# Advancing Global Cloud Detection in Satellite Imagery with Spatial and Spectral Awareness via Deep Learning

**Joseph Nied, Larry Di Girolamo - Department of Atmospheric Science, University of Illinois at Urbana-Champaign**

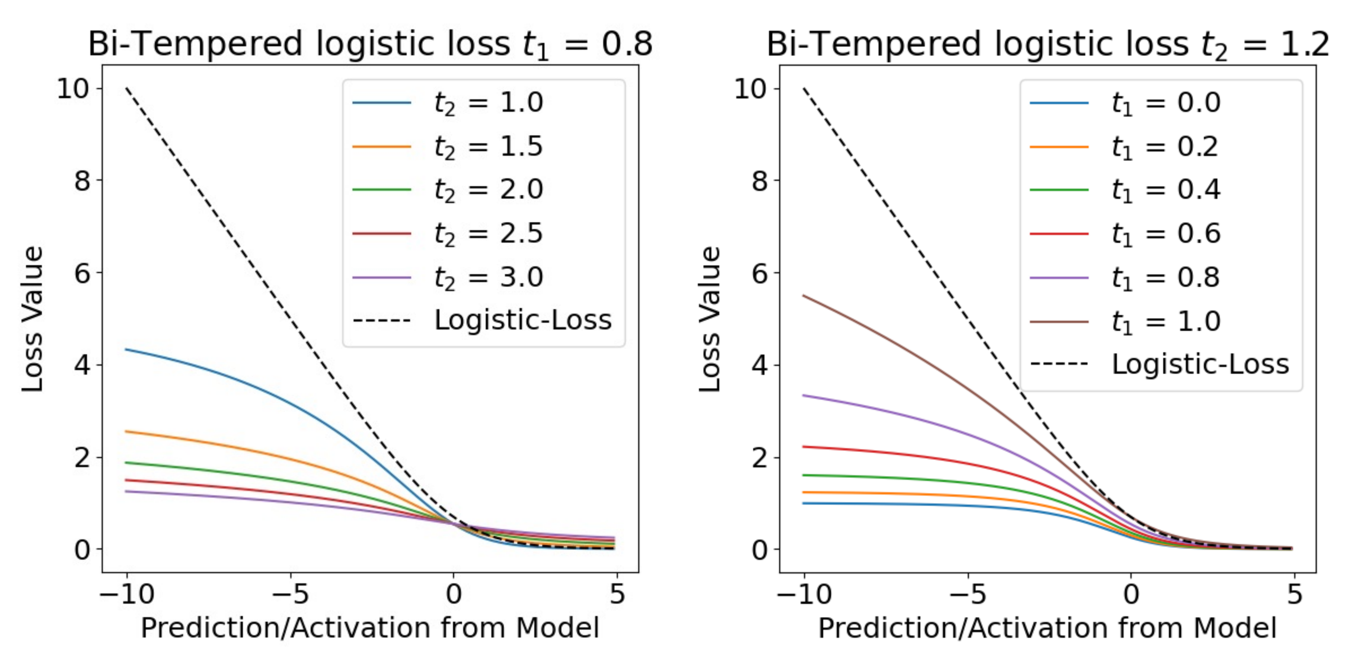
## PROBLEMS

- In remote sensing, one of the first processing levels that is performed is the cloud mask.
- Errors in the cloud mask propagate to uncertainties in geophysical data products (SST, AOD, etc.).
- Aboard TERRA, the Moderate Resolution Imaging Spectrometer (MODIS) and the Multi-angle Imaging SpectroRadiometer (MISR) observe earth's surface and atmospheric systems across multi-spectral bands.
- Cloud masks must fundamentally be purpose-driven in their design (Yang and Di Girolamo 2008). MISR aims for a clear-sky conservative cloud mask, while MODIS aims for a more cloud-fraction conservative cloud mask. Due to this there are disagreements between the cloud masks.
- It is not yet known how to design an algorithm of achieving multi-purpose cloud masks that are easy to train and are globally applicable.

## HOW TO IMPROVE

- A.I. and deep learning have made advancements to exploit data to solve non-linear problems.
- It is unclear how to operationalize these advancements for the purposes of global cloud detection.
- Deep learning models require exhaustive human labeled datasets.
- It would be beneficial to develop a modeling structure & training process that is robust to label-noise.
- For cloud detection, a model structure needs to be identified and examined to ensure it is robust to the robust to the wide range of cloud, surface and sun-view conditions.
- A model structure that is robust for cloud detection would be able to use both spatial and spectral patterns of clouds to decide the class confidence of a pixel.

