August 2021

# Computer Vision News

The Magazine of the Algorithm Community


Awesome Research Papers

InnerEye by Microsoft

AI for Medical Imaging

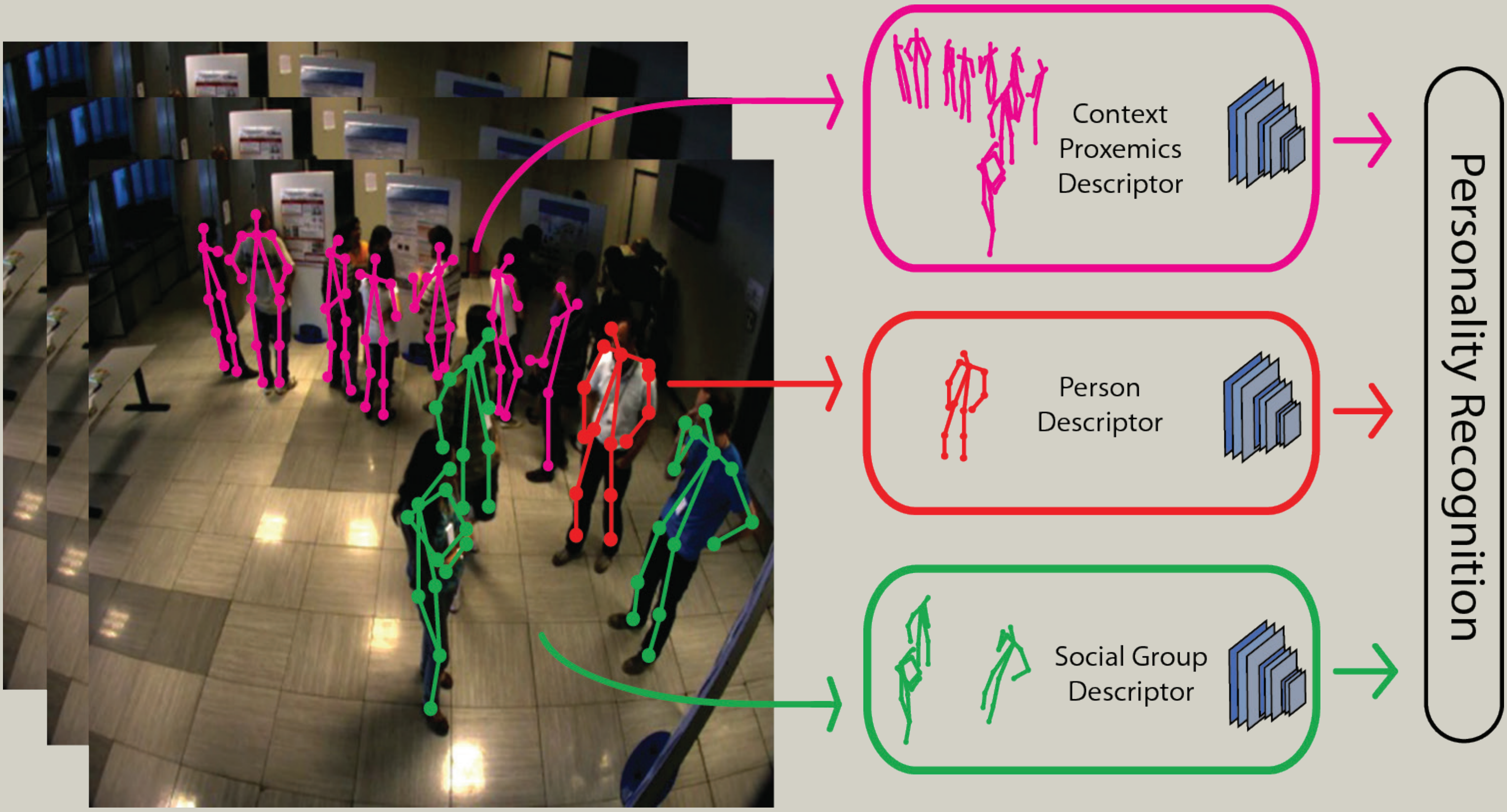
Computer Vision Projects

AI Spotlight News

Women in Science

Upcoming AI Events

Bay Vision Webinars






**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,

The **Medical Imaging with Deep Learning (MIDL)** conference, held virtually again this year, provided our community with another high point last month. In this August issue of **Computer Vision News**, we review the two exciting papers that scooped the **Best Paper and Best Paper Runner-up awards** at the event. There were so many innovative papers this year, so these really are the best of the best. You won't want to miss it!

Our computer vision Application of the Month is **DIDA**, a new image-based historical handwritten digit dataset. Read our fascinating interview with authors Amir and Hüseyin to find out more, including how they dealt with the elaborate **Gothic-style writing** and what separates DIDA from its modern counterparts.

Did you miss our great Bay Vision webinar with the awesome **Lena Maier Hein**? The video with the full talk and slides is available now: [Does Machine Learning Based Biomedical Image Analysis Require Domain Experts?](#) Join the group [here](#) and be the first to learn about the next meetup!

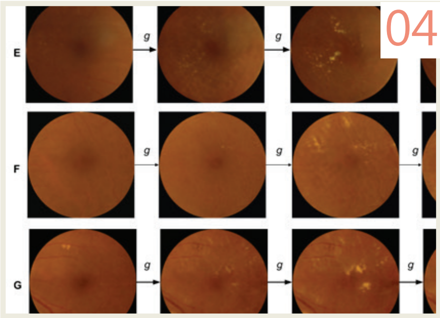
Away from MIDL, this edition of Computer Vision News is packed full of content to keep you enthralled, so get started now with the **AI Research of the Month**, on page 4.

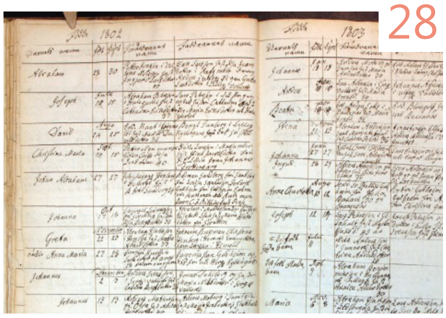
**Enjoy the reading and [subscribe for free](#)!**

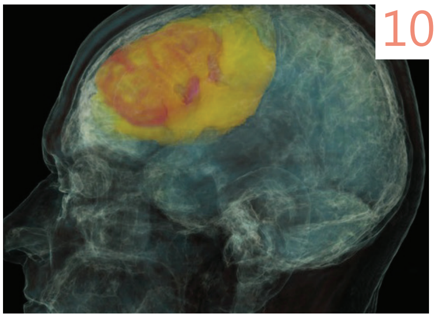
**Ralph Anzarouth**
Editor, **Computer Vision News**
Marketing Manager, **RSIP Vision**

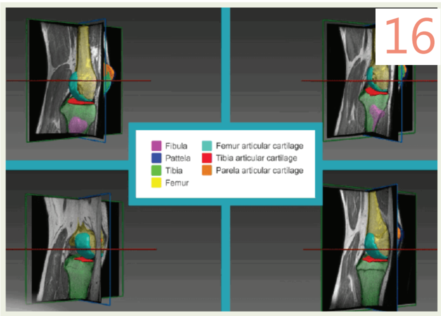
"It's great that you're featuring our work. I really appreciate it. I have been following your magazine for years and I enjoy it very much!"

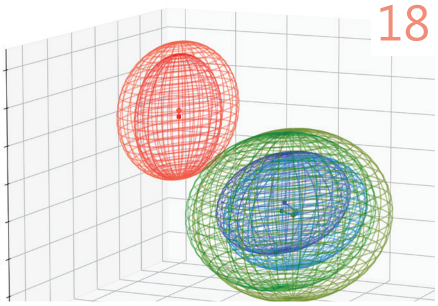
Piotr Dollár
Research Director
Facebook AI Research (FAIR)

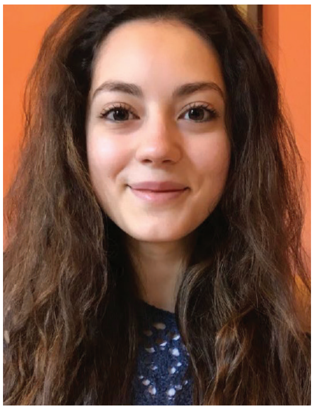





Fibula
Patella
Tibia
Femur
Femur articular cartilage
Tibia articular cartilage
Patella articular cartilage

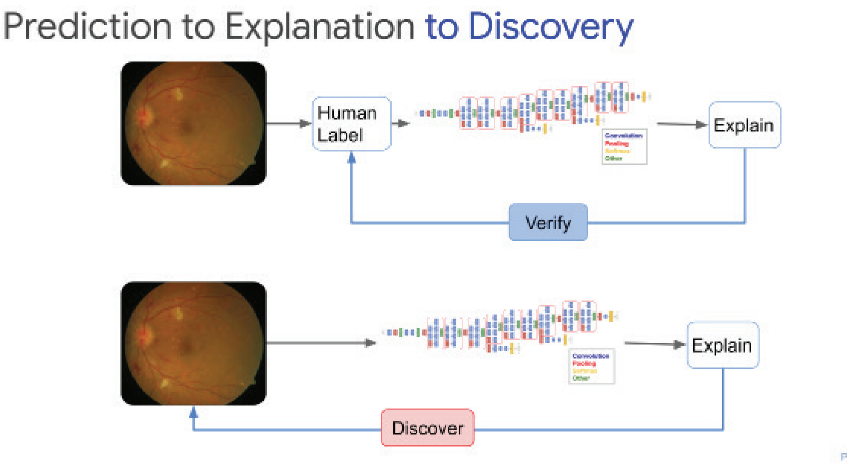


# Scientific Discovery by Generating Counterfactuals Using Image Translation



by Marica Muffoletto

Every month, Computer Vision News selects a research paper to review. This month's article is dedicated to **Scientific Discovery by Generating Counterfactuals Using Image Translation**, written by several authors from the Google Health research centre and in collaboration with the Rajavithi Hospital. This work was presented at MICCAI 2020 and it is based on preliminary work which appeared in Nature Communications at the start of last year! We are thankful to all the authors (**Arunachalam Narayanaswamy, Subhashini Venugopalan, Dale R. Webster, Lily Peng, Greg Corrado, Paisan Ruamviboonsuk, Pinal Bavishi, Rory Sayres, Abigail Huang, Siva Balasubramanian, Michael Brenner, Philip Nelson, Avinash V. Varadarajan**) for allowing us to use their images to illustrate this review and especially to Subha for providing us with extra illustrations and links. Their paper can be found here.




The scope of this paper, just like for many other papers these days, is to increase our knowledge of what causes the predictions of a deep learning model.

As visual recognition models are more and more employed, especially in medical imaging (where the tasks are more complex and require specific knowledge), the establishment of excellent explanation techniques is becoming crucial. This work aims to go beyond explanation techniques to provide means for scientific discovery. How to do that? By finding out **what** about the data leads to the predictive power of a classification model, rather than **where** in the image the model focuses to build its predictions.

This project is then based on creating a pipeline for explainability & discovery which is applied on **Diabetic Macular Edema (DME)** prediction models. The gold standard for the diagnosis of this disease is a 3D scan of the retina based on **Optical Coherence Tomography (OCT**) while a cheapest but less accurate option is the most used Color Fundus Photos (CFP), a 2D image of the retina. The screen is done by looking for the presence of lipid deposits called hard exudates (HE), a proxy feature for retinal thickening.
The dataset used is made of 7072 paired CFP and OCT images labeled by 2 medical doctors. Note that only the combination of CFP images and labels by CFP/OCT are employed for training.

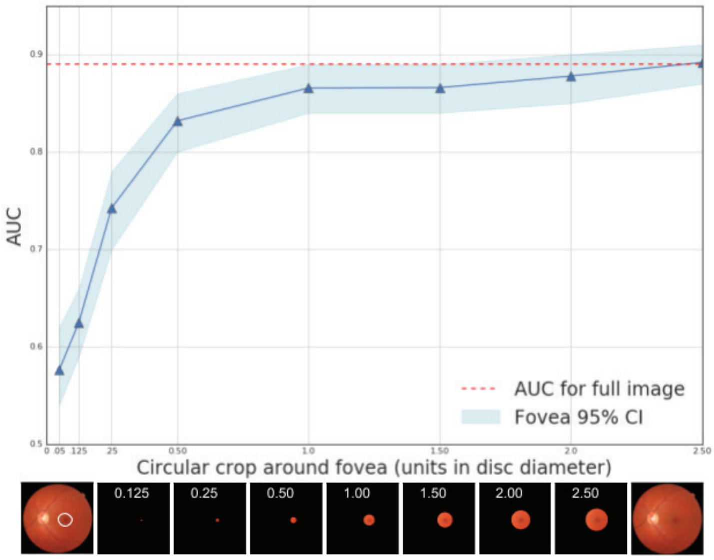
While some previous studies found that the same CNN architecture which uses labels derived from human experts grading OCT significantly outperforms HE labels based on CFP images, they do not explain why it is so. The authors fill in this knowledge gap by building a method based on generation of counterfactuals. *The outcome is a scientific discovery which answers the question of why OCT labels are more accurate.*

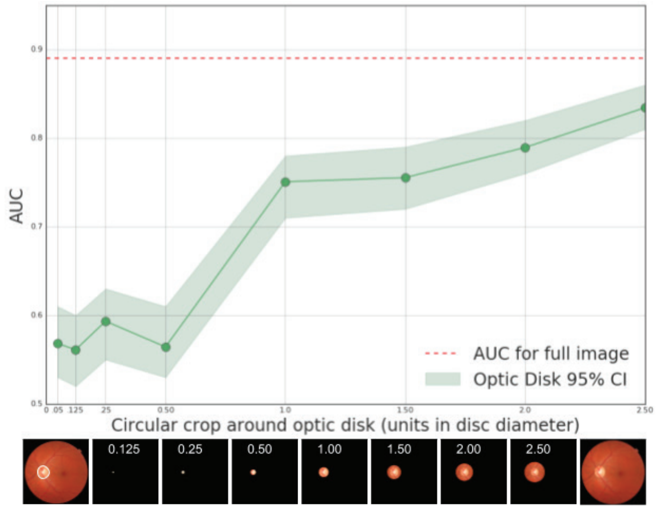
Briefly, the method firstly identifies salient regions; it creates image translation models which determine what in the region influences the predictions; it amplifies those modifications to enhance human interpretability; and finally it extracts a minimal set of hand-engineered features which are used to train an SVM classifier. The performance of this is compared with a CNN trained on raw images. Are the two outcomes going to be similar?

