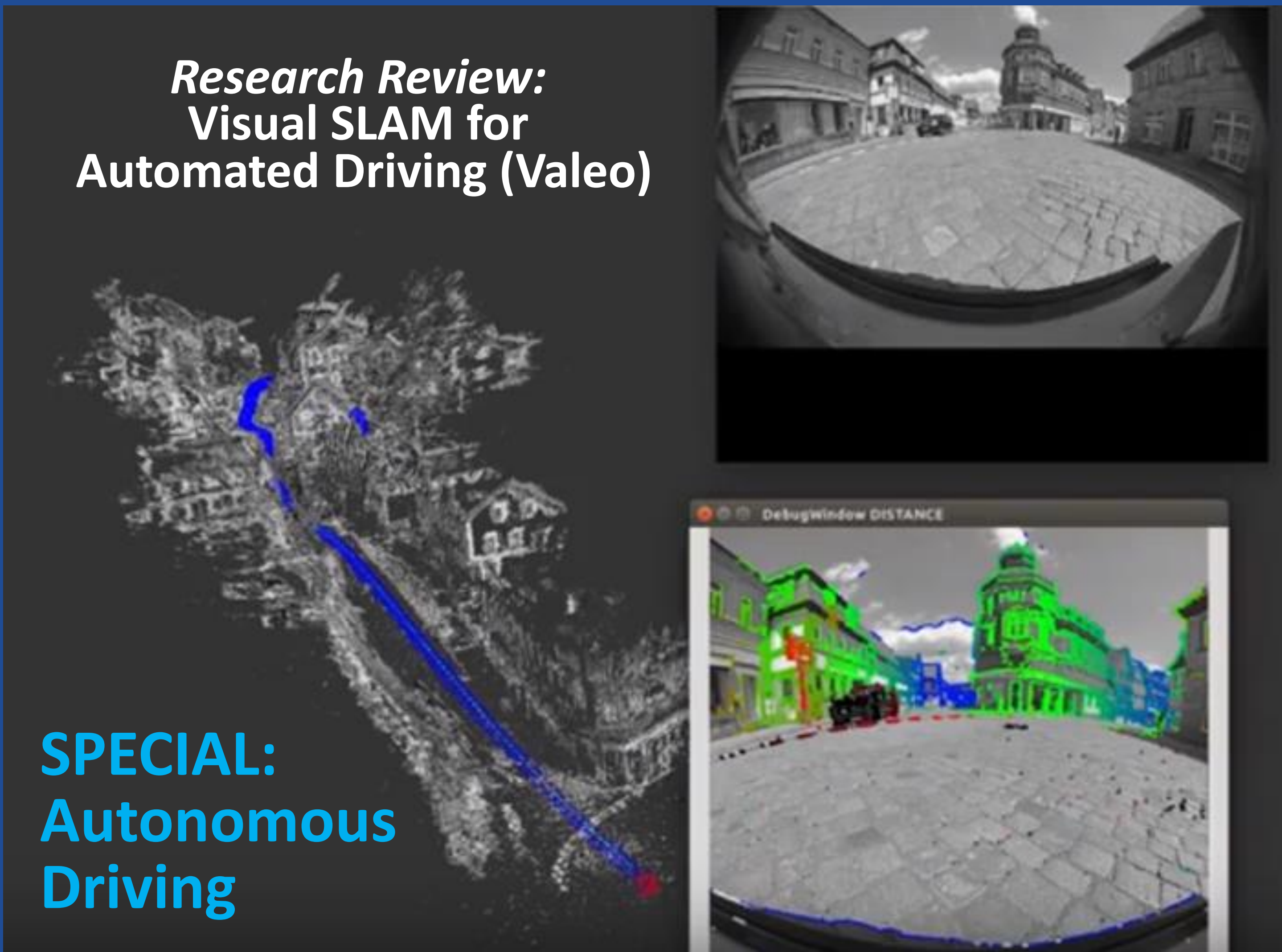
# Computer Vision News

The magazine of the algorithm community

A publication by RSIP Vision

## February 2019




*We Tried for You:*
OpenCV on Android (with codes!)

*Computer Vision Application:*
Video annotation

*Artificial Intelligence Spotlight News*

*Challenge:*
Medical Segmentation Decathlon

*Women in Computer Vision:*
Serena Young

*Project Management:*
Learning the Right Skills for ADAS Projects

*Upcoming Events*

*Computer Vision Project:*
Sensor Fusion Analytics for In-Cabin Monitoring

CES2019 Review

# Contents





BRATS_001 [2/4]
A
R
L A



We don't need no education.
We don't need no thought control.







if (item == mItemPreviewRGBA)
viewMode = VIEW_MODE_RGBA;
else if (item == mItemPreviewC
viewMode = VIEW_MODE_CANNY
else if (item == mItemPreviewSobel)
viewMode = VIEW_MODE_SOBEL;
else if (item == mItemPreviewZoom)
viewMode = VIEW_MODE_ZOOM;
else if (item == mItemPreviewPixelize)
viewMode = VIEW_MODE_PIXELIZE;
return true;



DISTRACTED

Did you subscribe to Computer Vision News? It's free, click here!

Dear reader,

**RSIP Vision** is very active in the field of **Autonomous Vehicles, having carried many projects in that area.** It is therefore natural for us to dedicate most articles in this issue of **Computer Vision News** to this subject. As usual, you will find articles on a wide range of topics. But just for this time, most of them will be in self-driving cars, ADAS, autonomous systems and the like. They include a report from **CES in Las Vegas**, the review of a project by RSIP Vision, a recent paper by **Valeo** on automated driving, our recommendation for training on project management, a gratifying testimonial by one of RSIP Vision's clients and more **AI for AV news** than ever.

You will also read on page 14 the wrap-up remarks of **Jorge Cardoso** concerning the **Medical Segmentation Decathlon**, a very popular challenge sponsored by **DeepMind**, **NVIDIA** and **RSIP Vision**. Results are quite surprising!

Finally, we know from readers' feedback that you love codes. One of the technical articles of this month will give you more code than you have ever found in our magazine: read it on page 24.

**Enjoy the reading!**

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

**Last Minute: the Bay Vision Meetup in Sunnyvale on the topic of Developments in Artificial Intelligence for Autonomous Vehicles - photos on pages 36-37!**

**Computer Vision News**

Editor:
**Ralph Anzarouth**

Engineering Editor:
**Assaf Spanier**

Publisher:
**RSIP Vision**

Contact us

Give us feedback

Free subscription

Read previous magazines



**Follow us:**

by Assaf Spanier


**CNNs are quickly becoming the go-to solution for vision tasks like object detection and semantic segmentation for Automated Driving**

Every month, Computer Vision News reviews a research paper from our field. This month we have chosen **Visual SLAM for Automated Driving: Exploring the Applications of Deep Learning**. We are indebted to the authors (**Stefan Milz, Georg Arbeiter, Christian Witt, Bassam Abdallah and Senthil Yogamani**), for allowing us to use their images. It is the Valeo's Visual SLAM pipeline for Autonomous Driving. Credit for the algorithm goes to the Valeo Kronach team. The paper is here.

Deep Learning has become the go-to solution for tasks like detection and recognition, and lately there have been advances in using CNNs for geometric tasks, depth assessment, optical flow prediction and motion segmentation. However, Visual SLAM (Self-Localization and Mapping) for Automated Driving systems is a field in which CNN methods are not yet sufficiently developed for production. In this article, the authors review the field and outline where and how Deep Learning can likely be applied to replace some of the stages of Visual SLAM.

The paper start by describing the challenges, building-blocks and stages of Visual SLAM. Next, the applications of Visual SLAM to Automated Driving is reviewed. And finally, opportunities for applying Deep Learning to improve upon the classic methods currently in use is discussed.



## The challenges of Visual SLAM

### Algorithm challenges:

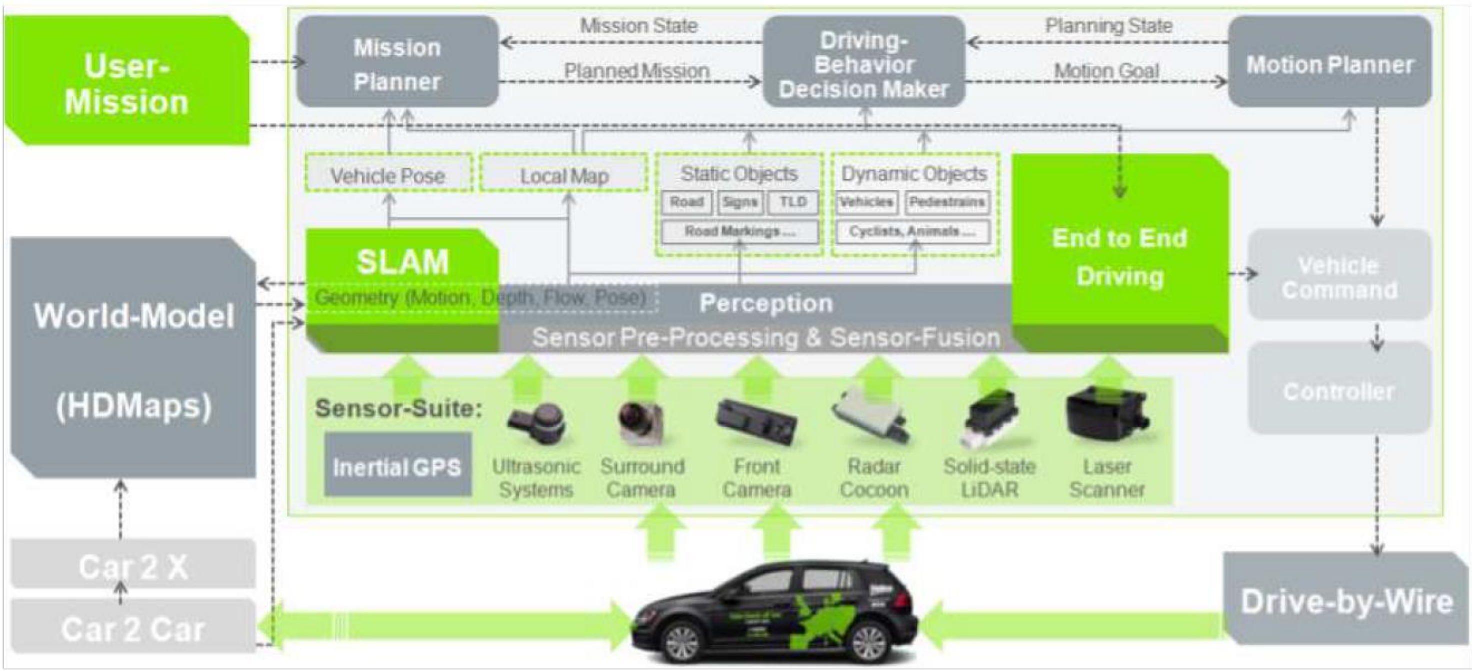
- Depth assessment: If the on-car camera rotates, depth-assessment from consecutive images is almost impossible.
- World Map and initialization: Most approaches to world-mapping start from a random initialization -- which creates many challenges as the rate of convergence is dependent on camera motion,
- Scale Ambiguity: A single camera SLAM can only estimate the World Map and Trajectory up to a certain scale, and requires a global reference to get a precise scale.
- Rolling shutter: This image acquisition method scans the frame horizontally or vertically, which results in not all parts of the input image being acquired simultaneously. This can create distortions. (as in the “bent” propeller blade in the image)

- Intelligent Loop-Closure Detection: Loop-Closure Detection recognizes the car has returned to a previously visited location. Current state of the art approaches use image processing to detect Loop-Closure, these are computationally intensive and very dependent on image quality and features.

