Deep Learning to Enhance Anomaly Detection in NOAA Geostationary Imagery

Haibing Sun[1], Rebekah Esmaili[2], Elizabeth Kline[3] ,Walter Wolf[2], L. Soulliard [1]

[1]GAMA-1 Technologies, Greenbelt, MD, USA [2] NOAA/NESDIS/ *Office of Common Services*, Silver Spring, MD, USA [3]NOAA/NESDIS/ *Geostationary Earth Orbit Observations*, Greenbelt, MD, MD, USA

SUMMARY:

The anomalies occur in satellite GOES ABI observation and NPP/J1/J2 VIIRS observation. Those image level anomalies are not sufficiently flagged in the pixel level QA and there is no enough information can be provided to downstream users. Some of these anomalies are predictable based on satellite maintenance schedules, observation time et al, but others may occur randomly. Many of these phenomena are so rare that most users, even ones who frequently analyze ABI imagery such as National Weather Service (NWS) field office forecasters, may not notice them. Unexpected and widespread data anomalies can disrupt NOAA operations and significantly impair hazard monitoring. The tradition method to detect and determinate of the anomalies is done by the manual image analyzing and with the help of ancillary data, inevitably this process is prone to errors and very complex, laborious, and time-consuming. An Automatic anomalies identification and classification method for NOAA satellite imagery is gaining interest day by day. And this project is now to develop a multilabel model for anomaly detection and classification basing on deep learning.

METHODOLOGY

Deep Learning GOES Image anomaly Classification

We define GOES-R image anomalies as widespread data quality degradation. Some type anomalies occur daily, some of which are expected and fleeting while others are unexpected and can persist for multiple images. Some causes include interference in the downlink from the satellite to the ground station or the transmission from the ground station to forecaster offices via the satellite broadcast network (SBN). Sensor malfunctions can cause persistent anomalies.

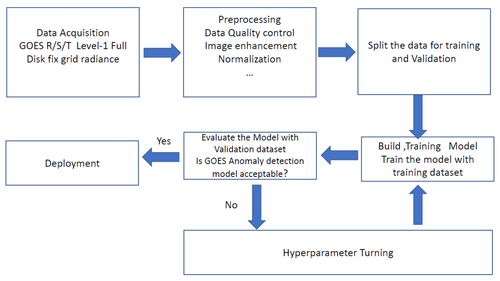
The tradition method to detect and determinate of the anomalies in GOES image is done by the manual image analyzing and with the help of ancillary data, inevitably this process is prone to errors and very complex, laborious, time-consuming and challenging. The advent of deep learning, in particular, Convolutional Neural Network (CNN), provide opportunities for computer vision-based data analysis and to overcome the problems associated with the manual analyzing. The CNN-based deep learning approaches have been widely adapted for an array of tasks such as image classification, object detection, object semantic segmentation, and pose estimation. Residual Network (ResNet) is a Convolutional Neural Network (CNN) architecture that overcame the "vanishing gradient" problem, making it possible to construct networks with up to thousands of convolutional layers, which outperform shallower networks, leading to better accuracy on most image classification tasks, but often requires more computational power to train. In the first stage of GOES anomaly detection model development, A binary anomaly detection successfully identified the existing of anomalies in ABI imagery using a ResNet-18 model with 99.2% accuracy. In present anomaly classification detection project, ResNet will be continuous used.

Transfer Learning in GOES Image anomaly Classification

It is relatively difficult to have a GOES anomaly training dataset of sufficient size. To avoid the problem from not sufficient training dataset size, transfer learning can/is adopted in GOES-R anomaly classification deep learning processing. The pretrained ResNet on ImageNet(which contains 1.2 million images with 1000 categories) is used in GOES anomaly detection transfer training. The pretrained ResNet either as an initialization or a fixed feature extractor and both are installed in GOES anomaly detection deep learning. Method 1: Finetuning the ConvNet: Instead of random initialization, initialize the network with a pretrained ResNet network. Rest of the training looks as usual; Method 2: Pretrained ConvNet as fixed feature extractor: Here, freeze the weights for all of the network except that of the final fully connected layer. This last fully connected layer is replaced with a new one with random weights and only this layer is trained.