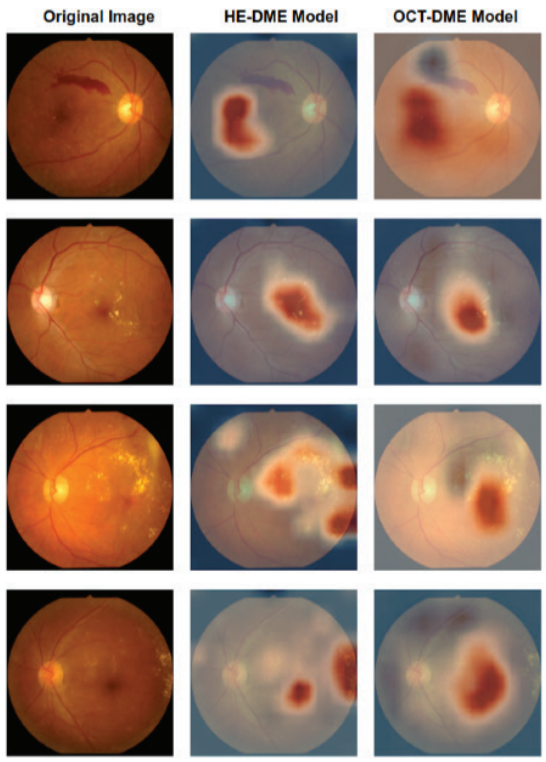
Let's look at all the steps more in depth:

0) Train **a CNN model M** based on inception-v3 on the DME dataset (obtained AUC of 0.847). A multi-task version of this model is also trained (AUC of 0.89).

1) Input ablation to evaluate the importance of known regions: two known landmarks (optic disc and fovea) are used. Circular crops of different radii (0.25 to 5) are cropped around both landmarks and the rest of the pixels becomes background. From this experiment, the authors can conclude that the model gets most of its information from the region surrounding the fovea (i.e. the macula). **The model is looking at the right region**.



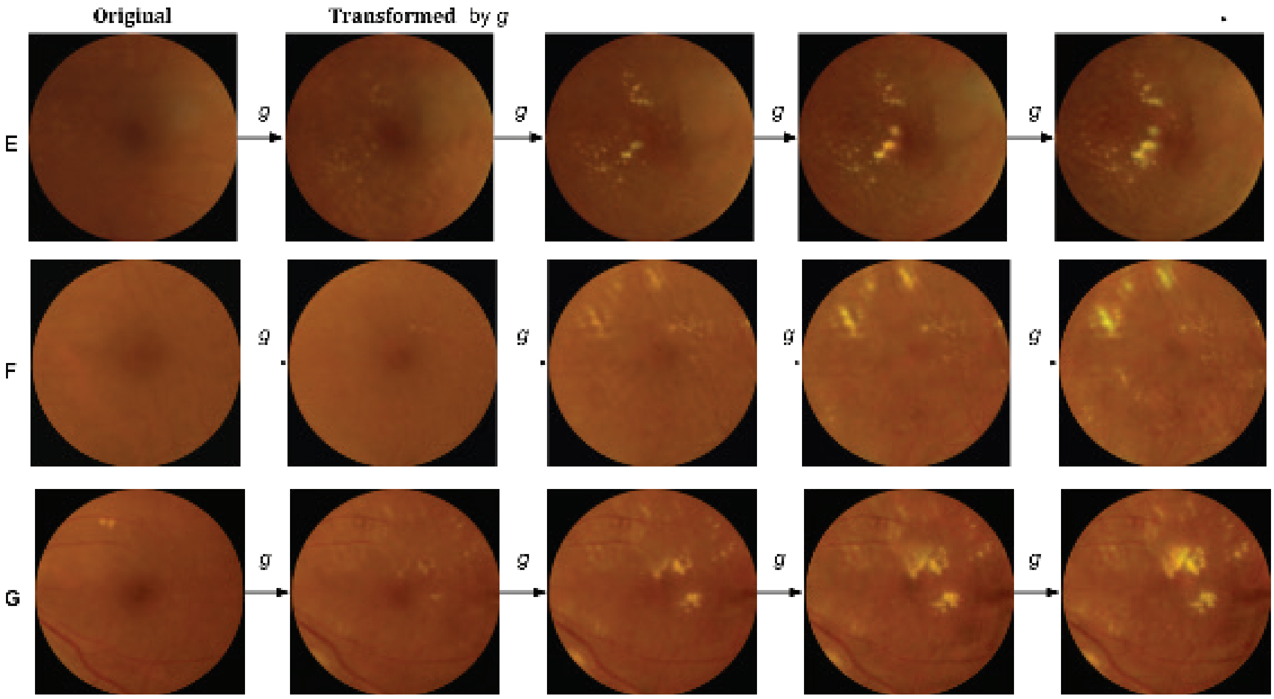

2) Create visual saliency maps which are employed to show visual support. The GradCAM method is trained to visualize differences between the HE-labels and OCT-labels. This shows that **the former focus on hard-exudates (yellow lesion) and the latter centers around the fovea**.

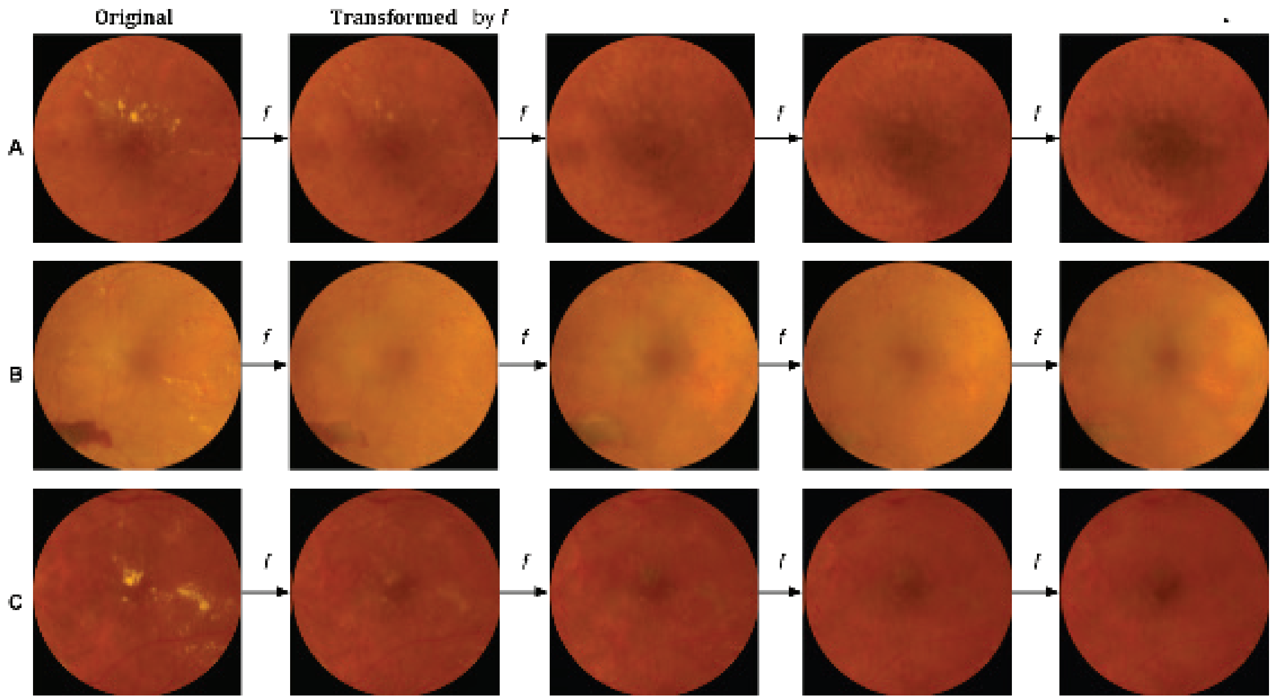
But the generated heatmaps can be tricky to interpret, especially in the medical field, as they estimate where the model is looking to make its predictions, but unfortunately the appearance of medical images is not as straightforward as natural images.

3) Train an unpaired image to image translation method. The authors modify the classic CycleGAN adding a 1x1 convolutional path from the input to the output and modelling functions f (for transformation from DME to no-DME predictions) and g (for transformation from no-DME to DME predictions) as residual networks. To verify that the CycleGAN works, a model M is trained on both original and translated images and the area under the curve (AUC) is recorded; the results show that the **CycleGAN is able to successfully fool prediction model M into thinking images are from the opposite class**, and it continues to improve with successive applications of f and g. The AUC for input source-target pair $(x,y)$ is 0.804 while the AUC for successive applications $(g^4(x),f^4(y))$ is 0.106.

4) Amplify differences from image transformation by successive applications of f and g. The two main changes are observed:

- **hard exudates are added (for g) or removed (for f).**
- **fovea region is brightened (for g) or darkened (for f).** This is visible only with successive applications, but a single application is enough to be picked up by **the classifier M** -> the enhancement of features is only needed for the purpose of explainability.



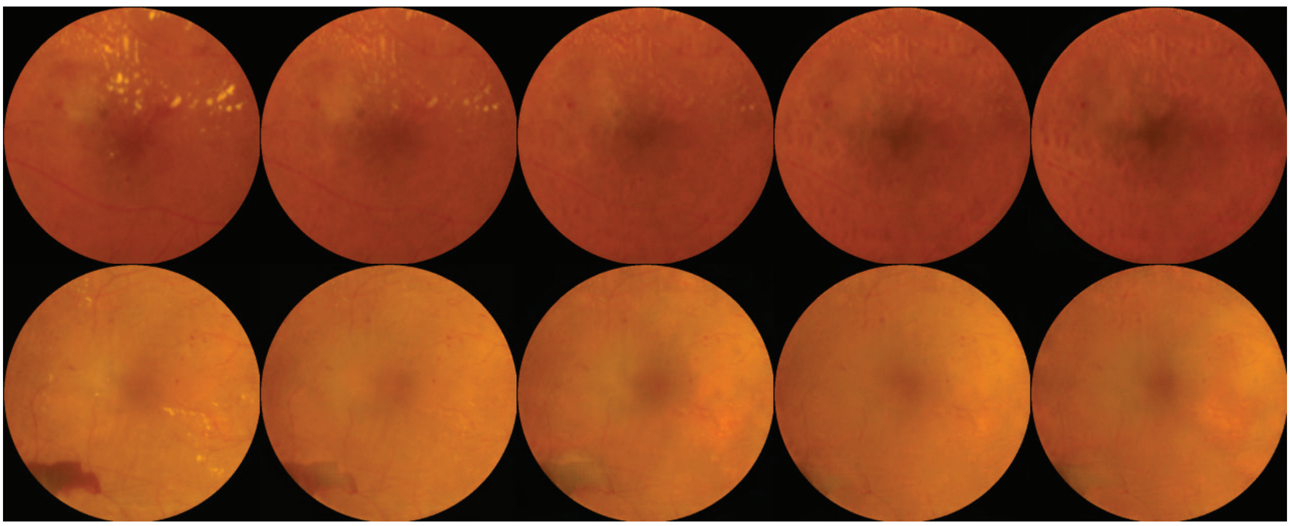
5) Identify and evaluate hand-engineered features. These are created by computing mean RGB intensities within 10 concentric circles (5mm apart) around the fovea. These are engineered with the purpose of evaluating previous hypotheses by the image translation method of the essentiality of the fovea area for the model predictions. When used in combination with the presence of hard exudates to train **two simple classifiers (a linear SVM with l1 regularization and an MLP with 1 hidden layer),** these **perform similarly to the model M (CNN on raw pixels single task on cropped image) for the task of DME classification**.

| Features | **SVM** (AUC) | **MLP** (AUC) |
|---|---|---|
| Hand-engineered features alone | $72.4 \pm 0.0$ | $76.3 \pm 0.3$ |
| Presence of hard exudates alone | $74.1 \pm 0.0$ | $74.1 \pm 0.0$ |
| Hand-engineered features + hard exudates' presence | $\mathbf{81.4 \pm 0.0}$ | $\mathbf{82.2 \pm 0.2}$ |
| $M$ (raw pixels single task on cropped image) | CNN: 84.7 | |

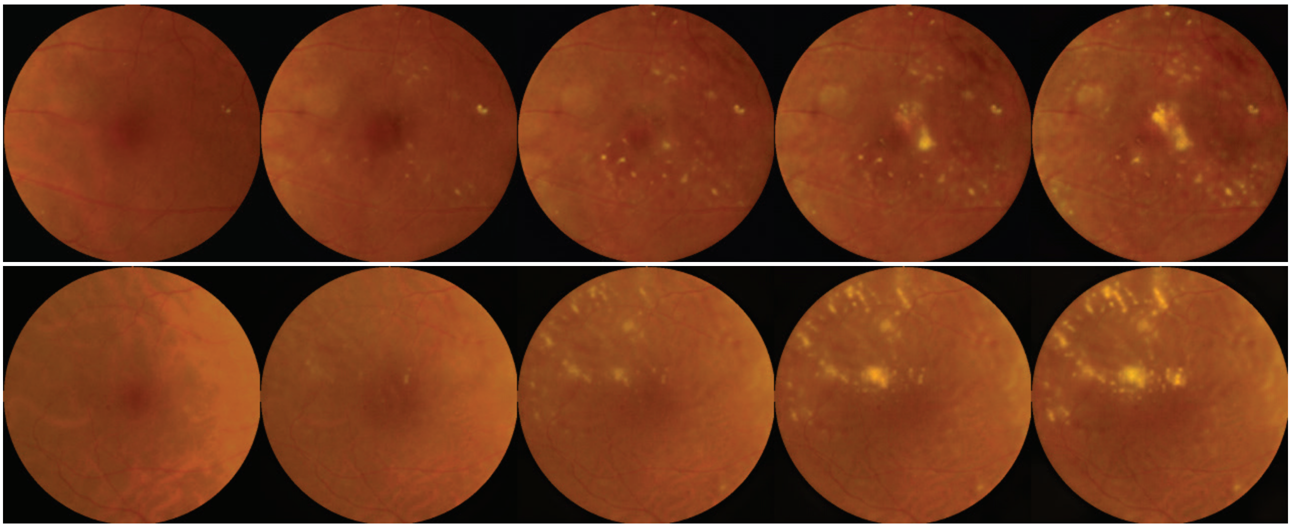
To summarize, the most important contributions of this work include 1) the creation of a framework to convert predictions from explanation techniques to a mechanism for discovery; 2) a proof that generative models in combination with black-box predictors can generate hypotheses without human priors that can be critically explained; and 3) a case study on classification for retinal images predicting DME.

