September 2021

Computer Vision News

The Magazine of the Algorithm Community

Awesome Research Papers

Computer Vision Projects

AI Spotlight News

Medical Imaging

MICCAI 2021

Computer Vision Tool

Upcoming AI Events

Congrats, Doctor!

Women in Science

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,

MICCAI 2021 is just around the corner, and once again, the International Conference on **Medical Image Computing and Computer Assisted Intervention** has gone virtual! We won't be meeting in Strasbourg (France) as planned, but the organizers have developed an alternative conference program to retain so much of the MICCAI that we all know and love. The presentations, workshops, tutorials, and challenges are all still there. You will not be disappointed!

We have asked MICCAI 2021 General Chair **Caroline Essert** to pinpoint the main highlights of what we expect to be a great Medical Imaging event. Find her exclusive interview in this magazine on page 18.

We recommend that you register with MICCAI 2021 now in order to **get the full experience** alongside the whole community. But if you can't join us, please do subscribe to the **MICCAI Daily magazine**, to receive news, updates and previews from the conference every day, directly to your mailbox. **Feel at MICCAI (almost) as if you were at MICCAI!**

Don't miss the many other articles that we have put together for you this month. Enjoy reading this exciting September issue of Computer Vision News **and bring us along for your next Deep Learning project!**

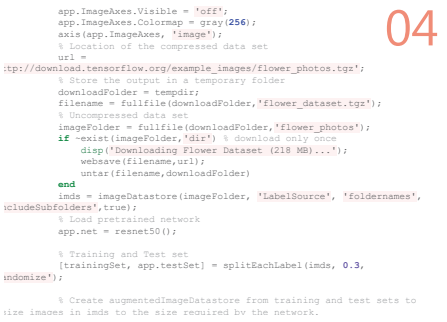
Enjoy the reading and subscribe for free!

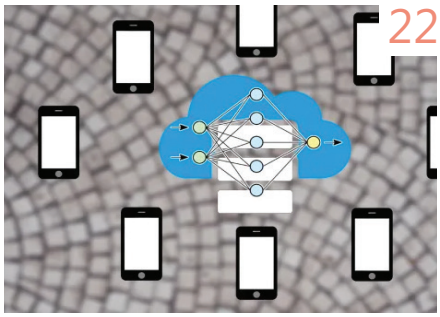
Ralph Anzarouth
Editor, **Computer Vision News**
Marketing Manager**, RSIP Vision**

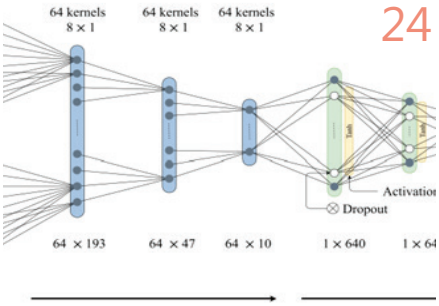
**"Thank you for all of the hard work you do in putting the newsletter together.
I hope you will continue to do this for years to come!
The community really benefits from it."**

Rahul Sukthankar
General Chair CVPR2021
Distinguished Scientist & Sr Director at Google Research

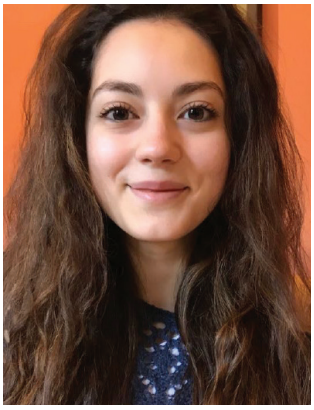
Summary 3







App Designer for MatLab

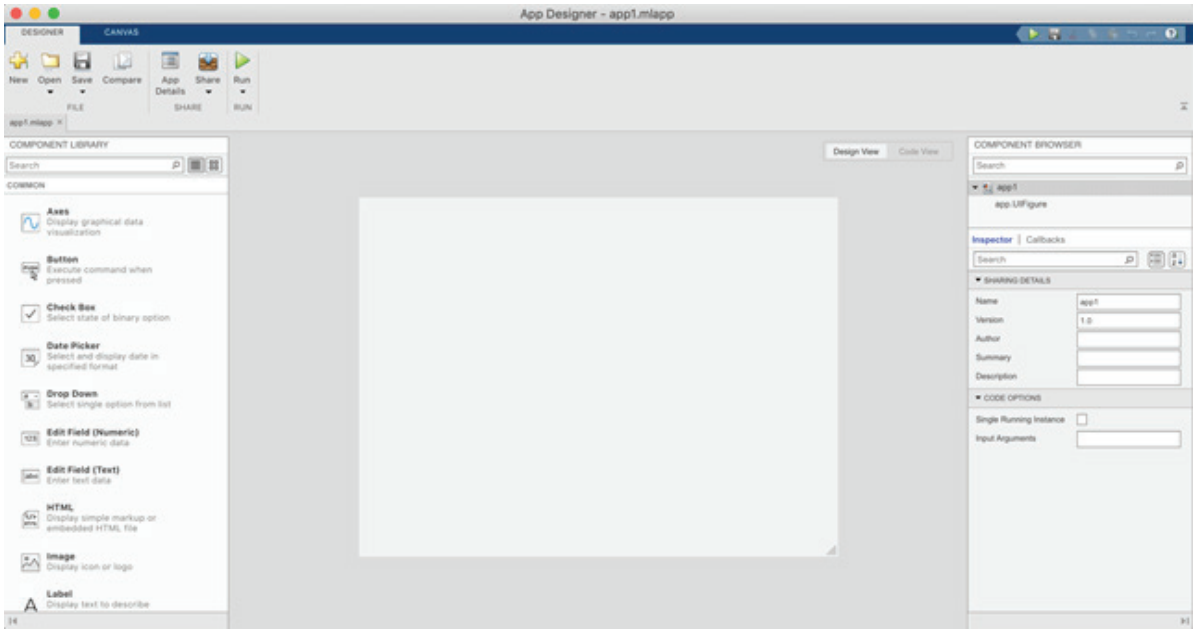


by Marica Muffoletto

Hi everybody, here we are again, slowly getting back to our routines after the summer.

This month I will keep on talking about GUIs and how to create a small app from a project of yours. While last time we looked at how to do that using **Streamlit for Python** (check out the article here!), today we will follow up with an introduction to **App Designer**, a **MATLAB** interactive development environment which will give you the opportunity to **easily translate your matlab code into an app**. For the readers who have been familiar with MATLAB GUI environments in the last years, App Designer is basically the new alternative of the outdated GUIDE environment. It provides a fully integrated version of the MATLAB Editor and a large set of interactive UI components. *"It also offers a grid layout manager to organize your user interface, and automatic reflow options to make your app detect and respond to changes in screen size. It lets you distribute apps by packaging them into installer files directly from the App Designer toolstrip, or by creating a standalone desktop or web app (requires MATLAB Compiler™)."*

There is a good documentation for App Designer online, but for some specific questions and tips, you might find the wonderful MathWorks community and its forums much more useful.



Let's now start looking at what App Designer has to offer to its users. To open it, you can click on the Design App button in the Apps tab or just type "app designer" on the command window.

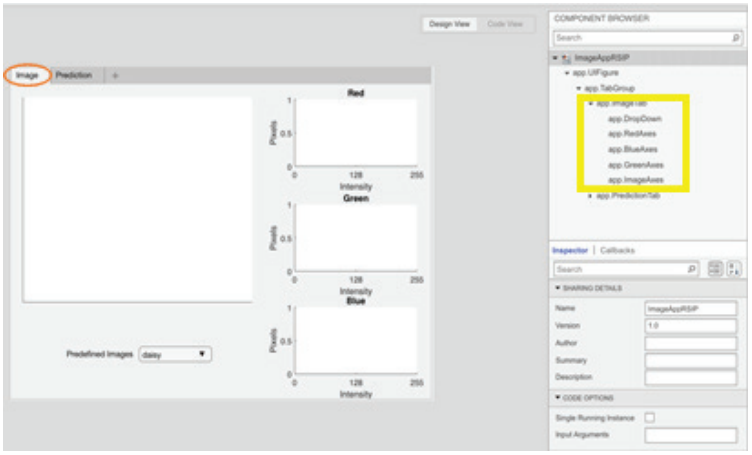
App Designer offers two main components, a Design View where you will be able to lay out the visual components of your GUI; and a Code View which regulates the programming app behaviour. In the figure above, we now see the Design View modality which allows to move over different components and drag them in your favourite location within the blank space at the centre (our new GUI project). If you press on Code View instead, you will be able to design your GUI using blocks of object-oriented code. Usual practice is to keep switching between the two modes to combine the visual and code behaviours of your GUI.
On the left, you can find a menu with common and extra components such as buttons, check boxes, trees, and drop-down lists, or controls such as gauges, lamps, knobs, and switches that let you replicate the look and actions of instrumentation panels. You can also use container components, such as tabs, panels, and grid layouts to organize your user interface.
On the right, you can modify and change positions, shape, and style of all your components trough the Inspector tab or add component callbacks, which regulate custom mouse and keyboard interactions that execute when a user interacts with your app.

Once you are happy with your design, you can decide to package your MATLAB app into a single file that can be easily shared with other users using **MATLAB Desktop** and **MATLAB Online**. When you package an app, MATLAB creates a single app installation file (.mlappinstall). The installation file enables you and others to install your app and access it with a single-click from the app's gallery. This can be shared on MATLAB Online and MATLAB Drive with other users who can collaborate to the design. Most importantly, App Designer also offers the opportunity to create standalone applications using MATLAB Compiler (you need to install this) and Simulink Compiler, or package them as interactive web apps. End-users can run the web apps directly from their browser without installing any additional software.

Let's now have a look at a practical example!