## BI-TEMPERED LOGISTIC LOSS




- Amid et al. (2019) proposes a new loss function based on Logistic-Loss but adding in tempered variables to dictate how quickly datapoints are learned, which provides robustness to mislabelled datapoints.
- $t_1$ relaxes impacts of datapoints far from the decision boundary.
- $t_2$ relaxes impact of datapoints near the decision boundary.

We replace exponentials and logarithms with the following tempered versions:

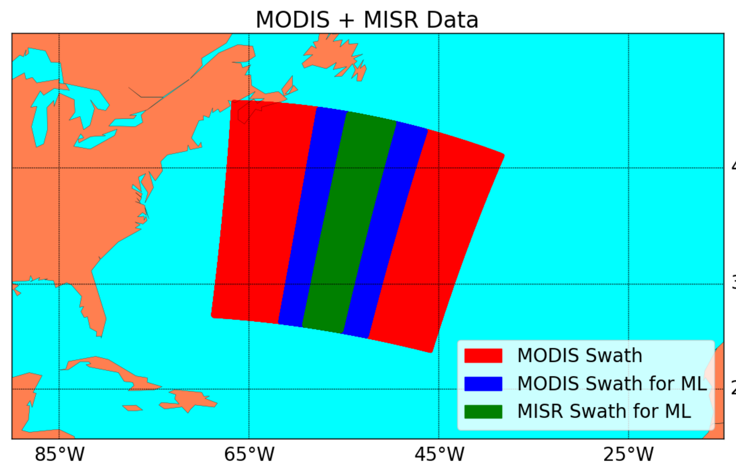
$$\log_{t_1}(x) := \frac{1}{1-t_1}(x^{1-t_1} - 1) \qquad (1)$$

$$\exp_{t_2}(x) := [1 + (1-t_2)x]_+^{\frac{1}{1-t_2}} \qquad (2)$$

As limit of $t_1$ and $t_2$ approach 1, the original logistic loss is recovered.