Some nice extra images from this work below…

Examples of converting CFP images from DME to no-DME predictions (model f).

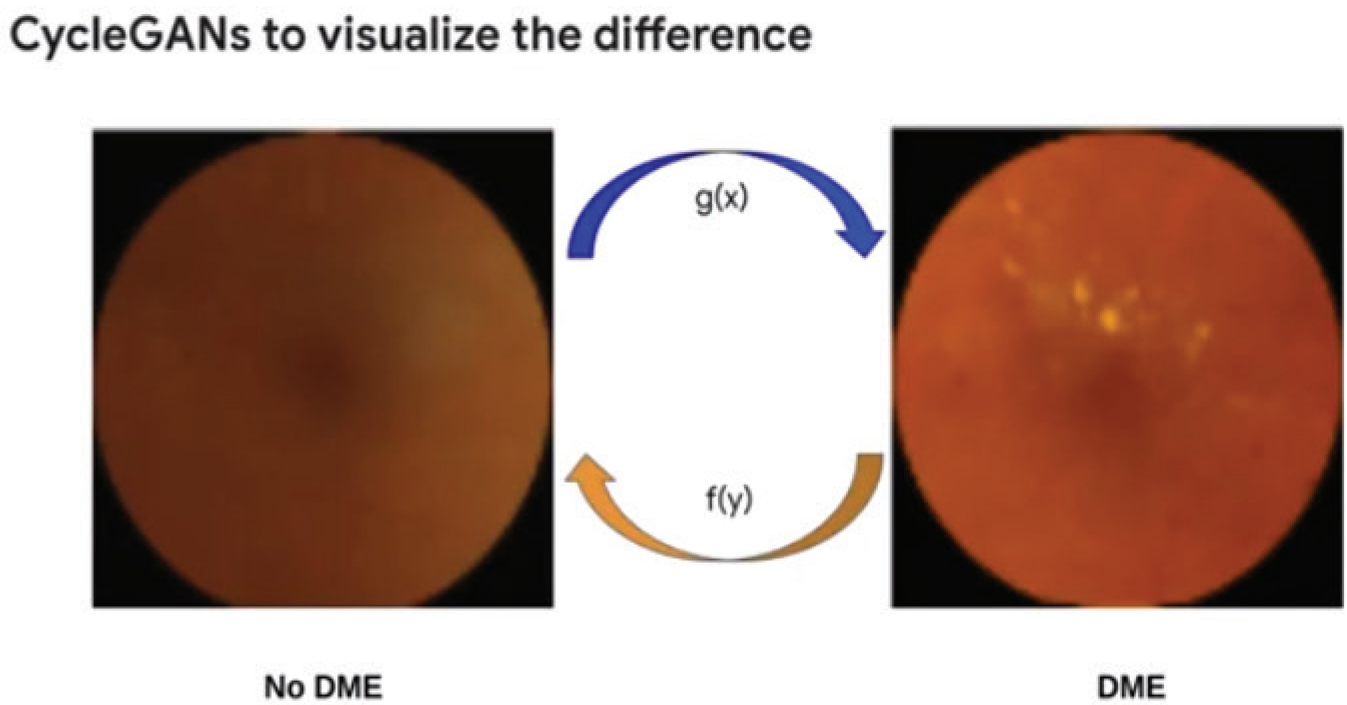


Examples of converting CFP images from no-DME to DME predictions (model g).



If you want to know more about this work, why don't you watch the comprehensive presentation of this paper at MICCAI (video below)?

While you wait for our next article, it's also definitely worth checking out the latest research lead by Google Health and the authors of this paper!
See you all next month ☺



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

IOANNIS VALASAKIS, KING'S COLLEGE LONDON

@WIZOFE

I hope you had a good month and you're again eager on learning about a new tool in the great landscape of medical AI. This month, I am going to present a tool, developed recently by **Microsoft Research Cambridge**, which aims at "democratize medical imaging AI". Sounds like a good bet, doesn't it?

Now, read further to find out how this is going to happen! I hope, I already have you with me ☺

## What is InnerEye?

InnerEye aims at providing *"automatic, quantitative analysis of three-dimensional medical images"* delivered as an open-source solution. The project is a collaboration between **Microsoft AI** and **NHS** as well as researchers. By using state of the art and automatic deep learning on medical imaging on the cloud, it provides a safe, coherent workflow which will enable *"researchers, hospitals, life science organizations, and healthcare providers to build medical imaging AI models using **Microsoft Azure**".*

One of the issues in the field of medical AI is the growing demand and the delivery of a coherent workflow, with great data management. That's important for multiple reasons: workflow acceleration, patient safety, data safety and data integrity. By using **a cloud service** (in this project that would be Azure), that has a very high level of security, enabled with extra secure measures for institutions, researchers and hospitals, the hope is to achieve great coherency in all those.

### Introduction

A recent paper published by Microsoft Research Cambridge and its evaluation shows how AI can augment and accelerate clinicians' ability to perform radiotherapy planning **13 times faster**. There is an interesting blog on the website about how the image segmentation model (**convolutional neural network based on a 3D U-Net architecture**), which has approximately 39 million trainable parameters, is used in the radiotherapy practice.

In the Fig 1. a comparison, taken from this paper, is shown with the main results showing that the ML model reduces the time it takes for end-to-end image segmentation and annotation in radiotherapy. Both the time to draw segmentation contours on the image and the time to correct inaccuracies in the automated (or semi-automated) system are included in this comparison.

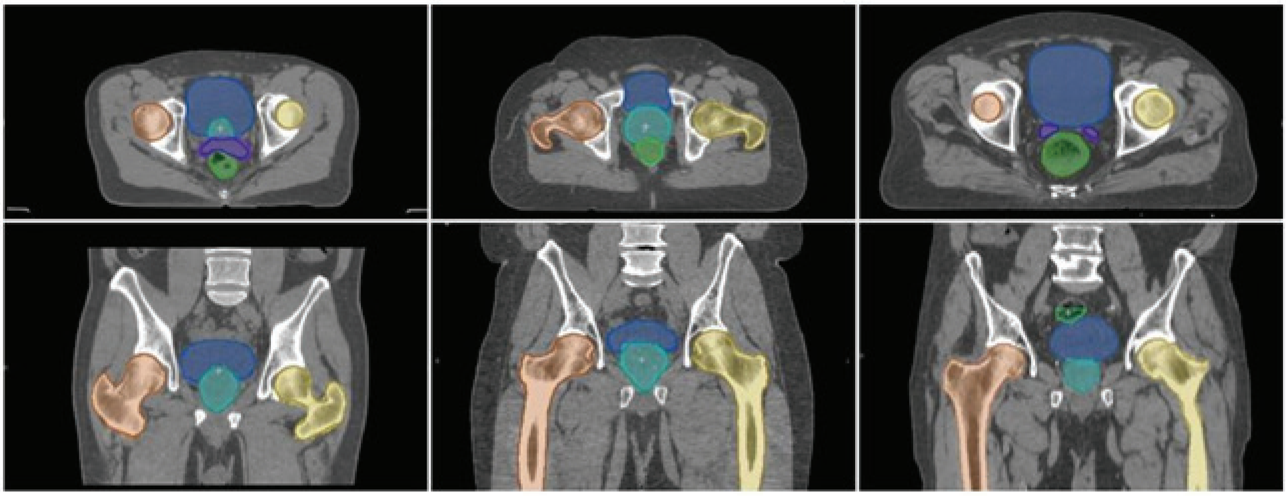


Figure 1: The dark colurs are the algorithm predictions which are compared to the ground truth annotations (lighter colors are the reader annotations). From left to right (columns displaying mid-axial and mid-coronal slices): first left scan is from the main dataset and the remaining two are from the external dataset. The differences between the datasets on patient anatomy and through-plane scan resolution are shown.

Hospitals and healthcare providers could utilize InnerEye as a toolkit to create personalized products and services using the layers of machine learning. They can also deploy them in hospitals and clinics using **Azure Machine Learning** and/or **Azure Stack Hub** (of course in that case regulation needs to be in place, e.g. **FDA clearance or in-house exemption controls**).

I am sure you are now curious to see some of the code and ideas on using it. So, let's go!

### Dataset creation

To create a dataset for segmentation or classification tasks, the following methods need to be used, followed by uploading the dataset as an **AzureML blob storage**. Here an example about the segmentation dataset will be given.

The structure expected by InnerEye is as follows:

- Each subject has one or more scans, and one or more segmentation masks. There should be one segmentation mask for each ground truth structure (anatomical structure that the model should segment)
- For convenience, scans and ground truth masks for different subjects can live in separate folders, but that's not a must.
- Inside the root folder for the dataset, there should be a file dataset.csv, containing the following fields at minimum: subject (unique int > 0), channel (scans/ground truth masks), filePath (either nifty, numpy or hdf5 formats).

An example would be:

```
dataset_folder_name
├──dataset.csv
├──subjectID1/
│   ├── ct.nii.gz
│   ├── heart.nii.gz
│   └── lung.nii.gz
├──subjectID2/
|  ├── ct.nii.gz
|  ├── heart.nii.gz
|  ├── lung.nii.gz
├──...
```

The dataset.csv for this dataset would look like:

```
subject,filePath,channel
1,subjectID1/ct.nii.gz,ct
1,subjectID1/heart.nii.gz,structure1
1,subjectID1/lung.nii.gz,structure2
2,subjectID2/ct.nii.gz,ct
2,subjectID2/heart.nii.gz,structure1
2,subjectID2/lung.nii.gz,structure2
```

**Building up models**

The next step would be to start building model. Of course, if there's a pre-existing model, it can be ported to the InnerEye system; let's see how!

A configuration settings.yml, a folder which includes model configurations (InnerEyeLocal) and a file InnerEyeLocal/ML/runner.py are needed. Below such a sample file is shown:

```
from pathlib import Path
import os
from InnerEye.ML import runner

def main() -> None:
    current = os.path.dirname(os.path.realpath(__file__))
    project_root = Path(os.path.realpath(os.path.join(current, "..", "..")))
    runner.run(project_root=project_root,
               yaml_config_file=project_root / "relative/path/to/settings.yml",
               post_cross_validation_hook=None)

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

Let's say a prostate segmentation is run. The model can be simply inherited from a base ProstateBase parent file and an example would be as follows:

```
from InnerEye.ML.configs.segmentation.ProstateBase import ProstateBase

class Prostate(ProstateBase):
    def __init__(self) -> None:
        super().__init__(
            ground_truth_ids=["femur_r", "femur_l", "rectum", "prostate"],
            azure_dataset_id="name-of-your-AML-dataset-with-prostate-data")
```

Similarly, for a head and neck model, the inheritance and sample code are shown below:

```
from InnerEye.ML.configs.segmentation.HeadAndNeckBase import
HeadAndNeckBase
```

```
class HeadAndNeck(HeadAndNeckBase):
    def __init__(self) -> None:
        super().__init__(
          ground_truth_ids=["parotid_l", "parotid_r", "smg_l", "smg_r", "spinal_cord"]
           azure_dataset_id="name-of-your-AML-dataset-with-prostate-data")
```

Now if the model is created (or adapted to run for InnerEye) how does the training work? If one wants to train on the Azure platform using multiple machines (nodes) the following command is to be run:

```
python InnerEyeLocal/ML/runner.py --azureml --model=Prostate --num_nodes=2
```

To test an existing model registered in AzureML:

```
python Inner/ML/runner.py --azureml --model=Prostate --cluster=my_cluster_name \
  --no-train --model_id=Prostate:1
```

Of course, as a last step, one would need to visualize the results. Assuming a python file which performs the cross validations and plots this, one needs to run the following command to get such results:

```
python InnerEye/ML/visualizers/plot_cross_validation.py --run_recovery_id ... --epoch ...
```

All the results from the model inference, any plots and validations are written in the output folder. A specific structure is followed (with checkpoints saved, metrics in CSV format and additional build information and much much more!). Feel free to investigate the documentation for more details of all the useful files that are created!

## Deployment

An overview of a deployment method is shown in Fig 2. For deployment of a segmentation model, all code that was used in training, plus the checkpoint(s) of the model, are packaged up into a folder and then registered in AzureML for later use.

The checkpoints are chosen from the checkpoints for the epochs specified in the epochs_to_test parameter: the model is evaluated on the validation set for those epochs. The best dice score is written in the checkpoint file (as described in the previous section).

In AzureML, there's a "Models" section where all the **source code** and **checkpoints** are saved. The structure is for model: name, a numeric version and tags and properties.

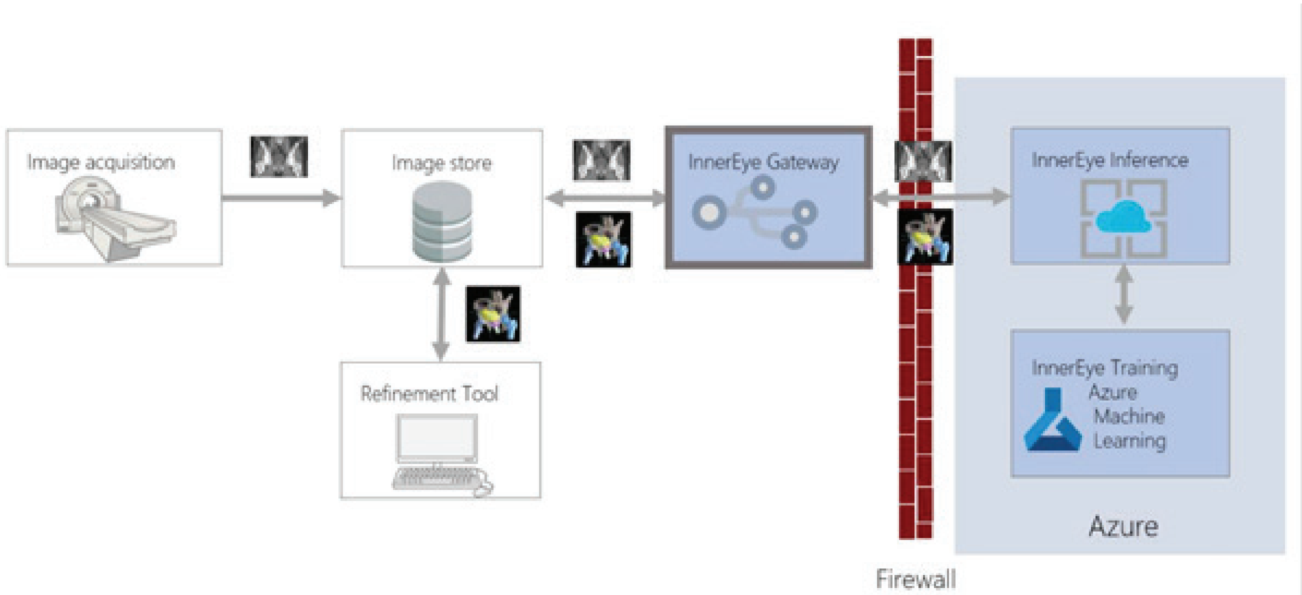



Figure 2: InnerEye segmentation models using a single DICOM series as input and producing DICOM-RT can be integrated with DICOM networks using: InnerEye-Gateway: a Windows service that provides DICOM AETs to run InnerEye-DeepLearning models InnerEye-Inference: a REST API for the InnnEye-Gateway to run inference on InnerEye-DeepLearning models. More information for this example can be found on the documentation.

## Wrapping up!

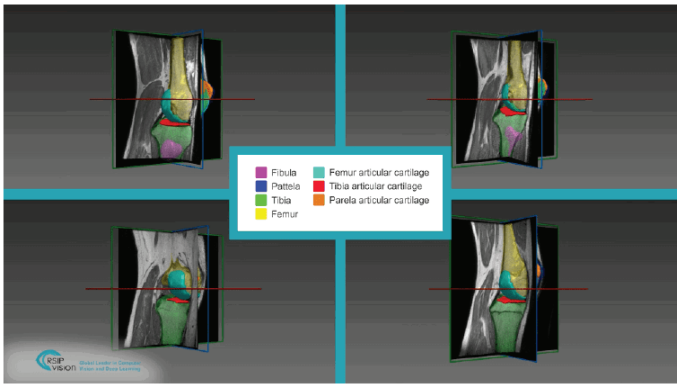
I hope that you discovered something new and interesting this month. Let me know what you are going to do with it!
As always, please let me know if you have any questions, or suggestions for the article! It would be great to hear more of you and what tools you would like to be presented; feel free to reach out to me in any of the social media ☺

# Assessing Articular Cartilage Health Using MRI Segmentation

## Hyaline cartilage and chondral lesions

Articular cartilage plays an essential role in protecting the end of bones and transmitting forces across joints, such as the hip, knee, or ankle joints. The surface of articular cartilage consists of **hyaline cartilage**, which provides a smooth, low-friction surface for the transmission of forces through joints. The name ‘hyaline’ cartilage is derived from the glass-like properties associated with its collagen-rich structure, including firmness, smoothness, and translucence.

Nonetheless, because articular cartilage is **avascular** (contains no blood vessels), its capacity for healing and regeneration is very limited. As a result, chondral (or osteochondral) lesions are a common occurrence, appearing in up to 60% of knees undergoing arthroscopy. **Chondral lesions** are particularly prevalent among athletes in sports that involve repeated stress and strain on the joints. Left untreated, these lesions can not only limit the athlete’s ability to participate in sports, but they can also affect daily activities and lead to degenerative changes and premature osteoarthritis.




## Treating cartilage damage

A variety of conservational treatments for chondral lesions exists, including the following surgical procedures:

- **Micro-fracture (MF)-** the goal of this surgical procedure is to restore the cartilage surface to its healthy state using bone marrow stimulation. MF is suitable for small chondral lesions (less than 25 mm in diameter) and it is usually the first line of treatment.
- **Mosaicplasty and osteochondral autograft transfer system (OATS)-** these surgical procedures entail harvesting and transferring small, cylindrical osteochondral grafts (2.7 to 8.5 mm in diameter). These grafts are harvested from non-weight-bearing areas and transplanted in the damaged or defected regions.
- **Autologous Chondrocyte Implantation (ACI)-** an advanced surgical procedure that entails harvesting cartilage grafts, growing

the sampled cells for several weeks in a specialized lab, and implanting the grown cells in the damaged regions.

For the physician to accurately assess the extent of the cartilage damage and select the most fitting conservational treatment, the cartilage must be imaged and segmented accurately and non-invasively. Articular cartilage is imaged using **magnetic resonance imaging (MRI)** and then segmented and evaluated for lesion shape, size (diameter), and boundaries. In addition, any healthy cartilage in non-weightbearing regions is mapped for possible cartilage harvesting and transfer.

## Cartilage segmentation using MRI

The novel cartilage segmentation tool developed by **RSIP Vision** uses **artificial intelligence technology** to provide accurate assessment of articular cartilage health in MRI scans. This new module provides **fully automated segmentation** and precise assessment of the integrity of damaged cartilage in a variety of joints, such as hips, knees, and ankles.

Using deep learning algorithms, RSIP Vision's module accurately measures the exact location, diameter, shape, and boundaries of both chondral lesions as well as healthy, non-weight bearing cartilage. *"It is very valuable to be able to accurately map chondral lesions preoperatively,"* says Dr **Shai Factor**, an orthopedic surgeon at **Tel Aviv Sourasky Medical Center** in Israel. *"Analyzing the parameters of the lesion and its boundaries allows the surgeon, along with the patient, to choose the ideal cartilage repair technique"*. Evaluating these crucial geometric features of both healthy and damaged cartilage supports the physician in selecting the most suitable **cartilage-sparing procedure**, resulting in **improved patient outcome** and **shortened recuperation downtime for athletes**. RSIP Vision's segmentation module offers not only an improved way to evaluate cartilage damage and select the best treatment, but also a way to follow-up postoperatively and assess the treatment efficacy.

This module follows RSIP Vision's novel module for segmenting bones in **computerized tomography (CT) scans with metal implants present**. This existing tool was designed and trained to deal with the presence of metals in CT scans, offering accurate and robust segmentation of both bones and implants. Together, these two novel modules offer better diagnosis and treatment for patients undergoing orthopedic joint procedures, including the many patients that undergo additional or follow-up procedures throughout their lifetime.

# Beyond pixel-wise supervision: semantic segmentation with higher-order shape descriptors

**Hoel Kervadec is a postdoc researcher at CRCHUM in Montréal. His work proposing a new approach to semantic segmentation has just won the coveted Best Paper award at MIDL 2021, which recognizes the highest quality full-length paper presented at the conference. We speak to him to find out more about it.**

In this paper, Hoel and the team explore a way to supervise a network by **describing where an object of interest should be and what shape it should have**, instead of micromanaging the network output by telling it the label of each pixel in an image.

Existing loss functions used to train networks, like cross-entropy or dice loss, treat image segmentation as independent pixel-wise classification and do not take the image space into account. You could shuffle the pixels and get the same computed value because it is not affected by the shape of the object.

***"We compute the location of the predicted segmentation and some descriptors of its shape***," Hoel explains.

*"Supervision based on a description of*

Hoel Kervadec

*the object rather than the pixels **can be reused more easily across scans or patients in the future**. However, with pixel-wise supervision for a specific scan, you might need to redo the complete annotation."*

This opens up **new ways to supervise networks and use labels**. For example, one could directly define a problem by encoding anatomical information about it, rather than relearning what is already known through pixel-wise annotation.

The work has been **three years in the making** and not without its challenges.

*"It is an extension of our previous work on **constrained deep networks**, but it took some time to reframe the problem,"* Hoel tells us.

*"We went back and forth between a few papers, which were ultimately all related to the same application and methodologies. We needed to improve on the supervision method with the constraints and it took us some time*

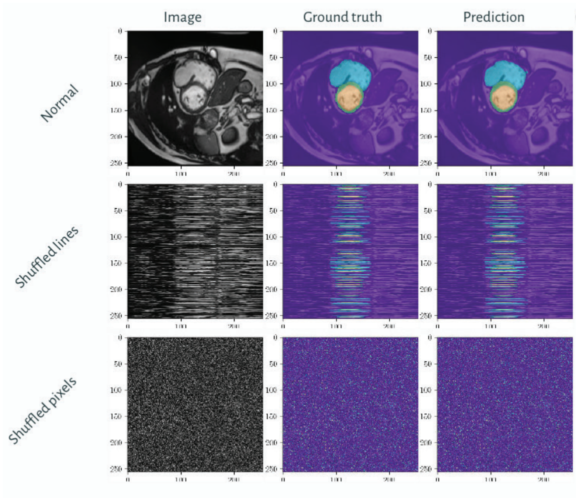
*to fully realize all the possibilities that were opened up by it.”*

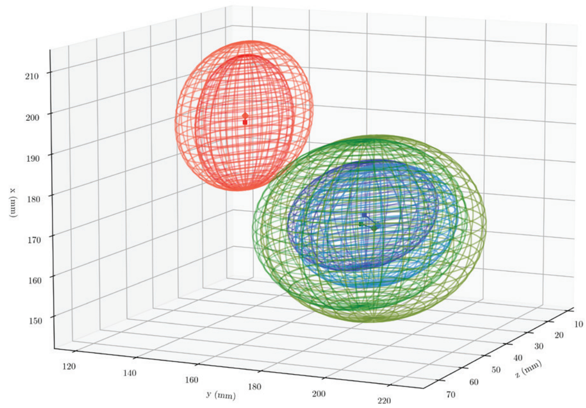
The paper focuses on **2D segmentation**, but it can be easily **generalized to 3D**. One example would be computing the 3D descriptors for the minimum and maximum size of a heart in its cardiac cycle, and the diameter of the different ventricles at those two extremes, meaning that any point in the cardiac cycle will fall between those two bounds. You then have a **valid description** which you could reuse to supervise future scans for that patient **without needing extra annotation**.