I decided to design a simple app with two windows: one where we visualise an image from an online dataset containing 5 different labels, and the other one where the

same image is predicted using a multiclass linear SVM trained on features extracted by a pre-trained **Convolutional Neural Network (CNN)**. Hence, the final design should include both the ground truth label image and, in a different window, the prediction generated by the SVM classifier and an overview of the extracted training features.



First, in the Design View, you can drag the desired components on the empty space in the middle. I chose to fill it with a Tab group, with two tabs called Image and Prediction. On the Image tab, I added a main Axes component, three lateral ones which display respectively red, green, and blue channels' histograms and a drop- down menu from which you can select one of the five labels in the dataset. Every time a value is selected from the menu, a different image from the testing dataset will randomly be chosen. At this stage, from the Code Browser (in Code View) on the left you need to add two callbacks: the startupFcn which will run as soon as you open the app; and one linked to the drop-down menu (a DropDownValueChanged callback).

Now, if you switch to the Code View, you should find some sections that are already filled in for you - app constructor and components function - and some others that you need to fill in. In the startupFcn, you can add everything that you want to run at startup time, such as download of the images, loading of the pretrained network, preparation of training dataset for the prediction and so on.

```
    % Code that executes after component creation
        function startupFcn(app)
            % Configure image axes
            app.ImageAxes.Visible = 'off';
            app.ImageAxes.Colormap = gray(256);
            axis(app.ImageAxes, 'image');
            % Location of the compressed data set
            url =
'http://download.tensorflow.org/example_images/flower_photos.tgz';
            % Store the output in a temporary folder
            downloadFolder = tempdir;
            filename = fullfile(downloadFolder,'flower_dataset.tgz');
            % Uncompressed data set
            imageFolder = fullfile(downloadFolder,'flower_photos');
            if ~exist(imageFolder,'dir') % download only once
                disp('Downloading Flower Dataset (218 MB)...');
                websave(filename,url);
                untar(filename,downloadFolder)
```

```
            end
            imds = imageDatastore(imageFolder, 'LabelSource', 'foldernames',
'IncludeSubfolders',true);
            % Load pretrained network
            app.net = resnet50();

            % Training and Test set
            [trainingSet, app.testSet] = splitEachLabel(imds, 0.3,
'randomize');

            % Create augmentedImageDatastore from training and test sets to
resize images in imds to the size required by the network.
            app.imageSize = app.net.Layers(1).InputSize;
            augmentedTrainingSet = augmentedImageDatastore(app.imageSize,
trainingSet, 'ColorPreprocessing', 'gray2rgb');
            augmentedTestSet = augmentedImageDatastore(app.imageSize,
app.testSet, 'ColorPreprocessing', 'gray2rgb');

            % Get the network weights for the second convolutional layer
            w1 = app.net.Layers(2).Weights;

            % Scale and resize the weights for visualization
            w1 = mat2gray(w1);
            w1 = imresize(w1,5);

            montage(w1,'Parent',app.FeaturesAxes);
            featureLayer = 'fc1000';

            % Update the image and histograms
            updateimage(app,'daisy');

            trainingFeatures = activations(app.net, augmentedTrainingSet,
featureLayer, 'MiniBatchSize', 32, 'OutputAs', 'columns');

            % Get training labels from the trainingSet
             trainingLabels = trainingSet.Labels;

             % Train multiclass SVM classifier using a fast linear solver, and
  set 'ObservationsIn' to 'columns' to match the arrangement used for training
  features.
             app.classifier = fitcecoc(trainingFeatures, trainingLabels, ...
                 'Learners', 'Linear', 'Coding', 'onevsall', 'ObservationsIn',
  'columns');

        end
```

You also need to add an extra function which takes as input a label name and displays the main image and the three channels' histograms.

```
    methods (Access = private)

        function updateimage(app,imagefile)
            % Find an instance of an image for each category
            labelIm = find(app.testSet.Labels == imagefile);
            % generate one random index
            randIdc = randperm(length(labelIm),1);
            % initialize R to be the four numbers of A
            R = labelIm(randIdc);
            try
                app.testImage = readimage(app.testSet,R);
            catch ME
                % If problem reading image, display error message
                uialert(app.UIFigure, ME.message, 'Image Error');
                return;
            end

            % Create histograms based on number of color channels
            switch size(app.testImage,3)
                case 1
                    % Display the grayscale image
                    imagesc(app.ImageAxes,app.testImage);

                    % Plot all histograms with the same data for grayscale
                    histr = histogram(app.RedAxes, app.testImage,
'FaceColor',[1 0 0],'EdgeColor', 'none');
                    histg = histogram(app.GreenAxes, app.testImage,
'FaceColor',[0 1 0],'EdgeColor', 'none');
                    histb = histogram(app.BlueAxes, app.testImage,
'FaceColor',[0 0 1],'EdgeColor', 'none');

                case 3
                    % Display the truecolor image
                    imagesc(app.ImageAxes,app.testImage);

                    % Plot the histograms
                    histr = histogram(app.RedAxes, app.testImage(:,:,1),
'FaceColor', [1 0 0], 'EdgeColor', 'none');
                    histg = histogram(app.GreenAxes, app.testImage(:,:,2),
'FaceColor', [0 1 0], 'EdgeColor', 'none');
                    histb = histogram(app.BlueAxes, app.testImage(:,:,3),
'FaceColor', [0 0 1], 'EdgeColor', 'none');

                otherwise
                    % Error when image is not grayscale or truecolor
                    uialert(app.UIFigure, 'Image must be grayscale or
truecolor.', 'Image Error');
                    return;
            end
                % Get largest bin count
                maxr = max(histr.BinCounts);
                maxg = max(histg.BinCounts);
                maxb = max(histb.BinCounts);
                maxcount = max([maxr maxg maxb]);

                % Set y axes limits based on largest bin count
                app.RedAxes.YLim = [0 maxcount];
                app.RedAxes.YTick = round([0 maxcount/2 maxcount], 2,
'significant');
```

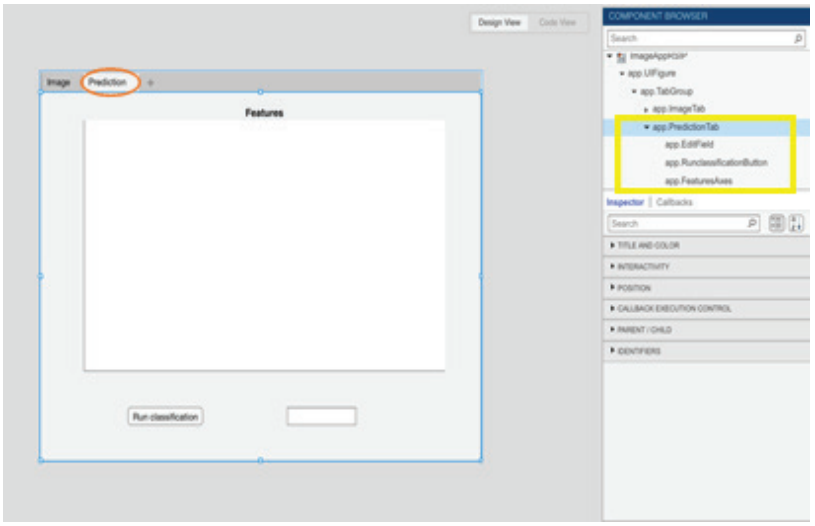
```
                app.GreenAxes.YLim = [0 maxcount];
                app.GreenAxes.YTick = round([0 maxcount/2 maxcount], 2,
'significant');
                app.BlueAxes.YLim = [0 maxcount];
                app.BlueAxes.YTick = round([0 maxcount/2 maxcount], 2,
'significant');

        end
    end
```

Finally, the DropDownValueChanged callback code should be filled with a call to the updateimage function with the selected value from the drop-down menu.

```
% Value changed function: DropDown
        function DropDownValueChanged(app, event)
            value = app.DropDown.Value;
            % Update the image and histograms
            updateimage(app,string(value));
    end
```

On the Prediction tab, I added an Axes component showing the weights for the second convolutional layer and a button which I named Run Classification. If this is pushed, the Edit Field component will be filled in with the prediction from the SVM classifier.



Below, the code linked to the button callback.

```
% Button pushed function: RunclassificationButton
        function RunclassificationButtonPushed(app, event)

            % Create augmentedImageDatastore to automatically resize the
image when image features are extracted using activations.
            ds = augmentedImageDatastore(app.imageSize, app.testImage,
'ColorPreprocessing', 'gray2rgb');

            % Extract image features using the CNN
            featureLayer = 'fc1000';
            app.imageFeatures = activations(app.net, ds, featureLayer,
'OutputAs', 'columns');

            % Make a prediction using the classifier
            app.predictedLabel = predict(app.classifier, app.imageFeatures,
'ObservationsIn', 'columns');
            app.EditField.Value = string(app.predictedLabel);
        end
```

A final note to remember when using App Designer: all variables which need to be passed between callbacks and functions must be listed as properties and accessed through the app handle.

Now, have a glance at the result of our app design below.

