February 2021

Computer Vision News

The Magazine of the Algorithm Community

Research: RAFT (best paper ECCV)

Imaging Challenge
EndoCV 2021

Upcoming Events

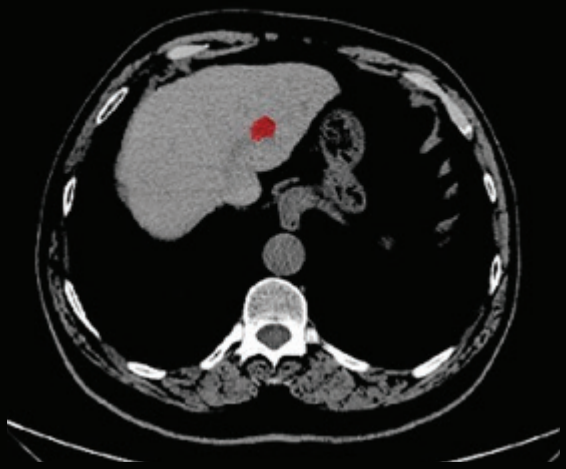
Women in C.V.

Artificial Intelligence
Spotlight News

Congrats, Doctor!

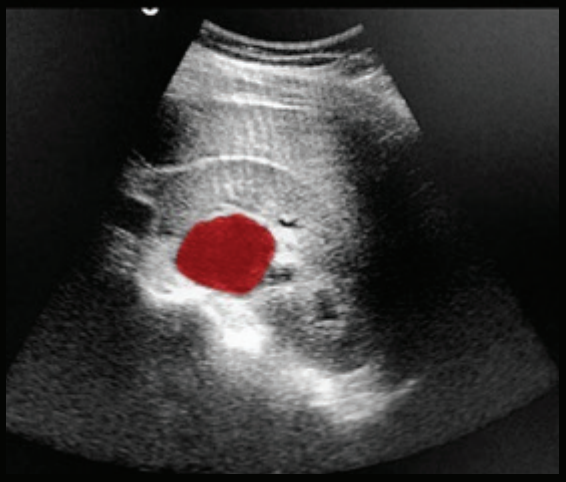
One-click Segmentation

Liver tumor, post Segmentation reading

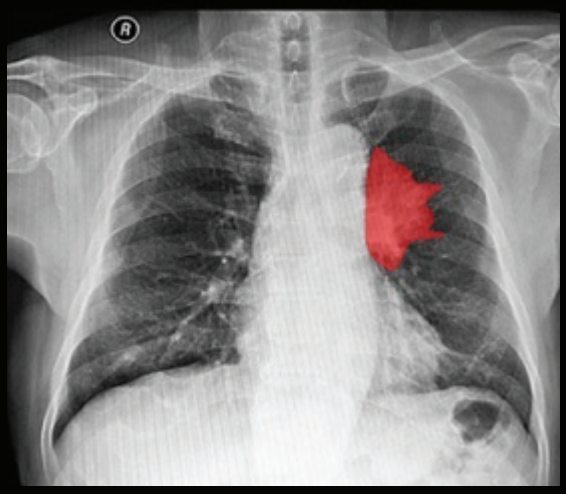


Lung tumor, post segmentation reading

Kidney Cyst, post segmentation reading



Lung Mass, post segmentation reading



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,

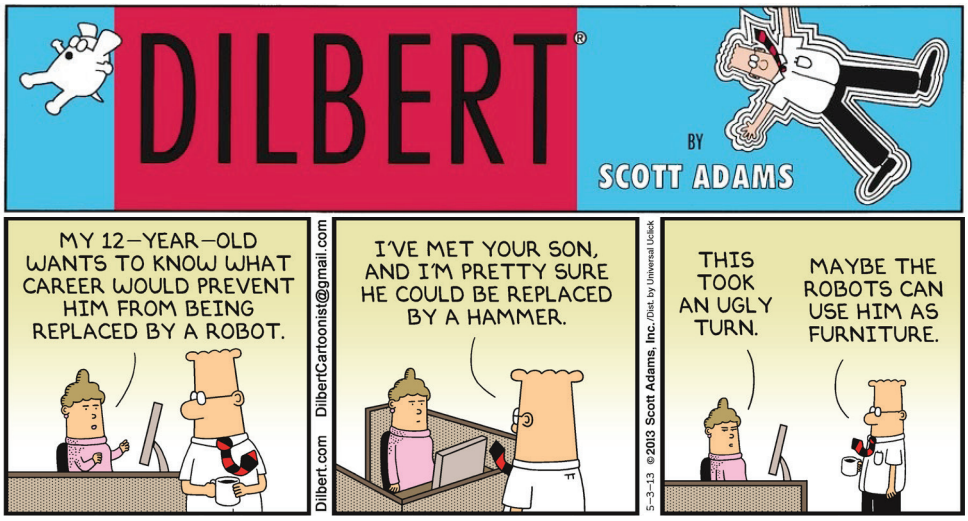
What do you get when you send a marketing guy to join a team of pioneers in computer vision? You get **Computer Vision News**, the magazine of the algorithm community, published by **RSIP Vision** since 2016!

First (on page 4), we introduce the **Best Paper winner** at ECCV 2020, the widely acclaimed **RAFT: Recurrent All-Pairs Field Transforms for Optical Flow**. Then, on page 10, we will see what AI is capable of doing on medical images, almost alone, **with just a single click**! Further on, we head to Canada with someone who has been in **MILA** leadership with **Yoshua Bengio** for 10 years, with great stories to tell. We will finally pay a short visit to my hometown Milan, Italy, where **Rita Cucchiara** has organized a very successful ICPR2020 meeting. BTW, it was held online (and it was already 2021).

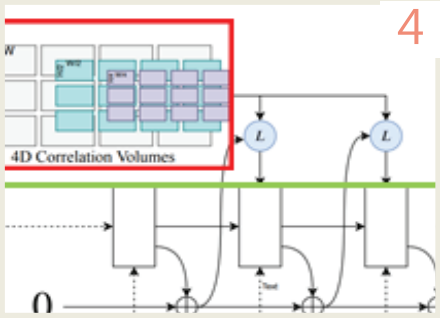
Save the date for a very special **meetup** in our series on **AI in Medical Imaging**: Julia Schnabel, Chair in Computational Imaging and Head of Research & Impact at the School of Biomedical Engineering & Imaging Sciences, **King's College London**, has accepted our invitation to speak at the meetup. She will talk about **AI in Clinical Applications: Machine learning in medical imaging - from image acquisition to interpretation**. Secure your place now and join us **on February 18 at 10am PT**. **Register here** to receive your personal access to the event.

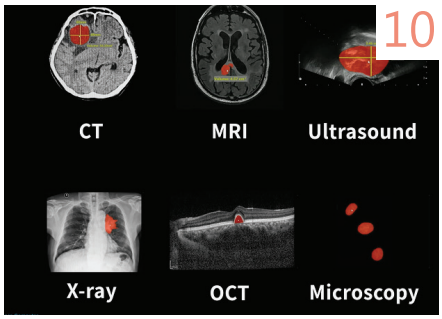
Of course, take us along for your next Deep Learning project!

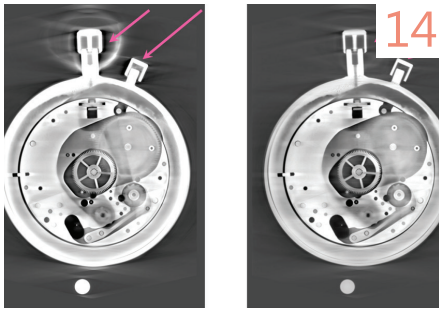
Ralph Anzarouth
Editor, **Computer Vision News**
Marketing Manager, **RSIP Vision**

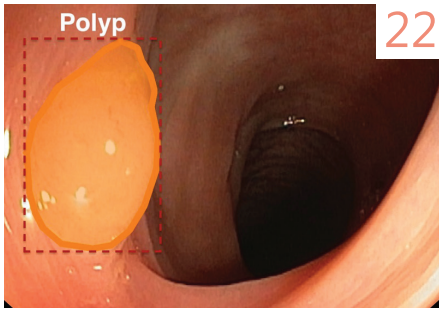


Summary 3

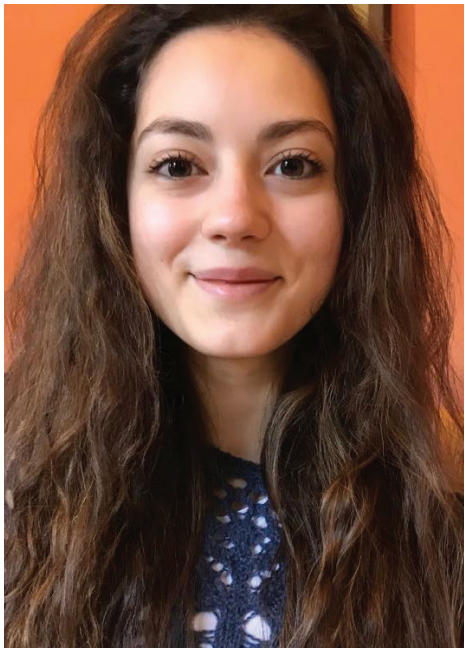








RAFT: Recurrent All-Pairs Field Transforms for Optical Flow



by Marica Muffoletto

BEST PAPER ECCV 2020

Every month, Computer Vision News reviews a brilliant research paper. We started the new year with a great January issue focused on medical imaging research. Let's just now have look at a completely different angle of computer vision, with a paper called **RAFT: Recurrent All-Pairs Field Transforms for Optical Flow**, written by a duo from Princeton (Zachary Teed and Jia Deng). We are indebted to the two authors for allowing us to use their images to illustrate this review.
Their paper can be found at this **link**.

RAFT is a new end-to-end trainable model dealing with the delicate task of **optical flow**. This refers to all those algorithms that have as ultimate scope to **estimate per-pixel motion between video frames**, which of course can become particularly difficult in real-life situations, where fast moving objects, occlusions, blurs and any kind of unpredictable changes are frequently found. As much challenging as it can be, optical flow is similarly useful. There is a wide range of computer vision tasks which could be substantially improved by **a good estimation of the flow between different frames**, or "instances in time".

Mathematically, the problem of optical flow can be expressed as:

Finding a dense displacement field (f^1, f^2) which maps each pixel (u, v) in an image to the corresponding coordinates $(u', v') = \left(u + f^1(u), v + f^2(v)\right)$ in the next image.

Before the **Deep Learning revolution**, traditional optimization-based approaches were hand-crafted and used a continuous formulation of the above, based on the iterative refinement of a single estimate of optical flow. On the contrary, more

recent deep learning methods aim at **directly predicting optical flow between a pair of frames**. This paper, together with a few others, works on an integration of these two aspects - a method defined as "**learning to optimize**" strategy - where the network uses a large number of update blocks to emulate the steps of a first-order optimization algorithm. A strong point of RAFT lies in the use of a **single high-resolution flow field, updated through a large number of lightweight operators**.

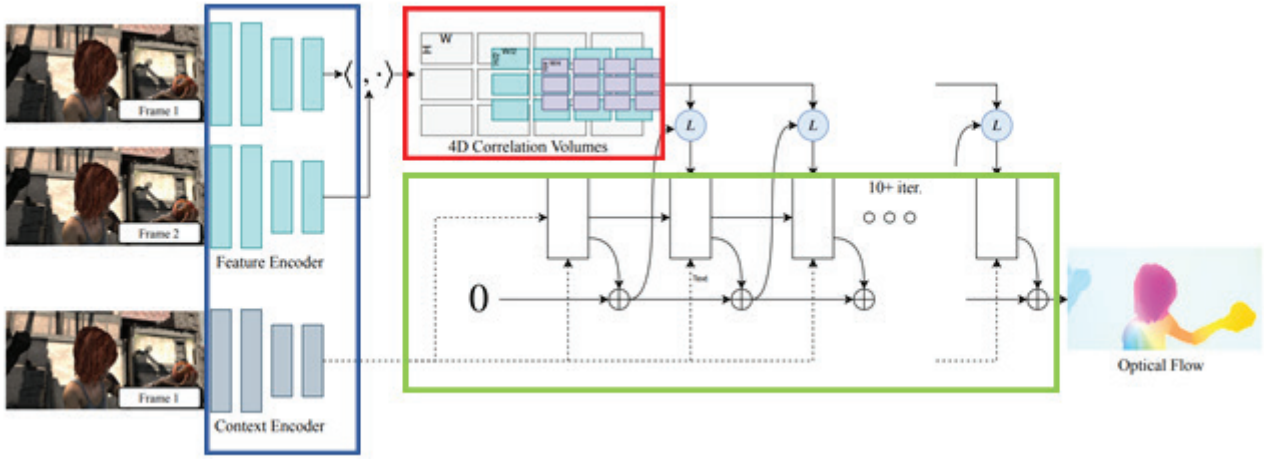


Figure 1: RAFT main components

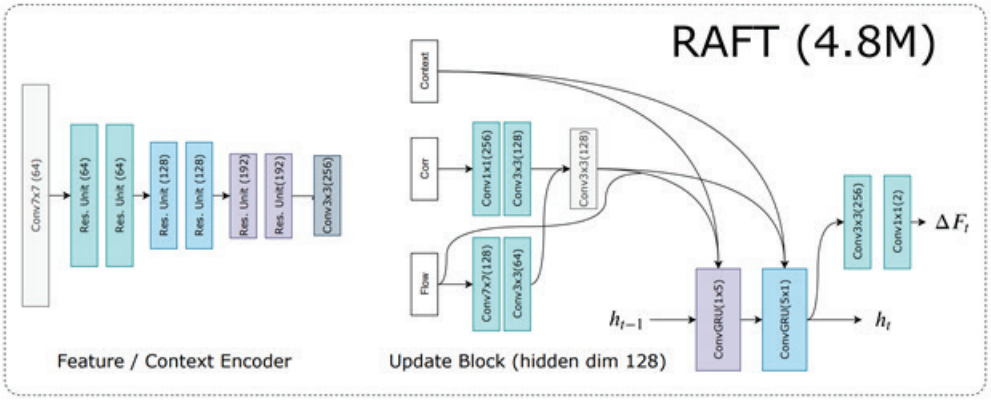
The model architecture is divided into:
- an encoder section that extracts 1) per-pixel features from the two paired frames, and 2) context information from just the first frame;
- a correlation layer that computes visual similarity between pixels by constructing a 4D W × H × W × H correlation volume by taking the dot product of all pairs of feature vectors. The last 2-dimensions of the 4D volume are then pooled at multiple scales to construct a set of multi-scale volumes which retains high resolution information on large and small displacements;
- a recurrent update operator which mimics the steps of the traditional optimization algorithms. It calculates updates of the flow field by estimating the descent direction based on features extracted by the correlation volumes. In this component, the GRU block is found, a gated activation unit with fully connected layers replaced with convolutions. The GRU block outputs a hidden state passed through 2 convolutional layers. These return the predictions of the flow update which need to be upsampled to full resolution.

