June 2021

Computer Vision News

The Magazine of the Algorithm Community

Women in Science

IBM Watson Health

Automatic Gesture Recognition in Surgical Robotics

Feel at CVPR!

Challenge: Segmentation and Registration in Fetoscopy

Artificial Intelligence Spotlight News

ConvNets using Julia

Bay Vision Meetup

RSIP VISION

Global Leader in Computer Vision and Deep Learning

DILBERT

BY SCOTT ADAMS

Computer Vision News

Editor:
Ralph Anzarouth

Engineering Editors:
Marica Muffoletto
Ioannis Valasakis

Head Designer:
Rotem Sahar

Publisher:
RSIP Vision
Contact us
Free subscription
Read previous magazines

Follow us:

Dear reader,

We have many reasons to be excited this month! First, **CVPR 2021** is just around the corner! The entire community will virtually meet starting **June 19**. It will be a very special edition, and we know that all is being done by the organizers to make it an unforgettable one. **RSIP Vision** will partner with **CVPR** for the 6th consecutive year to publish the **CVPR Daily**, the official magazine of CVPR. Do you want to stay in the loop? **Subscribe here to receive the CVPR Daily every day in real time and feel just like you were there!**

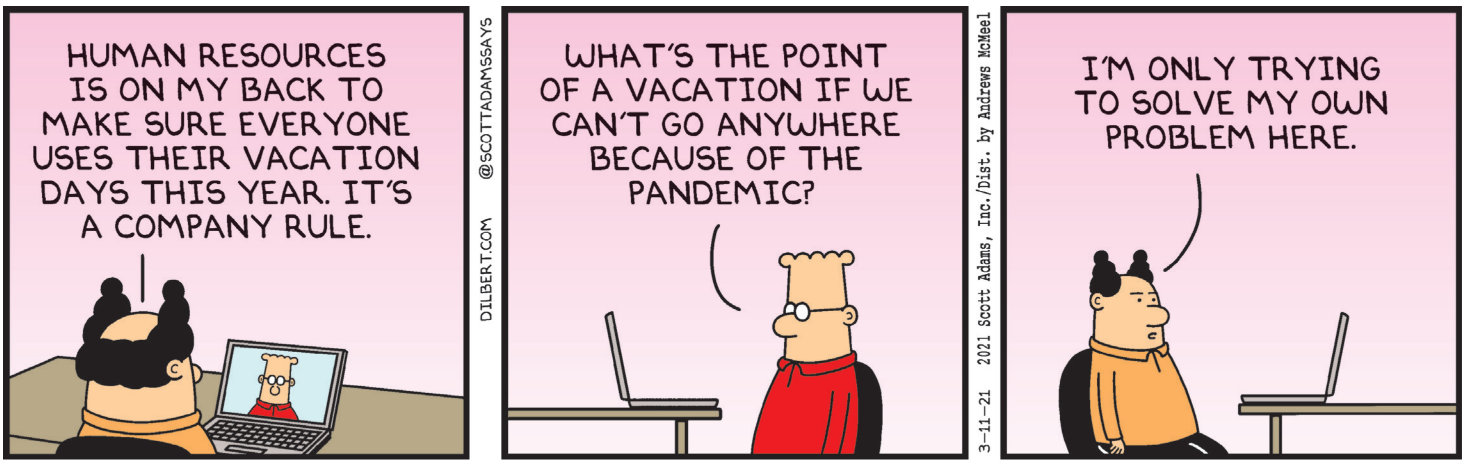
The next **Bay Vision Meetup**, which we sponsor, will also be unforgettable: **Lena Maier-Hein** will honor us with her talk: **"Does machine learning require domain experts?"** **Register here to obtain your personal link to access the webinar.**

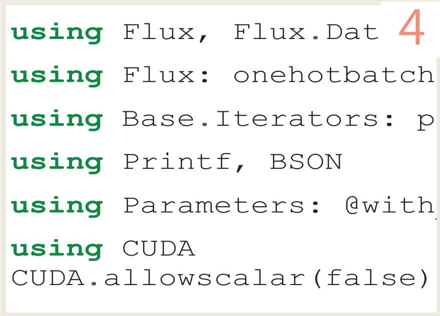
Don't miss the special section dedicated this month by **Computer Vision News** to a growing and exciting field: **Robot-Assisted Surgery**. We publish in this magazine some terrific (yet very different) articles, starting on page 14.
Enjoy the reading!

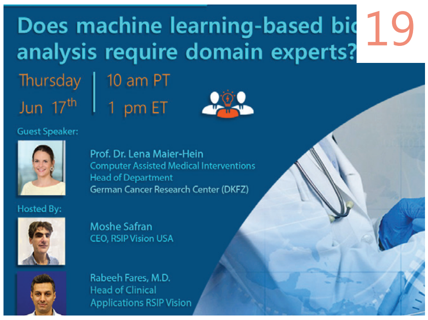
Ralph Anzarouth
Editor, **Computer Vision News**
Marketing Manager, **RSIP Vision**

Happy 5th Anniversary!!!

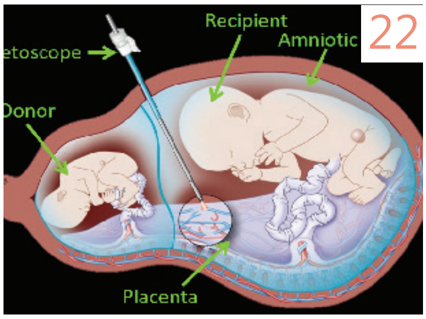
"I am a big fan of Computer Vision News!"
Andreas Maier
Head of the Pattern Recognition Lab
at the Friedrich-Alexander University Erlangen-Nürnberg

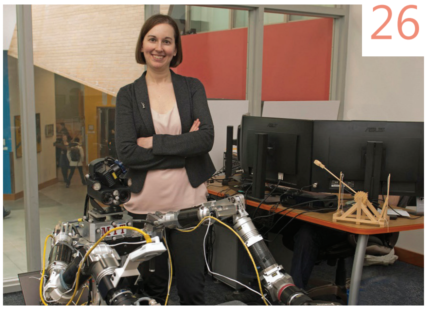


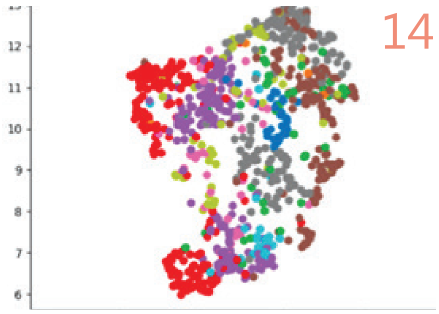

using Flux, Flux.Dat
using Flux: onehotbatch
using Base.Iterators: p
using Printf, BSON
using Parameters: @with
using CUDA
CUDA.allowscalar(false)


Does machine learning-based bi
analysis require domain experts?
Thursday
Jun 17th
10 am PT
1 pm ET
Guest Speaker:
Prof. Dr. Lena Maier-Hein
Computer Assisted Medical Interventions
Head of Department
German Cancer Research Center (DKFZ)
Hosted By:
Moshe Safran
CEO, RSIP Vision USA
Rabeeh Fares, M.D.
Head of Clinical
Applications RSIP Vision


CVPR
VIRTUAL JUNE 19-25


etoscope
Donor
Recipient
Amniotic
Placenta





ConvNets using Julia and the Flux machine learning framework

A technical review

IOANNIS VALASAKIS, KING'S COLLEGE LONDON

@WIZOFE

How is everyone doing? As one side of the world starts getting into the summer season and the coronavirus vaccinations are stronger than ever and create an immune world, let's remember that there are still places, like a big part of Africa that aren't vaccinated at all. It is a complex issue but definitely it is worth thinking about, especially when living in such a connected world.

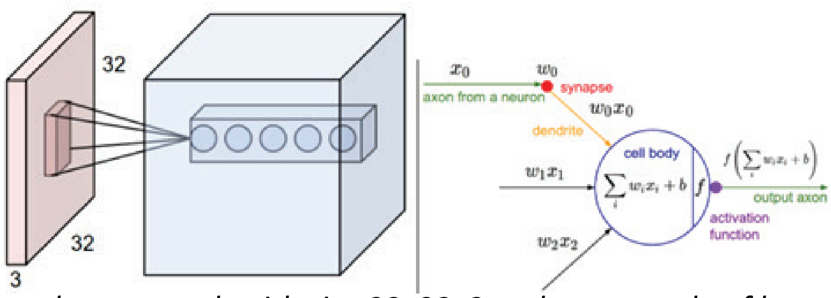
A few months ago, I created a tutorial to explore the Julia programming language. During the latest months, Julia is picking up in the scientific community by combining a very welcoming atmosphere and a great scientific place to work collaboratively. I would recommend using Pluto, as it's a great tool to experiment with.

Pluto is like the Jupyter Notebooks, just with a nicer interface and - most importantly - it re-executes the code that uses a variable, if that variable changes *anywhere* in the notebook!

ConvNets

Here is a video that will help with the visualisation and understanding of ConvNets. This is from MIT 6.S191 (which I totally recommend as an introduction to Deep Learning in general) and it is taught by Alexander Amini.

The main idea of a convolutional network for computer vision and recognition is that a ConvNet is composed of Layers. The communication protocol of each layer is also based on a simple idea: an input 3D volume is transformed to an output 3D volume using a differentiable function with optional parameters.



The input volume on red, with size 32x32x3 and an example of how neurons are composed in the first neural network layer [image from Stanford CS231n course].

Tutorial with Flux

This demo tutorial creates a Convolutional Neural Network (ConvNet) to classify the MNIST dataset. The architecture is made of three feature detection layers (Conv -> ReLU -> MaxPool) and followed by a final dense layer that classifies MNIST handwritten digits. For 20 epochs this has about 99% accuracy, which is impressive.

The model is also saved in the file `mnist_conv.bson`. and it shows how to create a model but also train, save and early exit, if that is needed. The following packages are needed:

```
using Flux, Flux.Data.MNIST, Statistics
using Flux: onehotbatch, onecold, logitcrossentropy
using Base.Iterators: partition
using Printf, BSON
using Parameters: @with_kw
using CUDA
CUDA.allowscalar(false)
```

With default values for learning rate, batch size, number of epochs, and path for saving the file `mnist_conv.bson`:

```
@with_kw mutable struct Args
    lr::Float64 = 3e-3
    epochs::Int = 20
    batch_size = 128
    savepath::String = "./"
end
```

6 Computer Vision Tool

To train our model, we need to bundle images together with their labels and then group them into mini-batches (makes the training process faster). We define the function make_minibatch that takes as inputs the images (X) and their labels (Y) as well as the indices for the mini-batches (idx):

```
function make_minibatch(X, Y, idxs)
    X_batch = Array{Float32}(undef, size(X[1])..., 1, length(idxs))
    for i in 1:length(idxs)
        X_batch[:, :, :, i] = Float32.(X[idxs[i]])
    end
    Y_batch = onehotbatch(Y[idxs], 0:9)
    return (X_batch, Y_batch)
end
```

make_minibatch takes the following steps:

- Creates the X_batch array of size 28x28x1x128 to store the mini-batches.
- Stores the mini-batches in X_batch.
- One hot encodes the labels of the images.
- Stores the labels in Y_batch.

