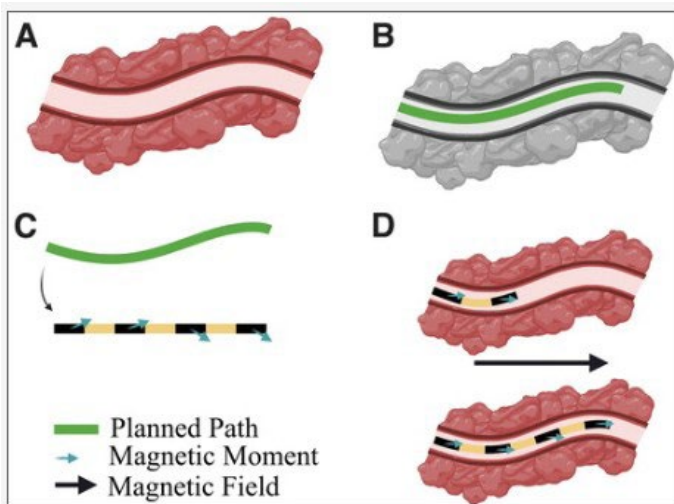


FIG. 1. Magnetic signature optimization for follow-the-leader navigation. (A) Target lumen to navigate. (B) Path planning on preoperative image. (C) Signature optimization based on desired path. (D) Navigation through the anatomy under applied field.

would with a rigid tool. Before magnets were used, tools had to be relatively rigid as flexible elastic tools could buckle in the anatomy.

In this work, Giovanni transitions from having one magnet at the tip of the catheter to incorporating magnetic elements along

its length. Standard tools may allow two or three different bends of the tip, but this one can bend anywhere, and the way it bends is specific to what you want it to achieve.

“Now, we have **multiple soft, magnetized silicone elements and controllability of multiple points along the length of the catheter,**” he explains. “Given the preoperative image, we say we want to go from point A to point B, and to do that, we need to follow a specific path. Then we program the magnetic signature or lengthwise magnetization of those soft elements to form to the shape we want in the anatomy once we apply a magnetic field.”

In this case, the soft catheter is 2.2 millimeters in diameter but does not currently have integrated vision. Cameras can be quite small, but the problem is a lack of light. **Fluoroscopy, an external imaging technique, can help guide here,** but X-rays are undesirable for the patient and medical team due to radiation exposure and

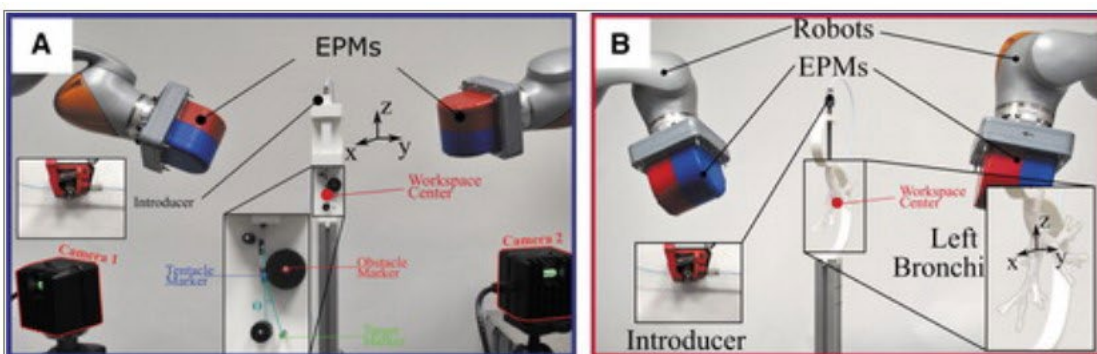


FIG. 4. (A) Experimental setup for comparison between tip-, axially-, and optimally magnetized catheters. Markers were placed on the four magnetic segments of the catheters, on the obstacles and the target. Infrared cameras were used to track these markers. The angle θ is referred to as heading error and measures the angle between the desired and the actual heading. (B) Dual-arm magnetic manipulation platform for navigational bronchoscopy. A 3D-printed anatomically accurate phantom, extracted from CT imaging, is used for demonstration purposes. CT, computed tomography; EPMs, external permanent magnet.

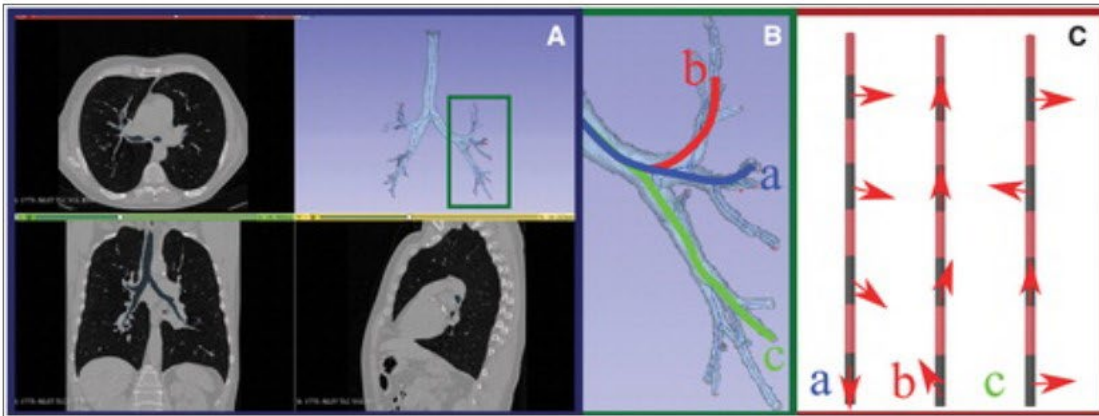


FIG. 5. From a preoperative CT image to an optimized magnetic catheter. (A) The slicer dataset from the CT scan. (B) Datapoints are extracted and aggregated into a connectivity matrix, and paths from proximal to distal nodes are extracted for three different targets. (C) Optimized magnetizations will follow-the-leader shape form along their respective desired navigations.

additional cost. It is a fascinating challenge and a part of the research still in progress. This work solves one problem but opens up new ways to explore the overall topic.

“We asked a list of questions and answered a few of those, but we’re still finding answers,” Giovanni agrees. *“Other researchers are doing amazing work on this topic too. When I started, there were relatively few, but now, new researchers appear every day, which is super exciting because it shows the interest is there. There are many great researchers in Leeds and worldwide seeking answers.”*

As well as winning a Best Paper award, the work has been published in *Soft Robotics*, the leading journal in this important and emerging robotics field. Following such prestigious recognition for his work, what does Giovanni think is the secret of its success?