## Application challenges:

- Map Self-Correction: The scene is constantly changing and needs to be constantly adjusted using mapping and localization methods.
- Local computation power: Map acquisition using the car's embedded processor without access to cloud infrastructure is challenging, especially so for CNN methods, which have heavy computational requirements.
- Unique signature for large-scale regions: Automated Driving maps are very large and similar objects are viewed with high frequency. This requires disambiguation, which is usually solved through semantic interpretation or global reference.

Automated Driving is a rapidly advancing field with a complex structure (see figure below). The application of Deep Learning to the field has also seen significant advances, but not yet achieved the performance required for production. Two main paradigms exist: 1) The two-stage approach -- which first constructs a semantic understanding (mapping) of the world, and then, at the second stage, makes driving decisions based on its semantic understanding. 2) The end-to-end approach -- which learns the driving decisions in a single stage.



# Visual SLAM approaches
## Fundamental Pipeline:

The term Visual SLAM includes all approaches to self-localization and mapping that take images as their input. The main difference between Visual SLAM and SLAMs based on other sensors is the need to produce depth data from camera images. There are two approaches to handling this challenge: 1) Depth estimation based on feature extraction, such as SIFT or ORB, and 2) Depth estimation from the image's pixels directly. Both approaches use the same pipeline.

**Basic Visual SLAM pipeline:**

Tracking → Mapping → Global Optimization → Relocalization

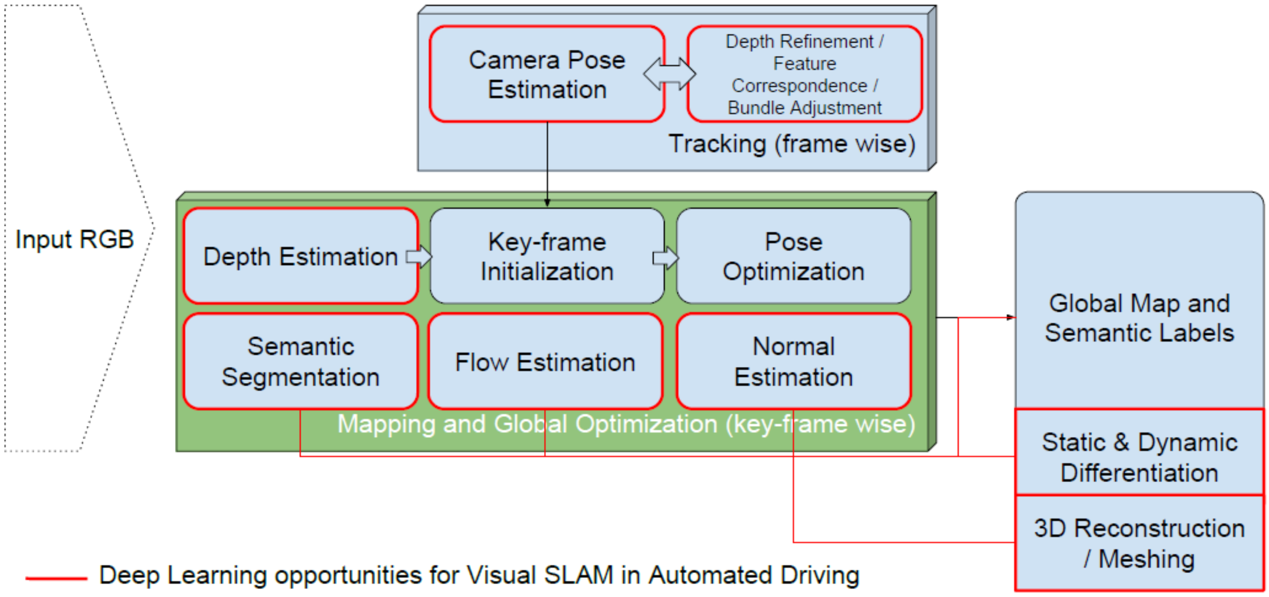
**Tracking** -- Consecutive-image tracking is used to deduce camera trajectory and depth information, usually using non-linear optimization. Most approaches use key frames as a basis for tracking, acquiring a new key frame whenever the tracking algorithm detects there is no longer sufficient overlap between the current camera image and the key frame.

**Mapping** -- The process of producing a map from the sensor data. At this stage, there is a significant difference between feature based and direct methods. The first create sparse feature maps, while the second creates semi-dense point maps. Some approaches use key frames that include depth and scale data.

**Global Optimization** -- The tracking process includes errors, and the Global Optimization step is used to correct the car's World Map. The optimization step is computationally intensive and therefore only invoked every couple of minutes. This step usually relies on recognizing previously seen locations or objects -- and so detects Loop-Closure by default. Some approaches also use 3D data for the optimization step.

**Relocalization** -- The process of placing a sensor at an unknown pose on the Map, and then trying to assess the pose. This is most often done by attempted matching of the actual sensor data with the Map. Many approaches use descriptive image features. The process of interpolating a sensor at an unknown pose on the Map, and then attempting to assess the pose

**... how Deep Learning methods and their specific capabilities can be used to replace individual elements / stages of the Visual SLAM pipeline [...] combining those solutions into an overall Deep Learning Visual SLAM**




## Visual SLAM approaches

Deep Learning is in the process of taking over the lead in the geometric vision field. We will now examine how Deep Learning methods and their specific capabilities can be used to replace individual elements / stages of the Visual SLAM pipeline, and consider the possibility of combining those solutions into an overall Deep Learning Visual SLAM.

### Localization and Depth Estimation:

Localization and depth estimation are crucial processes for Automated Driving. Three approaches exist: Supervised, Unsupervised and Semi-supervised. Supervised is the most common approach: This approach estimates depth using methods very similar to those of semantic segmentation, inspired by classification networks. The loss function is usually a continuous regression function, but some attempts have treated depth learning as a form of classification problem, giving different degrees of discreteness to depth learning. The unsupervised approach relies on running a projection function between different displays, for instance, by estimating movement between different views. Semi-i-supervised approaches combine elements of both.

The table below summarizes the results of three methods, each implementing a different approach, on the KITTI dataset.

| | Modality | RMSE (0-80m) |
|---|---|---|
| Eigen et al. | supervised | 7.156 |
| Godard et al. | unsupervised | 5.381 |
| Kuznietsov et al. | semi-supervised | **4.621** |

## Motion detection:

Motion detection (Optical Flow) is a challenging problem because the camera is in (fluctuating) motion simultaneously with the movement of objects the camera captures. With the detection of moving objects being of the utmost importance in Automated Driving to avoid accident casualties and fatalities of autonomous vehicles. This process uses motion cues to enable detection of a generic object, because (for the time being) there are no network to detect all the possible objects. Classic motion detection approaches focused on geometry-based methods, purely geometry-based approaches, however, suffer from many limitations. One such limitation is the problem of parallax -- the change in the perceived location of two points in relation to each-other, from the viewer's point of view, caused by a change in the viewer's location.

Leading work in this field includes: Fragkiadaki et al's proposal for segmenting moving objects using a separate proposal generation approach; however, this is computationally inefficient. Jain et al's motion fusion based method focused on a generic object. Tokmakov et al's use of a fully convolutional network with optical flow to estimate type of motion, which can work with either optical flow or concatenated image; note that concatenated image inputs won't benefit from the pre-trained weights, as those were acquired using RGB inputs only. Finally, Drayer et al's video segmentation work applied R-CNN methods followed by object segmentation using a spatio-temporal graph.



*Illustration of dense Motion detection (Optical Flow)*

## CNN Based Pipelines:

The nature of Deep Learning networks allows using a shared architecture to jointly learn different regression tasks - this is especially useful for critical real-time applications like Automated Driving. Here, too, there are supervised and unsupervised approaches. Tateno et al proposed a supervised approach with a

CNN jointly learning semantic segmentation and Deep Maps; this method uses CNN-based depth prediction with SLAM to overcome the traditional limitations of monocular reconstructions, combining geometric building blocks like depth estimation with semantic segmentation to improve upon the traditional Visual SLAM pipeline. Yin et al presented an unsupervised approach architecture for jointly deep learning optical flow and movement estimation, based on video all this without pre-segmentation, which currently achieved the top, state of the art, results for all tasks on the KITTI dataset. Their approach eliminates the need for annotation, based on the idea of using the strong interdependence of each geometric vision task (depth, motion and optical flow) to construct a joint loss function based only on consistency checks.

**In-map localization** is an essential task for SLAMs, with location described by DOF-6, which can be reconstructed using features from feature-based pipelines like SfM. CNN-based approaches map a single RGB image directly as captured by the camera. One such method is Kendall et al's PoseNet, whose memory requirement isn't linearly proportional to input data volume. PoseNet proved robust to motion blur, different lighting conditions and camera factors that tripped-up SIFT. Brachmann et al replaced direct regression of the 6-DOF camera pose with a sequence of less complex tasks: First, a network that learns to map limited image areas to 3D scene-space; Then, a differentiable RANSAC approach proposes a camera pose aligned with the previously-predicted scene coordinates. This approach utilizes geometrical constraints to improve performance, while still providing all the advantages that come with being an end-to-end trainable pipeline.



### Bundle Adjustment:

Before concluding, it is worth noting there is currently no CNN-based Bundle Adjustment solution to speak of. The last year saw some preliminary attempts - some approaches trying to model projection differently, some techniques attempting jointly learning parts of the Visual SLAM process using defined geometric part, but nothing concrete. This field still requires much thought and hard work.

## Conclusion

In practice, CNNs are quickly becoming the go-to solution for vision tasks like object detection and semantic segmentation for Automated Driving, they also show promising progress dealing with geometric vision algorithms, like depth estimation and optical flow. Nevertheless, progress on applying CNN-based approaches to Visual SLAM tasks is still slow. The paper provides an overview of Visual SLAM for Automated Driving and potential opportunities for using CNNs for different stages of the process.

# Learning the Right Skills for Autonomous Systems

**RSIP Vision's CEO Ron Soferman** has launched a series of lectures to provide a robust yet simple overview of how to ensure that computer vision projects respect goals, budget and deadlines. This month **Ilya Kovler** tells us about **Learning the Right Skills for Autonomous Systems Projects**. It's another tip by **RSIP Vision** for Project Management in Computer Vision.

## *"Learn the basics from those who are already engaged"*



**Artificial Intelligence** is impacting the automotive industry in more than one way. Of course, by developing **self-driving cars**, which might completely change our travel experience in only a few years. But also in many other ways: think at **driving assistance tools**, using sensors of one kind or another; **driver monitoring**, enabling to increase security by monitoring the driver's behavior; and many more driving and security features, some of which have to be invented yet. The project manager has to take into account the huge complexity of this new field and the colossal resources that it requires (for instance in terms of road testing and annotated data).