The test data are loaded from the MNIST database and this creates an array with the indices of the train images corresponding to each mini-batch.

```
function get_processed_data(args)
    # Load labels and images
    train_labels = MNIST.labels()
    train_imgs = MNIST.images()
    mb_idxs = partition(1:length(train_imgs), args.batch_size)
    train_set = [make_minibatch(train_imgs, train_labels, i) for i in mb_idxs]

    # Prepare test set as one giant minibatch:
    test_imgs = MNIST.images(:test)
    test_labels = MNIST.labels(:test)
    test_set = make_minibatch(test_imgs, test_labels, 1:length(test_imgs))
    return train_set, test_set

end
```

Here the model is built by creating three convolution layers and one classification layer. As the images are grayscale, only one channel is used (instead of three). When combined together, the convolutional layer structure would look like Conv(kernel, input_channels => output_channels, ...). To reduce the size of each image a ReLU and MaxPool operations are performed. The classification layer outputs a vector of 10 dimensions (a dense layer), which is the number of predicted classes.

```
function build_model(args; imgsize = (28,28,1), nclasses = 10)
    cnn_output_size = Int.(floor.([imgsize[1]/8,imgsize[2]/8,32]))

    return Chain(
    # First convolution, operating upon a 28x28 image
    Conv((3, 3), imgsize[3]=>16, pad=(1,1), relu),
    MaxPool((2,2)),

    # Second convolution, operating upon a 14x14 image
    Conv((3, 3), 16=>32, pad=(1,1), relu),
    MaxPool((2,2)),

    # Third convolution, operating upon a 7x7 image
    Conv((3, 3), 32=>32, pad=(1,1), relu),
    MaxPool((2,2)),

    # Reshape 3d tensor into a 2d one using `Flux.flatten`, at this point it
should be (3, 3, 32, N)
    flatten,
    Dense(prod(cnn_output_size), 10))
end
```

To chain the layers of a model we use the Flux function Chain. It enables us to call the layers in sequence on a given input. Also, we use the function flatten to reshape the output image from the last convolution layer. Finally, we call the Dense function to create the classification layer.
Before training our model, we need to define a few functions that will be helpful for the process:

- `augment` augments the data by adding gaussian random noise to our image to make it more robust:

```
augment(x) = x .+ gpu(0.1f0*randn(eltype(x), size(x)))
```

- `anynan` checks whether any element of the params is NaN or not:

```
anynan(x) = any(y -> any(isnan, y), x)
```

- `accuracy` computes the accuracy of our ConvNet:

```
accuracy(x, y, model) = mean(onecold(cpu(model(x))) .== onecold(cpu(y)))
```

Finally, we define the train function:

```
function train(; kws...)
    args = Args(; kws...)

    @info("Loading data set")
    train_set, test_set = get_processed_data(args)

    # Define our model.  We will use a simple convolutional architecture with
    # three iterations of Conv -> ReLU -> MaxPool, followed by a final Dense
layer.
    @info("Building model...")
    model = build_model(args)

    # Load model and datasets onto GPU, if enabled
    train_set = gpu.(train_set)
    test_set = gpu.(test_set)
    model = gpu(model)

    # Make sure our model is nicely precompiled before starting our training
loop
    model(train_set[1][1])

    # `loss()` calculates the crossentropy loss between our prediction `y_hat`
    # (calculated from `model(x)`) and the ground truth `y`.  We augment the
data
    # a bit, adding gaussian random noise to our image to make it more robust.
    function loss(x, y)
        x̂ = augment(x)
        ŷ = model(x̂)
        return logitcrossentropy(ŷ, y)
    end

    # Train our model with the given training set using the ADAM optimizer and
    # printing out performance against the test set as we go.
    opt = ADAM(args.lr)

    @info("Beginning training loop...")
    best_acc = 0.0
    last_improvement = 0
    for epoch_idx in 1:args.epochs
        # Train for a single epoch
        Flux.train!(loss, params(model), train_set, opt)

        # Terminate on NaN
        if anynan(Flux.params(model))
            @error "NaN params"
            break
        end

        # Calculate accuracy:
        acc = accuracy(test_set..., model)
```

```
        @info(@sprintf("[%d]: Test accuracy: %.4f", epoch_idx, acc))
        # If our accuracy is good enough, quit out.
        if acc >= 0.999
            @info(" -> Early-exiting: We reached our target accuracy of 99.9%")
            break
        end

        # If this is the best accuracy we've seen so far, save the model out
        if acc >= best_acc
            @info(" -> New best accuracy! Saving model out to mnist_conv.bson")
            BSON.@save        joinpath(args.savepath,        "mnist_conv.bson")
params=cpu.(params(model)) epoch_idx acc
            best_acc = acc
            last_improvement = epoch_idx
        end

        # If we haven't seen improvement in 5 epochs, drop our learning rate:
        if epoch_idx - last_improvement >= 5 && opt.eta > 1e-6
            opt.eta /= 10.0
            @warn(" -> Haven't improved in a while, dropping learning rate to
$(opt.eta)!")

            # After dropping learning rate, give it a few epochs to improve
            last_improvement = epoch_idx
        end

        if epoch_idx - last_improvement >= 10
            @warn(" -> We're calling this converged.")
            break
        end
    end
end
```

`train` calls the functions that were defined above and it trains our model. It stops when the model achieves 99% accuracy (early-exiting) or after performing 20 steps. More specifically, it performs the following steps:

- Loads the MNIST dataset.
- Builds our ConvNet model (as described above).
- Loads the train and test data sets as well as our model onto a GPU (if available).
- Defines a `loss` function that calculates the crossentropy between our prediction and the ground truth.
- Sets the [ADAM optimiser] to train the model with learning rate `args.lr`.
- Runs the training loop. For each step (or epoch), it executes the following:
    - Calls `Flux.train!` function to execute one training step.
    - If any of the parameters of our model is `NaN`, then the training process is terminated.

- Calculates the model accuracy.
- If the model accuracy is >= 0.999, then early-exiting is executed.
- If the actual accuracy is the best so far, then the model is saved to `mnist_conv.bson`. Also, the new best accuracy and the current epoch is saved.
- If there has not been any improvement for the last 5 epochs, then the learning rate is dropped and the process waits a little longer for the accuracy to improve.
- If the last improvement was more than 10 epochs ago, then the process is terminated.

Finally, to test our model we define the `test` function:

```
function test(; kws...)
    args = Args(; kws...)

    # Loading the test data
    _,test_set = get_processed_data(args)

    # Re-constructing the model with random initial weights
    model = build_model(args)

    # Loading the saved parameters
    BSON.@load joinpath(args.savepath, "mnist_conv.bson") params

    # Loading parameters onto the model
    Flux.loadparams!(model, params)

    test_set = gpu.(test_set)
    model = gpu(model)
    @show accuracy(test_set...,model)

end
```

`test` loads the MNIST test data set, reconstructs the model, and loads the saved parameters (`in  mnist_conv.bson`) onto it. Finally, it computes our model's predictions for the test set and shows the test accuracy (around 99%).

Wrap up

I hope you had fun with this article! Feel free to always ask any questions or send through your recommendations.

The resources and contributors for the Flux community: Convolutional Neural Networks (CNNs / ConvNets). Convolutional Neural Networks Tutorial in PyTorch. – Elliot Saba, Adarsh Kumar, Mike J Innes, Dhairya Gandhi, Sudhanshu Agrawal, Sambit Kumar Dash, fps.io, Carlo Lucibello, Andrew Dinhobl, Liliana Badillo

Enjoy your month and keep having fun 😊


CVPR DAILY
VIRTUAL JUNE 14-19
Computer Vision and Pattern Recognition
2020
Tuesday

Editorial
with Silvio Savarese
Program Co-chair

Workshop
Learning 3D
Generative Models

Women in Computer Vision
Johanna Wald

Picks of the Day
Francesca Babiloni

Exclusive interview
Trevor Darrell

Oral Presentation
Boyang Deng

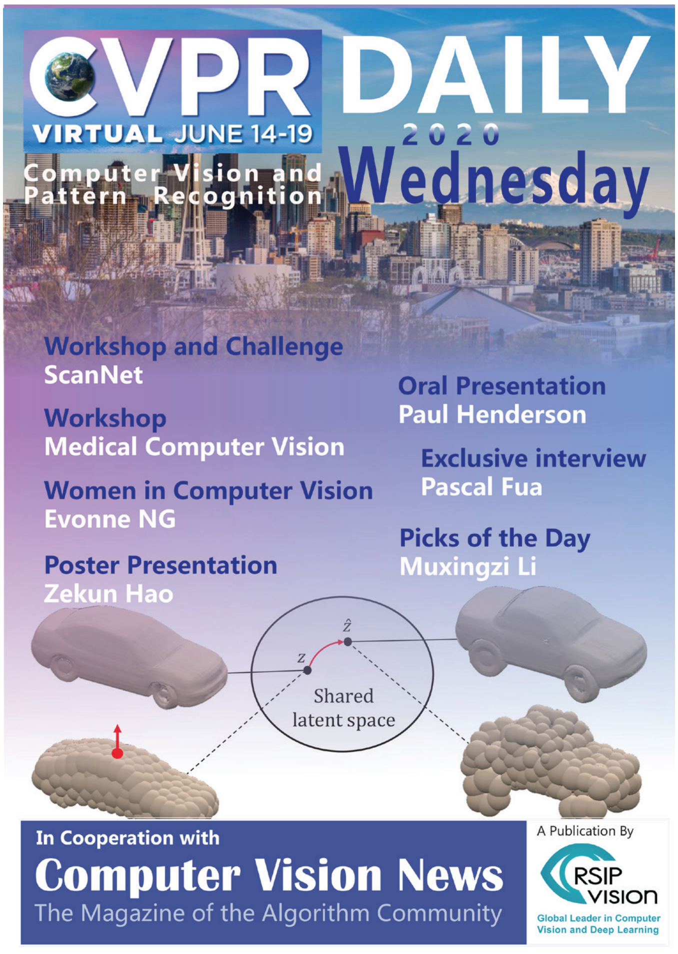
Workshop Preview
CVsports

Poster Presentation
Despoina Paschalidou

Image Wu, Nozawa, Perteneder, Koike

In Cooperation with
Computer Vision News
The Magazine of the Algorithm Community

A Publication By
RSIP VISION
Global Leader in Computer Vision and Deep Learning


CVPR DAILY
VIRTUAL JUNE 14-19
Computer Vision and Pattern Recognition
2020
Wednesday