“This paper was the culmination of four years’ work, not only

from me but many other colleagues and collaborators and all the knowledge and technology we’ve built together,” he reveals. *“It’s a very interesting and novel approach to performing these operations. When I graduated, the University of Leeds*

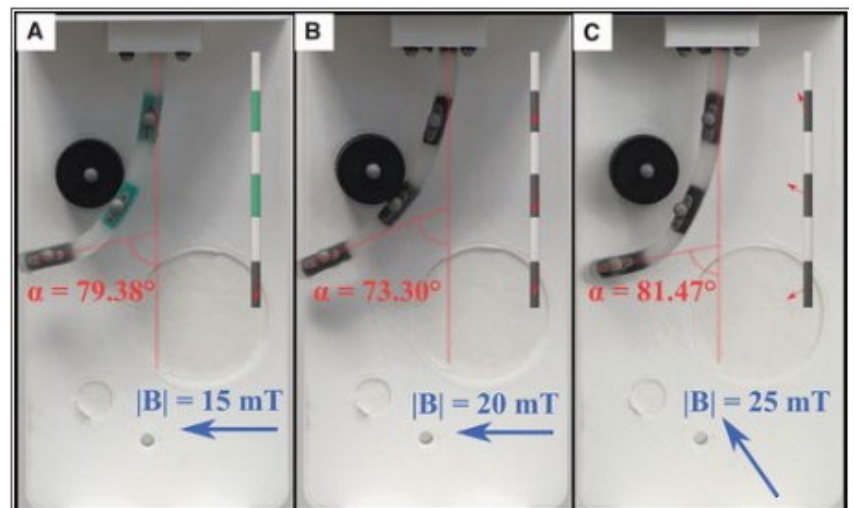


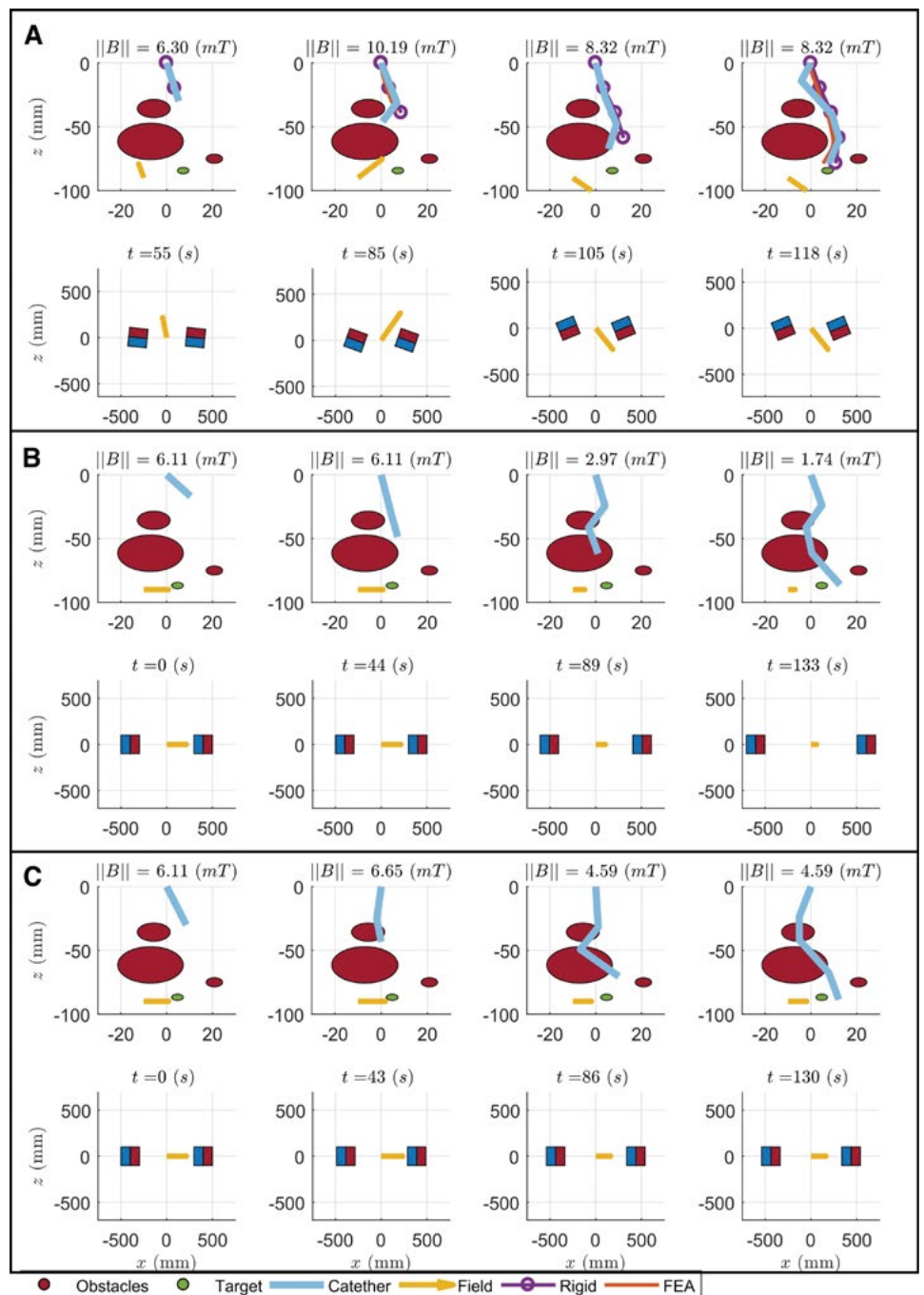
FIG. 6. Results of obstacle avoidance experiments. Comparison of tip-magnetized (A), axially magnetized (B), and optimally magnetized (C) catheters. The angle α represents the angle between the catheter in its actuated and unactuated poses (the unactuated pose runs parallel to gravity).

offered prizes, including Best Paper, so I took the opportunity and applied.”

Giovanni has donated his prize to **EMERGENCY**, a humanitarian NGO founded in Italy in 1994 by the late Gino Strada.

“It’s important for me that the donation is used to bring healthcare where it’s most needed, and EMERGENCY is doing an amazing job,” he declares. “I work in a children’s hospital, and EMERGENCY has helped so many children in the last few years, and it’s still helping them. It also helps save migrant lives in the Mediterranean, which is very important to me. I’ve been away from Italy for the past few years and have been somewhat of an immigrant to other countries, but I’m a privileged European. When everyone is closing the door, many people need EMERGENCY’s help to be welcomed to a country.”

Giovanni is based in the cardiovascular surgery department of Boston Children’s Hospital, working on medical robotics and catheters in general with his current PI, **Pierre Dupont**. His work mainly relates to **using catheters to treat conditions affecting the heart** and investigating the platforms and approaches that should best



be used for these procedures.

“I’ve always done my job because I find it fascinating, and it’s what I want to do in life. I’ve never been in a place where I’m doing it for the money,” Giovanni confesses. “**I want to push the boundaries of medicine! I want to push the boundaries of healthcare!** You do research and, if you do a good job, and if things align, hopefully, you’ll see what you’re working on applied in the future.”