Many startups are entering the industry, hoping to find a successful way to implement AI in the automotive industry. Many of them employ capable and skilled workers, who still lack the needed training to do their job in this new field. They do not know the techniques and they do not know the terminology. The most probable result is a lot of energy and time going to waste. My recommendation to these companies, in particular to the project managers and to computer vision engineers, is to avoid reinventing the wheel: they should properly understand the importance of making the first steps in this field only after **learning the basics from those who are already engaged in it**. I can bring a very appropriate example from our successful experience at **RSIP Vision**: the company has sent very early a group of computer vision engineers (which included myself) to the **Self-Driving Car Engineer Nanodegree program** offered by **Udacity**, one of the best programs that teaches the right skills and prepares for a fruitful work in the field of autonomous systems.

This program is a practical course, built in cooperation with **Waymo (Google)**, **Uber**, **Mercedes** and others. It enables to learn the theoretical background, as well as working hands on practical tools and on writing the algorithms, using the deep learning neural networks in Python and C++. This results in the engineer's readiness to work very efficiently in any of the many projects which are being carried on in the AI for autonomous vehicles arena. Deep learning, sensor fusion, LiDAR data and other important components of this work will be precious in a very large range of applications and products.

## *I recommend … to invest in this kind of training to acquire the skills needed for the success of the project*

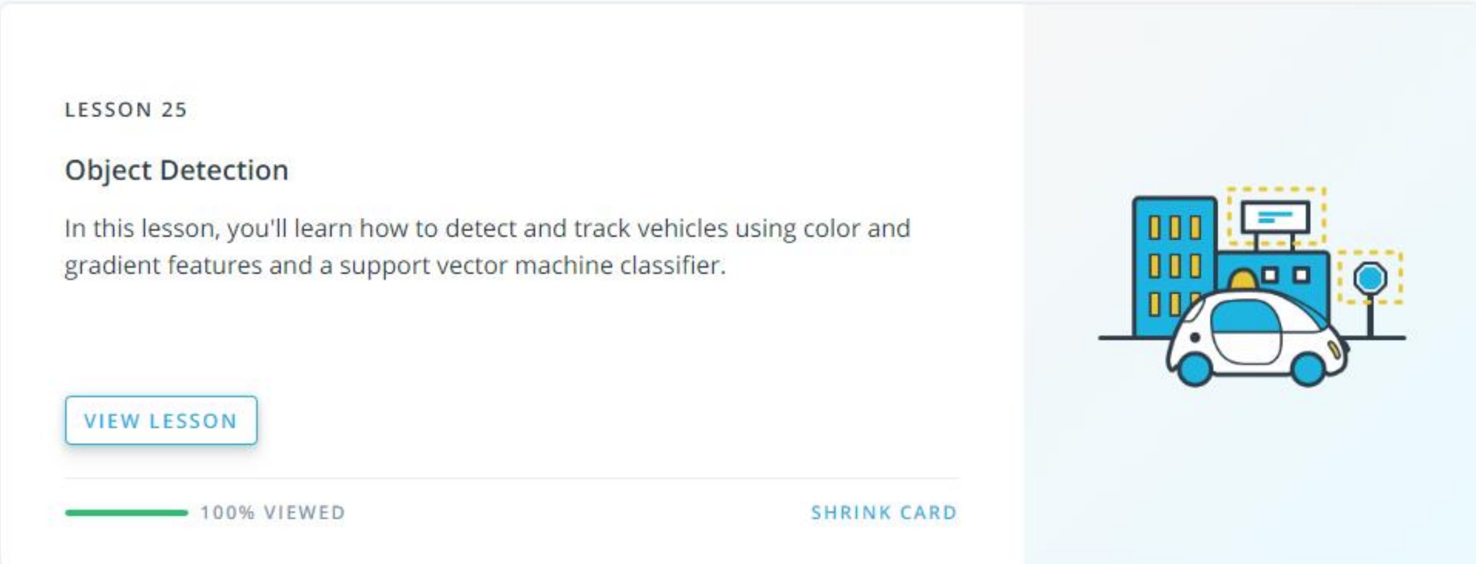
For example: a modern car includes today tens of sensors. In order to develop a specific application, it might be recommended to use sensor fusion, a very powerful and effective tool to combine data coming from multiple sensors. If you want to learn more about it, **read our article on page 12**.

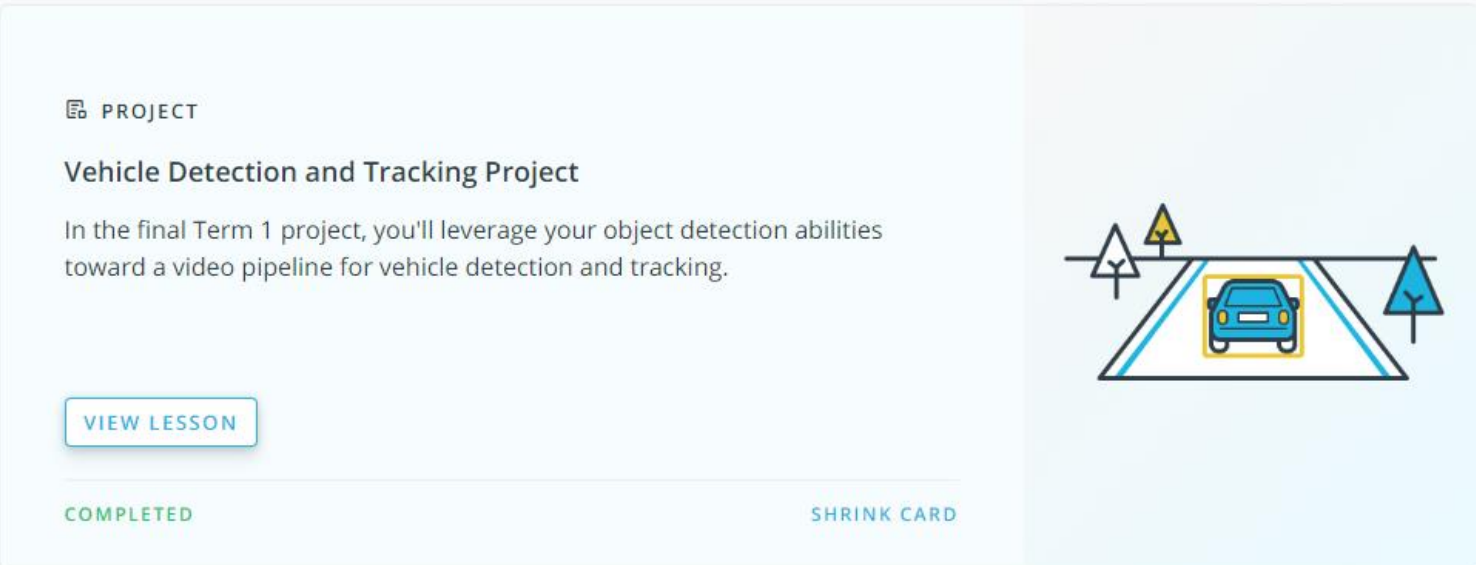
My personal experience from this course is very positive, since it enabled our engineers to acquire a very wide range of useful tools: from classic computer vision to deep learning, through road signs detection and lane detection. Each student collects data by training on the Udacity simulator and after that it uses that to teach the car to drive autonomously. And this is only the first part of a very varied and interesting course. I recommend all project managers in this field to invest in this kind of training to acquire the skills needed for the success of the project.

[Sensor Fusion Analytics](#)

[More articles on Project Management](#)





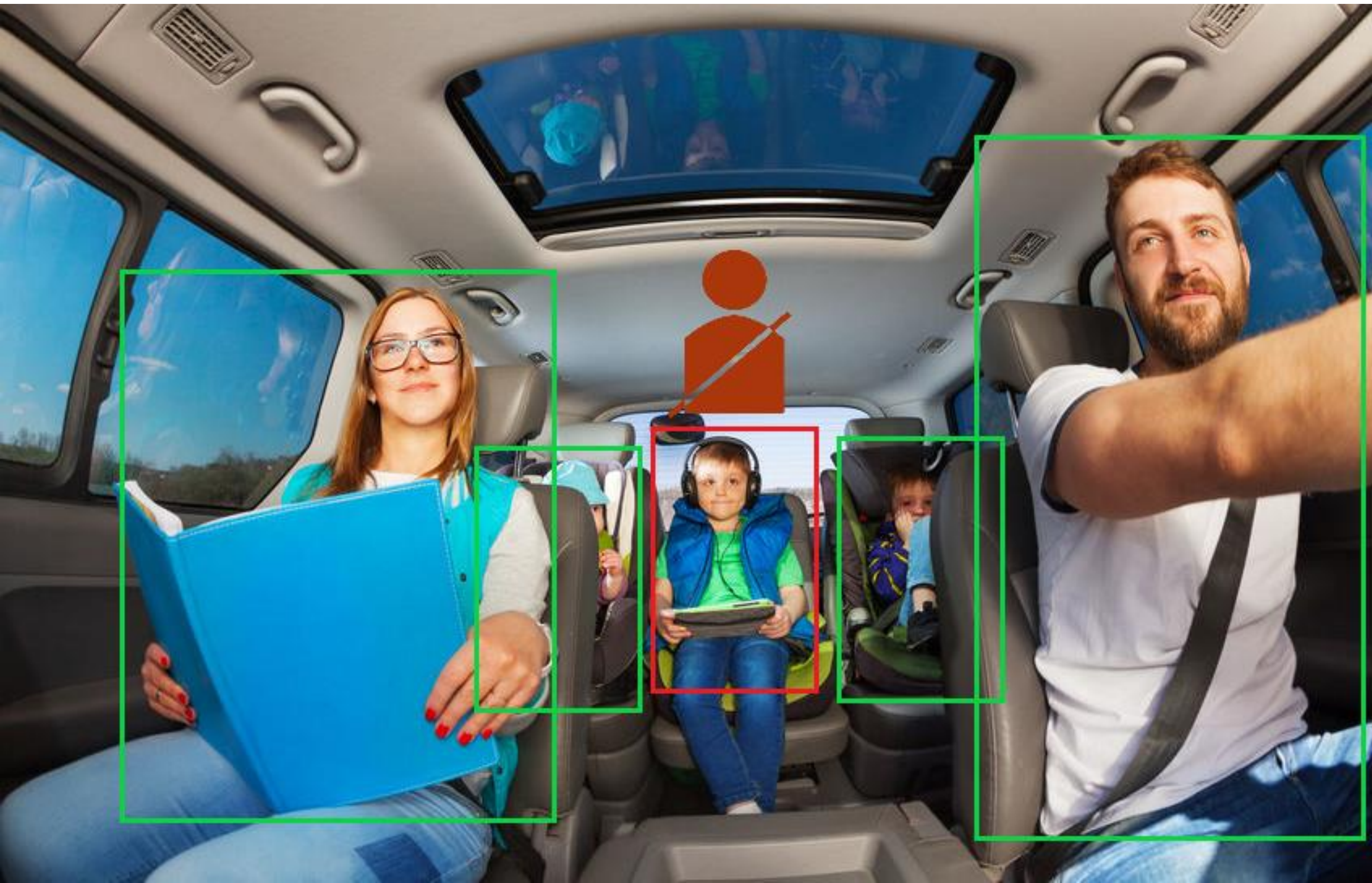



*Guardian Optical Technologies and RSIP Vision cooperated on the next-generation sensor fusion analytics for in-cabin monitoring systems*



Every month, **Computer Vision News** reviews a successful project. Our main purpose is to show how diverse image processing techniques contribute to solving technical challenges and real world constraints. This month we review a precision agriculture project by **RSIP Vision**: **sensor fusion analytics for in-cabin monitoring systems**. This work is the result of a cooperation between **RSIP Vision** and **Guardian Optical Technologies**, a market leader for in-cabin monitoring systems, winner of major accelerator competitions. Guardian Optical Technologies is dedicated to enabling **"passenger aware" cars** with cutting-edge sensor technology that makes cars **safer and more convenient**. Just one sensor combined with advanced 2D, 3D, and motion analysis protects drivers and passengers by constantly scanning and tracking occupants and objects anywhere in the vehicle. These technologies work with a car's seatbelts and airbags to sound immediate alerts.

