Reconstruction of Missing Data in GOCI AOD using a Deep Learning Algorithm

Yannic Lops[1], Yunsoo Choi[1] (ychoi6@uh.edu), Ahmed Khan Salman[1], Ebrahim Eslami[1], Sojin Lee[1], Jia Jung[1]
[1]Department of Earth and Atmospheric Sciences, University of Houston, Houston, TX 77204, USA

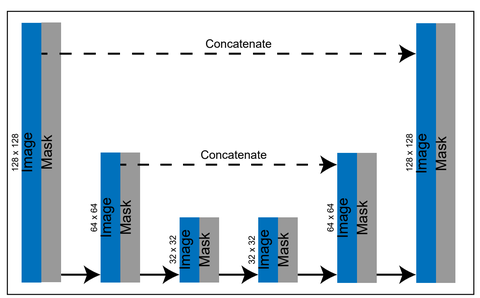
SUMMARY

Remote sensing Aerosol Optical Depth (AOD) data faces limitations of data occlusion due to cloud cover. Due to this, the Geostationary Ocean Color Imager (GOCI) satellite data cannot be directly used for extracting coherent patterns and mechanistic correlations of AOD. The only recourse is to spatially interpolate the data. Kriging is a widely used spatial interpolation approach, but has a computational cost that scales to the cube of the number of data points. Furthermore, Kriging experiences problems when interpolating image borders. We utilize a deep learning technique, Partial CNN (Liu et al. 2018), as a generalized model for reconstruction of GOCI AOD images. Results show the Partial CNN model can predict missing data with a lower mean absolute error (MAE) than Kriging (0.05 for Partial CNN and 0.06 for Kriging). The model has on average 2% higher Index of Agreement (IOA) and Pearson correlation coefficient than Kriging. The partial convolution model was achieving more stable results with less variance of predictions than kriging in all statistical evaluation methods.

Highlights

- Model implemented and tested for accurate spatial data imputation
- Model is more stable than kriging in predicting missing spatial data
- General improvement in accuracy and performance of Partial CNN over kriging

METHODS



Partial CNN is based on the U-net architecture (left). During the convolution, the mask shape is excluded, reducing the significance of mask during each encoding phase (bottom). The convolved image is upsampled without the mask.

Fig. 1: U-net style architecture of Partial CNN developed by Liu et al. 2018.

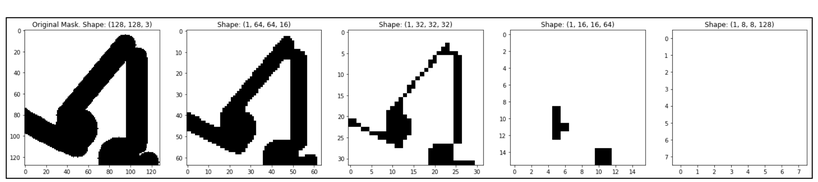


Fig. 2: Mask significance reduction during each encoding phase of the Partial CNN model.

RESULTS

The performance of Partial CNN and Kriging have been compared using randomly selected CMAQ AOD images with intended randomized missing area(s) representing the mask of missing data. Partial CNN achieved a 2% improvement in mean Index of Agreement (IOA) and Pearson correlation coefficient (COR) with less variance in accuracy than Kriging. Partial CNN achieved lower bias in predicting missing data with an average Mean Absolute Error (MAE) of 0.05, while Kriging achieved an average MAE of 0.06. Performance of Partial CNN and Kriging in Root Mean Square Error (RMSE) were relatively similar in average performance. Partial CNN achieved less variance in prediction accuracy in all performance evaluation cases (see Figure 3).

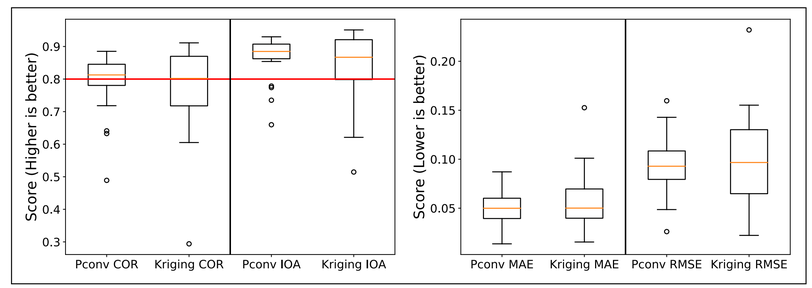


Fig 3: Overview of overall performance of Partial CNN and Kriging when predicting missing data based on intended randomized missing area(s).