SMART MEDICAL INFORMATICS LEARNING AND EVALUATION (SMILE) LAB

Ruogu “Reanna” Fang is an Associate Professor in the Biomedical Engineering Department at the University of Florida (UF). She is also associated with the Computer Science and Electrical and Computer Engineering Departments and is the Associate Director of the Intelligent Critical Care Center at UF.



Neurodegenerative diseases have been a growing healthcare burden worldwide for several years and are becoming an even more significant problem due to an aging population. 13-16 million people are expected to be living with **Alzheimer's disease** in the US alone by 2050. With global numbers soaring at an exponential rate, it is one of the **top healthcare challenges today**, with billions of dollars spent and tens of thousands of researchers working on it over the past decade.

The mechanism of Alzheimer's disease is still unclear. There are several hypotheses, for example, involving **tau and beta-amyloid proteins** and the role of the immune system, but no definitive answers.

“There is no cure today for Alzheimer's disease, and 99% of the clinical drug trials have failed,” Ruogu laments. *“So-called medication or treatments on the market*

*are questionable in their effectiveness. There are multiple reasons for that. One is that people don't capture the disease early enough. **When symptoms appear, brain cells are already damaged**, and it's too late for medication to reverse that trajectory. Even if patients are identified early, the current treatment plan is one size fits all.”*

SMILE lab aims to use the advanced tools of AI and deep learning to transform the early diagnosis of Alzheimer's disease and related dementia, as well as other neurodegenerative diseases, such as Parkinson's.

*“Alzheimer's is a very popular research topic globally, but **our work is unique because we're looking at the human eye to diagnose the disease,**”* Ruogu reveals. *“Existing research in animal models, as well as human data, shows that **the retina at the back of the eye is highly related to pathological***



changes in the brain. The eye and the brain come from the same embryonic stage and share many anatomical, functional, metabolic, hemodynamic similarities.”

The eyes are recognized as a means for understanding diseases of the brain, such as Alzheimer’s. Still, most researchers use a statistical method to show that Alzheimer’s patients and healthy controls have some retinal differences. This method is like a t-test and has no predictive powers.

SMILE lab proposes an explainable AI model that can diagnose someone with Alzheimer’s now or say whether they will have the disease in the next five years. It explains the decision using fundus images looking at the retinal vasculature and other anatomical and functional changes in the eye.

“The animal model shows that pathological changes appeared earlier in the retina than the brain,” Ruogu explains. *“In humans, it’s also likely that changes caused by Alzheimer’s will appear earlier in the eye than can be observed in the brain.”*

The lab uses the open-access and large-scale biomedical database **UK Biobank**, which has 500,000 subjects from the UK and has been recruited for the past 10 years. Among them, 4,000 people have already been diagnosed with Alzheimer’s disease. The lab uses a subset of the database containing subjects with a retinal fundus image at their baseline visit and then pairs them with age and gender-matched healthy controls **to develop its AI and machine learning platform.** It has built a **modular machine learning pipeline** to classify who has and doesn’t have Alzheimer’s disease from those fundus vasculature images.

“By looking into the eye, which is a window to the brain, we have a big promise to provide a unique, timely, accessible, and low-cost solution to perform Alzheimer’s disease screening for those at risk in the older population,” Ruogu tells us. *“This work is still in the explorative research stage. When it can be applied to clinical practice with high accuracy, precision, and confidence, that will need our and maybe*



even other researchers' efforts towards this, and the nation's support."

SMILE lab is growing fast. Currently, Ruogu has nine PhD students, five master's students, and eight undergraduates.

*"The unique characteristics of our group are that we're not researchers who work only in one domain like biomedical engineering or computer science, but we work on **the convergence of AI and healthcare**," she points out. "We have the domain knowledge from medicine, and with collaborators, we can design and apply advancements in today's technology to make high-impact applications in healthcare."*

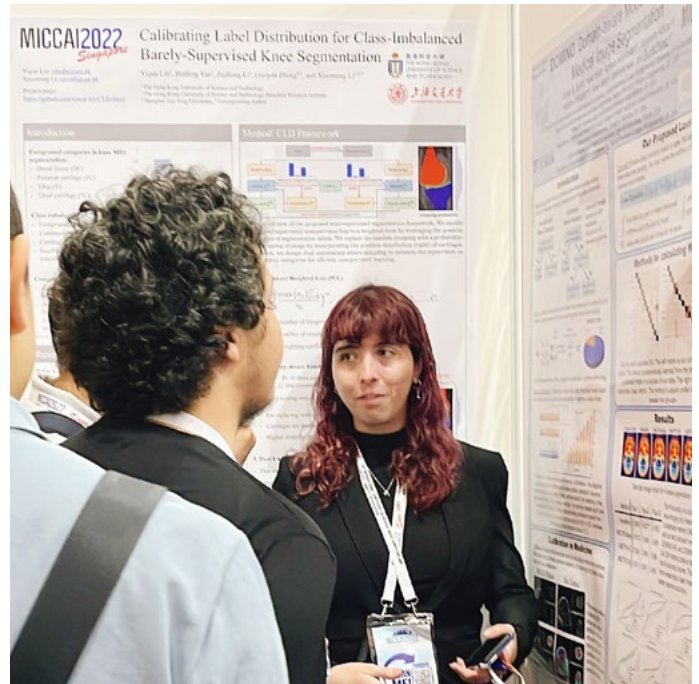


SMILE lab's ambitions know no bounds. It is also working on **designing next-generation AI using neuroscience or biological principles, theories, and findings because it believes today's AI is insufficient.**

*"This line of research is expanding our lab and is very inspiring," Ruogu smiles. "It may not be directly applicable to today's healthcare yet because it includes mathematical theories like the Riemann equation, but **if it develops well, it could have a huge impact on not only healthcare but on general computer vision and natural language processing!**"*



Skylar Stolte is a third-year PhD student in the Department of Biomedical Engineering and a Graduate Assistant in Ruogu Fang's Smart Medical Informatics Learning and Evaluation (SMILE) lab, which focuses on AI and machine learning for neurodegenerative diseases. They speak to us about the lab's revolutionary work on Alzheimer's disease and its exciting plans for the future.



