Atmospheric
Measurement
Techniques

Open Access

Discussions

EGU

Interactive comment on "Applying machine learning methods to detect convection using GOES-16 ABI data" *by* Yoonjin Lee et al.

Anonymous Referee #2

AMT-2020-420 Applying machine learning methods to detect convection using GOES-16 ABI data

This manuscript suggests the detection algorithm of convection using GOES-16 and convolutional neural networks. By using the encoder-decoder model, the suggested model can predict the convection with the same resolution of the input. Without the handcrafted input features based on the physical characteristics, the suggested model can produce the convection model well. This manuscript is generally well-written with the proper experiment and discussion. The advantage of two-step loss is interesting and intuitive. However, there are some major issues to be addressed to improve the manuscript.

1. Line 136: The whole images were divided into multiple tiles. Please provide the

reason for using tile rather than the whole image.

2. The resolution of MRMS is not explicitly provided.

3. Line 185: What is the 'simple transformations' of the batch normalization? Please add a more specific explanation.

4. Typically, binary cross-entropy is used for the binary classification. MSE is generally used for the regression because the target data is continuous. Even if MSE is successfully used in this study for two-step loss and the threshold, it is not the usual case for the classification. Hence, the rationale of using MSE should be briefly noted in the manuscript.

5. How two losses were trained? Is the model trained with the loss 1 and re-trained with the loss 2? Or, are two losses are in the same network? The experimental scheme should be suggested with more detail. Also, some key hyper-parameters such as epoch or optimizer should be added.

6. Figure 1: The final output is cracked. Please check the figure.

7. Discussion: The locations of the states should be in the figures if the paper is not just for the readers from the US. (line 299 and others).

8. Figure 5-10: It is recommended to crop the marginal area of the figures.

9. There is no discussion about channel 14. When the IR channel is not important to detect the convection, I think it can be removed from the model. Or, the discussion and analysis of the IR channel should be conducted with proper figures and discussion.

10. As the image is divided into tiles, how they are merged? How about the discontinuity at the edge of each tile?

11. The visible channel is not available at nighttime. Even if the convection occurs in the daytime more, it is the limitation of the visible channel and should be noted as the limitation of the suggested model.

12. The temporal sequences of VIS and IR are fed into the model. However, the discussion and analysis of the effect of the time series dataset are not covered. It is necessary to compare the results along the length of the sequence when using the temporal dataset.