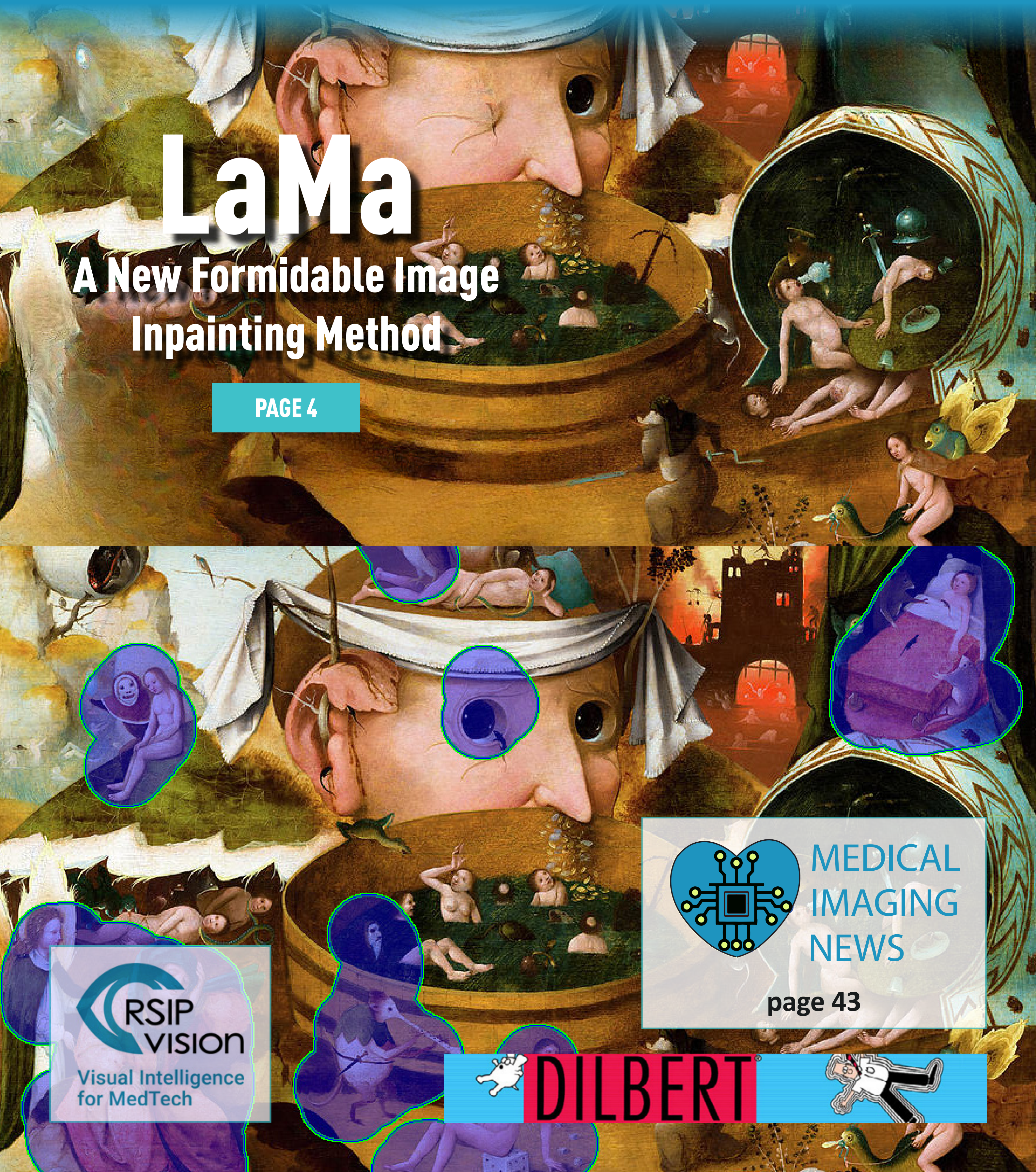
JANUARY 2022

Computer Vision News

The Magazine of the Algorithm Community



LaMa

A New Formidable Image Inpainting Method

PAGE 4

MEDICAL IMAGING NEWS

page 43

RSIP VISION

Visual Intelligence for MedTech

DILBERT

Computer Vision News

Editor:
Ralph Anzarouth

Engineering Editors:
Marica Muffoletto
Ioannis Valasakis

Publisher:
RSIP Vision

Contact us

Free subscription

Read Previous Magazines

Follow Us

Dear reader,

Happy New Year! On behalf of **RSIP Vision**, we wish all our readers and the community **a prosperous and successful 2022**.

We are kicking off the year with a bumper January issue of **Computer Vision News** including many features you will love! The first is a formidable **new image inpainting method** developed by a Samsung-led team. It is the Research of the Month by our writer **Ioannis Valasakis,** and we dare to call it formidable because of its results. Turn the page for the chance to try it out yourself.

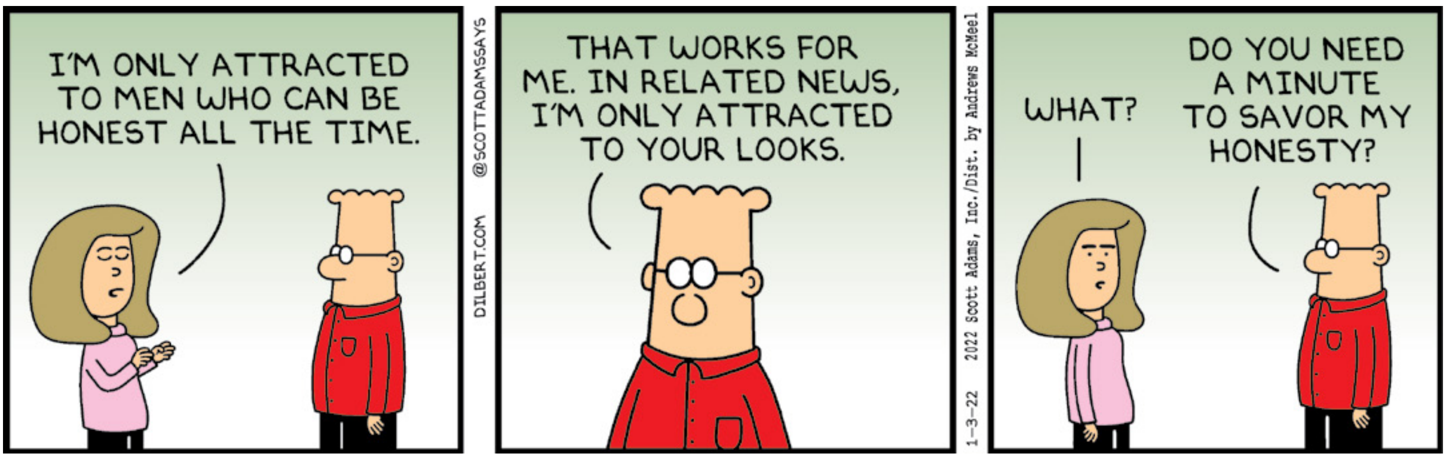
We have reviewed the well-deserved winner of the **Best Paper award** at the recent **British Machine Vision Conference (BMVC)**. **DISCO** is an innovative new approach to accurate scale-equivariant convolutional neural networks (SE-CNNs). The first author is an old friend of the magazine, and you might not be surprised to discover who he is. Find out more on page 10.

Our writer **Marica Muffoletto**'s monthly column is dedicated to one reader's enticing request. If we hear from more readers in the future, we are considering opening a new section called **YOU ASKED FOR IT!** This month, Marica presents a large panorama of how to draw neural networks, with code! Do not hesitate to enrich your skillset by following her guidance. Learn more on page 14.

We are also honored to host again, after more than three years, one of the pillars of the **MICCAI** community: Professor **Daniel Rueckert**. He speaks to us about his work at the intersection of computer science and medicine. You don't want to miss his exclusive interview on page 44 and a **surprise** on page 51.

Enjoy reading this exciting January issue of Computer Vision News - probably one of the richest ever - and **take us along for your next MedTech project!**

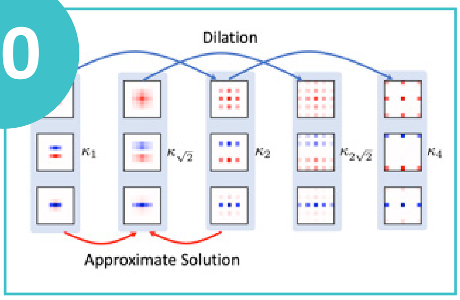
Ralph Anzarouth
Editor, **Computer Vision News**
Marketing Manager, **RSIP Vision**



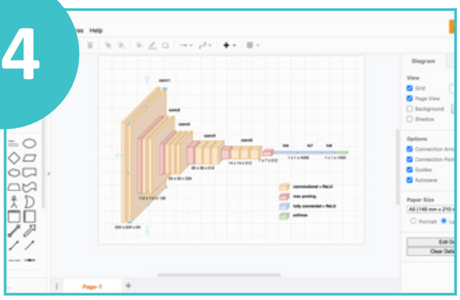
Computer Vision News

04

10



14

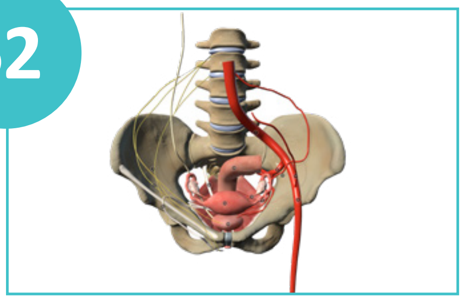


32

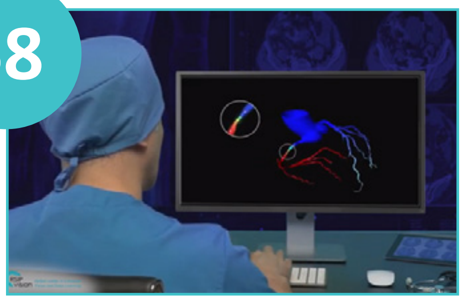
Medical Imaging News

44

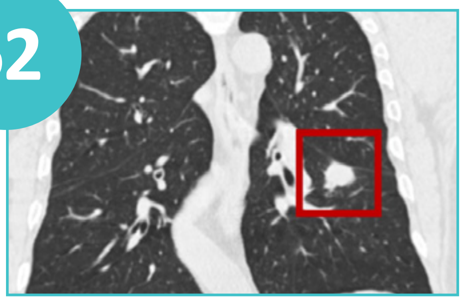
52



58



62



Resolution-robust Large Mask Inpainting with Fourier Convolutions

IOANNIS VALASAKIS, KING'S COLLEGE LONDON

@WIZOFE

Welcome to our last issue before the year, which for a lot of people comes to a change

During our articles this year we looked at topics ranging from neuroscience (I hope that you enjoyed especially the last article, which was a great introductory but also more in-depth programmatic approach to neuroimaging), to liver, cardiac disease and MR angiography, deep learning techniques as applied to biological imaging and medical imaging, to many many more.

What was your favorite? What does excite you? Feel free to contact and let us know of new and interesting topics that you enjoy, or you would like to see covered in the usual channels (social above or directly on Computer Vision News)!

The pandemic seems to have a come-back but this time with hopefully more vaccinated people and less losses. Let's hope that it will be normality, as with the flu and we'll learn to be more thoughtful and careful. For everyone celebrating, have an amazing new year!

Review

This month, we'll present **"Resolution-robust Large Mask Inpainting with Fourier Convolutions"** from Roman Suvorov and colleagues in Samsung Research, Skolkovo Institute of Science and Technology, Moscow, Russia and EPFL in Lausanne. We thank the authors (special thanks to Arsenii Ashukha) for authorizing Computer Vision News to use photos and extra material! I personally need to thank Ralph, for his amazing support, suggestion and help contacting me with his vast network!

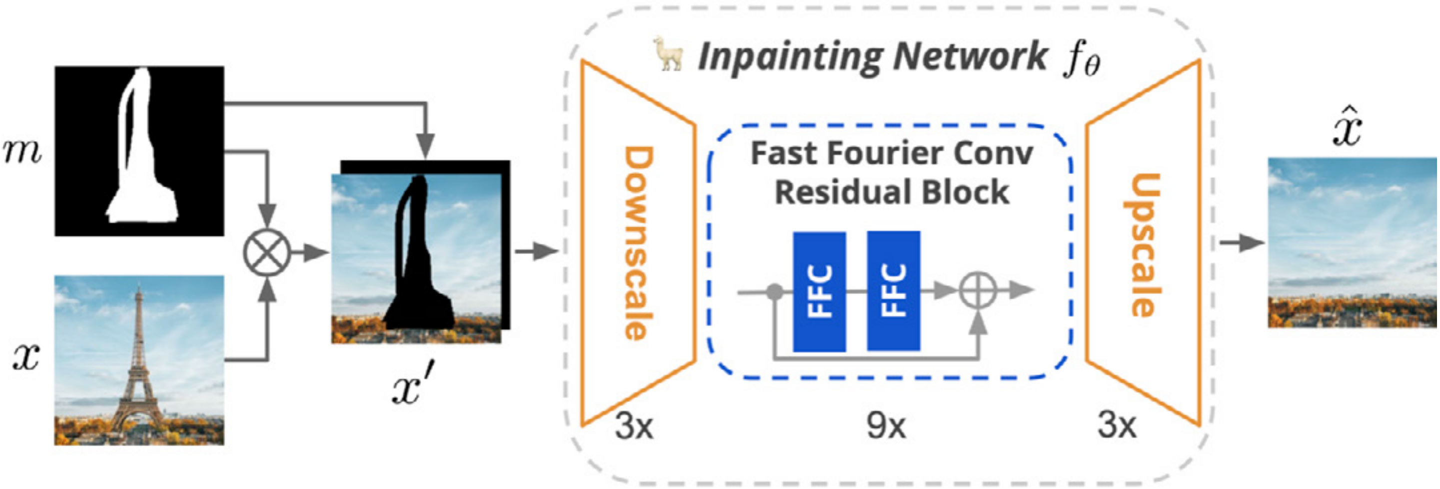
The target of this article and the recommended approach is to attempt and solve the image inpainting problem. I can hear you asking already, **whattt?** This is simply the act of filling the missing parts of an image with a "realistic" approach. It is important to "understand" the large-scale structure of natural images and to perform image synthesis. In the pre-deep learning times, this has been extensively studied (you can explore more about this in the references of the original paper).

Now how is this going to be achieved? The plan is to utilize **Fourier convolutions**, using a **single-stage** deep learning network. The method which was used by the authors to achieve this goal is to inpaint a color image x masked by a binary mask of unknown pixels. The mask m is stacked with the masked image x' m, resulting in a four-channel input tensor. A feed-forward inpainting network was used, named $f_\theta(\bullet)$, which is further referred around as the **"generator"**.

This method proposed is called large masked inpainting (LaMa) and is based on the following:

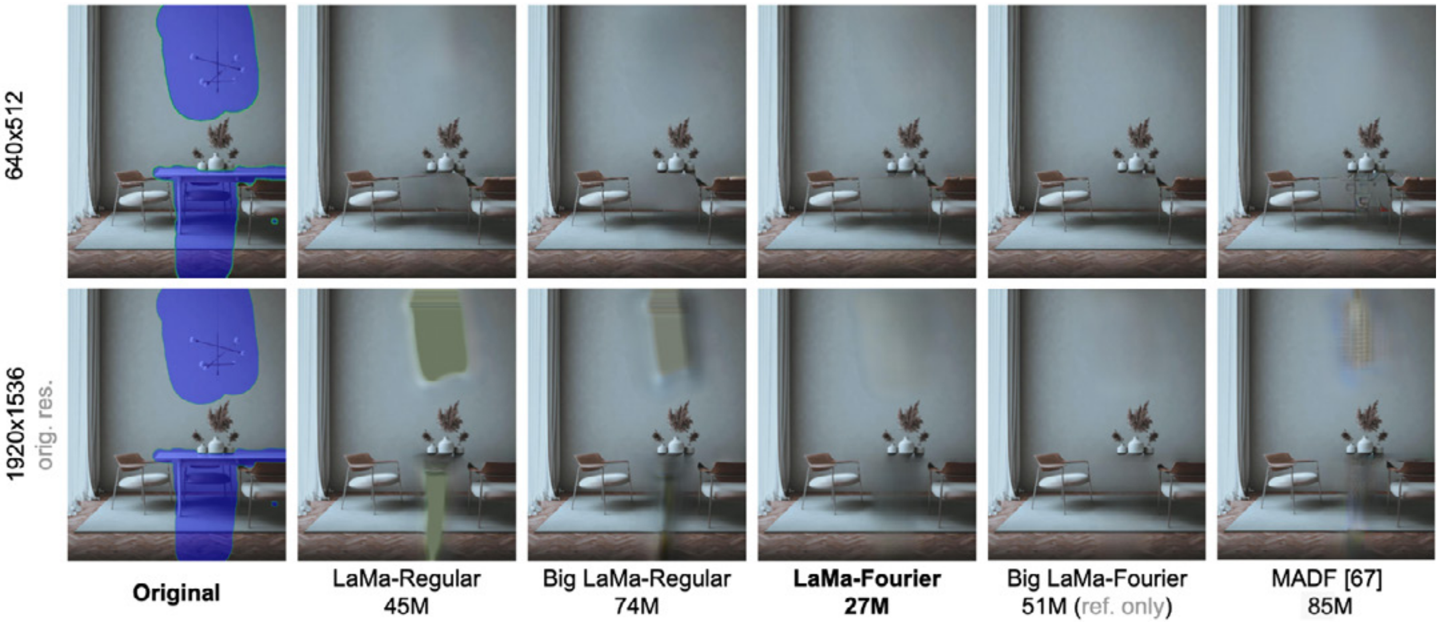
- A new inpainting network architecture that uses fast Fourier convolutions, which have the image-wide receptive field;
- A high receptive field perceptual loss;
- Large training masks, which unlocks the potential of the first two components.

A good architecture should have units with as wide-as-possible and the receptive field as early as possible in the pipeline. ResNet and other conventional fully convolutional models suffer from slow growth of effective receptive field which might be insufficient, especially in the first layers of



the network. Fast Fourier convolution (FFC) is the recently proposed operator that allows to use global context in early layers. FFCs are proven to suit the capture of periodic structures, which are common in human-made environments.

The inpainting problem is inherently ambiguous with many plausible fillings for the same missing areas. To design the components of the proposed loss the following ideas were taken into the game



The image shows the transfer of impainting model to higher resolutions. The LaMa models were trained on 256x256 crops from 512x512 and MADF was trained directly on the latter. With the increase of the resolution the models with regular convolutions produce critical artifacts with the FFC-based models generating finer details and being more consistent structurally. .

High receptive field perceptual loss

The focus of large mask inpainting is shifted towards understanding of global structure. Naive supervised losses require the generator to reconstruct the ground truth precisely. The visible parts of the image often do not contain enough information for the exact reconstruction of the masked part. In contrast, perceptual loss evaluates a distance between features extracted from the predicted and the target images by a pre-trained network. It does not require an exact reconstruction, allowing for variations in the reconstructed image.

Adversarial loss

Adversarial loss was used to ensure that inpainting model $f_\theta(x')$ generates naturally looking local details. A discriminator $D_\xi(\bullet)$ works on a local patch-level. Only patches that intersect with the masked area get the "fake" label.

The final loss function

The final loss function is the weighted sum of the discussed losses.

The last component of the system is a mask generation policy. The way the masks are generated greatly influences the final performance of the system and unlike the conventional practice, e.g. DeepFillv2, a strategy with aggressive large mask generation was used. Samples from polygonal chains were uniformly dilated by a high random width (wide masks) and rectangles of arbitrary aspect ratios (box masks).

The outcome of the network was shown to outperform a range of strong baselines on low

resolutions. The difference was shown to be even greater with wider holes and surprisingly, it can generalize to high, never seen resolutions!

You should have a look at the original paper, especially to see the interesting results of comparisons between the specific architecture and the baselines. An ablation study was also done to examine the influence of different components of the method.

Most real-world image editing scenarios require inpainting to work in high-resolution. FFC-based models transfer to higher resolutions significantly thus the hypothesis is made that FFCs are more robust across different scales due to

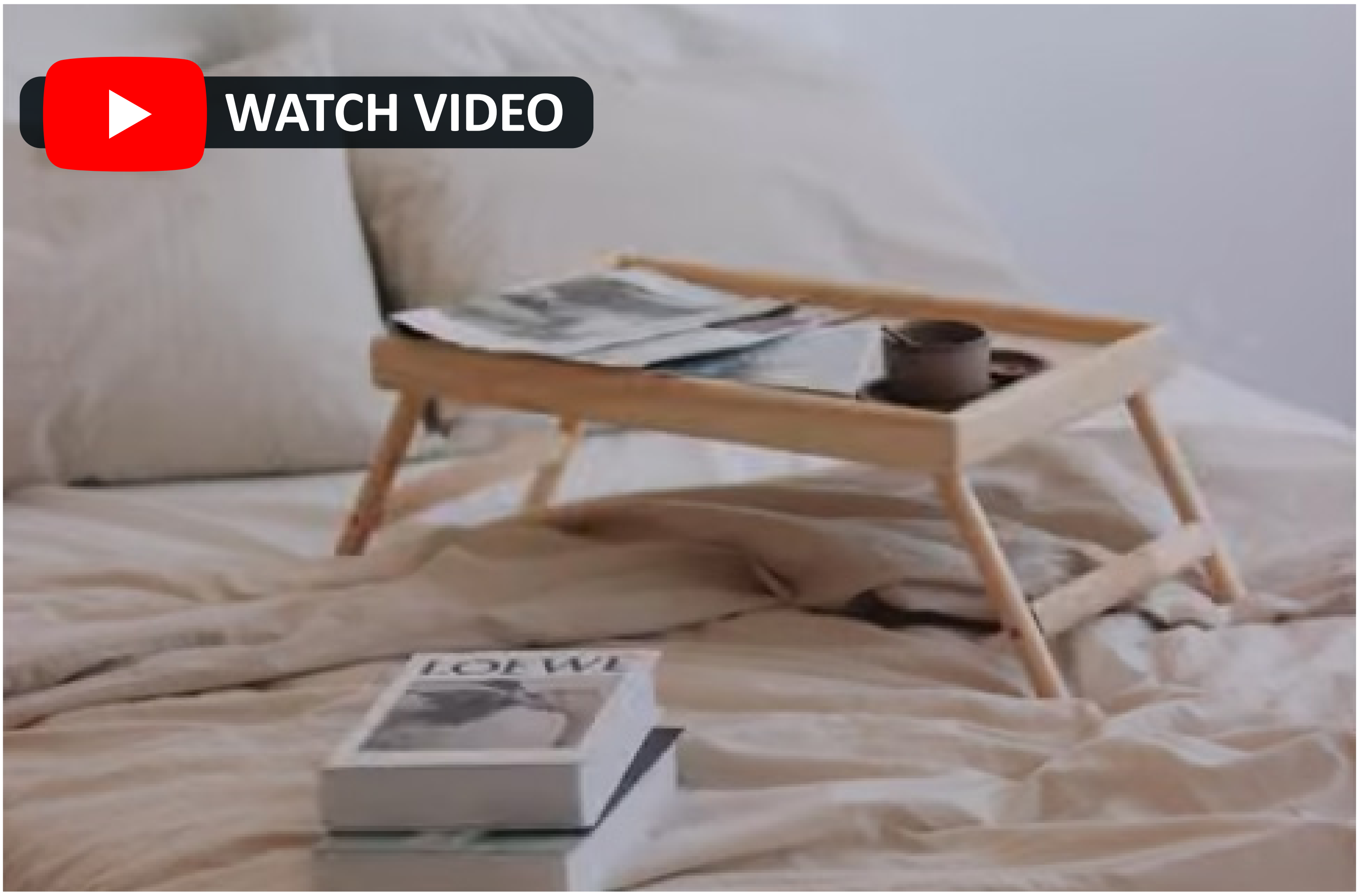
i) image-wide receptive field

ii) preserving the low frequencies of the spectrum after scale change, i.e. scale equivariance of 1×1 convolutions in the frequency domain.

Showcase of results

For this month we will present the results in a gallery format, as we were lucky the authors provided us with an amazing resource website. Feel free to browser, explore and try to explain based on the theory explained previously!