**RSIP Vision**'s and **Guardian Optical Technologies**' engineers have recently developed a new layer of AI algorithms enabling occupancy, weight and age detection in **one all-inclusive sensor**. Guardian's sensor capabilities range from depth and vibrations to visual data. Cooperation with RSIP Vision's engineering team gave birth to a new level of **AI algorithms** that leverage the



detection and classification capabilities of the system.

This hybrid solution combines **Artificial Intelligence deep learning algorithms** with advanced heuristic and non-neural networks to yield cutting edge classification in a very challenging environment, given by the variability of passengers' looks and poses, the occlusions and the use of a wide-angle camera with a strong distortion effect.

An additional challenge is given by the extreme variation in illumination and other spatial information, which is solved by pioneering sensor fusion technology. The Guardian sensor produces multiple channels of information which the system processes simultaneously by fusing those different data to support the complex decision.

The obtained performance is almost real-time, since the system provides an all-inclusive report about the status of the car's cabin in less than one second, regardless of all challenging conditions.

*"In the last 6 months, I have had the pleasure of working with RSIP Vision on the development of computer vision algorithms"* says **Yuval Gronau, Vice-President of R&D at Guardian Optical Technologies**; he adds that *"there were many technical challenges in the development of the algorithms and RSIP Vision came up with original solutions to overcome them. Thanks to this work, we have succeeded to significantly improve the performance of the algorithms."*

**RSIP Vision's CEO Ron Soferman** concludes saying that *"the combination of Guardian's advanced hardware with leading algorithms development is the ideal formula for the next generation's systems."*



## Feedback of the Month

In the last 6 months, I have had the pleasure of working with **RSIP Vision** on the development of **computer vision algorithms**.

There were many technical challenges in the development of the algorithms and **RSIP Vision** came up with original solutions to overcome them.

Thanks to this work, we have succeeded to **significantly improve the performance of the algorithms**.

**Yuval Gronau**
**Vice-President of R&D - Guardian Optical Technologies**

**Jorge Cardoso, Senior Lecturer at King's College London, served in the organizing team with the Medical Segmentation Decathlon (MSD) challenge, sponsored by DeepMind, NVIDIA and RSIP Vision. The challenge tested the generalizability of machine learning algorithms when applied to 10 different semantic segmentation tasks with the aim of developing an algorithm or learning system that can solve each task, separately, without human interaction. Michela Antonelli told us about it before the start of the decathlon. The challenge concluded recently and Jorge Cardoso shares with us very interesting results.**

**Jorge Cardoso:** The challenge ended up happening during MICCAI. There were two phases. The first phase was seven tasks where people tried their algorithms when tasks were available. The second phase was three new tasks where the researchers had not seen the tasks before.

During the event itself, what we did is a relatively long presentation which took almost two hours going through all the aspects of the challenge. That included what the datasets have and why the dataset is important in terms of how free it is from a licensing point-of-view that allows commercial companies to build on it… aspects related to the data itself.

Everything that we've done is on the website. You have the code that was used to validate; a document explaining the methodology for the statistical analysis; and you have the results of that analysis.

After all of that, what we did is we announced the winners of the challenge for phase 1 and phase 2. When we were reporting the results, we found that there were three large blocks of results. There was a very clear winner, which was surprising. I was not expecting a method to pretty much outperform all others on pretty much every metric. There was a group of two or three methods which formed a very good cluster of very well-performing methods, deserving honorable mentions. Then there was pretty much everyone else, with quite poor performances. This was on the training tasks, the seven tasks that were part of phase 1.

When we did the full statistical analysis of the results over multiple tasks over multiple regions over multiple metrics using non-parametric ranking. We ended up concluding that there was one method that was better than the others. The issue is how well did the method extrapolate to the other three tasks. The three new tasks were a little bit special in a way. One of them was relatively simple. It was large organs with a lot of training data. There was another one which was a little bit different because you were segmenting vessel and tumors. Previously, there were no tasks which had to do segmentation of any tumor structures. So that was slightly different.

The most impressive was the third task which was on rectal primaries, cancer in the colon. That is truly hard to see. If you open the images yourself, it’s pretty hard to see the areas that need to be segmented. Probably two were leaders in this area. We wanted to push the algorithms to see how they were able to perform. The surprising outcome from all of this was that the results seen in phase 1 translated to phase 2 without almost any changes whatsoever. The scale of the data that we had was very hard for people to fix, even on the training sets, because the data was so large. The data was pretty much the same in terms of proportions between different methods, the training and test sets, even though there was no organization of parameters, obviously, because that was a whole other task. That was exactly the purpose of the challenge. Again, we saw the same behavior.

The team from NVIDIA forfeited their award because NVIDIA was sponsoring the event: they decided to reward their own Titan V GPU card, giving it away to someone in the public. We did a really funny thing where we put a sticker underneath the chairs, and the person who found the sticker won the GPU. We have a picture of pretty much everyone in the room trying to search under their chairs.

**From a scientific point-of-view, what do we learn from the challenge?**

That’s the interesting part. We’re going through quite an expansive exploration of the data itself. The aim of the challenge was to ask the question: can algorithms learn to segment images without human interaction? Can we just develop a learner that can automatically solve a very large variety of tasks without a human? That question is very important because, in the medical domain, we have a huge number of tasks to solve. All of them have relatively small sometimes large datasets.

In the medical domain, there are many, many tasks that we could possibly solve. We at least need an image segmentation context. You can segment any organ or different types of tumors. It’s very hard to develop an algorithm that is independent and different for every single task. It would be ideal if there was an algorithm that performs very well on many different tasks at the same time. This was really the primary aim of the challenge.

Can an algorithm, without any human interaction, learn many different things even when the tasks are extremely different from each other? The tasks could be anything from segmenting tumors from CT, segmenting the hippocampus from the MR... pretty much every single semantic segmentation task you could possibly imagine with very large and very small objects, with a large number of datasets and very small numbers of datasets. The ten tasks were purposely selected to encompass very different types of problems that you see in the medical imaging domain.

What we found in the end is that yes, indeed, the better algorithms, the ones that won the challenge, were able to perform very well in a large amount of problems. They were actually able to achieve performance very close to state-of-the-art on a large number of problems.

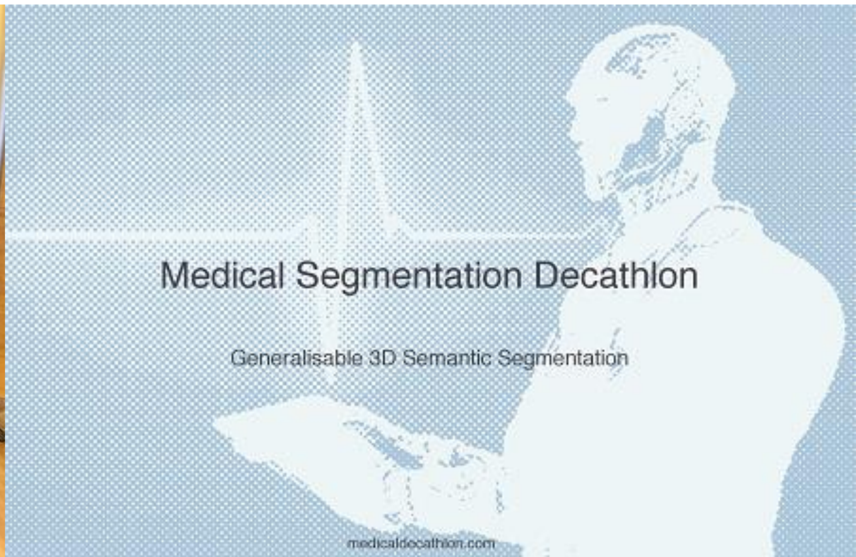
The methods that were proposed were general purpose. They were not task-specific. They were not designed specifically to solve a task. They were designed to solve many possible tasks with very general purpose. They should perform pretty much as well as algorithms tailored and designed for a specific task.

Even more interesting than that, the winning algorithm, the algorithm from DKFZ, got state-of-the-art performance in two tasks. Even though it was not developed on purpose for those two tasks, they performed better than the best algorithm until that day on that specific task. They had state-of-the-art performance just behind the submissions of that specific year from NVIDIA which was quite phenomenal.

They could demonstrate that you can have algorithms that are general, but also perform extremely well. It also demonstrates that you can get state-of-the-art performance on certain tasks even though those algorithms are general purpose.

Something that needs to be discussed and we need to understand is when we talk about generalizing these algorithms, we talk about algorithms as general learning systems, but they