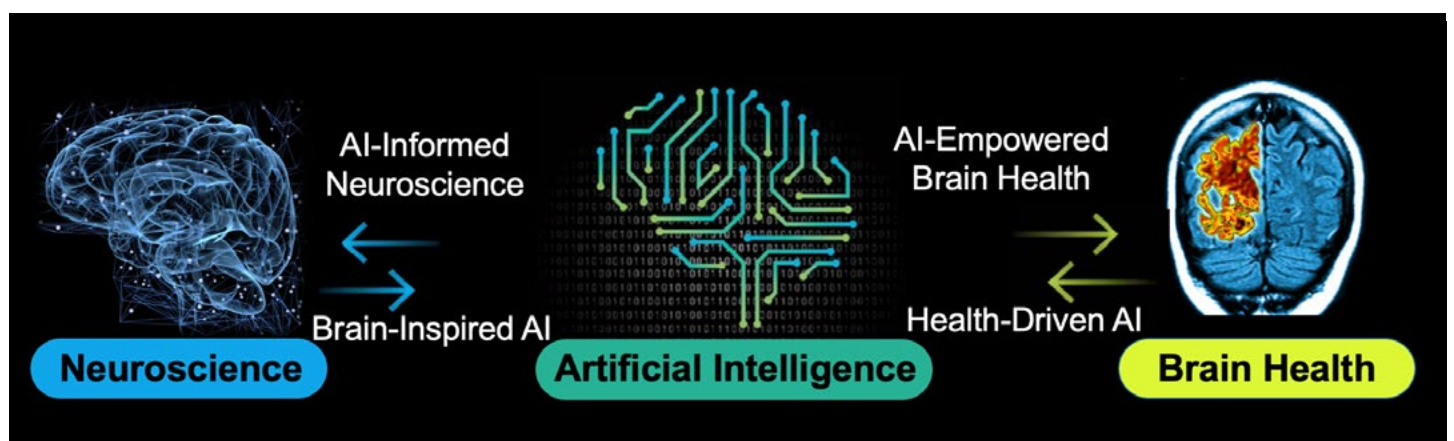
As well as providing a solution for finding people with the disease, SMILE lab is conducting unique research with a view to **developing precision interventions and personalized treatment plans to prevent at-risk people from developing Alzheimer's or dementia in the first place.**

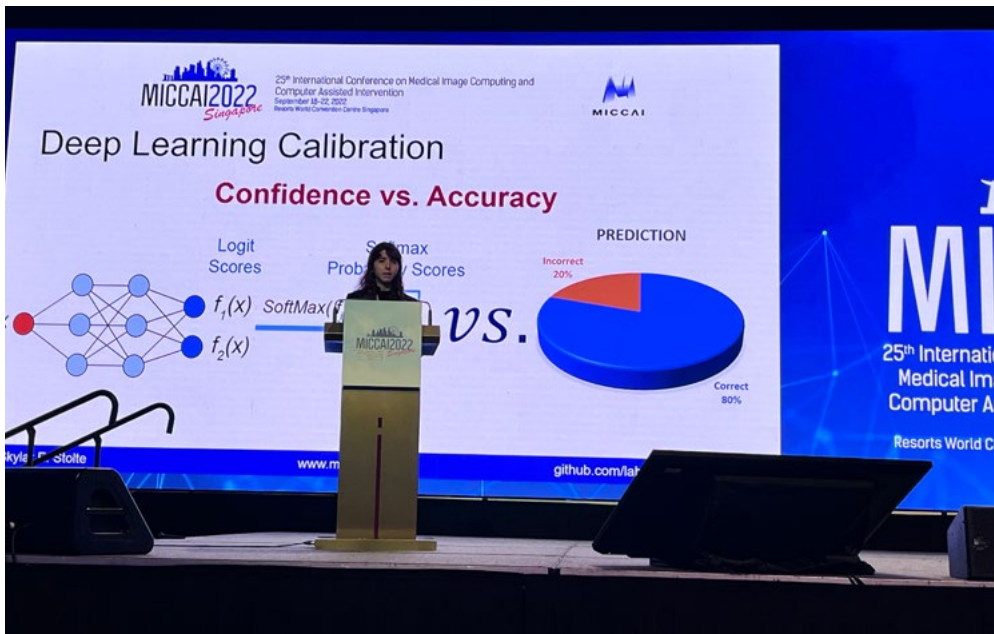
For the past three and a half years, the lab has collaborated with **Adam Woods** at University of Florida's **College of Public Health and Health Professions**. He runs a large-scale phase three clinical trial called ACT, with 379 older adult subjects recruited to **perform cognitive training paired with transcranial direct current stimulation**

(tDCS) to improve their cognitive function and possibly other aspects, such as anxiety and depression, which are risk factors for Alzheimer's disease and related dementia.

Building that pipeline involves **creating a personalized head model**, which requires 20 hours of manual segmentation to segment the whole T1 MRI into 11 tissue types with different conductivity for the electricity to model the current flow in the brain for a specific subject. **Everyone, especially older adults, has distinct brain atrophy, adipose distribution, and skull thickness.**

Skylar's work is an essential part of this, using **deep learning to build a rapid, accurate,**





robust, reliable, and trustworthy AI model that can segment the T1 MRI volume in only three seconds.

“A challenge we face is that our patient population is over 50 years old, whereas public segmentation tools have been built on younger adults around 20 to 30 years old,” Skylar tells us. “These public tools don’t account for the natural processes that occur with aging, such as white matter and gray matter degeneration in older adults. We want a tool that can work better for older adults and better for individualized cases since the human head is highly variable.”

Skylar points out that the segmentation output of many of the tools available is part of an overall pipeline that can take 11 hours or more to run. These tools are not particularly user-friendly, and someone unfamiliar with them may be unable to separate the specific segmentation approach from the rest and would have to wait the total time for each subject to

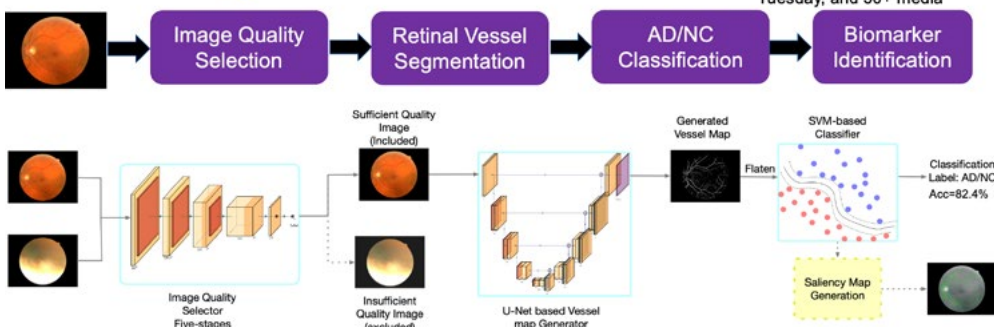
get the output they need. **Three seconds is clearly a massive improvement on that.**

