with the closest method, YOLOv4 [3], while with a similar inference speed. In addition, MSFANet gets much higher accuracy than other one-stage detectors, e.g. SSD[13], CSPNET[22], shows in table II. For two-stage methods, our MSFANet model achieves a 10.39 point promotion compared with Faster R-CNN[7] model based on ResNet101.

Conventional one-stage systems struggle with small-scale objects compared to multi-scale method. YOLOv4[3] extracts features from different scales using a similar concept to feature pyramid networks and predicts boxes at 3 different scales to improve the performance of small objects. On classes like rockfalls and warning signs, MSFANet scores 11.93% higher than YOLOv4. And, on other classes like truck and car MSFANet also achieves higher performance. Two-stage methods achieve better performance on small-scale object detection. We found our MSFANet get a 25.71 point promotion even compared with Faster R-CNN.

Consequently, the experimental results show the competitive -ness of our proposed MSFANet, as shown in appendix, especially on small-scale and hard samples, and the trade-off between accuracy and speed on SurMine determines it is more suitable for autonomous driving in surface mine.

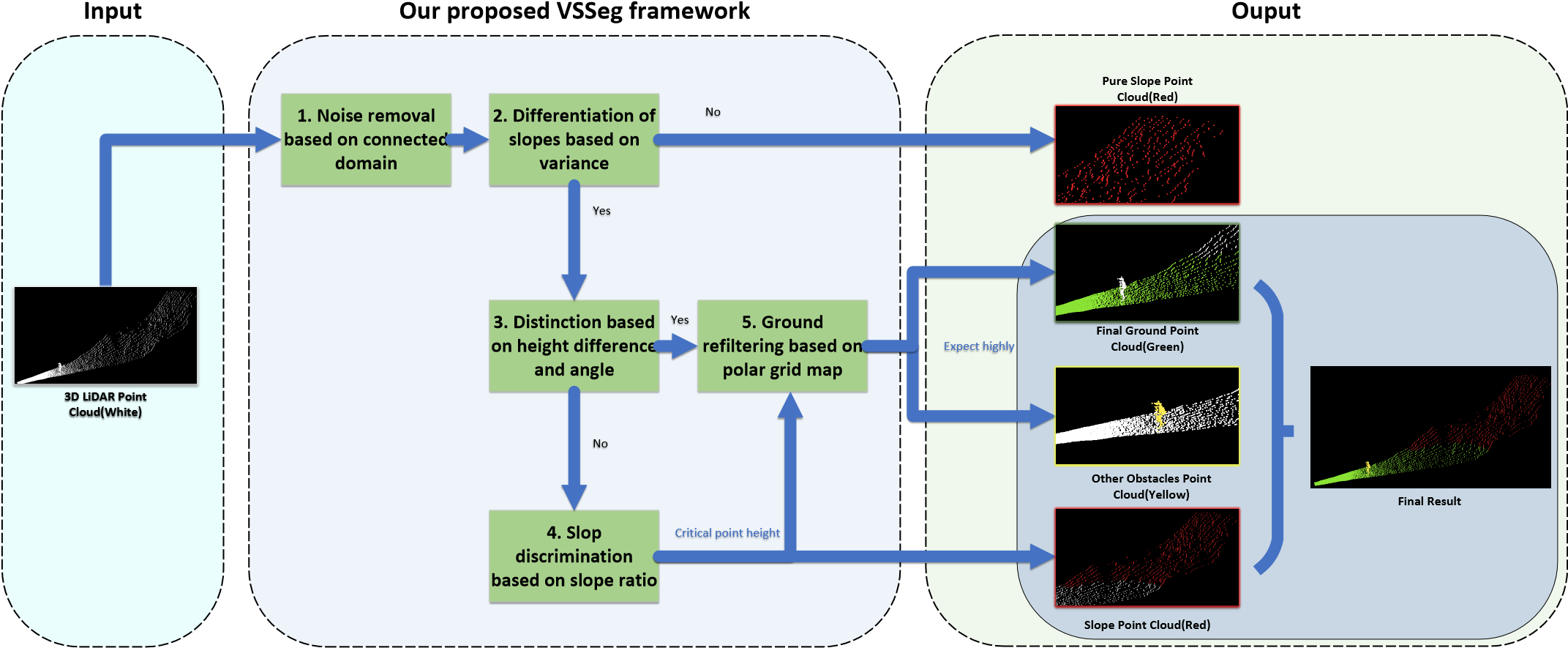


Figure 8. Some images with challenges of object detection in surface mine

## Experiment on KITTI

To make the experimental results more convincing, we perform several experiments on public autonomous driving datasets, KITTI. We test several advanced detection model on KITTI. MSFANet has good performance on KITTI and its AP is higher than other one-stage methods, like SSD, YOLOv3, YOLOv4. Our MSFANet produced small increases in mAP between 4.17 point and 20.51 point, see Table III for details. Compared to classical two-stage methods, we achieves a 12.75 point mAP promotion (92.57vs. 79.82) with the well known method, Faster R-CNN [9], while with a faster inference speed.

Consequently, the experimental results show the MSFANet not only have a good performance on SurMine, but also achieve robust performance on public dataset of autonomous driving . It gets a trade-off between accuracy and speed on both on SurMine and KITTI.