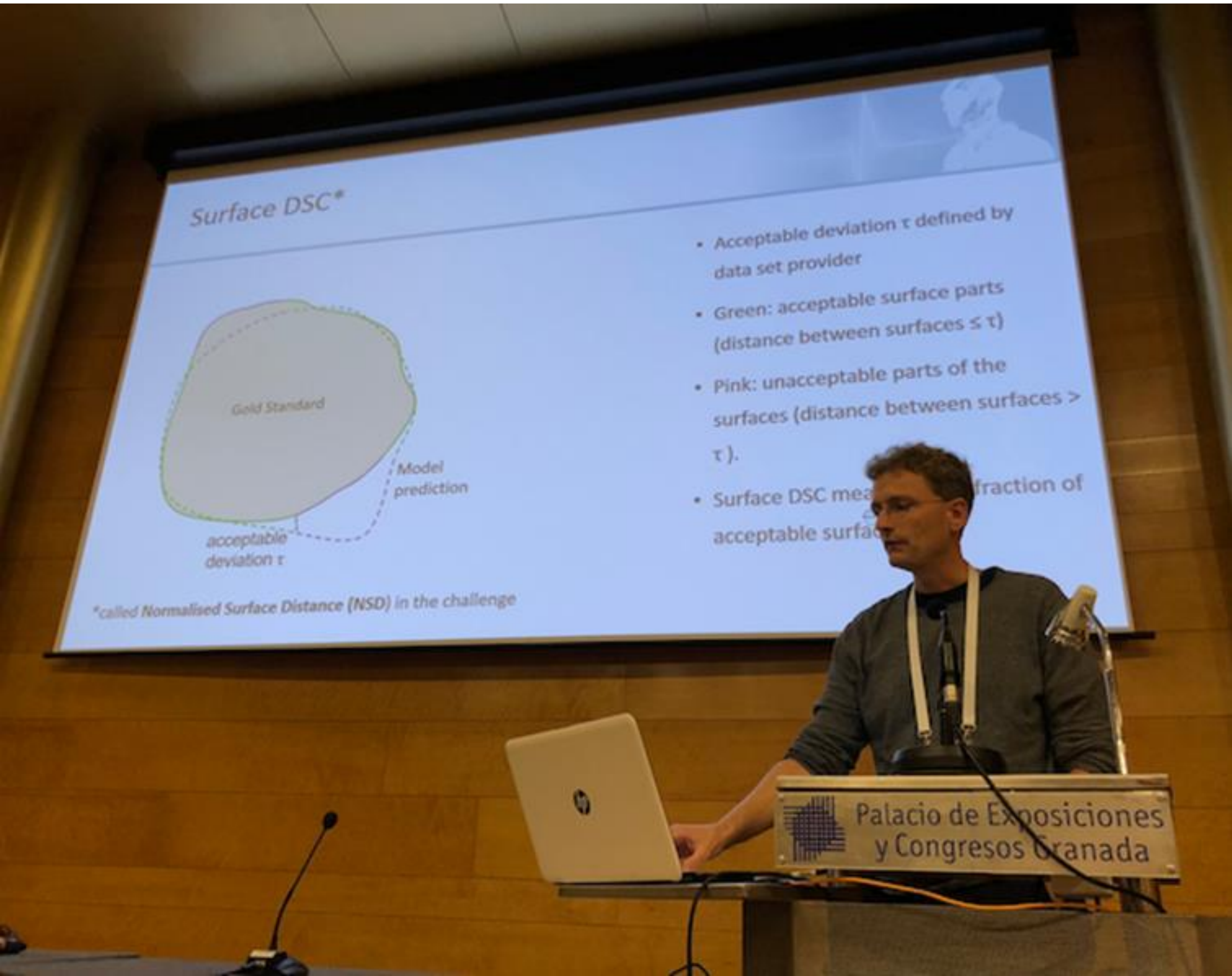


are optimizing a specific metric. Obviously, the metrics we use to assess are not necessarily what we care about for clinical purposes.

There is an interesting part that we don't know, and that's something that I would like to explore in a future challenge. We really hope that the next challenge will happen very soon, maybe not this year, but maybe the year after. We still haven't conversed on that. The idea would be how would we redesign the challenge in a way that is fairer where most teams have access to the same computational power, where the models themselves can be deployed and trained offsite. We have, for example, containerization of algorithms so that multiple teams can actually submit their models to be trained onsite in the same computational infrastructure.

Also, how can we design a challenge that is more clinically relevant? Something where the outcome of the challenge are metrics that are relevant from the point-of-view of the task that is being asked from a clinical point of view. How can we be sure that it's a statistical methodology that can put together multiple tasks with different metrics? Every task is going to have a metric which is appropriate for what you are trying to solve.





Women in Science

# Serena Yeung

**Serena Yeung recently received her PhD from Stanford University, where she was advised by Fei-Fei Li and Arnold Milstein. During the current academic year, she is visiting Harvard University as a TEAM postdoctoral fellow. In Fall 2019, Serena will start her position as Assistant Professor of Biomedical Data Science and, by courtesy, of Electrical Engineering at Stanford University.** [More interviews with women scientists]

**Serena, you are a postdoc at Harvard in between two Stanford stints.**

Yes, I did my PhD at Stanford, and I'm going to be returning as a faculty in the summer. In between, I decided that I wanted to take a year to go somewhere else and gain some different perspectives.

**And what better place than Harvard! When did you finish your PhD?**

In June 2018.

**Tell us something about Stanford that we don't know.**

Someone told me once that Stanford is the second biggest university campus in the world. I believe the largest was in Moscow. It's a really huge place. I include some of the hills around the main campus as well. I really love it as a place where you have a lot of space to walk, bike, see different views, and think about things.

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

My work is focused primarily at the intersection of AI and healthcare. I want

*"My work is focused primarily at the intersection of AI and healthcare"*

*photo: Timothy Archibald*

**"I think dissemination and communication are also an important part of research!"**

to develop AI algorithms that can enable better healthcare. At the same time, I also want to study and to formulate important challenges in real world applications like healthcare that can then help motivate useful research directions in AI. One of my main focuses has been creating AI assisted healthcare spaces where we have computer vision that's infused in physical spaces like hospitals in order to understand what is happening in an environment like a hospital at every moment in time and use that to assist in delivering quality care. Activity recognition is a huge part of that. It's something that I have worked on quite a lot in computer vision. In general, there are many challenges that you realize you need to solve when you want to build something like a smart hospital in the real world.

What will be the major benefit of your work in health care? I understand that it might help nurses have more time to treat patients or improve the quality of the care.

I think there are a number of things. At the first level, if you just have something like an automated system that can recognize and record the actual care activities and what's happening at every moment in time, all of a sudden you have a lot of data to actually understand what's happening, study correlations with outcomes, and develop best care practices. This is a kind of data that hasn't previously been available. I think that's actually at the first level of what I'm really excited that we can use this data and computer vision for. At the second level, as you said, once we're able to recognize what is happening, we can use that feedback to assist clinicians and nurses in real time. That means being able to document automatically what is happening, allowing clinicians and nurses to focus more on taking care of patients. We can also help them know what steps have happened so far, whether these are the intended steps of different types of care protocols.

*photo: Dan Taylor/Heisenberg Media*

We can make sure that they are actually delivering the intended care. Computer vision can also be used for the continuous monitoring of patients. When you don't have a healthcare provider that can be by a patient's side all the time, you can help keep monitoring for patient status. Another thing you can imagine is this being a smart space where you have an environment that is reactive and changing, predicting ahead of time what might be a doctor or nurse's next step to help make changes in the environment, to anticipate that, and assist them in real time.

**Where do you stand now in putting this into practice?**

We've had some early prototypes that we've set up at Stanford Children's Hospital as well as at Intermountain Healthcare in Utah. At these places, we've equipped pilot units, for example, one unit at each hospital with sensors capturing visual data and developed prototype pilot computer vision algorithm for these. We've had some exciting early results for these, looking at when anyone is performing hand hygiene, which is important for preventing infections that are acquired in the hospital. We've also looked at continuously monitoring for patient mobilization activities, healthcare workers working to move patients in and out of bed and mobilizing them a little bit when they're in ICUs. These mobilization activities are really important for preventing ICU acquired weakness from lying too long in bed without moving.

**Can you explain that?**

ICU acquired weakness is a condition you acquire from lying too long in bed without moving. It can cause complications or prolong your stay in the hospital.

**How have the hospital personnel reacted?**

One thing I've really noticed has been the interest from the medical personnel we've worked with in investigating and studying these new types of technologies that can help provide better patient care.

***"It shows how much the nurses and clinicians really care about their patients. They are enthusiastic to be the guinea pigs trying out and shaping the future of this technology."***

It’s been really evident that the first priority of doctors and nurses is trying to look at anything they can do to improve the care that they are able to give their patients. This sort of technology is very new. There’s no precedent for it. There are obviously a lot of questions that need to be answered, ranging from the technology, but also to the hospital and policy level about how this kind of technology will be used. It shows how much the nurses and clinicians really care about their patients. They are enthusiastic to be the guinea pigs trying out and shaping the future of this technology.

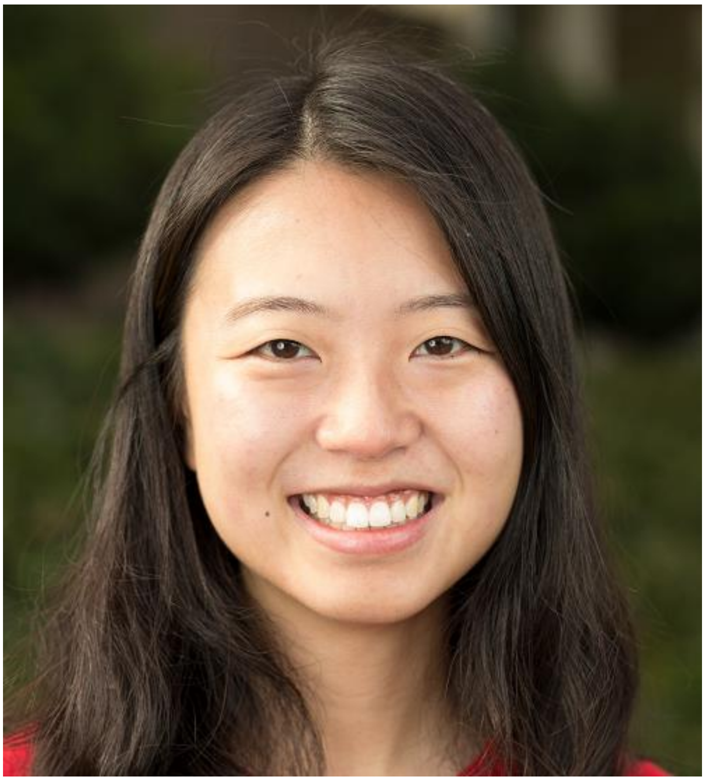
**We didn’t mention yet that in Stanford you were supervised by Fei Fei Li. Of all the positive aspects of working with Fei Fei, is there anyone that you would like to point out in particular?**

I think one thing that has stood out to me about Fei Fei is how much she genuinely cares about each of her students and their future. She really believes in her students and sees the best in each of them. She’s the kind of person that I’ve seen be extremely supportive and always believing in you. As a PhD student, sometimes you can go through tougher periods where you doubt yourself a little bit. I think having an advisor like Fei Fei has been amazing.

Jessica Sieren once told me: ***“If you are not having a crisis during your PhD, it means you are not doing it right!”***

[*laughs*] I agree with that! It’s a rite of passage!

***“Fei Fei really taught me to believe in myself and not to get discouraged”***



**Of course, being at Stanford and working with Fei Fei has given you a lot of visibility at conferences and on the internet. Do you like that exposure, or sometimes do you think it's too much?**

To be honest, as a PhD student, my goal was both figuring out myself as a researcher and getting to the end of my PhD. In some sense, I feel like I kept my head down and was just very focused on these goals. I'm not sure that I was thinking about that aspect of things too much. At the same time, I think that as a researcher it's important to, not only come up with and demonstrate your research ideas, but to also share them with others. I think dissemination and communication are also an important part of research. It is extremely important to have an advisor like Fei Fei who has been very helpful with that.

**You are going to teach at Stanford. Is it important for you to teach the new generation what you have learned?**

Over the course of my PhD, I started to have some experience working with younger students. This is also something Fei Fei encouraged me to do. I really enjoyed it. I think it's really fun to work with people. A lot of the students are super enthusiastic and eager to learn. They come from very diverse backgrounds and perspectives. It's been really fun to just be able to help shape a little of their growth. That's something I look forward to doing more of once I transition to being a full-time faculty at Stanford.

**Can you already see what your generation of teachers should take from the previous one?**

For me, it's realizing that every person is different. Everyone comes from different backgrounds, strengths, interests, and perspectives. One important aspect of being a teacher is to recognize what is unique and special about each student and try and nurture that.

**Like Fei Fei does!**

Yes, exactly!

**Would you like to tell us one specific thing you've learned that stands out? A lesson for life?**

A lesson for life? Hmm… I'm not sure that this is necessarily about being a teacher. In general, Fei Fei really taught me to believe in myself and not to get discouraged. Believe in what you are passionate about even if it's something that may not come easy. In the short term, it may seem like you are struggling. My PhD work, especially with AI in healthcare and trying to implement new technology in hospitals was not something where you could see a lot of quick and early gratification. She was helpful in encouraging me to be patient and keep on working toward what you believe in. I guess this means finding gratification about yourself and your own conviction that you did something right.