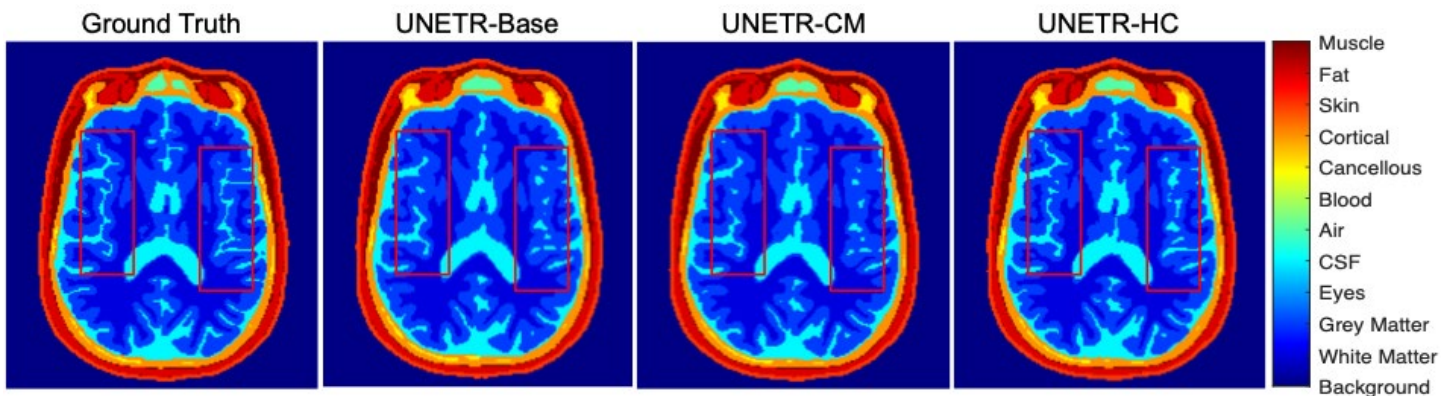
“The segmentation pipeline is used to get the electrical flow model, which doesn’t sound like something that’s directly translatable to practice, but it’s **a crucial part of non-invasive brain stimulation**,” Skylar explains. “The goal is to have more accurate parameters for non-invasive brain stimulation, so we can use this to treat people with conditions like Alzheimer’s. Not all heads are the same, so you want to give someone the best parameters possible to get the treatment that will most help them.”

One of the biggest challenges in head segmentation research matches one of the biggest challenges in deep learning: **a model often struggles with data that is different from the data on which it was trained.** This study uses data from two locations, and where it would train on data from one location, it was

Modular Machine Learning Pipeline for Alzheimer’s Disease Diagnosis via Retinal Vasculature

Featured by The Washington Post, Forbes, WJCB Tech Tuesday, and 50+ media





challenging to translate the performance to the data from the other location.

*“It’s fundamental in getting deep learning in medical practice that you have to be **more translatable to different locations and patient populations**,” Skylar asserts. “That is where the idea of the **calibration approach** comes in. Calibration allows you to get an output in addition to your classification that measures your confidence in your prediction. When we have this additional output that tells us how likely we are to be right or wrong, we can use it to take an output where we think that our model isn’t very certain and have manual segmenters help us improve the labeling, feeding that back into our model to train it to handle these harder cases better.”*

The lab has been using **MONAI (Medical Open Network for AI)** software, which should be familiar to our readers, and was developed from a project originally started by **NVIDIA** and **King’s College London**.

*“MONAI has been so helpful in my research because it comes with many frameworks already implemented in MONAI,” Skylar tells us. “You can take the U-Net from MONAI and use that in your code instead of a random PyTorch repository. It allows you to **take these different frameworks, easily change the parameters, and get your code running without too much effort.**”*

*A big struggle for coders is repositories that aren’t documented well. MONAI is really well documented, and it has all of these neural networks to work with and pre- and post-processing functions that work well on a dictionary of all of your inputs and can handle any problem you have. The libraries available under this framework are also **GPU-accelerated frameworks**, so we have this faster and more efficient software that allows us to **focus on solving future problems** rather than struggling with basic challenges.”*

When CVN spoke to the guys at NVIDIA about MONAI, they promised something almost revolutionary, and it seems they kept their word! Skylar says that because **MONAI has been so useful for this work**, the lab wants to give back by contributing to progressing the tools available for other researchers worldwide.

“We presented our loss term in our MICCAI oral talk this year, and we’ve been talking about the potential for implementing it in MONAI,” she reveals. “This would help researchers because although there are many efficient losses in MONAI, none are calibrated losses. Calibration is a field of deep learning that’s not talked about enough.”



Claudia D'Ettorre recently finished her PhD with the Surgical Robot Vision group at University College London. Her research aimed to partially automate subtasks of robotic assisted minimally invasive surgical operations.

Congrats, Doctor Claudia!

Robotic systems have an increasingly important role in facilitating minimally invasive surgical treatments. In robot-assisted minimally invasive surgery, surgeons remotely control instruments from a console to perform operations inside the patient. However, despite the advanced technological status of surgical robots, fully autonomous systems with decision-making capabilities are not yet available. Despite the technical challenges, adoption of autonomy could potentially introduce multiple benefits, such as

decreasing surgeons' workload and fatigue and pursuing a consistent quality of procedures. Ultimately, allowing the surgeons to interpret ample and intelligent information from the system will enhance the surgical outcome and positively reflect both on patients and society.

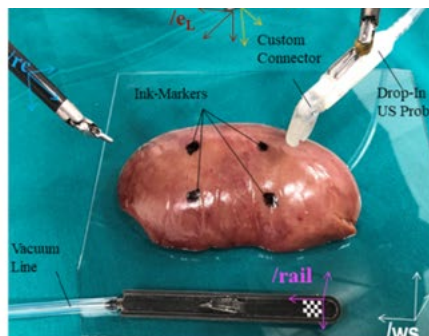
During my PhD, I have addressed some of the technological challenges encountered when trying to achieve partial automation of the pick and place task of surgical objects.

Three main aspects are required to introduce automation into surgery: the surgical robot must move with high precision, have motion planning capabilities, and understand the surgical scene. Besides these main factors, depending on the type of surgery, there could be other aspects that might play a fundamental role, to name some compliance, stiffness, etc. Control

Needle pick-up



US scanning



Organ repositioning

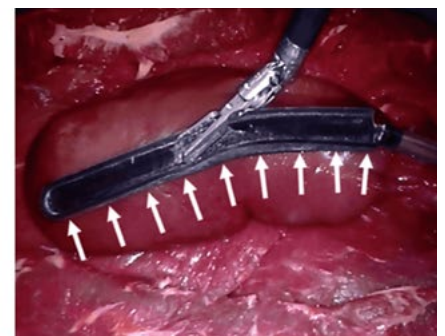


Figure 1: sample of surgical tasks addressed solved with automation algorithms.

algorithms, therefore, should be considering all these aspects when aiming to autonomously execute surgical tasks.

I worked with three main applications of the pick and place task (Figure 1): pick-up of a circular needle prior to suture, pick and place of a Pneumatically Attachable Flexible (PAF) rail for intraoperative Ultrasound (US) scanning and retraction for intraoperative organ repositioning using the PAF rail.

