

Using Intelligent Video Analytics against the Spread of COVID-19

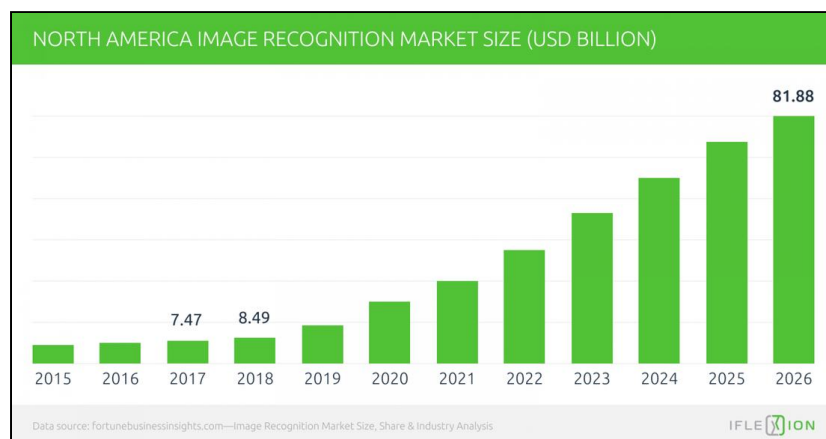
Originally published at <https://www.iflexion.com/blog/intelligent-video-analytics>

Published: July 2, 2020 | By Martin Anderson



Real-time object recognition and analysis has been one of the most compelling sub-sectors in machine learning and [AI software development](#) over the last ten years since the advent of GPU acceleration.

Besides increasing interest from government and civic departments, the transformative potential of [computer vision applications](#) has captured the imagination of the business world. The market for image recognition across all sectors in the US alone is forecast to rise to \$81.9 billion by 2026 according to Fortune Business Insights, with surveillance systems accounting for the largest section of the market.



Real-time object recognition and analysis holds a massive promise in fighting the spread of COVID-19 in public places.

The Value of Real-Time Segmentation in Intelligent Video Analytics

The ability to individuate objects in real time brings several logistical advantages to [political data science](#) and private sector video-analysis projects alike. Firstly, ongoing analysis with zero downtime means that statistical conclusions are drawn from the entirety of the data, and not a representative fragment of it. This changes the nature of the analysis from a study of averages to a data-complete and integral model.

Secondly, real-time segmentation enables real-time responses to real-life scenarios. For instance, an AI video analysis system that has been trained to recognize supine forms on a train platform is able to alert emergency response services immediately when a passenger collapses on a platform.

Creating Sustainable Business Environments Under COVID-19

2020 has brought an unusual and urgent use case for these techniques. At the time of writing, the world is attempting limited easing of the lockdown restrictions imposed by governments in the wake of the COVID-19 outbreak in early 2020. The economic impact of the lockdowns has been unprecedented, with a number of industries brought to their knees by the sudden disruption in trade:

- Airlines are set to lose \$113 billion in 2020, with experts predicting multiple bankruptcies in the sector.
- Global debt is set to soar from an already-staggering \$253 trillion in Q4 2019, with 53% held by private companies and corporations, and the further prospect of related bankruptcies as lockdown-related debt accrues.
- The falling price of oil, one of the world's key economic engines, remains in crisis, and has already necessitated a sub-cost sell-off at the height of the lockdowns.
- The radical fall in leisure and business travel and the 65% drop in flight capacity have had a massive impact on international and local business and tourism revenue, threatening further collapse for many businesses and stagnation or deficit for larger corporate entities.

The key to survival across most sectors is to develop at least a subsistence economy, where possible, relative to the pre-COVID-19 era.

For most sectors, this means creating new techniques and regulations designed not only to make some fraction of economic mobility possible, but to ensure that the incremental easing of lockdown regulations does not exacerbate a second wave of infection enough for the former, more restrictive measures to be re-imposed — sending ailing business models into free-fall again.

Computer Vision Systems vs COVID-19

Private and public sector cooperation is also unprecedented under the coronavirus-imposed conditions. Companies and corporations cannot institute effective anti-infection techniques in their headquarters if their workers are put at risk during their commute or by other civic factors related to re-establishing some degree of normal movement.

Consequently, the range of innovative vision-based approaches for mitigating COVID-19 infection rates and lockdown stasis are often cooperative efforts between industry and government, as well as public service departments.