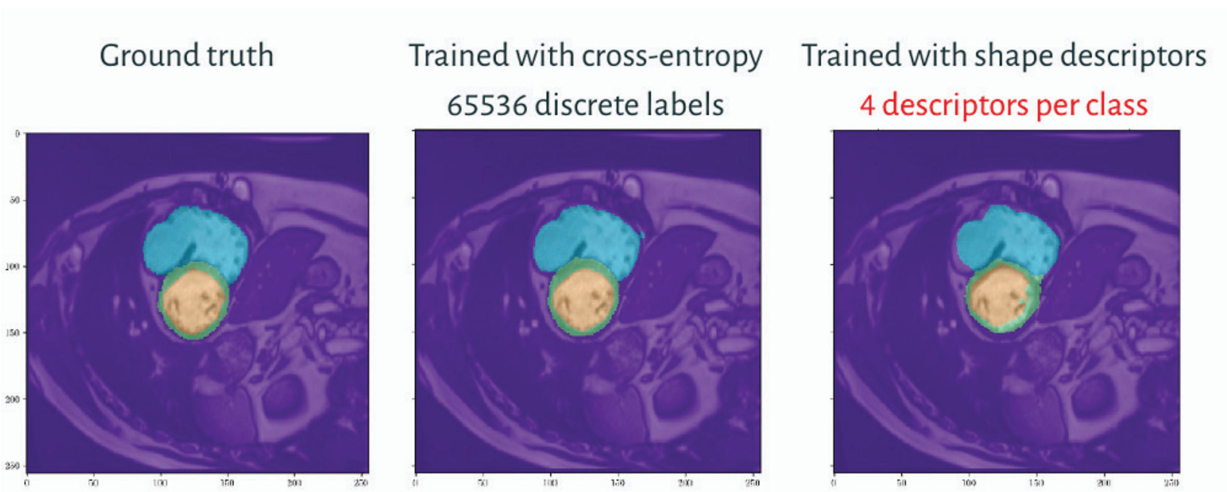
Originally from **France**, Hoel wanted to travel, and moved to **Canada** for his PhD when he was offered the opportunity to join **Ismail Ben Ayed**’s group. He is moving to a new postdoc position in **Marleen de Bruijne**’s group at **Erasmus MC** in September.

Hoel says it was only after submitting this work that he fully appreciated its potential impact.

*“What we show about generalization in 3D and defining bounds per patient would be valid for any scan in the future. It was a bit risky to submit it as we did not have a state-of-the-art result, but rather **a different way to think about the problem** and to make sense of the existing descriptors in the literature. **It opens up so many possibilities for new research**.”*

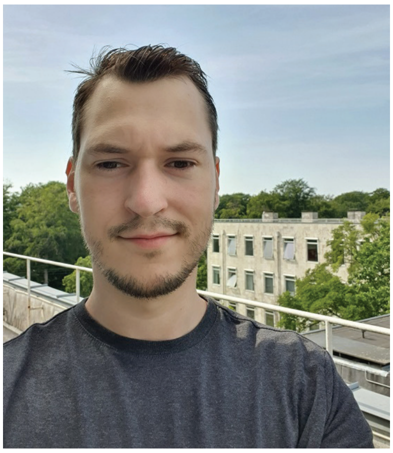




# Semantic Similarity Metrics for Learned Image Registration

**Steffen Czolbe is a PhD student in the Department of Computer Science at the University of Copenhagen, under the supervision of Aasa Feragen and Oswin Krause. He presented his work proposing a data-driven semantic similarity metric for image registration at MIDL 2021 last month, and the judges were so impressed with it that he took home the Best Paper - Runner-up award! He tells us more about his exciting research.**



Steffen Czolbe

Most recent progress in image registration has been achieved via deep learning methods. Modern research is much more likely to train a neural network to predict a transformation that can be applied to register two images, than it is to **iteratively optimize their similarity**.

**Comparing two images is a complex task**. One simple way to do it is to compare your aligned image to your target image by looking at individual pixel intensities using mean squared error. However, in this paper, Steffen proposes a better way.

*"When we align pixel intensities we ultimately say as long as the image intensities are the same between the images, they're well aligned,"* Steffen explains.

*"Instead, what we should align in image registration is areas of the image that have the same meaning. If you are aligning images of the brain, lungs, or another organ, you want to align the same organs and same parts of organs. That's how I got the idea of* ***using a more semantic representation of the images and focusing on aligning that****. I used representation learning to extract representations of the images and ultimately we used these to judge their similarity."*

***"That's how I got the idea of using a more semantic representation of the images and focusing on aligning that!"***

Steffen had worked previously with methods using **variational autoencoders and simple loss**

**functions**, which compare images during training. More recently, these have been superseded by networks trained on image classification – using VGG, for example, to **extract features for comparison**.

This idea of using a separate network to extract features and then judge the distance of images based on that was what inspired this work. However, an image classification network trained on **ImageNet** may work well when you generate faces because the datasets are somewhat similar, but **medical images are radically different, with a range of different modalities**.

To overcome this, Steffen trained specific networks to make sense of these images. He explored using both **an autoencoder and a segmentation model** – the autoencoder has the benefit of being unsupervised, while the segmentation network needs segmentation masks to train.

Looking to the future, Steffen tells us he is currently working on a journal version of this work which has already moved it on further.

*"There have been some developments around other baselines I want to compare against,"* he reveals.

*"For example, **mutual information** is a metric that traditionally people use on image registration, but when I first set out, it hadn't really been adapted to deep learning. Now, I'm looking forward to comparing against that. Overall, I want to **discuss the different similarity measures in more detail**."*

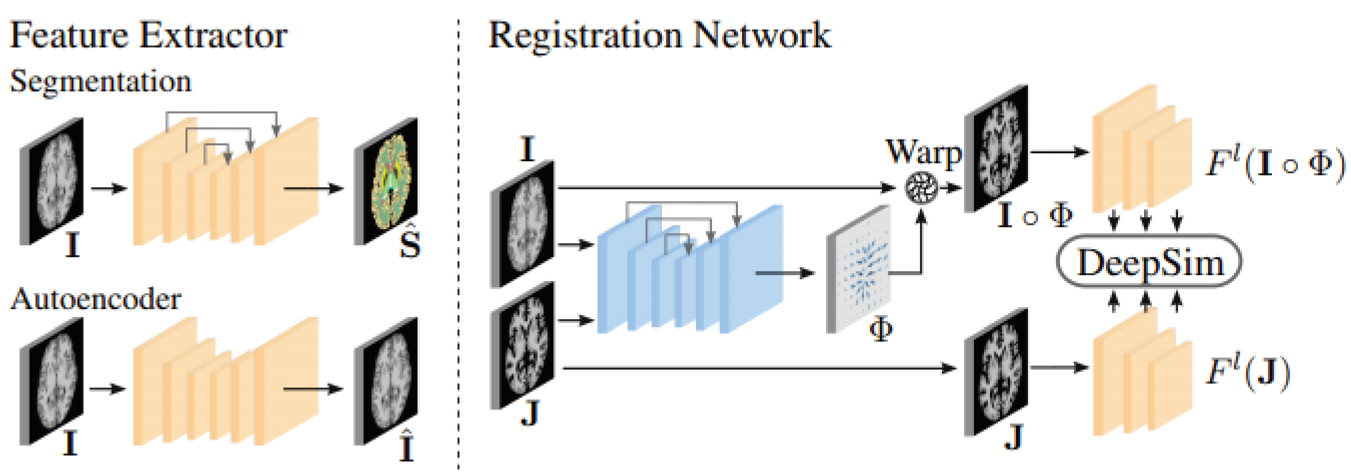



Figure 1: Two-step training: First, the feature extractor (yellow) is trained. We test both a segmentation model, and an auto-encoder. Next, the feature extractor weights are frozen and used in the loss computation of the registration network (blue).

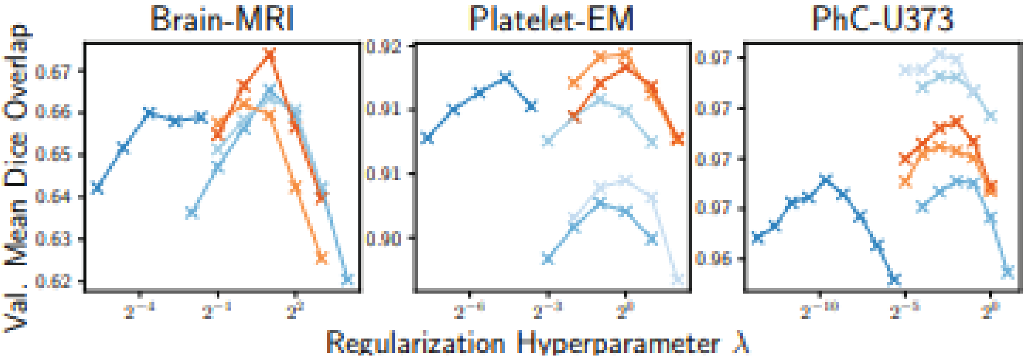



Figure 3: Hyperparameter tuning. For each model, we select the regularizer weighting factor λ with the highest validation mean dice overlap for further evaluation. Color scheme follows Figure 4.

He also plans to test whether images should be warped first before features are extracted, or vice versa. **Warping is a non-linear image alignment**, where all parts of the image are distorted.

Winning the runner-up prize at a prestigious conference like MIDL is no mean feat for anyone, but for someone in **the first year of their PhD**, it is extra special. Why does Steffen think the judges picked his work above so many others?

*"I think what makes it stand out is that it is generally applicable,"* he responds.

*"In the medical domain, a lot of research is very specific to a limited number of applications, whereas with this work I strive to be general. Also,* ***I evaluated on three completely different biomedical datasets, which is very unusual for this community. People generally stick to one dataset*** *and if it works that's enough. By going the extra mile, I think this is a really important contribution."*

The work was positively received by participants too, who were inspired by its novelty.

*"People really liked the idea of training these feature extractors specifically for the dataset, as opposed to what they do in other areas of computer vision where they'll just use VGG-net,"* he adds.

Originally from Germany, where he did his bachelor's degree, Steffen wanted to study abroad and chose the University of Copenhagen because its master's degree gave him the balance he was looking for between **computer science and machine learning**.

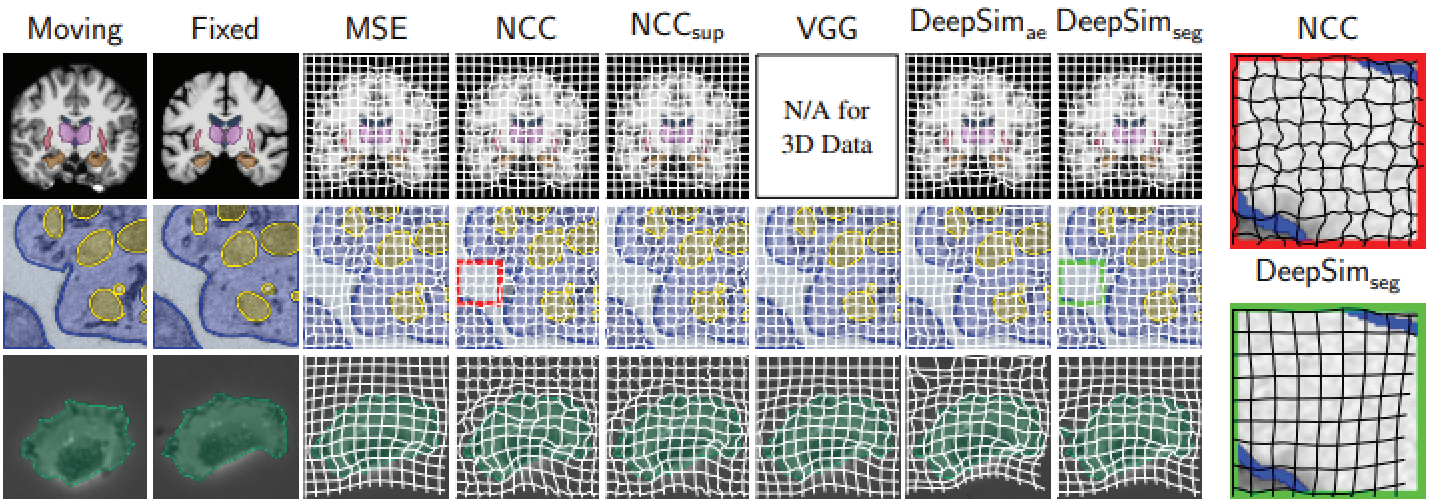
*"There's a very active machine learning research group here,"* he tells us.

*"The university is great, but it's also embedded in a really beautiful city with a very high quality of life and work-life balance."*

As we draw to a close, Steffen reveals that this work could have turned out very differently.

*"I actually didn't set out to do this,"* he declares.

*"The project initially had a completely different goal that didn't work out, but this part did. It spun out of a larger project that was riddled with challenges! **That changed the entire direction of the work**. In the end though, it was relatively straightforward."*




(*a*) Qualitative comparison (*b*) Detail view

Figure 2: Left: Qualitative comparison of registration models. Rows: Datasets Brain-MRI, Platelet-EM, PhC-U373. Columns 3-8: Registration models trained with various similarity metrics. Right: Detail view of highlighted noisy background areas on the Platelet-EM dataset. Select segmentation classes annotated in color. The transformation is visualized by morphed grid-lines.

# Self-Supervised Out-of-Distribution Detection for Cardiac CMR Segmentation

**Camila González is a second-year PhD candidate in the Medical and Environmental Computing Lab at the Technical University of Darmstadt in Germany, under the supervision of Anirban Mukhopadhyay. Her paper explores out-of-distribution detection, and she spoke to us ahead of her short oral discussion session.**

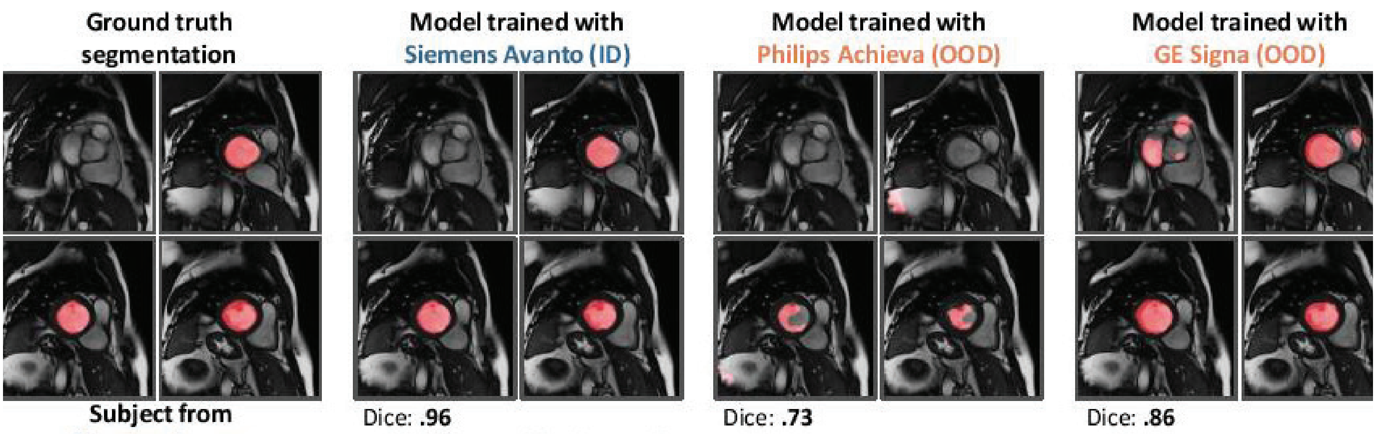
Camila Gonzalez

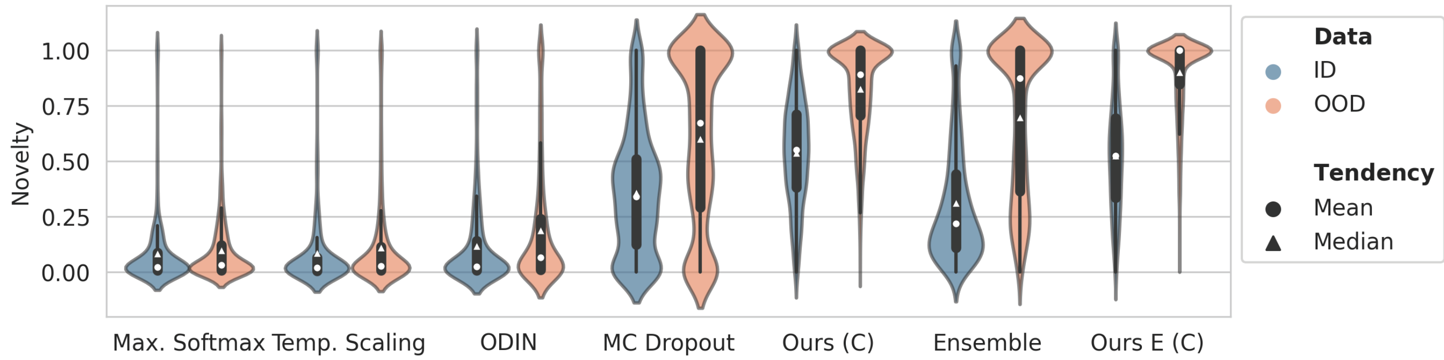
Camila begins by telling us the wider research topic for her PhD is **continual learning for multi-domain data**. Specifically for data that comes from different institutions, or data that was acquired with different pieces of equipment, such as CT or MRT data acquired by machines from different vendors.

When models are trained on this multi-domain data, they often do not generalize well to data from other places or vendors. **This causes deep learning models to fail silently**. For quality assurance, it is important to identify these out-of-distribution samples for which the trained model is unsuitable.

*"Regardless of whether you have a model that aims to generalize well to other domains, you still need methods to **detect when a sample is out-of-distribution,"*** Camila explains.

*"That's why in this paper **we look at self-supervised models and how to use self-supervision loss calculated during testing to detect out-of-distribution samples**. There are many different*






*models and methods that can be used for out-of-distribution detection or uncertainty estimation. As well as having a method that generalizes to out-of-distribution data, you would use all these different methods because you need these* ***extra quality checks****."*

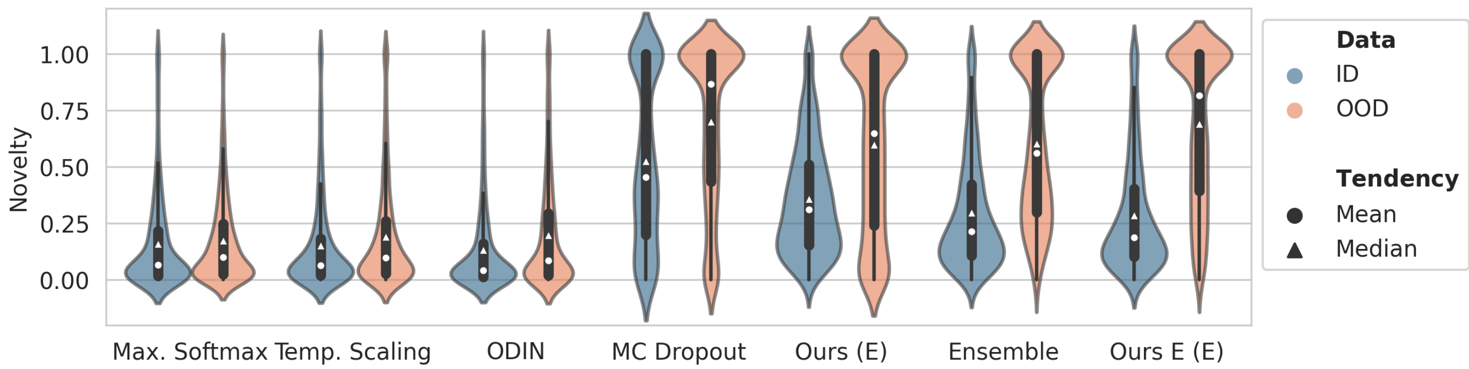