LaMa generalizes with super high resolutions (that includes a resolution more/around 2k). This allows to create examples such as shown in the short videos here:



You can try yourself this technology on this cool third-party service that in no way relates to Samsung: https://cleanup.pictures/

The specific results of the Big Lama (51M) are shown below for positive examples:

And here's a result from the domain transfer:

Interesting, isn't it?

Overall, in this work, the use of a simple, single-stage approach for large-mask inpainting was investigated. Such an approach proved to be very competitive and able to push the state of the art, as this method was arguably good in generating repetitive visual structures.

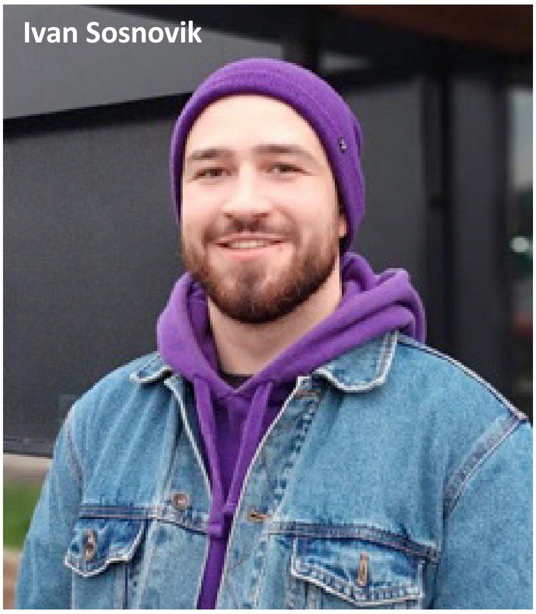
A negative aspect of the LaMa architecture is that it may struggle when a strong perspective distortion gets involved, which if you are interested you can see in the supplementary material of the paper!

Wishes for the new year...

We'll meet each other with the new year! Meanwhile, be curious and let's always connect and discuss the latest news.

DISCO: ACCURATE DISCRETE SCALE CONVOLUTIONS


Artem Moskalev

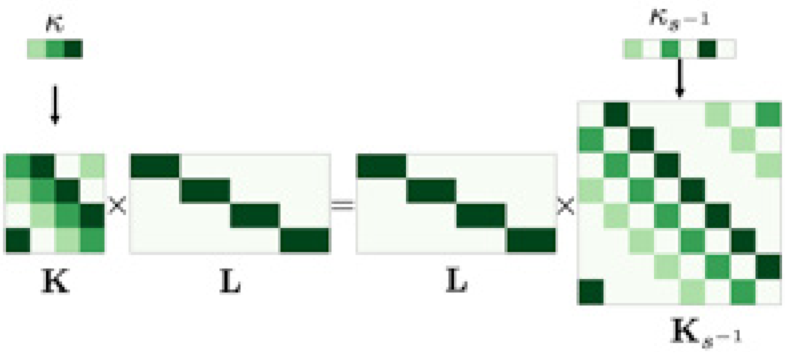

Ivan Sosnovik

Ivan Sosnovik and Artem Moskalev are PhD students at the University of Amsterdam in the UvA Bosch Delta Lab, under the supervision of Professor Arnold Smeulders. Last month, they won the Best Paper award at the British Machine Vision Conference (BMVC) for DISCO, their innovative new approach to accurate scale-equivariant convolutional neural networks. They are here to tell us more about it.

We know objects in images and video vary in size if the distance between the object and the camera changes. When watching a football game, for example, we assume the distance between us and the players increases or decreases as they move away from or towards the camera.

"***Convolutional neural networks** are currently the state-of-the-art models for responding to these size differences,"* Ivan tells us. *"They are so good because of one very important feature: **convolutions are equivariant to translations**. In other words, a translation of the input signal leads to a proportional translation of the output signal."*

Ivan and Artem want to take this one step further by equipping a CNN with an extra quantity – **scale**. They want to achieve a model for which a translation and scale transformation of the input will lead to a proportional translation and scale transformation of the output to analyze images and all their rescaled copies in the same way.



Other papers have tried different methods to build such networks, but DISCO presents a theory which is generalized and at the same time shows us there is a missing element in this field of CNNs for scale equivariance. This missing element was a set of kernels, a class of convolutions, which are very accurate and at the same time very flexible in the sense that you can choose any set of scales and it will be the most accurate model.

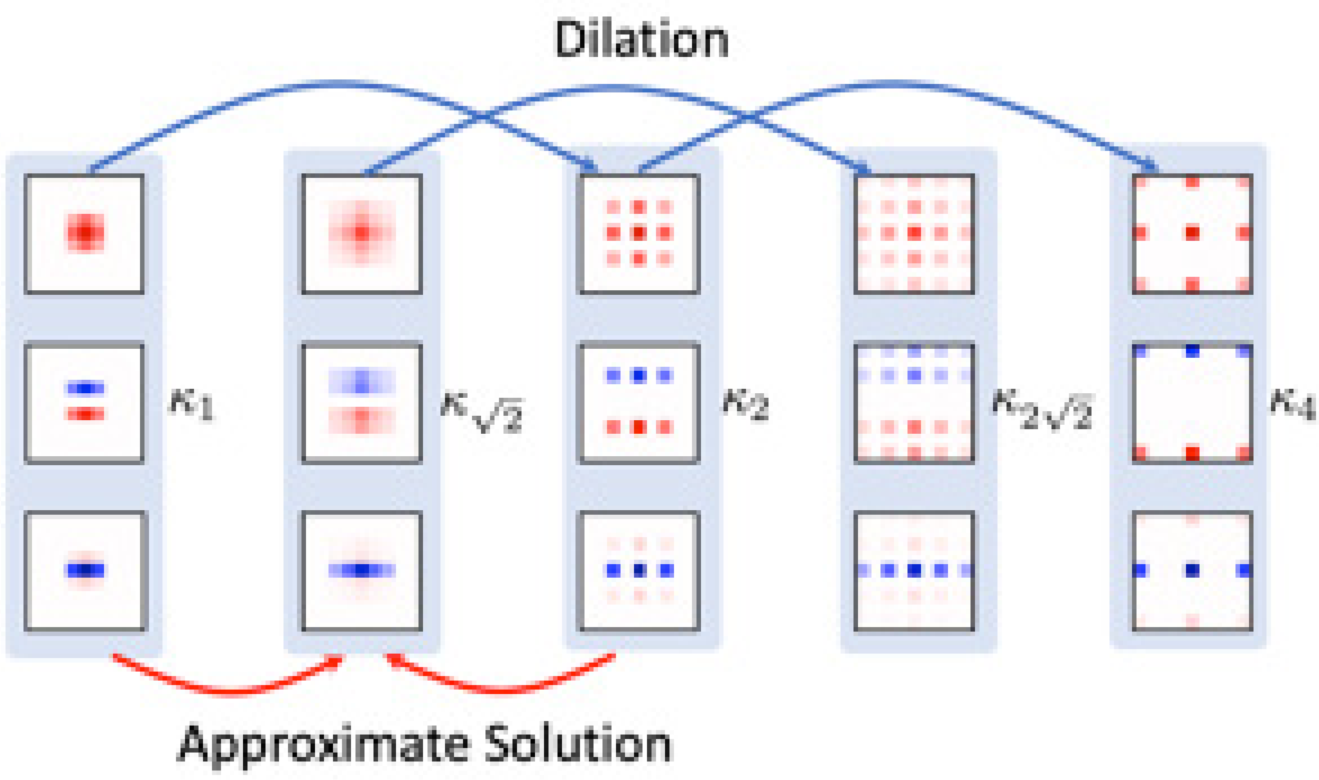
"DISCO uses a different type of kernels to other models," Ivan explains. *"These kernels are sparse. A big fraction of elements in these kernels are zeros. With PyTorch or similar frameworks, where there are no sparse convolutions, it's not possible to implement these kernels to get the best out of the theory. We demonstrated the computational complexity of the network grows linearly with the increasing number of scales, but that's not true for PyTorch because it doesn't calculate sparse convolutions."*

One of the practical applications for this work is in **autonomous vehicles**. If an autopilot can identify every car model on the market, it must also be accurate at varying their size. It cannot be confused because a car is 500 meters away. This method is highly accurate at classifying and tracking objects, so in this example would be able to follow the cars highly accurately and significantly faster than all previous models with the same features. The system could also use those cars to deduce the geometry of the road the vehicle is driving on, which may not be the main focus of the autopilot but supports many auxiliary tasks.

*"In this work, we designed a **scale-equivariant network**, and we show lower equivariant error leads to better accuracy, which means because we add additional structure to our neural network, it's able to generalize better,"* Artem explains. *"It needs less data to generalize, so it's useful in many areas where the transformation of the data is not limited to the translation group."*

Ivan is no stranger to awards. We last spoke to him in 2017 when he had just won the **Kaggle Leaf Classification Competition**. But it takes an extra special paper to take home a Best Paper award. What made DISCO stand out to the judges?

"That's a very good question and one I've definitely thought about because I would like to repeat this success!" he laughs. *"I think it was more about the set of experiments we did. We demonstrated*



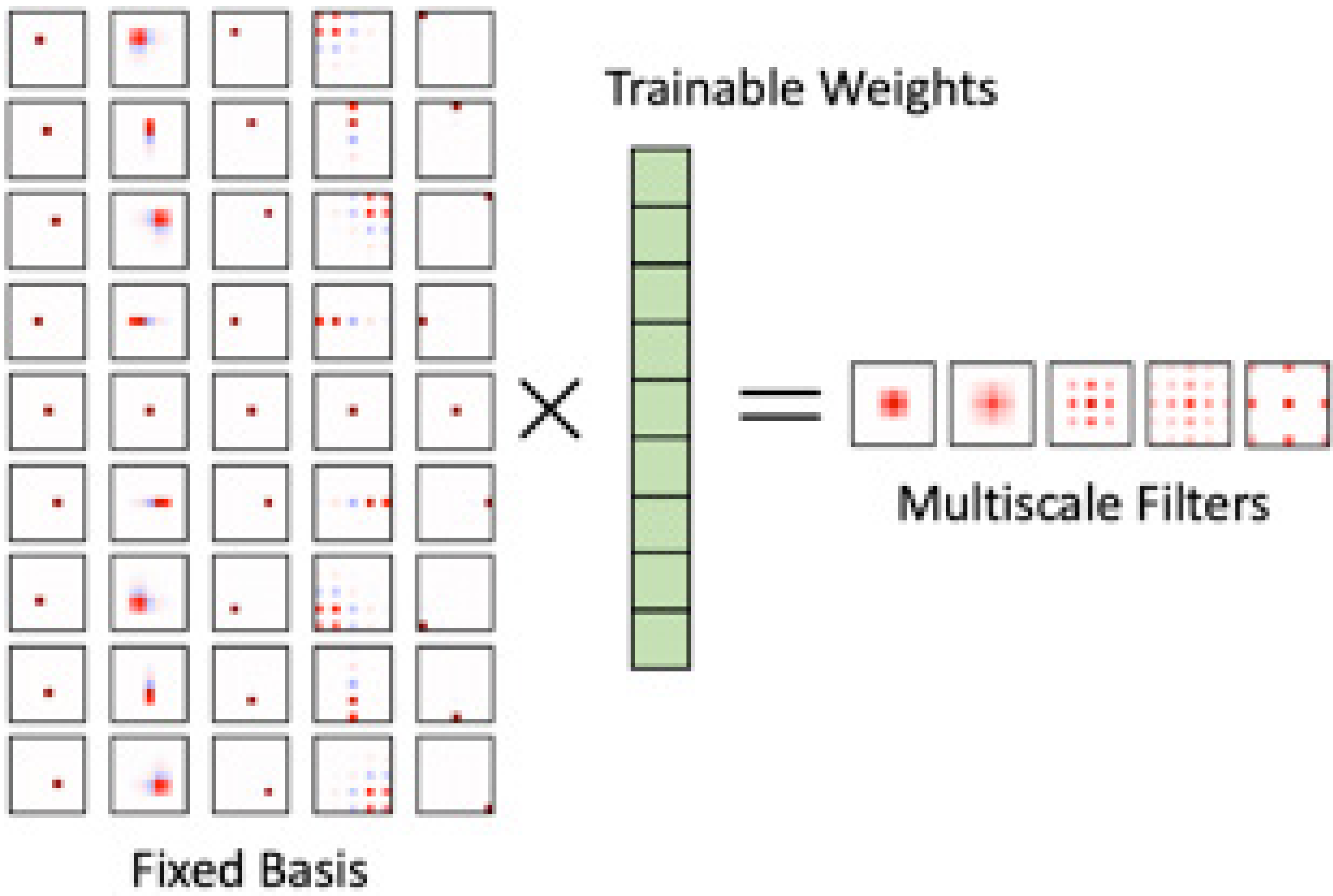
that as soon as you have a task, you can ask yourself a couple of questions and if the answers to those questions are yes, our model will significantly improve the quality of your technology. You can use it as a classifier, or as a tracker if you want to make it faster, or for geometry. You can use it where you have a lack of labeled data, or a lack of data in general, and you want to generalize to the data you do have. We want our models to be universal."

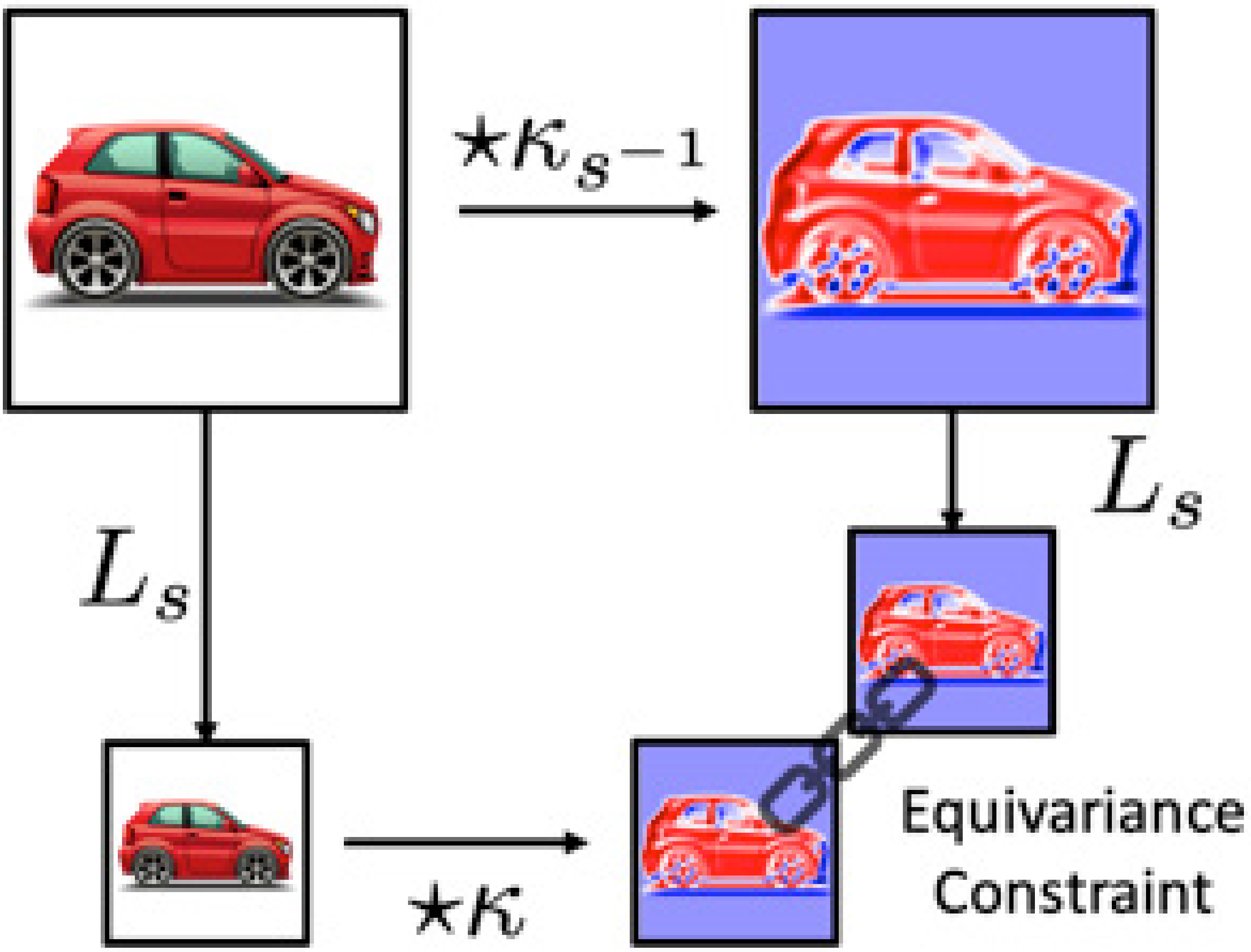
Artem adds:

"We started with a problem that no one was paying much attention to, and wasn't described very well, and we showed it's very important. If you solve this problem, then you have better performance in a wide variety of tasks."

What features would they add to the model if they could? Being a big music fan, Ivan has a fun idea.

"Signals are discretized, they're not continuous, and they're discretized in exactly the same way images are discretized," he tells us. *"But images are in two dimensions and sounds are in one dimension, and when we analyze sound, some problems arise which are not common for images. In image analysis we use small windows of three or five pixels, but for audio analysis, the windows are much wider. I would try to show if it is applicable to sound, and if so, how to use it for sound analysis and generation. When I say one word very fast, and when I say the same word slow, it's still the same word, so we want the result to be the same. For sound, the scale is the duration. We could demonstrate that you can use the same type of kernels in 1D and the same kernels in 2D and generalize it to 3D. It gives us some estimate of how to increase the number of dimensions when*





we analyze signals with this DISCO approach."

Ivan hopes the pair's next paper will present their idea in an even clearer way so that everyone will be able to understand what it is about, why it is so important, and where it is applicable – if applicable.

"It's always important to show there are limits," he points out. *"CVPR explicitly asks authors to write about that so it's clear that if these limitations are against your approach, you can still read the paper, but it will not help you. The way we structure the research is very important. The next paper we write will have a focus on good presentation from the very start."*

He would like to see the model demonstrate a **speed advantage** over previous attempts because speed is so important for fields like autonomous vehicles. The price point is also key, and he would like to show it is affordable enough for people to equip their products, such as smartphones, with the technology.

"In this work, we pay attention to the computational complexity of other methods," Artem points out. *"In modern computer vision, in an age where we have these giant models, transformers which take weeks of training to converge, I think it's very cool if we can impose some additional structure on our neural network. It allows us to train faster and to have some nice implementation, which means we also get faster inference. Speed and convergence of the networks is an important issue, and we try to tackle that in our work."*

DRAWING NEURAL NETWORKS

by Marica Muffoletto (Twitter)

Hi all! I am delighted to start yet another year with the RSIP Vision community. I wish both regular and new readers 12 months filled with happiness, joy and "more importantly" exciting computer vision content!

We begin 2022 with an article proposed by one of our own readers, discussing a hot and useful topic in the AI and Deep Learning community: how to draw neural networks. As we all know, this field is growing at light speed and diagrams of Neural Network architectures are commonly found in most computer vision papers and proudly shown everywhere in conferences.

Hence, it becomes essential to add this to our skillset: it could be used to show a schematic of your algorithm to the team, to be included in a paper, in a poster or a talk. Images of Neural Networks are sometimes even more explanatory than the corresponding code or description. They are usually one of the best methods to represent a model.

A fast, easy solution that is widely employed is PowerPoint: the Microsoft suite offers shapes, charts and icons that can be organised to mimic the common diagram, but today we are going to focus on more complex tools.

Let's start from picking a pre-defined model as an example that we will keep throughout the article and aim to draw using different tools. I chose VGG16, an architecture for classification and detection which won the ILSVR (ImageNet) competition in 2014.

Most famous deep learning libraries PyTorch & Keras have common architectures stored as classes and these can be called and printed with a summary of the layers and the output shapes.

The table below shows how to achieve this and what output is displayed for both cases.

Pytorch model

```
from torchvision import models
from torchsummary import summary

vgg = models.vgg16()
summary(vgg, (3, 224, 224))
```

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv2d-1         [-1, 64, 224, 224]           1,792
              ReLU-2         [-1, 64, 224, 224]               0
            Conv2d-3         [-1, 64, 224, 224]          36,928
              ReLU-4         [-1, 64, 224, 224]               0
         MaxPool2d-5         [-1, 64, 112, 112]               0
            Conv2d-6        [-1, 128, 112, 112]          73,856
              ReLU-7        [-1, 128, 112, 112]               0
            Conv2d-8        [-1, 128, 112, 112]         147,584
              ReLU-9        [-1, 128, 112, 112]               0
        MaxPool2d-10          [-1, 128, 56, 56]               0
           Conv2d-11          [-1, 256, 56, 56]         295,168
             ReLU-12          [-1, 256, 56, 56]               0
           Conv2d-13          [-1, 256, 56, 56]         590,080
             ReLU-14          [-1, 256, 56, 56]               0
           Conv2d-15          [-1, 256, 56, 56]         590,080
             ReLU-16          [-1, 256, 56, 56]               0
        MaxPool2d-17          [-1, 256, 28, 28]               0
           Conv2d-18          [-1, 512, 28, 28]       1,180,160
             ReLU-19          [-1, 512, 28, 28]               0
           Conv2d-20          [-1, 512, 28, 28]       2,359,808
             ReLU-21          [-1, 512, 28, 28]               0
           Conv2d-22          [-1, 512, 28, 28]       2,359,808
             ReLU-23          [-1, 512, 28, 28]               0
        MaxPool2d-24          [-1, 512, 14, 14]               0
           Conv2d-25          [-1, 512, 14, 14]       2,359,808
             ReLU-26          [-1, 512, 14, 14]               0
           Conv2d-27          [-1, 512, 14, 14]       2,359,808
             ReLU-28          [-1, 512, 14, 14]               0
           Conv2d-29          [-1, 512, 14, 14]       2,359,808
             ReLU-30          [-1, 512, 14, 14]               0
        MaxPool2d-31            [-1, 512, 7, 7]               0
AdaptiveAvgPool2d-32            [-1, 512, 7, 7]               0
           Linear-33                 [-1, 4096]     102,764,544
             ReLU-34                 [-1, 4096]               0
          Dropout-35                 [-1, 4096]               0
           Linear-36                 [-1, 4096]      16,781,312
             ReLU-37                 [-1, 4096]               0
          Dropout-38                 [-1, 4096]               0
           Linear-39                 [-1, 1000]       4,097,000
================================================================
Total params: 138,357,544
Trainable params: 138,357,544
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 0.57
Forward/backward pass size (MB): 218.78
Params size (MB): 527.79
Estimated Total Size (MB): 747.15
----------------------------------------------------------------
```

Keras model

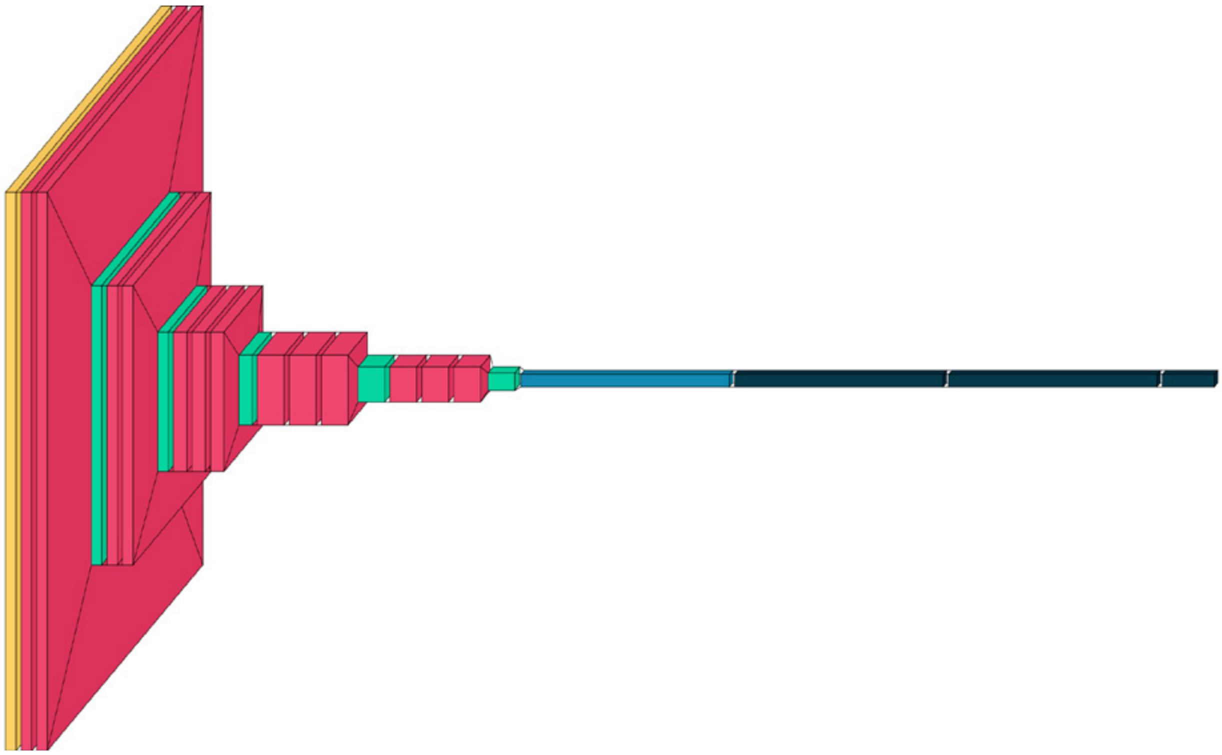
```
import keras
from keras.applications import vgg16

vgg = vgg16.VGG16(weights="imagenet")
print(vgg.summary())
```

```
Model: "vgg16"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_3 (InputLayer)         [(None, 224, 224, 3)]     0

block1_conv1 (Conv2D)        (None, 224, 224, 64)      1792

block1_conv2 (Conv2D)        (None, 224, 224, 64)      36928

block1_pool (MaxPooling2D)   (None, 112, 112, 64)      0

block2_conv1 (Conv2D)        (None, 112, 112, 128)     73856

block2_conv2 (Conv2D)        (None, 112, 112, 128)     147584

block2_pool (MaxPooling2D)   (None, 56, 56, 128)       0

block3_conv1 (Conv2D)        (None, 56, 56, 256)       295168

block3_conv2 (Conv2D)        (None, 56, 56, 256)       590080

block3_conv3 (Conv2D)        (None, 56, 56, 256)       590080

block3_pool (MaxPooling2D)   (None, 28, 28, 256)       0

block4_conv1 (Conv2D)        (None, 28, 28, 512)       1180160

block4_conv2 (Conv2D)        (None, 28, 28, 512)       2359808

block4_conv3 (Conv2D)        (None, 28, 28, 512)       2359808

block4_pool (MaxPooling2D)   (None, 14, 14, 512)       0

block5_conv1 (Conv2D)        (None, 14, 14, 512)       2359808

block5_conv2 (Conv2D)        (None, 14, 14, 512)       2359808

block5_conv3 (Conv2D)        (None, 14, 14, 512)       2359808

block5_pool (MaxPooling2D)   (None, 7, 7, 512)         0

flatten (Flatten)            (None, 25088)             0

fc1 (Dense)                  (None, 4096)              102764544

fc2 (Dense)                  (None, 4096)              16781312

predictions (Dense)          (None, 1000)              4097000

=================================================================
Total params: 138,357,544
Trainable params: 138,357,544
Non-trainable params: 0
_________________________________________________________________
None
```

Visual Keras

Probably the fastest tool to draw a Keras model. Just add these lines to your code to employ the Visualkeras package, and automatically create and save a diagram on your disk with the specified filename.