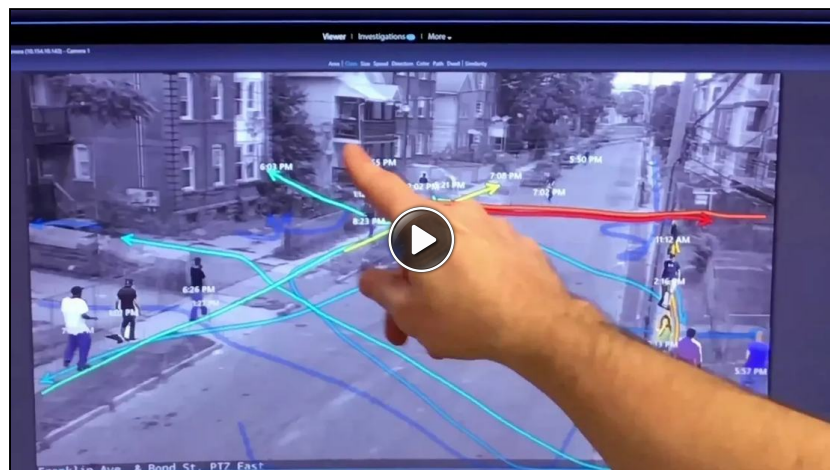
Segmentation and Tracking to Assess Mask-Wearing and Social Distancing

Since the outbreak of COVID-19, several systems have been devised to adapt real-time object segmentation and tracking techniques to the identification instances where social distancing rules are broken.

Social distancing infractions, or 'collisions', are calculated based on the proximity of tracking boxes that define the physical limits of each person recognized by the system.

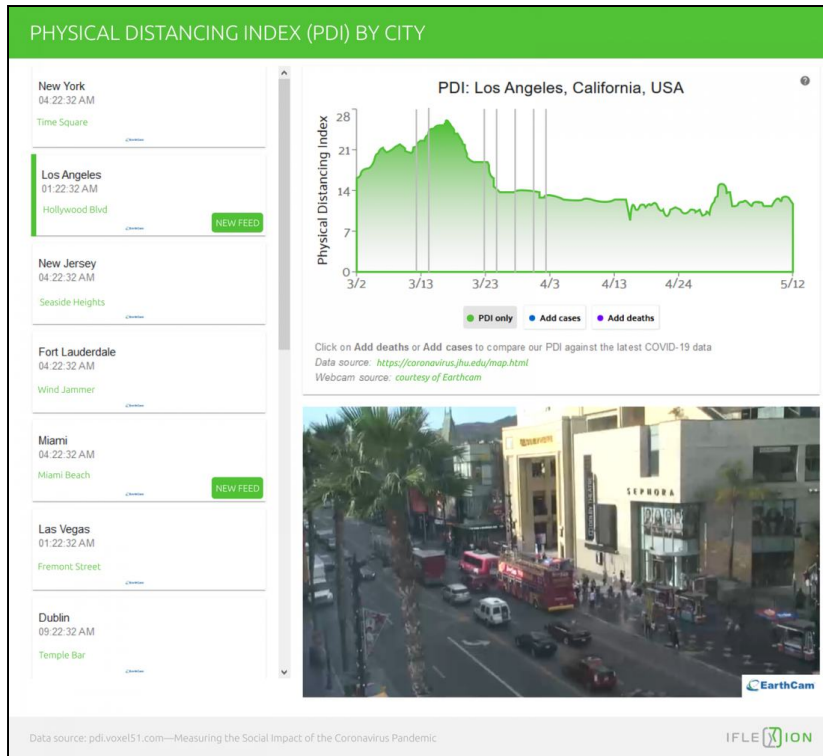
Without the facility for depth perception, the approach generally only works with high-set cameras which can clearly perceive the actual distance between pedestrians. Since such placements are rare, some innovation is necessary, as we will see.