TRAINING

RAFT is implemented using the **PyTorch** library.
The full architecture, including feature and context encoder (with slightly different

characteristics) and two convolutional GRU update blocks with 1x5 filters and 5x1 filters respectively, is shown below.

	FEATURE ENCODER	CONTEXT ENCODER
Input	(I_1, I_2)	I_1
Normalization	Instance	Batch



All modules are initialized from scratch with random weights and during training AdamW optimizer is used. As network loss, the I_1 distance between predicted and ground truth flow is employed:

$$\mathcal{L} = \sum_{i=1}^{N} \gamma^{N-1} \left|\left| \boldsymbol{f}_{gt} - \boldsymbol{f}_i \right|\right|_1$$

where $N = (number\ of\ frames - 1)$ and γ=0.8.

Several random augmentation methods are also applied to the inputs during training: 1) photometric augmentation (perturbation of brightness, contrast, saturation, and hue); 2) spatial augmentation (rescaling and stretching); 3) erosion of rectangular regions in I_2 with probability 0.5 to simulate occlusions.

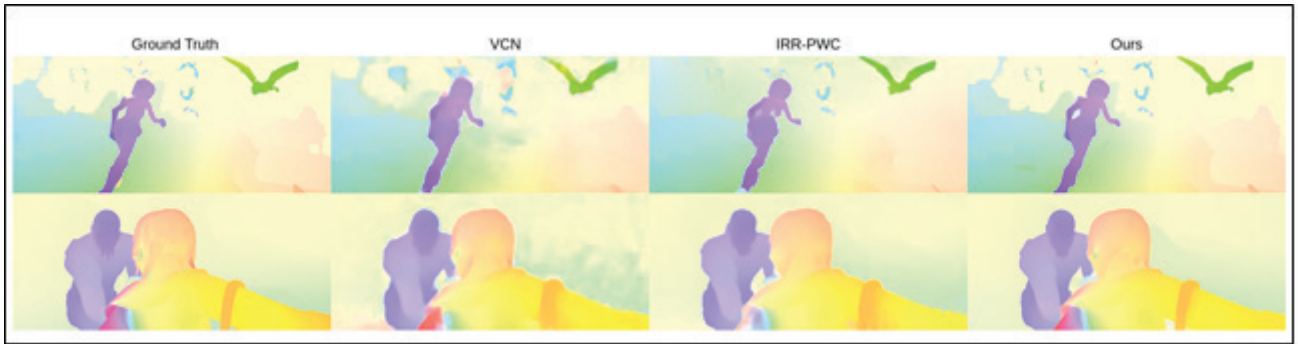
Further details of the training are summarized in the table below, including (first 3 columns) the combinations of data used for training and finetuning.

Stage	Weights	Training Data	Learning Rate	Batch Size (per GPU)	Weight Decay	Crop Size
Chairs	-	C	4e-4	6	1e-4	[368, 496]
Things	Chairs	T	1.2e-4	3	1e-4	[400, 720]
Sintel	Things	S+T+K+H	1.2e-4	3	1e-5	[368, 768]
KITTI	Sintel	K	1e-4	3	1e-5	[288, 960]

Remarkably, the final architecture and training options are the results of detailed ablation experiments conducted on several isolated components to make sure that the best setting was implemented.
The need for an **update operator** is justified by comparing the performance of a GRU block with a set of 3 convolutional layers with ReLU activation, and **tied weights** at all instances are found to provide better convergence than untied ones. The addition of the **context** encoder also results in better performance after tests. Features are extracted only at a **single resolution** because multi scale one over-complicates the architecture and it is not found to perform significantly better. The presence of the **correlation pooling operation** on the output features allows to capture both large and small displacements. **All-pairs correlation** is preferable and convenient over the computation of a local neighbourhood correlation. The **correlation volume** is compared to a layer warping the features from I_2 onto I_1 by use of the current estimate of optical flow, and found superior. The **upsampling module** results in better performance than the bilinear upsampling. Finally, a total number of 32 inference updates is picked after relative experiments.

CONCLUSION

RAFT was evaluated on 3 datasets: **Sintel**, **KITTI** and **DAVIS**.
Examples of the results obtained on them are shown below. The images speak for themselves! The top row includes comparisons with state-of-the-art methods similar to RAFT, and the differences in performance are noticeable especially at the boundaries of fine details. This is due to the presence of the convex upsampling module in the architecture which improves accuracy near motion boundaries, and also allows RAFT to recover the flow of small fast-moving objects such as the birds shown in the figure.



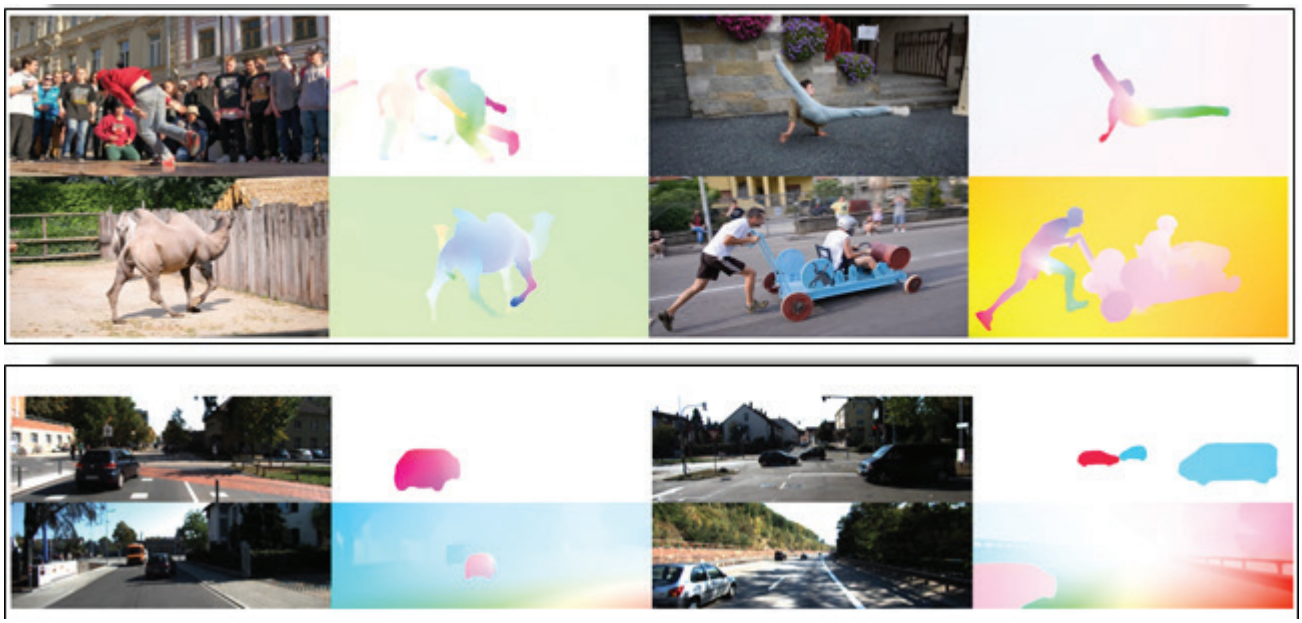
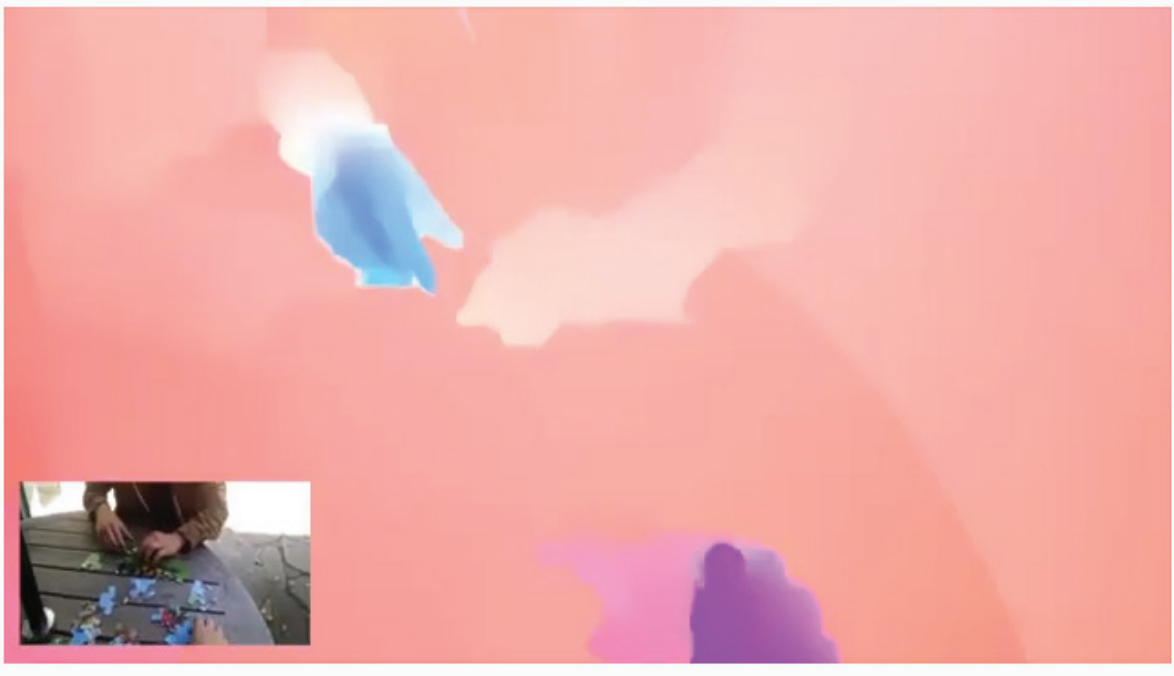


Figure 2: Top row: flow predictions on the Sintel test set; Middle row: Flow predictions on the KITTI test set; Bottom row: flow predictions on 1080p (1088x1920) video from DAVIS (550ms x frame).

The accuracy obtained in these practical optical flow examples have made this paper a success of the field already. The **Best Paper award at ECCV 2020** and the high interest in this work are proof of a novel and well-carried out research. Moreover, this paper is linked to a github page available to the community which could be object of one of the future tools of the month and which is already extensively employed for experiments.
We leave you with one of such beautiful live demonstrations of RAFT!



RAFT Optical Flow - EgoHands

Bay Vision Meetup 9

RSIP Vision invites you to its upcoming webinar about AI in Clinical Applications: Machine learning in medical imaging - from image acquisition to interpretation. For the first time, our invited speaker will be Julia Schnabel, one of the leading professors in the field of medical imaging. The webinar will be held... on February 18 at 10am PT.

Register now to join us!

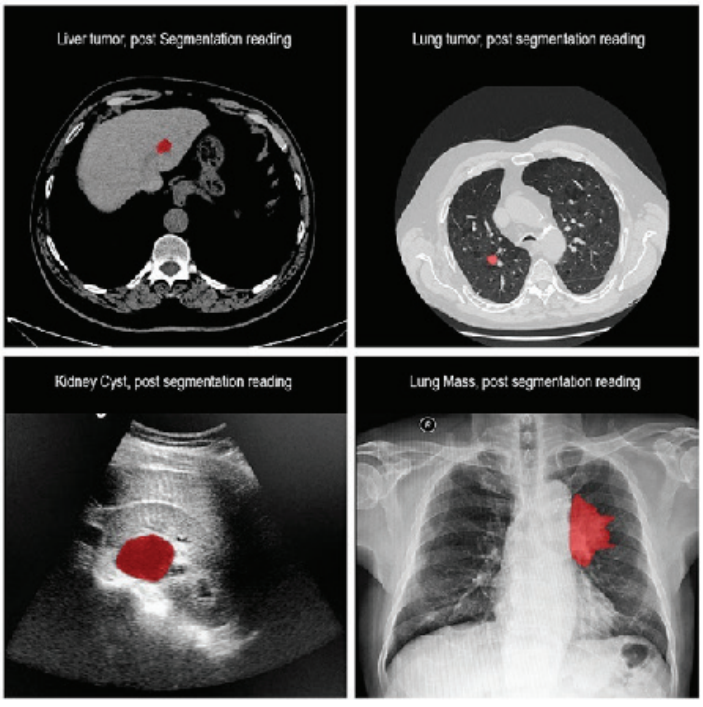
One-click segmentation of multimodal medical images

The challenges of medical image segmentation

Image segmentation is a fundamental and essential process in image analysis and computer vision, which involves partitioning an image into **multiple meaningful segments** for future analysis and use. In the medical field, image analysis plays an exceptionally vital role in both the **diagnosis process**, as well as in **interventional procedures**.