Previous research exists based on classification models, but Camila says it was a challenge to implement existing methods for semantic segmentation and make sure that they are working properly. There are two proxy tasks in the paper – **contrastive learning and edge detection** – chosen because they encourage the learning of geometric information that is relevant for the segmentation models.

*"We weren't sure of course when doing the experiments what the results would look like,"* Camila tells us.

*"We had a strong assumption because it's quite a basic expectation that the models will also fail at proxy tasks, but the most exciting part was when we got good results!"*

There was one challenge that Camila was not expecting, but which could be a lesson for younger researchers out there.

*"Two of the datasets that we used in our work are from the M&Ms Challenge, and I didn't realize until quite late in the day that we could ask the authors for additional details of this data," she reveals.*



*"In the end, we did this as part of the rebuttal process at the review stage. I contacted the authors, and **Victor Campello** from the challenge responded quickly and gave me lots of information. The actual pieces of equipment that had been used for extracting the data didn't appear on the website of the challenge. It made the paper so much better! I felt a little awkward at first, but I'm more confident now and I've encountered other situations since where I have needed to contact people responsible for datasets, so feel more comfortable doing it. **Everyone is friendly in the community!**"*

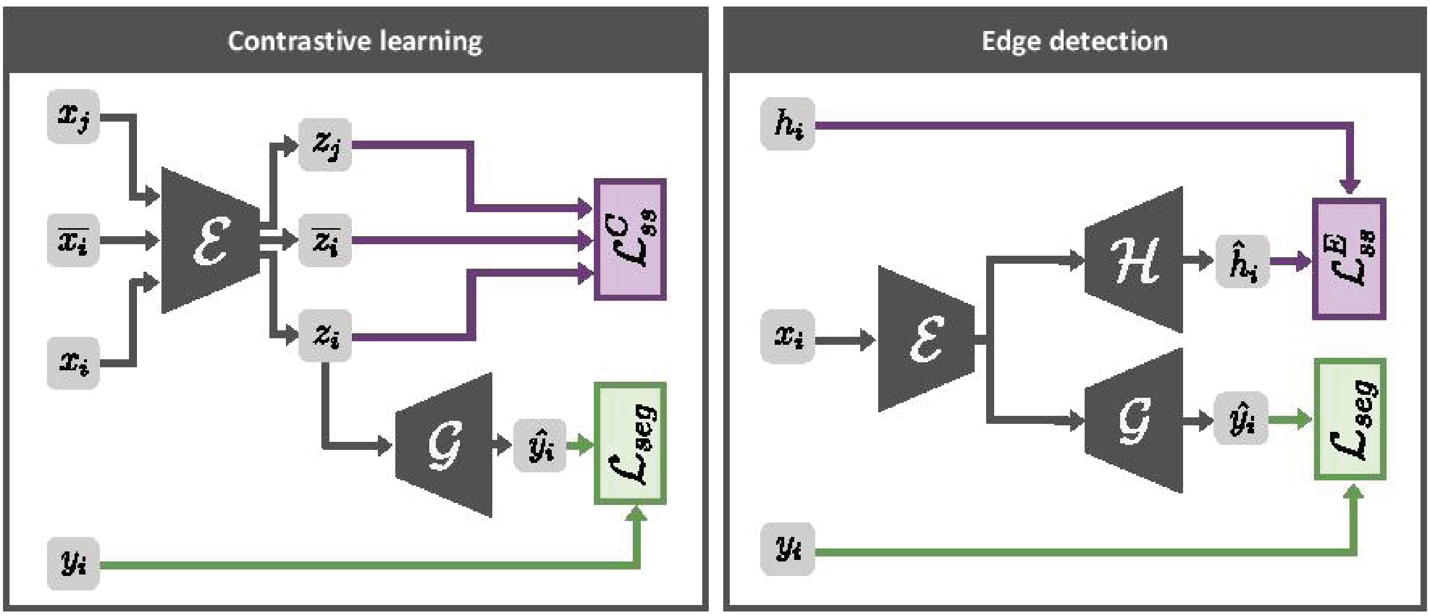
Looking ahead, Camila wants to **implement different self-supervisedtasks**. She would like to know how good the detection signal from the self-supervised loss is by a wide array of proxy tasks.

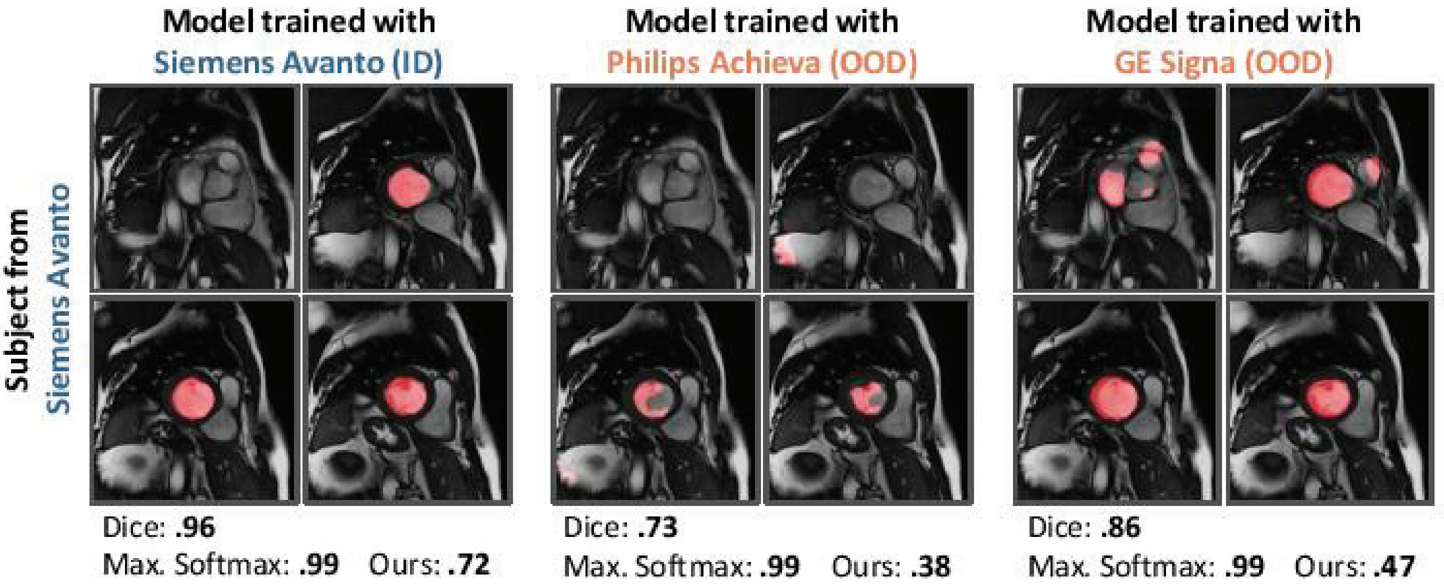
Another desire is to have different uncertainty estimation methods. This work uses **Deep Ensembles and Monte Carlo Dropout**, and Camila tells us they are extending that for TreeNets – architectures where you have one main body and many different heads. She would also like models that include some **variational inference components**.

We ask Camila, if she had a magic wand, which one of the above would she add to her model right now?

*"I suppose if I had to make a choice to implement one then it would be variational inference,"* she decides.

*"Maybe a very nice PhD student could be that magic wand!"*






Camila has another couple of years to go before she has to think about what comes next after her PhD. She is torn between academia and industry but would love to continue working with different types of models and machine learning.

*"I really like this direction of quality assurance or just solutions to make these models usable in practice in a way that is safe,"* she says.

*"We are in a moment in time where we are very focused on improving performance, but then because of safety concerns and regulations we can't use any of what we are doing in practice. We need to find a middle ground where we can use these things in a safe way."*

Originally from **Buenos Aires, Argentina**, Camila came to Germany for her bachelor's degree, but had already attended a German school an

*"That was a bit of a coincidence!"* she laughs.

*"I went to a German school because it was very close to my home, and I learnt German there. Once I finished my high school degree I wanted to go abroad and because I knew German, it was quite easy to apply for universities and get a student visa here. Germany is a very friendly country for students. It's a really nice environment – and inexpensive too!"*

Camila is a member of the **MICCAI Student Board** and can be found on Twitter @camgbus.

## DIDA and DIGITNET

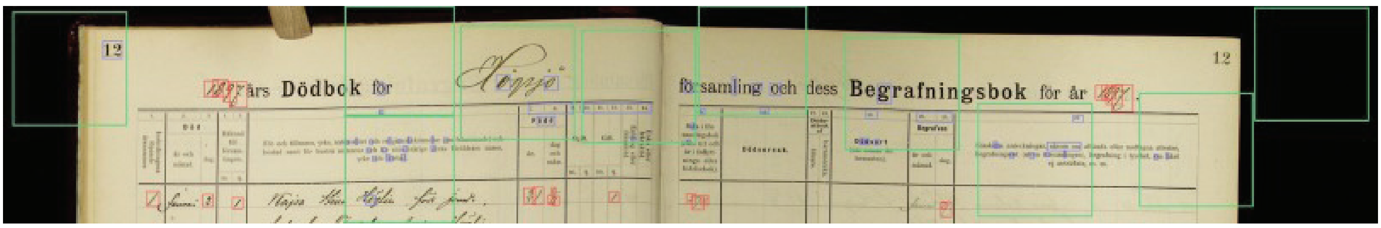
**Amir Yavariabdi is an assistant professor in KTO Karatay University in Turkey. Hüseyin Kusetogullari is a Senior Lecturer in the Department of Computer Science at the Blekinge Institute of Technology in Sweden. They are speaking to us about DIDA, a new historical handwritten digit dataset, and DIGITNET, a deep handwritten digit detection and recognition method.**

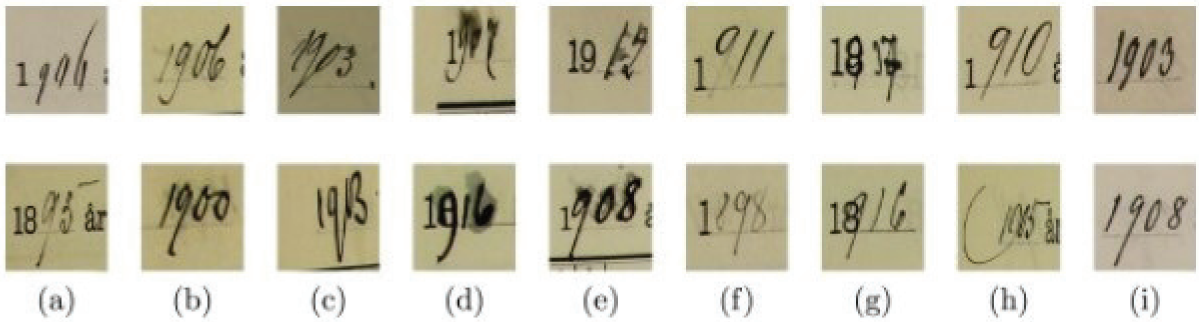

Amir Yavariabdi

**DIDA** is the largest **image-based historical handwritten digit dataset**, with 250,000 single digits and 200,000 multi-digits in RGB color space, collected from historical handwritten document images. Around 75,000 document images were used, written by different priests with different handwriting styles in Sweden in the 1800s and 1900s.

Hüseyin Kusetogullari

**DIGITNET** uses a **YOLO-based detection algorithm network**, combined with an external recognition algorithm for higher accuracy, to detect and collect these digits inside the document images.






Download : Download high-res image (219KB)
Download : Download full-size image

Illustration of different challenges for digit segmentation and recognition: (a) broken digits, (b) single-touch digit, (c) multi-touch digits, (d) degradation, (e) ink flowing, (f) faint, (g) overlapping, (h) size variability, and (i) digits without artefacts.

*"We want to be able to detect the digits and four-digit year strings to support document indexing,"* Amir tells us.

*"Historical handwritten documents like these can be complex, particularly in terms of their writing style and appearance, so reading them with the computer is very difficult."*

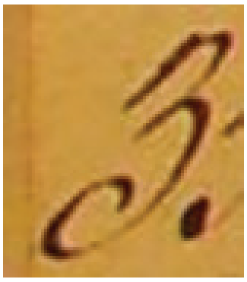
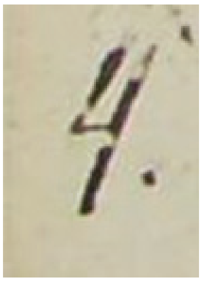
*"Some priests use Gothic-style writing, which is more elaborate,"* Hüseyin adds.

*"There is a world of difference between the way a zero is written, for example, which is generally easy to recognize, and the number two, which can be much more complicated."*

To recognize these digits, they used different models and different available handwritten digit datasets, including **MNIST** and **USPS**. However, they found that the characteristics of these relatively modern datasets limited their application for historical texts.

*"One of the issues with MNIST is that it has a fixed size of 28 x 28 pixels,"* Amir explains.

*"By resizing all these historical digits to 28 x 28 you decrease the quality. Many of them have already lost their colour over the years and degraded in quality; by downsampling them, we lose all their important features."*

Because of this, they came up with the idea to generate a new additional dataset, DIDA, which can cope with a variety of image sizes. DIDA contains **25,000 samples of each digit from 0 to 9**.

*"This also allows the community to apply different types of CNN models,"* Amir points out.

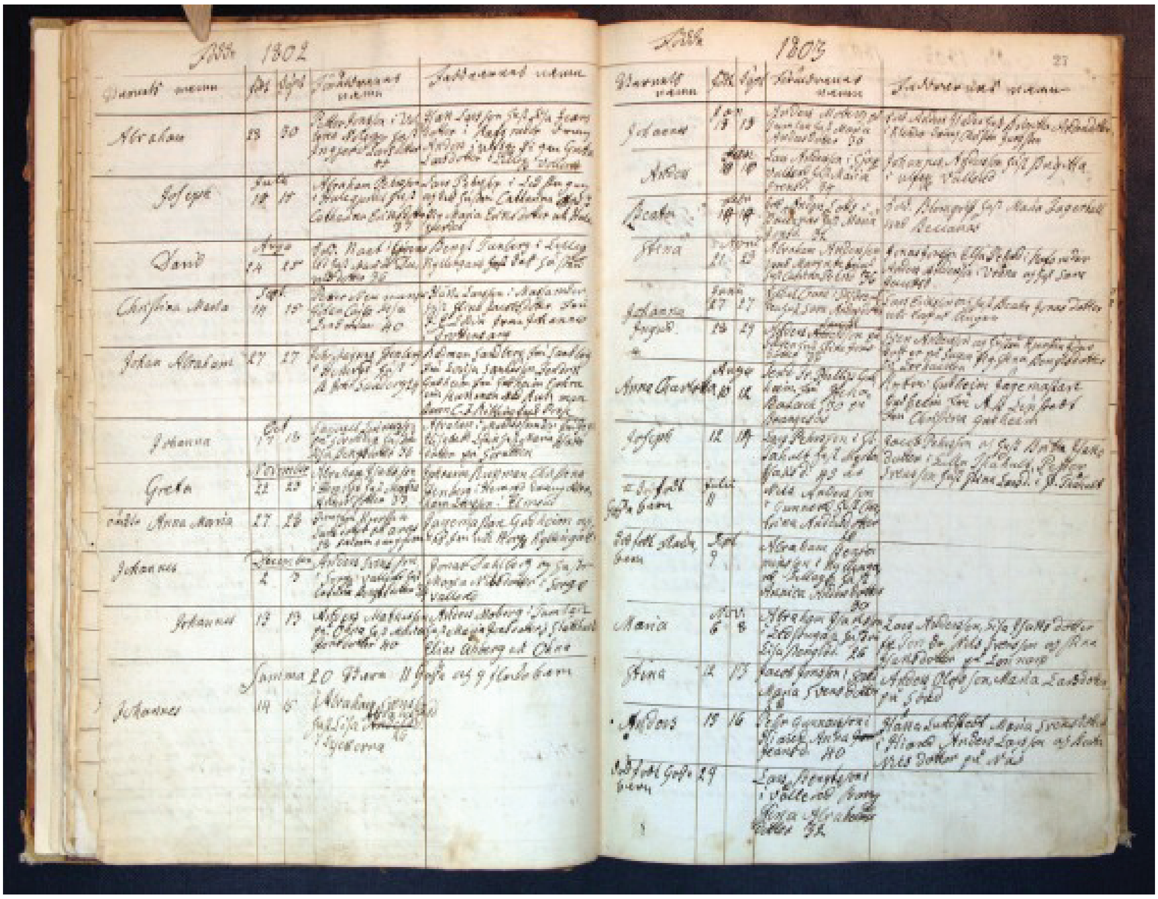
*"With the 28 x 28 MNIST dataset the network contains a maximum of three to four convolutional layers, which is a problem. The more convolutional layers you have, the more information you can take from your images. This is another important characteristic of our dataset. Now, recognition accuracy is around 96-97%."*

DIGITNET is a sub-project of a Swedish big data project funded by the **Knowledge Foundation** in collaboration with **ArkivDigital**. ArkivDigital has collected and digitalized over **90 million historical Swedish handwritten document images** from between the 13th and 21st centuries. It is keen to provide an application like DIGITNET for its customers to improve the search capabilities on its database.

Out of the 75,000 document images Amir and Hüseyin used for DIDA, they have only currently got permission to publish around 15,000 whole document images of historical handwritten birth records, but they hope to be able to publish more soon. These birth records were used for another paper and dataset: **SHIBR – The Swedish Historical Birth Records**.

Within the next year, they are hoping to expand their dataset further. They are preparing character datasets to use for word spotting to understand and label the document images as fully as possible and to make searching the database faster and more powerful.

*"We are now working on a new project,"* Hüseyin tells us.



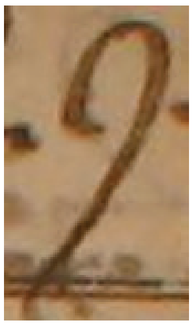
*“ArkivDigital are going to share some more document images with us which were written during the 17th and 18th centuries. Reading, recognizing, and understanding the characters, digits, and sentences in there is going to prove even more difficult!”*

Practically, they see this work as being useful for anyone who is working on historical document analysis, particularly for **Latin-based languages**.

Also, anyone looking to create stronger recognition and detection algorithms will benefit from the huge variety of writing styles available by combining DIDA with MNIST and USPS.

**The DIDA single digits dataset is publicly available for the research community at [https://didadataset.github.io/DIDA/](https://didadataset.github.io/DIDA/).**




**Katja Heuer is a Postdoc in the Center for Research and Interdisciplinarity at the Université de Paris and at the same time a PhD student at Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig.**

**Can you tell us more about your current work?**

I am finishing my PhD, but have already started a position as a Postdoc in a new lab. I am very interested in the development and evolution of the brain.

