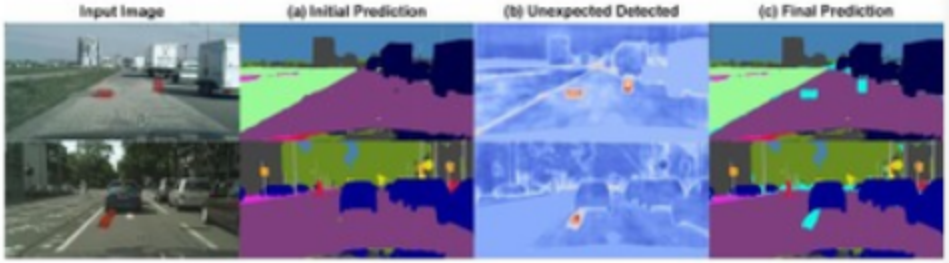
Standardized Max Logits: A Simple yet Effective Approach for Identifying Unexpected Road Obstacles in Urban-Scene Segmentation

Sanghun Jung (left) and Jungsoo Lee (right) are Master's students at KAIST AI in South Korea, under the supervision of Professor Jaegul Choo.

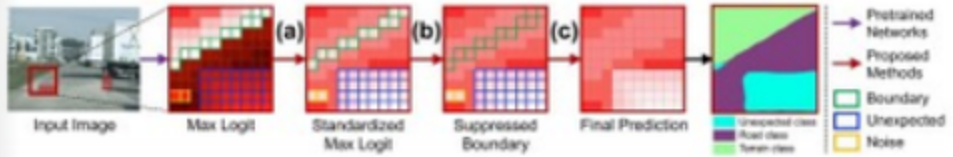
Their work, which tackles out-of-distribution detection in semantic segmentation, has been selected as an oral presentation this year. They speak to us ahead of their live Q&A session today.



Semantic segmentation of urban-scene images usually works with a pre-trained model that predicts all the pixels in an image as one of the predefined classes, such as cars or roads. However, there is no training data for **out-of-distribution detection**. Even if such data could be collected, it would always be restricted because you cannot collect data for every unexpected object in the real world.

By detecting **anomaly pixels or objects in an image**, this work provides an initial point to treat them differently from the training classes.

"There are well-known metrics to measure anomaly scores of the objects in an image," Sanghun tells us.