Medical image segmentation poses many **challenges** in several different areas:

- **Modality**- a wide variety of imaging modalities are in use today, depending on availability as well as the task at hand. Such modalities include microscopy, X-ray, ultrasound (US), computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET) and many more. For some tasks, several imaging modalities are used simultaneously, for example preoperative MRI with live US-guidance during prostate biopsies.
- **ROI**- in medical imaging, the region of interest (ROI) varies widely depending on the task at hand, be it identifying the location and extent of an arterial blockage, or accurately estimating the size and boundaries of a tumor.
- **Dimensionality**- the resulting medical image may be a two-dimensional image of a slice made up of pixels, or a three-dimensional image of a volume made up of voxels. At times, 3D images of volumes are even reconstructed from a set of 2D slices. Accordingly, the ROI can be represented as a slice or a volume.



Automation levels in segmentation

A wide variety of image analysis and artificial intelligence methods can be implemented in order to **segment a medical image**. When choosing a segmentation method, several approaches can be taken:

- **A manual approach**- in manual segmentation the user has full control over the segmentation process. Consequently, manual methods may offer the user control of the results, but they require meticulous, time consuming work by a specialist that may still be prone to user bias.
- **A fully automated approach**- in a fully automated segmentation the process is completely computerized and involves no user control. Despite requiring no human input, fully automated methods remain limited by their inability to make corrections in the process and accordingly deliver suboptimal results at times.
- **A semi-automated approach**- in semi-automated segmentation, the user is provided with a means of interaction and keeps some control of the process. The main challenge in semi-automated segmentation is to offer a good balance between the precision level and the required user interaction. In other words, achieving accurate segmentation with minimal user input in the interactive settings.

The benefits of one-click segmentation

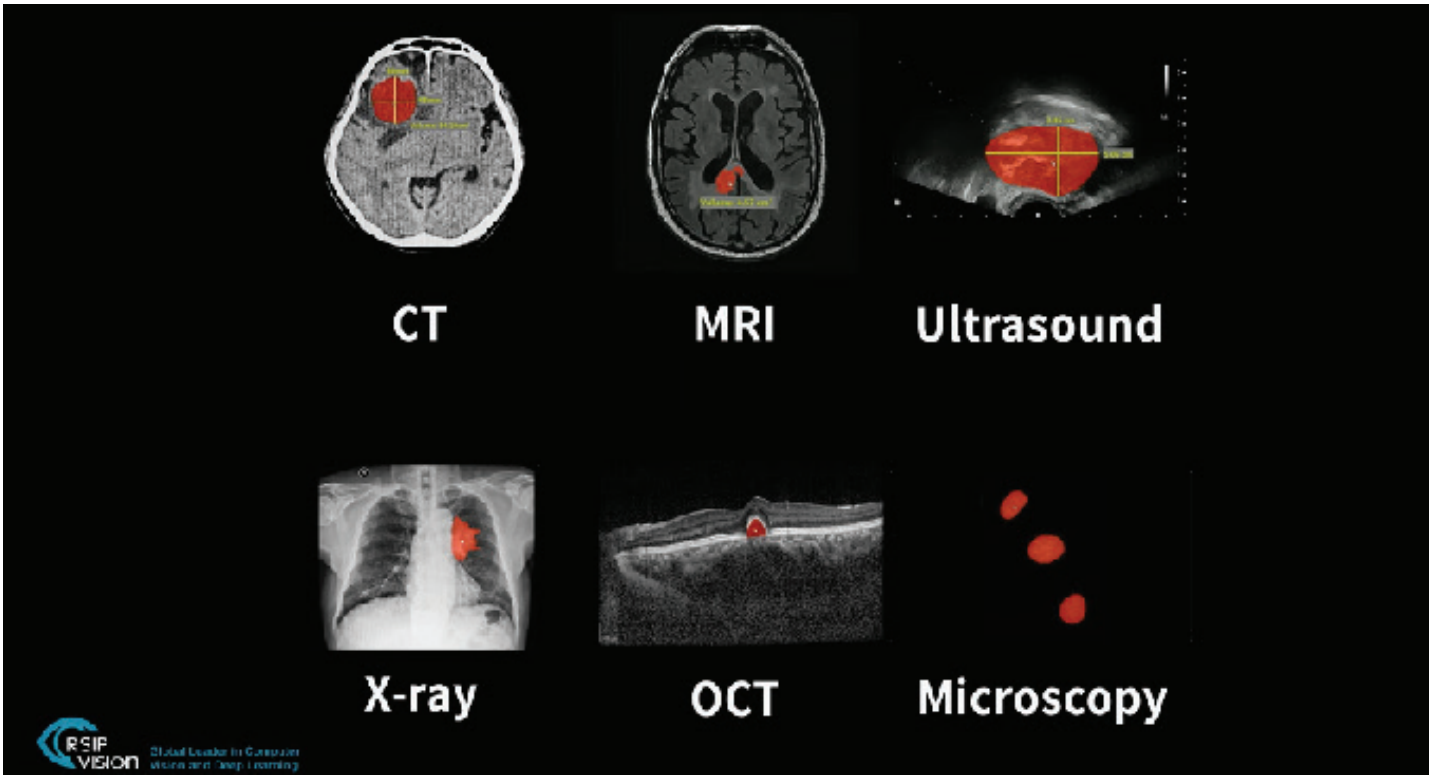
One-click segmentation is a semi-automated, AI-based segmentation and measurement tool for **detecting selected regions of interest and their boundaries quickly and automatically**. In medical imaging especially, this segmentation tool has many benefits:

- **High accuracy**- one-click segmentation not only delivers accurate segmentation of an ROI, but it also enables easy and intuitive corrections of selections. The physician or researcher can select and correct his/her choice of an ROI.
- **Generalization**- one-click segmentation can be generalized to "corner cases", including previously unseen anatomical structures. When dealing with anatomical structures with an unlimited combination of sizes, shapes and locations, generalization is key.
- **User impact**- one-click segmentation allows the user input to impact the segmentation process in a more direct way. The physician or researcher has full control of the choice of region to segment.

- **Minimal effort**- one-click segmentation requires minimal input (a single click) in order to deliver fast segmentation. The physician or researcher is unburdened by the arduous, time-consuming task of manual segmentation, resulting in more time left for diagnosis and treatment.

Have a peek at this short video and you will see how simple this is!

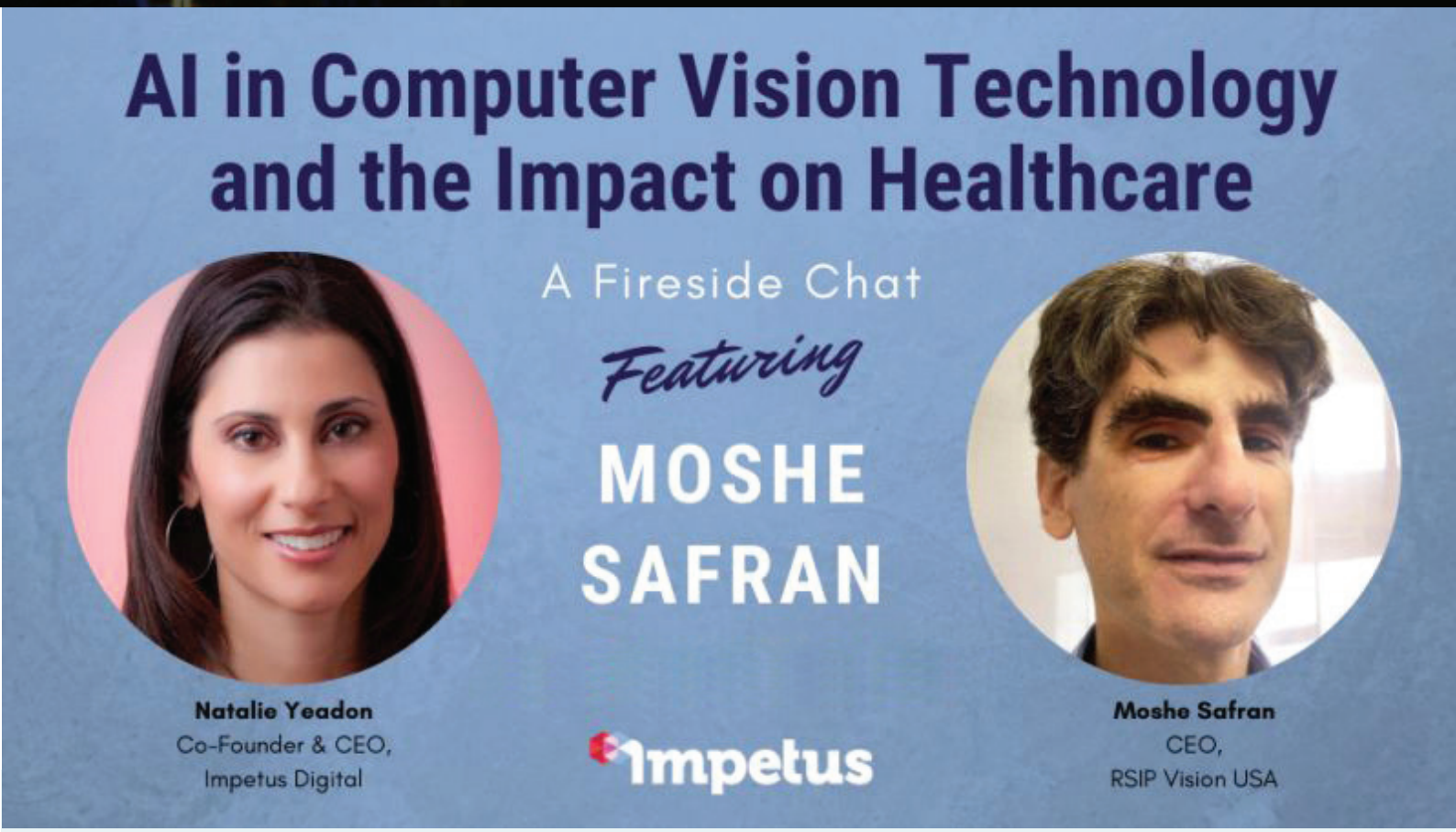
of selected ROIs using AI technology that is available **across all modalities**, including X-ray, CT scans, MR, surgical robotics, and pathology. For example, the tool can be used for one-click segmentation of lesions in multiple organs such as the lungs and the liver, across patient populations and cohorts. RSIP Vision's one-click segmentation delivers accurate andefficient anatomical measurements and



RSIP Vision, a leader in medical imaging innovations using advanced AI and computer vision solutions, has developed **a versatile one-click tool for medical image segmentation**. This AI-based, domain-agnostic tool delivers repeatable measurements leads to better overall treatment.

Read more about this innovative tool!

Watch now the "One-click segmentation" video above!



Here is the video of the fascinating fireside chat with **Moshe Safran**, CEO of RSIP Vision USA, about **Artificial Intelligence and its impact on healthcare.** Thank you to **Natalie Yeadon** of Impetus Digital for hosting this talk.

14 Congrats, Doctor!

Tobias Würfl (@wuerflts) recently completed his PhD **at Friedrich-Alexander University Erlangen-Nürnberg**, where he developed raw-data-based artifact compensation algorithms for Computed Tomography (CT). He currently works as an application developer for Magnetic Resonance Imaging on deep-learning-based reconstruction at **Siemens Healthineers**.

Our readers might remember Tobias telling us about the **workshop on Machine Learning for Medical Image Reconstruction (MLMIR)** at the latest **MICCAI**.

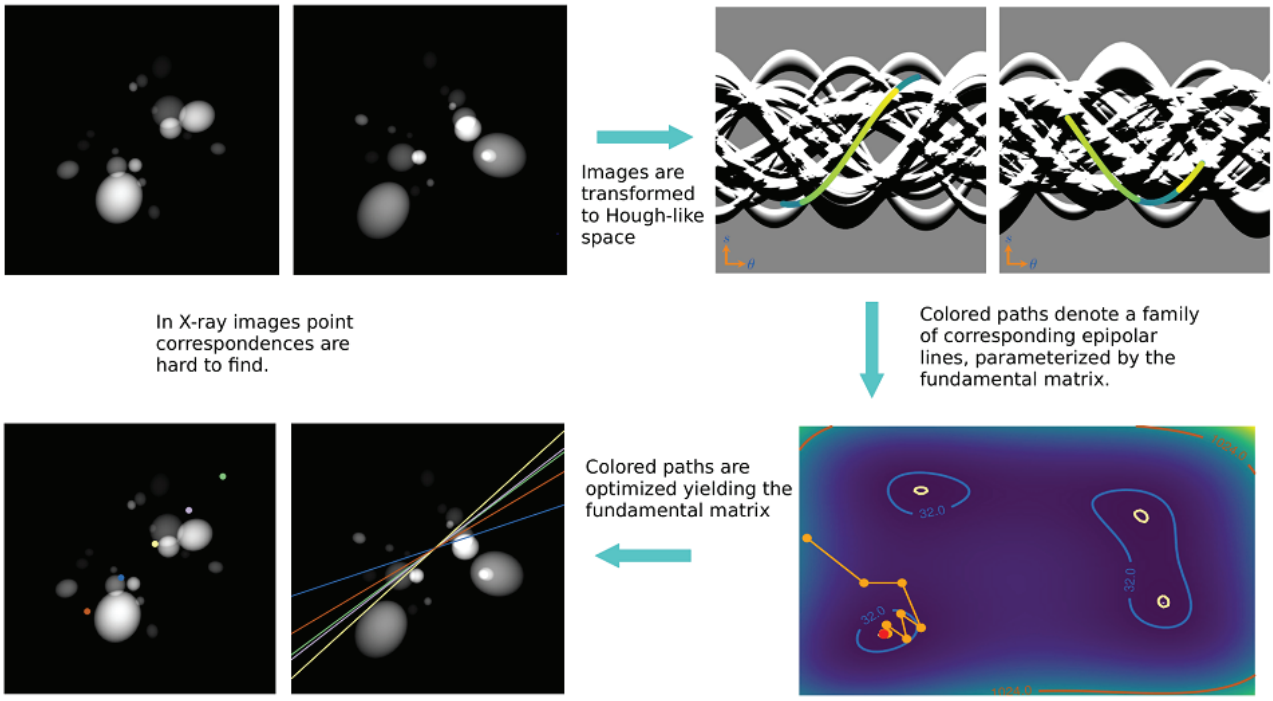
Congrats Tobias and best of luck from all of us!