Workshop and Challenge
ScanNet

Oral Presentation
Paul Henderson

Workshop
Medical Computer Vision

Exclusive interview
Pascal Fua

Women in Computer Vision
Evonne NG

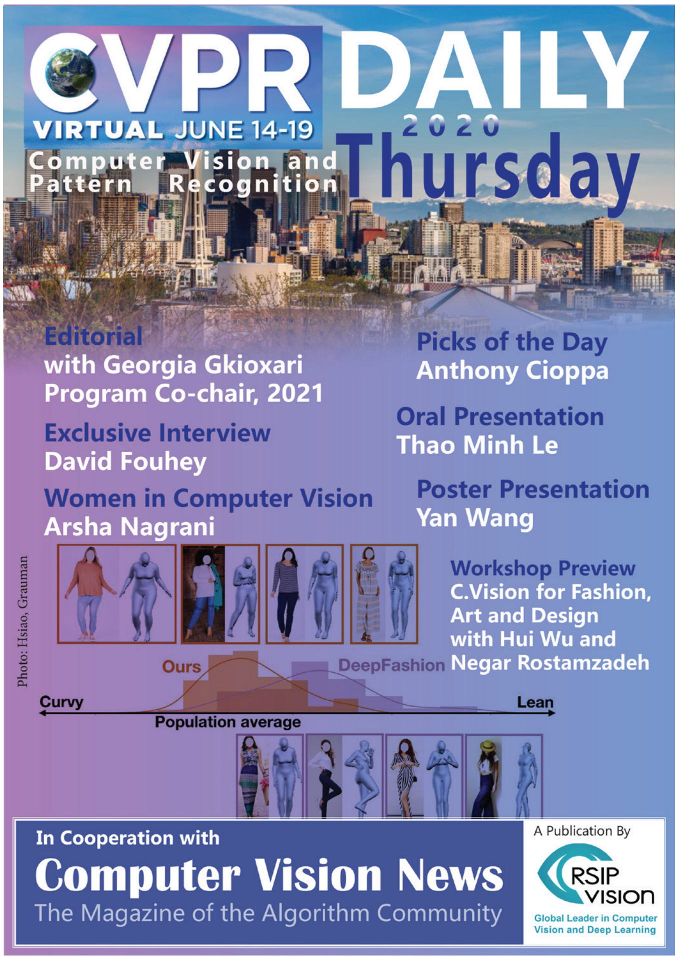
Picks of the Day
Muxingzi Li

Poster Presentation
Zekun Hao

Shared latent space

In Cooperation with
Computer Vision News
The Magazine of the Algorithm Community

A Publication By
RSIP VISION
Global Leader in Computer Vision and Deep Learning


CVPR DAILY
VIRTUAL JUNE 14-19
Computer Vision and Pattern Recognition
2020
Thursday

Editorial
with Georgia Gkioxari
Program Co-chair, 2021

Picks of the Day
Anthony Cioppa

Oral Presentation
Thao Minh Le

Exclusive Interview
David Fouhey

Women in Computer Vision
Arsha Nagrani

Poster Presentation
Yan Wang

Workshop Preview
C.Vision for Fashion, Art and Design
with Hui Wu and Negar Rostamzadeh

Photo: Hsiao, Grauman

Ours
DeepFashion
Curvy
Population average
Lean

In Cooperation with
Computer Vision News
The Magazine of the Algorithm Community

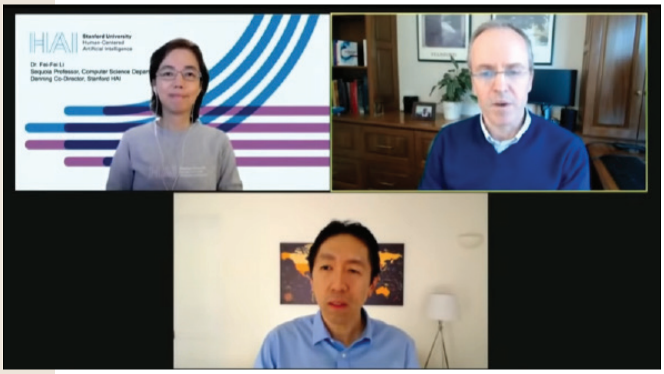
A Publication By
RSIP VISION
Global Leader in Computer Vision and Deep Learning

AI SPOT LIGHT NEWS

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

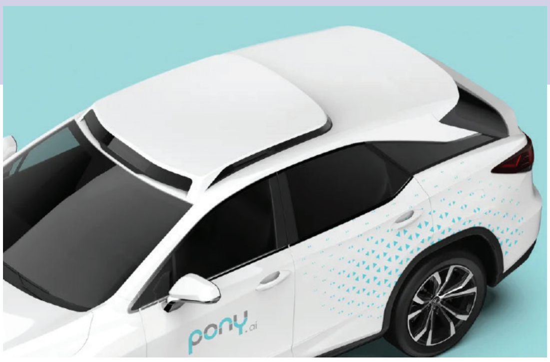
The Future of AI in Healthcare (according to Fei Fei and Ng)

Two AI luminaries, **Fei-Fei Li** and **Andrew Ng** got together on YouTube, to warn us that it might take much longer than thought to see the results of AI progress in healthcare! Apparently, they compare the current state of AI in MedTech to the excitement around and investment in autonomous vehicle technologies a few years ago, as successful implementations of AI-based products and services in healthcare may not be just around the corner. With all due respect, we can witness that there are scores of AI-based solutions which are already implemented in doctors' practices and in the hospital, including in the OR! Read More



Pony.ai Unveils its Next-gen Robotaxi with LIDAR from Luminar

And since we mentioned autonomous vehicle technologies, let's talk a bit about **robotaxis**. We are all used to see big **LIDAR** devices on the roof of most autonomous vehicles. Well, **Pony.ai's next-generation robotaxi** is distinctive because it appears to be missing the cone-shaped LIDAR sensor. They are indeed teaming up with **Luminar**, to use the fast-growing LIDAR company's sleek new sensors that are more flush with the vehicle's roof. The new vehicles will not be up and running until 2022, but Pony.ai founder and CEO **James Peng** said preparation was already underway for mass production. Read More

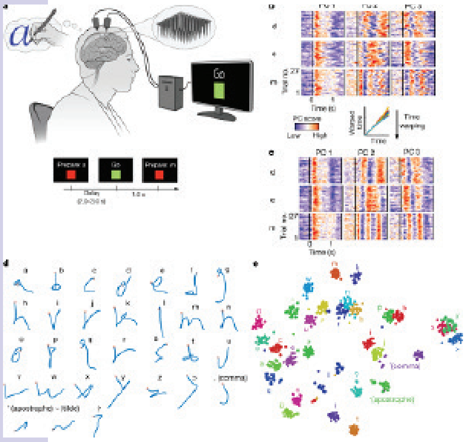


insideBIGDATA
Your Source for Machine Learning

AI Gets the Glory but ML is Quietly Making Fortunes

While **Artificial Intelligence** is no doubt booming in so many fields, we must admit that there is a lot of hype around it, especially in the media. This nice article claims that AI is talked about in inflated terms, because we still lack a good definition of what it means for a computer to be "intelligent." On the other hand, Machine Learning is already changing the world and, like all technology eventually, is getting cheaper and more accessible all the time. In other words, **Machine Learning** is already being used at its best to **augment human capabilities**, not replace them like AI is expected to do. Read More

[Artificial Intelligence Translates Mental Handwriting into Text]



Back to the amazing things that AI already does. Apparently, an **AI software**, paired with a **brain-computer interface**, decodes neural signals to convert mental handwriting to text. This is per se amazing; but you will ask, what is this good for? That's the point: it enables communication for patients with spinal-cord injuries, strokes, and other dire conditions. According to a study conducted at **Stanford** and published in **Nature**, a man with full-body paralysis was able to communicate at a speed of 18 words per minute. Compare that to able-bodied people who can type about 23 words per minute on a smartphone! [Read More]

[Seeing-Eye Shoes Pair Computer Vision with Haptic Feedback]

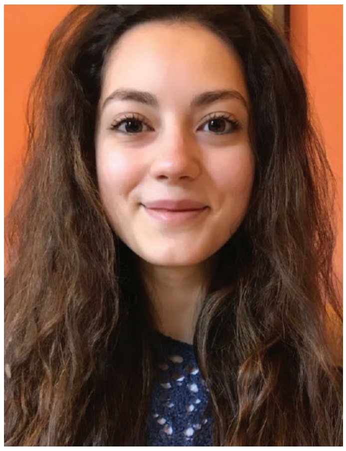
Something cute now: if you wonder why ultrasonic or time-of-flight sensors have been placed on the wrist or head or anywhere else, except shoes - here, somebody did it. The purpose is of course to send pulses and measure the time it takes for the signal to bounce off the object and come back, in order to give visually impaired people a prompt notification of obstacles. They get the info through haptic feedback in the shoes as well as an audible phone notification via Bluetooth. The company is called InnoMake and their shoe was designed in partnership with Austria's Graz University of Technology.
[Read More]

What can AI researchers do to help prevent Lethal Autonomous Weapons?

A terribly serious subject now: more and more scientists and researchers question the potential (or actual) use of their AI as **lethal weapons by governments** and/or by organized crime. It is not difficult to imagine how dangerous could be autonomous weapons sent by armies, dictatures and actually anybody on enemy grounds. Shifting the burden of conflict even further on civil populations using autonomous weapons is a danger that a panel of scholars decided to address in the video here. If you are concerned too, have a look at the [Stop Killer Robots website], where things are explained much better than I can do.

Multimodal and Self-Supervised Representation Learning for Automatic Gesture Recognition in Surgical Robotics

by Marica Muffoletto



Every month, Computer Vision News selects a research paper to review. This month we review "Multimodal and Self-Supervised Representation Learning for Automatic Gesture Recognition in Surgical Robotics". We are indebted to the authors (Aniruddha Tamhane, Jie Ying Wu, Mathias Unberath) for allowing us to use their images to illustrate this review. You can find their paper at this link.

Today we'll examine an extremely recent work by the trio of researchers from the John Hopkins University in USA who is looking at the exciting world of surgical robotics. Their paper focuses on developing a method to distinguish steps of the surgical process. They worked on a self-supervised multi-modal representation learning algorithm, trained on a combination of videos of surgeries and kinematics data. This model is able to learn task-agnostic surgical gesture representation from both these sources that can generalise well across multiple tasks.

The idea of this work seems to come from a well-posed analysis of the state-of-the-art methods with their drawbacks and strengths. Majority of SOA methods considered are found to analyse surgical data that are based on task-specific, supervised learning from a single modality. The fact that they are trained with specific tasks leads to very narrow learning on surgical processes. The choice of using a supervised method of course enforces dependence on expert annotations, which can be tedious and defeat the purpose of having an automatic tool. Finally, some of them usually ignore multiple modalities of information, which prevents from learning generalizable, feature-rich representations.

From these weaknesses, the main objectives of this paper are defined: these include the development of a DL architecture which is self-supervised and effectively learns gesture representations, the quantitative demonstration of SOA accuracy in gesture/skill-recognition, and the visualisation of these learnt representation and the formation of semantically meaningful clusters.