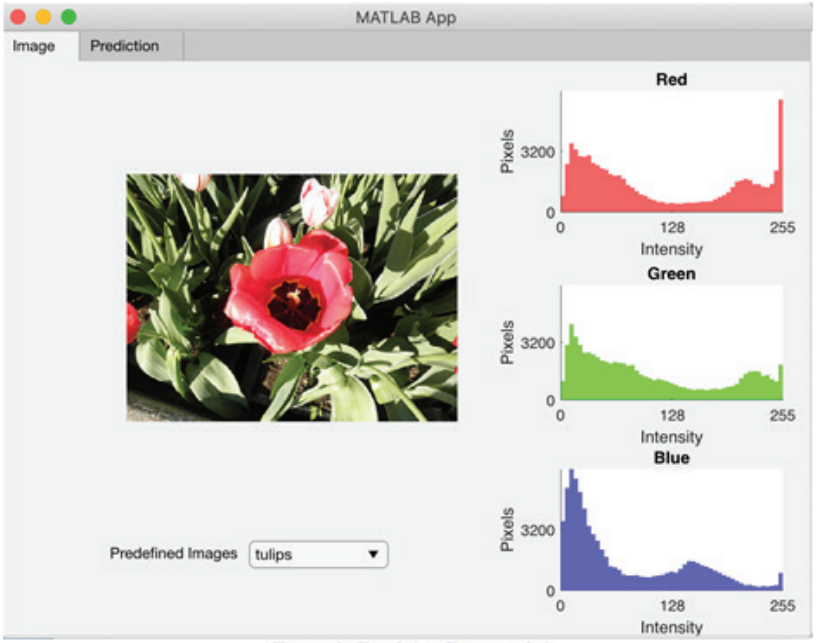


Figure 1: Final App Design tab 1

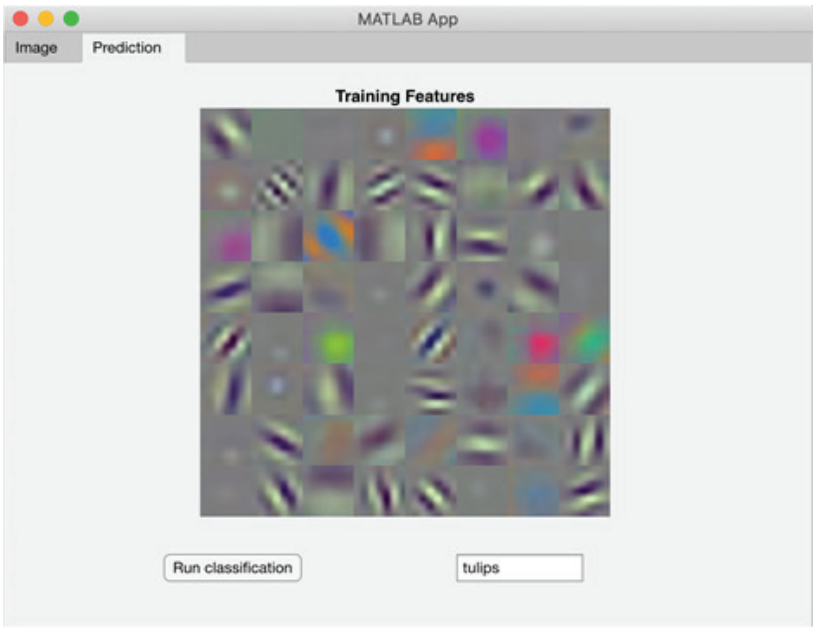


Figure 2: Final App Design tab 2

This article will hopefully guide your first steps into the world of App Designer and give you some ideas on how to transform your MATLAB code into a little standalone project.
Keep following our next articles to discover other interesting computer vision topics & tools!

MICCAI 2021

Feel at MICCAI as if you were at MICCAI
September 28-29-30

Keep in contact with the community!

Click here to subscribe for free to MICCAI Daily 2021!

Intraoperative AI-based Registration for Orthopedic Surgeries

Total joint arthroplasty

Total joint arthroplasty is a common orthopedic procedure that involves the replacement of a damaged joint with an artificial prothesis. Total joint arthroplasty is one of the most common types of procedures in orthopedics, and within this category, hip and knee joints are the most frequently replaced. For instance, according to the **Agency for Healthcare Research and Quality**, more than **450,000 total hip replacements** are performed annually in the United States alone. Other, less frequently replaced joints, include ankles, wrists, elbows, and shoulders.

Surgical planning and implementation

As both the number of orthopedic procedures performed and the quality of their outcome continue to rise, surgeons are becoming more cognizant of the importance of patient-specific surgical planning and functional component positioning during surgery. *"These days, we see more and more patients with active lifestyles undergoing orthopedic surgeries. Accordingly, the need for* ***accurate and successful orthopedic surgeries*** *is increasing. These patients have years and years of active life ahead of them and so every millimeter in the*

final positioning counts", says MD **Shai Factor**, an orthopedic surgeon at **Tel Aviv Sourasky Medical Center in Israel**. *"Achieving such a high level of accuracy is not easy, but it is vital. Today, we mainly rely on our personal experience and 'eyeballing' to get that done*".

Recently, orthopedic surgeons have started utilizing new types of emerging technologies with **outcome accuracy in the range of sub-millimeters** with the intent of accommodating younger and more active patients and extending the longevity of the orthopedic implants. Intraoperative registration aims to provide the surgeon with real-time indications by matching a preoperative CT (computerized tomography) to an intraoperative X-ray fluoroscopy. Using automated landmark detection and measurements of anatomical relationships, the surgeon can progress in the procedure while the registration system identifies and updates any changes from the original surgical plan.

AI-based registration

Registration is a process that involves transforming multiple imaging datasets into a single coordinate system. Intraoperative registration aims to bridge the gap between the imaged surgical plan and real-life in the operating room. The registration process demands overcoming multiple challenges:

- **Field of view (FOV)**- the space captured in the imaging process

varies between modalities, and sometimes even within a single modality. Registration aims to match the FOVs captured in the images by cropping one of the images (or both).

- **Resolution**- a resolution discrepancy usually exists between the per-operative CT image and the intraoperative X-ray fluoroscopy image. Registration aims to reach a pixel-to-pixel transformation by matching the resolutions using either down sampling or up sampling of the images.

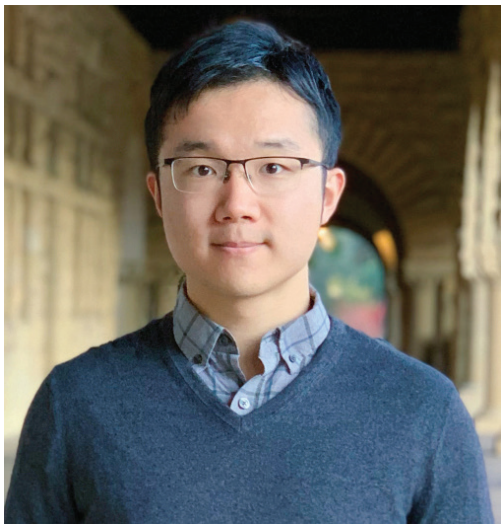
- **Dimensionality**- the preoperative CT is a three-dimensional, high-resolution image, while the intraoperative fluoroscopy is two-dimensional with a lower resolution. Registration aims to match dimensionalities by reducing the three-dimensional image to a two-dimensional image.

- **Orientation change**- the orientation of the different bones and tissues captured changes between the preoperative and the intraoperative imaging. Registration aims to match the orientation to the preoperative plan using an approximation of the movement.

RSIP Vision keeps innovating in the field of visual intelligence for medical devices: it has developed a **patient-specific, intraoperative registration module for orthopedic surgery**. This novel AI-based module performs 2D to 3D registration that improves the clinical outcome of orthopedic surgeries and can easily be implemented in procedures such as hip replacement. The module provides accurate, quantitative measurements of bones and implants during the surgery, it supports the optimization of implant positioning, and improves patient outcomes.

Using an **AI-based approach to image registration**, RSIP Vision's module utilizes **a proprietary neural network** that enables accurate, automated, and patient-specific assessment of the real location and positioning of the implants and bones. *"This technology will assist and guide us, the surgeons, throughout the procedure, providing the essential feedback regarding anatomical measurements and implant positioning"*, says MD Shai Factor. *"I believe it will **improve the accuracy of our surgeries** and lead to better clinical outcome"*. This innovative module joins RSIP Vision's wide suite of **orthopedic image analysis tools for preoperative planning and intraoperative navigation during orthopedic surgeries**.

16 Congrats, Doctor!



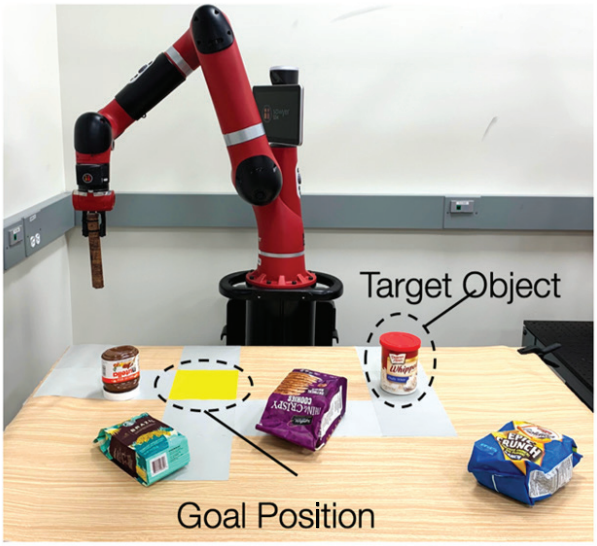
Kuan Fang recently completed his Ph.D. at Stanford University. His research interest lies at the intersection of robotics, computer vision, and machine learning, with the goal of enabling robots to learn effective physical interactions in the real world through visual reasoning. He received his B.S. degree from Tsinghua University and has spent time at Microsoft Research Asia, Google [x] Robotics, and Google Brain. He will start as a postdoctoral researcher at UC Berkeley in October 2021. Congrats, Doctor Kuan!

Kuan would like to thank his advisors Fei-Fei Li and Silvio Savarese for their support and guidance during these years. This work wouldn't have been possible without them.

