

Converting LiDAR to Photo-Real Imagery With a Generative Adversarial Network

Originally published at <https://www.unite.ai/converting-lidar-to-photo-real-imagery-with-a-generative-adversarial-network/>



Earlier this week, footage was [released](#) showing a Tesla autopilot system crashing directly into the side of a stalled vehicle on a motorway in June of 2021. The fact that the car was dark and difficult to discern has prompted [discussion](#) on the limitations of relying on computer vision in autonomous driving scenarios.



[CLICK TO VIEW ANIMATED GIF](#)

Footage released in December 2021 depicts the moment of impact. Source: <https://twitter.com/greentheonly/status/1473307236952940548>

Though video compression in the widely-shared footage gives a slightly exaggerated impression of how quickly the immobilized truck 'snuck up' on the driver in this case, a higher-quality video of the same event demonstrates that a fully-alert driver would also have struggled to respond with anything but a tardy swerve or semi-effective braking.

The footage adds to the controversy around Tesla's [decision to remove radar sensors](#) for Autopilot, announced in May 2021, and its stance on [favoring vision-based systems](#) over other echo-location technologies such, as LiDAR.

By coincidence, a new research paper from Israel this week offers an approach to straddle the LiDAR and computer vision domains, by converting LiDAR point clouds to photo-real imagery with the use of a Generative Adversarial Network (GAN).



In the new project from Israel, Black cars identified in LiDAR footage are converted to a 'daylight' scenario for computer vision-based analyses, similar to the task of developing its Autopilot system. Source: <https://arxiv.org/pdf/2112.11245.pdf>

The authors state:

'Our models learned how to predict realistically looking images from just point cloud data, even images with black cars.'

'Black cars are difficult to detect directly from point clouds because of their low level of reflectivity. This approach might be used in the future to perform visual object recognition on photo-realistic images generated from LiDAR point clouds.'

Photo-Real, LiDAR-Based Image Streams

The [new paper](#) is titled *Generating Photo-realistic Images from LiDAR Point Clouds with Generative Adversarial Networks*, and comes from seven researchers at three Israeli academic faculties, together with six researchers from Israel-based Innoviz Technologies.

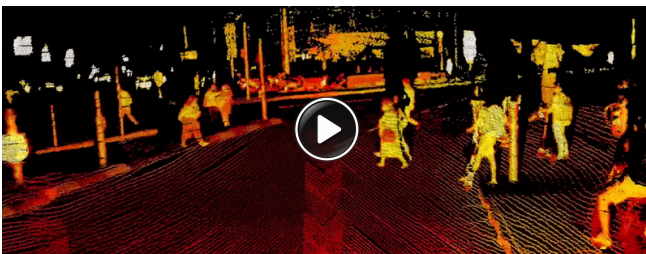
The researchers set out to discover if GAN-based synthetic imagery could be produced at a suitable rate from the point clouds generated by LiDAR systems, so that the subsequent stream of images could be used in object recognition and semantic segmentation workflows.

Data

The central idea, as in so many [novel \[x\]>\[x\] image transliteration projects](#), is to train an algorithm on paired data, where LiDAR point cloud images (which rely on device-emitted light) are trained against a matching frame from a front-facing camera.

Since the footage was taken in the daytime, where a computer vision system can more easily individuate an otherwise-elusive all-black vehicle (such as the one that the Tesla crashed into in June), this training should provide a central ground truth that's more resistant to dark conditions.

The data was gathered with an InnovizOne LiDAR sensor, which offers a 10fps or 15fps capture rate, depending on model.

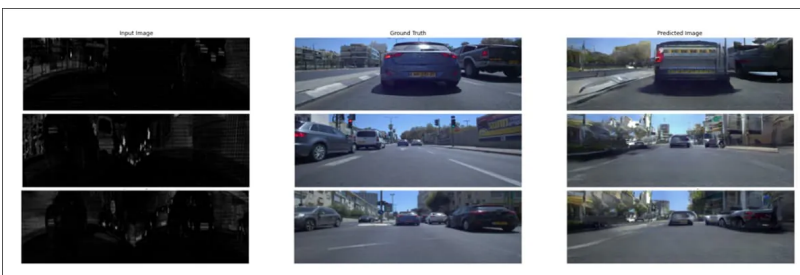


CLICK TO VIEW ANIMATED GIF

LiDAR data captured by an Innoviz device. Source: https://www.youtube.com/watch?v=wmcaf_VpsQI

The resulting dataset contained around 30,000 images and 200,000 collected 3D points. The researchers conducted two tests: one in which the point cloud data carried only reflectivity information; and a second, in which the point cloud data had two channels, one each for reflectivity and distance.