Cone-beam (CB) Computed Tomography (CT) refers to modern CT systems which use flat-panel detectors acquiring two-dimensional X-ray images instead of a single line of values. They are commonly used e.g. in image-guided surgery as so called C-arm systems, in radiation therapy or in industrial computed tomography for inspecting wares.

The current challenge with CBCT is that the added flexibility comes at the price of reduced data quality compared to standard diagnostic CT. This reduced quality leads to artifacts in the reconstructed images since the different X-ray images are inconsistent with each other.

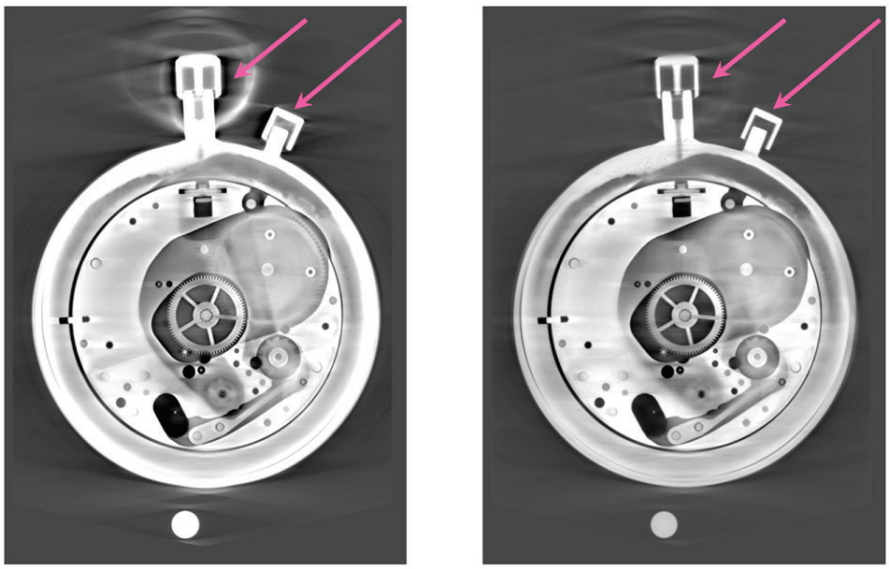
In my PhD work, I used this inconsistency between X-ray images to develop novel algorithms to compensate for artifacts in the reconstruction solely based on the raw measurement data itself.

I addressed two different classes of artifacts: 1) Artifacts which are caused by the polychromatic spectrum of conventional X-ray tubes called beam-hardening artifacts; and 2) Artifacts which arise due to motion of the object or inaccurate system calibration which can be subsumed as geometry artifacts.



The geometry artifacts can be solved by estimating the system geometry image-based, similar to SfM algorithms in Computer Vision. However, the key challenge in CBCT is that point correspondences are extremely hard to establish. Therefore, I developed a method to estimate the fundamental matrix without point correspondences. The fundamental matrix can then be used as input to a SfM method.

Follow the links for the full paper on fundamental matrix estimation and the work on beam hardening compensation or my thesis.



Uncompensated

Beam hardening compensated

Using JAX to accelerate deep learning research

by Ioannis Valasakis, King's College London
(Twitter/LinkedIn/GitHub as @wizofe)

Welcome to this month's code discussion. I hope that you are all staying safe and are aware of any local restrictions due to the pandemic. It's really important to be careful and keep the numbers now as in many cases (such as in the UK) the death toll is constantly rising.

This month, I am presenting a tool that can accelerate deep learning research using **Autograd and XLA. JAX** is developed by **Google Research** teams for **high-performance machine-learning research**.

In this article, I am going to give a short introduction to JAX and present some tools to use with it!

Differentiation

Autograd is a project which is no longer developed (although actively maintained). It can handle a large subset of **Python's features**. Importantly, **it supports reverse-mode differentiation i.e. backpropagation**. It takes gradients of scalar-valued functions with respect to array-valued arguments and forward-mode differentiation, and the two can be composed arbitrarily.

Here's an example from the documentation:

```
>>> import autograd.numpy as np  # Thinly-wrapped numpy
>>> from autograd import grad    # The only autograd function you may ever
need
>>>
>>> def tanh(x):                 # Define a function
...     y = np.exp(-2.0 * x)
...     return (1.0 - y) / (1.0 + y)
...
>>> grad_tanh = grad(tanh)       # Obtain its gradient function
>>> grad_tanh(1.0)               # Evaluate the gradient at x = 1.0
0.41997434161402603
>>> (tanh(1.0001) - tanh(0.9999)) / 0.0002  # Compare to finite differences
0.41997434264973155
```

JAX supports a very similar API like the one from Autograd while it extends it with more automated features. Here's the respective example from JAX:

```
from jax import grad
import jax.numpy as jnp

def tanh(x):  # Define a function
  y = jnp.exp(-2.0 * x)
  return (1.0 - y) / (1.0 + y)

grad_tanh = grad(tanh)  # Obtain its gradient function
print(grad_tanh(1.0))   # Evaluate it at x = 1.0
# prints 0.4199743
```

Using the function grad one can differentiate to any order of gradient, e.g.

```
print(grad(grad(grad(tanh)))(1.0))
# prints 0.62162673
```

Hessian matrices can be computed, essentially by using the hessian function, which supports nested Python containers as inputs and outputs.

```
>>> import jax.numpy as jnp
>>> f = lambda dct: {"c": jnp.power(dct["a"], dct["b"])}
>>> print(jax.hessian(f)({"a": jnp.arange(2.) + 1., "b": jnp.arange(2.) + 2.}))
{'c': {'a': {'a': DeviceArray([[[ 2.,  0.], [ 0.,  0.]],
                               [[ 0.,  0.], [ 0., 12.]]], dtype=float32),
             'b': DeviceArray([[[ 1.    ,  0.    ], [ 0.    ,  0.    ]],
                               [[ 0.    ,  0.    ], [ 0.    , 12.317766]]], dtype=float32)},
       'b': {'a': DeviceArray([[[ 1.    ,  0.    ], [ 0.    ,  0.    ]],
                               [[ 0.    ,  0.    ], [ 0.    , 12.317766]]], dtype=float32),
             'b': DeviceArray([[[0.    , 0.    ], [0.    , 0.    ]],
                               [[0.    , 0.    ], [0.    , 3.843624]]], dtype=float32)}}}
```

Compiling expressions

To compile JAX expressions XLA can be used (end-to-end compilation) with `jit`. There are two ways to use `jit`: either as a decorator (`@jit`) or as a higher-order function.

```
import jax.numpy as jnp
from jax import jit

def slow_f(x):
  # Element-wise ops see a large benefit from fusion
  return x * x + x * 2.0

x = jnp.ones((5000, 5000))
fast_f = jit(slow_f)
%timeit -n10 -r3 fast_f(x)  # ~ 4.5 ms / loop on Titan X
%timeit -n10 -r3 slow_f(x)  # ~ 14.5 ms / loop (also on GPU via JAX)
```

Debugging

One great feature of JAX, which is very often missing in the field of deep learning (or at least it's not equally represented in relation to the broader software-engineering perspective) is "debugging".

18 Computer Vision Tool

It is often very difficult to debug functions or even simple event compilations. JAX provides a **very simple API** to trace where NaN's are occurring in the functions or gradients. A documentation example is the following:

```
In [1]: import jax.numpy as jnp

In [2]: jnp.divide(0., 0.)
---------------------------------------------------------------------------
FloatingPointError                        Traceback (most recent call last)
<ipython-input-2-f2e2c413b437> in <module>()
----> 1 jnp.divide(0., 0.)

.../jax/jax/numpy/lax_numpy.pyc in divide(x1, x2)
    343     return floor_divide(x1, x2)
    344   else:
--> 345     return true_divide(x1, x2)
    346
    347

.../jax/jax/numpy/lax_numpy.pyc in true_divide(x1, x2)
    332   x1, x2 = _promote_shapes(x1, x2)
    333   return lax.div(lax.convert_element_type(x1, result_dtype),
--> 334                  lax.convert_element_type(x2, result_dtype))
    335
    336

.../jax/jax/lax.pyc in div(x, y)
    244 def div(x, y):
    245   r"""Elementwise division: :math:`x \over y`."""
--> 246   return div_p.bind(x, y)
    247
    248 def rem(x, y):

... stack trace ...

.../jax/jax/interpreters/xla.pyc in handle_result(device_buffer)
    103         py_val = device_buffer.to_py()
    104         if np.any(np.isnan(py_val)):
--> 105           raise FloatingPointError("invalid value")
    106         else:
    107           return DeviceArray(device_buffer, *result_shape)

FloatingPointError: invalid value
```

The generated NaN was caught as an invalid value above. What if we need a debugger for the functions under `@jit`? The solution is to run `%debug` as shown below.

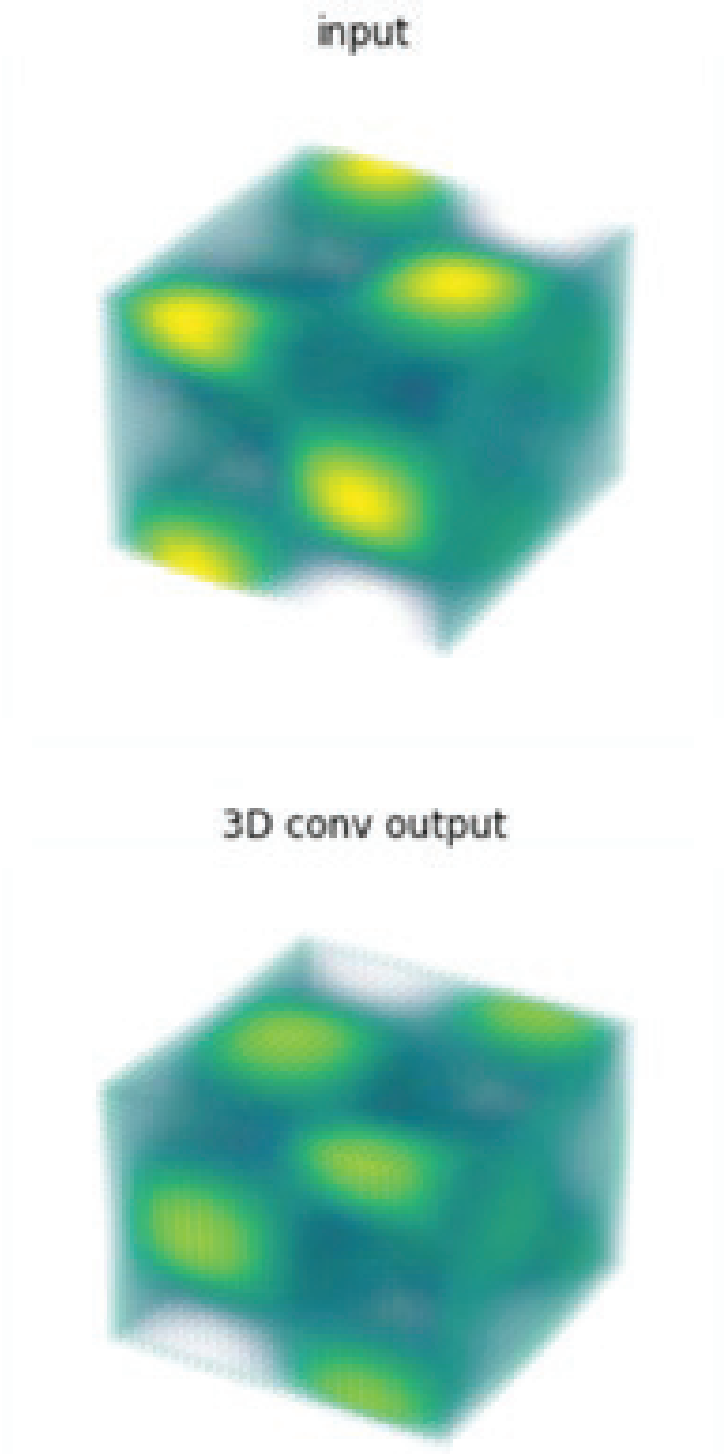
```
In [4]: from jax import jit

In [5]: @jit
   ...: def f(x, y):
   ...:     a = x * y
   ...:     b = (x + y) / (x - y)
   ...:     c = a + 2
   ...:     return a + b * c
   ...:

In [6]: x = jnp.array([2., 0.])

In [7]: y = jnp.array([3., 0.])

In [8]: f(x, y)
Invalid value encountered in the output of a jit function. Calling the de-optimized version.
---------------------------------------------------------------------------
FloatingPointError                        Traceback (most recent call last)
<ipython-input-8-811b7ddb3300> in <module>()
----> 1 f(x, y)

 ... stack trace ...

<ipython-input-5-619b39acbaac> in f(x, y)
      2 def f(x, y):
      3     a = x * y
----> 4     b = (x + y) / (x - y)
      5     c = a + 2
      6     return a + b * c

.../jax/jax/numpy/lax_numpy.pyc in divide(x1, x2)
    343     return floor_divide(x1, x2)
    344   else:
--> 345     return true_divide(x1, x2)
    346
    347

.../jax/jax/numpy/lax_numpy.pyc in true_divide(x1, x2)
    332   x1, x2 = _promote_shapes(x1, x2)
    333   return lax.div(lax.convert_element_type(x1, result_dtype),
--> 334                  lax.convert_element_type(x2, result_dtype))
    335
    336

.../jax/jax/lax.pyc in div(x, y)
    244 def div(x, y):
    245   r"""Elementwise division: :math:`x \over y`."""
--> 246   return div_p.bind(x, y)
    247
    248 def rem(x, y):

 ... stack trace ...
```