Building **robotic systems that can autonomously solve challenging tasks in the physical world** is a long-standing goal in artificial intelligence. Despite the encouraging progress in robotics in the past decades, most successes occur in task-specific systems that are programmed to follow fixed routines in well-controlled environments. To enable robots to perform challenging tasks in unstructured environments, we investigate methods for **learning generalizable perception and control from rich interactions**. In particular, we focus on three essential themes: (i) **Utilization of rich interactions**: How do we design structured models and learning algorithms to endow robots with generalizable perception and control? (ii) **Collection of rich interactions**: How do we provide sufficient and suitable interaction data to scale up learning. (iii) **Close the loop of utilization and collection**: How do we jointly adapt the algorithms and the data without extensive human intervention?

Learning Tool Use from Simulated Self-Supervision.

We develop a method that learns to use novel objects as tools through **simulated self-supervision**. A neural network work policy is designed to jointly predict the task-oriented grasps and the manipulation actions given visual observations of the object. To improve the generalization capability of the policy, we collect large-scale data through self-supervision using procedurally generated tool objects in the simulation. We demonstrate that the learned policy can successfully perform tool use tasks with a variety of novel objects in both simulation and the real world.

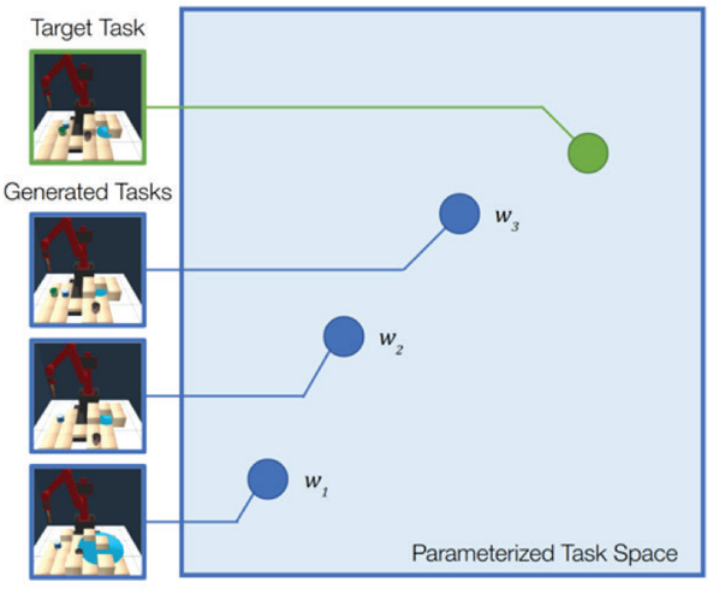


Learning Sequential Tasks from Task-Agnostic Data.

We present a method that learns to solve sequential tasks by **learning from task-agnostic interaction data**. To effectively find plausible plans, we propose to hierarchically generate subgoals and actions using learned latent representations. We propose cascaded variational inference to learn the latent representations fromtask-agnostic interactions between the robot and the environment. We demonstrate that **our method successfully solves sequential manipulation tasks under various semantic and physical constraints in cluttered environments**.

Generating Tasks for Curriculum Learning.

We present a framework to progressively generate tasks as the curricula to **facilitate reinforcement learning in hard-exploration problems**. At the heart of our approach, a task generator learns to create tasks from a parameterized task space via a black-box procedural generation module. To enable curriculum learning in the absence of a direct indicator of learning progress, we propose to train the task generator by balancing the robot's performance in the generated tasks and the similarity to the target tasks. Through adversarial training, the task similarity is adaptively estimated by a task discriminator defined on the robot's experiences, allowing the generated tasks to approximate target tasks of unknown parameterization or outside of the predefined task space.



Caroline Essert is Associate Professor of Computer Science at the University of Strasbourg – CNRS. She is also General Chair for MICCAI 2021, the 24th International Conference on Medical Image Computing and Computer Assisted Intervention, being held later this month (starting on September 27). Caroline speaks to us ahead of the big event.

Caroline, with MICCAI 2021 just around the corner, what can we all expect from this year's conference?
Obviously, the big news is that due to uncertainty around the pandemic, we have had to cancel the physical event in Strasbourg, which is of course disappointing. However, we have worked very hard to modify the program and update everything so that we can offer an amazing virtual experience for participants.

Tell us more – why should people register?
There are so many reasons! The first one is of course the scientific content. We've had so many excellent submissions this year. I just saw the program and it's very exciting. The second thing is that we have some exceptional keynote speakers. This year we have four keynotes instead of three, allowing us to achieve a perfect gender balance, which I am really pleased about. The topics have been chosen very carefully.


Caroline Essert

That sounds very intriguing! Can you give us a sneak preview?
Pierre Jannin will be speaking about sustainable research, which is a subject that's important to me. Strasbourg is a green city, and we are very concerned about the environment. We are keen cyclists and have the first bicycle network in France. Pierre's speech will be so interesting, especially with deep learning these days consuming so much energy.

Pierre Jannin

Who else do you have on the agenda?
Alyson J. McGregor will be delivering a keynote that will enlighten us all about a vital topic which is very close to my heart – how gender bias can affect science and medicine research. Fei-Fei Li, the researcher who famously built ImageNet, will be giving a talk about ambient intelligence in medicine, which will offer an insight into some of the medicine of the future. Last but not least, we have Richard M. Satava, who will be speaking about technology transfer for non-invasive surgery.

Fei-Fei Li

I note that Richard will be speaking on the day of the inaugural CLINICCAI event. Can you tell us more about that?
The idea of the clinical day (CLINICCAI) is to bring clinicians to our conference to share their translational research experiences and hopefully become active within the community. My hope is that this event will be renewed in the years to come and will be the first of a successful series.

Is this similar to the clinical day at CARS?
Yes, it's in the same spirit. It has been organized by a team of clinicians at the IHU Strasbourg, who are our partners as well. It will bring an interesting perspective to our attendees and give the clinicians who attend the full conference a fascinating overview of the breadth of our activities.

I'm sure you have convinced most of our readers to sign up now. Is there anything else you can reveal to win over those final few?
We have many other events, such as those traditionally organized by the MICCAI Student Board. Students are usually very excited about that – but not only students! Faculty members and other participants enjoy attending their events, including the morning yoga sessions, and the academia and industry event. They will also be holding two webinars, including one on burnout.

And for those of us who are no longer students?
You can still attend! There is no limit on numbers, so they will be open to everyone. We also have the MICCAI Startup Village with the pitch competition back for another year. Ihope many startups will participate in this event and deliver an exciting speech to present their company.

Alongside this, we will be featuring one startup every day in MICCAI Daily.
That's fantastic news! Thank you. What

a great collaboration!

Is that everything?
Not quite! We have a few surprises that I can't tell you about right now, but which will give you a flavour of Strasbourg and France, including the local cuisine and the sights. I think everyone will enjoy those.

What can you tell us about the technical side of the conference?
Like last year, we are using the Pathable platform, but in addition to that we will have a second platform called SpatialChat, which we will use for the interactive sessions, such as the poster sessions, networking sessions, and industrial booth. It offers a very smooth and engaging experience. You can move your avatar to go and discuss something with a group or visit a poster, for example. That will be really nice.

How are you supporting participation this year?
We had a couple of different calls this year – one was for participation in the main conference and another for the satellite events. We had many applications from all over the world and are so happy that the MICCAI Society can offer registrations to 50 students for each, which will be great for our younger colleagues who are just starting out.

Finally, the well-established Women in MICCAI network is a highlight every year, and we hear you have a new network this year. What is that?
Yes, we have a new network this year called the RISE-MICCAI network. The RISE network aims to bring together people from countries currently underrepresented in the community. This is so important. We want as many people as possible to join the network to open up new and urgently needed collaboration around the world.

If you would like to attend MICCAI 2021, early bird registration is open until 3 September, so sign up today!

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!**

How Can AI Help Fight Climate Change?

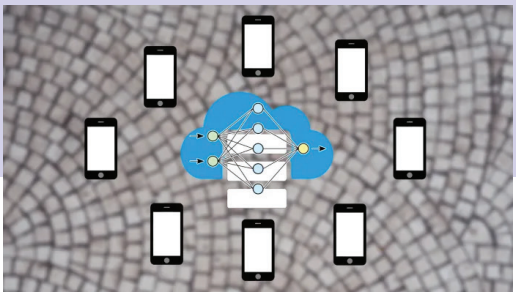
Good question, right? If you are not yet worried by **climate change**, we can point you at a few experts who really are. Apparently, AI can help predict and limit **devastation from storms**, **wildfires and droughts**: it is already used to send alerts in Japan, in China, in the Amazon Forest. AI applications could also help design more energy-efficient buildings, improve power storage and optimize renewable energy deployment by feeding solar and wind power into the electricity grid as needed. They could also help households minimize their energy use. The downside is how much energy AI consumes on the way. Read More

M.Video uses artificial intelligence to offer customers personalized pricing

As you have read many times on Computer Vision News, the retail is one of the most active fields in AI adoption. Today's news is that the M.Video retail chain (Russia's leading e-commerce and consumer electronics retailer), has launched personalized pricing powered by AI for users of its mobile platform. The new service offers customers guaranteed discounts based on purchase history and available bonuses. Does it work? Apparently, the conversion rate for the customer mobile app has risen 1.5x. Personalized pricing for M.Video website customers will be launched at a later time. Read More

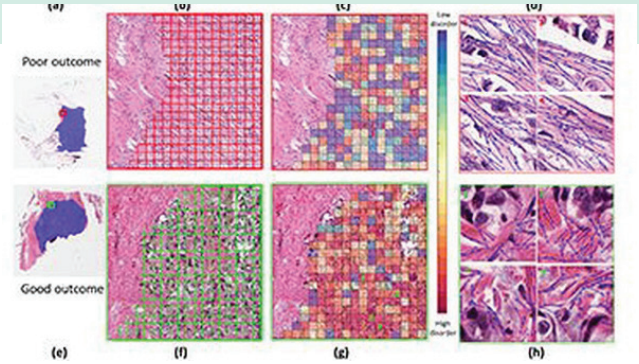
What is federated learning?