Spatial calibration is a critical issue with an AI-assisted social distancing tracker, a challenge almost unique to COVID-19. In recent years, thanks to the rise in [machine learning](#) and the drive to automation of some civic and policing services, public CCTV feeds have served an increasing number of purposes besides general security recording. These include [facial recognition](#) of known criminals; gait recognition; and, in one case, determining that a domestic dwelling is being used for illegal purposes, based on the volume and frequency of visitors:

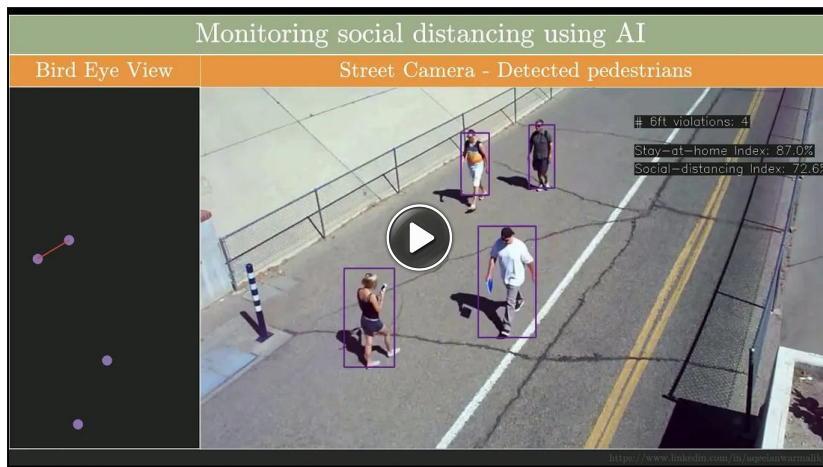


However, none of these novel pursuits have ever been required to estimate how far people in the video are standing from each other, or whether their walking pattern suggests that they are 'together' in the sense of cohabiting, which would indicate that their proximity should not be counted as an infraction of social distancing.

Voxel51 has developed a Physical Distancing Index (PDI) that dedicates specific machine learning approaches to the challenge of distance estimation. Though the source code is publicly available to clone from GitHub, it is composed of wrappers and API hooks, with tracking-related code remaining proprietary.



Aqeel Anwar from the Georgia Institute of Technology has created a similar system. In order to calibrate it, the user has to define a rectangle along the central converged plane on which the camera is trained. The four points of the rectangle are then interpolated by the system into a flat aerial view, which creates an index ratio to approximate the distance between the pedestrians.

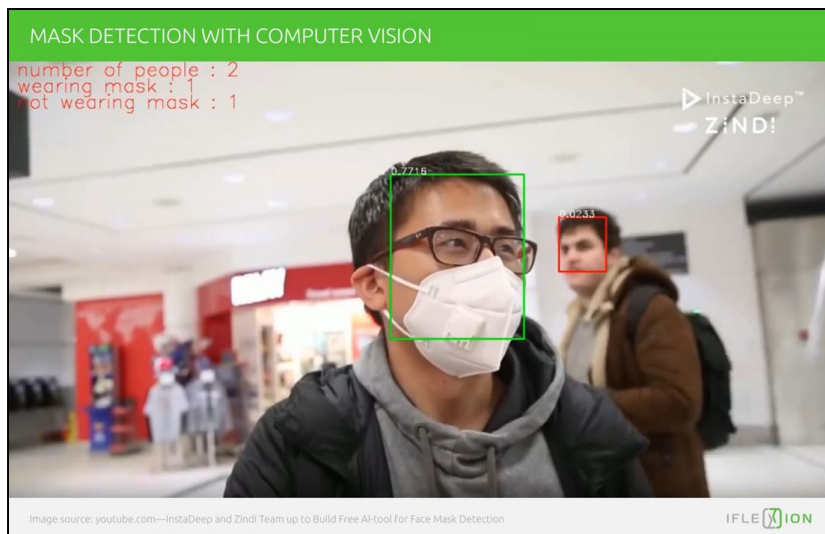


The social distancing data-points output from the system are:

- Frequency of distancing violations (6ft, 2 meters, or whatever is applicable in context) on a per-pedestrian basis.
- Volume of pedestrians as measured against pre-coronavirus levels, where this is possible, or else against earlier frequencies at various stages of the lockdown, or at other times of the day or week.
- Generation of a Social Distancing Index (SDI) which counts the number of violations, either as an aggregate number of violations (in which case one pedestrian may represent many offences), as an average number of violations per person, or as a granular index showing a gradient curve of pedestrians committing the most to least violations (though without facial recognition or some other identity token, a system naturally cannot account for subjects that re-appear later in a monitoring session).