Convolutions

The final example shows how to use a random 3D kernel and display a 3D convolution (in the NHWDC pattern).

```
# Random 3D kernel - HWDIO layout
kernel = np.array([
  [[0, 0,  0], [0,  1,  0], [0,  0,   0]],
  [[0, -1, 0], [-1, 0, -1], [0,  -1,  0]],
  [[0, 0,  0], [0,  1,  0], [0,  0,   0]]],
  dtype=jnp.float32)[:, :, :, np.newaxis, np.newaxis]

# 3D data - NHWDC layout
data = np.zeros((1, 30, 30, 30, 1), dtype=jnp.float32)
x, y, z = np.mgrid[0:1:30j, 0:1:30j, 0:1:30j]
data += (np.sin(2*x*jnp.pi)*np.cos(2*y*jnp.pi)*np.cos(2*z*jnp.pi))[None,:,:,:,None]

print("in shapes:", data.shape, kernel.shape)
dn = lax.conv_dimension_numbers(data.shape, kernel.shape,
                                ('NHWDC', 'HWDIO', 'NHWDC'))
print(dn)

out = lax.conv_general_dilated(data,    # lhs = image tensor
                               kernel,  # rhs = conv kernel tensor
                               (1,1,1), # window strides
                               'SAME',  # padding mode
                               (1,1,1), # lhs/image dilation
                               (1,1,1), # rhs/kernel dilation
                               dn)      # dimension_numbers
print("out shape: ", out.shape)

# Make some simple 3d density plots:
from mpl_toolkits.mplot3d import Axes3D
def make_alpha(cmap):
  my_cmap = cmap(jnp.arange(cmap.N))
  my_cmap[:,-1] = jnp.linspace(0, 1, cmap.N)**3
  return mpl.colors.ListedColormap(my_cmap)
my_cmap = make_alpha(plt.cm.viridis)
fig = plt.figure()
ax = fig.gca(projection='3d')
ax.scatter(x.ravel(), y.ravel(), z.ravel(), c=data.ravel(), cmap=my_cmap)
ax.axis('off')
ax.set_title('input')
fig = plt.figure()
ax = fig.gca(projection='3d')
ax.scatter(x.ravel(), y.ravel(), z.ravel(), c=out.ravel(), cmap=my_cmap)
ax.axis('off')
ax.set_title('3D conv output');
```

```
in shapes: (1, 30, 30, 30, 1) (3, 3, 3, 1, 1)
ConvDimensionNumbers(lhs_spec=(0, 4, 1, 2, 3), rhs_spec=(4, 3, 0, 1, 2), out_spec=(0, 4, 1, 2, 3))
out shape:  (1, 30, 30, 30, 1)
[45]:
Text(0.5, 0.92, '3D conv output')
```



Wrapping up!

This was a very quick introduction to JAX! If you are interested in learning more, you can always find exciting projects (GitHub). In the magazine of April, I plan to continue with some more interesting projects used by companies such as **DeepMind**, which are part of the JAX ecosystem (or utilise it). Stay tuned!☺

EndoCV2021 Challenge

Addressing generalisability in polyp detection and segmentation

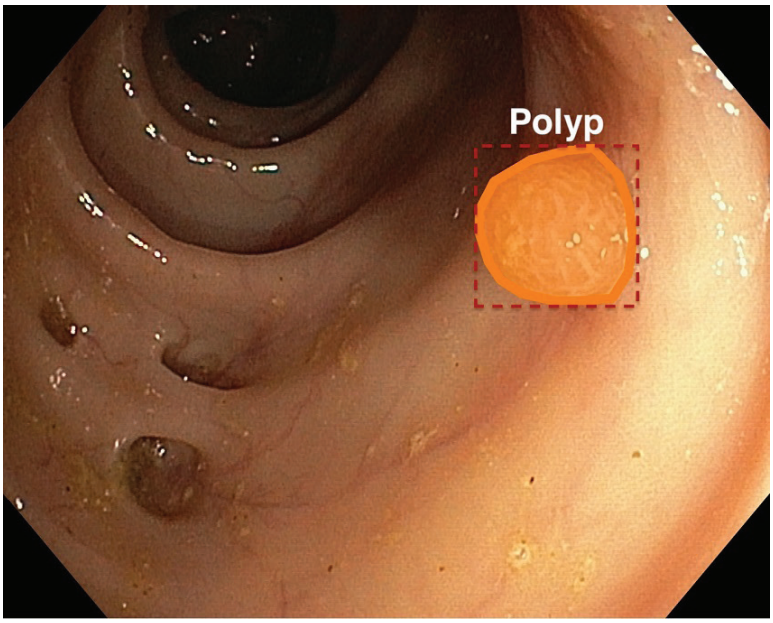
The 3rd **EndoCV2021** International challenge and workshop will be held in conjunction with **IEEE ISBI** on 13th April 2021. It is aimed at "**Polyp detection and segmentation focusing generalization ability of deep learning methods**" and is now open for participation. Polyps in the colon are the **early cancer precursors** that need to be detected and removed during colonoscopy. Automated systems using deep learning techniques for example can help to localize and assist in removal.

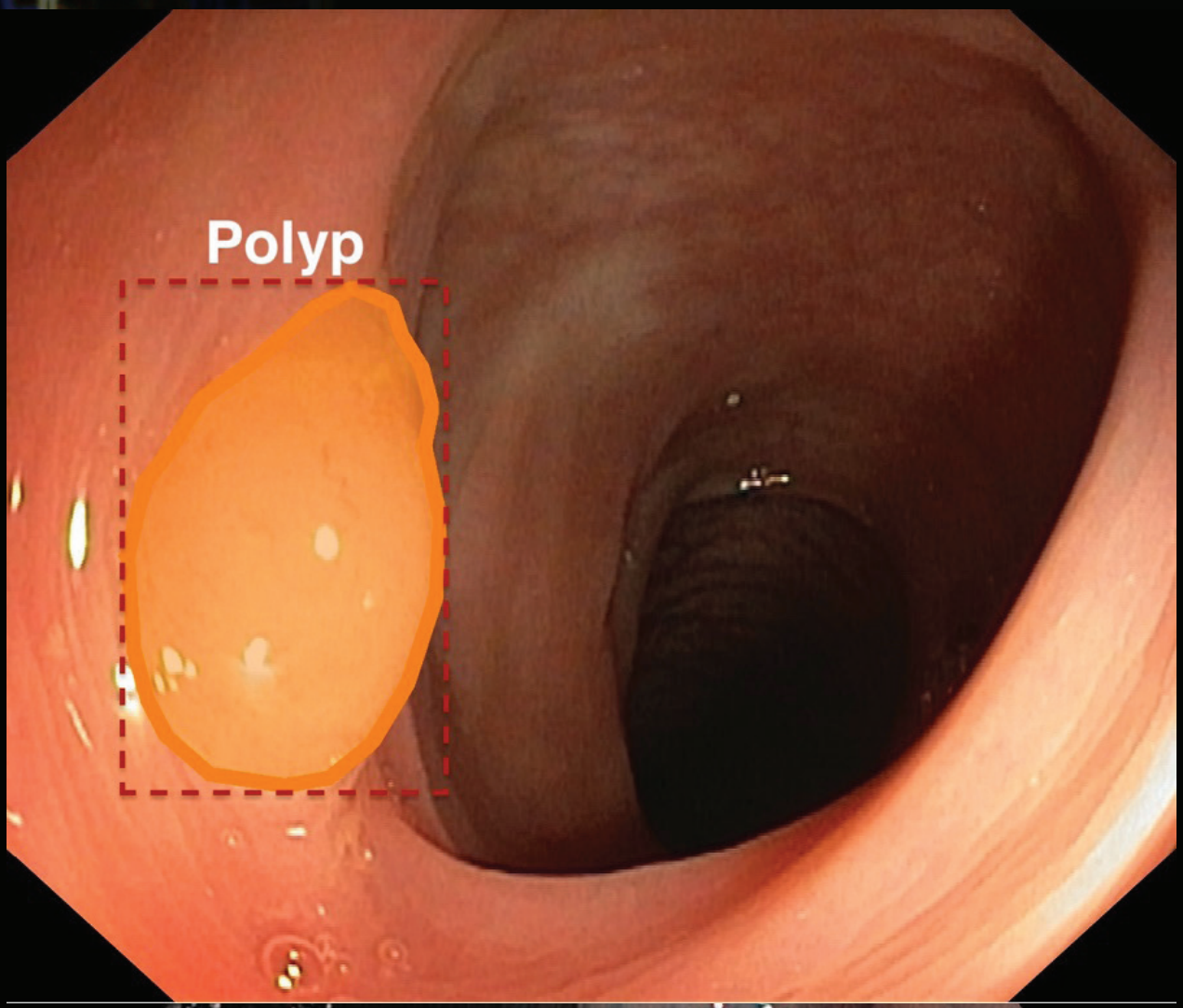
With this challenge, the organizers (**Sharib Ali**, **Debesh Jha** and **Noha Ghatwary**) aim to release annotations of more than 2000 endoscopy video frames collected from 6 different hospitals. This challenge will reveal the potential of **deep learning methods for clinical translation**. The organizers say, "*This is a major bottleneck and necessary for clinical deployment of currently existing or new methods.*"

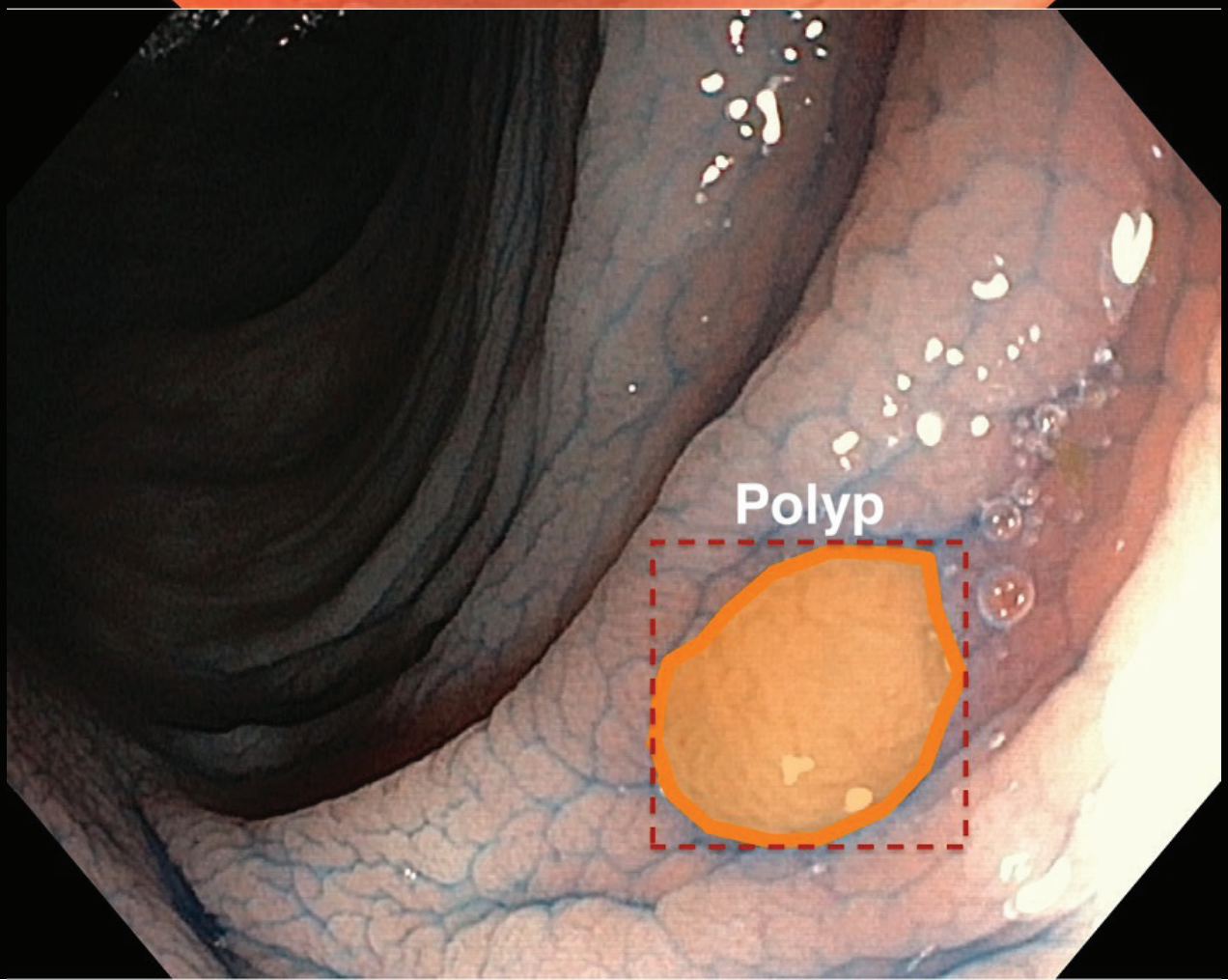
Every year EndoCV invites distinguished keynote speakers for the workshop. Sharib Ali says, "*we encourage interested researchers to participate in this challenge and submit their work to the workshop.*" The leaderboard submission for EndoCV2021 starts on February 20 while the training data is due to be released on February 2.