Some of our readers certainly already know. They can skip to the next news. For all the others, this is a solution to the key challenge of machine learning for **large amounts of data**, the processing risks of which organizations would rather avoid. Do you run your model on the cloud? Then you need to send the data to the server where the ML model resides. Advances in **edge AI** have made it possible to avoid sending sensitive user data to application servers, **making the model fit on smartphones and other user devices**. Do you want us to write a full feature about federated learning? Tell us! Read More



AI Aids in Discovery of New Prognostic Biomarkers for Breast Cancer

Scientists at **Case Western Reserve University** have used AI to identify new biomarkers for **breast cancer** that can predict whether the cancer will return after treatment - and which can be identified from routinely acquired tissue biopsy samples of early-stage breast cancer. We suppose this is not an exact prediction, however a very educated guess can be precious to determine correct diagnosis and treatment. Without entering into a discussion about correlation and causality, apparently the chances for **an aggressive tumor and a likely recurrence** are written in the arrangement of **collagen** in breast tissue. **Read More**



Designing Machine Learning Systems by Chip Huyen

As our readers know, **Chip Huyen** is not only an awesome friend of our magazine, a very successful ML engineer and a popular ML lecturer at Stanford: she has also published a few traveling bestsellers in Vietnam. If you still don't know her, read about her career in her own words. This is to tell you that US publisher O'Reilly is going to publish her English writing in this amazing guide: **Designing Machine Learning Systems**, a framework for designing real-world ML systems that are quick to deploy, reliable, scalable, and iterative. **The first 3 chapters are already available here!**

GFP-GAN: Towards Real-World Blind Face Restoration with Generative Facial Prior

Alessandro Ferrari kindly beckons we have a look at this: he says that **Tencent AI** just opened the source code of **GFP-GAN, the most advanced photo restoration system ever!** Blind face restoration usually relies on facial priors, such as facial geometry prior or reference prior, to restore realistic and faithful details. However, very low-quality inputs cannot offer accurate geometric prior while high-quality references are inaccessible, limiting the applicability in real-world scenarios. GFP-GAN could jointly restore facial details and enhance colors with just a single forward pass. **Project Page**

A novel deep learning method for predictive modelling of microbiome data

IOANNIS VALASAKIS, KING'S COLLEGE LONDON
@WIZOFE

The topic for this month is a little bit different than usual; it's taking a deep dive into microbiome data and how it is used for drug response predictions and disease diagnosis. Microbiome (and its research using bioinformatics techniques) is an important scientific area, which is really hot at the moment; both in the scientific sense but also in the entrepreneurship world, where many companies are investing in the field. Our microbiome is a great source of lot of information about us; how we eat, we digest and even how we feel. Let's dive into that! We are indebted to the authors of this paper (Ye Wang, Tathagata Bhattacharya, Yuchao Jiang, Xiao Qin, Yue Wang, Yunlong Liu, Andrew J Saykin, Li Chen) for allowing us to use their images. The paper was published in Briefings in Bioinformatics, 22(3), 2021, 1-14.

"The way to succeed is to double your failure rate."
- **Thomas J. Watson**, pioneer in the development computing equipment for IBM

Hi everyone! For people who took a summer break I hope that you enjoyed it; and for everyone else, keep going strong!

Let's begin with an inspirational quote that may shift your mood. Failure shouldn't always be seen as an obstacle, but - when that's possible - as a learning event to fight for adversities and create a new plan. Failure in experimentation or in science should be as well a motivator to try new ideas, explore what went wrong and better understand the problems. By looking closer into the problems, one may find more solutions ☺.

Microbiome

An exact definition isn't easy. The fact that in the terminology, "microbiota" (human-associated microbia) and "microbiome" (catalogue of microbes and their genes) are often used interchangeably, makes it even more complex.

The human microbiota consists of the **10-100 trillion symbiotic microbial** cells which can be found in each human's gut (mostly); the **human microbiome** consists of all those genes living inside the human cells. That's the reason many different projects have been started to understand the role and its impact on human health (in a primary level) but also use them as a guide of what's happening inside a human body.

The drastic reduction in sequencing costs experienced over the past few years has made it possible to identify specific microbial taxa found within the human gut that are difficult or impossible to culture. Researchers are now able to generate millions of sequences per sample to assess differences in microbial communities between body sites and individuals.

You can read more and a bit of historical info, in the great article "Defining the Human Microbiome" by Luke K Ursell et al in doi: 10.1111/j.1753-4887.2012.00493.x

Role

Previous research on the microbiota has shown to be useful in the following distinct ways:

1. Aids the human digestive system as it can be very efficient in extracting energy from food and harvesting nutrients, as microbiota provides humans with enzymes and biochemical pathways produced by versatile metabolic microbial genes that are far more than found in human genome.
2. Protects the system against invasive pathogens by providing a physical barrier, producing antimicrobial substances, or involving in competitive exclusion.
3. It is essential in the induction, training, and function of the host immune system.

Modern regression techniques have been applied to solve the problem of microbiome prediction using the high dimensional features. They usually incorporate a **sparse penalty** they tend to be computationally expensive and require a significant parameter tuning. More importantly, they are limited in exploring the phylogenetic relationship among taxa. To explain, with some diseases (e.g., inflammatory bowel disease) the overall composition of the microbiota is significant rather than granular marker taxa. The opposite is true for diseases like

colorectal cancer, which creates a need to develop an optimized, robust model across varied cluster size and density.

Deep learning approaches such as CNN provide a more granular correlation of features but, as discussed in earlier works, this doesn't apply to a small training dataset. In this work, **a novel architecture** is proposed, which attempts to be computational efficient and utilize the phylogenetic tree to predict both continuous and binary outcome, as can be seen in Fig 1.

Methodology

The network is effectively a CNN which regularizes the phylogeny of the microbiome prediction. The taxa are clustered based on a phylogeny-induced correlation structure. To achieve high accuracy, convolutional layers are designed to include the phylogenetic correlation across different phylogenetic depths as much as possible. With a clever use of convolutional layers, it achieves both the dimensionality reduction but also the encoding of the phylogenetic information, leading to **the recognition of spatially local input patterns**.

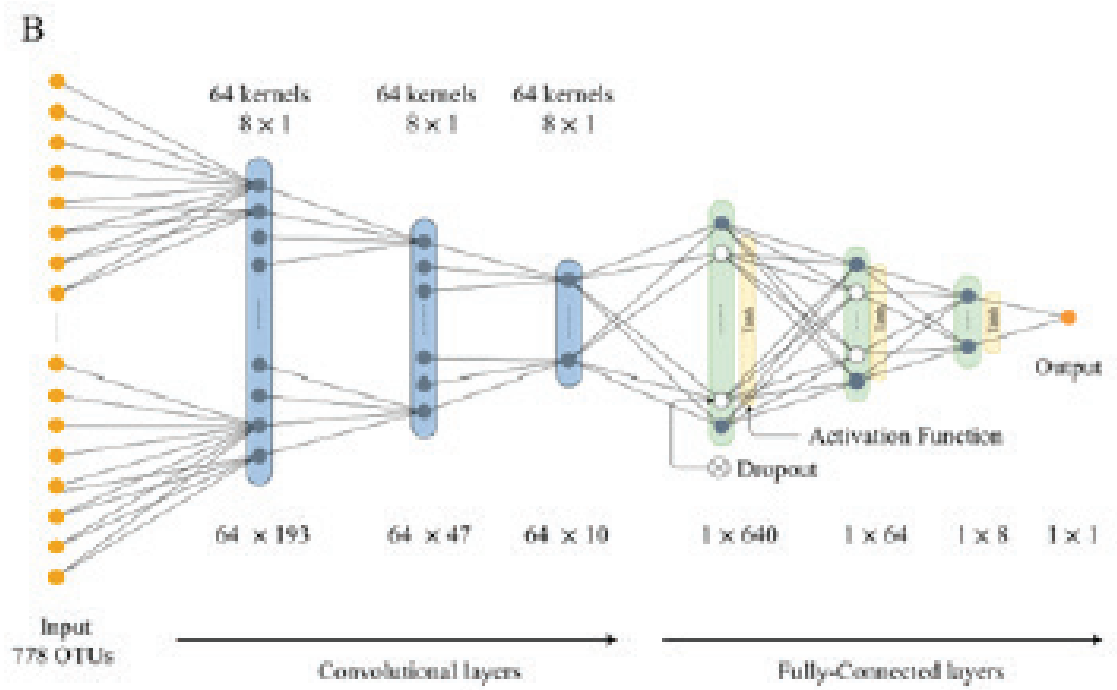


Figure 1. The architecture of the proposed MDeep. You can see the input, convolutional layers, fully connected and the output of the network, together with the dimensions of each. Notice, that in the last fully connected layer, a single node for continuous outcome and two nodes for binary outcome are shown.