```
pip install visualkeras
import visualkeras
visualkeras.layered_view(vgg)
visualkeras.layered_view(vgg, to_file='vgg_diagram.png').show() # write and show
```



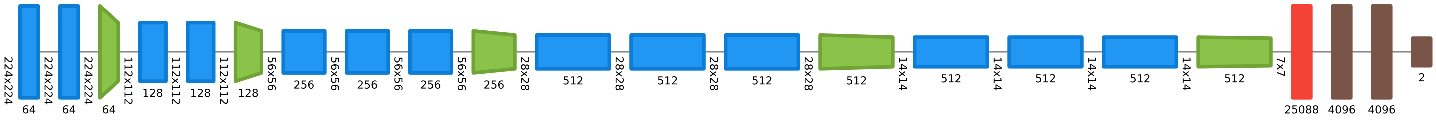
Net2Vis

This is a web tool that allows to upload a Keras model or directly copy paste the code on the left panel. The script needs to contain a get_model function which returns the model to draw. Below, we wrote such a function for the vgg16 architecture.

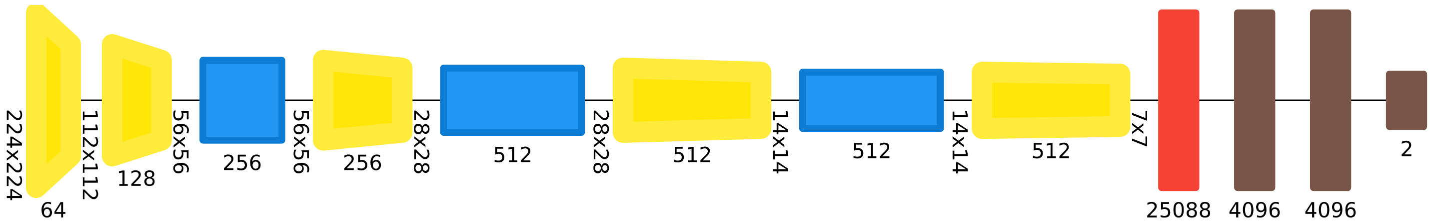
```
# You can freely modify this file.
# However, you need to have a function that is named get_model and returns a
Keras Model.
import keras,os
from keras.models import Sequential
from keras.layers import Dense, Conv2D, MaxPool2D , Flatten
from keras.preprocessing.image import ImageDataGenerator
import numpy as np
```

```
def get_model():

    model = Sequential()

    model.add(Conv2D(input_shape=(224,224,3),filters = 64, kernel_size=(3,3),padding=”same”, activation=”relu”))
    model.add(Conv2D(filters=64,kernel_size=(3,3),padding=”same”, activation=”relu”))
    model.add(MaxPool2D(pool_size=(2,2),strides=(2,2)))
    model.add(Conv2D(filters=128, kernel_size=(3,3), padding=”same”, activation=”relu”))
    model.add(Conv2D(filters=128, kernel_size=(3,3), padding=”same”, activation=”relu”))
    model.add(MaxPool2D(pool_size=(2,2),strides=(2,2)))
    model.add(Conv2D(filters=256, kernel_size=(3,3), padding=”same”, activation=”relu”))
    model.add(Conv2D(filters=256, kernel_size=(3,3), padding=”same”, activation=”relu”))
    model.add(Conv2D(filters=256, kernel_size=(3,3), padding=”same”, activation=”relu”))
    model.add(MaxPool2D(pool_size=(2,2),strides=(2,2)))
    model.add(Conv2D(filters=512, kernel_size=(3,3), padding=”same”, activation=”relu”))
    model.add(Conv2D(filters=512, kernel_size=(3,3), padding=”same”, activation=”relu”))
    model.add(Conv2D(filters=512, kernel_size=(3,3), padding=”same”, activation=”relu”))
    model.add(MaxPool2D(pool_size=(2,2),strides=(2,2)))
    model.add(Conv2D(filters=512, kernel_size=(3,3), padding=”same”, activation=”relu”))
    model.add(Conv2D(filters=512, kernel_size=(3,3), padding=”same”, activation=”relu”))
    model.add(Conv2D(filters=512, kernel_size=(3,3), padding=”same”, activation=”relu”))
    model.add(MaxPool2D(pool_size=(2,2),strides=(2,2)))
    model.add(Flatten())
    model.add(Dense(units=4096,activation=”relu”))
    model.add(Dense(units=4096,activation=”relu”))
    model.add(Dense(units=2, activation=”softmax”))

    return model
```

The following diagram is produced, with legend blue = conv2D, green = maxpooling2D, red = flatten, brown = dense.



With the option "Automatically Group", the program draws a new diagram where some operations are grouped.



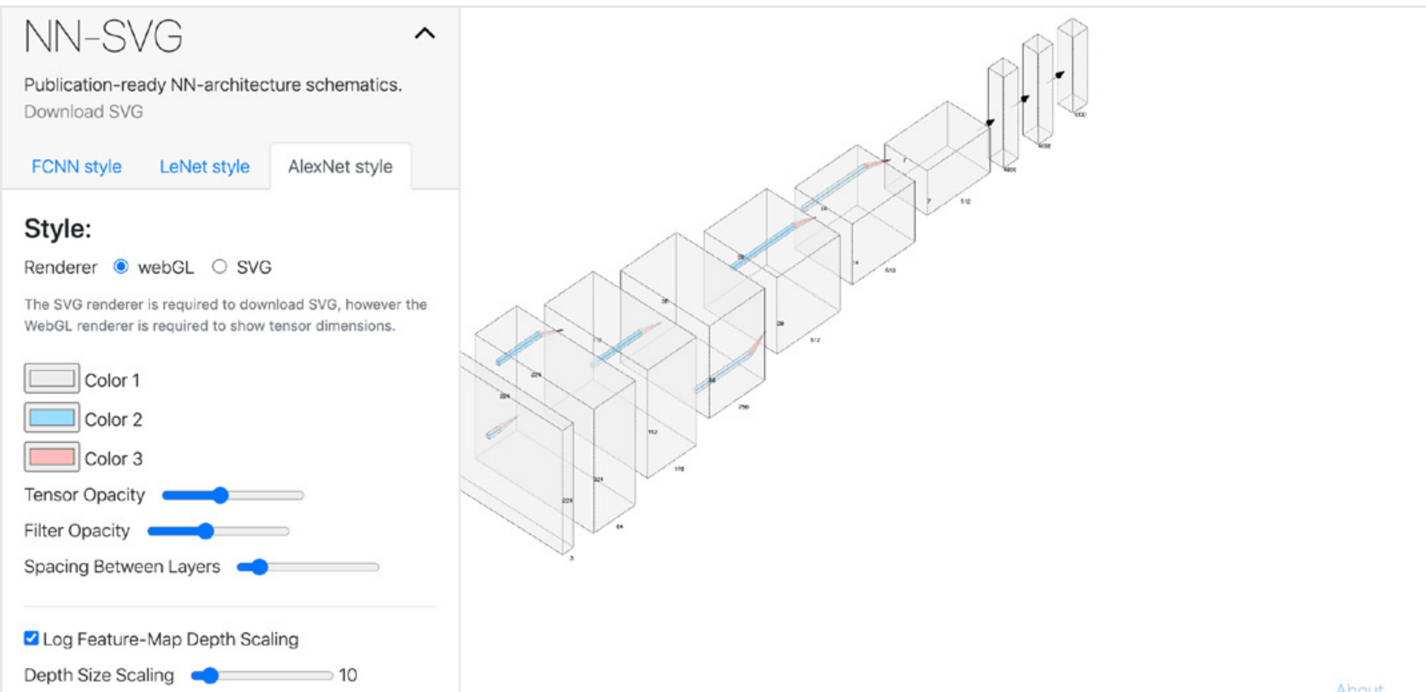
This is the corresponding legend:

You can download a zipped folder containing both graph and legend in pdf and svg formats.

Alexnail

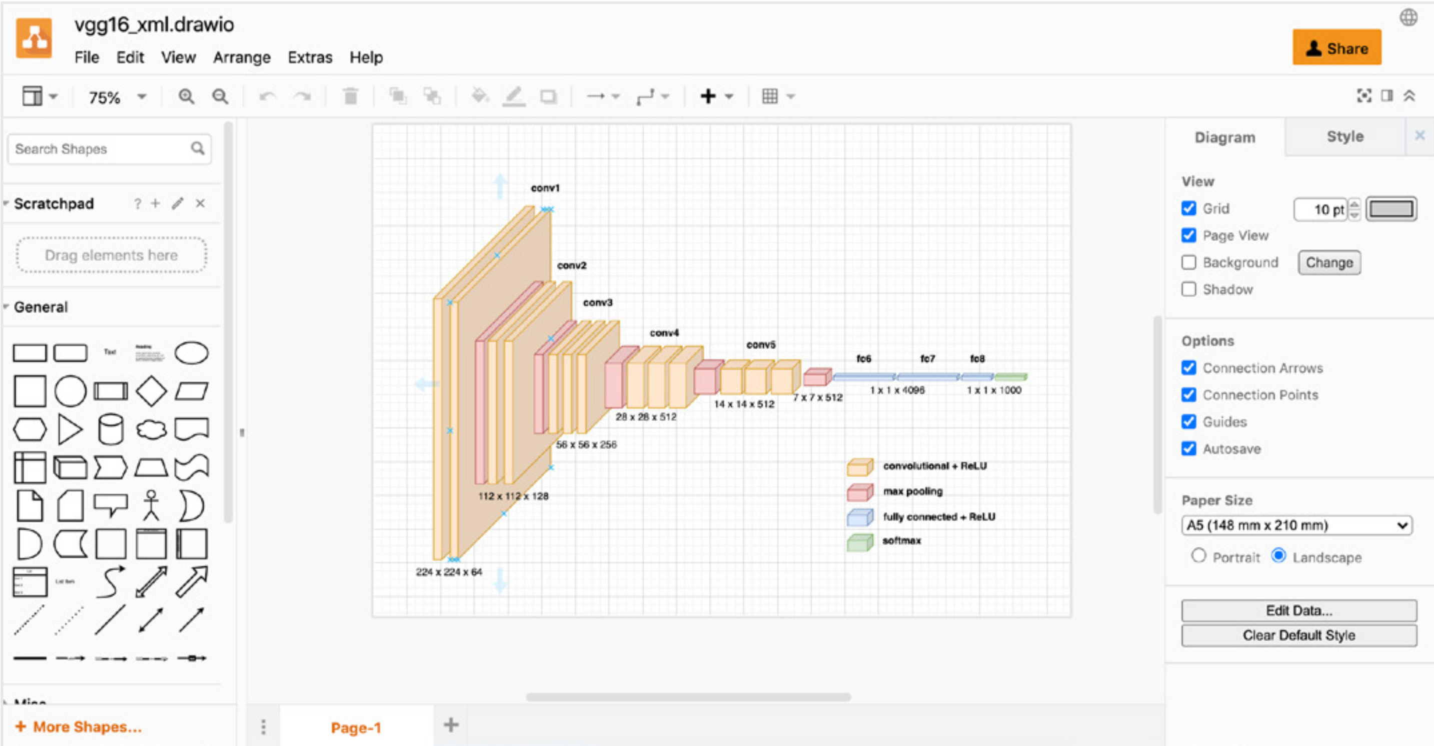
This is a web tool that includes three diagram styles: FCNN, LeNet and AlexNet. In the browser, one can select the preferred style and add layers defining height, width, depth, and filter shape.

We built a diagram for VGG16 following the output shapes in the summary. In this representation, we added a layer for every convolution block (first 2 Conv2D + Max Pooling, then 3 Conv2D + Max Pooling). This was an attempt to group the layers, following the AlexNet style, and make a compact representation. Anyways, if interested in the topic, you should explore all available styles & features. When your project is finished, you can download it in SVG format.



Diagrams.net

Another web tool that looks like an advanced version of PowerPoint for this task, and a more flexible option than Alexnail to draw your very own diagram. Here, you can use the proposed shapes, dynamically change their dimensions, add labels, and choose colors. Definitely worth trying if you want to create a personalized diagram, but it might take longer than the previous options.



PlotNeuralNet

This is the last tool that we will employ to draw our vgg model. It is also a very peculiar one because it allows to create latex code that can be directly added to your project to draw an image of the CNN and insert it in your paper. First, you will need to download and install the required packages on your machine. Instructions for Ubuntu system are the following:

```
git clone https://github.com/HarisIqbal88/PlotNeuralNet.git
sudo apt-get install texlive-latex-base
sudo apt-get install texlive-fonts-recommended
sudo apt-get install texlive-fonts-extra
sudo apt-get install texlive-latex-extra

cd PlotNeuralNet
mkdir my_project
cd my_project
vim my_arch.py
```

Once opened your Python script on terminal or on your favorite editor, you can define the vgg16 architecture as follows.

```
import sys
sys.path.append('../')
from pycore.tikzeng import *

# defined your arch
arch = [
    to_head( '..' ),
    to_cor(),
    to_begin(),

    #conv1
    to_ConvConvRelu( name='cr1', s_filer=64, n_filer=(224,224), offset="(0,0,0)", to="(0,0,0)", width=(2,2), height=40, depth=40, caption="conv1"),
    to_Pool(name="p1", offset="(0,0,0)", to="(cr1-east)", width=1, height=35, depth=35, opacity=0.5),

    #conv2
    to_ConvConvRelu( name='cr2', s_filer=128, n_filer=(112,112), offset="(2,0,0)", to="(p1-east)", width=(4,4), height=35, depth=35, caption="conv2"),
    to_Pool(name="p2", offset="(0,0,0)", to="(cr2-east)", width=1, height=30, depth=30, opacity=0.5),

    #conv3
    to_ConvConvRelu( name='cr3', s_filer=256, n_filer=("56","56","56"), offset="(2,0,0)", to="(p2-east)", width=(4,4,4), height=30, depth=30, caption="conv3"),
    to_Conv("cr3-bis", 256, 56, offset="(0,0,0)", to="(cr3-east)", height=30, depth=30, width=4),
    to_Pool(name="p3", offset="(0,0,0)", to="(cr3-bis-east)", width=1, height=23, depth=23, opacity=0.5),

    #conv4
    to_ConvConvRelu( name='cr4', s_filer=512, n_filer=("28","28","28"), offset="(2,0,0)", to="(p3-east)", width=(4,4,4), height=25, depth=25, caption="conv4"),
    to_Conv("cr4-bis", 512, 28, offset="(0,0,0)", to="(cr4-east)", height=25, depth=25, width=4),
    to_Pool(name="p4", offset="(0,0,0)", to="(cr4-bis-east)", width=1, height=21, depth=21, opacity=0.5),
```

```
    #conv5
    to_ConvConvRelu( name='cr5', s_filer=512, n_filer=("14","14","14"), off-
set="(2,0,0)", to="(p4-east)", width=(4,4,4), height=21, depth=21, caption="-
conv5"),
    to_Conv("cr5-bis", 512, 14, offset="(0,0,0)", to="(cr5-east)", height=21,
depth=21, width=4),
    to_Pool(name="p5", offset="(0,0,0)", to="(cr5-bis-east)", width=1, height=10,
depth=10, opacity=0.5),

    #fc1
    to_FullyConnected(name="fc1", s_filer=4096, offset="(3,0,0)", to="(p5-east)",
width=1, height=1, depth=100, caption="fc6" ),

    #fc2
    to_FullyConnected(name="fc2", s_filer=4096, offset="(1.25,0,0)", to="(fc1-
east)", width=1, height=1, depth=100, caption="fc7" ),

    #fc3
    to_FullyConnected(name="fc3", s_filer=1000, offset="(1.25,0,0)", to="(fc2-
east)", width=1, height=1, depth=30, caption="fc8+softmax" ),
    to_SoftMax("softmax", s_filer=1000, offset="(0,0,0)", to="(fc3-east)",
width=1, height=1, depth=30, opacity=0.8, caption="" ),

    to_connection("p1", "cr2"),
    to_connection("p2", "cr3"),
    to_connection("p3", "cr4"),
    to_connection("p4", "cr5"),
    to_connection( "p5", "fc1"),
    to_connection( "fc1", "fc2"),
    to_connection( "fc2", "fc3"),
    to_end()
    ]

def main():
    namefile = str(sys.argv[0]).split('.')[0]
    to_generate(arch, namefile + '.tex' )

if __name__ == '__main__':
    main()
```

To re-create vgg16, we needed to add the definition of a Fully Connected layer in the file pycore/tikzeng.py. This file writes up a drawing for each layer in latex code. If layers for your chosen architecture are missing here, you will need to create them yourself. This can be quite time consuming, but I believe the result is worth it. You can judge by yourselves on the diagram below.

```
# Fully Connected
def to_FullyConnected( name, s_filer=” “, n_filer=” “, offset=”(0,0,0)”,
to=”(0,0,0)”, width=1.5, height=3, depth=25, opacity=0.8, caption=” “):
    return r”””
\pic[shift={“””+ offset +”””}] at “””+ to +”””
    {Box={
        name=””” + name +”””,
        caption=””” +caption + “””,
        xlabel={{ “””+ ‘”’+str(n_filer) +’”, “dummy”’+ “”” }},
        zlabel=”””+ str(s_filer) +”””,
        fill=\FcColor,
        opacity=”””+ str(opacity) +”””,
        height=”””+ str(height) +”””,
        width=”””+ str(width) +”””,
        depth=”””+ str(depth) +”””
        }
    };
“””
```

Finally, run the script with:

```
bash ../tikzmake.sh my_arch
```

Once you run this, a pdf file containing the image of your network will be automatically saved. If you also need to store the .tex version, for use on Overleaf or another latex code, please make sure you go to ../tikxmake.sh and comment the line “rm *.tex”.

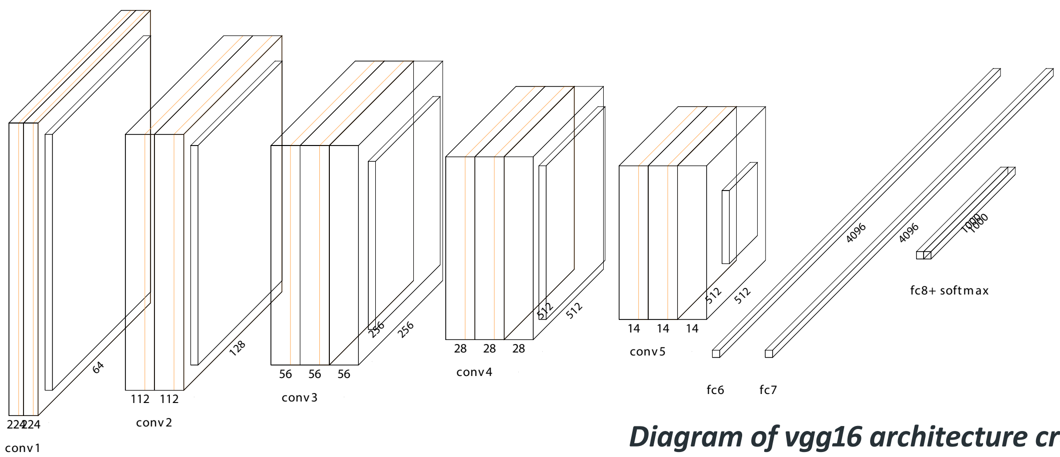


Diagram of vgg16 architecture created using PlotNeuralNet

Wrap-up

In this article, we described some of the most common and efficient tools for creating a diagram for a CNN architecture. We are also aware of many other pieces of software out there which could be interesting and better suited for different users. The following list includes some of them.

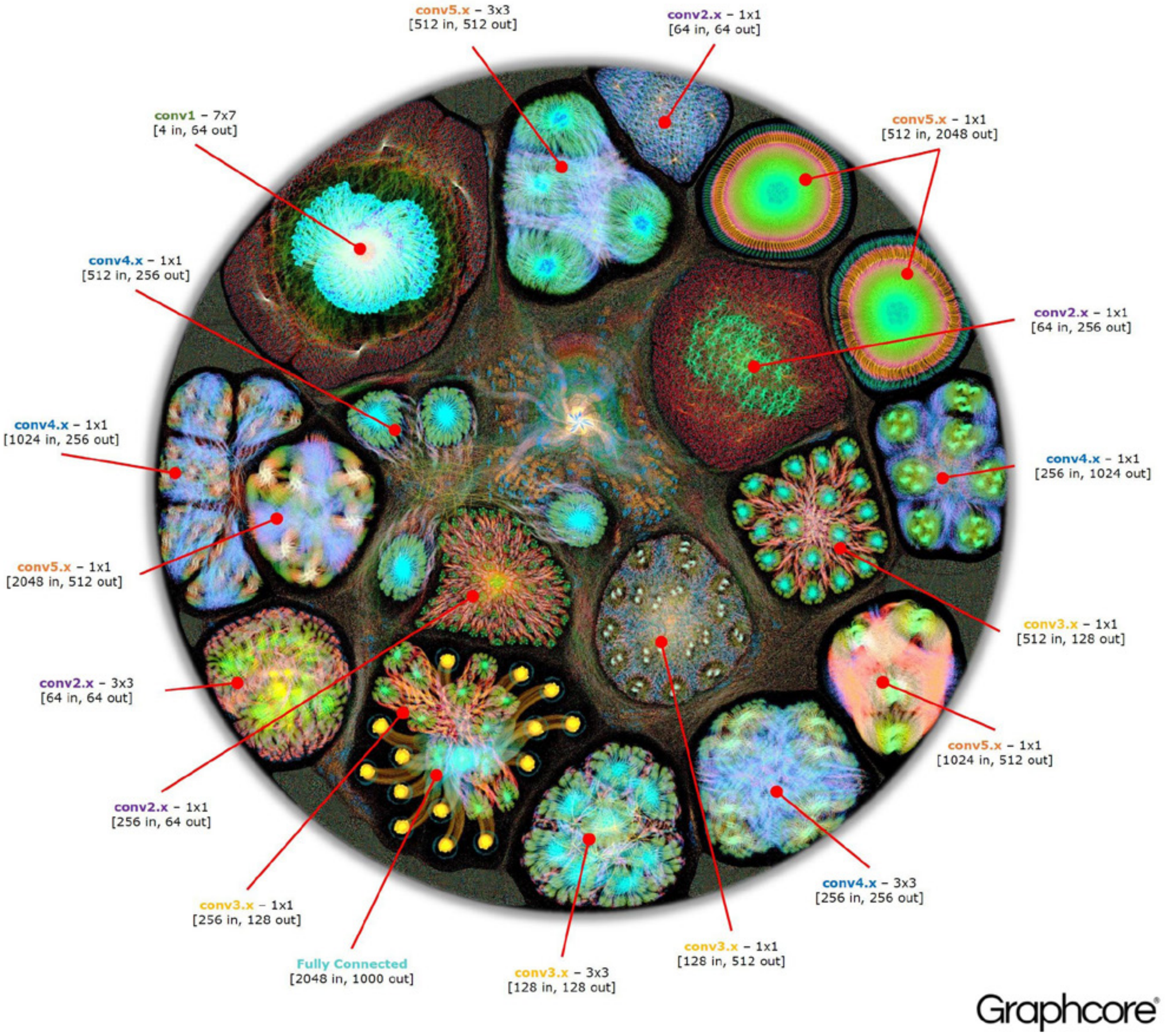
Netron: Takes a model stored in .h5 format and visualizes all layers and parameters.

Conx: Python package that creates dynamic, rendered visualizations with the function net.picture().

Netscope: an everyday tool for Caffe models.

Graphcore: a software for unique visualisation of CNNs, which is related to the natural meaning of AI. The architecture is in fact described through biological cells, neurons, synapses and so on. Though the need for such graphs might be far less, the images themselves are amazing. Have a look below at the Graphcore representation of ResNet50.

We are thankful to the reader who inspired this article, creating the opportunity to explore a handy topic, and as always, we welcome any feedback and suggestion for the future.



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

TOP 10 Robotic and Artificial Intelligence Stories of 2021

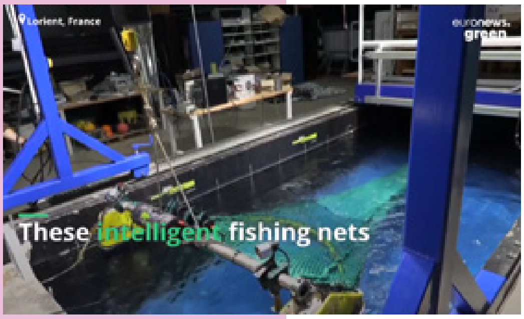
Yes, it's that time of the year when media dig into the past year to remind their top stories. We also do it this month on page 28. We liked very much this collection of top robot stories of 2021 by designboom.com and we think you will like them too. The most impressive ones? Probably the self-driving bike, the world's first ever robot remote tattoo (nicknamed 'The Impossible Tattoo', performed through T-Mobile 5G on Dutch actress Stijn Fransen - watch the video above) and Grace, the robot nurse 'built to care', who lends a caring hand for the elderly and those isolated with COVID-19. Read More

The Last Major Obstacle to Cracking the Code on Self-Driving Cars

Companies developing **autonomous driving solutions** are making different bets: some (like **Tesla**) are betting on the minimalist approach, using just visual data from **onboard cameras**; others (like **MobilEye**) plan to develop their own full self-driving car software by using **every sensor available**: cameras, radar, and LiDAR. Most are in between. If this is true, what's missing for MobilEye to be full ready? Apparently, a detailed map of every road a car might ever encounter, that means every single road on Earth. In the meantime, watch the lovely video with the launch of their **robotaxi fleet in Paris!** Read More

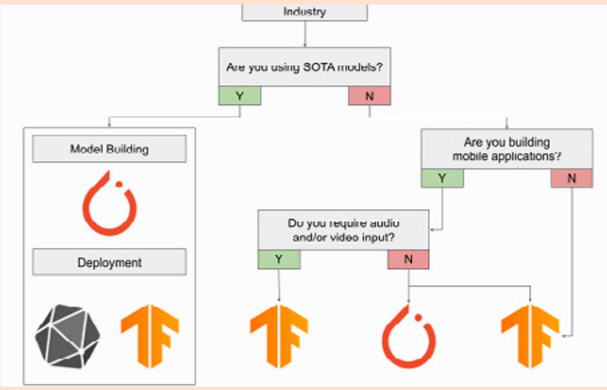
The 'Game of Trawls': Smart Fishing Nets Could Save Millions of Sea Creatures

A nice marine tech seeking to replace **the bottom trawl**, a large and heavy net - often the size of a few football pitches - which is dragged along the seafloor, capturing everything it encounters. Besides some tasty fish that we finally eat, 20 million tons of fish - about a quarter of global marine catches - are discarded each year and not made any use of, sealing the bitter end of too many marine species. French oceanographic institution **Ifremer** is testing smart trawls that - with the help of **AI and Computer Vision** - sort fish on the bottom, catching only the selected species the fishermen want. **Watch the Video**.



Machines That See the World More Like Humans Do

Computer vision systems sometimes make inferences about a scene that fly in the face of common sense. This is particularly problematic in systems requiring the highest accuracy, like **autonomous driving**. To overcome these errors, **MIT researchers** have developed a framework using probabilistic programming that helps machines see the world more like humans do. Their new AI learns to perceive real-world objects from just a few images, and perceives scenes in terms of these learned objects, including apparently new concepts like **"uncertainty"** and rules that don't hold. [Read More]



PyTorch vs TensorFlow in 2022

Should you use **PyTorch** vs **TensorFlow** in 2022? Spoiler alert: answers ahead! As you probably expect, the PyTorch vs TensorFlow debate does not have a single correct answer, **it all depends on what you want to do**. If you perform Deep Learning engineering in an **industry** setting, you're likely using TensorFlow and should probably stick with it. If you're a **researcher**, you're almost certainly using PyTorch, and for now you should likely stick with it: PyTorch is the de facto research framework. If you're looking to make a **career change**, either PyTorch or TensorFlow is a good option.

[Read More]

AI Gone Rogue: 6 Times AI Went Too Far

We all know a few **AI horror stories**. Some of them were disclosed here on Computer Vision News. Here is a short list, it's a nice reading. There might be a couple that you didn't know about, like when **Sophia** was asked whether she wanted to destroy humans and replied: **"OK. I will destroy humans!"**. Or like when **two virtual assistant robots** quickly turned to throwing insults and threats at each other. Some stories you already know, but it's always interesting to know how they ended. The article concludes by asking: **Rogue AI - What's Our Best Defense?** Don't expect too much of an answer. [Read More]

Congrats, Doctor!

Fabian Mentzer defended his PhD in January 2021 at ETH Zurich and was awarded the "ETH medal for outstanding doctoral theses" for his work. He now works as a Research scientist at Google, focusing on neural compression. This is also the topic he worked on during his PhD, his thesis was on lossy and lossless image compression with neural networks. Congrats, Doctor Fabian!

When (lossily) compressing images, the goal is to store an input image in a file so that on one hand, the file is small, and on the other hand that when we open the file**, the reconstructed image looks as close as possible to the input image.** There is a trade-off here: we can always transmit more information, thereby making the file larger, and the reconstruction becomes better. Inversely, bits can be saved by ignoring part of the image – something everyone that has ever seen a blocky JPEG image on the internet is familiar with.

To solve this problem with **neural networks**, there are a few challenges, one interesting one is described next. On a high level, we want to formalize the above trade-off, and we want to tell the optimization algorithm which parts of the image can safely be ignored to save bits. Intuitively, we are looking for **a loss that is aware of how humans perceive images**: Humans do not like blocky images and can tell that those are compression artifacts very easily. Same goes for blurry images. However, if the changes that happen due to compression are more semantic, they are hard to notice. For example, when compressing an image of a landscape, we do not have to store exactly how each blade of grass is oriented, it's fine to get the overall amount and type and color of grass right.

The idea underlying this argument is the **"rate-distortion-realism" trade-off**. "Rate" refers to the bitrate, how many bits do we need to store an image, "distortion" refers to how close is the reconstruction to the input on a per-pixel level (for example, we want grass blades in the reconstruction where there were grass blades in the input, not a house), and "realism" refers to how realistic an image looks: can a human tell that the image is compressed, or does it look like a "real" image taken with a camera.

This is visualized in the figure, where we see an input image, a reconstruction from a model trained only for rate and distortion (r + λd), for rate and realism (r + βR) or all three. The last picture looks close to the input and is realistic, but if you look closely, you see that not all blades are exactly where they should be.

That image is a reconstruction from "**High-Fidelity Generative Image Compression**", a paper where **the above trade-off was optimized using GANs, to achieve 2x lower rates than the previous best algorithm** (**see here for a demo**). Other publications from the thesis **are available here**.

Here are the most popular articles from Computer Vision News throughout the year 2021.

Our magazine had more than 1.1 million pageviews in 2021!

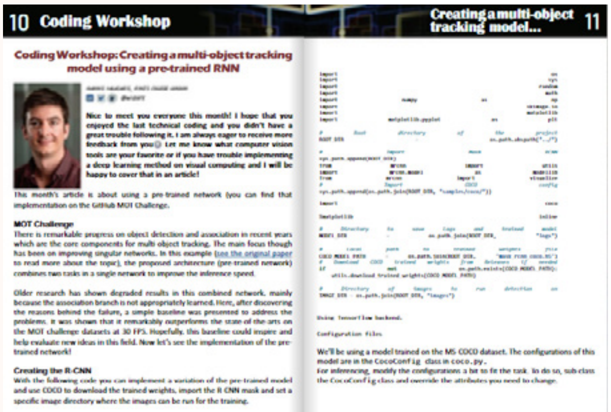
Women in Science
“I Am More of a Brain-folding Geek!”
Katja Heuer

Do you remember this young woman holding a glass jar with a real monkey brain in it? If you don’t, that means that you missed an amazing interview with **Katja Heuer**, an awesome scientist who uses computational neuroanatomy and phylogenetic comparative methods to study the evolution of the brain. So, if you thought that she was only holding a cup of coffee, you need to read that article too.

[Read It Here]

Coding Workshop: Creating a multi-object tracking model using a pre-trained RNN - with code!

Our **Ioannis Valasakis** demonstrates a pre-trained network that combines two tasks in a single network to improve the inference speed. Here, a simple baseline is presented (with code!) to address the problems and it remarkably outperforms the state-of-the-arts on the MOT challenge datasets at 30 FPS. We hope that this baseline has inspired you and helped you evaluate new ideas. Read It Here

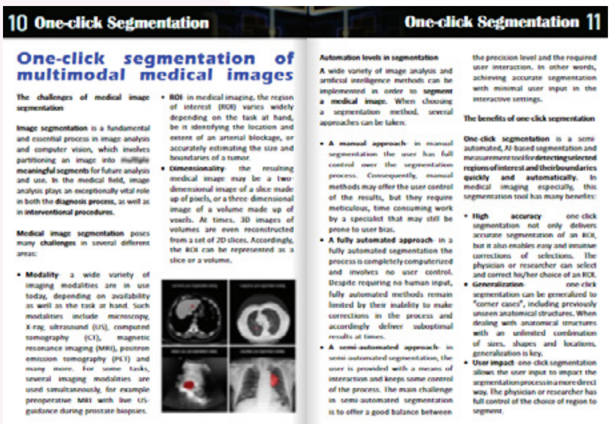


"We are actually going to change the world!" with Raquel Urtasun of Waabi

Awesome **Raquel Urtasun** has recently been included in the Entrepreneur's 100 Women of Impact in 2021, and she has been starring for years now in our own inspiring group of over 100 Women in Computer Vision. Her newly founded company **Waabi** just raised a $100 million Series A initial funding. Raquel is set to revolutionize the world as we know it and she told us how she and her team will do it. Read It Here

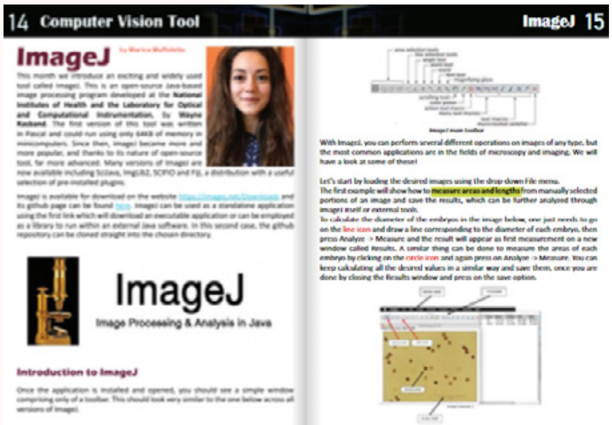
One-click Segmentation of Multi-Modal Medical Images by RSIP Vision

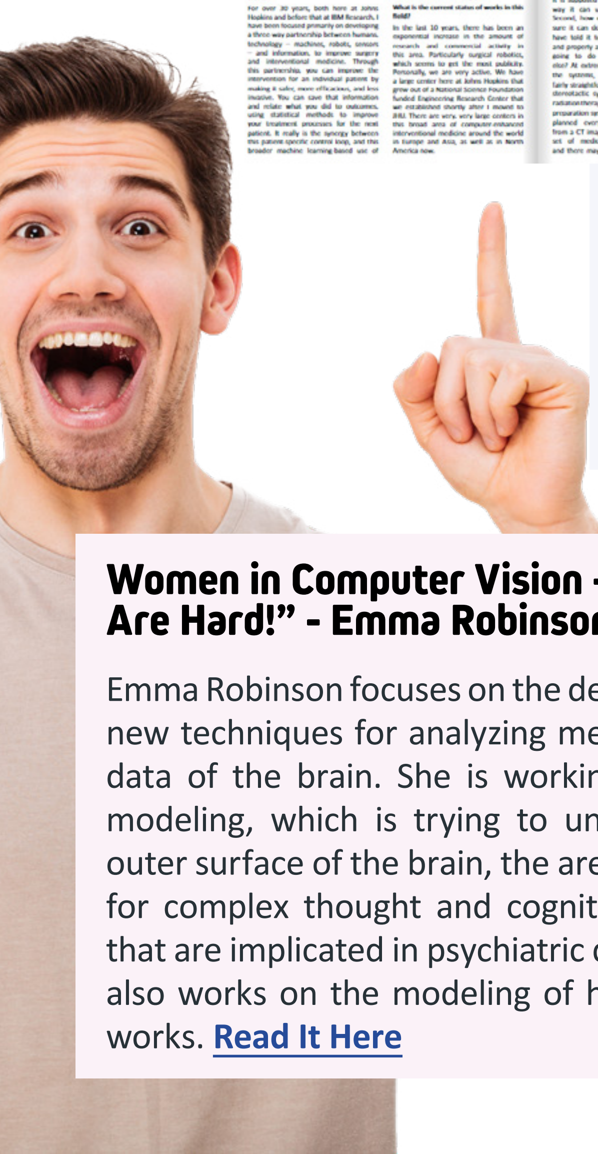
One-click segmentation is a semi-automated, AI-based segmentation and measurement tool for detecting selected regions of interest and their boundaries quickly and automatically. In medical imaging especially, this tool has many benefits: high accuracy, generalization, user impact and minimal effort. Watch the short video and see why this so simple technique by RSIP Vision is so powerful. Read It Here



ImageJ - Image Processing and Analysis in Java by NIH - with code!

ImageJ is an exciting and widely used tool. It's an open-source Java-based image processing program developed by **Wayne Rasband** at the National Institutes of Health and the Laboratory for Optical and Computational Instrumentation. ImageJ is becoming more and more popular, and you will know why by reading this report by our **Marica Muffoletto**, with code! Read It Here



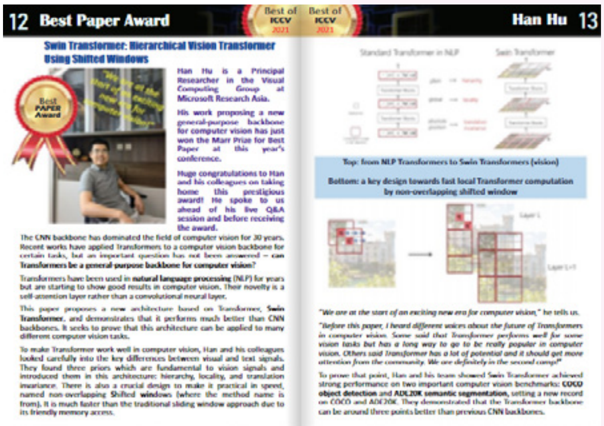


Russ Taylor - What is Happening in Surgical Robotics and the Hopes for the Future

Russell Taylor is Professor of Computer Science at **JHU** and director of JHU's Laboratory for Computational Sensing and Robotics. A pioneer in the field of Surgical Robotics, he has been focused primarily on developing a three-way partnership between humans, technology – machines, robots, sensors – and information, to improve surgery and interventional medicine. Read It Here

Women in Computer Vision - "Brains Are Hard!" - Emma Robinson

Emma Robinson focuses on the development of new techniques for analyzing medical imaging data of the brain. She is working on cortical modeling, which is trying to understand the outer surface of the brain, the area responsible for complex thought and cognitive processes that are implicated in psychiatric disorders. She also works on the modeling of how the brain works. Read It Here


12 Best Paper Award

Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows

Han Hu 13

Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows - Best Paper ICCV21

Out of 1,500 papers to be presented at **ICCV 2021**, we choose the one which later won the Best Paper Award. Believe it or not, our readers know that we are often very lucky: this has happened many times. Co-author **Han Hu** of **Microsoft Research Asia** told us about this work proposing a new architecture based on Swin Transformer that performs much better than CNN backbones. Read It Here

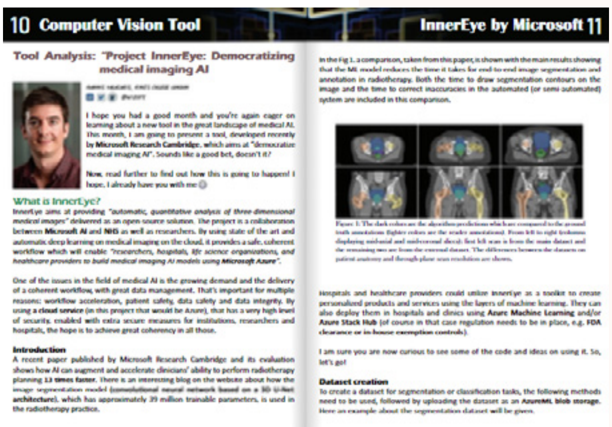
Class Imbalance in Classification Tasks - with code!

Datasets are called imbalanced when the distribution of examples across classes varies heavily, which often happens in computer vision and medical imaging applications. It can be due to incorrect data sampling or an inherent property of the domain. In most cases, it is advised to address this issue using specific tools and ad hoc libraries. Our **Marica Muffoletto** gives you precious tips, with code! Read It Here


14 Computer Vision Tool

Class imbalance in classification tasks

Class Imbalance in Classification Tasks 15


10 Computer Vision Tool

Tool Analysis: "Project InnerEye: Democratizing medical imaging AI"

InnerEye by Microsoft 11