From reviewing recent papers in the field of surgical robotics and computer vision, the authors also analyse and get inspired by the following points:

1. From SOA in Supervised Surgical gesture recognition, they borrow the notion that optical flow is an important source of information for learning to classify surgical gestures. This knowledge is fundamental for the decision to extract optical flow from the data as a source of domain-independent visual information. Moreover, from reviewing papers in this field, they also derive the use of integrating multimodal information to improve supervised gesture recognition results.
2. From papers in Unsupervised Surgical gesture recognition, they analyse the possibility to capture the latent information for surgical gesture recognition using RNNs in an unsupervised manner, and the proof that these embeddings naturally cluster corresponding to distinct higher-level activities. This turns out to be a very important factor in this research as well, even if RNNs are used instead of CNNs.
3. The field of Video activity recognition helps them shaping their final architecture, based on an encoder-decoder network. This is after reviewing different approaches with parallel CNN-based video and optical flow streams or visual attention-based models for surgical activity recognition.
4. From a Multimodal self-supervised learning study, they also get another example of corresponding embeddings for audio and video which cluster close to each other.

To combine this knowledge and develop the novel algorithm discussed here, the main data is taken from the JIGSAWS dataset and combines a series of annotated clips of surgical activity (further subdivided in the three categories reported below) and kinematics in the shape of a 76-dimensional vector that includes the x,y,z coordinates of the left and right tool tips, the corresponding linear and angular velocities, the rotation matrix and the gripper angle velocities.

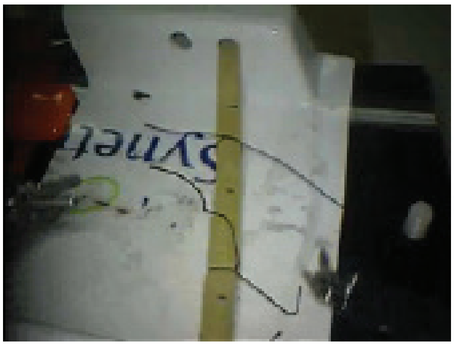


Figure 3: suturing

Figure 2: needle passing

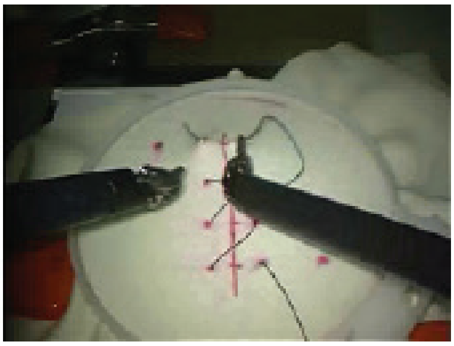
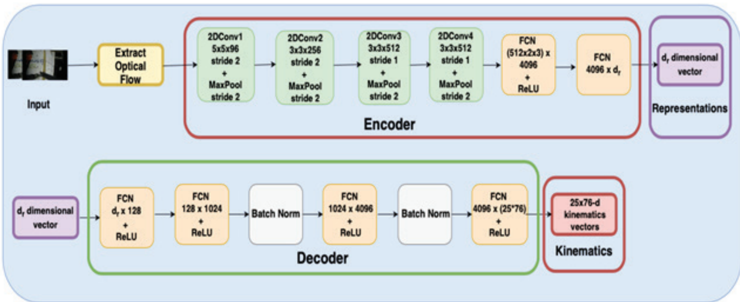


Figure 3: suturing

These data are paired to train an encoder-decoder model.
The choice of this specific architecture comes after the review of RNNs, CNNs and other architectures in similar applications, but it is also motivated by the consideration that training a model on the alignment-based task which learns the one-to-one mapping from video to kinematics might be more complex than employing an encoder-decoder task with the objective to extract the corresponding kinematics vectors from the optical flows. Hence, the final training objective in this deep learning algorithm is to minimise the information loss (chosen as L2 norm of the difference) between the decoded representations and the kinematics. This is expressed in the function below.

$$\min_{\theta,\phi} \frac{1}{n}\sum_{i=1}^{n} ||\mathcal{D}(r(T(\mathcal{V}_i);\theta);\phi) - \mathcal{K}_i||_2^2$$

Here, V_i and K_i are the two sources of data that are input to the model, namely videos and kinematics, T is a transformation on V that extracts the optical flow and r corresponds to the encoder function, parametrized by θ, while the decoder function D is parametrized by ϕ.
After the model is trained, the encoded representations in r(T(V_i);θ) should retain all the critical information such as: the exact surgical gesture, the identity of the surgeon and the skill with which the segment of surgery was performed.



The model is trained with the encoder-decoder architecture shown above, which includes 1) an optical-flow extraction step which filters out domain-specific information such as video quality, contrast, details about the surgical instruments, 2) an encoder made of 2D CNNs which encodes information from the videos (optical flow) into the representation and parses them to 3) a decoder built as a simple FCN with ReLU activations, which is kept shallow to maximize information retention in the representations yielded by the encoded network. This outputs the kinematics that are compared through an MSE loss with the ground truth vectors provided in the JIGSAWS dataset.

Hence, the final training protocol includes dividing the frames in each video clip according to the gestures, sampling 25 frames, encoding a gesture context (1.67s), extracting the optical flow. An end-to-end deep encoder-decoder network is trained, and the information encoded in the representations is shown using a U-MAP algorithm which reduces the dimensionality to a 2D plane. Then, the representations below are shown to cluster into two distinct skill-based clusters which correspond to beginners and experts. From these representations comes the first finding of the paper: **each gesture has a unique representation depending on whether it has been performed by an expert or a beginner.**

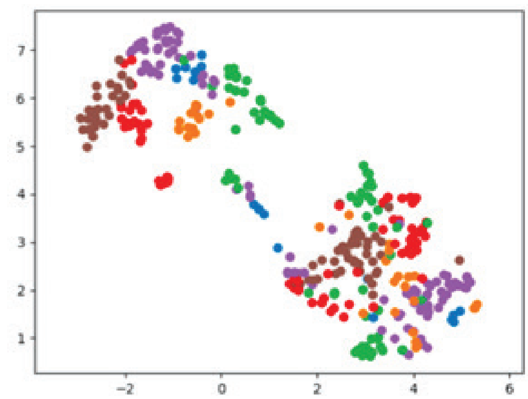


Fig. 2: Knot Tying representations (gesture)

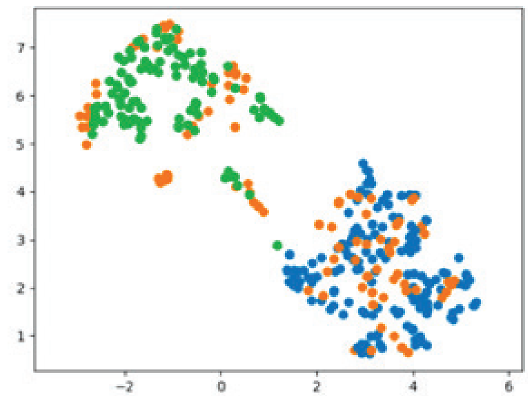


Fig. 3: Knot Tying representations (skill)

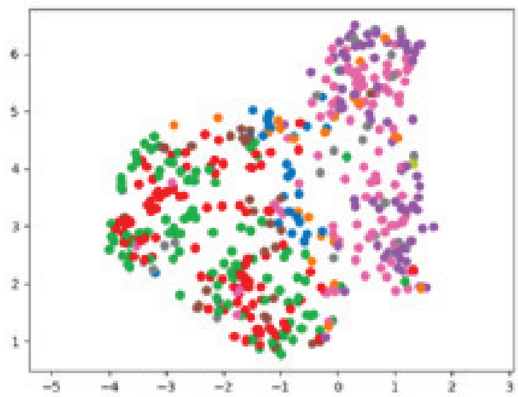


Fig. 4: Needle Passing representations (gesture)

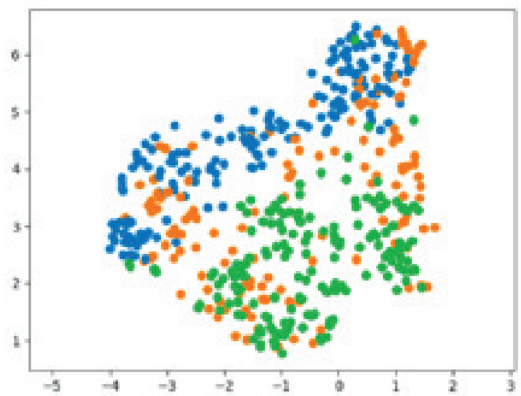


Fig. 5: Needle Passing representations (skill)

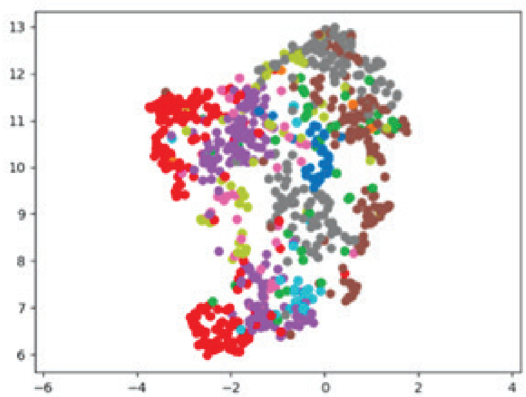


Fig. 6: Suturing representations (gesture)

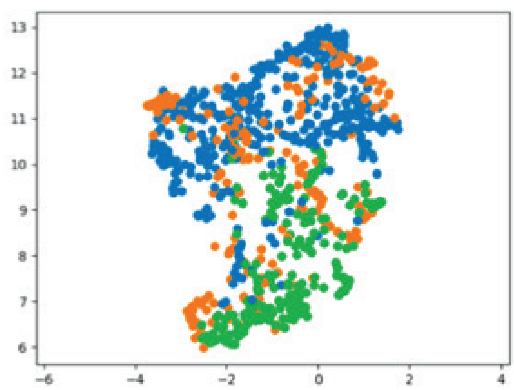


Fig. 7: Suturing representations (skill)

Three other experiments are led to evaluate the efficacy of the representations. These are measured using the accuracy, precision, recall and F-1 classification metrics.

The first one is a skill classification experiment done by training a 3-class gradient boosting-based classifier with the representations obtained previously by the encoder-decoder architecture. The result of this experiment indicates that **there is a significant retention of information related to surgeon skill** and that the inter-class variability between tasks remains low with highest performance across the suturing task (0.812 ± 0.0228) and lowest across the Knot-Tying dataset (0.768 ± 0.0303). Since a major portion of the classification error is contributed by surgeons having an “intermediate” skill, this also shows that there is a grey line between expert and beginner levels.