I use computational neuroanatomy and phylogenetic comparative methods to study the evolution of the brain. I try to understand the origins of the variability that we observe, what is normal variability and what is pathology, and try to put that in the perspective of the human brain evolution. Thanks to collaboration with Roberto Toro who I'm working with in Paris and the Natural History Museum in Paris, I have access to a large collection of brains from many, many different species, which gives me the opportunity to look at a large frame of brain evolution. We have a particular interest and focus on brain folding, to understand the role of brain folding in the organization of these complex structures in our brain that enable us to think, to feel, and to communicate.

**You spoke about the evolution of the brain. Are our brains becoming more sophisticated or less sophisticated?**
That's a very interesting question, and there is no simple answer to that. We use comparative methods looking at >60 different primate species through evolution. We've seen that there is not one trend or one direction that, for example, brain volume will evolve. If you look at evolution, what we observe in primates is that you can observe brain volume decreasing in some branches, in primates that are still living. In other branches, such as the humans, we see increases in volume. In other branches, we see a mix of both. Evolution is not always trending toward the bigger brain.

**Are you telling me that our ancestors, 100 generations ago, had less developed brains than ours?**
No (and 100 generations ago is a very short time if we think in terms of brain evolution). I'm saying it can go both ways. For example, our own species long time ago, they had a larger brain than humans today. How do we quantify sophistication? What's important is that there's an adaptation for everything that we want to do in every species we observe. What is the best adaptation that we want to strive towards? More sophisticated and less sophisticated really depends on what we want to do.

**So, size does not matter?**
At a large scale, there is a slight correlation between the larger the brain

and an observed higher intelligence among humans. But it doesn’t help when you look at the individual level. Now, across species, it’s very difficult to find a way to compare. How we define intelligence and then the framework to compare would be very difficult.

**Let’s not compare. Sometimes results are different from what you expected. Did that ever happen to you?**

Getting different results from what you expected is super interesting. What’s important is communicating them and interpreting them well in the context of what we know. There has been a problem for reproducibility in the field. I encourage people to share data, to work together, to share code, to reproduce their own studies. When you make everything available, you can discover where something has gone wrong and move the field forward altogether.

**You are certainly doing this for a purpose. What will humankind be able to do with the results of your research?**

I’m in basic research, not in applied research. I’m trying to understand better why brain folding emerges, for example. What we observe is an amazing relationship between the folding of the brain and its underlying organization. It's super interesting that sometimes species that are far apart in the phylogenetic tree have a similar folding pattern because they have a similar brain volume. In addition to genes having a causal role on the organization of our brain, there could be physical constraints playing a causal role on the organization of the brain. It's very important to take that into consideration more and more. It gives us a different perspective for looking across evolution to understand variability at a larger scale.

**You are obviously very passionate about the subject. What have you discovered that excites you the most?**
We have been studying brain development and brain folding using real data and computer simulations. When we looked at primate species, it was the first time that we could clearly see species that were genetically further apart, and yet, their folding patterns were so similar. We developed a method that we call fold graphs. It's the first time that we tried to quantify how folding patterns look, how can we compare folding patterns – not the number of the folds but the organization of the folds across different brains within one species, within humans or across species. We looked across species, and

OHBM Hackathon in Rome


Standing on shoulders of giants

what we found there is that even if genetically they are more similar they may have different folding patterns just because their brain volume is different. It was very interesting when we found that out. We also developed a method to match the brains of different species. And when we match them based on volume similarity, jumping between the branches in the species tree, we get a super smooth trajectory of evolutionary expansion. It's fascinating!

**Does your study influence the way that you relate to your own brain? Do you apply your observations in your life?**
I never thought about it that way. I find it super interesting and encouraging when you look at brain plasticity. People

do functional studies on how plastic the brain is. If you have an accident, that can have a tremendous effect on your performance. If these accidents happen earlier in life, they have been observed to have less dramatic effects: the brain can learn to compensate.

**You are also interested in graphic arts and dance. How did this come to be?**
I was very lucky to be in a lab environment where people encouraged my artistic projects. They very much appreciated spending a little time in data visualization but also bringing data to venues which were rather artistic, where you get a very different view and appreciation of the data. In the beginning, we made an installation where you could walk into the brain: you were walking through the slices as if you were walking through your own brain. Now we have developed methods that allow you to have a 360-degree view of the structure of the brain. People are reaching out to us to use these images for different things, journal covers, even for a music album cover. It's real data just visualized in a beautiful way. I'm very happy about that. In a recent project, we analysed the brains of many different primate species, we reconstructed the brain surfaces in 3D. This is an aye-aye for example:

People have been using this figure in supersized posters on their walls in their offices.

**You are making a lot of projects to communicate about the brain. What are you trying to communicate?**
This really depends on the context. I was so lucky to be introduced to the community we call Brainhack very early. It's a community of open scientists that

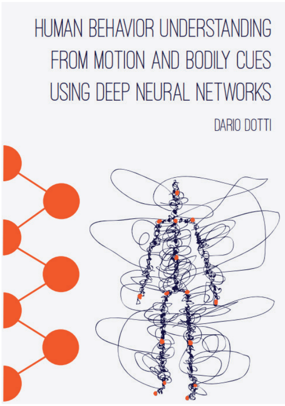
really accommodate everybody. They are so kind to everyone that wants to join, sharing data, sharing code, sharing ideas. I enjoy collaborating with people very much. With Roberto we have built neuroimaging tools to work online with scientists and citizen scientists – for example BrainBox. You just go to a web browser, and you open a webpage where you can browse through the brain. You see the brain slices and we can work together in segmenting different parts. There's a chat, so that you can also come and say, *"How can I help?"* Everybody around the world can help.

**What about your love for dance?**
It's always been really important for me to dance, but it's not really related to my PhD. We have a super nice tool that can translate brain folding into sound! That may be a first step into that direction! [*laughs*]

**What message would you like to share?**
Working together is such a pleasure. I learn so much through the collaborations that we have. Studying the brain is super interesting, and you have more and more questions the further you go. It's beautiful to have this opportunity to be curious and try to dive deeper and deeper and drive your own research.

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


**Dario Dotti completed his PhD in the Department of Data Science and Knowledge Engineering at Maastricht University (The Netherlands) in June 2021. His research focused on developing novel computer vision models for behavior understanding in different scenarios. His passion for this topic began in the framework of the European project ICT4Life, where, in a multidisciplinary team, they built smart services to support elderly living independently. Congrats, Doctor Dario!**




Human behavior analysis from video data is one of the most complex challenges in the computer vision community as movements are difficult to define and lack clear semantic structures. As automatic human behavior understanding is a research topic that can potentially be used to support several fields of our society, during my PhD, with my promoter Stelios Asteriadis and co-promoter Mirela Popa, we focused on two research fields: Video Surveillance and Affective Computing.