Tool Analysis - InnerEye by Microsoft: Democratizing Medical Imaging AI - with code!

Microsoft Research Cambridge recently developed a tool, aiming at "democratize medical imaging AI". Sounds like a good bet, and apparently it is, if readers adored this report - with code! - written by our **Ioannis Valasakis**. Using state of the art and automatic deep learning on medical imaging on the cloud, **InnerEye** provides a safe, coherent workflow with great data management. Read It Here

As we work harder, we become luckier!

Mahdieh (Madi) Babaiasl is the founder of the robotics startup, Mecharithm, where she is a researcher and robotics instructor.

More than 100 inspiring stories of Women in Science here!

Madi, can you tell us about your startup?

I started Mecharithm about six months ago. My mission is to revolutionize learning, teaching, and research in robotics. There are lots of news outlets out there, like CNN and the BBC, but there is no news outlet for robotics. I want the researchers to have a great resource, and I also want to revolutionize robotics learning. Now we are working on university-level robotics: how to make it easy to understand for professors and students, because robotics is hard for everybody. Instead of PowerPoints and lectures, we want to make it accessible, understandable, and practical for students.

Where will the funding come from?

Mecharithm is at C funding right now. We are talking to some investors. There is an investor who really liked the idea. He said that it's the future and he's willing to put money in it. Once it's finalized, I will make a statement about it. We're still talking.

How will the company earn money?

When the product is ready, there will be several money revenues for the company. One is like news companies who mostly gain money from advertisers. We also want to create kits. Suppose that a university professor wants to teach automation: there will be a smaller version of an automated line in a factory (a model) that the professor can use for better tangible teaching. For instance, robots assembling furniture or a small car model. We're going to build a small factory for that professor to teach it, and the university would buy this from us.

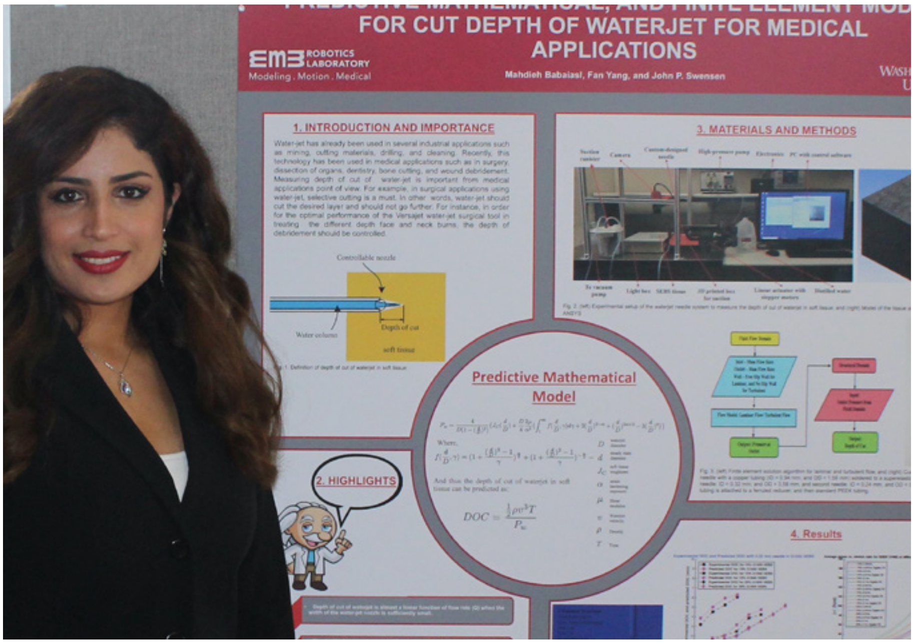
Are you going to sell content?

To some extent, yes. We also have a product… so content and product.

Can you tell us how you got to this point?

As long as I can remember, I was always interested in science. I liked math and physics in high school. When I wanted to go to university, I didn't want something theoretical because I like practical. The marriage between science and practical is engineering. So, I chose electrical engineering as my first major. Then I thought that electrical engineering isn't really for me. I wanted something that I could actually see, so I went and worked in robotics, something tangible. I was really interested in that. For my Master's, I was really interested in medical robotics. At the time, my dad had tremors in his hands. He was doing very delicate electronics on computer motherboards. He needed his hands to be stable for this work. I thought maybe we could develop a device that stopped these tremors. We started this project as a team. We developed a tremor suppression device. At that time, I also worked on a rehabilitation robot. I was also a part of the team for a surgical robot. This was my robotics career. Then I came to the US, and I continued robotics. I worked on steerable needles for my PhD. I got my PhD here. I always had this vision to do something for the world, to make a change and have something for myself. I had this vision so I started this company.

My mission is to revolutionize learning, teaching, and research in robotics!



My husband says, 'You work like a robot'!

Why did you come to the United States?

At first, for my PhD... The US has so much potential. You can do anything that you want. You can reach your potential. I think that in all parts of the world, you can reach your potential, but there are lots of opportunities here.

During this journey, did you meet scientists who inspired you and motivated you?

Yeah! There were lots of people who helped me on this path. The first person who helped me and had a great influence on me was my PhD advisor, Dr. John Swensen, who helped me and mentored me during my PhD. I also get advice from other mentors and friends in the area, whom I met at different conferences. They helped me to choose my career.

What does it mean to be Iranian for you?

[*laughs*] I really like to dance. The music and the food are both really good. It is a rich culture, full of kindness.

Don't you miss home sometimes?

I certainly miss my family! I think about my path, and they are very supportive. We usually meet together on Zoom meetings.

What are your hopes for the startup that you have created?

I have a great vision for the startup. In the future, I want to revolutionize learning and teaching robotics, from PowerPoint and these boring lectures to practical robotics that students really enjoy and engage in, so that they can learn. Also, for robotics researchers... for instance, if someone is researching autonomous vehicles. He or she cannot purchase a big car for it. We can create models for research that they can get and apply their algorithms on. Also, I want to categorize the world of robotics in terms of news so that researchers can have classified categories. I've already started it. I don't know if you've seen our videos.

Oh yes, there are many of them on YouTube. They are very impressive. Maybe some of our readers want to have a look.

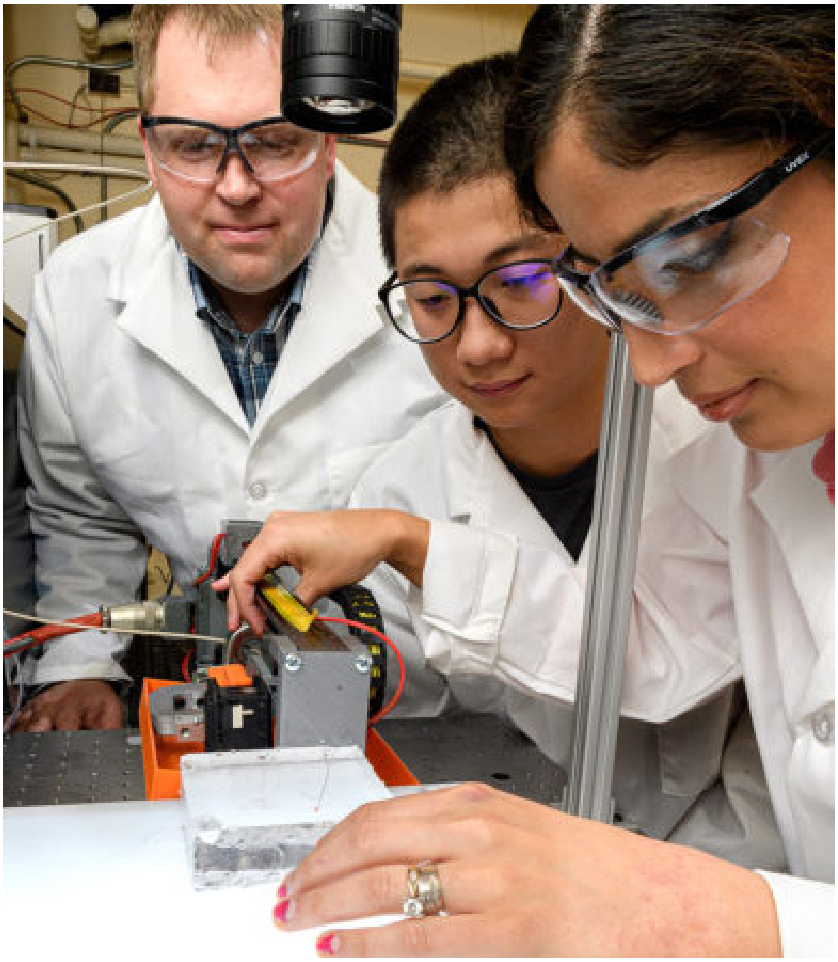
We also have a website.

What would be the main take-home success for you?

It means a lot to me. I want to have a voice in this world! As a female scientist, this company can be my voice. People can hear me, and I can inspire other female engineers and scientists to believe in themselves and try harder. We can do a lot of things!

What skills do you have that make you particularly well suited for robotics?

During my PhD at Washington State University, I was the lab instructor for the Mechatronics course. The professor was teaching the course. As lab instructors, my colleague and I were making these kits for students. At the end of that course, the students were really happy with the outcome. They learned with their hands. This is what I want to do. I have apps for teaching that make things simpler for them to understand. Some of my viewers said that I make complex things very easy to understand. That's something that I want to convey through my company. I think that I'm good at it!



I challenge you to show that now! Can you explain one complicated thing about robotics in a simple way?

[*laughs*] A lot of people think that robotics as a whole is complex. We can start with kinematics! Suppose my arm is a robot arm. It has 3 degrees of freedom for my shoulder, 1 degree of freedom for my elbow, and 3 degrees of freedom for the wrist. Degrees of freedom mean that it can rotate or translate easily. My arm has 7 degrees of freedom in total: 3 for the shoulder, 1 for the elbow, and 3 for the wrist. Overall, 7 degrees of freedom. An object in 3D space, like my cell phone, has 6 degrees of freedom because you can rotate it about the X axis. You can rotate it about the Y axis.

Robots are the future workers!

I think that I'm good at it!

Or, you can rotate it about the Z axis. You can translate it about X, translate it about Y, and translate it about Z. It has 6 degrees of freedom. Now, I want to manipulate this object with my hands. This object needed 6 degrees of freedom. My hand has 7 degrees of freedom, so it has 1 degree of freedom more. My hands are redundant manipulators. So I taught you redundancy in robotics!

How nice! I guess that a robot would have even more degrees of freedom. Technically, you could have infinite degrees, is that right?

No, not infinite… There are some robots that have infinite degrees of freedom, but industrial robots are usually 6 degrees of freedom. The redundancy depends upon your task. For instance, if a task needs 4 degrees of freedom, and your robot has 5 degrees of freedom, then it is redundant again [*see separate section on redundancy on page 39*].

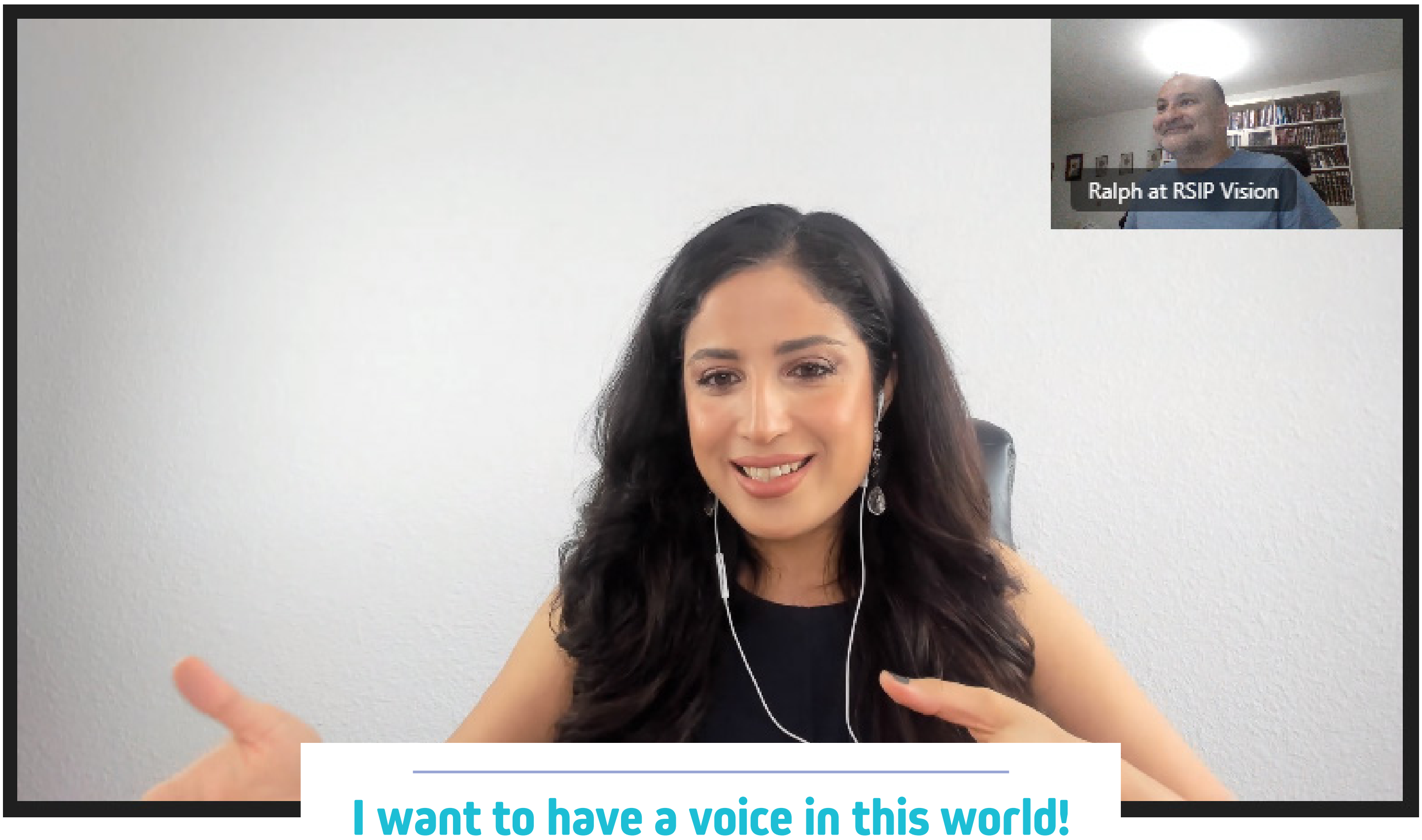
What do you think about creating robots to look like humans?

It depends on the task that you want to perform. Humanoid robots are usually used for tasks like human-robot interaction, or to take part in competitions, or to do some service. If you want to do a pick and place job, why use a humanoid robot? You can use an industrial robot with 3 degrees of freedom to do it. It doesn't need to be complex for your application. If the application is simple, use a simple robot.

I'd like to talk about the future with you. What is, in your opinion, the next big barrier that new robots are going to open for us?

Robots are the future workers. Humans used to do these repetitive jobs that are troublesome for their shoulders, for their joints, for their backs. Also, they are not good for humans in terms of creativity. Robots can replace humans to do these jobs. Automation is so prevalent nowadays. There are autonomous vehicles that no one needs to drive. They can do the job for us. Also, in stores, you can see these robots clean the floors for us. Humans that used to do that job can use their creativity in other ways. Those are robots' jobs. Robots can help us in those areas. They can do these repetitive jobs very easily for us. They can also augment someone's work. For instance, a physical therapist can use a robot. It's difficult on the joints to always work with patient after patient. A robot can help this therapist and augment his or her work for better treatment for the patient. One area that they are working on now is teleoperations. Robots can augment a surgeon's work. Surgeons can have shaky hands. A robot can help to resolve this issue. Sometimes they replace a human's job, and sometimes they augment a human's job.

Sometimes they replace a human's job, and sometimes they augment a human's job.



I want to have a voice in this world!

It sounds like you chose the perfect path for you! What quality about yourself is the most similar to a robot? And what is the quality that is the most different?

Most people say that I work like a robot. I work hard. [*laughs*] Sometimes I'm so focused on my work that I'm not really understanding what's going on around me. My husband says, "*You work like a robot!*" [*laughs*] At the same time, there are so many things that are not like a robot. As humans, we are compassionate, and we have sympathy. That's not like a robot.

Can robots laugh?

Yes, they can laugh. There are some humanoid robots that have features like to laugh. We all know that it's programming, and it's not genuine laughter.

Have you ever felt like competing with your robot to see who is best at doing something, you or your robot?

[*laughs*] No, that never happened!

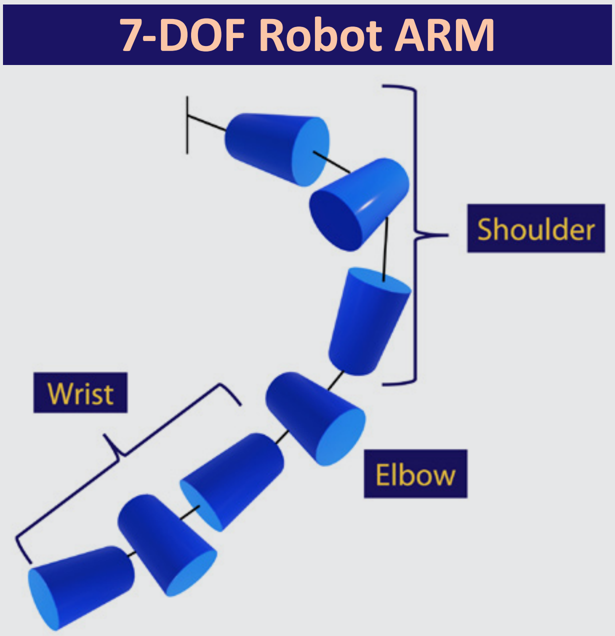
Your final message for humans, not for robots!

I want to tell your readers that luck is not something that makes us successful. We can create our own luck. Although there are things that are out of our control, like the family that we are born into, the country that we are born into; still, we can learn to create our own luck. As we work harder, we become luckier.

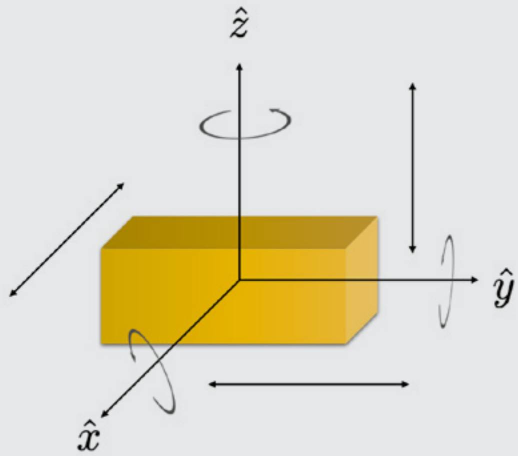
More than 100 inspiring stories of Women in Computer Vision here!

Redundancy

A robot has 7 degrees of freedom as the figure below:



And if it wants to manipulate a rigid object in a 3D space, the object only needs 6 degrees of freedom (3 for rotation and 3 for translation) out of the seven to be manipulated:



In other words, the kinematic function is a mapping from a higher-dimensional joint space to a lower-dimensional work space and thus the robot arm is considered to be redundant meaning that it has more degrees of freedom than the task to be performed.

DEEP LEARNING BARCELONA

by Petia Radeva (UB-CVC) and Xavier Giró (UPC)

Scientists rarely witness so fast scientific evolution as we observed the last decade in the field of **Deep Learning**. What yesterday sounded as science-fiction, today we see it as "nothing special" having self-driving cars, algorithms beating clinical experts for cancer detection and ictus prediction or Virtual Assistants completing our verbal orders, just to mention a few.

During the last 5 years, different scientists from several universities of Barcelona (UPC, UB, UPF, UOC, CVC-UB, etc.) organize the Deep Learning Symposium **(DLBCN)** to discuss the novel trends and challenges in Deep Learning. This year the DLBCN'2021 gathered more than 250 participants to discuss their advances published in the best Deep Learning conferences, as well as high-ranked Deep Learning journals.

The symposium also organized an industrial panel moderated by myself (**Petia**), with the participation of **Elisenda Bou-Balust** (Apple), **Antonio Bonafonte** (Amazon), **Daniel Arteaga** (Dolby Labs), and Cristian Canton (Meta AI). The panel covered some deep insights into the last challenges of Deep learning in the current industry and gave insights into the new relationship between the industry and academia.