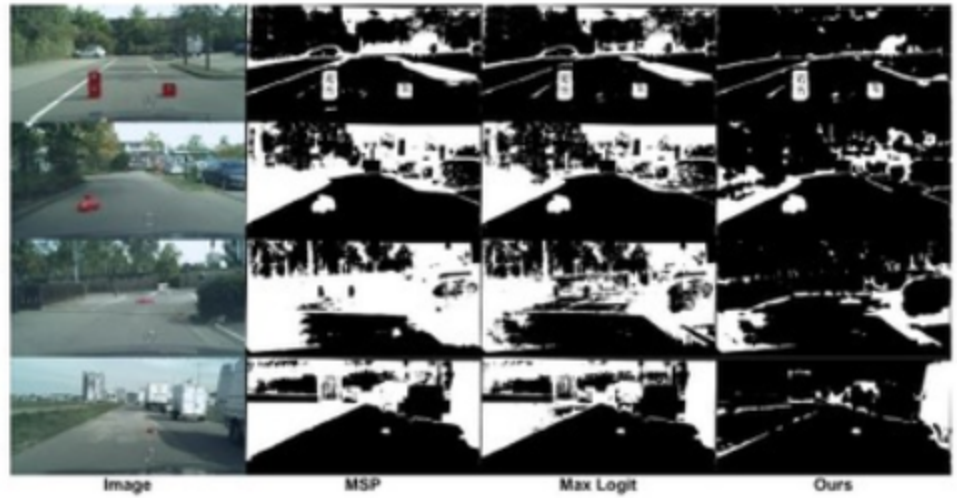


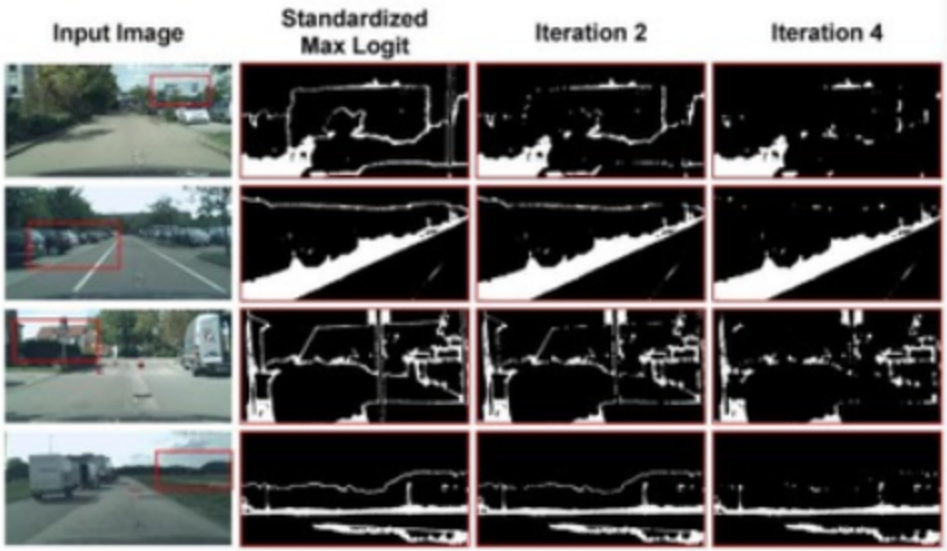
"We improve those metrics for semantic segmentation, especially in the urban sense. Our work starts from the intuition that the distribution of prediction scores is significantly different from each other for predicted classes. By normalizing them to have the same semantics, that's the starting point of our work."

There are three main contributions of this work. The first is trying to **align the differently formed prediction scores**, which is very commonly observed in semantic segmentation. Previous work didn't tackle this problem.

"The second contribution is that in semantic segmentation, the model detects the in-distribution pixels as the out-of-distribution pixels in the boundary regions," Jungsoo explains.

"One of the reasons is because the boundary regions are where the classes change, like from cars to road. Because of that characteristic, there are a lot of false positives occurring in the boundary regions. We try to reduce the false positives with a module named boundary-aware pooling. For the third contribution, there are some noises still existing in semantic segmentation after applying the first and second parts, so we reduce those noises using a Gaussian filter."





"One of the reasons is because the boundary regions are where the classes change, like from cars to road."

All the techniques used are based on **computer vision techniques**. The work uses a semantic segmentation model which shows state-of-the-art performance on segmenting the classes in an image. After that, as a post-processing approach, Gaussian smoothing is applied, which uses a **Gaussian kernel** for removing the smooth noises in an image.

One of the assumptions in the work is that the model has a reasonable capacity to discriminate between in-distribution and out-of-distribution pixels. If the model cannot discriminate between the two in the first place, the rest of it does not work. In terms of next steps, the team want to improve its capability to discriminate between those in-distribution and out-of-distribution pixels before applying the three models.

"Our method achieves state-of-the-art performance on the public leaderboard," Sanghun says proudly.

"It's a very simple and effective approach. It doesn't need any additional training or additional out-of-distribution data set."

Sanghun and Jungsoo were both present at the last in-person ICCV two years ago in Seoul.

"Our professor provided that opportunity to participate in the conference," Jungsoo remembers.

"That was a very honorable opportunity for me. I first met my lab mates there. It is such a happy memory walking around with my lab mates and listening to the posters and the sessions. I'd never had that kind of opportunity before. Undergrad students don't have those opportunities normally. I was very lucky. It was a great opportunity to broaden my horizons."

Jungsoo wants to share a final personal message:

"This is the first deep learning paper I've had accepted. I just want to thank my co-workers, Sanghun Jung, the first author, and especially our professor, Jaegul Choo, for advising us. It was really good to work with them all!"

To learn more about Sanghun and Jungsoo's work [Paper ID 2171], you can view their pre-recorded oral presentation and PDF poster now and join their live Q&A sessions today [12A] at 19:00-20:00 EDT and Friday [12B] at 12:00-13:00 EDT.