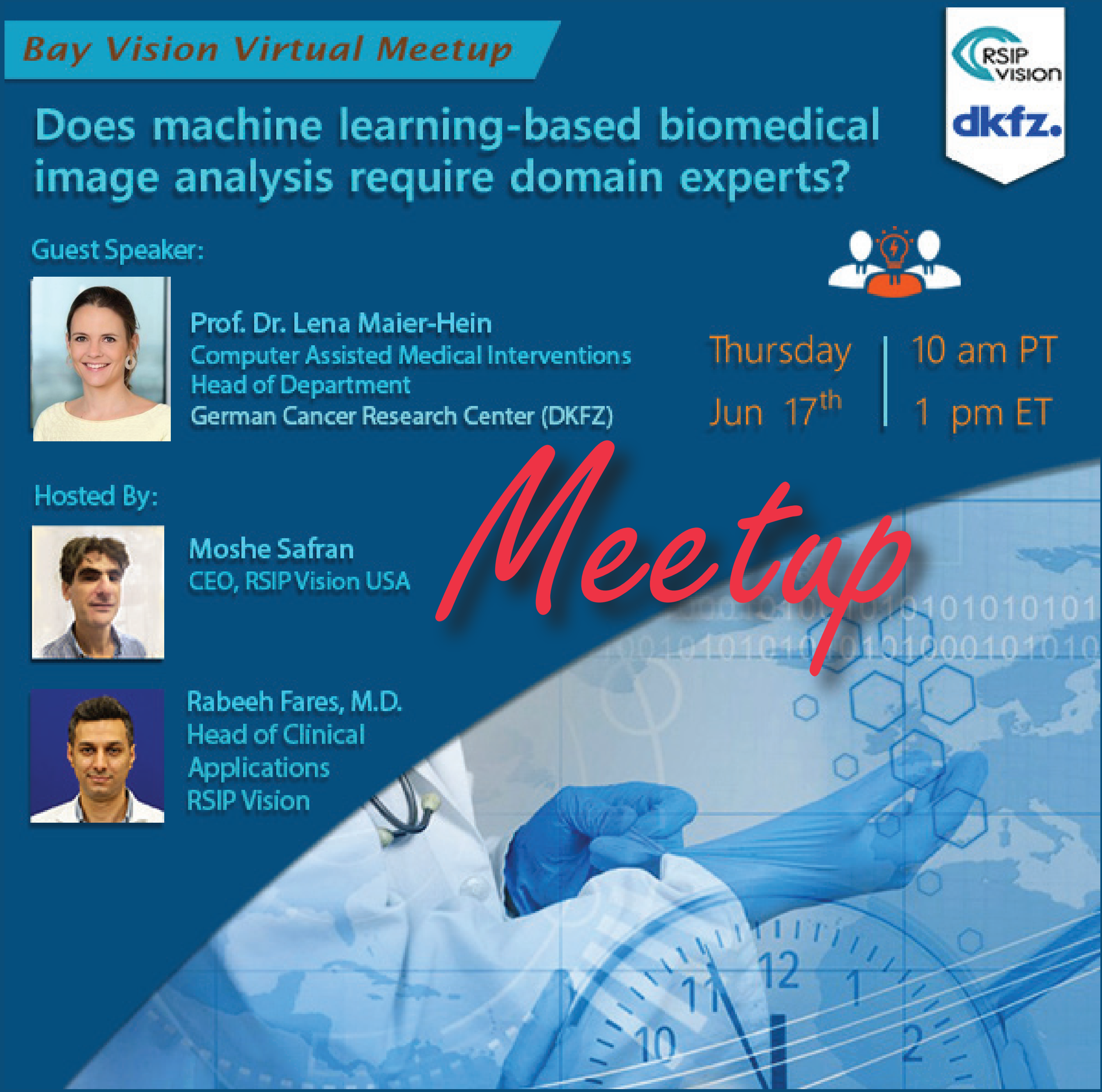
The second experiment is similar to the previous one, except that it focuses on gesture classification. This is also done by training a 3-class gradient boosting-based classifier with the representations. Promising results from this task are also found with an average accuracy across datasets of 0.745.

Finally, the authors try a transfer learning-based gesture recognition task. This aims at classifying gestures on two datasets based on the representations from the other one. This is repeated for all three combinations of representations used. Although a slight decrease in performance is found, this experiment concludes that **the representations are sufficiently robust across multiple task**s and might facilitate transfer learning.

Here we are, at the end of another paper review, which opens up this month’s section on surgical robotics. This is a demonstration of interesting research going on in this field and of a novel application of an encoder-decoder architecture. We hope to keep discovering together more papers using deep learning architectures for novel applications and, as suggested by the authors, we share their same curiosity to see one day soon **a universal technology that tracks surgical progress in real-time, giving feedback regarding possible mistakes, surgical scene depth, next gesture suggestion etc. with high accuracy.**

Join a great Meetup with Lena Maier-Hein, a MICCAI Fellow and one of the most authoritative scientists in the field of Medical Image Analysis!

don't miss out!



Computer Vision in Robotic-Assisted Surgeries

WHAT DOES THE FUTURE HOLD?

As AI technology improves and becomes smarter and more autonomous, we expect more assimilation of this technology in **Robotic-Assisted Surgeries (RAS)**. In a similar manner to other fields (e.g. laparoscopic surgeries), RAS will grow to be more interactive, and depend on computer vision modules to provide better and safer procedures. The robots will **autonomously perform operations**, notify the staff of procedural progress, and alert when the procedure does not progress as expected. Below we describe several **computer-vision features** which can make this ideal RAS come to life.

There are several areas where AI can improve and change RAS, these areas include the surgery planning, the procedure itself, and post-op surgical analysis.

Prior to the surgery, it is essential that the surgeon is **well trained** on the procedure, as well as prepared for the specific upcoming patient.
Acquiring the skills for proper RAS performance requires intensive

training, which is a costly and time-consuming process. Designing a **simulator for RAS** can overcome the difficulties of training. This simulator will combine real simulations using **virtual reality (VR) technology** and haptic feedback to mimic RAS as accurately as possible.
This simulator can also be used to prepare for specific procedures, using data from previous similar procedures done by that surgeon as well as data that is relevant to the patient specific anatomy. Using the capability to automatically analyze the workflow in RAS, we can easily procure **multiple videos of relevant procedures**, without the need to sift through the data. These can be used for training of surgeons, as well as training of machines to perform these tasks.

During the procedure, the robot can take on tasks from the surgeon.

Navigation is an essential part in any RAS. It is not always a straight-forward task, and different modules (endoscopies, laparoscopies) have different challenges in navigation. **Deep learning algorithms** can learn the desired paths and give real-time navigational instructions.

During specific steps like **biopsy** or **tumor removal**, often additional real-time imaging modalities are used, such as US or fluoroscopy. These modalities are limited and have relatively poor image quality. Usually, a better-quality pre-op CT or MRI scan is available. **Fusion of the real-time modality with the higher quality images** can utilize the advantages of each and provide better guidance during the procedure.
Simultaneously, the system can be trained to **recognize and classify** abnormal changes in anatomy, warn of areas which should be avoided and recommend preferred routes. The robot can highlight specific anatomies in the body, recommend

which tools to use, what procedure is coming up, and give specific warnings relevant to the patient/process/ stage of the RAS. Additionally, it can provide **diagnostic information and treatment suggestions**. This robot can shorten the procedural time and significantly increase accuracy, as well as reduce the need for additional procedures.

Further down the road, this ability can be extended to provide autonomous surgery. Similarly, to the concept of an autonomous car, the robot can perform **the entire surgery** based on the clinical input, while the surgeon supervises in case of need.

Another aspect of quality training is reviewing procedures post-op to **determine success and find weaknesses**. Using the workflow analysis tool this step can be conducted quickly and efficiently, saving time for the surgeon.

Overall, harnessing the power and advancements of AI and computer vision will significantly improve RAS. AI can improve performance during the procedure, and preparation, training, and post-op analysis can benefit from it. The assimilation of AI in Robotic-Assisted Surgery brings much hope and excitement to this field, and we look forward to it.

FetReg: Placental Vessel Segmentation and Registration in Fetoscopy

FetReg: Placental Vessel Segmentation and Registration in Fetoscopy is a sub-challenge of the popular endoscopic vision challenge EndoVis, which will be held at MICCAI 2021 in Strasbourg later this year. Sophia Bano, Sara Moccia and Alessandro Casella are three of its co-organizers and they speak to us about their exciting plans.

The EndoVis challenge has been a regular feature at **MICCAI** since 2015. Back then, there was a lack of good, annotated data for computer-assisted surgery. Since that time, the field has advanced, with new technology emerging in the operating room, including **devices, sensors, and robotic systems, generating a huge amount of data**. However, much of this data is still processed qualitatively by the surgeon, and there continues to be a need for common annotated datasets which enable consistent evaluation and validation of endoscopic vision algorithms.

EndoVis addresses this gap and encourages researchers in the field to work together to provide high quality data and new methodologies that can be translated into surgical practice to improve interventional healthcare. This is the main rationale behind surgical data science.

Sophia is a Senior Research Associate at **Wellcome / EPSRC Centre for Interventional and Surgical Sciences (WEISS) at University College London**, working in the fields of surgical vision and surgical data science.

Sophia Bano

Alessandro Casella

Sara Moccia

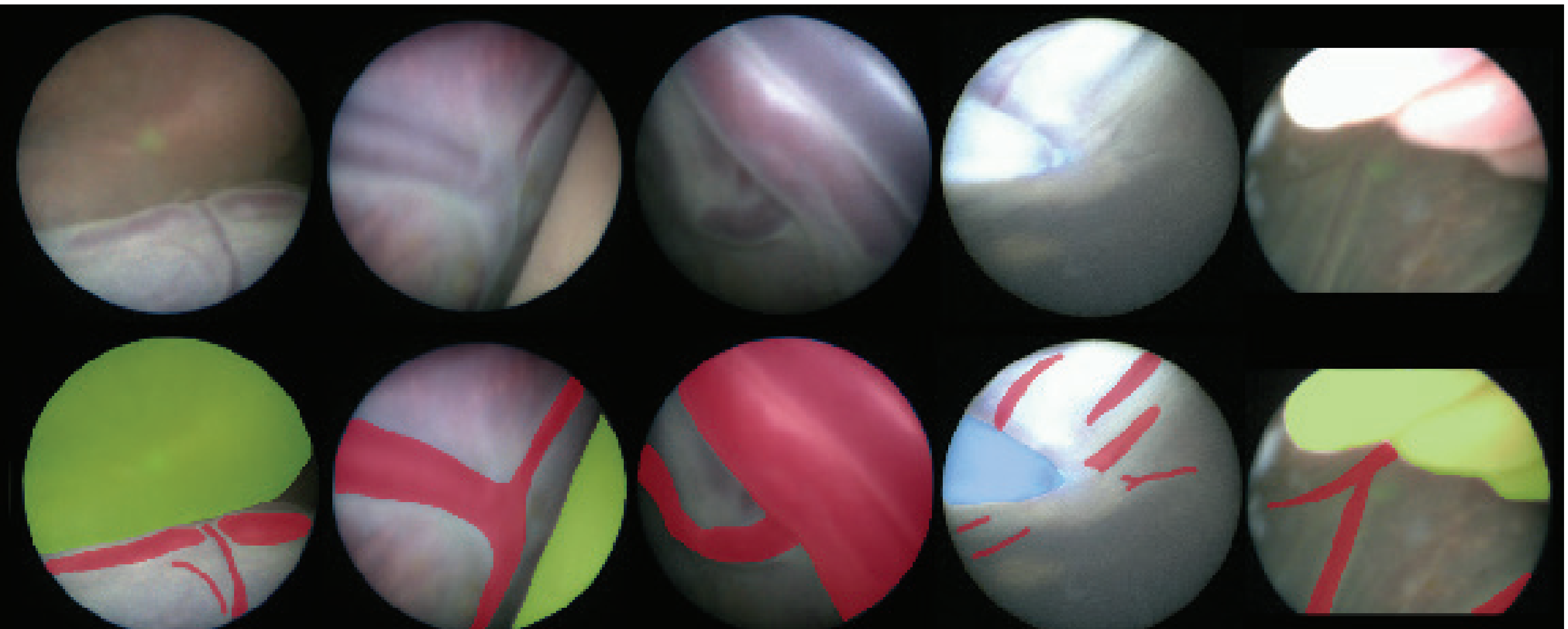
"With surgical vision, we are looking at multiple anatomies. Each imaging modality that we look at has a different anatomy," she explains. *"For example, if we are talking about upper GI, it has a totally different anatomy to colonoscopy, fetoscopy, or laparoscopy. For each of these, there is data available, which may or may not be annotated, but for which we want to apply computer-assisted intervention. Every year, EndoVis has a number of sub-challenges based on different modalities and procedures with new problems for the participants have to solve."*

