

Supplementary Material

Detecting Underwater Discrete Scatterers in Echograms with Deep Learning-Based Semantic Segmentation

Rhythm Vohra*, Femina Senjaliya*, Melissa Cote*, Amanda Dash*[†], Alexandra Branzan Albu*,
Julek Chawarski[†], Steve Pearce[†], Kaan Ersahin[†]

* Electrical and Computer Engineering, University of Victoria, Victoria, Canada

[†] ASL Environmental Sciences, Victoria, Canada

Email: {rhythmvohra, feminasenjaliya, mcote, aalbu}@uvic.ca, {adash, jchawarski, spearce, kersahin}@aslenv.com

1. Additional results

Fig. 1 to 3 show additional visual results obtained by the three compared semantic segmentation networks (U-Net, Attention U-Net, and DeepLabV3+) for the two best fusion strategies, i.e., early and late. For an analysis of the figures, we refer the reader to the discussion of Fig. 4 in the main text, as a similar analysis applies here, and to the captions in this document.

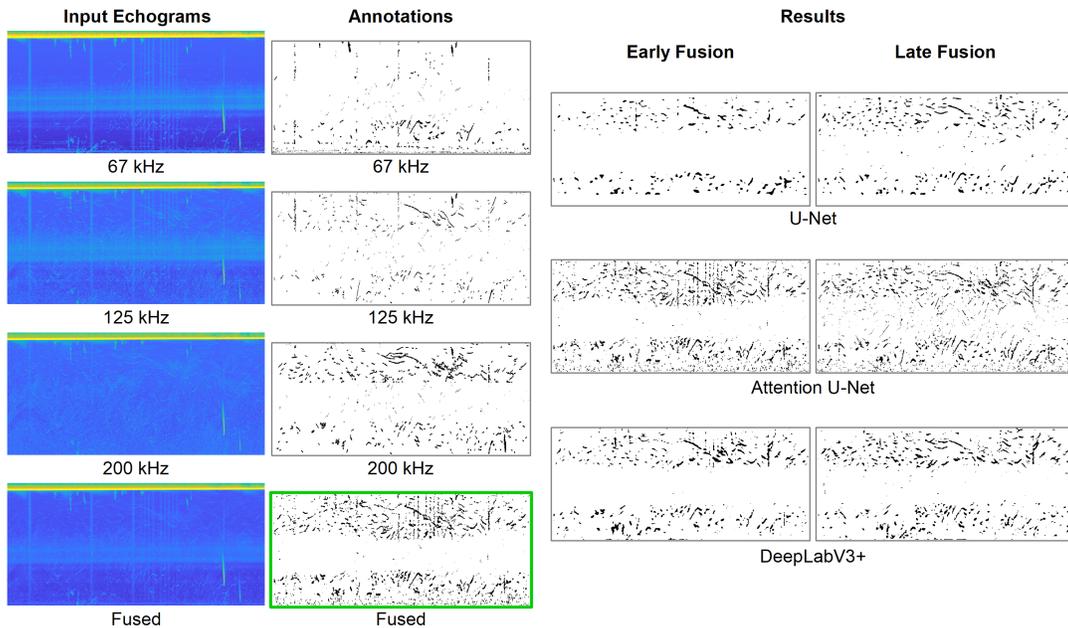


Figure 1. Additional sample results for one echogram from the OCUDS test set, for the best performing strategies (early and late fusion): 27-JUL-2015 10:00-11:00. All results should be compared to the fused annotations (with green bounding box). In the annotations and results, black pixels represent discrete scatterers. Attention U-Net yields more detected pixels with many smaller detections, and all networks yield additional and smaller detections in the late fusion case. Also, the networks seem to outperform the annotations with respect to the vertical noise bands (parts of the annotations). See discussion in manuscript for possible explanations.