RSIP VISION

Global Leader in Computer Vision & Deep Learning

# Women in Computer Vision

*by Ralph Anzarouth*

Women in Computer Vision (also called Women in Science) is a series of interviews conducted by Ralph Anzarouth. New interviews are regularly published on all RSIP Vision's publications: Computer Vision News and the Daily magazines (CVPR Daily, MICCAI Daily and many more).

Find now on the project page the direct links to almost 100 interviews... **at the click of a button**

Leadership

Mentoring

Competence

Confidence

Community

“The only way to succeed is to really start believing in yourself!"

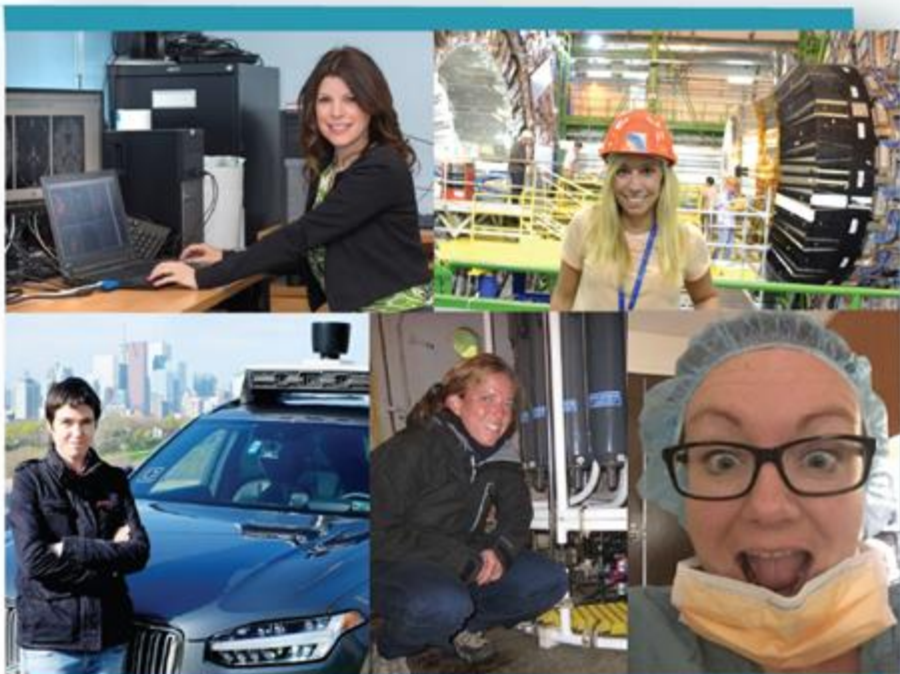
Michela Paganini

"Most of all, you have to believe that you can do it!"

Laura Leal-Taixé

"It may look like a long list of names, but behind each name there is a fascinating world in which we were let in."

Ralph Anzarouth



Did you miss an interview? No worries, you can find them all in the **Women Scientist section of RSIP Vision’s website** by Assaf Spanier


**... a little app that takes the video captured by the phone's camera as input and processes it according to the option selected by the user and displays it online as a continuous video.**

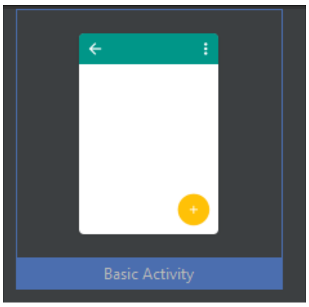
This month, **We Tried for You** to install and use OpenCV on Android app – and we will tell you how. Based on the demo included with OpenCV, we'll create a little app that takes the video captured by the phone's camera as input and processes it according to the option selected by the user (we will show how to define 4 options: pixelize, zoom, sobel or canny) and displays it online as a continuous video.

Let's get started. We assume you already have **Android Studio** installed; if not, you can install it from here.

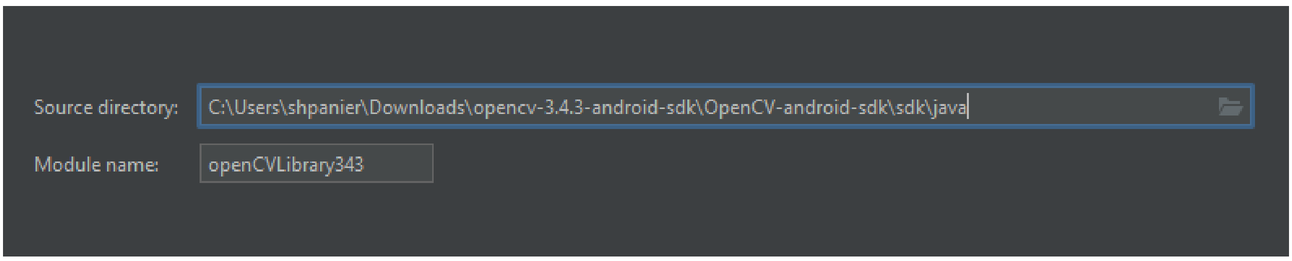
First, you need to **download Android's SDK from** here (pick your favorite Android version).

We recommend the OpenCV-3.4.3 SDK. Let's say we download it to the folder "C:\Users\shpanier\Downloads\" and from there unzip the file into the folder "opencv-3.4.3-android-sdk"

Now we go to Android Studio to write our app: Create a new project, by selecting File → New → New Project and picking a Basic Activity project. We'll call it OpenCVwCamera and select Api28 (the latest version).



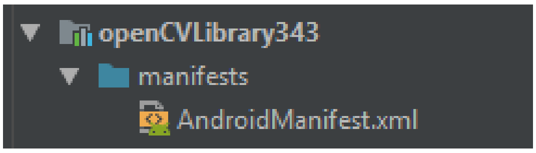

Next, import the OpenCV SDK into the project by going to File→ New→ import Module and picking OpenCV-android-sdk\sdk\java under the location we opened the file -- a folder titled opencv-3.4.3-android-sdk.




Focus on

Once you reach this window, click `next` twice.

You need to define the following additional settings: Open the file "AndroidManifest.xml" you'll find under `manifest` under `openCVLibrary343\` as seen in the figure below.



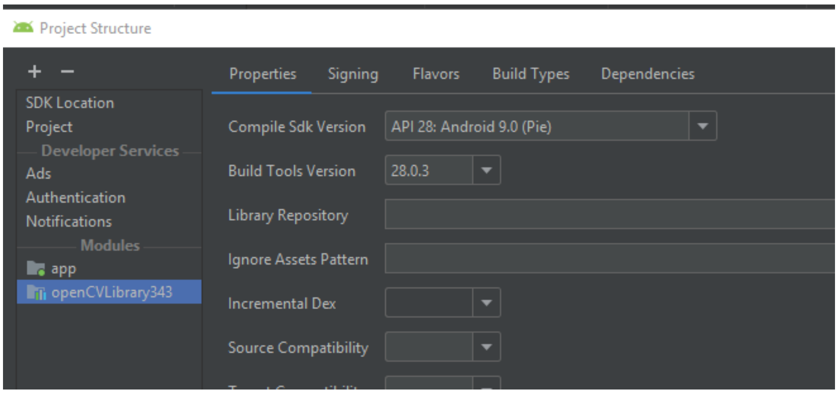

And in AndroidManifest.xml please comment out the line "

<!-- <uses-sdk android:minSdkVersion="8" android:targetSdkVersion="21" /> →

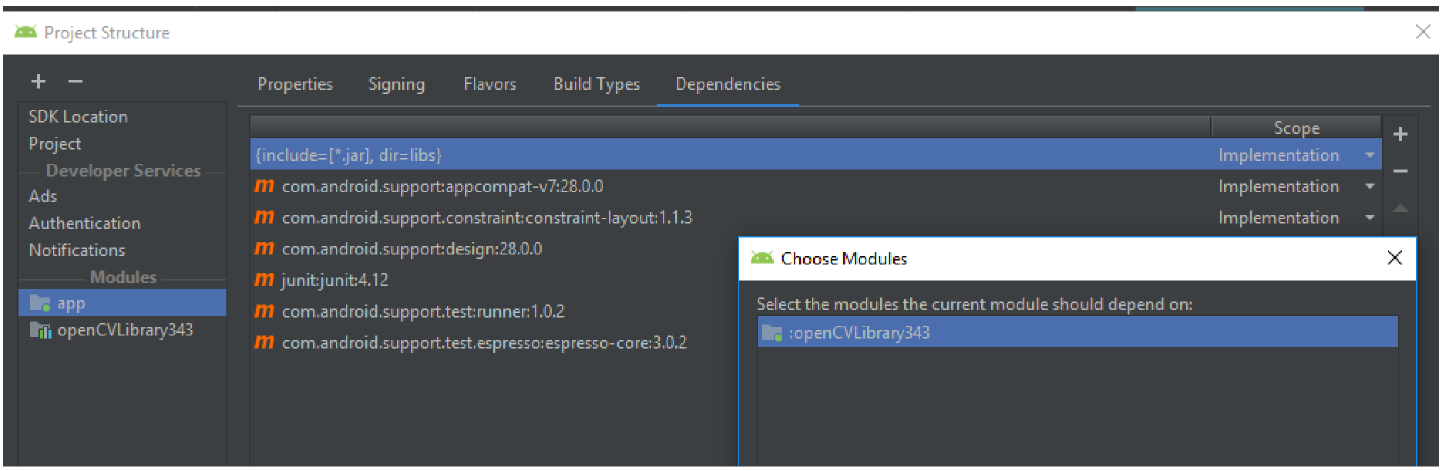
We are doing so since the SDK version will be taken from the Manifest of the project itself (we don't need it twice).

The next step is to import the module, which is done as follows:

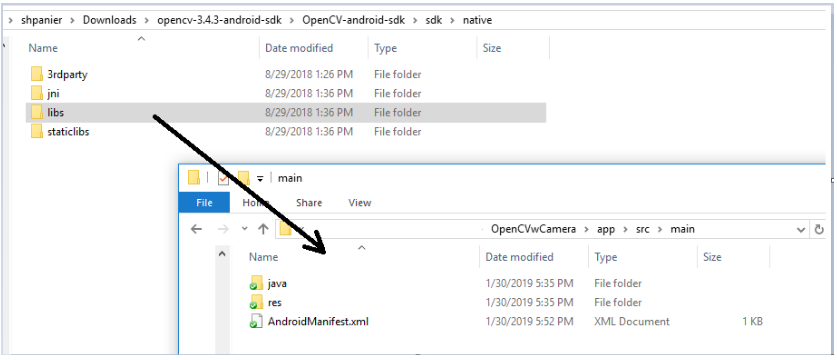
Go to: File → "Project Structure" , on the left menu select OpenCVLibrary343 and under the "Properties" tab choose "API 28" in the compile SDK Version field, as demonstrated in the figure below:



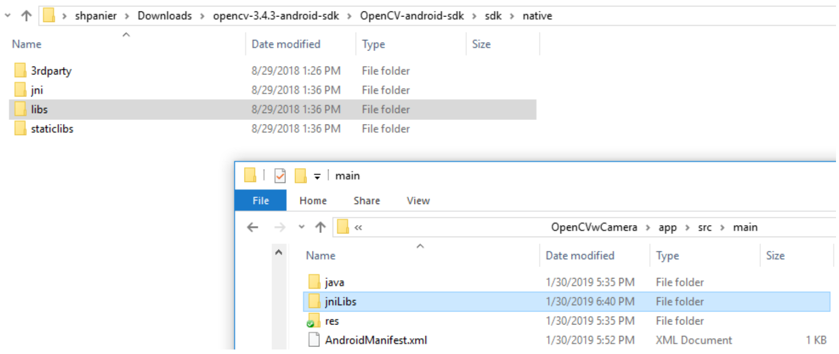

Next go to "app" and under the "Dependencies" tab click on the "+" button. When "choose Modules" window pops up, select openCVLibrary343 and click "ok" twice.