Important dates:

- Leaderboard submission start: 20 February 2021
- 4-page paper submission: 20 March 2021 (peer-reviewed) published in CEUR online
- Challenge and workshop day: 13 April 2021




Polyp


Polyp

24 Women in Computer Vision

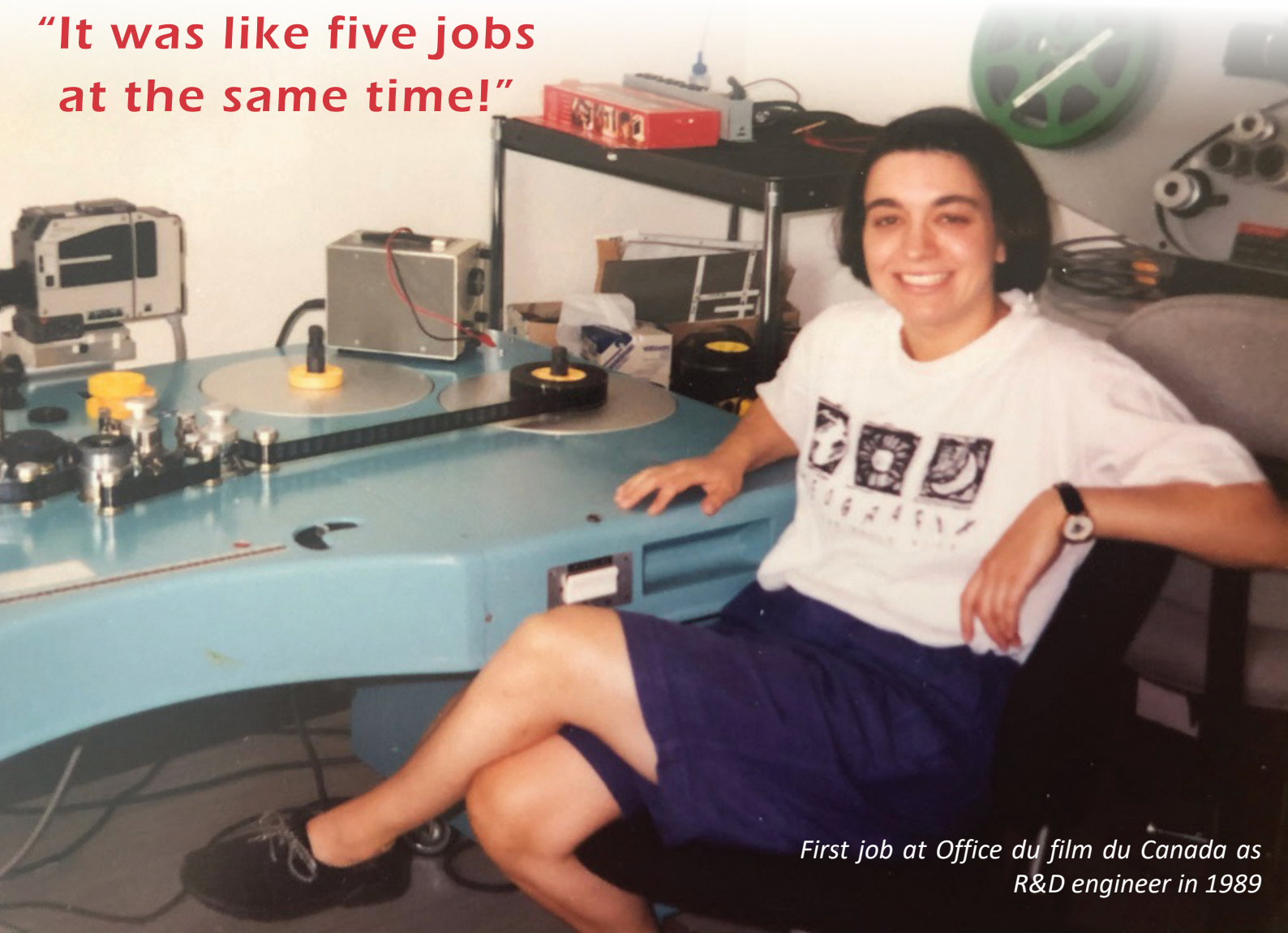
Currently a consultant, Myriam Côté has served as the Director of Projects and Industrial R&D at the University of Montreal for the past several years. More recently, she became the Director of AI for Humanity at MILA, the Montreal Institute for Learning Algorithms founded by Yoshua Bengio. She holds a PhD from what is now called Télécom Paris. This interview was conducted in French. Thank you to Yoshua Bengio for putting us in touch with Myriam. Click here to find 100 more interviews with Women in Computer Vision!

Myriam, how did you get into this impressive career?

I first studied Music, but I was very curious and I loved science, so I became an engineer in photonics optics and AI. I got my PhD in AI and went to work in the industry in the field of technology transfer. I wanted to do something practical with science: I had the romantic vision that an engineer must be creative, to invent things and find solutions. I needed to express my creativity and imagination.

We were 12 girls in a class of 100! I do not believe that women are attracted

"It was like five jobs at the same time!"



First job at Office du film du Canada as R&D engineer in 1989

to technique itself: they want to be able to do something with it and have an impact on people.

I taught as a professor and I worked in research. Then I joined a small and very innovative enterprise as VP of Engineering: SPG Data. They were specialized in vision and non-destructive analysis: they scanned aircraft parts for maintenance purposes and turned them into 3D objects. There, I did inspection and characterized fault detection.

How did that fit with your motivations at the time?
My professorship was very demanding. It was like five jobs at the same time. So I wanted to be an engineer again who works with real things. Team management and the human aspect of business were among the many things that I learned doing that. Managing people and working with them was a whole new world to me. This is not something that you learn doing a PhD or a post-doc, and I enjoyed dealing with that.

Did this experience make you progress in this human aspect?
Yes, of course! I learned extraordinary things!

We want to know them!
In a private business, you must learn to communicate ideas and be able to understand the needs of clients,

> *"I had the romantic vision that an engineer must be creative, to invent things and find solutions."*

manage their expectations and explain what you are doing: you need to listen a lot. All the soft skills that you do not learn in a classroom and by which you can see the impact of your technical work. You need to deliver something which fills the client's needs. When this works well, you can see the impact of your work, which is very rewarding. You see the effects, and you make people happy. Doesn't everybody enjoy making people happy?

What else did you learn then?
I also liked to mix the creative aspect of my work with technology and IT. I worked for Avtoma, a small Italian company with a research office in Montreal. I developed algorithms for object recognition in movie sequences to create special effects, like integrating synthetic characters in real world scenes. Working for small companies is so different from academia. University projects take years, while companies compete 24 hours a day with the whole planet and with researchers that work on the same subjects as you do. At the

Party at parents home in Quebec - Myriam's father is a MD, but she's the doctor!

time, the market for synthetic images was just beginning and the company went bust. I had to learn to adapt to market chaos and land on my feet.

As a young engineer, who wanted to express herself and have an impact, I was very naïve. It would have been easier if I had been trained to understand markets and trends, all things that an engineer is never taught.

How did you get to work with Yoshua Bengio?

I learned about his work while doing my thesis at Télécom Paris. Later on, when I needed to do defect classification, I thought that his neural networks might give me a very robust solution. I spoke with Yoshua in Montreal and shared my work with him to get his advice. Everything I told him worked very well when he tried it, so he offered me to work with him. So I joined MILA; at that time it was still called LISA, and it employed 30 people: professors, post-docs, students and staff. I was already experienced in research, in industrial project management, and I had a doctorate in AI. Yoshua wanted

help with the contracts that the lab had with private corporations. I managed the organization until it reached 300 people.

> "I want the last years of my career to count, to be relevant and to give me pleasure."

I think all my readers want to know what it is like working with Yoshua. We worked ten years together, so I got to know him well. What always struck me is his creative, artistic side, to which I can relate. His imagination is limitless. I think he owes it to the education that he received from his parents, who immigrated to France from Morocco. They lived through the freedom movement of May 1968 in Paris. They enrolled him as a very young child in a math learning program, and he became very fond of math in a creative environment, where you're encouraged to think outside of the box. That made us always get along very well. I focused on organization and management, which freed his time for what he really likes: research. He always sees opportunities to grab fast. He thinks quickly, he's brilliant. His memory is extraordinary! Yoshua is also very generous with colleagues: he shares his ideas, and he supports free access to research, a humanistic approach inherited from his parents' values. He knows how to talk to people's emotions and how to encourage them to get involved.

At Mila in 2019, a workshop led by Amir Banifatemi (CEO XPrize) - Standing and smiling in the center of the photo, Yoshua Bengio

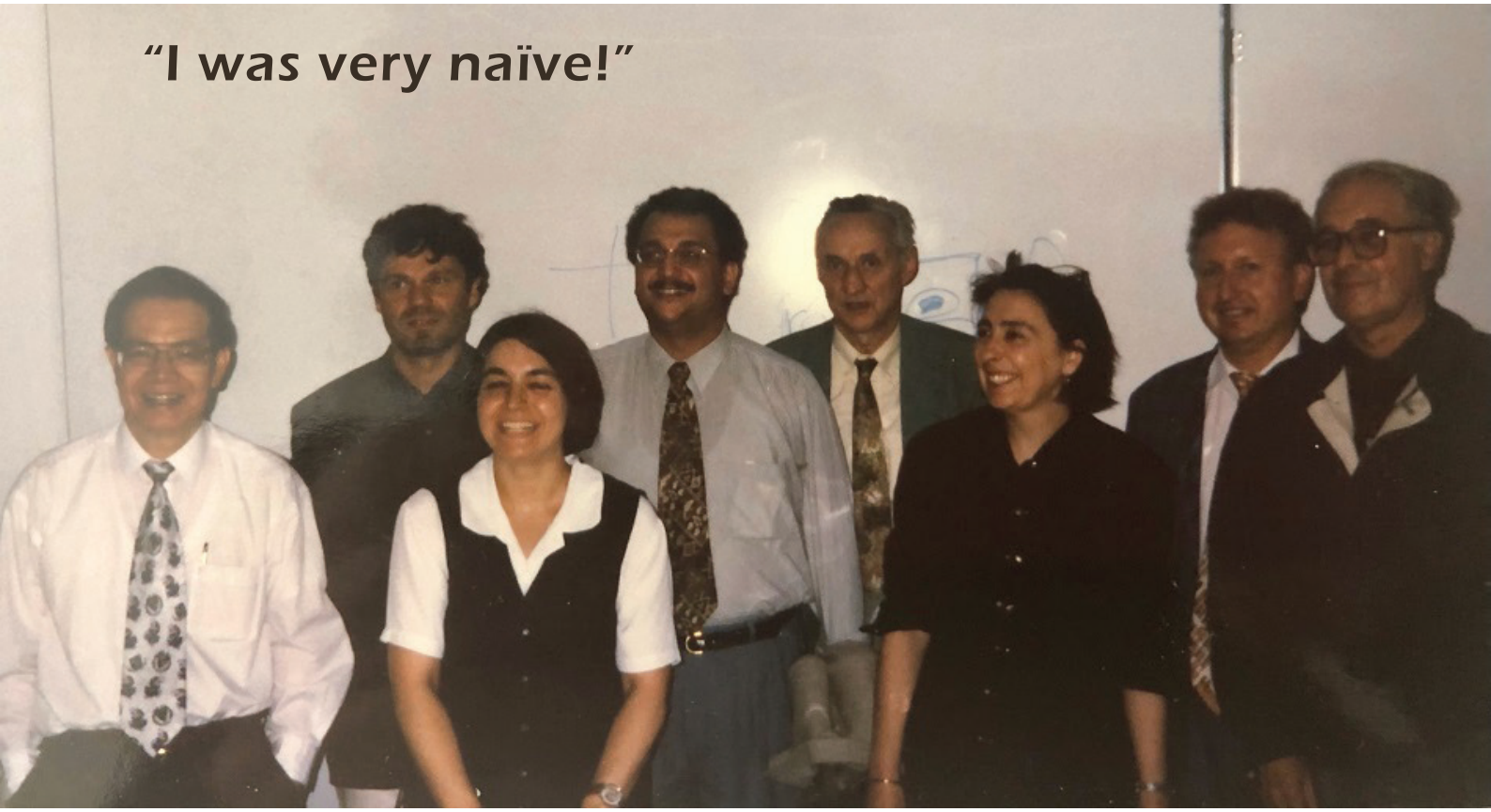
Back to you: what is your focus now?
MILA has changed a lot over the years, and I felt it was time for me to let a new management take the lead. Montreal has become an artificial intelligence hub, and I thought that I could help AI develop in Europe. I joined a Belgian company with a wonderful, philanthropic mission: help language learning in order to bring people together. This was very much in line with my last position at MILA, as Director of AI for Humanity. But as soon as I got my visa for Belgium, the Covid-19 pandemic arrived, and that was a bummer. I am now in a transition doing consulting work: I like to build things, make things start. I help with technology transfer, and I coach employees with all the things that I learned through my career. I enjoy my perpetual challenge to adapt and to be flexible, building on the self-confidence that I learned through the years.

What's next?
I want the last years of my career to count, to be relevant and to give me pleasure. That is the time when I can have the most impact and the largest contribution to the community.

Find out 100 more interviews like this with Women in Science!

Read our fascinating interview with Yoshua Bengio!

PhD defense at Paris Télécom in 1997



W4PR: Women at ICPR Workshop

by Marcella Cornia

The 1st Women at ICPR Workshop (W4PR) was held during the International Conference on Pattern Recognition (ICPR) 2020, with the aim of creating the opportunity for young and senior women scientists to interact and connect with each other. This workshop was built upon past social events (coffee break and lunch) for women organized at **ICPR 2016** (Cancun) and **ICPR 2018** (Beijing).