I developed a calibration pipeline embedded in a control algorithm that allows precision movement of the robot, capable of grasping a few millimetre-thick needle [1].

I used learning from demonstration approaches to plan the pick and place of the PAF rail inside a surgeon-in-the-loop control algorithm for execute us scanning in case of partial nephrectomy [2].

I also worked on planning in dynamically changing environments for organ retraction using gradient-based methods to trigger a smoother object repositioning phase during intraoperative procedures [3]. This procedure becomes extremely useful when surgeons need to access areas of the patient’s abdomen that are not visible because covered by other organs or tissues.

As last, to improve scene understanding, I developed a simulation environment where multiple tasks can be learned based on the surgical scene using Reinforcement Learning and then transferred to the real robot (Figure 2) [4].

My experiments proved that automation of the pick and place task of different surgical objects is possible. The robot was successfully able to autonomously pick up a suturing needle, position a surgical device for intraoperative ultrasound scanning and manipulate soft tissue for intraoperative organ retraction.

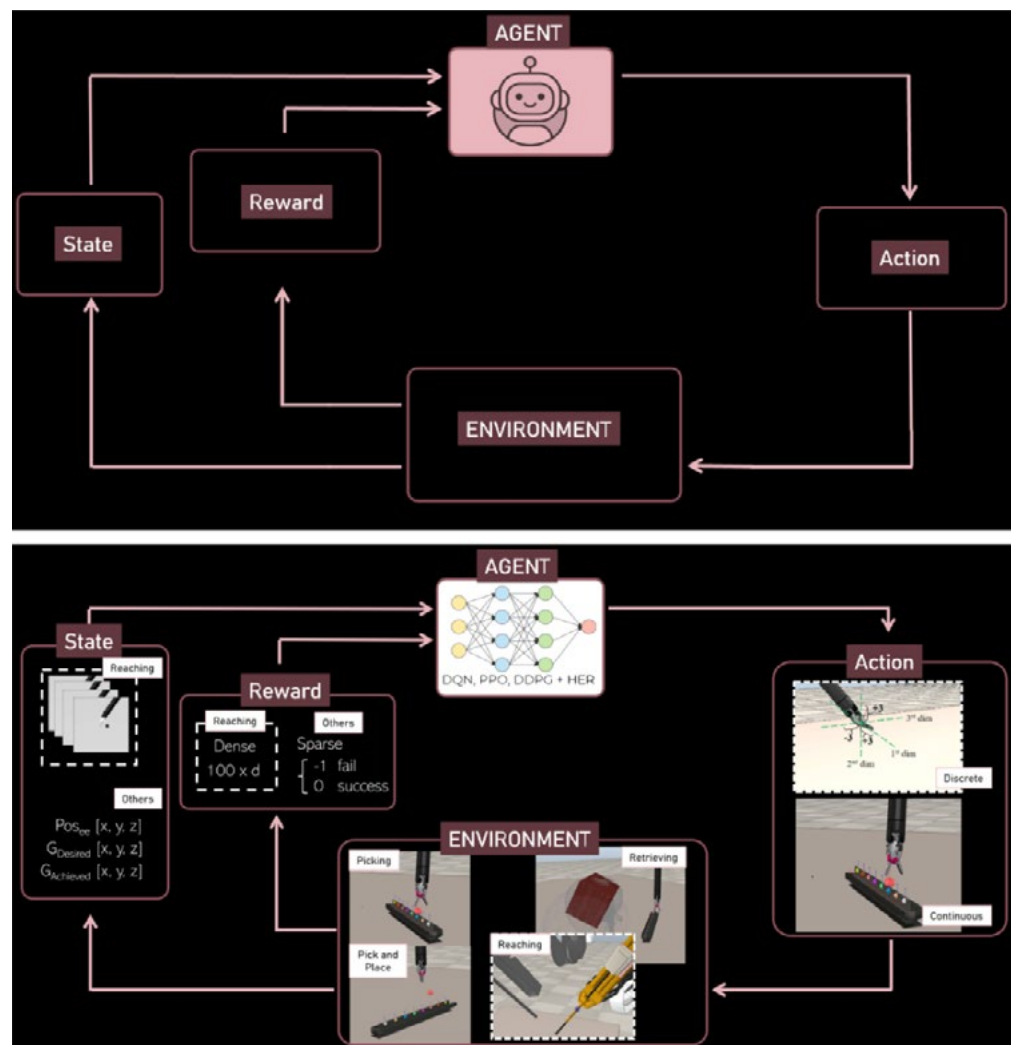


Figure 2: representation of the MDP control loop on the top, and on the bottom how the model has been applied to solving surgical tasks.

HAMLYN WINTER SCHOOL

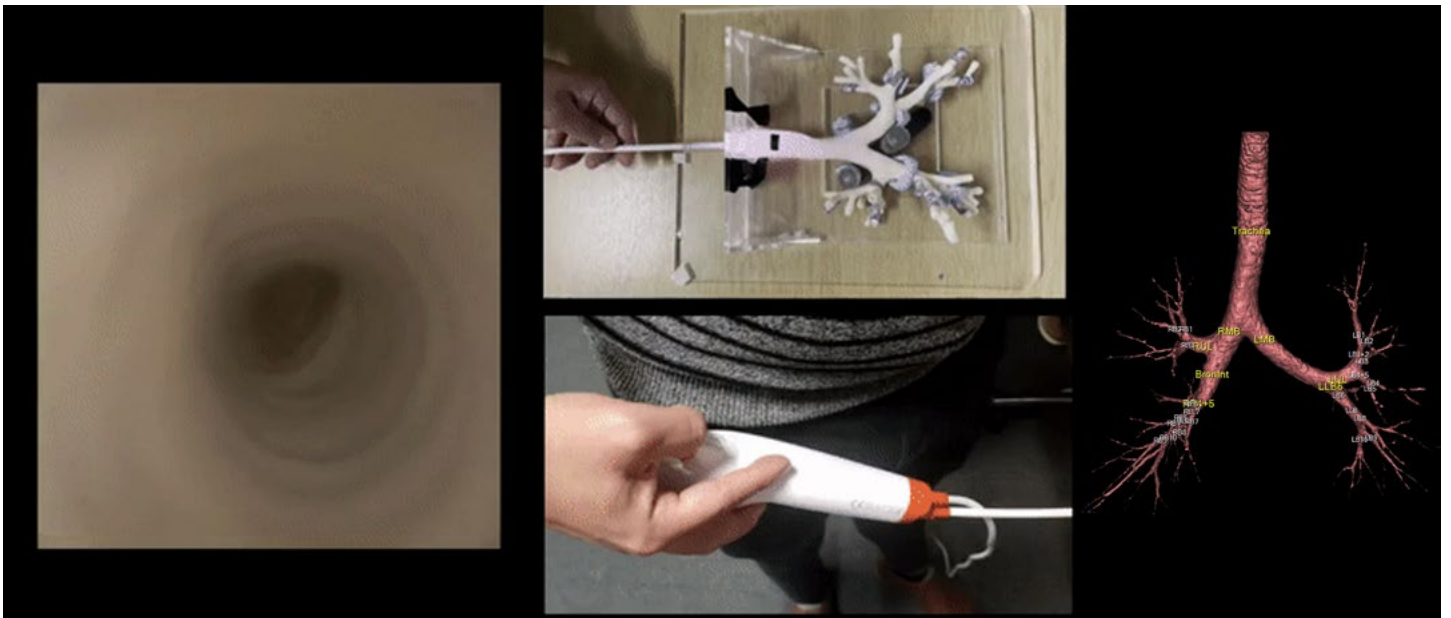
Stamatia (Matina) Giannarou works at the Hamlyn Centre for Robotic Surgery, Imperial College London. She is also a Senior Lecturer in the Department of Surgery & Cancer and holds a Royal Society University Research Fellowship. She speaks to us as the organizer of the Hamlyn Winter School on Surgical Imaging and Vision, which just took place in London.