GOES anomaly deep learning workflows

The main phases in GOES anomaly deep learning workflows include data collection, data pre-processing, building datasets, model training and refinement, evaluation, and deployment to production.



Data Acquision: ABI-L1b-RadF: GOES 16 /ABI: 2017-2024 GOES 18:2022- 2024. (aws, S3 Bucket) ABI-INST-CAL: (The STAR Central Data Repository(SCDR) , Anomaly Labeling &Quality control). GOES ABI Calibration Events Log, Anomaly Labelling)

Preprocessing: Convert ABI 16 band radiance to PNG and normalization.

Data Quality control: human identification, cross-check with DQF ,, GOES Event Log , Calibration data , related document and paper et al.

Image enhancement: Option for Lunar intrusion Anomaly

Hyperparameter Turning: ResNet number of layers, Learning rate, epoch number et al.

haibing.sun@noaa.gov

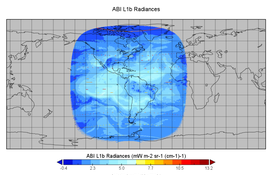
Geostationary Imagery : Geostationary Imagery Anomaly

The Geostationary Operational Environmental Satellite (GOES) – R Series is the nation's most advanced fleet of geostationary weather satellites. Designed to operate in geosynchronous orbit, 35,786 km above the equator, thereby remaining stationary relative to the Earth's surface. The Advanced Baseline Imager (ABI) is the primary instrument on the GOES-R series spacecraft for imaging Earth's weather, oceans and environment. ABI views the Earth with 16 different spectral bands, including two visible channels, four near-infrared channels, and ten infrared channels, ABI provides critical atmospheric, hydrologic, oceanic and climatic data, significantly improving the detection and observation of environmental phenomena. With the ABI present operational scanning routines(Mode-6), ABI acquires one full disk image (FD), two contiguous United States images (CONUS), and 20 1000 km by 1000 km mesoscale (MESO) images in 10 minutes. The Full Disk is defined as a circle, with a 17.4 degree angular diameter from the perspective of the ABI centered at the instrument's nadir that reaches the Earth's limb. The ABI observations for atmosphere-earth target, internal blackbody target, space and stars looks are transmitted to ground station and radiance calibrated, the Calibrated detector values are navigated to Earth location, re-sampled to form pixels and product are generated in the GOES fixed grid angular coordinate projection. The 16 bands Full Disk ABI Level 1b fixed grid products are stored in NETCDF format.

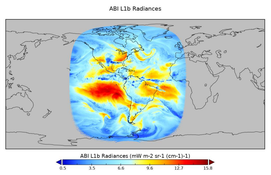
Unexpected and widespread data anomalies can disrupt NOAA operations and significantly impair hazard monitoring. The anomalies in ABI radiance image is result of anomalies of satellite platform, ABI instrument, ground processing or interference/signal contamination from un-expected target. Some of these anomalies are predictable based on satellite/instrument maintenance, ground processing algorithms schedule or time of year, but others may occur randomly. Many of these phenomena are so rare and very difficult to be identified. Rapidly identifying an ongoing anomaly in GOES satellite is important for monitoring data outages. OSPO provides alerts for satellite data outages to inform data users through email and a web portal. The data quality flags identify the image pixel level quality, however it can not identify patterns and trends of anomalous data in image. At the same time, Some anomaly type cannot predicted, or are difficult to be identified and thus not be reported in present web portal and not with the correct quality flags.

Geostationary Imagery Anomaly Target

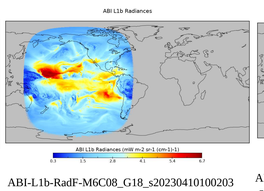
For ABI FD imager on aboard active GOES satellites(GOES16/18), the unpredictable and no specified flag Anomalies include: Bad Detector Stripes ; Latched Detector Response Stripes; Lunar Intrusion into Spacelooks Type I&Type 2; Incomplete images; Bright Object Avoidance ("Cookie Monster"); Stray light; Stray light in the visible/NIR/IR channel for earth /space look ; Nonuniform detectors responds Stripes; Cold Pixels Around Fires; Solar Farms and Ringing; Solar radio frequency interference (RFI); Caterpillar Tracks and Shark Fins; HBT Thruster Flush; Yaw Flip. Figure 1-6 give several types of anomalies which on DQF is not set and /or No specified DQF flags are defined .