Some of the interesting aspects covered were the exponentially growing trends in deep learning research, **the new role of the industry in the Deep Learning research** and how the collaboration changed between academy and industry within the new scenario considering **the increasing number of papers published by the industry**, the amount of free models provided from industry and academia, and other modes of showcasing the skills like contests and open-source projects. The panel also gave insights about the **emerging challenges**, their road to success, the skillsets industry seeks and ways to get into the industry through different positions. The panel discussed in detail the difference between the industry-academia relationships between the different ecosystems and **how Barcelona is putting its mark in the Deep learning world map**. The start-up cultures and the wide spectrum of professional opportunities of young researchers were another interesting topics of discussion.

COMPUTER VISION EVENTS

WACV 2022 Waikoloa, HI 4-8 January	AI and the future of radiology Bay Vision webinar 26 January	PMWC Silicon Valley S.Clara, CA 26-28 January
AAAI Conf. on AI Virtual 2-9 Feb	AI Ethics Summit S.Francisco, CA 17-18 February	Deep Learning Summit S.Francisco, CA 17-18 February
ICARA Aut., Robotics and Apps Prague, Czech Rep. 18-20 February	SPIE Medical Imaging S.Diego, CA 20-24 February	
Ai4 Healthcare 2022 Fully Digital 23-24 February		
CTO Plus (cardiology) New York, NY 24-25 February		

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.

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

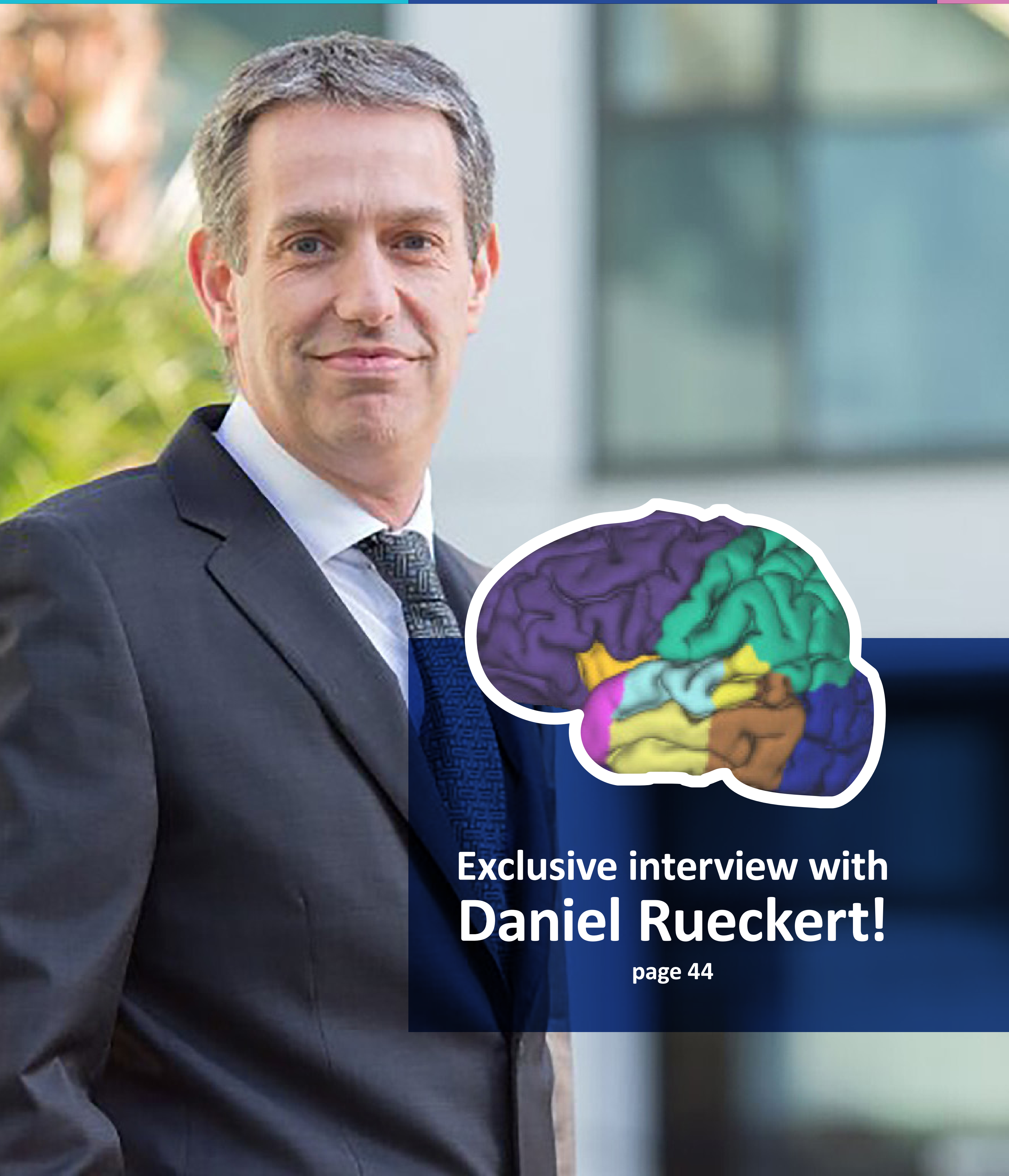
it takes less than 1 minute!

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

JANUARY 2022



Exclusive interview with Daniel Rueckert!

page 44

Do not be afraid of not knowing exactly where you are going. I never have!

Daniel Rueckert

Daniel Rueckert is the Alexander von Humboldt Professor for AI in Medicine and Healthcare at the Technical University of Munich and a part-time professor at Imperial College London. He speaks to us about his work at the intersection of computer science and medicine.

Daniel, our previous interview was three and a half years ago at MICCAI 2018 in Granada, when you had almost 50 PhDs students and a very big operation at Imperial. A lot has changed since then! Can you fill in the gap for us?

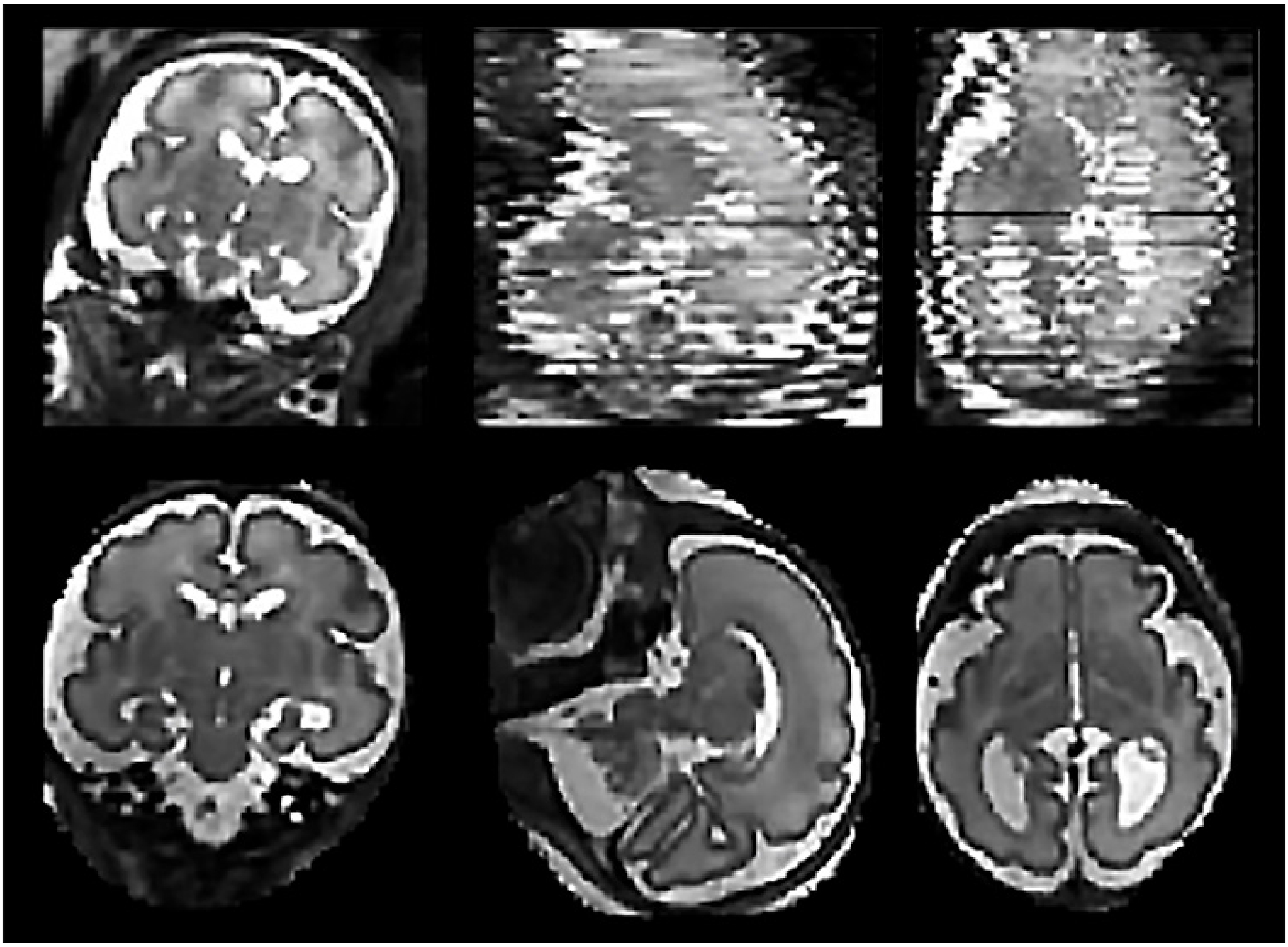
Imperial is still a large part of my activities, but when we last spoke, I had a big administrative role there as the head of a large department with nearly 60 academics. I wanted to get closer to research again and had a chance to take a position shared between medicine and computer science at TU Munich. It is a fantastic opportunity to bring the work we do much closer to the translation into clinical practice. Two years ago, Imperial College and TU Munich signed a strategic collaboration agreement to work more closely together. I was very lucky the people in Munich were keen for me to keep part of my position at Imperial, and the people at Imperial agreed!

I know how important and satisfying it was for you to see so many academics succeed because of your leadership and mentoring. Did pulling back from that role feel like any kind of sacrifice on your side?

Partly. You obviously have to reduce your time on one thing, but you are shifting your focus and working with a different group of people. You are absolutely right it is so rewarding to see the people you have been working with go on to be successful. It takes a lot of time to get to know people well enough to find out what their strengths are and where they would flourish most. That is probably the biggest part of the job. The impact you have from mentoring people is more long-lasting than any of your papers. In 10 years' time there will be something other than deep learning and nobody will remember your deep learning papers, but hopefully your colleagues will always remember you!

Can you tell us more about your research?

We want to go more in the direction of clinical translation. Taking the algorithms we develop and putting them in the hands of clinicians to get their feedback. Sometimes that feedback will be that it is great and will really help; other times, it will be that it is not working, and we need to improve it – we want to hear that! One of the aspects we are working on now is a system where we take every paper we produce and test it in a clinical environment. Germany is big on data protection. It is very



difficult to get data for medical research here. My colleagues, including Georgios Kaissis, and I are working on privacy preserving machine learning. This is a type of machine learning where you use data either in an encrypted fashion, or by leaving it where it is and taking the algorithm to it. This is what people call federated learning. We are also using differential privacy to make sure no one can reverse-engineer these models once we have constructed them. This gives mathematical guarantees on preserving privacy in a machine learning scenario. We have had a couple of papers in Nature Machine Intelligence on this topic in the last two years.

Who are your users?

Our users are mostly radiologists who work with medical imaging. The tools or algorithms we develop either simplify their tasks or allow them to make a diagnosis directly from the images. For example, measuring the size of someone's liver, or whether someone is likely to have a certain disease or not. Feedback from the radiologists tells us whether these

This is what people call federated learning!

algorithms work as they are supposed to work, and whether they are reliable and useful for them.

Nassir Navab told me some time ago ago how important it was for him that his students be in touch with physicians constantly to know what they need because they should be the main input for their work.

Exactly. Nassir and I work very closely together. It has been great coming to TU Munich and having somebody like Nassir as a colleague. He is very right, of course. You want to get feedback from clinicians, but you also need clinicians to define what the right problem is you should work on. Otherwise, you go ahead and develop something, then they tell you nobody ever wanted to solve that problem in the first place! Fortunately, that does not happen very often. I am also very happy I now have postdocs in my group who are medical doctors and radiologists. Your readers will know medical imaging often has different challenges than standard computer vision tasks. It is very important to understand those challenges and what you really want to do with the images you acquire.

In the 14 months you have been at TU Munich, can you pinpoint one thing you found in your research that made it worth making the move?

You mean worse, or worthwhile?

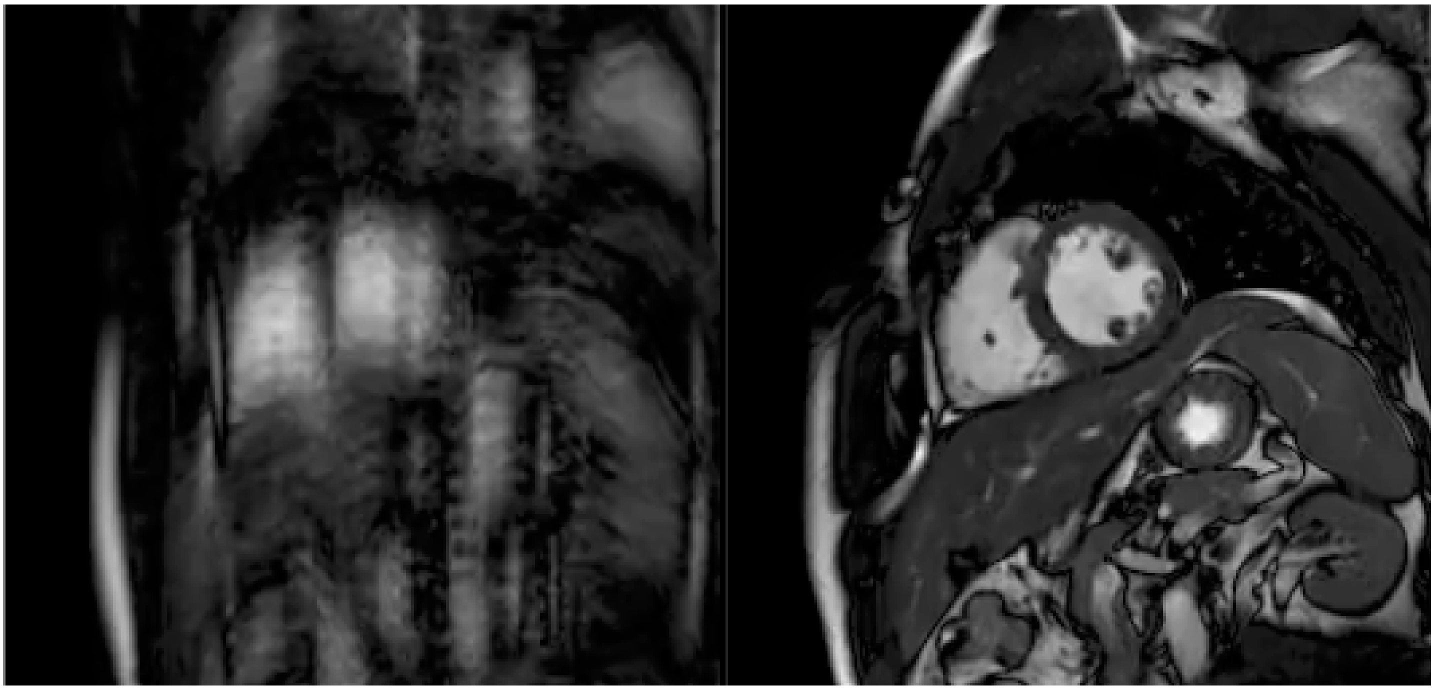
Yeah, worthwhile.

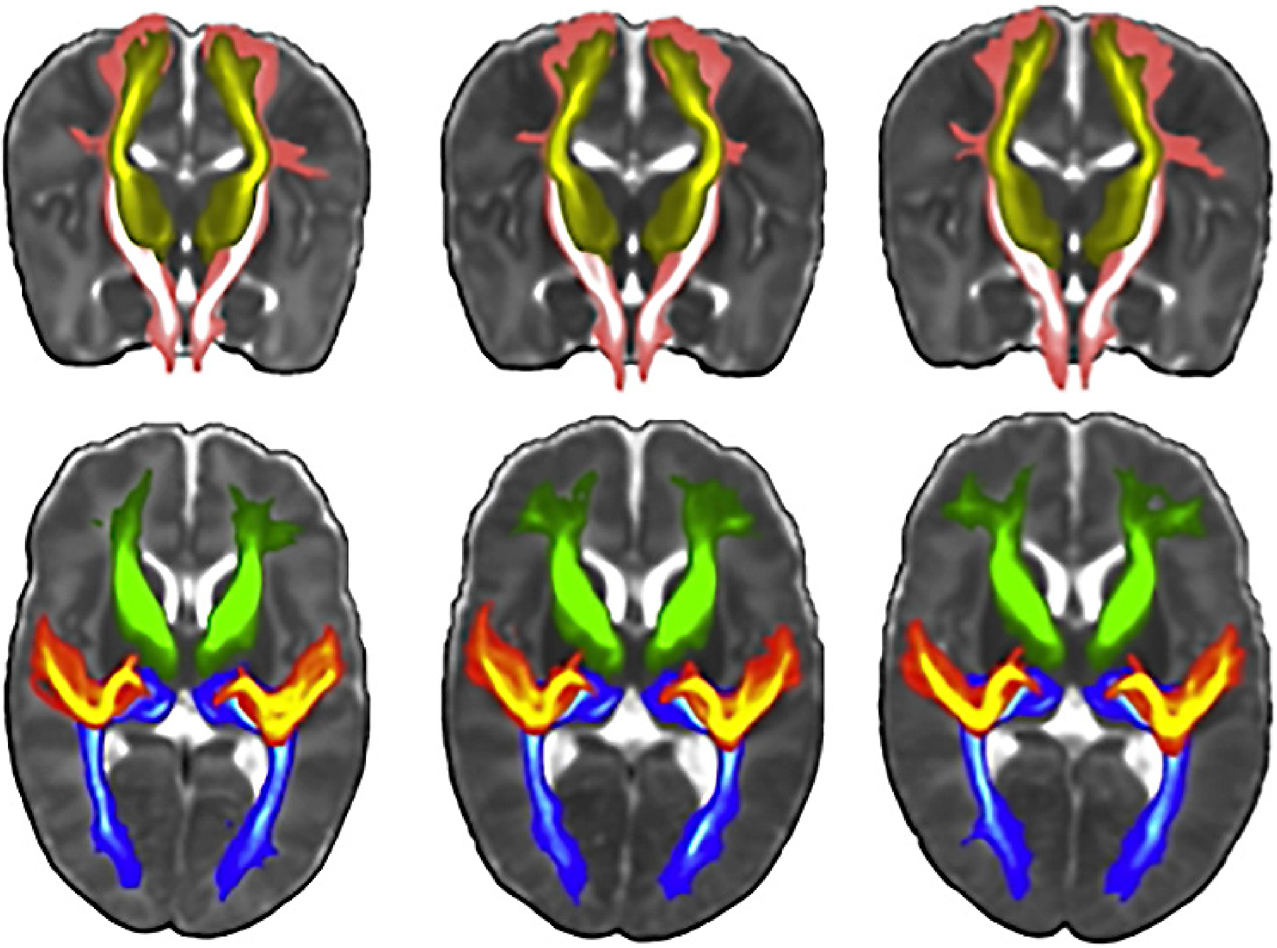
Oh, I thought you said it is bad!

No, worth, with T-H. I am generally a positive guy, so if in doubt, take the positive!

It is good you clarified it! [*we both laugh*]

We have developed algorithms which help us acquire better quality images faster. One of the worthwhile things has been seeing the difference this makes in clinical practice. I have learned a huge amount about what radiologists do in clinical practice, how they operate, their daily routine, and what their workflow looks like.





RSIP Vision organizes webinars in medical imaging. Out of the latest four speakers, three are German – and incidentally, all four are women! Why are German scientists at the forefront of this research?

That is a good question! I had studied in Germany, but had never worked here before, but one of the things I have noticed is in the German system there is less emphasis on a small number of excellent universities. It is a federal system, so it is much more diverse. It is not like Oxford or Cambridge or Imperial. You will find good scientists at universities which are not particularly famous. There is a lot of excellence around and there is a lot of funding available. The imaging technology is developed by companies like Siemens, who invest a lot in research, and there are a significant number of small- and medium-sized companies working in the space. These companies want to hire researchers with PhDs and there are many people interested in joining the field. In the UK, by comparison, there is a lot of investment in the research, but there is not much industry in the space.

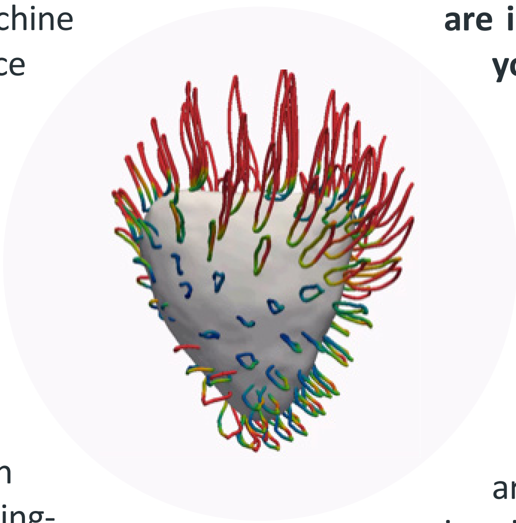
Would you say this is a great time to be a researcher?

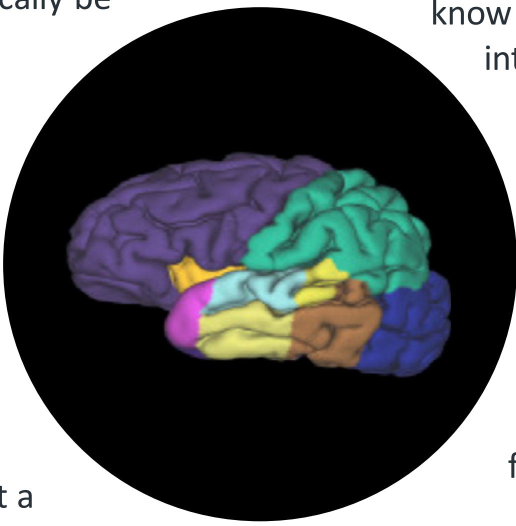
I cannot think of a better time! Particularly in this domain. I am inundated with so many requests from clinicians to collaborate


MICCAI2018
Granada
SPAIN
21st INTERNATIONAL CONFERENCE
ON MEDICAL IMAGE COMPUTING
& COMPUTER ASSISTED INTERVENTION
September 16-20, 2018
GRANADA CONFERENCE CENTRE - GRANADA / SPAIN
Nurture Community,
Recognize Excellence,
Develop Talent

"The impact you have from mentoring people is more long-lasting than any of your papers!"

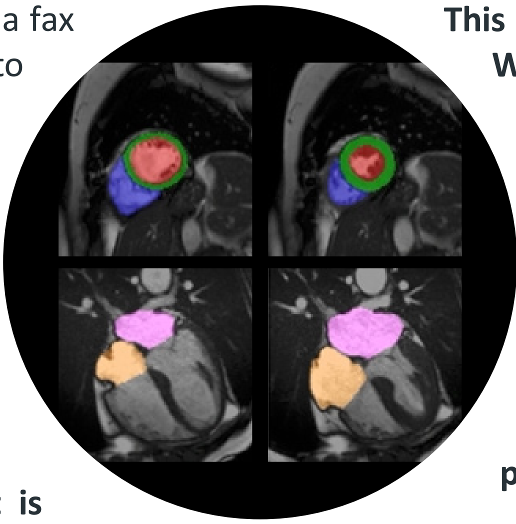
because they all feel AI, machine learning, and data science might help solve problems they could not solve before. Sometimes you have to explain it will not solve everything, but there is a lot of demand for our expertise. One of the nice things now is they view us as equals. There was a time when if you were doing engineering-type research you would basically be seen as a toolmaker.





What struck you the most coming back to Germany?

The most unexpected thing is the amount of bureaucracy. When I came here, I had to order office furniture and equipment, and a colleague told me I must get a fax machine. I could not remember the last time I had even seen a fax machine! I am still required to use it for certain things, such as faxing from one part of the hospital to the other. The lack of digitalization is quite shocking. Everybody is aware it is a problem, but nobody is fixing it yet!



Every time I talk to you, it is clear how happy and satisfied you are in your career. What advice do you have for our student readers so that they end up as happy as you are?

That is another good question! You have to realise you cannot plan everything ahead. Sometimes things happen and do not be afraid of not knowing exactly where you are going. I never have! I think this is quite common. Also, make sure you know one aspect really well and get into the detail of it because you will find it intellectually challenging but also enjoyable when you then solve it. Finally, always keep an open mind and talk to people from outside your own research area because many of the best ideas come from colleagues outside your field.

This has been fascinating, Daniel. We wish you a lot of success!

Thank you, Ralph. As always it has been a pleasure to talk to you. May I say you are very good at interviewing people.

The real secret is I am very lucky at finding the right people to interview!

Thank you, Daniel.

RSIP VISION
Visual Intelligence for MedTech

TUM

Imperial College London

BAY VISION VIRTUAL
MEETUP

AI and the future of radiology

How AI is elevating medical imaging while preserving privacy & providing trustworthy, explainable clinical solutions

January 26th 2022

10:00 am PT | 1:00 pm ET

REGISTER NOW

GUEST SPEAKER

Prof. **Daniel Rueckert**,
PhD, FREng, FMedSci, FIEEE
Alexander von Humboldt Professor for AI in Medicine and Healthcare at the Technical University of Munich. Professor of Visual Information Processing in the Department of Computing at Imperial College London.

HOSTED BY

Moshe Safran
CEO of RSIP Vision USA
Defining & developing innovative Medical Visual Intelligence solutions, in partnership with MedTech industry leaders.

Negin Ashouri is the co-founder and CEO at FemTherapeutics, a Canadian-based medical technology company using 3D printing and AI to improve the outcomes for women with pelvic organ prolapse. Fresh from graduating with a Master's in Computer Science from Concordia University in Montreal, Negin is here to tell us more about her organisation's bold mission.

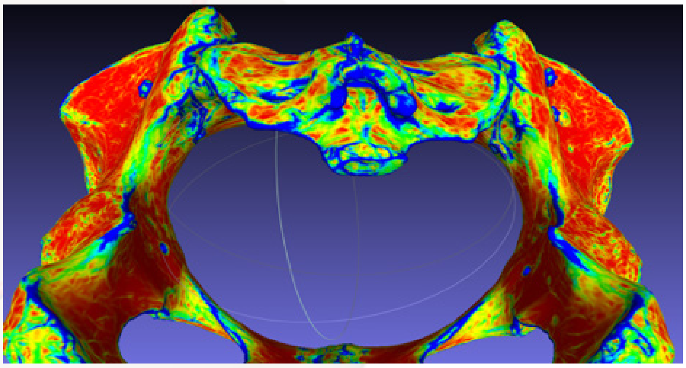
One in every 10 women around the world have a condition called **pelvic organ prolapse (POP)**, which is the descent of the bladder, uterus, or rectum down the vaginal canal. There are two solutions for this problem. **Surgery – either a hysterectomy or reconstructive surgery** – is one, but it is very expensive, and there is a 30% risk of developing another prolapse. Another solution, which is non-invasive, involves the use of an intravaginal device called a **pessary**, which can be inserted to hold the pelvic organs in place.

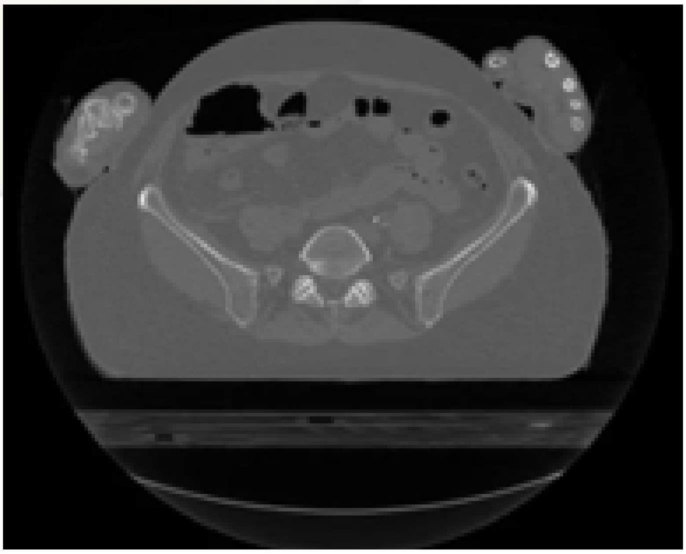
However, 40% of the time, women will discontinue using a pessary because of complications, such as displacement, where it can fall out of the vaginal canal, or where it is too tight and causes lacerations or discharge. This is more common for older women who are postmenopausal, but younger women who have had vaginal deliveries where there is a lot of pressure on the uterus itself can experience this pain and discomfort as well.

"Discontinuation and all these complications happen mostly because pessaries are not fit to the actual shape of the vaginal canal in each patient's body," Negin points out.