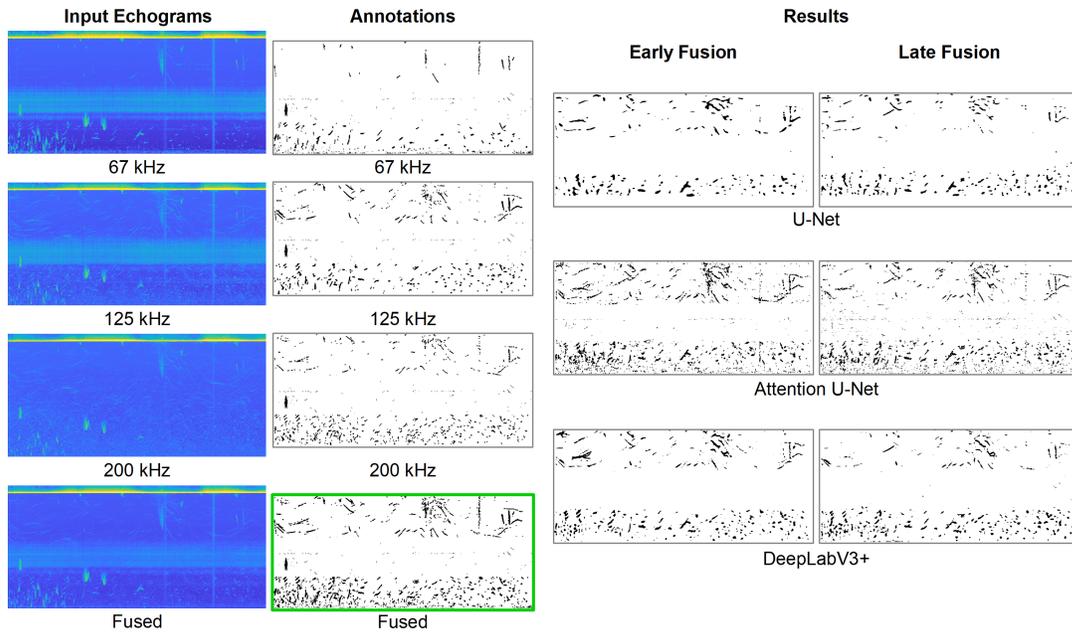


Figure 2. Additional sample results for one echogram from the OCUDS test set, for the best performing strategies (early and late fusion): 24-JUL-2015 05:00-06:00. All results should be compared to the fused annotations (with green bounding box). In the annotations and results, black pixels represent discrete scatterers. As in other visual examples, Attention U-Net yields more detected pixels with many smaller detections, and networks tend to yield additional and smaller detections in the late fusion case. The networks seem to outperform the annotations with respect to the vertical noise bands and the aggregation in the middle part towards the left handside (part of the annotations but not detected). See discussion in manuscript for possible explanations.

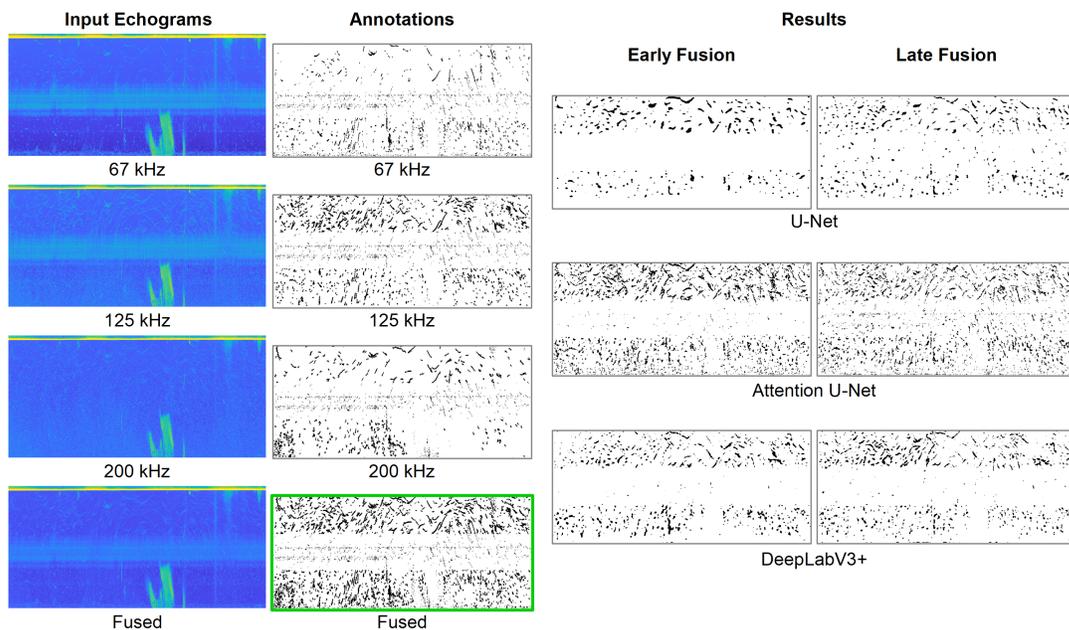


Figure 3. Additional sample results for one echogram from the OCUDS test set, for the best performing strategies (early and late fusion): 29-JUL-2015 14:00-15:00. All results should be compared to the fused annotations (with green bounding box). In the annotations and results, black pixels represent discrete scatterers. As in other visual examples, Attention U-Net yields more detected pixels with many smaller detections, and networks tend to yield additional and smaller detections in the late fusion case. The networks successfully deal with the presence of large aggregations and bubbles (in yellowish-green).