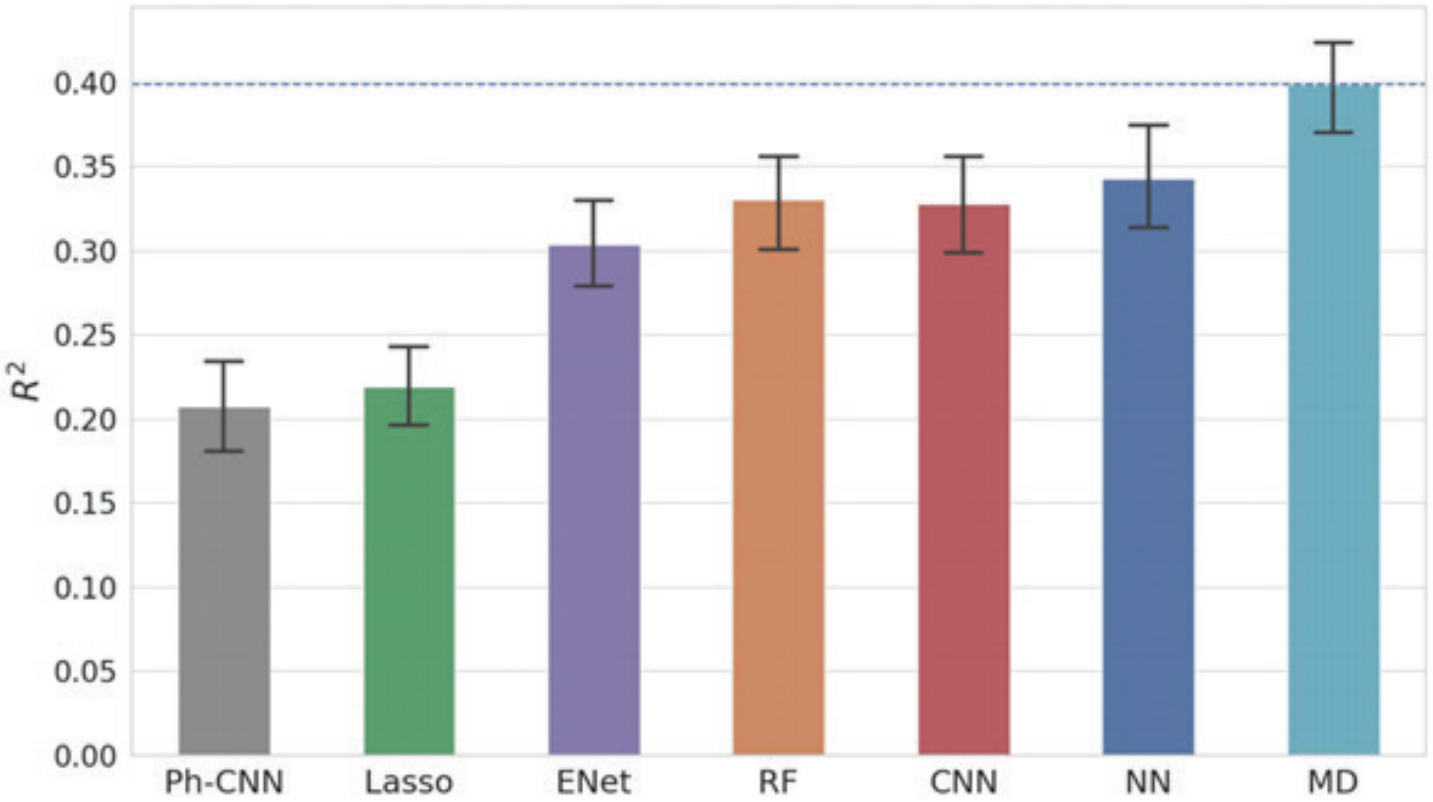


Figure 2. The comparison of the R2 metric with different architectures (proposed). It's visible that the currently proposed architecture performs considerably better than already successful ones such as the CNN.

The operation of multiple convolution filters allows local smoothing and feature maps between the adjacent convolutional layers are restricted by the local connectivity. This is equal to the concept of nearby taxa on the lower phylogenetic/taxonomic level being more likely to be closer to each other on the higher phylogenetic/taxonomic level.

The output layer contains a single neuron for continuous outcome and two nodes for the binary output.

Moreover, a few words about the simulation strategy. You can see that in Fig 3.

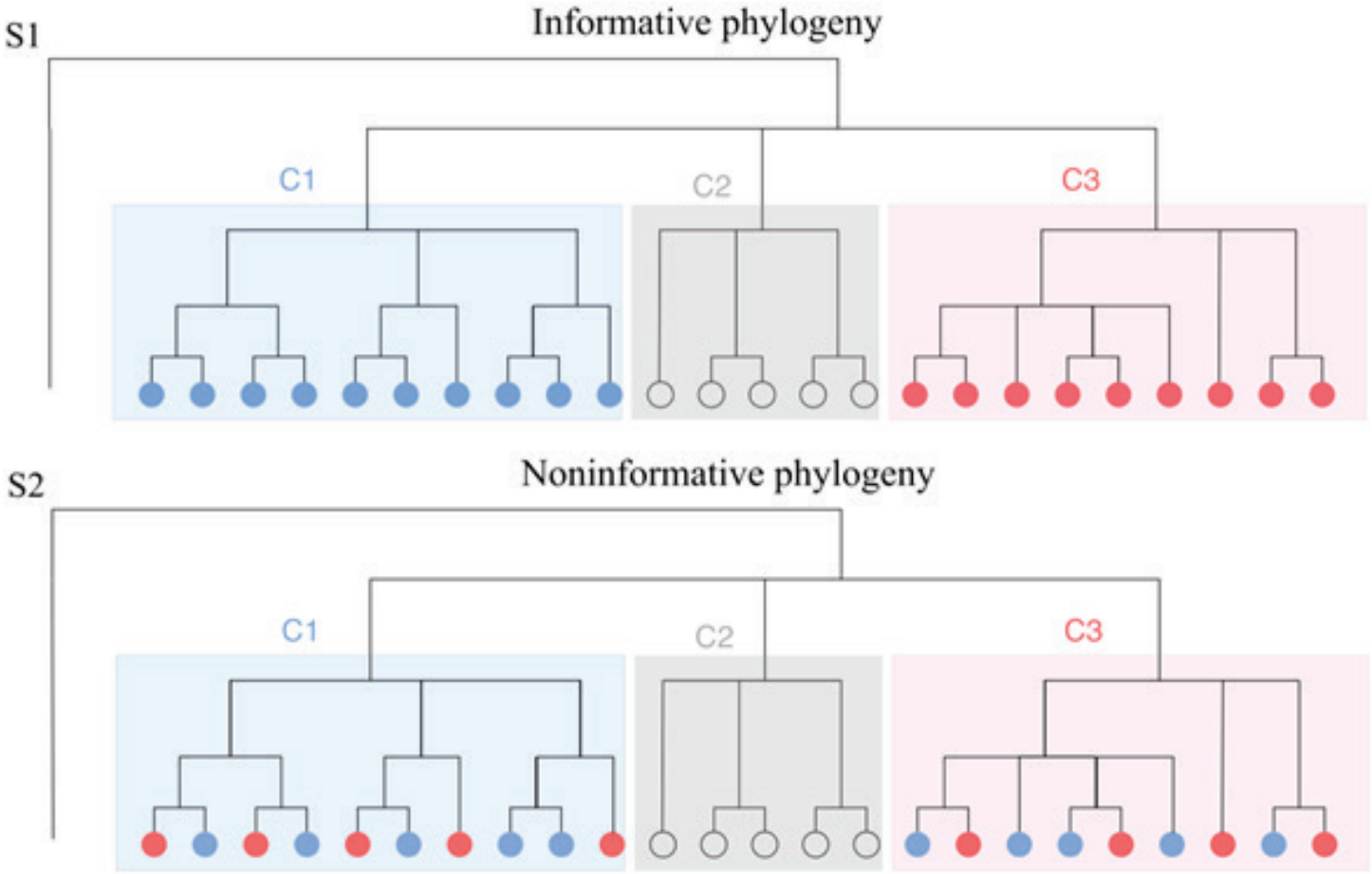


Figure 3. The simulation strategy for the experiment. S1 signifies the informative phylogeny with all OTUs in C1/C3 having the same size effect in the same direction of the outcome. S2 is the non-informative phylogeny with the adjacent OTUs having a different outcome in C1 vs C3. The red circles represent aOUTs with positive effects to the outcome while the blue circles represent aOTUs with negative effects to the outcome.

The prediction performance of the algorithm for population in the United States can be seen in the Table 1 below. You will notice on the table that the 'BC' represent 'Baby versus Child' while 'CA' is 'Child versus Adult'.

As always, please read the article, support the scientists and try to ask more questions ☺.

	Type	Lasso	Enet	RF	Ph-CNN	CNN	NN	MDeep
Sensitivity	BC	0.9856	0.9784	0.9856	**0.9936**	0.9904	0.9920	0.9896
	CA	0.7531	0.7808	0.7669	0.7108	0.8123	0.7977	**0.8254**
Specificity	BC	0.8018	0.8636	0.8327	0.7545	0.8764	0.8691	**0.8891**
	CA	0.7736	0.8112	0.7712	0.6896	**0.8176**	0.7760	0.8160
Accuracy	BC	0.9294	0.9433	0.9389	0.9206	0.9556	0.9544	**0.9589**
	CA	0.7631	0.7957	0.769	0.7004	0.8149	0.7871	**0.8208**
Precision	BC	0.9216	0.9438	0.9329	0.9048	0.9493	0.9469	**0.9544**
	CA	0.7826	0.8165	0.7803	0.7152	0.8276	0.7917	**0.8286**
MCC	BC	0.8334	0.8674	0.8568	0.8155	0.896	0.8933	**0.9041**
	CA	0.5317	0.5957	0.5405	0.4488	0.634	0.5774	**0.6459**
F1 score	BC	0.9516	0.9601	0.9577	0.9463	0.969	0.9684	**0.9712**
	CA	0.7634	0.7956	0.7717	0.6932	0.8172	0.7924	**0.8242**

Table 1. The table shows the representation for different metrics on the US population. Read the paper for more detail on the method of those results.

Wrapping up!

I hope you enjoyed again this month's article. Even though the topic was a little bit different than usual, being able to get inspiration and example from other subjects is more than worthy. Sometimes, similar problems have been solved in other disciplines in various ways, and it would be unworthwhile to ignore this!