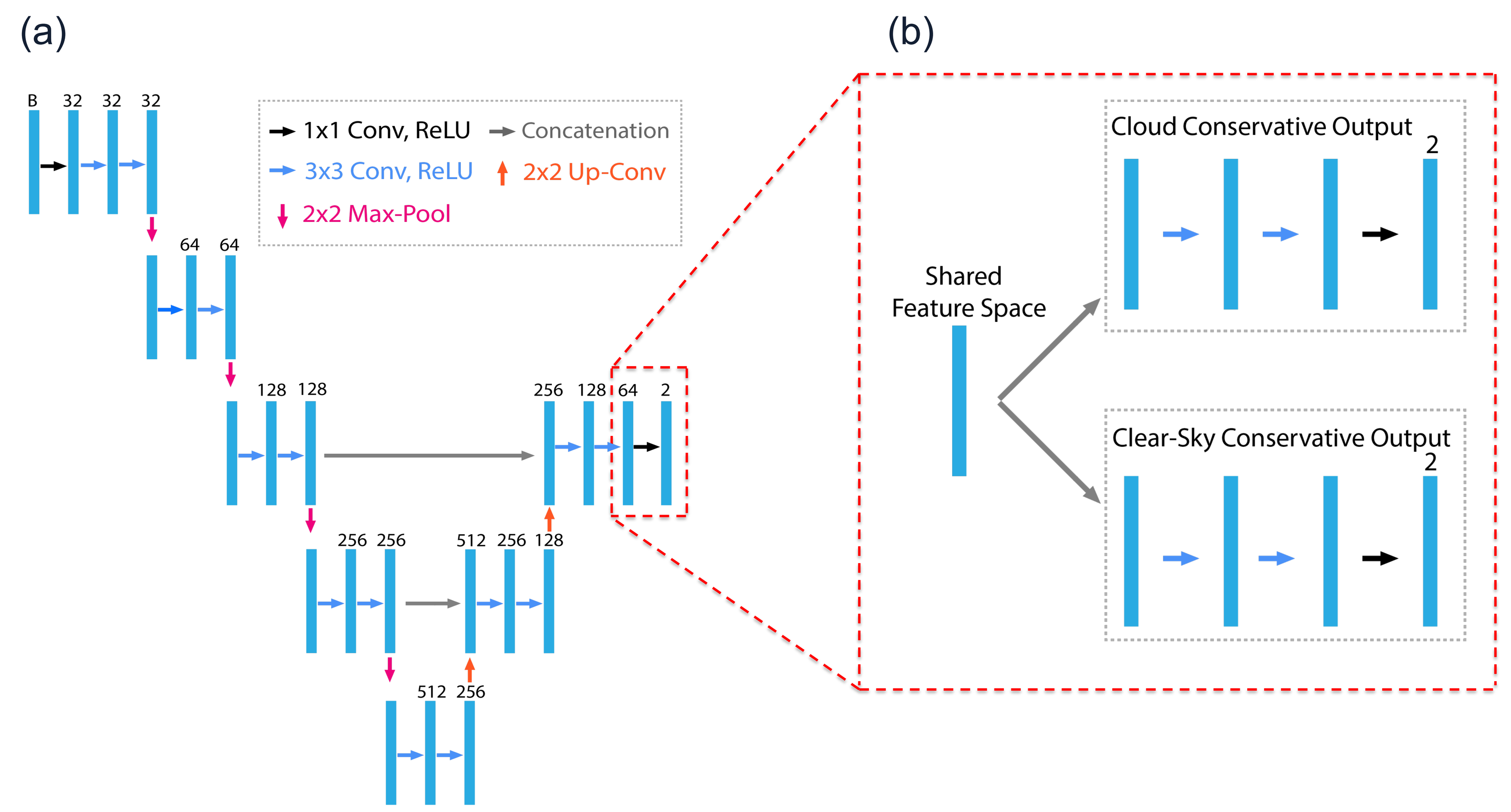
## REMOTE SENSING DATASET

| Instrument | Center Wavelengths | Spatial Resolution | Cloud Mask | Sampling Resolution | Training Data |
|---|---|---|---|---|---|
| MODIS | 0.65, 3.75, & 11 µm | 250 m, 500 m, 1 km | MOD35 (1 km) | 64 x 64 km | 634 total – 26 Cumulus Samples |
| MISR | 446, 558, 672, & 867 nm | 275 m | RCCM (1.1 km) | 281.6 x 281.6 km | 22 Cumulus samples |




*Training data for MODIS was obtained by manually selecting MODIS cloud masks to obtain examples of where the algorithm performs well and poorly. In total, 634 64x64 km scenes were obtained. From this dataset, 26 samples were selected, that particularly contained cumulus and sub-pixel clouds. These examples were used to assess the effects of Bi-Tempered Logistic Loss for training, as well as to observe the advancements of using 250 m resolution data. Of these 26 cumulus samples, 22 of them were obtained from MISR.*

## NEURAL NETWORK MODEL




*(a) U-Neural Network Structure used to train MODIS & MISR Models. Blue boxes denote learned feature channels for cloud and clear-sky textures. Numbers above each box denote the number of feature channels produced. The B above the first layer denotes the input imagery, and the number of bands provided.*

*(b) Multi-output augmentation of the U-Net model architecture. This allows for both MODIS or MISR to train and produce cloud & clear-sky conservative cloud masks from a single model. The dashed-red line depicts where this would be substituted into.*

### U-NEURAL NETWORK MODEL:

- Ronneberger et al. (2015) developed the "U-NET" model, which uses series of convolutional layers, pooling, up-convolutions, and concatenation to detect and decipher objects in imagery based on multi-scale textures & spectral information.
- Multi-resolution 'layers' of the model allows for direct assimilation of multi-resolution spatial textures to be exploited directly, such as the case with MODIS, unlike previous cloud masking algorithms.
- The model's large parameter size allows for many cloud features to be learned and retained which helps the model be robust for global cloud detection.