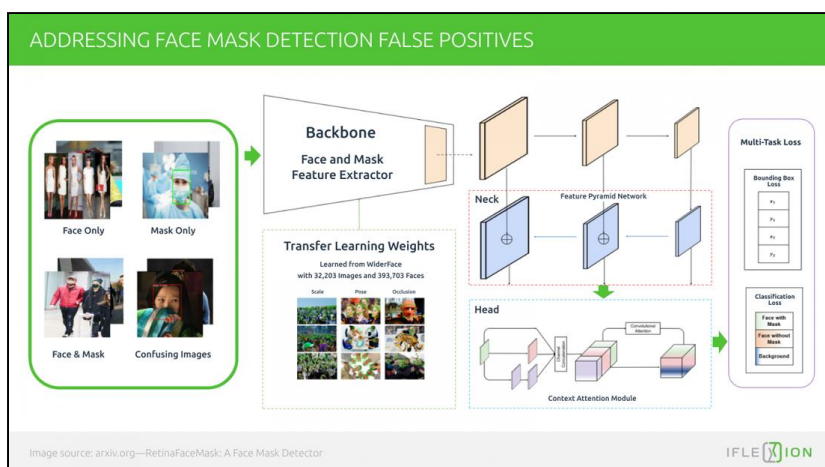
Detecting Mask-wearing with Computer Vision

Ironically, intelligent video analytics systems that were devised to detect mask-wearing as a sign of suspicious activity are now sometimes being used to detect it as an obligation.



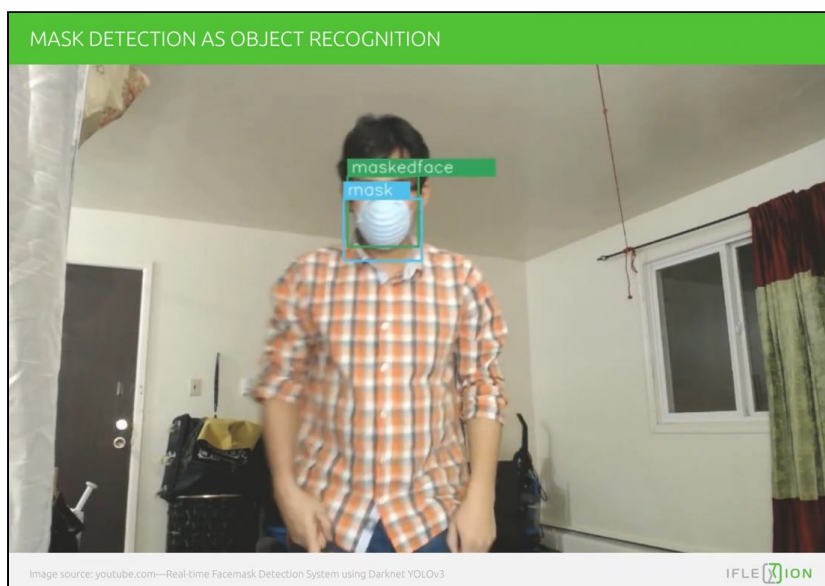
In [this Python project](#), a test subject is recognized as 'masked' as soon as she completes the act of putting on a mask. A trained facial recognition model is imported via TensorFlow into a Keras pipeline, and additional code added to analyze the processed face and monitor it for a 'masked' state.

As the study called *RetinaFaceMask: A Face Mask Detector* out of the University of Hong Kong shows, false positives are possible depending on the sophistication of the program and the detected traits that are used as a key indicator.



For instance, if a face is detected where eyes are apparent but no mouth appears, some people might have slight enough physiognomy to appear masked when they are not. Likewise, where a shift in color across the top and bottom of the face is the key indicator, a mask color similar to the flesh-tones of the subject's face could also indicate that they are not wearing a mask, even though they are.

A surer way would be to detect the lineaments of a mask itself, an approach that mixes facial recognition with object recognition, as one project has attempted:



In this case, the AI is running effectively on a lean NVIDIA GTX 1050ti with 4gb VRAM.