Specifically, the use of the **combined inference** could be of great use for models using **generative and discriminative approaches**. With a combined/joint distribution in a **computer vision setting**, usually a complex, high-dimensional function, simplifying assumptions need to be made to arrive to solutions that are explainable in their duration.

Examples of other discriminative networks on which this approach can be taken: linear regression, support vector machines (SVMs) and boosting techniques. This is applicable in case as of traffic sign recognition (with a significant use case of modern/"smart-cars"), driver monitoring, parking management and similar applications in other industries where visual recognition and pattern recognition is an important element.

Thank you Ye for this very interesting model. Research like this can bring forward our understanding of microbiome and our bodies relations and understanding.

Be mindful about the people around you; let's think and take care of others ☺.

"You need to be tenacious, to be ambitious!"

Dorsa Sadigh is an Assistant Professor in Robotics at Stanford University in the Computer Science and Electrical Engineering Department.

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

Dorsa, can you tell us about what you do with robots?
Sure, I work in the sub-area of human robot interaction. We are in general very interested in designing learning and controlling algorithms for robots that can safely and reliably interact with people and work with people.

You have a full group dedicated to this project.
Yes, we have the ILIAD Group. It's the Stanford Intelligent and Interactive Autonomous Systems Group. We have a good number of graduate students and undergrads working on interesting problems in this area.

What are the most interesting problems that you are tackling?
One of the most interesting directions that we are looking at is the study of repeated interactions, and how new norms or new conventions emerge when agents interact with each other over a period of time. In general, when people think about robotics, oftentimes

they don't think about interactions with other agents. If you think about building an autonomous car: for the longest time building an autonomous car was about the functionality of the vehicle perception. It wasn't as much about how this car was going to drive next to other drivers.

That is understandable. At the beginning, you want to think about the functionality of the system. Whereas at some point, people realize that they can't forget about the interaction. The interaction problem is one of the key challenges of having autonomous cars. People started worrying about the interaction between autonomous cars and human driven cars. I'd argue that a lot of interaction problems these days think about one step or few steps short term interactions. We need to think about long term, repeated interactions and what sort of behaviors emerge when we have, let's say, an autonomous car interacting with other drivers over long periods of time, or a robot interacting in peoples' homes over repeated, long-term interactions.

How did that mistake come about?

It's not a mistake. When you think about repeated interaction problems, you imagine things working. Maybe things are not working yet. Maybe it's a next step type of problem. Like we saw in autonomous driving, we weren't able to make progress until thinking about interactions between autonomous cars and human driven cars. Repeated interactions fall into the same category. If you put a robot in people's road, it guides the system towards behaviors that are undesirable. That can be pretty challenging. That is something that we really need to study. We need to worry about it now.

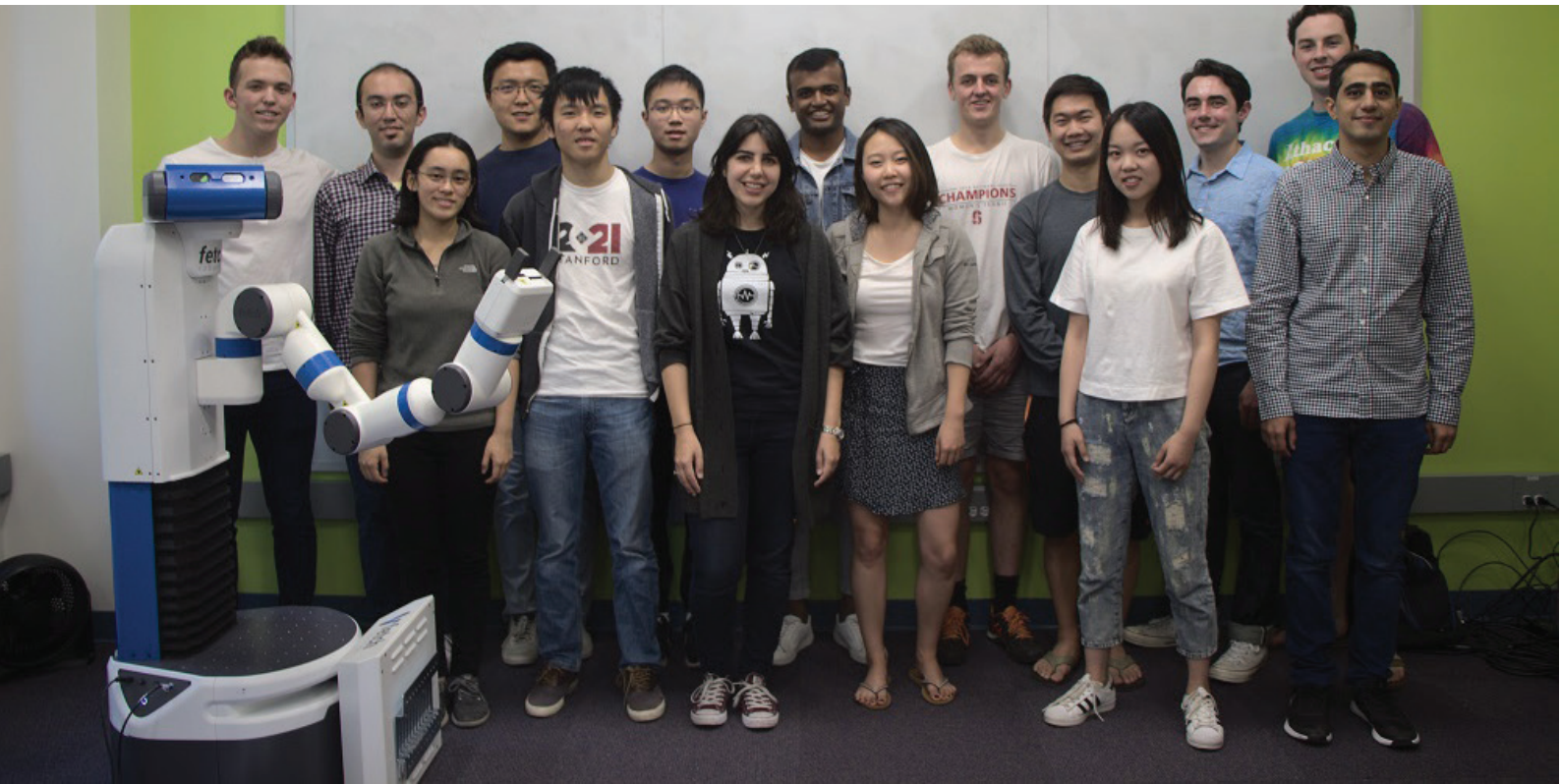
You are mainly working towards applications?

Yes, we are working on a spectrum of things. We are doing some more theoretical things. We are also very interested in real applications. I'm personally very driven by real applications and some of the challenges that exist in real robotic systems today. One problem that we are looking at, which is another favorite that I didn't talk about yet, is the problem of assisted feeding. We're looking at how we can have a robot bring food to a person's mouth, to put it inside of a person's mouth in a way that is safe and interactive and easy.

In which kinds of situations would anyone need that?
In general, the core application for that is in an assisted setting for people with disabilities who cannot feed themselves. They often use robot arms on a wheelchair. This robot arm can help them on various types of tasks, like feeding themselves or putting clothes on. All of these are pretty challenging tasks, and the person still has some control over the robot. But it's really difficult to use a joystick or to use the devices that patients with disabilities have, like sip and puff devices to control these robot arms. You need to have good algorithms that are intuitive to control and easy to operate to help the person.

It seems like you really want to leave your mark on what you are doing.
As I said, I personally get motivated by real problems. This is truly a real problem that we need to address. It has an easy valuation. I don't have robots that can feed people today, so I can see that it doesn't exist. It's really exciting to push in this area.

How do theoretical researchers put in so much effort but never really see the fruits of their labor?
Actually, my husband is in theory. He works in areas that are exactly like that. [*laughs*] They sometimes just prove the existence of things. They don't even have a constructive way of finding it. But it's an interesting puzzle. It's interesting to push on that puzzle even though you might not see the effects of that in an application today. There's an excitement over long term effects. If I solve this puzzle, maybe in 20 years, 40 years, it would actually be used in practice. That's true in a lot of algorithms that you see



today. After many years, you start seeing the effects of it. That is also exciting, but it's hard to predict these things. You got to be really excited about the puzzle to actually make progress!

Who has more chances to win the Nobel prize, you or your husband?
[*laughs*] I'd say my husband! I'm more interested in more practical things and seeing the effects today. I feel like the Nobel Prize is usually for research.

If I'm asked, I will vote for you anyway!
[*laughs*] Thank you!

What do you enjoy the most about working with ILIAD?
They are incredibly creative! That always blows my mind, and that makes talking to them every day really exciting. Our conversations end up leading to directions that at the beginning of the conversation, I wouldn't really think about. The things that come to mind when I think about the group are, first of all, incredible curiosity and creativity, which I really value. Beyond that, another thing that comes to mind is they're very collaborative and nice. On Friday, we had a lab social and played some board games.

I'm very creative and collaborative too. Does that give me a chance to be accepted one day?
[*laughs*] You should join! We'll have fun!

Do you think that it's because of this creativity and collaboration that your group has been able to write so many good papers?
That's a good question. It's a combination of things, to be honest. Of course, creativity needs to be there for ideas to develop. But it's not the only thing. You need to be tenacious! You need to actually be driven, be ambitious! They have all of these qualities; they really are awesome students. All of these things come together. There's a little bit of peer pressure in terms of the number of papers that come out these days, at all the conferences and so on.

"That is what makes me different and unique and creative, when you start walking outside of your boundaries."

Maybe a bit too much?
I do think it is a little too much. If you take six papers, that is what would have been like two papers in the old days. Now they're in smaller pieces.

