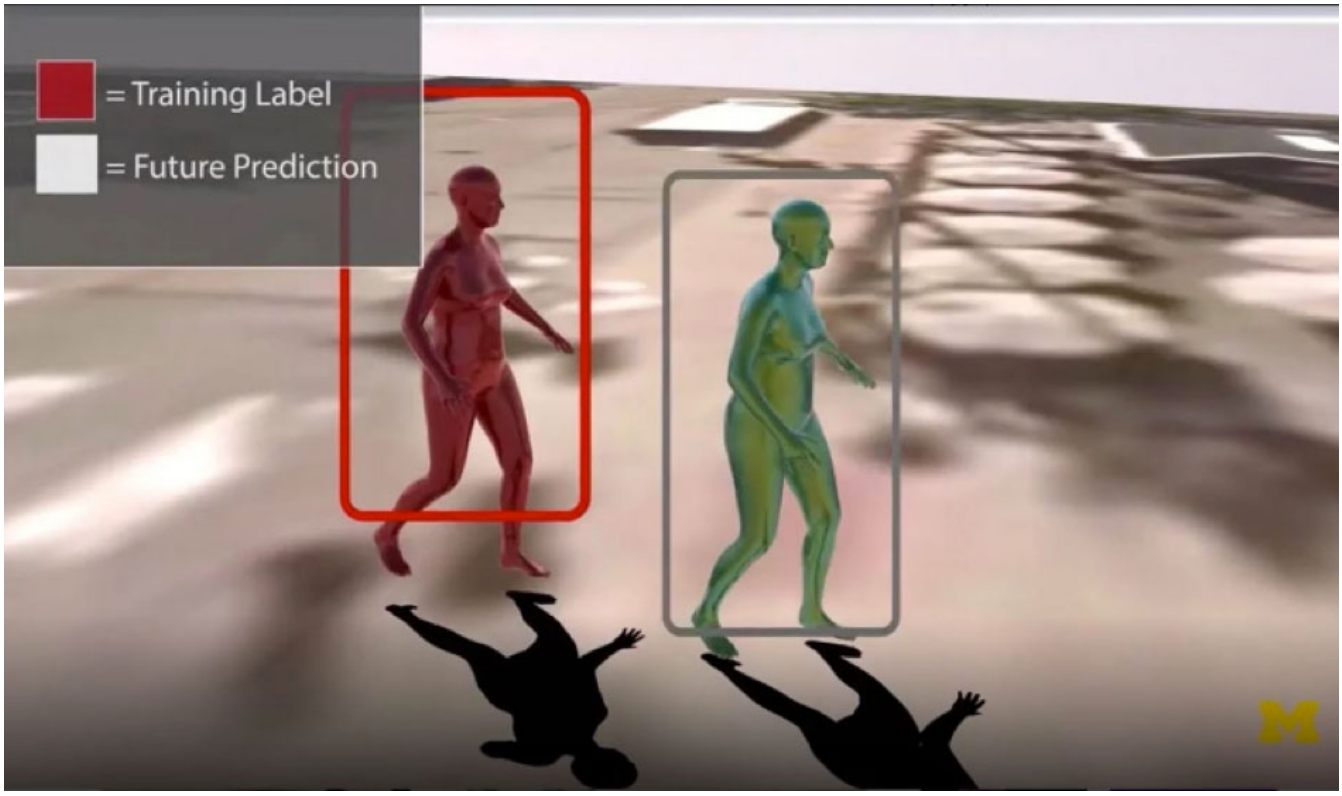
By studying humans' gait, body symmetry, and foot placement, researchers at the University of Michigan are teaching self-driving cars to recognise and predict pedestrians' motion with greater precision than current technologies. Data collected by vehicles through cameras, lidar and GPS allows the UM researchers to capture video snippets of humans in motion and then recreate them in 3D computer simulation. With that, they have created a 'biomechanically inspired recurrent neural network' (Bio-LSTM) that catalogs human movements. With it, they can predict poses and future locations for one or several pedestrians up to about 46 metres from the vehicle, which is about the scale of a city street junction.

Equipping vehicles with the necessary predictive power requires the network to dive into the minutiæ of human movement: the pace of a human's gait, the mirror symmetry of limbs, and the way foot placement affects stability during walking.

Much of the machine learning used to bring AV technology to its current level has dealt with 2D images. A computer shown several million photos of a stop sign, for example, will eventually come to recognise stop signs in the real world and in real time. But by using video clips that run for several seconds, the UM system can study the first half of the snippet to make its predictions, and then verify their accuracy with the second half. Supported by a grant from Ford, the team's results have shown that this new system improves a driverless vehicle's capacity to recognise what is most likely about to happen.

To reduce the number of options for predicting the next movement, researchers applied the physical constraints of the human body, such as our fastest possible speed on foot. To create the dataset used to train UM's neural network, researchers parked a vehicle with level-4 autonomous features at several Ann Arbor intersections.



With the car's cameras and lidar facing the intersection, the AV could record multiple days of data at a time. Researchers bolstered the real-world field data from traditional pose data sets captured in a lab. The result is a system that will raise the bar for what driverless vehicles are capable of. "Prior work in this area has typically only looked at still images. It wasn't really concerned with how people move in three dimensions", says UM assistant professor of mechanical engineering Ram Vasudevan. "But if these vehicles are going to operate and interact in the real world, we need to make sure our predictions of where a pedestrian is going doesn't coincide with where the vehicle is going next."

UM associate professor Matthew Johnson-Roberson adds, "Now, we're training the system to recognise motion and making predictions of not just one single thing, but where that pedestrian's body will be at the next step and the next and the next. The median translation error of our prediction was approximately 10 cm after one second and less than 80 cm after six

seconds. All other comparison methods were off by up to 7 m. We're better at figuring out where a person is going to be."