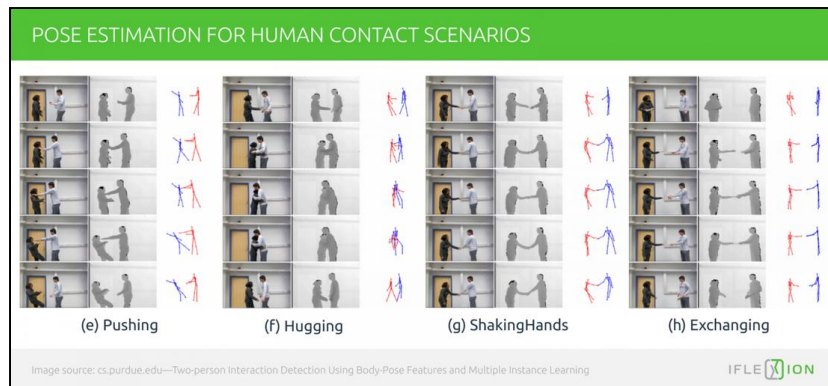
This project, along with most of the numerous other COVID-19 mask-identification initiatives, uses a variation on YOLO, the most popular real-time object detection engine used by [computer vision software developers](#) for machine learning deployment.

Recent research out of the University of California has also [proposed YOLACT++](#), a more efficient instance segmentation implementation of YOLO that uses parallel processing to improve recognition quality and speed.

Further Applications of Computer Vision for Infection Analysis

If social distancing turns out to be a long-term prospect, and track-and-trace systems get a chance to become more sophisticated, there are extended possibilities for the application of computer vision techniques.

For instance, pose estimation and gesture analysis technologies are capable of recognizing a handshake when it occurs:



In the context of a lockdown, this is a pretty reliable high-risk transmission event, since two cohabiting people are unlikely to make this semi-formal gesture in any circumstances.

Though no direct action is practicable in response to a public handshake, being able to recognize and log 'contact events' through AI-driven intelligent video analysis would be useful in correlating public behavior to fluctuations in transmission rates.

The increased interest in real-time pose estimation over the last ten years has come both from the fact that GPU acceleration and open-source research has made it practicable, and that it offers so many possibilities for detecting 'violent' action in the streams of generally under-utilized CCTV systems.



There are extended possibilities for computer vision applications if social distancing turns out to be a long-term prospect.

Detecting Traces of COVID-19 in X-Rays

MSc students on the UK Cranfield Aerospace research team have used machine learning to develop algorithms capable of identifying signs of the disease in X-rays.

In cases of pneumonia caused by COVID-19, traces left in the lung area are distinct enough to be specifically associated with the disease, thus making it easier to distinguish common annual infections from those related to the coronavirus through [medical image analysis](#).

Though initially challenged by the lack of public domain imagery, the team assembled enough data to utilize in machine learning frameworks and develop an effective image recognition algorithm.

Using Thermal Imaging Analysis to Detect COVID-19 Symptoms

Thermal imaging is a fairly blunt instrument in the context of COVID-19. Facial temperature can be notably lower than internal body temperature depending on the individual, while exertion, exposure to heat and sun prior to scanning, and general disposition to heat can all give false positives.

AI-based fever detection imaging systems usually concentrate attention on the area around the eyes, where the insulation between the interior and exterior body temperature is smallest, at least in a non-medical environment where only the face will be available for analysis.

A number of thermal video analysis systems have been implemented around the world, both to identify itinerant carriers and those who are not wearing masks where regulations demand it.

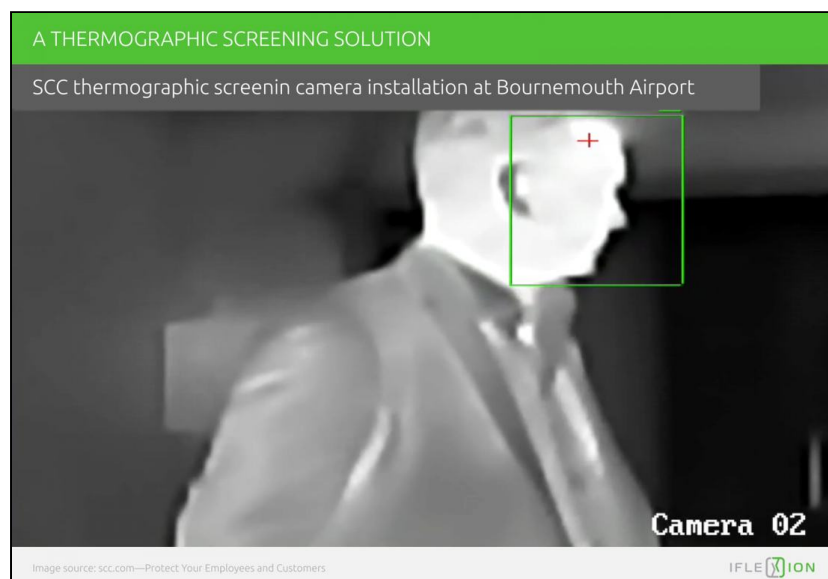
However, China's efforts in this regard have received the most media attention: in Wuhan, police drones were re-equipped to provide video material for machine learning-based algorithms to detect citizens going outdoors without masks.