...Which makes it less painful when they are rejected by a conference or a journal.
That's a very good point! Because maybe you are not as invested in it. You can go back and improve it. It's a faster cycle…

We did not mention that half of your scientific career was spent at another well-respected institution. You must have half your heart at Stanford and the other half at Berkeley.
[*laughs*] That is really true! I really like Berkeley. I did my undergraduate and graduate studies at Berkeley. Go Bears! (For anyone who is out there from Berkeley) It was wonderful being at Berkeley. Stanford also has its perks. I have a different experience at Stanford. Both of them are interesting. When we have big football games, it's hard to decide on a team.

You have 13 pages on your CV, starting in 2009. Can you tell us something about your life before those 13 pages?
For me, it started in middle school. I grew up in Iran. I went to this special middle school which was actually pretty fun. Pretty much anyone in academia from Iran went to the same school.

It's like a gifted children's school?
Kind of, it does have an exam. It's a different school outside of the usual education system. The organization is called the National Organization for Development of Exceptional Talents. I went to this middle school, and it kind of changed my life to be honest. Middle school is this period when you make ideas about what you want to be when you grow up. I met my best friends. We started taking physics, math, chemistry and some of these very advanced topics that are even taught at colleges. We started learning them in middle school. That blew my mind. We started learning about philosophy and religion. That was so interesting. The middle school continues so I went to that same high school. Mid high school, around 2006/2007, we got our green card for the United States through chain migration: our aunt applied for our family when I was 3, and it took 13 years. When I was like 16, we got our green cards. At that time, I remember visiting the United States. I went to Irvine, which is where my mom's side of the family lives. If you've been to the Irvine/Newport Coast area, it's this pretty Hollywood movie type place. [*laughs*] Everything looks so nice! My parents asked if I wanted to take the entrance exam at the colleges in Iran or move to the US. Of course, I want to move to the United States! That was like moving in the movies! [*laughs*]

The choice was easy.
I moved to Irvine. I was there for about a year, year and a half. Then I moved to Berkeley in 2009.

I have already interviewed many female Iranian scientists, all very impressive. What are the Iranians doing to generate so much incredible talent among their female students?
I think about that very good question too! When I was in Iran, I never thought about the gender inequality that we see in the US when we look at STEM. I was thinking that gender inequality was worse in Iran than in Western countries, but never about sciences or jobs. If you look at colleges, they are actually pretty much 50/50 in all STEM fields, which is pretty incredible. There isn't that mentality that engineering is not for girls. Because of that, it's actually popular for Iranian girls to go into engineering.
There are two things that come to my mind. One of them is, in general, Iranian women are pretty strong, and they want to prove that they are strong. There is

kind of a factor of trying to prove that you can do it too, and you're as good as everyone else. No one can tell you how to live your life. There is a bit of trying to prove yourself. The other angle, which is conflicting for me, is that Iranian schools are all girls or all boys. They are gender-segregated, because of religious reasons. I always thought that's a stupid idea. Now, looking back, I actually think that made a difference: in a class of all girls, someone has to be good at math, at physics, at chemistry! You end up getting a full spectrum of people. You never think girls are not good at math. You never build these stereotypes.

What else you'd like to share?

One thing that comes to my mind, an advice that I personally follow myself: every week, try to do one thing that makes you happy, genuinely happy. Try to do one thing that you are afraid of. Try to walk outside of your comfort zone. Find that one thing that pushes you outside of your comfort zone. Find that one thing that makes you happy.

Life is already full of difficulties. If you also go outside of your comfort zone voluntarily, you make it even more difficult.

Yes! But that's where the exciting things happen! That's where you get creative! I try to learn something outside of robotics every week. That is what makes me different and unique and creative, when you start walking outside of your boundaries.

More than 100 inspiring interviews with successful Women in Computer Vision await for you in our archive!

COMPUTER VISION EVENTS

Medical Augmented...
Online
30 Aug -10 Sep

IMVIP 2021
Dublin, Ireland
1-3 Sep

IoT Tech Expo Global
London, UK
6-7 Sep

MedTech Conference
DC - Minneapolis - Virtual
27-30 Sep

MICCAI 2021
Virtual
27 Sep -1 Oct
Meet us there

NLP SUMMIT
Online Event
5-7 Oct

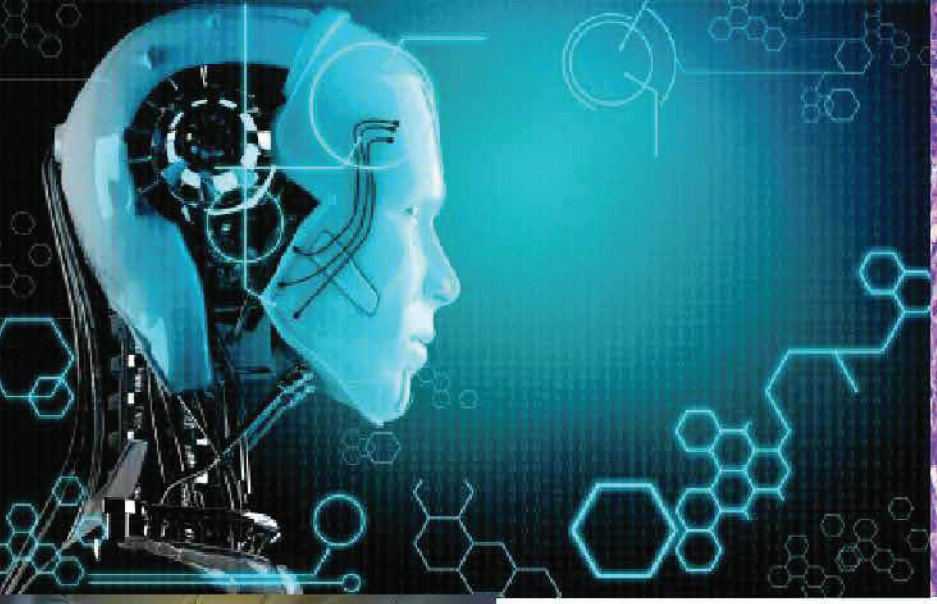
ICCV 2021
Virtual
11-17 Oct
Meet us there

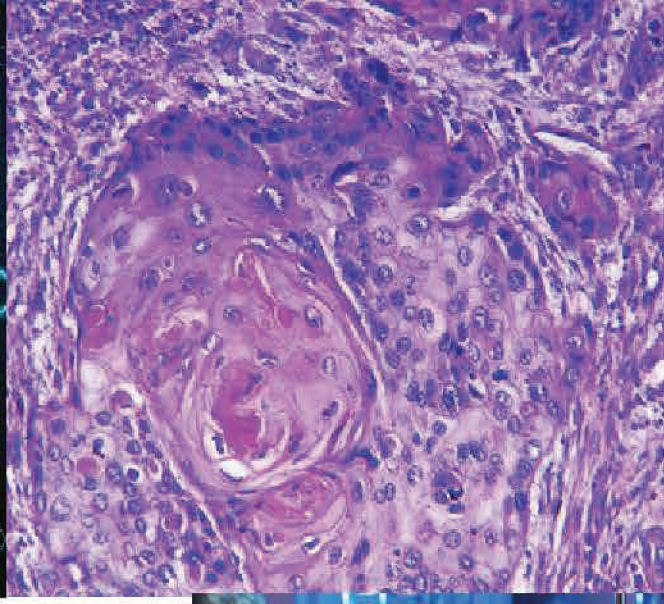
IMVC
Tel Aviv, Israel
26 Oct

SIPAIM 2021
Campinas, Brazil
17-19 Nov

BMVC 2021
Online
22-25 Nov

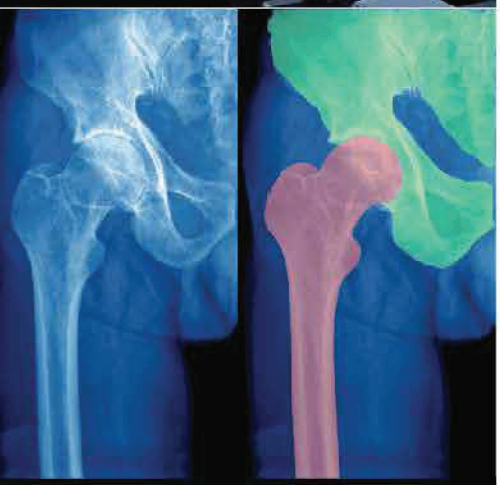
Due to the pandemic situation, most shows are considering to go virtual or to be held at another date. Please check the latest information on their website before making any plans!





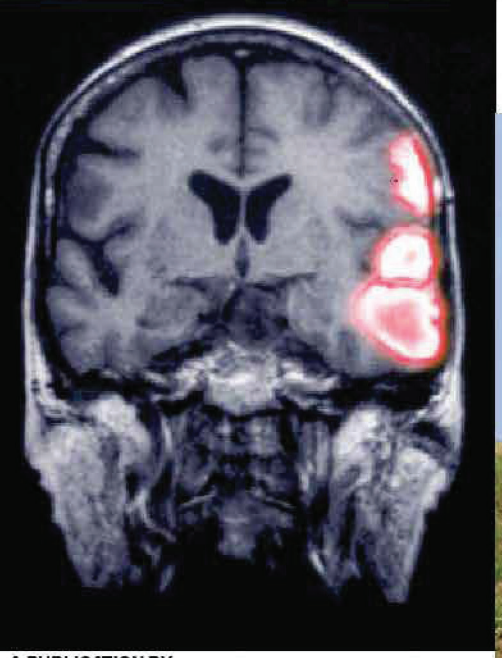
IMPROVE YOUR VISION WITH

Computer Vision News



SUBSCRIBE
CLICK HERE, IT'S FREE

The magazine of the algorithm community



A PUBLICATION BY

RSIP VISION

Global Leader in Computer