The next step is to copy the compiled OpenCV shared library into your project. For that you need to copy the “libs” folder from the opencv-3.4.3-android-sdk to your project under (OpenCVwCamera\app\src\main) as demonstrated in the following figure:




And rename it to “jniLibs” as you can see here:






Since we want to have an app with some image processing capabilities based on the camera input, we will need to add camera permissions; this is done by adding the following lines into AndroidManifest.xml under app->manifests

```
<uses-permission android:name="android.permission.CAMERA"/>
<uses-feature android:name="android.hardware.camera" android:required="false"/>
<uses-feature android:name="android.hardware.camera.autofocus" android:required="false"/>
<uses-feature android:name="android.hardware.camera.front" android:required="false"/>
<uses-feature android:name="android.hardware.camera.front.autofocus" android:required="false"/>
```

Next, let’s add JavaCameraView to content_main.xml; here are the lines to add for this:

```
<org.opencv.android.JavaCameraView
android:layout_width="fill_parent"
android:layout_height="fill_parent"
```

You've produced the code for a small app based on the demo included with OpenCV, taking the cellular camera image as input and processing it online, according to the option selected.

Paste the code bellow into "MainActivity.java". This is the main code of the app:

```
package com.example.opencvwcamera;

import android.os.Bundle;
import android.support.v7.app.AppCompatActivity;
import android.support.v7.widget.Toolbar;
import android.util.Log;
import android.view.Menu;
import android.view.MenuItem;
import android.view.WindowManager;

import org.opencv.android.BaseLoaderCallback;
import org.opencv.android.CameraBridgeViewBase;
import org.opencv.android.LoaderCallbackInterface;
import org.opencv.android.OpenCVLoader;
import org.opencv.core.Mat;

import org.opencv.android.CameraBridgeViewBase.CvCameraViewFrame;
import org.opencv.core.Core;
import org.opencv.core.CvType;
import org.opencv.core.Point;
import org.opencv.core.Scalar;
import org.opencv.core.Size;
import org.opencv.imgproc.Imgproc;

public class MainActivity extends AppCompatActivity implements
CameraBridgeViewBase.CvCameraViewListener2 {
   private CameraBridgeViewBase mOpenCvCameraView;
   private static final String  TAG            = "OCVSample::Activity";
   private Mat mIntermediateMat;

   public static final int      VIEW_MODE_RGBA     = 0;
   public static final int      VIEW_MODE_CANNY    = 2;
   public static final int      VIEW_MODE_SOBEL    = 4;
   public static final int      VIEW_MODE_ZOOM     = 5;
   public static final int      VIEW_MODE_PIXELIZE = 6;

   private MenuItem             mItemPreviewRGBA;
   private MenuItem             mItemPreviewCanny;
   private MenuItem             mItemPreviewSobel;
   private MenuItem             mItemPreviewZoom;
   private MenuItem             mItemPreviewPixelize;
   public static int           viewMode = VIEW_MODE_RGBA;
   private Size                 mSize0;

   private BaseLoaderCallback mLoaderCallback = new BaseLoaderCallback(this) {
       @Override
       public void onManagerConnected(int status) {
           switch (status) {
               case LoaderCallbackInterface.SUCCESS:
```

Focus on

```
                {
                    mOpenCvCameraView.enableView();
                } break;
                default:
                {
                    super.onManagerConnected(status);
                } break;
            }
        }
    };

    @Override
    public void onResume()
    {
        super.onResume();
        if (!OpenCVLoader.initDebug()) {
            OpenCVLoader.initAsync(OpenCVLoader.OPENCV_VERSION_3_0_0,
this, mLoaderCallback);
        } else {

mLoaderCallback.onManagerConnected(LoaderCallbackInterface.SUCCESS);
        }
    }

    @Override
    public boolean onCreateOptionsMenu(Menu menu) {
        getMenuInflater().inflate(R.menu.menu_main, menu);
        mItemPreviewRGBA  = menu.add("Preview RGBA");
        mItemPreviewCanny = menu.add("Canny");
        mItemPreviewSobel = menu.add("Sobel");
        mItemPreviewZoom  = menu.add("Zoom");
        mItemPreviewPixelize  = menu.add("Pixelize");
        return true;
    }

    @Override
    public boolean onOptionsItemSelected(MenuItem item) {
        Log.i(TAG, "called onOptionsItemSelected; selected item: " + item);
        if (item == mItemPreviewRGBA)
            viewMode = VIEW_MODE_RGBA;
        else if (item == mItemPreviewCanny)
            viewMode = VIEW_MODE_CANNY;
        else if (item == mItemPreviewSobel)
            viewMode = VIEW_MODE_SOBEL;
         else if (item == mItemPreviewZoom)
            viewMode = VIEW_MODE_ZOOM;
        else if (item == mItemPreviewPixelize)
            viewMode = VIEW_MODE_PIXELIZE;
        return true;
    }
    public void onDestroy() {
        super.onDestroy();
        if (mOpenCvCameraView != null)
            mOpenCvCameraView.disableView();
    }
   @Override
   public void onCreate(Bundle savedInstanceState) {
```

```
        Log.i(TAG, "called onCreate");
        super.onCreate(savedInstanceState);
       getWindow().addFlags(WindowManager.LayoutParams.FLAG_KEEP_SCREEN_ON);
        setContentView(R.layout.activity_main);

        mOpenCvCameraView = (CameraBridgeViewBase)
findViewById(R.id.image_manipulations_activity_surface_view);
        mOpenCvCameraView.setVisibility(CameraBridgeViewBase.VISIBLE);
        mOpenCvCameraView.setCvCameraViewListener(this);

        Toolbar toolbar = findViewById(R.id.toolbar);
        setSupportActionBar(toolbar);
    }

    public Mat onCameraFrame(CvCameraViewFrame inputFrame) {
        Mat rgba = inputFrame.rgba();
        Size sizeRgba = rgba.size();
        Mat rgbaInnerWindow;

        int rows = (int) sizeRgba.height;
        int cols = (int) sizeRgba.width;

        int left = cols / 8;
        int top = rows / 8;
        int width = cols * 3 / 4;
        int height = rows * 3 / 4;

        switch (MainActivity.viewMode) {
            case MainActivity.VIEW_MODE_RGBA:
                break;
            case MainActivity.VIEW_MODE_CANNY:
              rgbaInnerWindow = rgba.submat(top, top + height, left, left + width);
                Imgproc.Canny(rgbaInnerWindow, mIntermediateMat, 80, 90);
                Imgproc.cvtColor(mIntermediateMat, rgbaInnerWindow,
Imgproc.COLOR_GRAY2BGRA, 4);
                rgbaInnerWindow.release();
               break;
            case MainActivity.VIEW_MODE_SOBEL:
                Mat gray = inputFrame.gray();
                Mat grayInnerWindow = gray.submat(top, top + height, left,
left + width);
                rgbaInnerWindow = rgba.submat(top, top + height, left,
left + width);
                Imgproc.Sobel(grayInnerWindow, mIntermediateMat,
CvType.CV_8U, 1, 1);
                Core.convertScaleAbs(mIntermediateMat, mIntermediateMat,
10, 0);
                Imgproc.cvtColor(mIntermediateMat, rgbaInnerWindow,
Imgproc.COLOR_GRAY2BGRA, 4);
                grayInnerWindow.release();
                rgbaInnerWindow.release();
                break;
            case MainActivity.VIEW_MODE_ZOOM:
                Mat zoomCorner = rgba.submat(0, rows / 2 - rows / 10, 0,
cols / 2 - cols / 10);
       Mat mZoomWindow = rgba.submat(rows / 2 - 9 * rows / 100, rows / 2 +
9 * rows / 100, cols / 2 - 9 * cols / 100, cols / 2 + 9 * cols / 100);
```

```
                    Imgproc.resize(mZoomWindow, zoomCorner, zoomCorner.size(),
0, 0, Imgproc.INTER_LINEAR_EXACT);
                    Size wsize = mZoomWindow.size();
                    Imgproc.rectangle(mZoomWindow, new Point(1, 1), new
Point(wsize.width - 2, wsize.height - 2), new Scalar(255, 0, 0, 255), 2);
                    zoomCorner.release();
                    mZoomWindow.release();
                    break;
                case MainActivity.VIEW_MODE_PIXELIZE:
                  rgbaInnerWindow = rgba.submat(top, top + height, left, left + width);
                    Imgproc.Canny(rgbaInnerWindow, mIntermediateMat, 80, 90);
                  rgbaInnerWindow.setTo(new Scalar(0, 0, 0, 255), mIntermediateMat);
                  Core.convertScaleAbs(rgbaInnerWindow, mIntermediateMat, 1./16, 0);
                  Core.convertScaleAbs(mIntermediateMat, rgbaInnerWindow, 16, 0);
                    rgbaInnerWindow.release();
                    break;
            }
            return rgba;
        }

        public void onCameraViewStopped() {
            if (mIntermediateMat != null)
                mIntermediateMat.release();
            mIntermediateMat = null;
        }

        public void onCameraViewStarted(int width, int height) {
            mIntermediateMat = new Mat();
        }
}
```

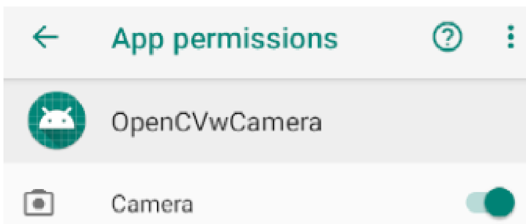


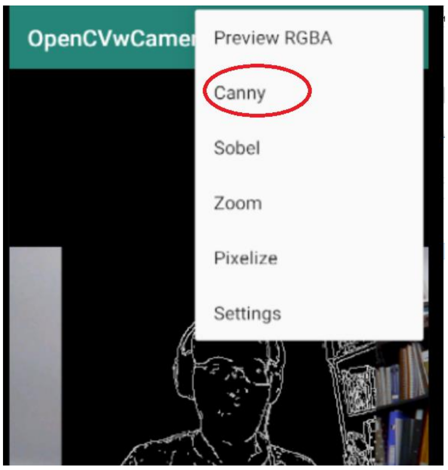
The code's main functions are:

- `onCreateOptionsMenu` and `onOptionsItemSelected` are the functions responsible for creating the menu and changing menu options according to user selection.
- Is responsible for receiving camera images as input and processing them according to the option selected and displaying the processed image on screen.

*see main switch of the function.

Don't forget to give the app the permissions it needs to function.