For the first experiment, the GAN was trained to 50 epochs, beyond which overfitting was seen to be an issue.



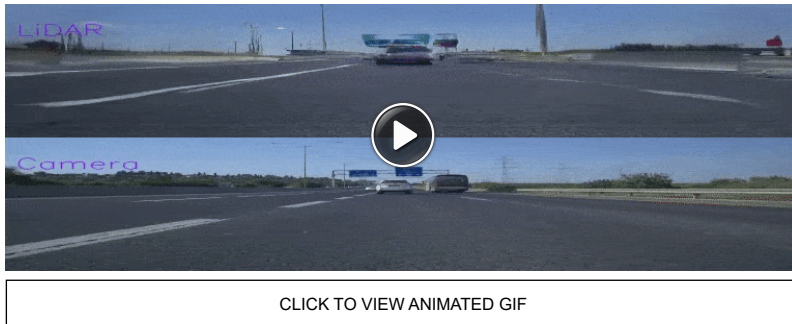
GAN-created images from the first experiment. On the left, point cloud data; in the middle, actual frames from captured footage, used as ground truth; right, the synthetic images generated by the Generative Adversarial Network.

The authors comment:

'The test set is a completely new recording that the GANs have never seen before the test. This was predicted using only reflectivity information from the point cloud.'

'We selected to show frames with black cars because black cars are usually difficult to detect from LiDAR. We can see that the generator learned to generate black cars, probably from contextual information, because of the fact that the colors and the exact shapes of objects in predicted images are not identical as in the real images.'

For the second experiment, the authors trained the GAN to 40 epochs at a batch size of 1, resulting in a similar presentation of 'representative' black cars obtained largely from context. This configuration was also used to generate a [video](#) that shows the GAN-generated footage (pictured upper, in the sample image below) together with the ground truth footage.



Evaluation

The customary process of evaluation and comparison to existing state-of-the-art was not possible with this project, due to its unique nature. Instead the researchers devised a custom metric regarding the extent to which cars (minor and fleeting parts of the source footage) are represented in the output footage.

They selected 100 pairs of LiDAR/Generated images from each set and effectively divided the number of car images present in the source footage to the number present in the synthetic data produced, producing a metric scale of 0 to 1.

The authors state:

'The score in both experiments was between 0.7 and 0.8. Considering the fact that the general quality of the predicted images is lower than the real images (it is more difficult in general to detect objects in lower quality images), this score indicates that the vast majority of cars that present in the ground truth present in the predicted images.'

The researchers concluded that the detection of black vehicles, which is a problem for both computer vision-based systems and for LiDAR, can be effected by identifying a lack of data for sections of the image:

'The fact that in predicted images, color information and exact shapes are not identical to ground truth, suggests that that prediction of black cars is mostly derived from contextual information and not from the LiDAR reflectivity of the points themselves.'

'We suggest that, in addition to the conventional LiDAR system, a second system that generates photo-realistic images from LiDAR point clouds would run simultaneously for visual object recognition in real-time.'

The researchers intend to develop the work in the future, with larger datasets.

Latency, and the Crowded SDV Processing Stack

One commenter on the much-shared Twitter post of the Autopilot crash estimated that, traveling at around 75mph (110 feet a second), a video feed operating at 20fps would only cover 5.5 feet per frame. However, if the vehicle was running Tesla's latest hardware and software, the frame rate would have been 36fps (for the main camera), which sets the evaluation rate at [110 feet per second](#) (three feet per frame).

Besides cost and ergonomics, the problem with using LiDAR as a supplementary data stream is the sheer scale of the informational 'traffic jam' of sensor input to the SDV processing framework. Combined with the critical nature of the task, this seems to have forced radar and LiDAR out of the Autopilot stack in favor of image-based evaluation methods.

Therefore it seems unlikely that a system using LiDAR – which in itself would add to a processing bottleneck on Autopilot – to infer photo-real imagery is feasible from Tesla's point of view.

Tesla founder Elon Musk is no blanket critic of LiDAR, which [he points out](#) is used by SpaceX for docking procedures, but considers that the technology is 'pointless' for self-driving vehicles. Musk suggests that an occlusion-penetrating wavelength, such as the ~4mm of precision radar, would be more useful.

However, as of June 2021, Tesla vehicles are [not outfitted with radar either](#). There do not currently seem to be many projects designed to generate image streams from radar in the same way as the current Israeli project attempts (though the US Department of Energy [sponsored](#) one attempt for radar-sourced GAN imagery in 2018).

First published 23rd December 2021.

Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai



Discover More