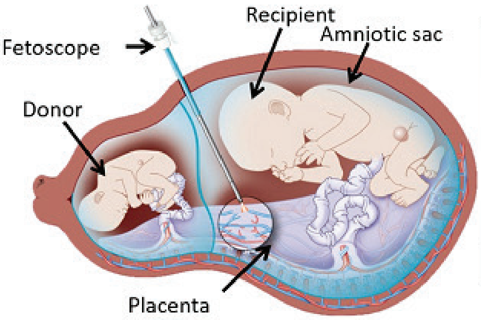
This year, the **FetReg sub-challenge** explores placental vessel segmentation and registration for mosaicking in clinical fetoscopy for the treatment of **Twin-to-Twin Transfusion Syndrome (TTTS)**. The team hope it will advance work to enhance surgeons' vision by expanding the fetoscopic field of view and providing better visualization of the vessel map.

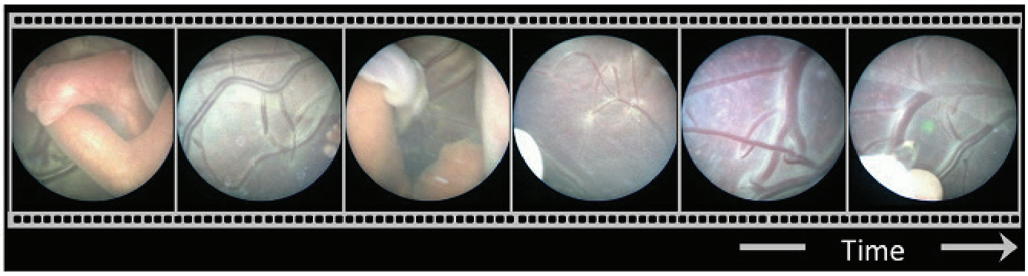
Alessandro is a PhD student at the **Italian Institute of Technology and Politecnico di Milano**. Computer-assisted surgery and surgical data science are his main interests, and his PhD research is focused on computer vision and deep learning techniques for computer-assisted fetal surgery, so he has already done some work in this area.

"My first work on TTTS was the identification of an anatomical structure that could be useful for a surgeon to ***orientate and navigate inside the placenta,****"* he tells us. *"Fetal surgery is very poorly explored. We hope with this work that we can push things forward so that in a few years we could have a complete framework for TTTS and be able to extend that to all fetal surgery practice."*

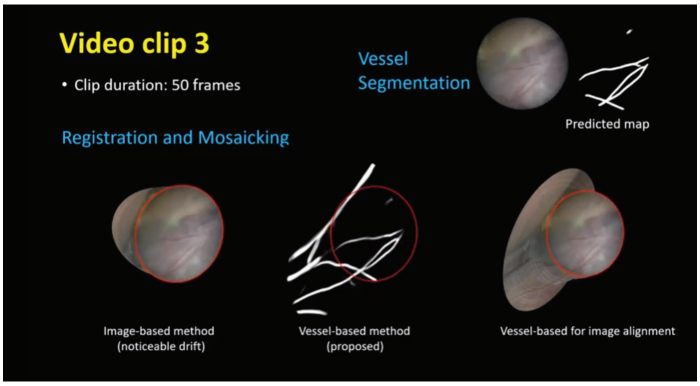
The research dataset for the challenge is not currently available online, but the team will release it publicly afterwards







and it will be a major contribution to the community. It has been reviewed by expert surgeons and features **18 fetoscopy videos annotated for placental vessel segmentation**. They will also be releasing sequences from these videos for registration and mosaicking, which means creating an image of the increased field of view that can be considered a panorama image inside the human cavity.



It will be the first large-scale multi-center fetoscopy dataset. Three fetal medical centers have contributed: **University College London Hospital; Istituto Giannina Gaslini in Genoa, Italy; and University Hospital Leuven in Belgium.**

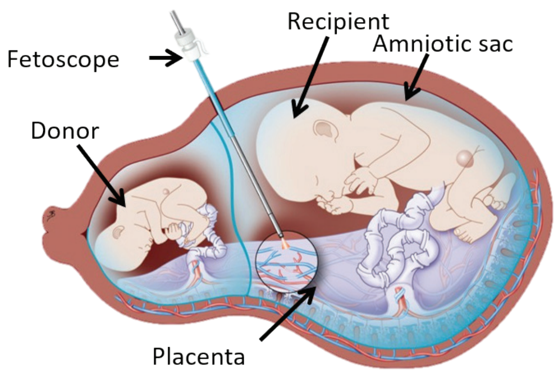
Sara is an Assistant Professor at the **BioRobotics Institute of Scuola Superiore Sant'Anna in Italy**. Her main research interests are in surgical data science and deep learning methodologies applied to the analysis of medical images and videos.

"If you develop a deep learning algorithm that works well for one surgeon, or one endoscope, or one center, that isn't particularly useful as it creates a bias problem," she explains. *"We are so proud to share this **multi-center dataset**, which is something new and will pave the way for future challenges."*

They aim to bring this technology into a real operating room within the next couple of years.

Finally, we asked the team to tell us what fascinates them most about working in this field.

"The most magical part of our work is you never know what ideas other researchers may have based on what you have done," Sara smiles. *"You do*



something and it's not perfect, certain things don't work, but then if someone has another idea based on what you did, I think that that is the best thing. It's like a chain. These things are complex. You are part of something bigger where somewhere, somehow, ***someone will solve it****!"*

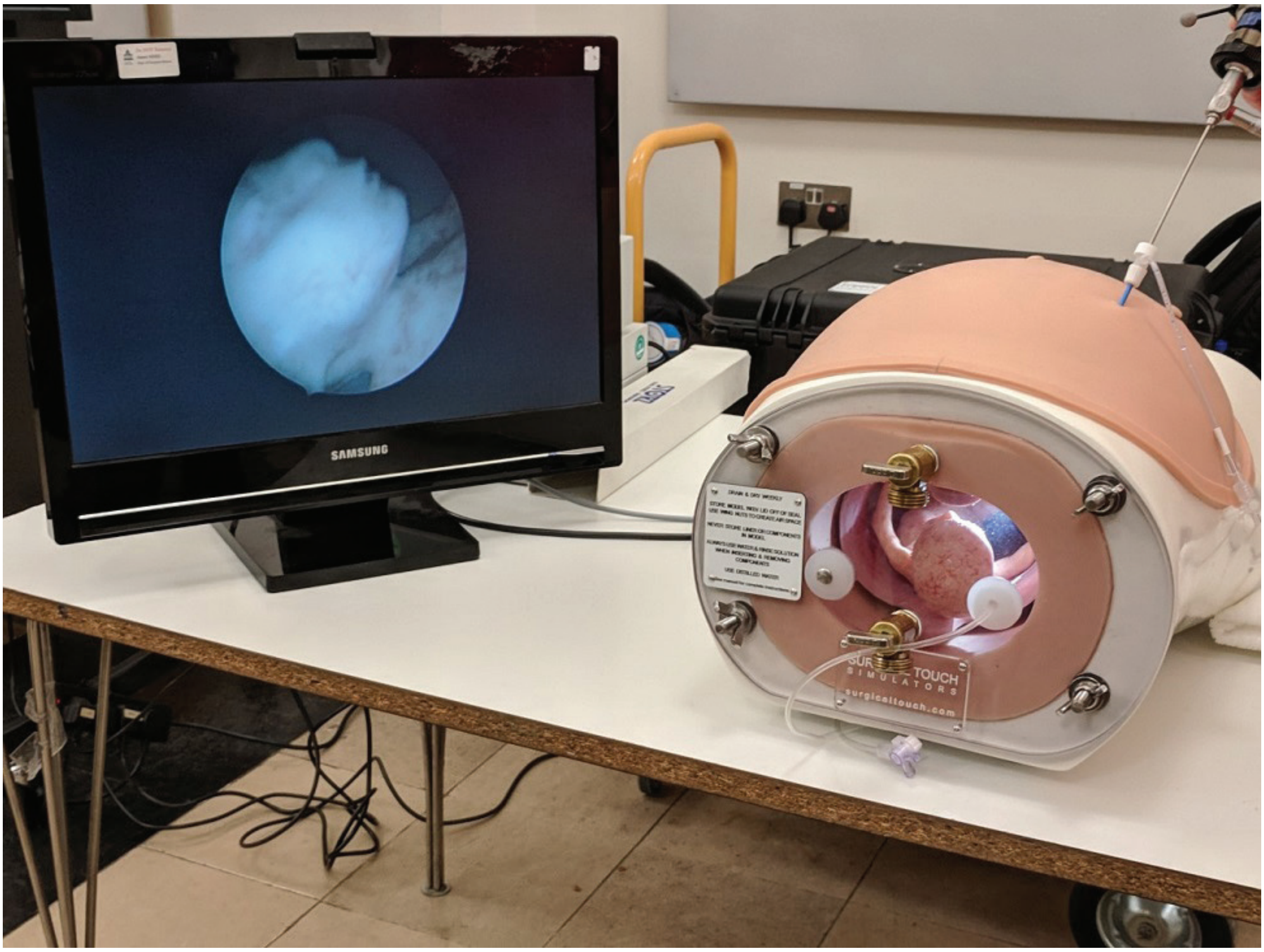
Alessandro shares the same feeling.

"I really hope it will pique the curiosity of other researchers," he adds. *"It feels like such an achievement for me as a student to work on something like this which isn't such a common field. We all share the same motivation, which is to improve patient outcomes."*

Sophia agrees: *"I feel we can make a difference to the healthcare system in the long run. That keeps me going and keeps me coming up with new ideas to broaden my vision in the field."*

The team are tight-lipped when it comes to giving away any tips for participants in case it gives them an advantage, but they are keen to point out that the purpose of any challenge is to **push your limits and discover something new.**

"We also have prizes for the winning team," Sophia adds, laughing. *"That will be an added bonus for the participants!"*



Julie Shah is an Associate Professor in the Department of Aeronautics and Astronautics at MIT. She is also the Principal Investigator at the Interactive Robotics Group of the Computer Science and Artificial Intelligence Laboratory. In addition, she is Associate Dean of Social and Ethical Responsibilities of Computing within MIT's Schwarzman College of Computing.