The Hamlyn Winter School is a week-long course on the technical and clinical aspects of **Surgical Imaging and Vision**. Its program features invited lectures from academics and clinicians, industry talks, hands-on sessions, and group projects.

“We split the delegates into small groups, and each group comes up with an idea they want to work with,” Matina outlines. *“We open our labs, and they have access to our*

*equipment. This year, we had five stations, and each group could take a look at the station and decide what they would like to do. It’s a mini project. On Monday and Tuesday, we mainly have lectures; then, on Wednesday, half of the day is lectures, and the other half is data collection. The groups process their data on Thursday, and on Friday, they pitch their presentation. **What’s the idea? What’s the clinical motivation?***



Why is their work valuable?"

The school first opened its doors in 2013. This year, it hosted 32 delegates, including PhD and master's students with engineering backgrounds, clinicians, and people from the industry. A limited number of scholarships are available each year. Participants come mainly from Europe, but also Asia and the US. Most notably, this time, there was an almost 50-50 gender split.

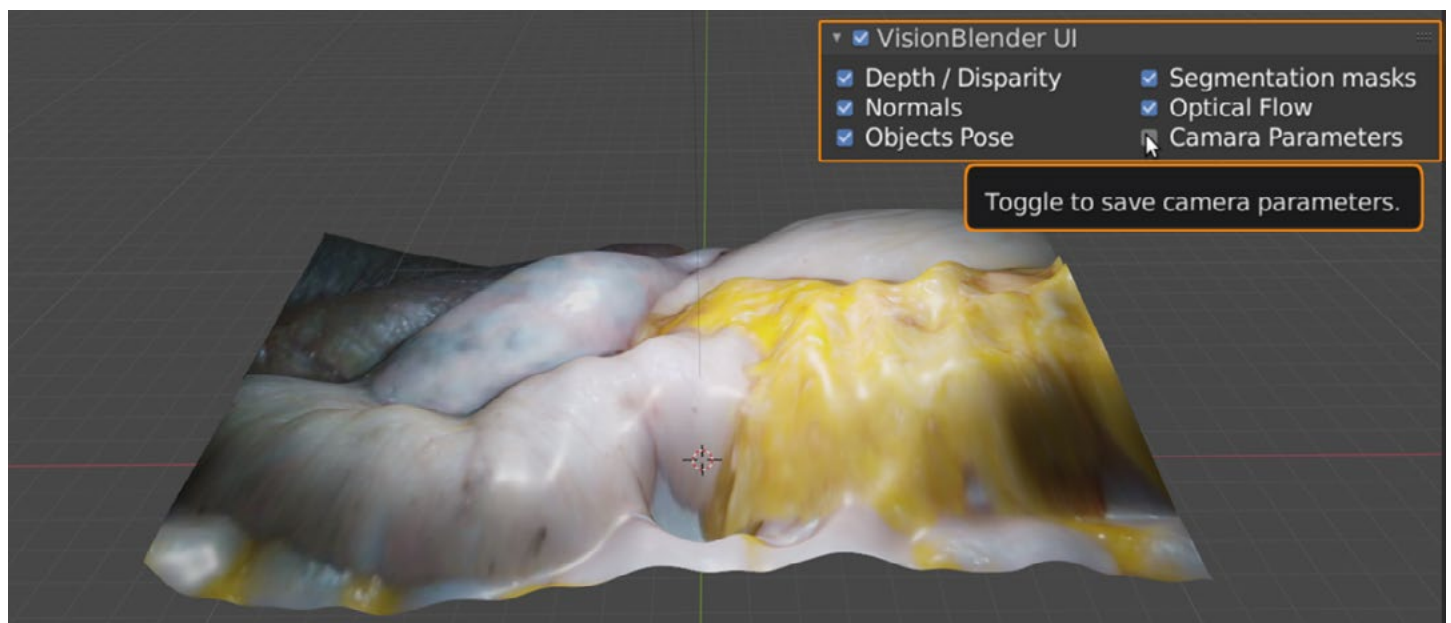
"For our lectures, we invite the best academics in surgical vision and imaging, and each present on their area of expertise," Matina continues. *"We talk about fundamentals and **the current state of the art in surgical imaging**. We cover methods including **3D reconstruction, surgical navigation, surgical data science, intraoperative registration, deep learning methods, multimodal image fusion, augmented reality, and robotic vision** – how we can use surgical vision to guide robots."*

For the clinical talks, neurosurgeon **Hani Marcus** spoke about the challenges in neurosurgery and the robots and navigation tools he's developing to assist. Breast

surgeon **Daniel Leff** talked about using imaging to guide him during operations and identify the margins of tumors. General surgeon **Pietro Mascagni** spoke about surgical data science, skill analysis, and how these can help train surgeons.

*"We had talks from academics working purely on medical applications as well as people like **Federico Tombari**, who works a lot on natural scenes but applies some of his expertise to medical applications,"* Matina tells us. *"Delegates get ideas about other domains and how they can transfer methods from other domains to their work. They see the clinical requirements and gaps and understand what surgeons actually need. This knowledge can help them to translate what they have developed. It's nice to do blue sky research, but if you work closely with clinicians, you can make your work more applicable."*

Industry talks included Intel, who sent a team of engineers and clinicians to present their **new computer vision AI platform, Intel Geti**. Between them, Eugene Lui, Claire Hogg, Thomas Burgoyne, and Mathieu Bottier showed both the platform's technical





and clinical sides and spoke about how it is being used in clinical practice. Other talks came from Pierre Berthet-Rayne, who described his ascent from PhD student to CEO of surgical robotics start-up **Caranx Medical** in France, and Patrick Schrempp from **Canon Medical Research**, who talked about their work on **medical image processing**. Other topics discussed included intellectual property and patents, how this can be done, and whether it is worth it. Industrial partners also presented openings they had for positions, and delegates were able to make crucial connections.