Unfortunately, due to the pandemic situation, this year ICPR was held online and all workshops and tutorials had a dedicated Zoom room to conduct all talks and interactions between participants. W4PR took place on the first day of the conference on January

Ingela Nyström

Linda O'Gorman

10. For me, while contributing to the organization of the workshop, it was a great opportunity to learn from amazing and inspiring professors and researchers who put a lot of effort into organizing the best possible event.

The workshop started with a welcome from **Rita Cucchiara**, one of the ICPR General Chairs and professor at the University of Modena and Reggio Emilia, and continued with two invited talks from **Ingela Nyström**, professor at the Uppsala University, and **LindaO'Gorman**, the IAPR (International Association of Pattern Recognition) Secretariat. The first talk was related to equal opportunities in academia with a specific focus on the Sweden situation, while the second focused on the evolution of the role of women at the IAPR. Both provided an interesting perspective on women's careers in two different important institutions and gave a starting point for analyzing the reasons behind the imbalance between the number of men and women in academia. I found it particularly interesting to see that in recent years there has been a growth in the presence of women both in international research institutions and in the internal boards of important associations such as the IAPR.

To foster the interaction between young and senior researchers, each talk was followed by a breakout room session where all participants were divided into small groups to discuss the most important points that had emerged before and connect

people with each other. It was a great opportunity to know the academic situation in different countries and interact with researchers from all over the world. The workshop ended with a **panel discussion** on recent successes and upcoming challenges for women in a male-dominated field.

events is a special occasion to highlight the technical contributions of women to the fields of computer vision and pattern recognition and to create a large network in which all researchers, both men and women, can exchange ideas and interact with each other.

My personal take-home message from the workshop: while there is still a discrepancy between male and female researchers in faculty positions, wehave done and are still doing a lot to foster the presence of women in our field. Promoting this equality is extremely important, but to do so, we cannot simply apply the strategy of treating women as a "**protected species in pattern recognition**": this is too simplistic and thus, not viable. For this reason, the organization of these

Hope to see you at the next ICPR in Montreal, hopefully in-presence and with the 2nd Women at ICPR Workshop!

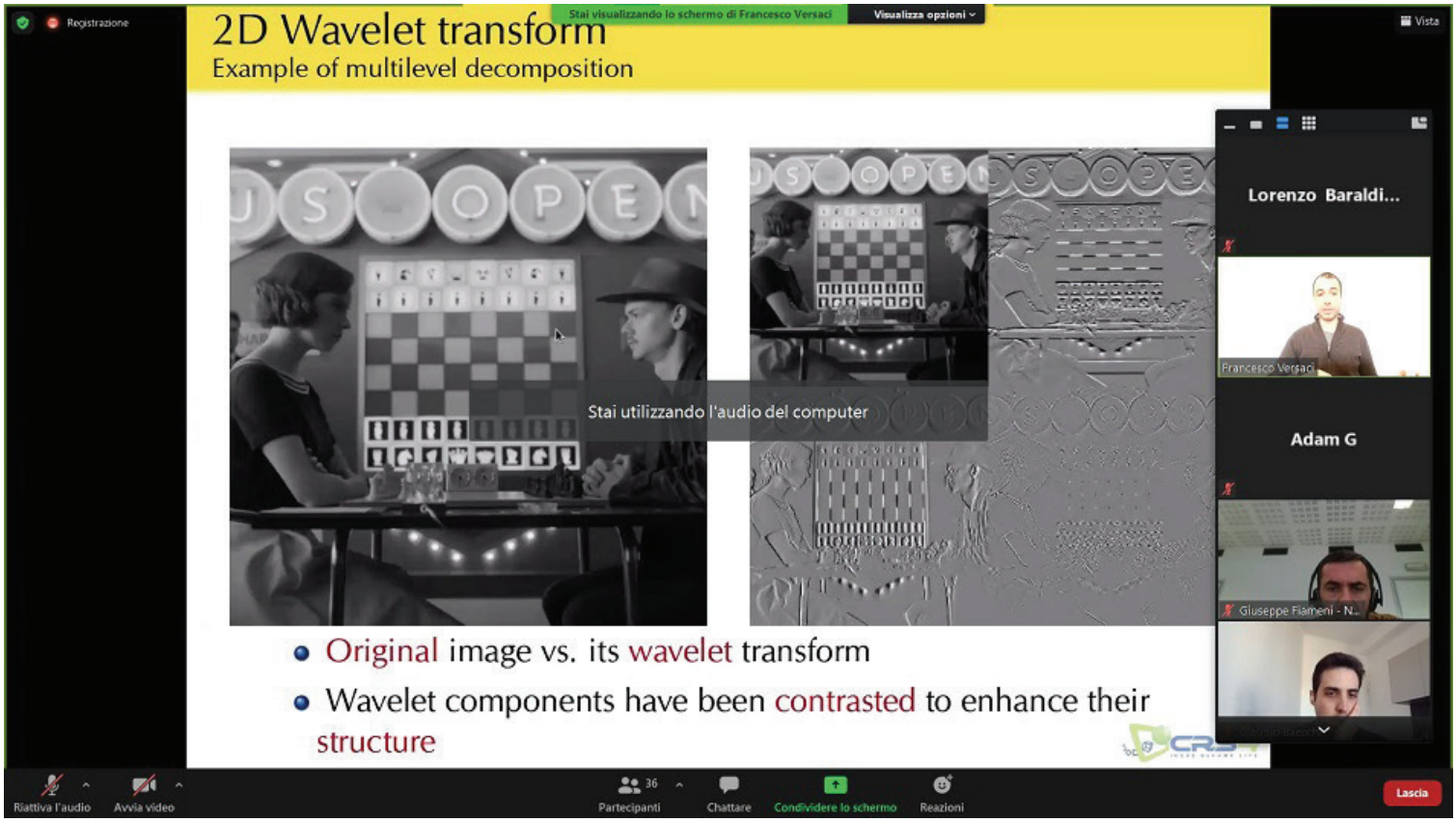
Do you believe that women can be great AI scientists?

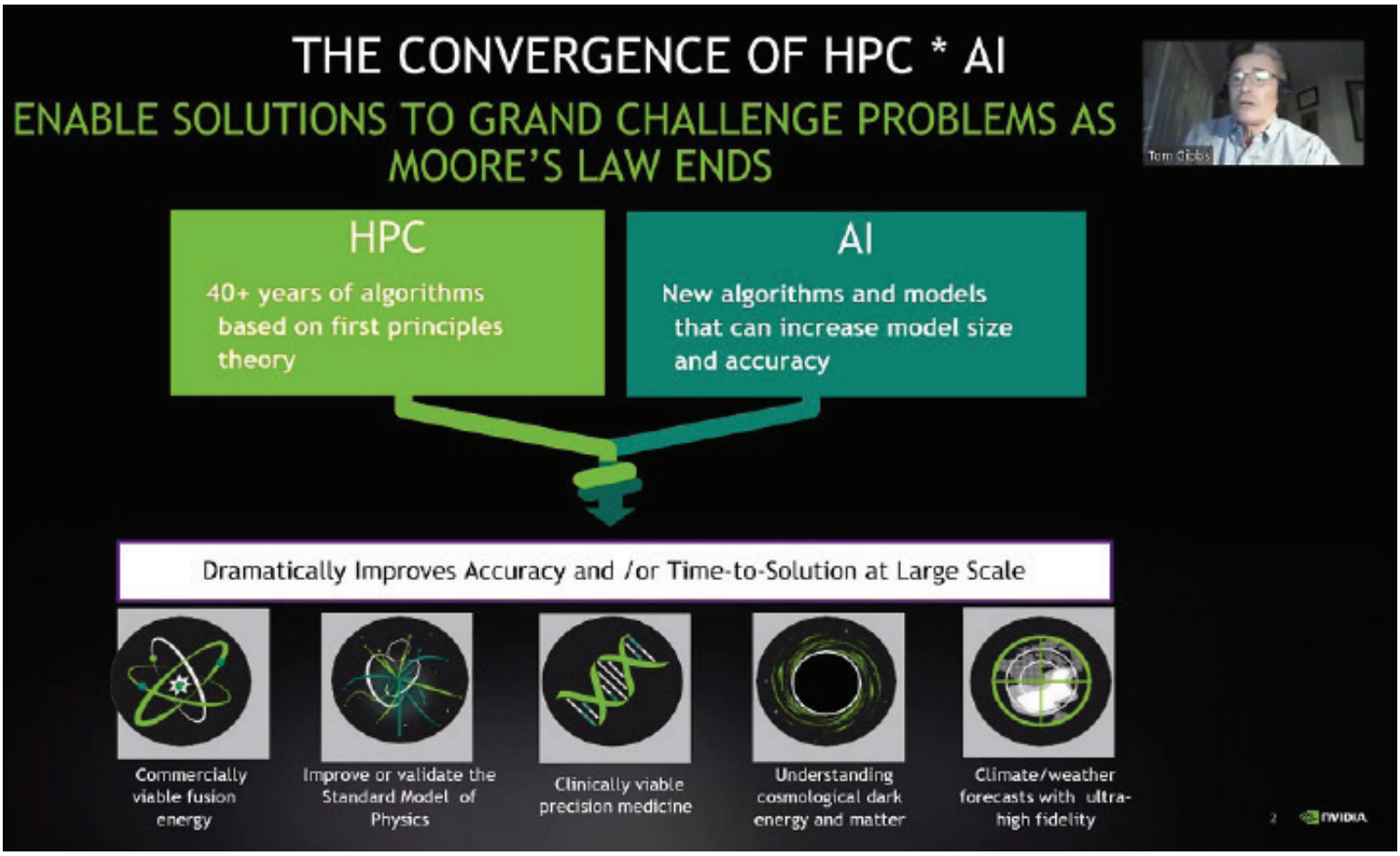
We also do. Here are 100 of them, in their own words!

The first edition of the Workshop on Computational Aspects of Deep Learning (CADL 2020) has been held virtually on January, 11. The workshop has been organized in conjunction with ICPR 2020 by Lorenzo Baraldi (University of Modena and Reggio Emilia), Frederic Pariente (NVIDIA, Deputy Directory of NVAITC in EMEA), Claudio Baecchi (University of Florence), and Iuri Frosio (NVIDIA). All this was in 2021, of course!

by Lorenzo Baraldi

Deep Learning has radically changed our research methodology towards a data-oriented approach, in which learning involves all steps of the prediction pipeline. In this new reality, the optimization and careful design of neural architectures play an increasingly important role which affects the research pace and the effectiveness and applicability of Deep Learning models. The **CADL**





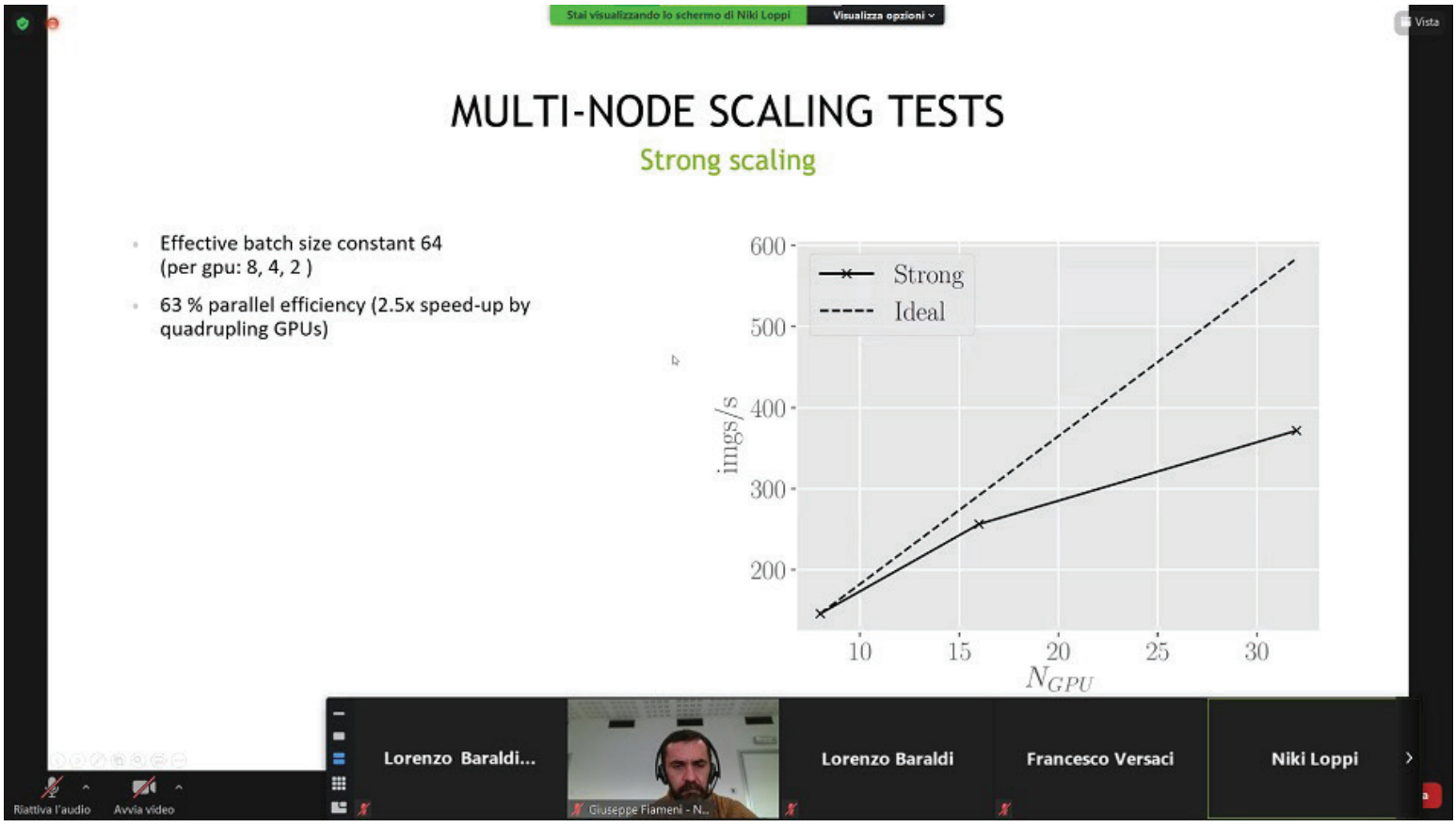
Workshop at ICPR2020 was organized with the goal of bringing the focus of the community on computational challenges, and as an opportunity to present solutions and research works in this direction. This first edition has focused on the **development of optimized architectures**, particularly on highly scalable systems, tackling the training on large-scale datasets, the **design of novel architectures** and operators, and the training on multi-**node systems** such as HPC clusters.