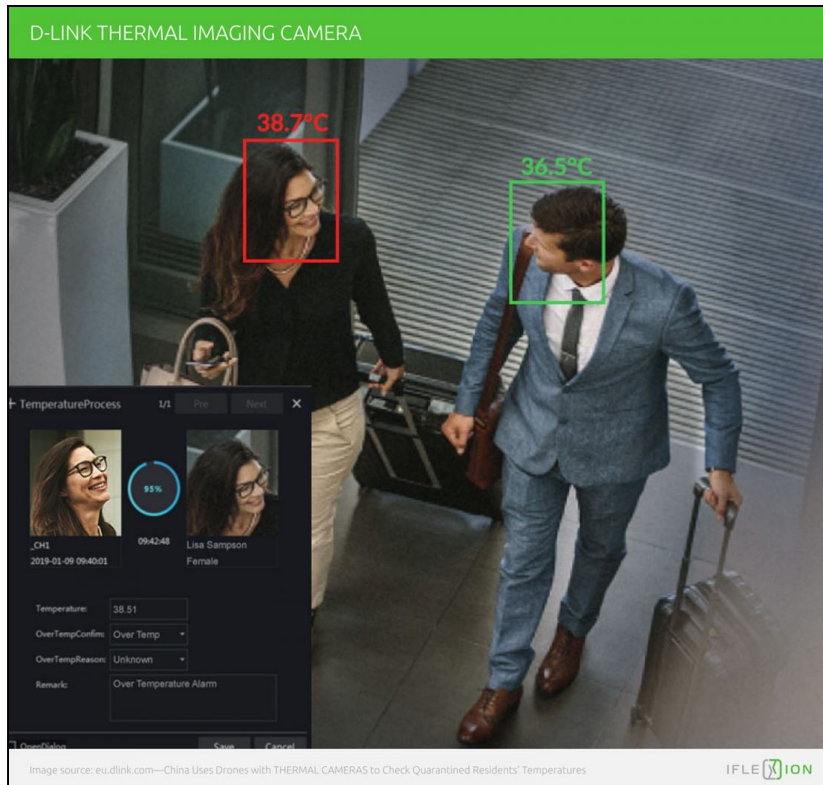
Also in Wuhan, the Chinese government implemented a network of infrared thermal imaging cameras to detect individuals in the station that appear to have a fever-grade body temperature. The system triggers when its AI-driven algorithm identifies a person whose visual signature matches a temperature of 37 degrees or above, based on the training data fed into the neural network that created the algorithm.

For citizens wanting to get in on the action, Wuhan-based Guide Sensmart Tech Co., Ltd offers a thermal imaging camera that can be attached to a smartphone. The device is intended to locate hidden cameras and leaky ventilation systems, but can be used to identify high-temperature individuals too.

Across the world, AI-based thermal imaging systems have been either implemented or adapted with a view to making travel safer. At Bristol Airport in the UK, a facial recognition system has been re-tooled as a thermal imaging system, with AI-driven triggers if a subject in a video stream appears to have a fever.



Dedicated fever-detection hardware is also beginning to emerge, such as a dual-lens thermal imaging camera intended for airports and other major public spaces.



One AI thermal imaging manufacturer suggests the possibility that temperature-screening cameras may become more covert in the long term, which indicates that the personal nature of temperature screening may eventually galvanize privacy rights supporters.

A number of thermal video analysis systems have been implemented around the world, both to identify itinerant carriers and those who are not wearing masks where regulations demand it.

Conclusion

In this context of the current public health crisis, passive and contactless systems of public control and health management are required more than ever before — a practically perfect use case for the new raft of innovations in computer vision, object detection and segmentation over the last ten years.