***Automatic human behavior understanding for Video Surveillance.***
Human movements are generated in different forms and levels of complexity. Movements can be generated by the full body that moves coherently while performing an activity such as running or jumping, or movements can be seen solely as motion through space towards a destination. In the video surveillance context, we considered the latter definition of movements to build a hierarchical framework for modeling real-time motion trajectories (Figure 1).


Figure 1. Overview of our trajectory prediction (colored in yellow) in the New York Grand Central dataset. Our framework embeds direction, orientation, as well as speed information from trajectory data (using our new feature representation on the right). Orientation and speed are critical behavior cues in surveillance scenarios.

The hierarchical architecture of Autoencoders was designed to capture short spatio-temporal trajectory patches in the lower levels. Short motion patches are highly variant, containing short deviations and turns caused by moving obstacles all around. Then, in the higher levels, we combined motion patches using a grid structure where we smoothed short deviations of individual patches and learned longer and more meaningful spatio-temporal motion trajectories. Finally, we modeled temporal motion transitions using Bayesian probability, by inferring the future trajectory step given the current motion information.

The model was tested in both indoor and outdoor public spaces for short-term as well as long-term path prediction. Furthermore, the presented approach was also suitable for abnormality detection in private homes to support independent living in the senior population.

***Automatic human behavior understanding for Affective Computing.*** Human behaviors can be verbal, where words expressed by the voice represent the communication channel, and nonverbal, where the body language is the main communication channel. Psychological studies showed that the nonverbal channels are much more unconscious than the verbal channel and, therefore, nonverbal signals can be more valuable in revealing the true internal state and the true personality. Hence, during my PhD, with my promoters and colleagues, we focused on nonverbal behavioral cues such as body posture and expressive movements and their relation with personality attributes.

In this context, we presented a novel CNN-based framework for personality recognition (Figure 2). Our model analyzed the scene at multiple levels of granularity: firstly, we encoded spatio-temporal descriptors for each individual in the scene. Secondly, we extracted spatio-temporal descriptors from social groups, and thirdly, we encoded the global proxemics (interpersonal distances) between every individual in the scene. Experimental results demonstrated that modeling simultaneously Person-Context information significantly improves the individual features on personality recognition tasks. Finally, the Analysis of Person-Context interactions provides useful information for several real-world applications such as social role detection (i.e., leadership) and social event understanding and prediction.

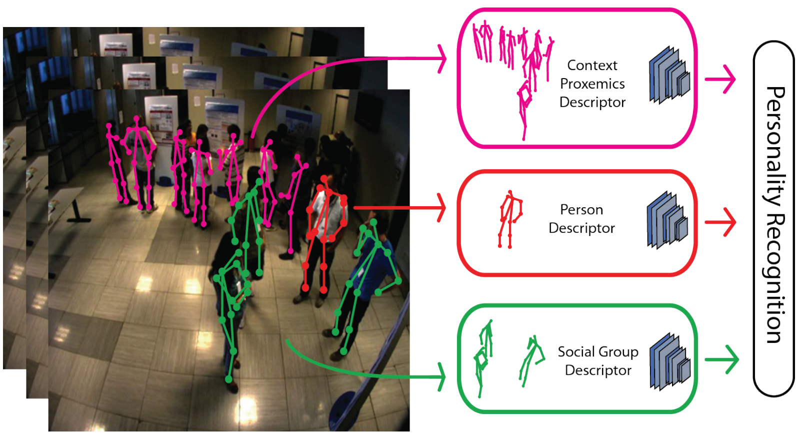



Figure 2. Individual motion descriptors as well as two Context descriptors are learned in a novel CNN framework for personality recognition. Individual motion descriptors (red color) indicate the engagement level of every person in the scene, social group descriptors (green color) indicate the engagement level of individuals in conversational groups, and finally, the context proxemics descriptors (purple color) indicate the global attitude of each individual with respect to the others.

AI SPOTLIGHT NEWS

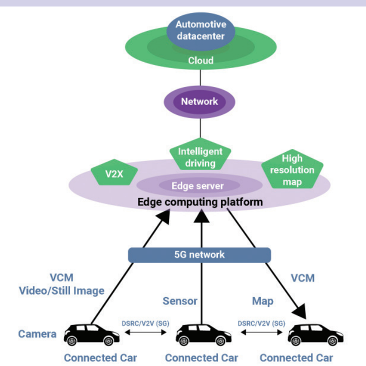
**Computer Vision News** has found great new stories, written somewhere else by somebody else. We share them with you, adding a short comment. **Enjoy!**

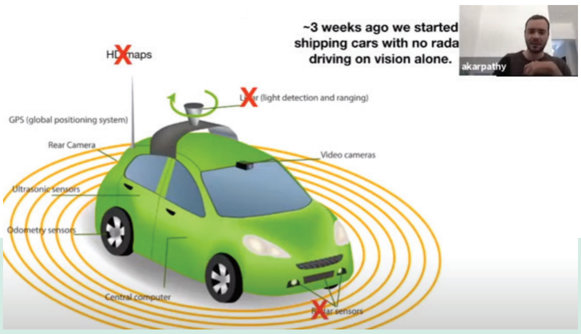
## CERN Sparks Podcasts Explore Artificial Intelligence

This brand-new podcast by **CERN** (the **European Organization for Nuclear Research**) brings together some of the delegates from the **Sparks! Serendipity Forum** in September. These are thought leaders among which computer scientists, neuroscientists, philosophers, physicists and psychologists. Why is it interesting? Because they are going to ask the same questions we are asking today about AI - and they will also try to answer. The first episode, **Brainy AI**, will collide the neurons of machine learning expert Stuart Russell and Tomaso Poggio, one of the founders of computational neuroscience. Read More

## Autonomous Vehicles: AI Must Accelerate

This very nice article is authored by a gentleman who is CEO of a company developing high-performance, low-power and low-cost Artificial Intelligence accelerators. Thus, it is not by mere chance that he recommends innovations in **edge computing** for autonomous vehicles to realize their full potential. Apparently, he has studied the problem from very close. You will read why it is highly impractical to send even a fraction of the **1 GB of data per second** (this is roughly what self-driving cars generate) for analysis to a centralized server because of the **processing bandwidth and latency**. Read More




## Tesla Places Big Bet on Vision-Only Self-Driving

More on autonomous vehicles. This is from **IEEE Spectrum**, a blog about the sensors, software, and systems that make cars smarter, more entertaining, and autonomous. Just like many of us, these gentlemen were at the **CVPR workshop on Autonomous Driving**, at which **Andrej Karpathy** gave a keynote and explained **Tesla**'s decision to drop radars. For those who didn't follow, he said that after training on more than 1.5 petabytes of video augmented with both radar data and human labeling, the **vision-only system** now significantly outperforms their previous tactic. The blog also quotes Marc Pollefeys!

## Robot-assisted Dressing - A New Algorithm by MIT CSAIL

One word now about robots. Robots are fun! They can also be useful, like helping people with limited mobility. In the past, algorithms have prevented robots from making any impact with humans for safety reasons. Which of course limited most of the interactions! Scientists from **MIT CSAIL** developed an algorithm to help a robot find efficient motion plans to ensure physical safety of its human counterpart. This lets the robot make non-harmful impact with a human to achieve its task. Have a look at the video! Of course, this idea could be used in areas other than dressing.


## How PepsiCo Uses AI to Create Products Consumers Don't Know They Want

**Pepsi: That's What I Like**! Apparently, the **PepsiCo** guys (maybe including Leo Messi himself) decided that people answering polls and manning consumer panels are too polite to provide credible input to unbiased market research. So they decided to bet on Artificial Intelligence instead: they use an AI tool called **Tastewise** which deploys algorithms to analyze massive quantities of food and menu data online (in the billions). Pepsi also builds on social predictions and has collected huge consumer datasets to let AI detect core targets and evaluate digital ads. **Read More**

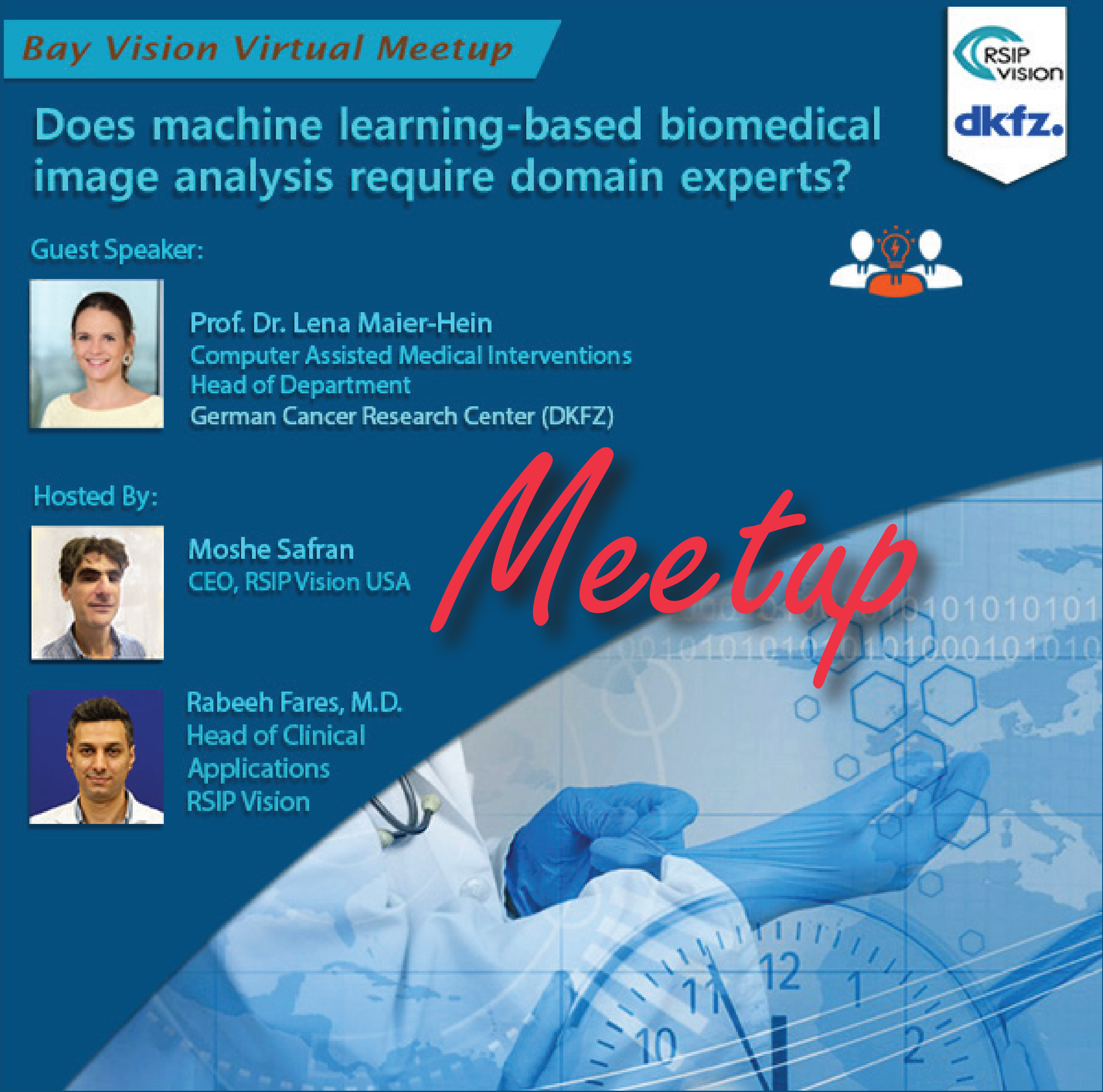
## MICCAI - the Medical Image Computing and Computer Assisted Intervention

This is not an article, but it is a very important community message. Did you register for **MICCAI 2021**? It will be held will be held from September 27 to October 1 as a virtual event. For the first time, the **ClinICCAI** event will be co-hosted. Three days of scientific content, including **oral presentations**, **poster sessions**, **workshops**, **tutorials**, and **challenges**. For the sixth consecutive year, RSIP Vision will publish the **MICCAI Daily**, the daily magazine of the MICCAI Conference. This is what **General Chair Caroline Essert** told us last year about MICCAI 2021. Click on the logo and be with us! **Read More**

MICCAI2021

## An awesome talk at the latest Bay Vision meetup with MICCAI Fellow Lena Maier-Hein.

## If you missed the meetup, don't miss the video!



# COMPUTER VISION EVENTS

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

**Ai4 2021**
*fully online*
**17-19 Aug**

**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**

## SUBSCRIBE!

Join thousands of AI professionals who receive Computer Vision News as soon as we publish it. You can also visit our archive to find new and old issues as well.

We hate SPAM and promise to keep your email address safe, always!

**FREE SUBSCRIPTION**

(click here, its free)

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

Fill the Subscription Form
- it takes less than 1 minute!

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!