More than 100 interviews with Women in Science here!

Julie, I have a number of subjects I would like to talk about: flying machines, non-flying machines, work at MIT as well as your involvement in ethics. Where would you like to begin?
Oh! Whichever one you want!

Let's start with flying machines. What draws you to that?
I'm faculty at the AeroAstro department at MIT. Aerospace has been a dream and a love of mine for as long as I can remember. Since I was 3 or 4 years old, I've been watching

planes. My dad has videos of me when I was 4 years old talking about rocket ships.

I would love to see that! [*laughs*]
My parents really encouraged an early love of science and engineering. I remember watching the space shuttle launches when I was little, the Mars rover that landed in the 90s. I've always desired to be a part of developing new technologies that can expand our knowledge as humans. The other aspect that I really love about aerospace is that it has to be perfect! [*laughs*]

I **remember times when it did not work perfectly, and it was a disaster.**
Exactly! In my work currently, I specialize in human/machine collaboration and human/machine interaction for safety critical environments.

Do you also fly airplanes?
I worked on my pilot license while I was a grad student. I have flown airplanes. It's not an everyday part of my life currently, but I think it's important to understand how it really works, rather than just in the abstract.

What does sometime go wrong, in terms of the interaction with humans and machines?
I think there are interesting parallels between the challenges that we've had to address in aviation and pilot

The most wonderful part of being a researcher and being a professor is that it's my daily job to be "wowed"

interaction with cockpit automation. There are lessons that we've had to learn the hard way about how to design human capabilities and machine capabilities to really fit together like a puzzle piece effectively. In aviation in the United States, we have one of the safest air transportation systems in the world. It's through a lot of hard work and lessons learned the hard way. Every time you have a new generation of aircraft introduced, including things that you expect to make flying safer, you see a large spike in incidents and fatal accidents. One of the reasons we have been able, over the decades, to achieve such a safe air transportation system, is by honing the design of the human/machine as a system, rather than thinking about engineering technology as independent from human capability. If you think about what is happening in our world today, we have autonomous vehicles on our roads, sidewalk delivery robots, security guard robots... We have robots that go up and down our grocery stores looking for spills. None of these technologies really function independently. They are a part of a complex system. We co-design them with our social norms and policy as well, so that we can achieve a large-scale introduction of these systems. There's a lot we can learn from aviation. For example, things that you think should be clumsy aspects of interaction with a system, maybe the interface doesn't allow you to navigate to find the right window quickly, to understand some particular subsystem. Building a person's mental model of a system is critical. We've had many aviation accidents. There was an interplay of this effect with the Boeing 737 MAX mode confusion and not understanding the current state of the system, how the actions will impact the behavior of that system. The mode confusion has resulted inmany aviation accidents. We actually see it in the interaction with drivers, with intelligent vehicles like the Tesla. These are challenges that are not just specific to aviation, but the lessons that we've learned there can translate and can help us design systems that are safer for us broadly in society.

In general, what are we doing right and what are we doing wrong as humans interacting with machines?
I spent much of my current career in robotics for manufacturing, helping to deploy collaborative robots that work alongside people who build planes and build cars. The interesting thing about deploying a robotic system that moves and works around people is that largely with these systems, they just see people as obstacles. We see this with sidewalk delivery robots or robots in grocery stores. We're really no different than any other obstacle that systems might encounter, and that's a major limitation for these systems to integrate effectively with people. Just like the robot needs to understand what other cars are, what trees are, they need to understand more about people than us just being obstacles in their environment. So that's a key focus of my lab's research.

Can you tell us a bit more about robots in healthcare, like in robotic surgery?
This is a very exciting time in the introduction of technology and AI in healthcare. We have really excellent examples of research at MIT that can help diagnose cancer in images. What we see there and what we know from many other contexts is that humans and machines in AI have different capabilities. The errors that humans make are going to be different from the errors that an AI system will make. By cleverly architecting the system, a system that enhances human capability rather than replacing it, that's where you can achieve something better

than the human expert that diagnoses cancer. That can help fill in the gaps of the errors that a machine will naturally make, which will be different than people. Think of it like a swiss cheese model! [laughs] Each of us has our biases and errors that we make. If we put it together the right way, we can cover each other's gaps. That's how you can achieve more effective capability. We have a really interesting turning point in surgery with the introduction of the new robotic surgery systems, the Da Vinci and other systems. They are still primarily tele-operated by the surgeons. The surgeon is sitting at a console, removed from the patient. They are basically controlling every joint of that robot through this fancy, remote control interface.

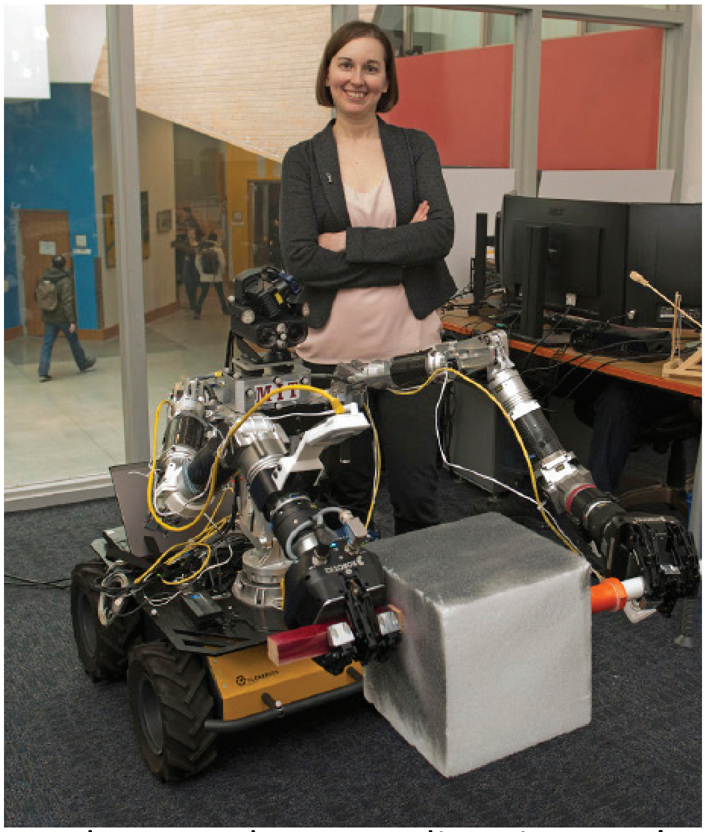
They might be in another continent.
Yes, with a time delay. There are researchers that are working very actively on that very remote form of conducting surgery. There are many interesting challenges on the stability of the control system under time delay.

These new surgical systems introduce a computer between the surgeon and the patients.
With surgical robots, you have a computer between the surgeon and the patient. This is the beginning of an enormous opportunity in enhancing how we go about surgery, improving safety and improving outcomes. That's quite exciting. Computer vision is a key enabling technology in that direction of work.

Did you do any research in this area yourself?
My own work in healthcare focuses more on supporting decision-making. For the past few years, I've worked on a project with Beth Israel Deaconess Hospital in Boston. We studied the decision-making strategies of very senior nurses on a labor delivery floor. These nurses are called charge nurses or resource nurses. They're like the nurse managers of the labor floor. It was very interesting for me to learn and understand that these nurses are essentially doing the job of an air traffic controller. They're deciding which patient goes to which room,

which nurses are assigned to which patients. They control aspects of the operating room schedule. When you study the decision-making tasks, they are actually doing a job that's more complex than an actual air traffic controller. Currently, they do it today without any decision support.
Air traffic controllers were sort of born to do that job. Some people just can't do it. It's like at the edges of human capability, these types of jobs, no matter how much training you have. Nurses are trained to do that job in that way too. How do you support them in decision-making? We have a lot of information that's captured about our air transportation system. We have digital tools. We have substantial national investments in upgrading the structure to support airtraffic control. But nurses today do an equally demanding and challenging job with a piece of pen and paper. The question is: how do you reverse-engineer what it is they are doing, to be able to offer support?

Let's talk about ethics. It seems that the more and more we let machines interact with us humans, instead of solving ethical problems, we are creating more ethical problems. It's not that we create more ethical problems with the introduction of technology. The ethical challenges are there, but technology and computing



accelerate whatever direction we're going down. I think there needs to be even more emphasis and importance placed on understanding the implications of decisions, both early in the conception of the technology and its implementation, and then thinking at the beginning of the implication for widespread use of the technology for different stakeholders, whether they directly interact with it or for society or communities more broadly. I'm not a philosopher. My training is as an engineer. However, it's equally my responsibility as a person who has the deepest insight into the technological decisions that we make, to be able to engage and communicate with those of other disciplines and bring that back into the design of the technology. Is my decision support system for the nurse taking over the job of a nurse? Or: is it uplifting the nurse to make better decisions? Providing the right information, the right time, making certain suggestions to reduce cognitive load, freeing up the nurse's capacity to do more complex aspects of their job - all of these are very much squarely in the domain of the engineer and computer scientist. Very passionately, we need to be training our engineers and computer scientists on all of these questions as sort of the bread and butter of what they do with technology.

Let's stop at the one word you just used but in a different context: passionately. You obviously are passionate about the things that you do. How do you continue to feel so passionate about something you do all the time?

How is this passion fed? For me, and I tell this to my students and my researchers, it's critically important to live a happy, well-balanced life. That's the only way you do your best, most creative work. Taking care of yourself, being well rested, eating well, having hobbies and things you enjoy outside of work - all of this adds energy and helps make you more successful, not just in your work but in all parts of your life. Where do I get my energy from? I do not know! I have a two-and-a-half-year-old and a four-and-a-half-year-old. They keep me busy! I don't know where their energy comes from! I just work to keep up!

What was the most fascinating moment in your career? Which moment of discovery really gave you that "wow" factor?

That's a great question!

Thank you! That's what I'm here for! [*both laugh*]

The most wonderful part of being a researcher and being a professor is that it's my daily job to be "wowed". I did my undergrad degree in aerospace engineering, also at MIT. Right up until my senior year, I was a double major in physics. I needed one more lab class to finish that double major in physics. I really loved physics. I liked the beauty of it. One of my motivations is to connect technology with humankind, to understand what it is we know about the world and the universe. Then my senior year, instead of taking that last physics lab class, I took basically an intro to artificial intelligence class. I wasn't sure if it was the right decision. There were PhD students and faculty that came through and presented the latest ongoing work from their lab. I remember looking at a video of a walking robot, balancing and jumping. Now we have real robots in the world. But that didn't exist yet. It was just in simulation. I remember when Cynthia Breazeal came, a pioneer in social robotics: she came to this class to talk about how she was pushing the field of how robots interact in a socio-emotional context. From each of those talks, I just felt that what they were doing was magic. It was indistinguishable from magic. I knew it wasn't magic! I knew I could probably learn how to create systems with those capabilities. The wonderful thing about my job is that I feel some version of that wonderful feeling of "wow" every week, if not every day, interacting with my students when they show me results on their latest direction of work, being inspired by colleagues in another meeting. I definitely chose the right career path overall. That's why it's so hard for me to pick one!

Of all the things people expect robots to do in the future, which one is an illusion that will never happen?