The workshop started in the morning with an invited tutorial, organized by **Adam Henryk Grzywaczewski**, on the parallel training of **Neural Networks**. Adam showed, through a critical reviewof the literature and practical examples, how to deal with the challenges of very large batch training, a setting that is very frequent in HPC and highly distributed Deep Learning.

The second part of the day has instead been dedicated to the presentation of **ten research works**, which focused on optimization and computational challenges. Beyond research papers, CADL also featured two exciting invited talks in the middle of the program. The first talk was given by **Tom Gibbs**, Vice President of Business Development at **NVIDIA**, who has been recently awarded the **ACM Gordon Bell prize** this year. He

gave a successful and exciting talk on the convergence of AI and HPC, and its exploitation to solve grand challenges, like molecular simulation. The second invited talk was instead given by **Mirko Cestari**, who leads the HPC and Cloud Technology team at **CINECA, the Italian Supercomputing Centre**. Mirko revealed an exciting preview of **Leonardo, the Next EuroHPC SuperComputer**, which will be employed to fuel the Italian andEuropean research on HPC, and which will be available in the next year. Leonardo will have a peak performance of **200+ PFlop/s, 3 PB of RAM, and 150 PB of storage**. It will host more than **14.000 GPU accelerators**. The organizing committee of CADL has also decided to award one of the ten orally presented papers with an NVIDIA Titan RTX GPU, to encourage further research in the field of optimized Deep Learning and HPC. The winner was **Francesco Versaci** from CRS4, who presented a fast library to perform the 2D wavelet transform on standard Deep Learning libraries.

The first edition of CADL ended with a discussion on the directions and aspirations for the community of AI and HPC, and with ideas for the next edition of the workshop. The success of CADL has demonstrated that the community at the intersection of AI and HPC is growing and that the interest in **highly scalable systems to solve grand challenges and tackle research problems** is quickly becoming important.



COMPUTER VISION EVENTS

ICAART 2021
Online Streaming
Feb 4 - 6

ICPRAM 2021
Online Streaming
Feb 4 - 6

VISIGRAPP 2021
Online Streaming
Feb 8 - 10

BIOSTEC 2021
Online streaming
Feb 11-13

SPIE Medical Imaging
Online forum
Feb 15 - 19

Ai4 Finance
Taking place digitally
March 3-4

AI & Big Data Expo Global 2021
Fully Virtual Conference
Mar 17 - 18

AI in Healthcare & Pharma
Online
March 25-26

IEEE ISBI 2021
Full virtual conference
April 13-16

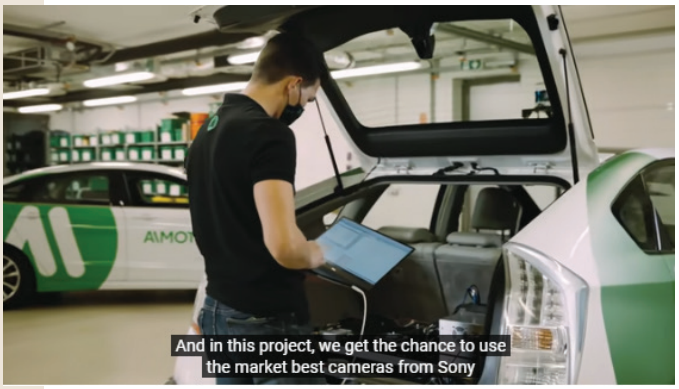
AI Applications
Virtual Summit Online
April 14-15

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!

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!**

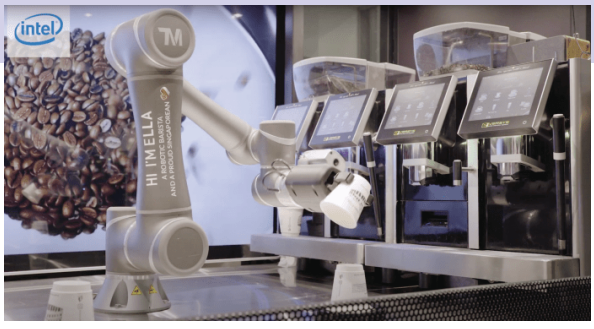
Collaboration of Almotive and Sony Evolves ADAS Capability



Our readers certainly remember **Almotive**, a promising autonomous driving company from Hungary whom we spoke about three years ago. They have news: they are working with **Sony** to advance their **automated driving software** stack with Sony's VISION-S Prototype. According to Sony, this prototype - which they just announced at **CES** - is conceived to integrate the advanced ADAS and automated driving systems with the latest **electric vehicle technology** and an innovative **entertainment system**. Security and fun plaited together! In the meantime, they have grown to almost 200 people and they look forward to solving new challenges. **Read More**

A Singapore-Based Coffee Bar Builds 'Ella': A Fully Autonomous Robot Barista Using Intel Technology

Apparently, you can get coffee served by a **fully autonomous robot** barista named "Ella". This happens in downtown Singapore, at a small coffee bar named Crown Coffee. Ella is a **six-axis robot** that takes your order remotely via an app, makes your coffee, notifies you when it is ready, serves it to you, finalizes your bill and charges your credit card. We don't know yet if Ella can already spill coffee on your jacket too! Crown Coffee's CEO/ founder, Keith Tan (whose wife is not surprisingly called Ella), did not do this alone: Intel helped him realize this idea of a robot barista and partnered with him to design and develop prototypes. **Read More**



HireVue drops facial monitoring amid AI algorithm audit

HireVue is a leading company in the "**hiretech**" space. That means software enabling companies to record videos of candidates and then sorting them based on those responses. Which is sort of touchy ground. They announced in January that they are scrapping one of the touchiest features of their software: their algorithms will stop using a candidate's facial expressions in video interviews as a factor to consider. In any case, the author of this great **Fortune** article was told by the company's Chief DS that nonverbal data didn't provide anyway much predictive power compared to the content of a candidate's answers. **Read More**

Artificial Intelligence of COVID-19 Imaging: A Hammer in Search of a Nail

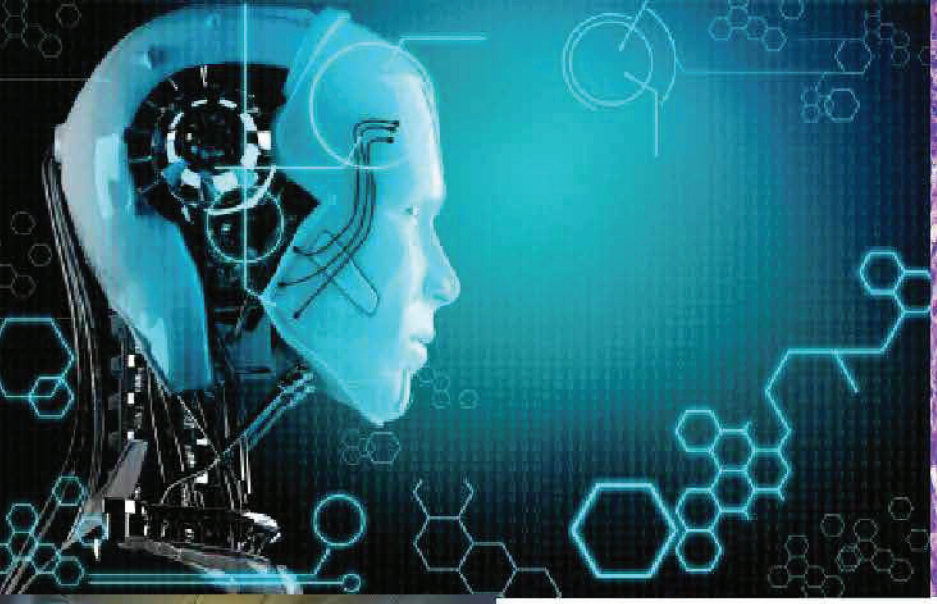
Ronald Summers of NIH, an old friend of our magazine, has found that the pace of publication of AI articles on **COVID-19** is increasing: there are more than **500 such manuscripts on arXiv** and more than **200 articles on PubMed**. The articles are appearing not only in radiology clinical journals but also in technical and general interest scientific journals. There is a clear appetite for such research despite its repetitiveness and unclear path to clinical utility. He so decided to finally put this deluge of articles into context and tell us what is apparently working. Many will be surprised... **Read More**

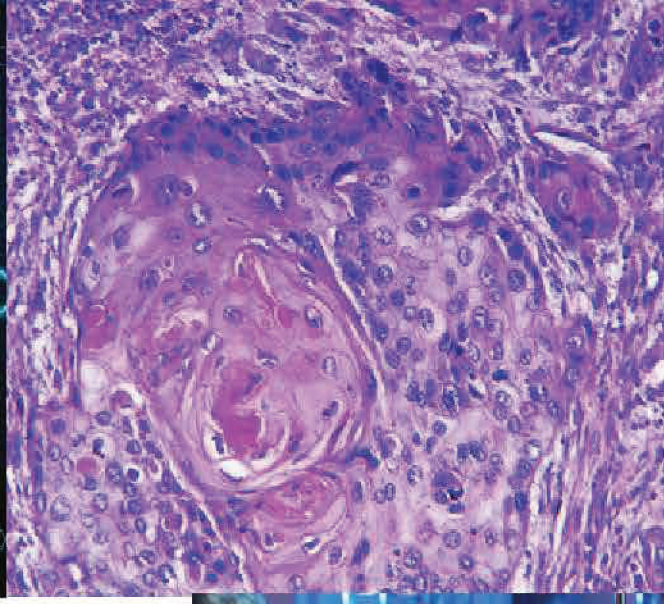
Professor Antonio Torralba Elected 2021 AAAI Fellow

Great news about another old friend of our magazine! Antonio Torralba, a professor at MIT in the Computer Science and Artificial Intelligence Laboratory, has been selected as a **2021 Fellow** by the Association for the Advancement of Artificial Intelligence (**AAAI**). **AAAI Fellows** are selected "in recognition of their significant and extended contributions to the field (contributions which typically span a decade or more), including technical results, publications, patent awards, and contributions to group efforts". MIT is justly proud of Antonio, who is at MIT since 2000, and we are proud too! ¡Felicidades Antonio! **Read More**

How Computer Vision Helps to Influence Customer Behavior

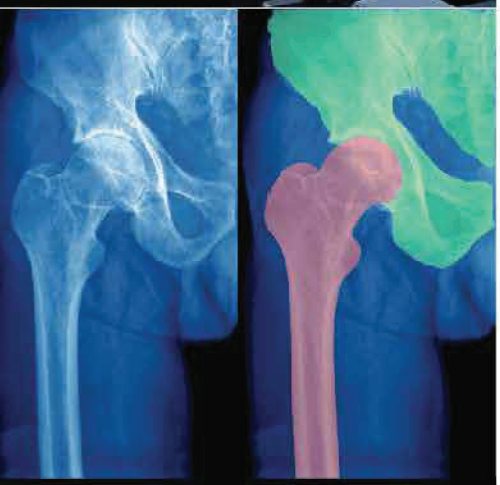
AI and **Computer Vision** help retailers to analyze visitors' behavior, store personnel's performance, checkout lines, availability of goods on shelves, their popularity, and more. Among the tasks that AI solves is the management of purchasing behavior using **heat maps**. The technology makes it possible to analyze the behavior of employees and customers as well. More importantly, heat maps help retailers to influence buyers' behavior: placing ads, visitor messages, and promotional items in "hot zones". They suggest placing unpopular goods in the hot zones or (this is very brave!) relocating popular products to heat up the cold zones. **Read More**





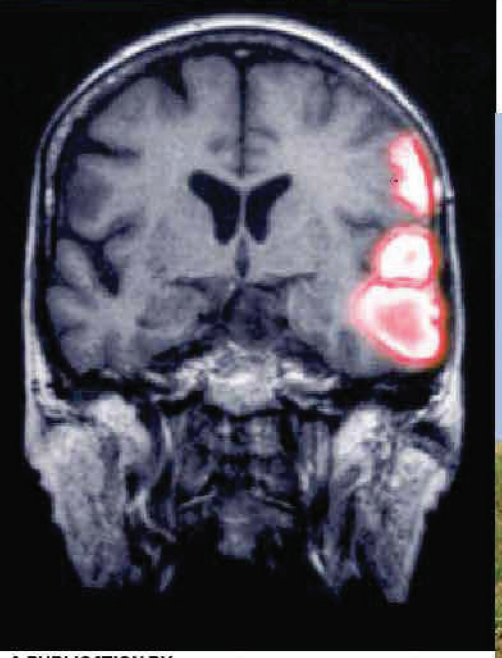
IMPROVE YOUR VISION WITH

Computer Vision News



SUBSCRIBE
CLICK HERE, IT'S FREE

The magazine of the algorithm community



A PUBLICATION BY

RSIP VISION

Global Leader in Computer