If all goes as planned, the application will look as follows: a menu on the right for selecting the image processing option the user wants the app to apply. And the output of the selected-option processed at the center of the screen. In the figure you see me after Canny Edge detection "processing", **don't I look more handsome post-processing**?

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

**Will NVIDIA's Automotive Platform Widen its Moat vs Intel:**
**NVIDIA** introduced at recent **CES2019** (see pages 34-35) **Drive AutoPilot**, a new reference platform that lets automakers add automated driving features to their vehicles, thus improving current advanced driver-assistance systems (**ADAS**) with features that bridge the gap between semi-autonomous vehicles and fully autonomous ones. **The Fool** thinks this is a direct move against **Intel's MobilEye**, which provides ADAS platforms for over 90% of the world's automakers. Read More...

**11 Big Autonomous Vehicle & AI Stories:**
It's the month of **Autonomous Driving**, so here are more stories from somewhere else. Big or not, I don't know: certainly interesting. Enjoy!

**We Don't Need no Education! We Don't Need no... :**
Yes, like in the song: we are talking about Computer Vision for classroom "Engagement Detection". It is said that engagement is raised as a major concern by educators, especially (but not exclusively) in online learning contexts. So why not tracking each pupil's attention, both in the classroom and in the student's home? Is 1984 here? Who can keep singin', after this...



**AI predictions for 2019 from Yann LeCun, Andrew Ng ... :**
...and more... We already know many of Yann LeCun's plans from our exclusive interview. Here is what other AI celebs ( like Cloudera's Hilary Mason and Accenture's Rumman Chowdhury) have to tell us about AI in 2019. Read Here...

**This Cool Tool Deletes the Background from Images:**
Looks like a very cool tool, so you don't have to use Photoshop. It is called Remove.bg and it can clear out the background of an image in seconds. It uses AI trained on images of people, so as of today it only works on photos of humans. Read More...

**Deep Learning:** the confluence of Big Data, Big Models, Big Compute. Interesting...
2018 proved that computer vision is the most powerful manifestation of AI. Nice Read
This bot judges **how much you smile** during your job interview. Another 1984ish app!
Self-driving car startup Zoox gets permit to transport passengers in California. More...

**Guy Kohen is the cofounder of Clay Sciences, a company that develops web-based tools for annotating videos, images and text.**

*“We added a tool that tracks objects automatically using image processing algorithms like optical flow”*

**You have a new platform for annotation. What’s special about it?**

At Clay Sciences, we built a platform for annotating data. We specialize in video annotations, annotating directly on the video and not on single frames. First of all, one of the things that we utilize is the continuous aspect of video. Annotation on videos does not require annotating each and every frame. You can, instead, annotate single keyframes and interpolate between them. This is something that happens on our platform and saves a lot of time annotating long videos.

One other thing we added to our platform is automatic tracking. We added a tool that tracks objects automatically using image processing algorithms like optical flow. This saves you even more time than annotating only keyframes. We have one video with 10 seconds - that is, 300 frames - of cars driving on Brooklyn bridge, where one annotation on the first frame gets updated automatically to the rest of the 300 frames, so that's all the work the annotators have to do. We developed tools that will annotate bounding boxes, lines, nonrigid, and rigid polygons. We have tools to annotate all of them. Our tools are very easy to use and very easy to configure. We have an interface that enables the manager to configure classes and attributes in a text box. We don’t annotate only labels. We also annotate attributes that change over time, like whether the car is braking, what’s the color of the person’s purse, their names, things like that. Both classes and attributes configured very easily through JSON format. We also manage the process of distributing the work with multiple workers. You just need to register all of the users, and everybody gets access to all of the annotation tools. Managers also get reviewing tools including contribution reports.

**These features are particularly precious in the autonomous vehicle industry. Can you tell us more about it?**

Sure, right now all of the autonomous vehicle companies record the road using videos, but still, they need to convert the videos to images because annotating videos is harder and tend to get messy. With the right tools annotating videos can get even easier. We want to move forward. We want to be more progressive instead of annotating single frames. As I said before, there are some advantages in video, temporal features or even just the continuous movement of the object. You can take advantage of it by annotating videos, and you cannot do it in images. Also, if you want to track some specific object, a long video, even if it goes outside of the frame and comes back, you can do it only in video, and not in images. We specialize in video and we are the only ones that have automatic tracking in video, which is the biggest state-of-the-art right now.

**...Which (again) is very precious for self-driving cars.**

Yes, basically because it saves time annotating video. Somebody (I think from Drive.ai) said once, that "For every one hour driven, it takes approximately 800 human hours to label it." This is very time-consuming.

**What are your next steps for this platform?**

We are very satisfied with the automatic tracking, but this is not enough. We can do better. We can plug in machine learning models into our platform that will train and infer on new videos while annotating the videos. Think of active learning methods. You start by annotating very short segments of video, send results immediately to your model, train it, and on the next segment you can start from the model's inferences. As you proceed with this method, the model gets better, and the data annotation process gets easier.

**Ron Soferman's comment:**

*Not only is this a nice platform for machine learning, but also a system that can use machine learning to guess the annotation and let the user fix it if necessary. The new system can save a lot of time and use the manual work only for non-trivial cases.*

*Let's consider an interface where you start the annotation, then after a few hundred images, the system uses transfer learning over known networks like ImageNet. Most features in natural images are already trained by ImageNet so that the marginal work of retraining should yield a good classifier in most cases.*



**by Shmulik Shpiro**

**Shmulik Shpiro is RSIP Vision's EVP Global Business Development & Strategy. Shmulik visited this year's CES in Las Vegas. Here is what he found out regarding the autonomous driving solutions and the Advanced driver-assistance systems (ADAS) systems field as a whole.**

CES features innovators and breakthrough in consumer technologies: self-driving cars and ADAS systems represented an especially hot topic at this year's event.

During the show, many vendors presented the newest cars and innovative concepts set to come out over the next few years, starting with autonomous cars. It has been an amazing show! It's definitely the place to see the latest updates and newest innovations in the automotive industry.

This year, we had the pleasure of being a part of the event: several of our Automotive clients participated and presented new technologies that we took part in developing. Our teams worked very hard for the last three months, side by side with the vendors internal R&D teams, to make sure everything was in place for the show. I am happy to say, after actually seeing the demos, that it worked beautifully. It's exciting to see the knowledge and know-how that we have in our company actually translating into products with very positive feedbacks




**One of the impressive booths at CES2019. Read the full article about sensor fusion at Guardian Optical Technologies on page 12.**

from the market. Hopefully, in the coming months, we'll see a lot of deals closing for our customers.

Even without having a booth at CES, we feel like one of the innovators at the show, working together with our clients and vendors as a team, pushing technology and the Automotive industry forward. Actually, that's the way we work with our clients. We become an extended arm of their R&D team, working with them to achieve their goals, meet the roadmap and develop innovative solutions.

This year, some exhibitors took things a step up, showcasing different kinds of applications and user experiences inside the car, which raise the question: if you don't need to drive, what will you actually do when driven in the car?

Many vendors deal with different kinds of media, entertainment and information systems that allow the passenger to do a lot more when "not driving". We saw technology that allows people to operate these systems without actually touching any buttons. The passenger uses vision, hand poses, and facial movements to easily control the systems.

*"The autonomous revolution is going to happen. Although it will take a bit more time than originally anticipated, the new mobility revolution is coming"*

Besides the large manufacturers, I noticed a decent number of smaller companies attending this year, mainly from the United States and China. They produce technologies and applications that just two or three years ago seemed very far away.

For example, companies manufacture smaller and more efficient LiDAR solutions, which were very complicated to build only a few years back. Now, instead of just taking images and video, different vision-based solutions (like smart dash cameras) actually embed AI technologies and add ADAS or ADAS-like technologies into relatively simple and inexpensive units.

Autonomous driving could even revolutionize public transportation. I have seen representatives from many municipalities exploring the booths. After all, self-driving vehicles have the potential to both change the transportation services to the public and provide additional added value to municipalities, such as: monitoring road conditions, smart alerting systems and increasing public safety.

Some manufacturers, like VOLVO, decided to push forward the Level 2+ ADAS systems, taking the very current, advanced technologies and embedding them as is into the market. They are starting to sell them to test the market, which has matured in the last few years, showing that one mature system can give a stable solution for today's modern driving challenges, while waiting to the full autonomous car revolution to happen. We also see, a wide market acceptance that one sensor will never be enough and there is a need to fuse certain kinds of sensors and combine technologies together, to achieve the end goal.

It is clear that the autonomous revolution is going to happen. Although it will take a bit more time than originally anticipated, the new mobility revolution is coming. For a practical example, **see the project on page 12**.

**Among many other initiatives, RSIP Vision sponsors the Bay Vision Meetup. We invite great speakers in the Silicon Valley. The January 31 meeting held in Sunnyvale focussed on the topic of Developments in Artificial Intelligence for Autonomous Vehicles. Below, Ron Soferman (CEO of RSIP Vision) says a few words of welcome to more than one hundred Meetup participants.**


**On top, the first speaker: Abhijit Thatte, the new VP of Software at [AEye](), who spoke about "Perception of 3D point clouds using deep learning". Below, the second speaker: Modar Alaoui, Founder and CEO at [Eyeris](), who spoke about In-Vehicle Scene Understanding AI Enabled with Automotive-grade AI Chip. Thank you both for being so kindly available for our audience!**




**BIOS 2019 - SPIE Photonics West**
S.Francisco, CA — Feb 2-7 — Website and Registration

**SPIE Medical Imaging**
S.Diego, CA — Feb 16-21 — Website and Registration

**ICPRAM Intl. Conf. Pattern Recognition Applications & Methods**
Prague, Czech Republic — Feb 19-21 — Website and Registration

**BMVA: British Machine Vision Association - Deep Learning in 3D**
London, UK — Feb 20 — Website and Registration

**VISIGRAPP - Computer Vision Imaging and Computer Graphics**
Prague, Czech Republic — Feb 25-27 — Website and Registration

**TRI-CON - International Molecular Medicine Tri-Conference**
S.Francisco, CA — Mar 10-15 — Website and Registration — MEET US!

**IAPR Computational Color Imaging Workshop**
Chiba, Japan — Mar 27-29 — Website and Registration

**ISBI 2019 - IEEE International Symposium on Biomedical Imaging**
Venice, Italy — Apr 8-11 — Website and Registration

**European Symposium - Artificial NN, Comput. Intelligence and ML**
Bruges, Belgium — Apr 24-26 — Website and Registration

**AI & Big Data Expo Global**
London, UK — Apr 25-26 — Website and Registration

Did we forget an event?
Tell us: editor@ComputerVision.News

## FREE SUBSCRIPTION

Dear reader,

Do you enjoy reading Computer Vision News? Would you like to receive it **for free in your mailbox** every month?

**Subscription Form**
**(click here, it's free)**

You will fill the Subscription Form in **less than 1 minute**. Join many others computer vision professionals and receive all issues of Computer Vision News as soon as we publish them. You can also read Computer Vision News in **PDF version** and find in **our archive** new and old issues as well.


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

**Did you read the Feedback of the Month? It's on page 21**

## FEEDBACK

Dear reader,

If you like Computer Vision News (and also if you don't like it) we would love to hear from you:

**Give us feedback, please** (click here)

It will take you only 2 minutes. Please tell us and we will do our best to improve. Thank you!