### MULTI-OUTPUT MODELING:

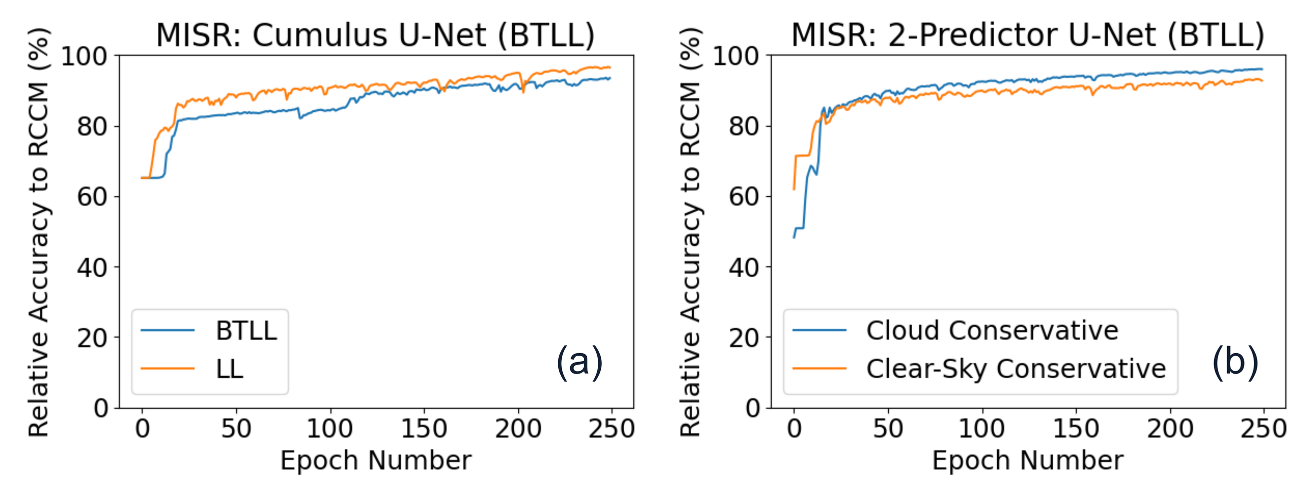
- Multi-output neural networks have increased in popularity to explore shared back-propagation learning, across classification problems (Xu et al. 2020).
- From the U-Net model, we can augment the output path to learn two cloud masks in parallel, while sharing a 'body' model of shared cloud & clear-sky textures.
- By doing this, cloud & clear-sky conservative cloud masks can be obtained for a single instrument in a shared model. This allows us to retain confidence in both methodologies for a single instrument.

## ONLINE RESULTS

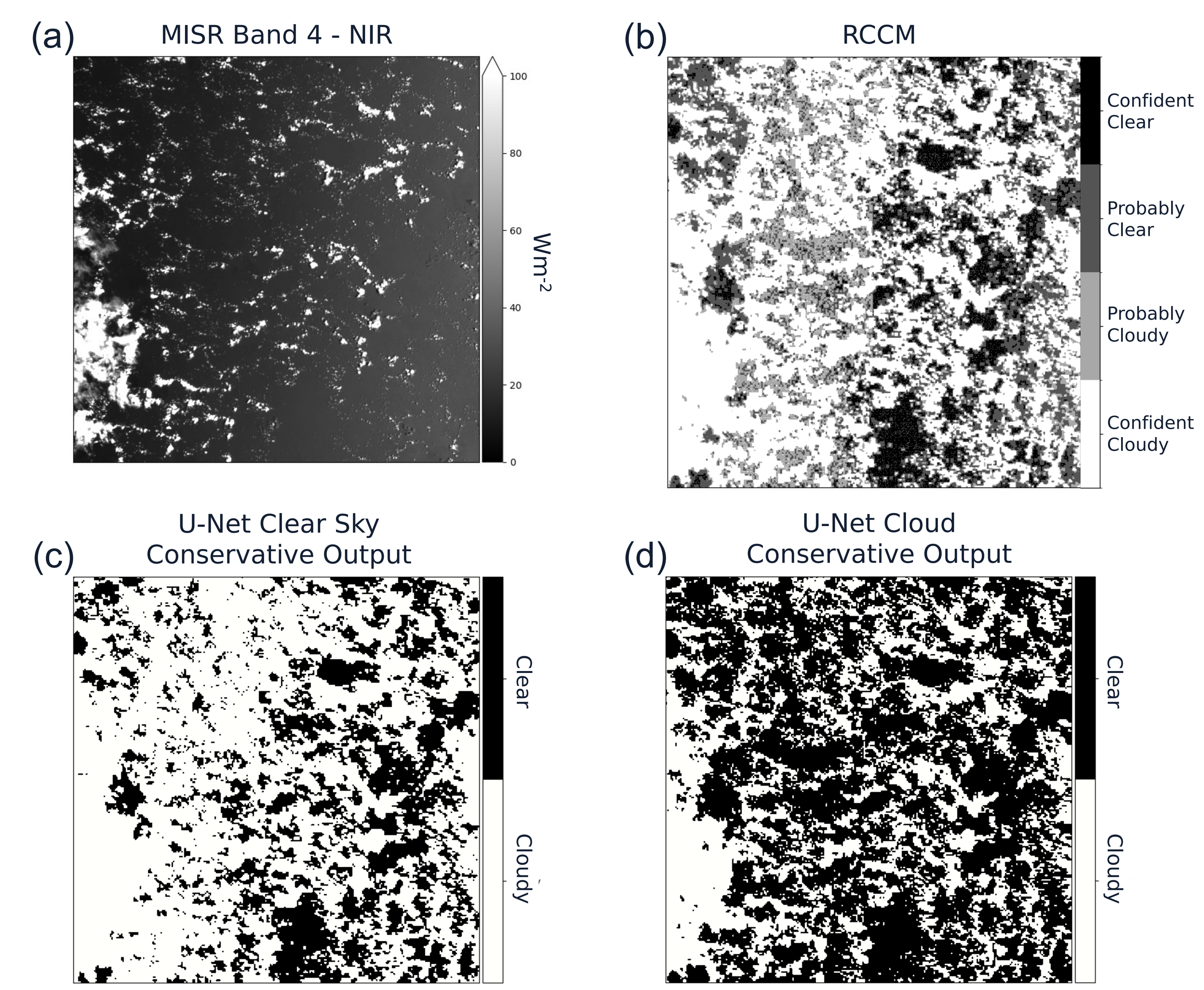
For animations depicting the learning process of our various models, please visit this GitHub:

https://github.com/jdn8608/AGU-2023-Joseph-Nied-Poster/

## RESULTS




*(a) Training accuracies for 2 U-Net models, weights initialized the to be the same, using Logistic-Loss or Bi-Tempered Logistic-Loss, relative to the RCCM cloud mask.*
*(b) Training accuracies for a single U-Net model, with the multi-output augmentation, to produce 2 separate cloud masks. Accuracies are relative to the binary binning of the RCCM to be more cloud or clear-sky conservative.*




*Visual comparison of (a) MISR NIR imagery, (b) the MISR RCCM cloud mask at nadir, the (c) clear-sky & (d) cloud conservative cloud mask produced by the same multi-output U-Net, trained with Bi-Tempered Logistic Loss. The RCCM was used to produce a cloud conservative cloud mask by grouping cloudy as a binary cloud classification and grouping all other labels as clear-sky. The opposite grouping was done for a clear-sky conservative mask. These masks were used as the ground truth during training for each output.*

## CONCLUSIONS

- Bi-tempered Logistic Loss parameterization can be used to train neural networks with unsupervised labels in training.
- U-Nets can exploit spatial and spectral textures across multiple resolutions to allow for large- & small-scale features of clouds.
- Multi-output neural network models gives a unique solution to gain confidence in cloud masking by producing ranging confidence cloud masks for a single instrument, that can be used by the scientific community.

## FUTURE WORK

- Compare different neural network algorithms from literature with current modeling techniques, to establish confidence in our methodology.
- Investigate how transfer-learning techniques can be utilized to translate learned cloud features for future instruments.
- Re-analyze how we use multi-output modeling, to be more robust to clear or cloud sky cloud masking purposes across multiple instruments.
- Incorporate geometric data, such as sun-view geometry, latitude & longitude, and time of year, into model creation and training to achieve optimizations of the model for global, year-round cloud detection.
- Expand training data to unsupervised MODIS and MISR data to achieve different purpose-driven cloud mask outputs under a single unified A.I. framework for operational implementation and traceability, and whose quality is robust to the inherent label-noise of the various MODIS and MISR cloud masks.


## ACKNOWLEDGEMENTS

Partial support for this research was provided by the MISR and MAIA projects under contracts 147871 and 1586704, respectively, with the Jet Propulsion Laboratory, California Institute of Technology.


## References


1. Yang, Y., and L. Di Girolamo (2008), Impacts of 3-D radiative effects on satellite cloud detection and their consequences on cloud fraction and aerosol optical depth retrievals, J. Geophys. Res., 113, D04213, doi:10.1029/2007JD009095.
2. Amid, E., Warmuth, M. K., Anil, R., & Koren, T. (2019). Robust Bi-Tempered Logistic Loss Based on Bregman Divergences. (arXiv:1906.03361). arXiv. https://doi.org/10.48550/arXiv.1906.03361
3. Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation* (arXiv:1505.04597). arXiv. https://doi.org/10.48550/arXiv.1505.04597
4. Xu, D., Shi, Y., Tsang, I., Ong, Y., Gong, C., & Shen, X., (2020). "Survey on Multi-Output Learning," in IEEE Transactions on Neural Networks and Learning Systems, vol. 31, no. 7, pp. 2409-2429, July 2020, doi: 10.1109/TNNLS.2019.2945133.