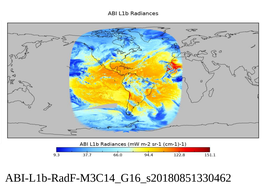
ABI-L1b-RadF-M6C08_G16_s20200631030173
Bad Detector stripes DQF = 0 Fig 1



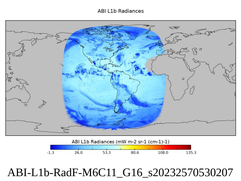
ABI-L1b-RadF-M3C09_G16_s20181550715491
Latched Detector Response Stripes DQF=0 Fig 2



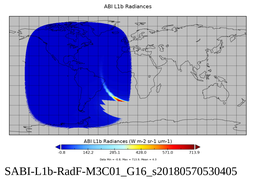
ABI-L1b-RadF-M6C08_G18_s20230410100203
Lunar Intrusion into space look DQF=0 Fig 3



ABI-L1b-RadF-M3C14_G16_s20180851330462
Shark Fin(Missing data) Pixel level DQF=-1 Fig 4



ABI-L1b-RadF-M6C11_G16_s20232570530207
Bright Object Avoidance DQF=-1 Fig 5



SABI-L1b-RadF-M3C01_G16_s20180570530405
Stray Light Pixel level DQF=-1 /2 Fig 6

Training dataset

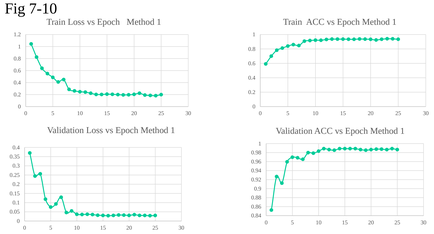
Deep learning requires labeled data for learning. Supervised learning is the process of gaining knowledge or extracting insights from data labeled by humans. Image classification, particularly supervised classification, is also reliant hugely on the data fed to the algorithm. A well-optimized classification dataset works great in comparison to a bad dataset . with data imbalance based on class and poor quality of images and image annotations.

Stripes	Fatnt broad stripes from Lunar intrusion	Incomplete Image (Losing data in sampling , processing and Telecommunication)	BOA related Incomplete Image	Signal contamination	Artificial in Ground processing
Bad detector 1605	154 , Lunar intrusion (Type 1 SPL)	Drop block 1000+	Bright Object Avoidance, 1000+	Stray Light 300+	
Latch-up artifact 1180	30+ , Lunar intrusion (Type 2 SPL)	Drop swath in Level 1 processing 300+			
		Losing section for other reason. (Section losing, Multiple swath) 300+		Solar Farms and Ringing 5-10	Cold Pixels Around Fires
		Caterpillar Tracks 50+ Shark Fins 30+			Solar Farms related CPAF
		RFI missing data (Missing data)			Other related CPAF

Table 1: GOES-R ABI Anomaly detection & classification Training Dataset .

Lunar intrusion/latched detector response/Bad detector /Normal observation classification.

Method 1 Fig 7-10



Training labelled dataset:

1: GOES-R ABI FD radiance: Data from NOAA's GOES-R series satellite is available on Amazon S3.

2: Anomaly Classification Label :

2.1: Basing on the GOES ABI Calibration Events Log (Lunar intrusion stripes/ Latched Detector Response Stripes/ Bad Detector Stripes

2.2 Basing on GOES ABI DQF pattern analyzing.(Incomplete images , Bright Object Avoidance , Stray light et al

2.3 Basing on ABI Calibration/Engineer file information.

2.4 Manual image analyzing.

3: Predictable Anomalies and very rare Anomalies are not include in present training dataset.

Conclusion:

Preliminary test show that Deep Learning can enhance Anomaly Detection. Figure 7-10 give the train loss, and classification accuracy for training dataset and validation dataset. This training and validation are for anomalies related to lunar intrusion, latched detector response and Bad detector stripes which are difficult to be identified with human analyzing.

Multiple label anomalies, not enough training data for several types of anomalies and data imbalance effect are the potential problems existed and need more attention.

There exist some type of "un-known" Anomalies. The connection between the Anomalies and the factors that generate the anomalies need to be established