I know to never say never! There are many things that are not likely truly impossible. I can tell you what I spend most of my time worrying about. I don't worry about singularity. I don't worry about generalized artificial intelligence. What I worry about are the technologies that are here and now, or that will be here in the next ew years, and the thoughtfulness that we're putting into guiding and shaping the direction of that technology development and the effort and thoughtfulness we're putting into those technologies right now. Maybe those other concerns are multi-decade concerns. Autonomous vehicles have the potential to drastically improve the safety of our roads. Tens of thousands of people die in the US every year. Or it could make it worse! It could increase congestion. I worry about issues of access, like who has access to these technologies. There are concerns around equity. Who are they developed for? Who will benefit? Those are the questions that currently keep me up at night, and I am very grateful to the researchers and scholars who think about those longer-term questions.

This interview is exceptionally long because it is fascinating. But I don't want to abuse your time.
Is this enough?

It's never enough! I am Italian: I can ask you questions until tomorrow! [*laughs*] Tell me just one thing about MIT that we don't know.
MIT has various different academic, degree, and certificate programs. The one I love most is the MIT pirate certificate through our physical education program. There is a physical education requirement. Students take different phys-ed courses through their undergrad or grad program. If students finish archery, hunting, pistoling, rifling, and sailing, they get an official MIT pirate certificate.
It's a fun place! A lot of people think of MIT as a very competitive environment. I've been here since I was 18. I've done all my degrees at MIT. I've chosen to stay here, time after time after time. It's anything but a competitive environment. This is a place where if you have a question, from a freshman to a full professor, you will find a friend or a colleague who is so generous with their knowledge to help with something you don't know or something that you're interested in. I think it's just a very, very special place. And maybe one indication of the specialness is that you can become a pirate at MIT!

More than 100 interviews with Women in Science here!

Dr. Mehran Anvari is Professor of Surgery and Chair in Surgical Innovation at McMaster University. He is also the Scientific Director and CEO of the Centre of Excellence in Canada for Commercialization of Medical Robotics. Dr. Maria Victoria Sainz de Cea is a Data Scientist at IBM Watson Health. They speak to us about a partnership that could revolutionize the world of surgical robotics.

Mehran Anvari

Maria Victoria Sainz de Cea

Dr. Anvari's Centre develops precise robotic capabilities that can navigate in a variety of imaging modalities, including MRI. The technology can accurately navigate to the site of any abnormality, however small within the body, so long as the system detecting it can provide it with the correct coordinates.

His group have been working with a global team from **Merge Healthcare**, part of **IBM Watson Health** in Canada, to develop AI platforms that connect with the Centre's robotics to build a truly autonomous medical robot which can accurately detect, locate, and intervene.

"Imaging analysis is very fast now," Dr. Anvari explains.

"The systems we're building will significantly improve on current practices, reducing time even further between analysis and diagnosis. ***AI can analyze images in a matter of seconds*** *and we're working on platforms that can* ***add histopathology in minutes****. They consider not only the usual diagnosis but look at patients' outcomes and the decision-making which led to that outcome."*

Where surgeons make decisions based on personal experience of perhaps hundreds of previous patient interactions, these systems use machine learning to base decisions on analysis of millions of interactions from around the world. They are constantly evolving and enhance patient treatment by taking surgical intelligence to a higher level than what one surgeon or even a group of surgeons could achieve.

The team are on the cusp of a huge breakthrough, building tools for **screening, diagnosis, and treatment of cancer at a much earlier stage**. Proof of principle testing was completed a couple of years ago, supported by the **Canadian Space Agency (CSA)**, to ensure that the AI and robotics platforms can communicate seamlessly without human intervention.

"We are even working with the CSA and NASA to potentially build an autonomous medical robot for the Mars mission," Dr. Anvari adds.

"You can see that the same type of ideology behind this can not only improve care of patients on earth, but potentially on long space missions too. We are partnering with the main manufacturer of all the robotics for the space program - probably the most advanced robotic company in the world - to build a series of systems which will work with IBM Watson Health in various situations. But that's as much as I can tell you!"

IBM Watson Health have a history of partnering with different stakeholders. Back in 2018, **John R. Smith** told us about its partnership with **MIT** on the **MIT-IBM Watson AI Lab**.

"The way we push science forward is by partnering, hearing other opinions, and bringing other minds into play," Maria tells us.

Maria works in the imaging branch of IBM Watson Health, developing deep learning algorithms for detection, segmentation, and classification of medical images.

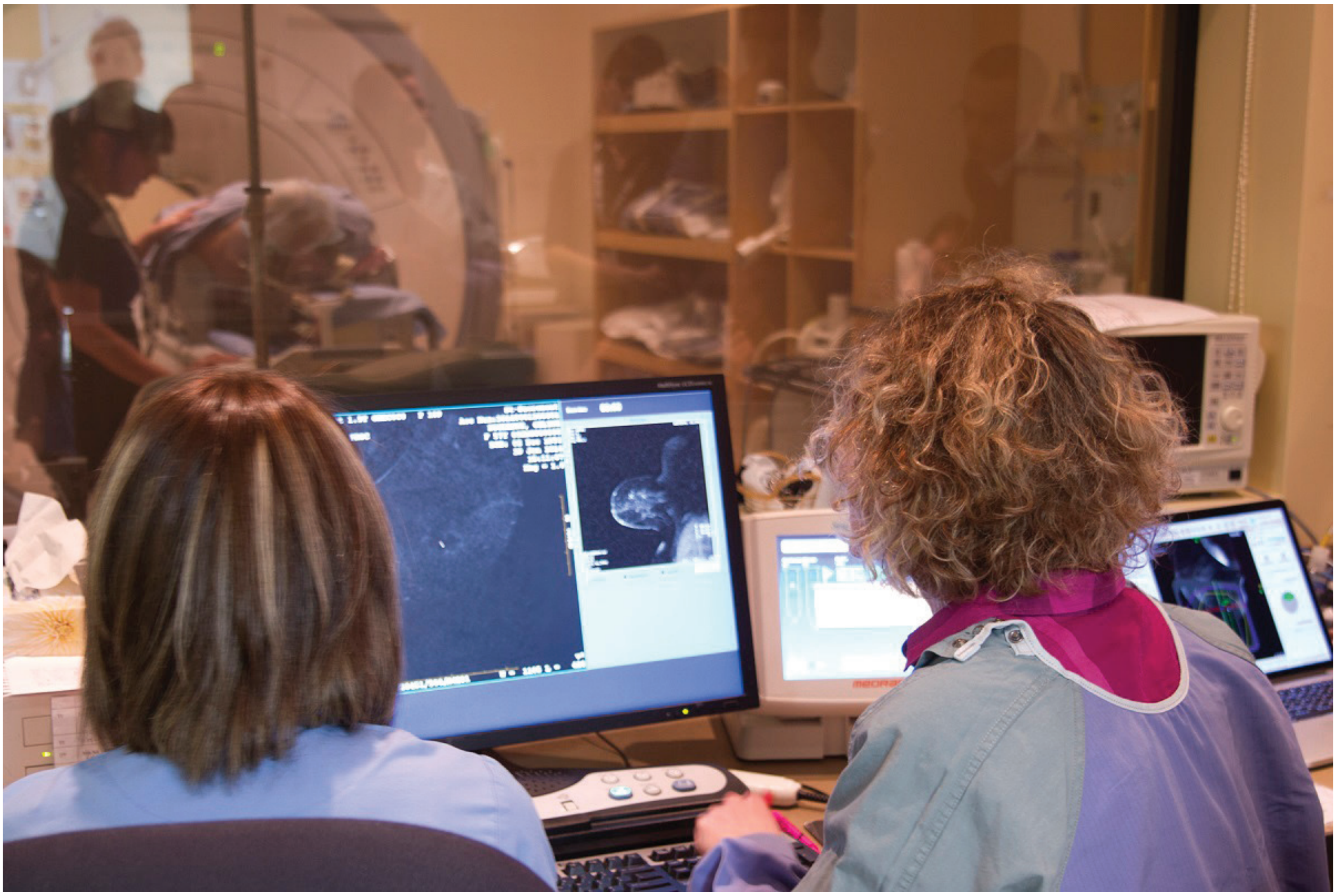
"We deal mostly with radiology images - right now, I'm working on prostate MRI, but we work on a wide variety of modalities. The best thing

about working here is the fact you get to work with very smart people on cutting edge technologies. We have access to a huge amount of data, and that makes our algorithms highly effective."

Dr. Anvari agrees:
*"Any organization is only as strong as the people it employs and the vision that motivates it. I've been working with a very collaborative group of engineers at IBM Watson Health under **Dr. Marwan Sati** who have really grasped our ideas. The motivation that I see in the group of people I work with is not only to improve patient care, but to improve access to care also."*

Remote access to care is another benefit that AI can bring, meaning any patient anywhere with a disease or who has had an accident or emergency can get **the same treatment as they would receive at the very best hospitals in the world.**

However, with some high profile ethical and societal issues related to the use of AI in the news recently, has this affected patient confidence at all?



"Physicians, who are traditionally very conservative, have to accept the fact that for some tasks, AI can enhance their abilities to diagnose and treat; the public have to agree that it is the best thing for them; and hospitals have to understand that AI is going to save them money. I have no doubt that we will satisfy all three!"

"My experience with patients is that they're very willing to accept technology," Dr. Anvari reassures us.

"When they see that AI improves the care they receive, and that they can get the care closer to home, they are very happy. 17 years ago, I started a program operating on patients remotely. Everybody said, 'Who would allow you to operate on them when they've never seen you?' The first patient was obviously very brave, but once everybody had heard that it was a success, there was no shortage of people saying I'd rather have the surgery remotely than have to travel 400km!"

AI and robotics have already been combined successfully in many industries, but in medicine, Dr. Anvari says it is **revolutionary**, although there is some work to do.

"There are three groups that you need to satisfy," he points out.

"Physicians, who are traditionally very conservative, have to accept the fact that for some tasks, AI can enhance their abilities to diagnose and treat; the public have to agree that it is the best thing for them; and hospitals have to understand that AI is going to save them money. I have no doubt that we will satisfy all three!"

Did you like our section
about Robotic Assisted Surgery?

Learn more about it
in many more articles
on RSIP Vision's website.
Find it here!

COMPUTER VISION EVENTS

PMWC
Virtual
June 14-18

Women in AI
Online
June 15

CVPR 2021
Virtual
June 19-25
Meet us there

FIMH 2021
Stanford University
(virtual or hybrid)
June 21-24

IPMI
fully virtual
June 27-July 2
Read More

MIDL
fully virtual
July 7-9

Mass Data Analysis...
New York, NY
July 11-13

MIUA
Virtual
July 12-14

SPIE Optics + Photonics 2021
S.Diego, CA
(options for remote)
Aug 1-5

Ai4 2021
fully online
Aug 17-19

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 to go virtual or to be held at another date. Please check the latest information on their website before making any plans!


IMPROVE YOUR VISION WITH
Computer Vision News
SUBSCRIBE
CLICK HERE, IT'S FREE
The magazine of the algorithm community
A PUBLICATION BY
RSIP VISION