Partial CNN does not experience interpolation errors as Kriging when missing data is located at the borders of images (Fig 4). Hence, the stability in the overall performance of Partiall CNN over Kriging.

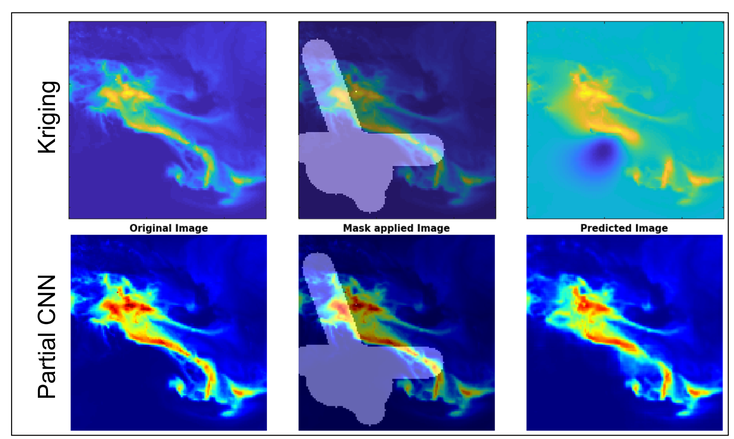


Fig 4: Sample case of Kriging and Partial CNN predicting same AOD images with same mask shapes. Kriging suffers in performance due to missing data at border of image.

RESULTS

Conclusion: Process performance of Partial CNN is highly beneficial over Kriging. Once Partial CNN has been trained, the model can predict images much quicker than Kriging. In our research, Kriging required 20 minutes to process each image. Partial CNN required an estimated 5 hours of training time, but can predict each image in less than 1 minute. Figure 5 shows the estimated time to predict missing data between Kriging and Partial CNN (with training time) for a number of images.

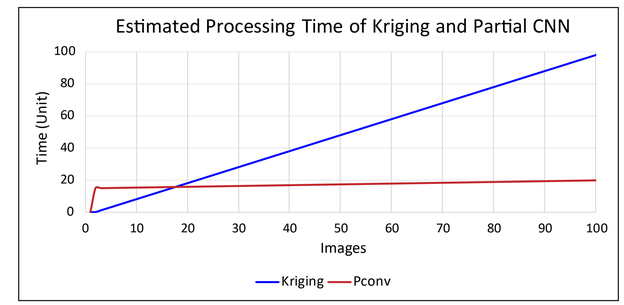


Fig 5: Estimated processing time of Kriging and Partial CNN. Processing time is based on relative Kriging time. While Partial CNN requires considerable time to train, it out-performs Kriging when predicting more than 16 images at a time.

We apply Partial CNN on the GOCI image data over the Korean Peninsula (see Figure 6). Partial CNN is able to reconstruct missing AOD data with IOA generally over 0.8.

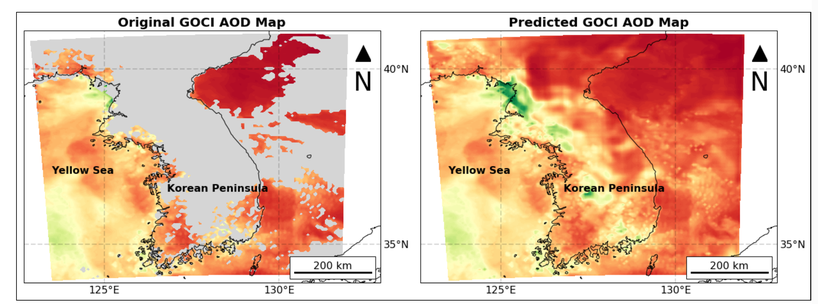


Fig 6: Sample case of Partial CNN predicting (right) missing GOCI AOD (grey) image (left) over the Korean Peninsula.

Acknowledgements:
This research is partially supported by funding from the Department of Earth and Atmospheric Sciences (EAS Research Grant) of the University of Houston.

References:
Liu, Guilin, et al. "Image inpainting for irregular holes using partial convolutions." Proceedings of the European Conference on Computer Vision (ECCV). 2018.