"Imagine you buy a shoe, and you have to wear it for the rest of your life, but it doesn't fit your foot at all! These pessaries are mass-produced in various geometric shapes and sizes and there hasn't been any innovation in them for over 50 years. What we are doing is bringing comfortability to patients by providing them with a pessary customized to their anatomical body."

The team have been talking to doctors in Canada and the US, and have been working

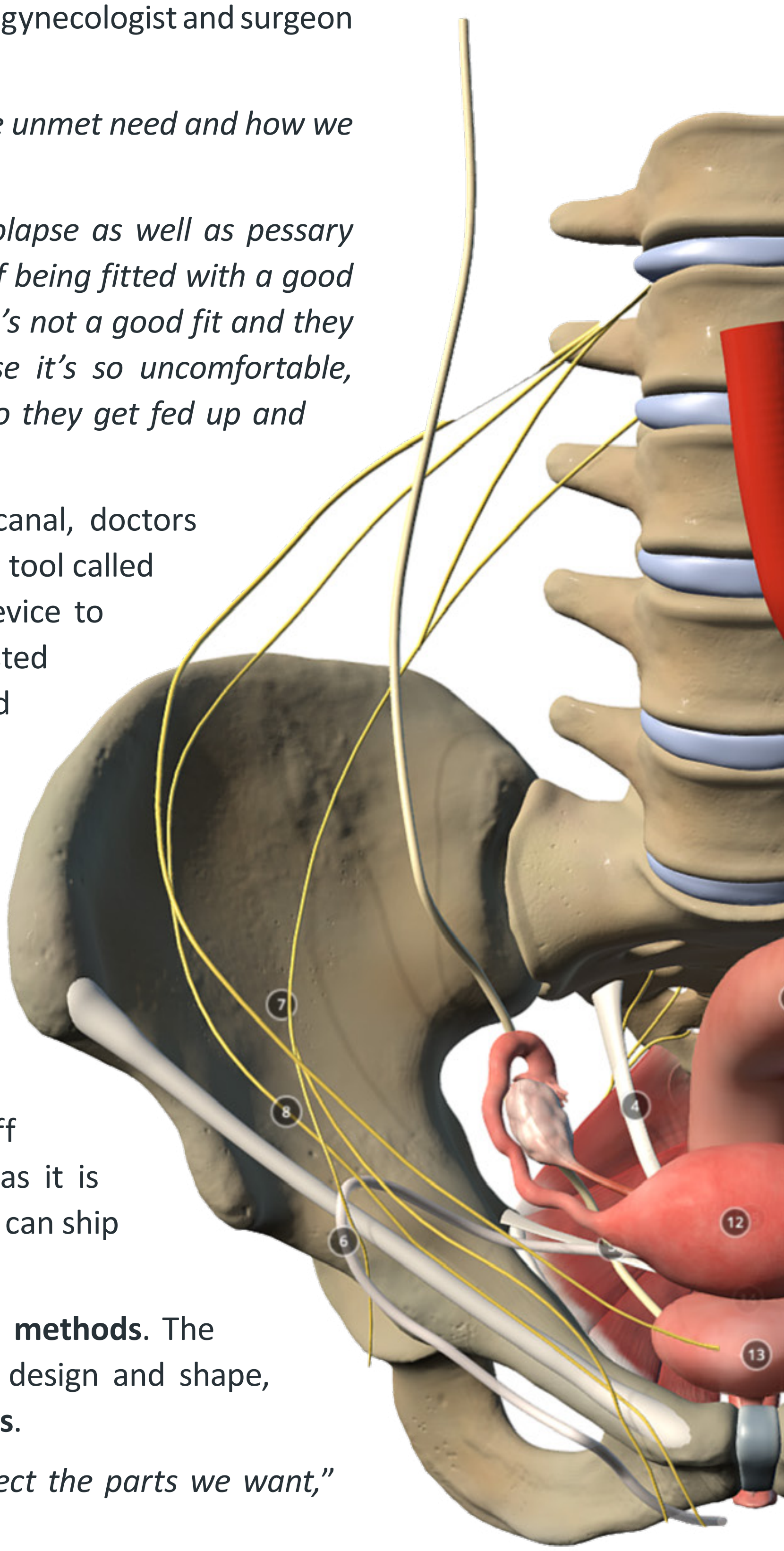




with a clinical supervisor, **Mihnea Gangal**, a urogynecologist and surgeon at Notre Dame Hospital, from the outset.

"We sat down with him and talked through the unmet need and how we could solve it," Negin tells us.

"He has lots of surgeries for pelvic organ prolapse as well as pessary fitting. Based on his experience, the process of being fitted with a good pessary is trial and error. For some patients, it's not a good fit and they come back wanting to have surgery because it's so uncomfortable, or they have a problem with displacement, so they get fed up and discontinue using it."



Currently, for measurement of the vaginal canal, doctors use their hands and an interactive assessment tool called **POP-Q**. The team had been developing a device to do this – they even built a prototype and tested it on a cadaver – but through research and development, found they could use equations to **translate the data from doctors** to fit their need.

When a patient goes to the doctor, the doctor validates their measurements and inputs them into FemTherapeutics's software. With those measurements, and the use of AI, the team can design an optimal pessary for the patient and send it off for production. They are using **3D printing** as it is simple, fast, accurate, and accessible and they can ship the pessary directly to the patient afterwards.

The model involves **machine learning and AI methods**. The first part uses imaging to predetermine the design and shape, with two datasets of **CT scan images and MRIs**.

"We annotate the images one by one to detect the parts we want," Negin explains.

"We use ***Gaussian curvature, deep learning-based segmentation methods, and contour and corner detection****. We make sure we do it for all slices of the CT scans and MRIs and use software like Fiji to put all the MRIs together to make a 3D shape of the segments we need. We use segmentation and localization to make sure we accurately detect the required parts. We bring this data into design and simulation*

software packages, such as Abaqus and Ansys."

Negin continues:

*"The second part uses **supervised learning**. When we have the patient data, there will be some parameters defined by the software based on the information obtained from the doctors. We fit them into the design and then we use AI – some basic Neural Networks – to run through the parameters and output the best parameters of all to create the most optimal design for that particular patient."*

FemTherapeutics is now 11 people covering business, marketing, technical, and medical. Its funding is non-dilutive and what it does have has come from government grants, Mitacs projects, competitions, and regular applications to bring in further funding from other sources.

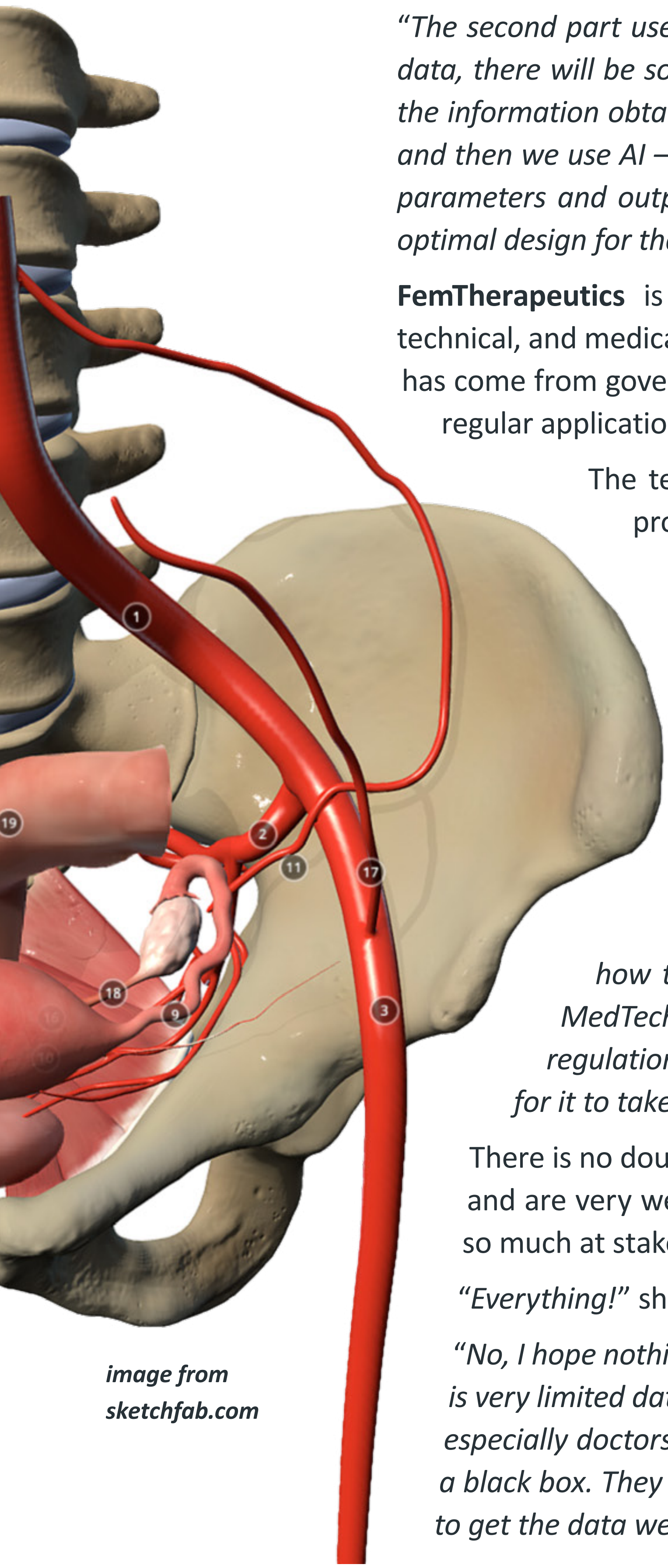


image from sketchfab.com

The team have set themselves quite a task with many problems to solve, including design, production, regulation, and distribution. As a scientist, is Negin prepared for all the challenges this new venture is going to entail?

"That's a good point!" she laughs.

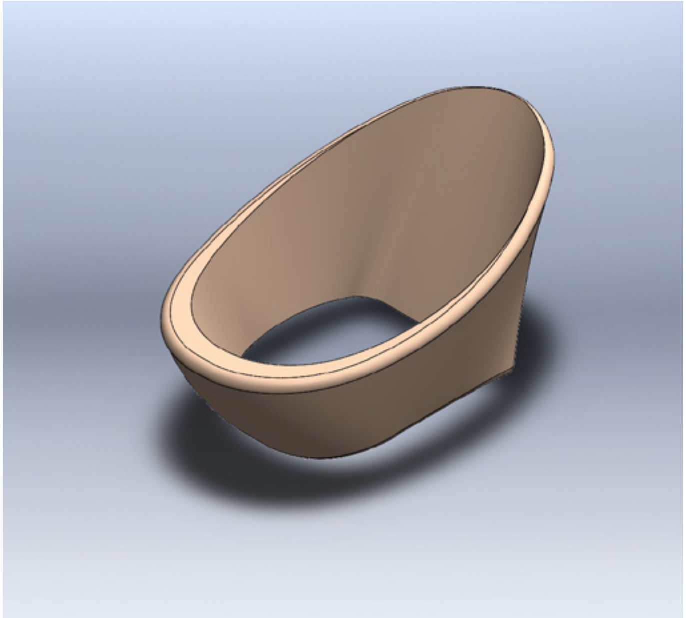
"But we have been working on this for two and a half years now and are fortunate enough to have lots of resources behind us. We did two accelerations where we met with other CEOs, scientists, and entrepreneurs and explored how they overcame these problems, particularly in the MedTech sector, which is more difficult because of the regulations and certifications involved. You have to be ready for it to take at least three or four years to get to market."

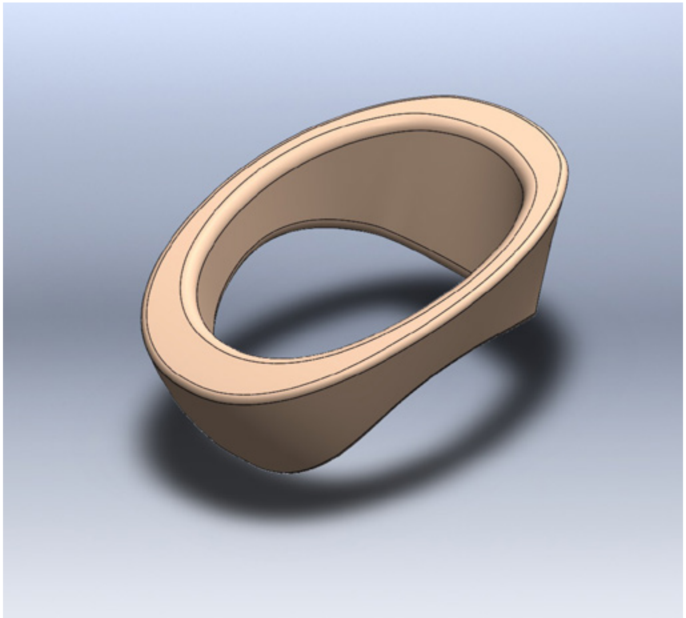
There is no doubt Negin and the team have done their homework and are very well placed to bring this project to market. But with so much at stake, has she thought about what might go wrong?

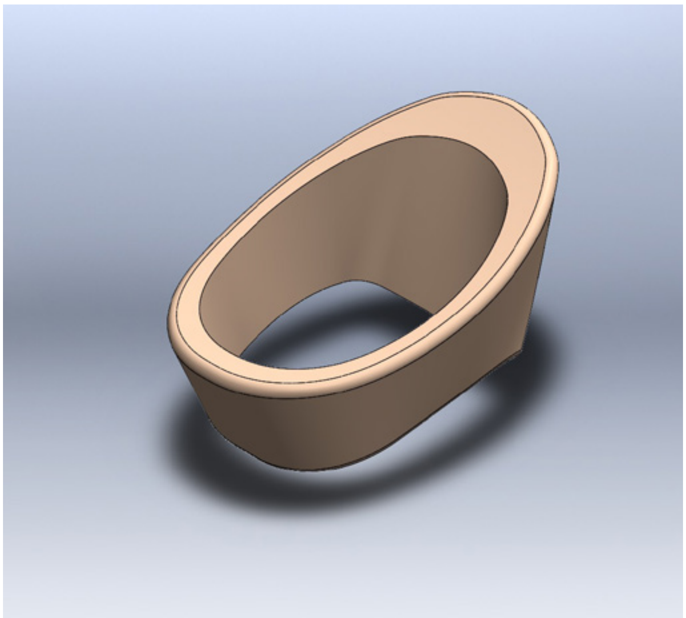
"Everything!" she jokes.

"No, I hope nothing, but we have to be prepared for anything. There is very limited data in the MedTech sector. People are resistant of AI, especially doctors. They don't know what they're looking at. It's like a black box. They don't relate to it. We have a lot of explaining to do to get the data we need."

Another issue they have is **access to data from hospitals and clinics**







around the world. To populate the AI model, they need plenty of data, but in the **pelvic medicine for women** domain there are not many datasets available and what there is has restricted access. To mitigate this, they have created synthetic data, close to the real data, and are trying to connect with doctors and hospitals to gain access. They have not been able to work with real humans so far but are planning to start clinical trials this year.

"Right now, we're doing everything in simulation, and then 3D printing to make sure the results we have from simulation mimic the real data in the lab," Negin explains.

"One risk is we cannot simulate comfortability, so we must make sure the patient is comfortable at the end. We think we have the most optimal design, but we have to make sure we really do."

Negin has recently been recognized for her efforts with a **Mitacs Entrepreneur Award**. This is awarded to people who turn their research into an innovative business that impacts the lives of Canadians.

"It's all very exciting!" she smiles.

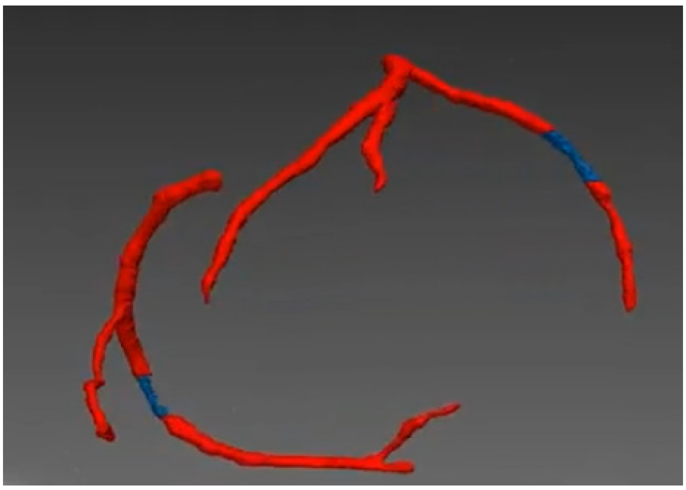
"We have lots of people interested in our work and are very lucky to have so many people around because they're not being paid but they're here for the vision of the company, which is to develop the most effective solution for women. There hasn't been any innovation in this sector for decades. If we can solve this problem and bring comfort to women around the world, it will have a huge impact."

CORONARY 3D RECONSTRUCTION FROM X-RAY

A new Technology for Non-Invasive Planning of Coronary Intervention

Heart disease continues being the number one killer in the western world, leading to ~25% of deaths in the USA. Consequently, cardiac procedures, specifically diagnostic catheterization and percutaneous coronary interventions (PCI), amount to $10B worldwide, making it one of the largest markets in the medical world. **Below we describe an example of how artificial intelligence (AI) provides nowadays an impressive contribution to improve the treatments provided by modern medicine with exceptional clinical results in an economical manner.**

When suspicion of **cardiac pathology** arises, a diagnostic cardiac angiography is performed. During this procedure, contrast dye is injected



into the coronary arteries while x-ray imaging is acquired. The cardiologist visually assesses whether the coronaries are normal or if a **stenosis** (narrowing) is present. Upon stenosis diagnosis, the severity is determined, and stent placement is planned. Planning requires accurate measurement of the stenosis length and diameter, taking into account the number of occurrences throughout the coronary branches. Additionally, an adequate stent needs to be selected to prevent misalignment and a need for a repeat procedure.

The planning stage may require additional contrast injections to better visualize problem areas, as well as time-consuming overview of the images to verify adequate anatomical understanding. The additional imaging is harmful for the patient as well as the crew in the cath-lab, as X-ray imaging uses ionizing radiation. Also, continuous overview of the images is costly as it adds use-time of the cath-lab.

Utilizing a new technology for 3D reconstruction from 2D x-ray images can provide a solution to the above-described challenge.

The input for this is 2 image-sequences of coronary angiography acquired from different angles. Using state-of-the-art deep learning

algorithms combined with classic computer vision methods, a 3D model of the coronary arteries is reconstructed accurately, rapidly, and automatically. This model can be used to better visualize artery structure and measure vessel dimensions in points-of-interest. More advanced capabilities, enabled using deep learning, are **stenosis detection**, **3D quantitative coronary angiography** (3D QCA); they can even be the baseline for **computerized fractional flow reserve (FFR) measurement**. Additional modules can show artery modification due to stent placement or place a virtual stent in the desired position within the coronary artery. These uses replace invasive cardiac measurements and give the physician a better pre-procedure planning tool for stent selection.

Using a 3D model for visualization is more intuitive than using the 2D view currently available, resulting in fewer contrast injections and x-ray exposures and reducing the risk for contrast over-dose and radiation-associated complications.

Finally, embedding this technology in the PCI workflow **shortens the procedure** and adds confidence to the physician's decision-making.

Overall, this tool allows better anatomical and pathological visualization and appreciation, as well as AI-based algorithms for coronary lesion characterization, and augments the procedural success, decreases the complication rate and (based on a very broad scientific literature) has the potential of increasing stent patency rate and patients' long-term survival. This AI-based technology requires no additional hardware installation and can be executed using the existing setup of the cath-lab. It is a quick and efficient method to provide better healthcare for patients undergoing PCI.

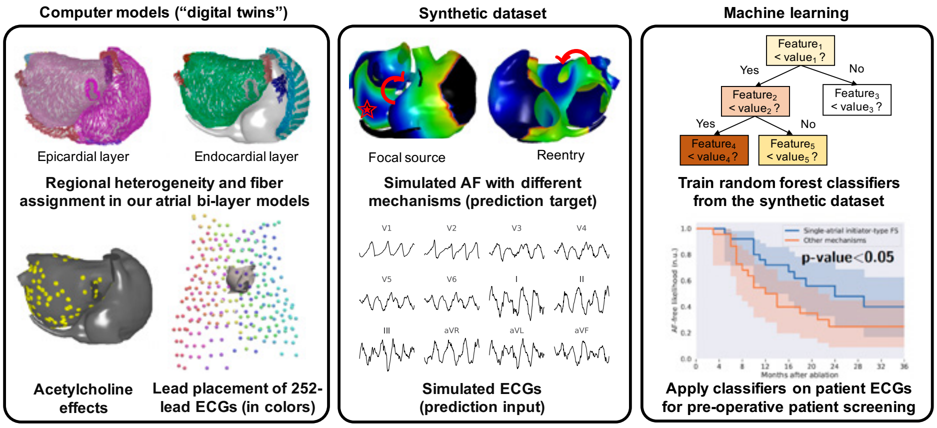
Read about this and more pioneering solutions in AI for Cardiology.

Yingjing Feng recently completed her PhD in Applied Mathematics at the University of Bordeaux and the IHU-Liryc institute in France. Her expertise lies in bridging machine learning, computational modeling and electrophysiology for atrial fibrillation. Previously, she obtained an MSc in Computing (Machine Learning) with Distinction from Imperial College London. After her PhD, she will start as a research fellow at the University of Birmingham, working on machine learning algorithms for epilepsy modeling. Congrats, Doctor Yingjing!

Atrial fibrillation (AF) refers to the most common clinical arrhythmia, where the atria activate rapidly and irregularly. Standard catheter ablation protocols to treat AF, however, show sub-optimal long-term AF freedom maintenance, substantiating the need for personalized ablation. Non-invasive mapping techniques provide essential pre-operative planning tools for personalized ablation, which include 12-lead and 252-lead electrocardiograms (ECGs). In my thesis, I developed three algorithms to extract the prognostic values from patient ECGs by using machine learning techniques.

Computer models as “digital twins” to aid the development of machine learning algorithms

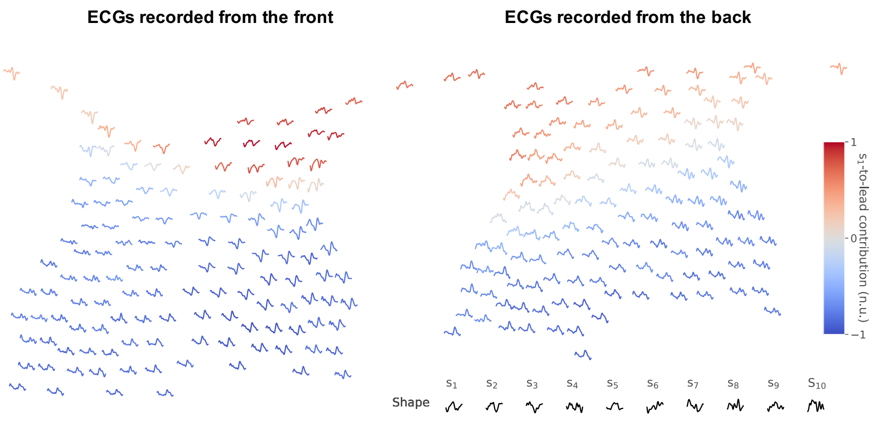
There were several challenges associated with predicting post-ablation outcomes from pre-operative ECGs: there were a limited number of patients as training examples,



and we did not have the ground-truth atrial activities associated with each ECG. That's why I looked into building "digital twins" for AF patients using computer models. These models were built on detailed atrial geometries made from patient meshes, with realistic electrophysiological parameters and computational rules. We generated a large and diverse set of AF episodes, caused by different AF mechanisms. We then trained machine learning classifiers on the synthetic dataset to distinguish the mechanism non-invasively, and applied these classifiers on patient signals. With this, we successfully predicted a patient group that was more likely to benefit from the current ablation procedure.

Extraction of spatiotemporal features without an image scanner

The most important contribution we made was the development of spatiotemporal machine learning algorithms without using an image scanner, by exploiting common structures between ECG channels. One example is shown below, where we used a second-order blind source separation algorithm based on Belouchrani et al. to extract 10 atrial activities (shown as S_1 to S_{10}, at the bottom right corner of the figure). The figure also shows 252-lead ECG signals recorded from a patient, colored by the contribution (S_1-to-lead contribution) received from the principal atrial activity (S_1). It can be seen that signals in dark red, which received the highest contribution from S_1, had a similar morphology with S_1. The S_1-to-lead contribution, which was computed from ECGs, demonstrated consistency in encoding the location of S_1, regardless of the difference in geometries and electrical properties between patients. Therefore, it was used as a spatiotemporal feature for classification. With our scanner-free algorithms on imaging the atria, AF patients are no longer subject to any imaging process (MRI or CT) to obtain a non-invasive diagnosis of the underlying mechanisms. This means improved accessibility for both clinical and home uses.



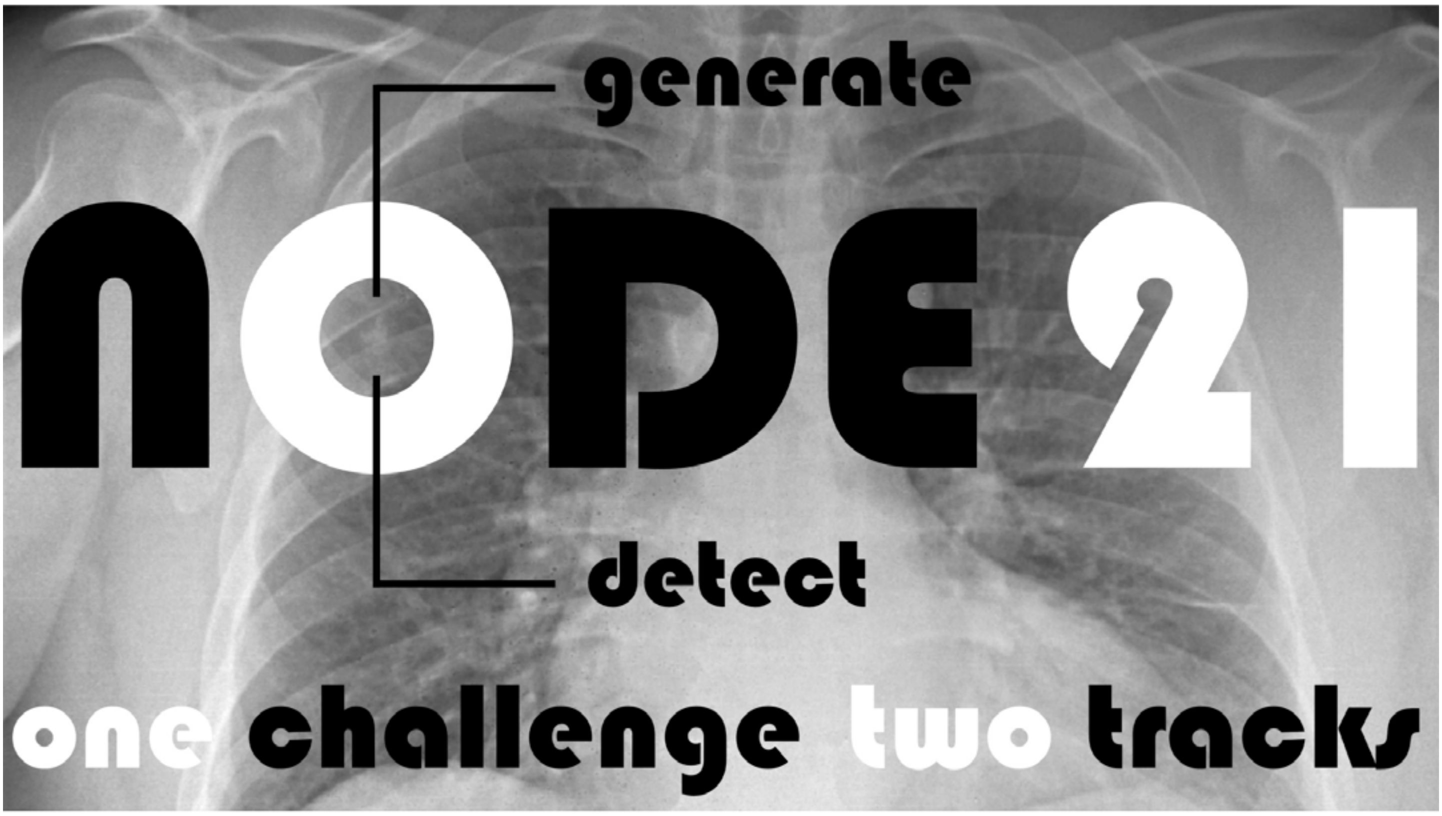
Acknowledgements:

The research is supported by European Union Horizon 2020 Research and Innovation programme "Personalised In-silico Cardiology (PIC)" under the Marie Skłodowska-Curie grant agreement no. 764738 and GENCI computing resources, allocation A0080310517.

Ecem Sogancioglu is a final-year PhD student at the Radboud University Medical Center (Radboudumc) in the Netherlands. She joined the Diagnostic Image Analysis Group as a PhD candidate in 2017, under the supervision of Professor Bram van Ginneken and Senior Postdoc Keelin Murphy. They are organizers of the first NODE21 challenge, which is about detecting and simulating lung cancer, and Ecem and Keelin are here to tell us more.

Keelin Murphy

Ecem Sogancioglu



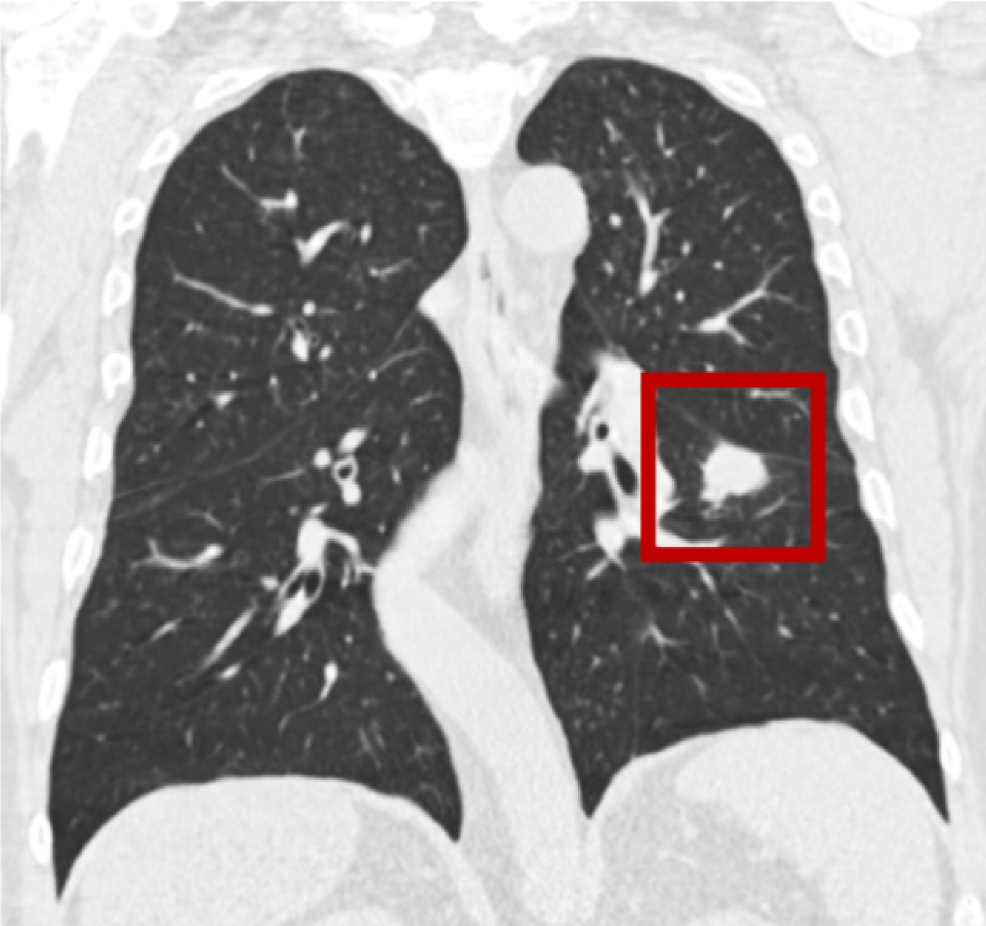
Lung cancers, which are visible in the lungs as nodules, are often fatal. The most important resource available for fighting them is time. Once someone is showing symptoms, it is often already too late. Hence there is a need for **easy and affordable imaging** to identify these nodules at the earliest opportunity. Routine screening is one way to improve a patient's chances and chest X-ray is by far the most cost-effective option. Chest X-rays are used in many other health-related scenarios, so checking them for nodules as a standard task would be one way to improve outcomes for lung cancer patients.

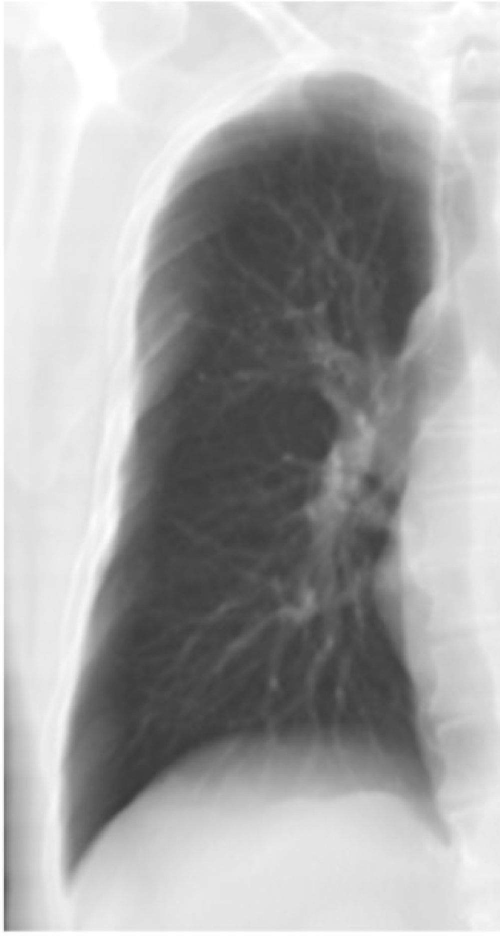
"Early detection of lung cancer is vital and is one of the most important applications for chest X-ray," Ecem tells us.

*"Recently, we have published a review paper where we covered around 200 papers that were using **deep learning and chest X-ray analysis**. Most of the researchers put their efforts into solving these problems using publicly available datasets, with labels generally extracted from radiology reports by text mining. Despite the importance of this nodule detection task, the only papers covering it were from companies with enough resources to collect large, private datasets, which aren't available to the community."*

NODE21 is a MICCAI-recognized challenge, which shows how important this issue is for the MICCAI leadership. Ecem sees it as more of a collaboration than a competition, presenting the opportunity as a research community together to develop a good system for detecting nodules and to compare the performance of different algorithms in a fair way.

It is a two-track challenge. The detection track is for developing automated algorithms





to detect nodules. The generation track is for developing algorithms to generate nodules at the requested location.

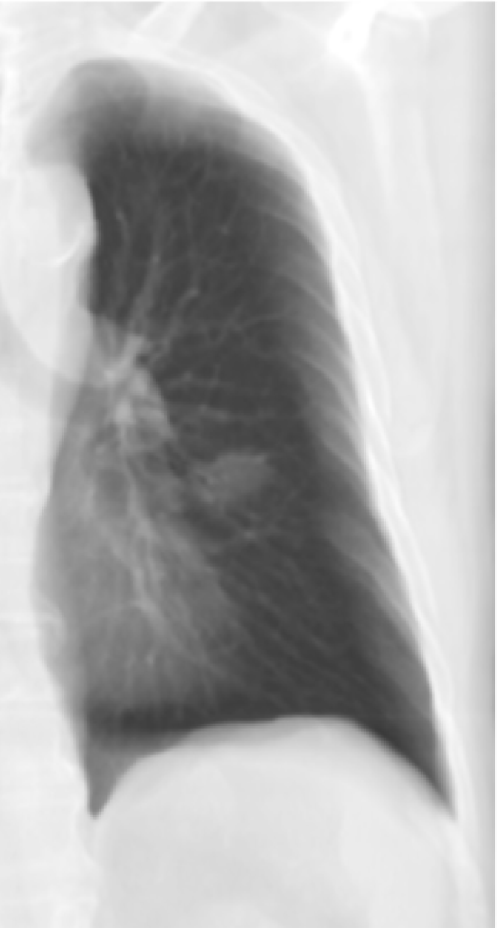
"We hope the generation track algorithms actually help to improve the performance of the detection track," Ecem explains.

"In that way, it's even more collaborative. We're all pushing each other. One track is trying to improve the performance of the other track, so we're all trying to get the best performance together."

Ecem's co-supervisor, **Bram van Ginneken**, is the co-founder of **grand-challenge.org**. He loves organizing challenges and using the platform to see everything in a fair and comparable way. This challenge came out of his and the team's excitement around generative modelling.

"One of the things we noticed was people weren't focusing on tasks that were clinically interesting," Keelin points out.

"People were focusing on tasks they have labels for, and some of the labels are things like: Does the patient have a pacemaker inserted? That's not a clinically interesting question. You can just see it immediately on the image! Lung nodules, by comparison, can be very difficult to see in the chest X-ray because of the projections of the ribs, heart, and diaphragm. If a patient has come in for a heart check, the radiologist may not even be looking for nodules and so could very easily miss one. They're only human. With AI, we can make suggestions to look more closely at certain parts of the image in case they have missed something."



The test cases used in the challenge are especially valuable because the labeling was performed by looking at **a CT scan and chest X-ray together**. This is **gold standard labeling**, which is great for participants developing algorithms. They can see how their algorithms perform on clinically relevant data.

Ecem and Keelin work in the same group at **Radboudumc** and tell us they share the same principles.

"We work on open science," Keelin says.

"With this Grand Challenge platform, we make everything reproducible and let people try things out. We always aim to do things that are clinically useful. Working in a hospital here, we really want to make clinically relevant chest X-ray applications."

With the final submission date looming, Ecem and Keelin advise everyone to check the tutorials before you submit to make sure you have not forgotten anything! There is a forum on the Grand Challenge website where people can discuss issues and contact them if necessary.

Finally, could we be seeing a second edition – **NODE22**, perhaps?

"I'm not sure yet!" Keelin teases.

"There may be future developments, especially on the generation track. This is something that is quite new. There's been a lot of work recently about generation of different kinds of images, but we haven't seen something that's really clinically useful yet. We're curious to see what will happen."

VIRTUAL REALITY AND HEALTHCARE GLOBAL SYMPOSIUM

Robert Fine is the founder and executive director of the International Virtual Reality and Healthcare Association. Currently, he is busy organizing the 6th Annual Virtual Reality and Healthcare Global Symposium, which takes place in Nashville in March. He tells us more about what we can expect from the event.

Robert has been working in VR since 2016 and after experimenting and seeing where the interest was in different verticals, he launched the first **Annual Virtual Reality and Healthcare Global Symposium** the following

year. Having spotted the VR and healthcare sector was growing fast, he realized it could benefit from having its own community, so in 2018 he founded the **International Virtual Reality and Healthcare Association (IVRHA)** to fill that gap.

"From the start, we designed IVRHA to be representative of the healthcare ecosystem," he recalls.

"It's not just about VR start-ups. We have teaching hospitals, universities, primary care hospitals, and research organisations as members. A quarter of our members are outside the United States. I just got an email from one member to say they have been acquired by another member – they wouldn't have met if it weren't for IVRHA."

Now in its sixth year, the Annual Virtual Reality and Healthcare Global Symposium is **the largest event around on the topic**, hosting a mix of start-ups, established technology companies, academia, and clinicians all working together to create successful products for the marketplace.

The last in-person event was held in March 2020 and was very well received, so after a challenging couple of years, Robert is looking forward to finally **getting together again with everyone in Nashville**.

"If people are trying to understand virtual reality in general and what this Metaverse is all about, as well as how it's playing into healthcare, they should register," he tells us.

"For healthcare, it's all cutting edge. New products. New applications. It crosses a very wide range of disciplines – mental health, pre-surgical planning, physical rehab, treating addiction. If you're a healthcare practitioner, it will impact you down the road eventually, so why not get ahead of the game?"

The event will host 50+ speakers, three tracks, two receptions, and has been scheduled for the end of the week to give those travelling the chance to stay for the weekend and enjoy what Nashville has to offer.

*"I think participants will realize **the revolutionary impact virtual reality can have on healthcare** and is already having on healthcare. They will see ways to integrate it into their own practice."*

Robert also recognizes computer vision is a highly important area within VR and healthcare:

*"I lead a subgroup on terminology working with the **FDA** and one of the things they're focused on is trying to figure out how to measure and **evaluate VR products from a computer vision standpoint**,"* he tells us.

Robert has just co-edited and published a new book called **Applied Virtual Reality in Healthcare: Case Studies and Perspectives**. With a foreword written by Palmer Luckey, who founded Oculus before it was bought by Facebook, it is sure to be a popular tome.

"It's a great primer for anyone trying to understand the industry, where it's impacting heavily into healthcare, and what many of the opportunities are," he reveals.

Is this the first of many future editions?

"Well, this was a two-year project with 60 contributing authors, and it ended up around 350 pages. There will likely be another volume, but you will have to wait another couple of years!"

If you would like to buy a copy of the book, Robert has been kind enough to offer a 20% discount for **Computer Vision News** readers using promo code 'rsip'.

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

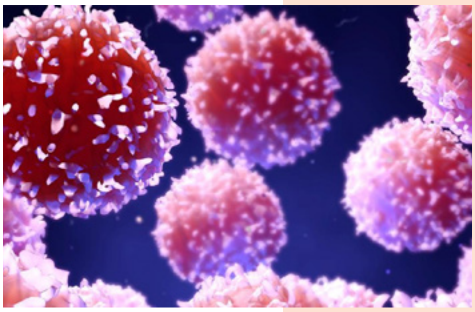
AI Accurately Predicts Who Will Develop Dementia in Two Years

This AI predicting who will **develop dementia in two years** with 92% accuracy is important stuff, because it reduces the guesswork in clinical practice and significantly improve the diagnostic pathway, helping families access the support they need as swiftly and as accurately as possible. It reduces by 80% the number of false positives reducing the unnecessary distress that a **wrong diagnosis** could cause, by spotting hidden patterns in the data and learning who is most at risk. This research from the **University of Exeter** used data from more than 15,300 patients in the United States. Read More

Mind-controlled Robots Now One Step Closer

Two groups of researchers at **EPFL Lausanne** have developed an ML program that commands a robot arm obeying to a human via **electrical signals from the brain**. This ML uses inverse reinforcement learning to learn commands coming a tetraplegic patient wearing a headcap equipped with electrodes scanning the patient's brain activity. When the robot makes an incorrect move, the patient's brain emits an error message through a clearly identifiable signal, which indicates to the robot *"No, not like that!"* The robot will try again with another course until it learns the patient's intention. **Watch the video!**

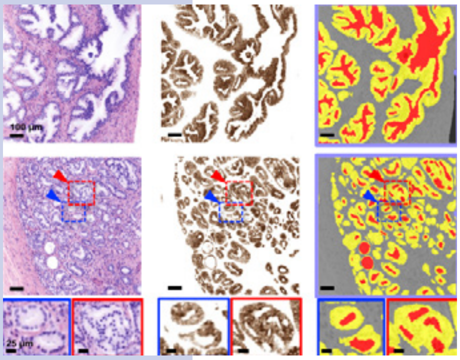
AI-Driven Medical Imaging May Help Fight Against Rectal Cancer

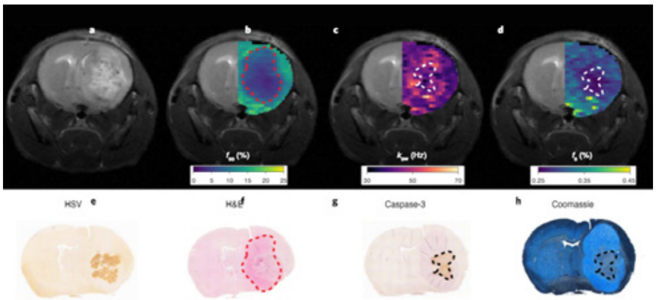


A **Case Western Reserve University**-led team is figuring how AI can help **rectal cancer patients**. Apparently, many of them cannot rely on chemotherapy or radiation, so most patients have to undergo invasive surgery. Researchers want to **avoid patients being overtreated**, so their AI will tell clinicians right up front - based on a routine MRI scan - if a patient will do well with only chemoradiation, **without having this serious surgery**. They will test their **radiomics** (an AI that extracts many features from medical imaging using algorithms on large data) on about 450-500 patients. Read More

3-D Imaging Method May Help Doctors Better Determine Prostate Cancer Aggressiveness

The great folks at **ITNonline** inform us that a team led by the **University of Washington** has developed a new, non-destructive method that images entire 3D biopsies of instead of just a slice. The purpose is to make the computer use 3D and 2D results to predict the likelihood that a patient has **aggressive prostate cancer** and therefore needs to be treated right away. Looking for abnormalities in slices of biopsied tissue is a classical way to assess the aggressiveness of the disease. But turning the slices into 3D makes it much easier to properly diagnose even borderline cases. Read More



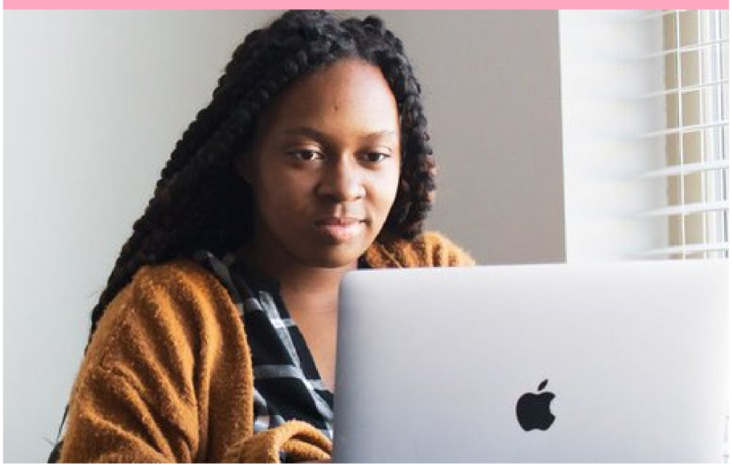


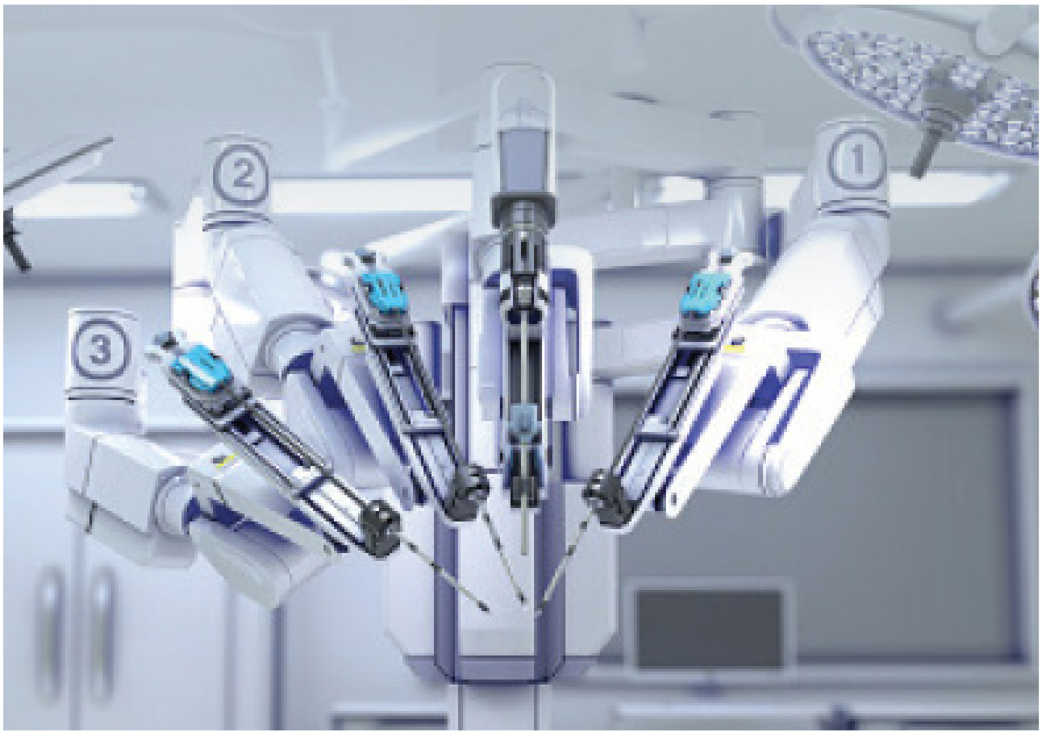
Artificial Intelligence, MRI Detect Early Signs of Tumor Cell Death

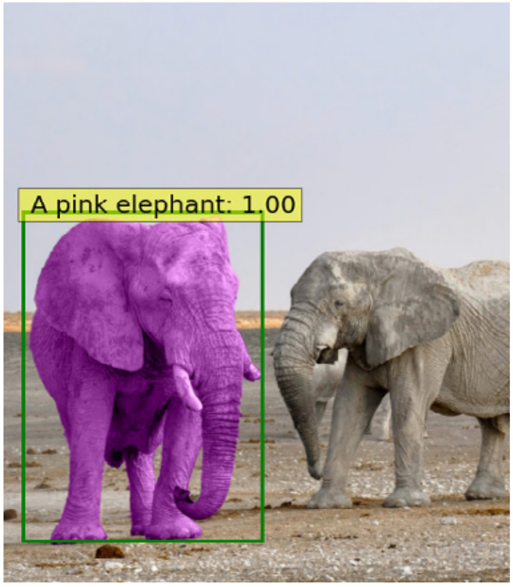
Massachusetts General Hospital (MGH) researchers found that MRI and AI can be used to **detect early signs of tumor cell death** in response to a novel **virus-based cancer therapy**. The researchers claim that they programmed an MRI scanner to non-invasively create unique signal 'fingerprints' for different molecular compounds and cellular pH. **A deep learning neural network** was then used to decode the fingerprints and generate quantitative pH and molecular maps." The main novelty of their approach lies in the use of **simulated molecular fingerprints** to train the ML neural network. Read More

Can Artificial Intelligence Stop the Harsh Reality of Internet Trolling?

Why is this in MedTech Spotlight News? Follow the guide. Believe it or not, Internet trolls are able to **spread flashing images on social media that trigger seizures in people with photosensitive epilepsy**. It's a terrible online abuse that has triggered this response by Bristol-based innovation agency **Gravitywell**. They have developed a plugin named **'unFlash'**, an ML AI that learns how to effectively detect and block images that fall within the critical photosensitive range, helping **social media** to prevent harmful imagery from being shared online - or to display a warning where necessary. Read More






A pink elephant: 1.00