Networking is a vital function of the school that Matina is keen to highlight. Over the past decade, she has seen delegates form ongoing friendships and collaborations that

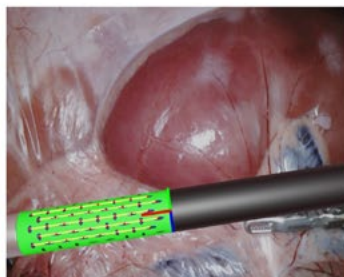
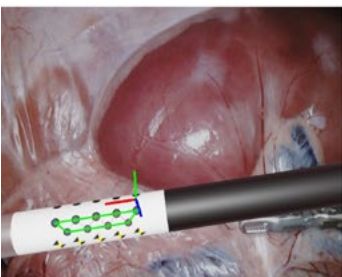
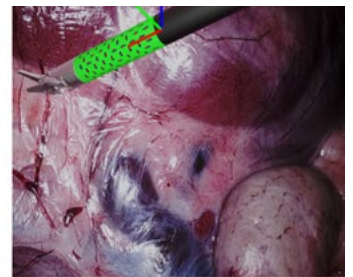
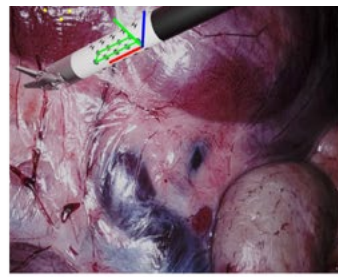
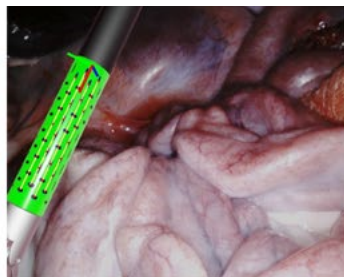
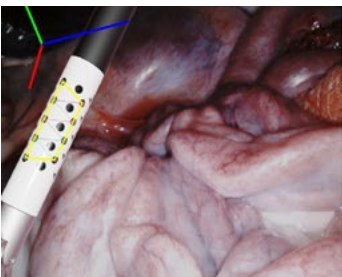
have lasted long past the closing ceremony.

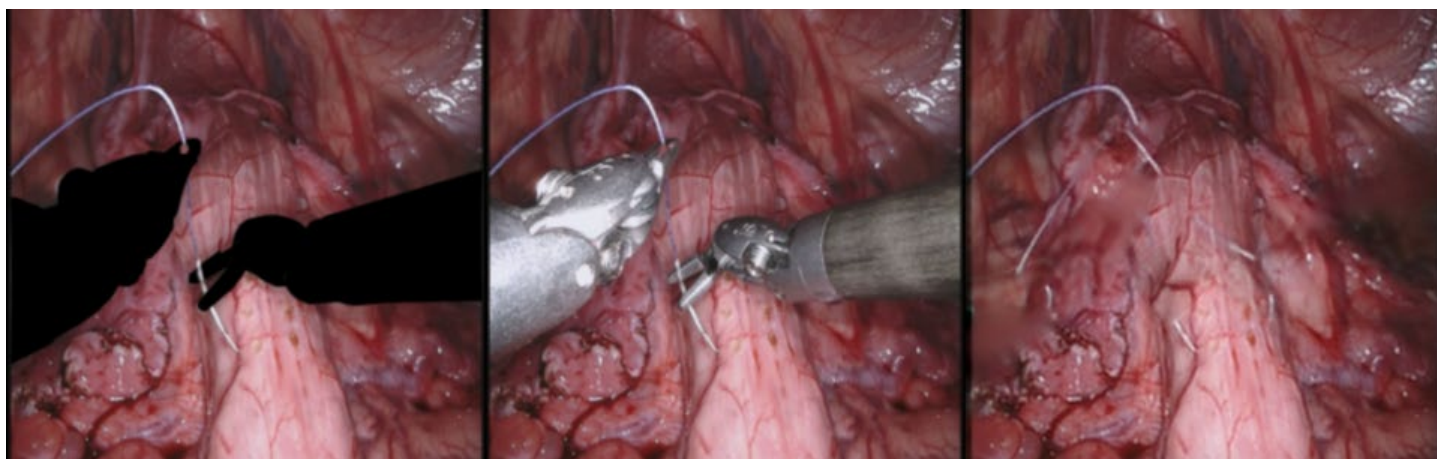
“Every year, people from different countries and universities get to know each other and continue working together,” Matina reveals. *“They even have papers together. Networking with the presenters and lecturers is important too. For example, one group had an idea and got in touch with one of the clinicians who gave a talk. They ended up incorporating his feedback in their presentation.”*

After a decade of successful events, is there anything Matina thinks the school could add or do differently in future?

“There is always room for improvement, and we always consider feedback,” she responds. *“I was talking to Pietro, who provides the material before the school so that students are familiar with the specific areas that will be covered. Perhaps we could give them more details about our stations in advance so they can start thinking about and preparing for their projects. Then they can go straight into the data collection when they get here.”*

Luckily for the students, the Hamlyn Centre has enough equipment and phantoms for each station to have a different phantom.





Stations and phantoms are paired and matched well, so the surgical microscope, for example, has a brain, and the bronchoscope has a lung. Delegates have access to the **da Vinci Surgical System**.

The week is not all work and no play. The Winter School Dinner is a welcome fixture away from the campus in the middle of the program, where participants can have some fun over a relaxed meal at a nearby Italian restaurant.

“For me, the best day is the last day because the students present their work, and it’s a very festive atmosphere!” Matina smiles. *“You can see them competing, trying to create nice slides. Presenting gets them excited, and they feel involved. We also give awards for the best project and runner-up. Afterward, we have the closing ceremony with drinks and nibbles.”*

Away from the school, **Matina leads the Cognitive Vision in Robotic Surgery group at the Hamlyn Centre for Robotic Surgery**. Her research aims to develop computer vision techniques to assist **intraoperative surgical navigation**. The group is developing machine learning models to provide diagnosis support and methods to control robotic platforms like the da Vinci and handheld robots for autonomous robotic task execution. It is also enhancing

the visualization using mixed reality like **HoloLens** and is working on analysis of the surgical workflow and the objective assessment of surgical skills.

Looking ahead to next year, Matina has some helpful advice for prospective Winter School applicants.

“Don’t think twice about applying!” she declares. *“I believe it’s a very nice school because it’s not very big, so you’ll get the opportunity to talk to each other, work with each other, and talk to the lecturers. We get excellent feedback every year. You will get many things from the clinical and engineering aspects. Also, you get to see London at the best time!”*





**IMPROVE YOUR
VISION WITH
Computer Vision
News**

SUBSCRIBE

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

