An Offline EO Data processing Challenge using Open source packages

Automatic CLOUD and SHADOW mask generation from Resourcesat-2/2A Liss4 Satellite Images

Expected Participants :

1. Earth observation data processing Enthusiasts developing Computer vision and AI-ML based algorithms .
2. Students and Researcher's from Universities, Academic Institutions and enthusiastic developers.
3. Team of 2-4 .

Guidelines on procedure :

1. On registration you will be get a registration Id and details of FTP Login to download Data to generate Training sets . Evaluation Test data will be provided in the last 7 days to generate output.
2. Ensure code gets executed in an independent environment with all opensource dependent packages. Any opensource AI-ML , Computer vision image models/ libraries can be used .
3. A mentoring WebEx session will be scheduled based on requirement for the registered teams .
4. **On Completion of the Challenge** , the team member should send email to coding_challenge@nrsc.gov.in , after which a new Login will be provided to upload the outputs .
5. The shortlisted teams need to give a presentation to NRSC experts.
6. The shortlisted teams may be asked to demonstrate the execution of the code at NRSC.
7. NRSC Decision on evaluation is final.
8. For any Queries feel free to mail: coding_challenge@nrsc.gov.in

About the Challenge:

Clouds and Shadows hinder the analysis of optical satellite images as it obscures the capture of features of the Earth's surface. Clouds appear bright white in RGB satellite images and delineating clouds from look-alike features like snow, buildings tops, deserts are a challenge. Similarly delineating Cloud shadows from water bodies, building shadows, etc are also challenging. Another challenge for Liss-4 sensor Image is it does not have a SWIR band. Generally, the steps for the solution are

- Convert DN to reflectance
- Sun incident angle corrections
- Generate labelled training sets by Thresholding on various indices or band ratios, Morphology based, Machine learning and Deep Learning using labelled datasets or pre-trained models

But you can come up with any other innovative methodology.

Expected Output:

- Model should be trained and validated with evaluation matrices i.e., IOU, F-1 Score, Precision, Recall and accuracy.
- Model output on test dataset should be a georeferenced tiff file (8 bit) registered with the input image with following class values.
  - NOCLOUD: 0
  - CLOUD: 1
  - SHADOW: 2

- CLOUD and SHADOW mask as shapefiles (ESRI format: .shp , .shx, .dbf,.prj)
- Trained model weights and test data inference code along with help in Report.pdf should be provided to run the inference code in an independent computer system.
- Model training graphs on training and validation data with all evaluation metrices should be shared.
- Model should be trained and optimized on diverse datasets (different seasons, different terrains – snow clad mountains, deserts, water bodies, forests, crops,.) to have better generalizability.
- Using the FTP Login(provided after completion of Project) to Upload the output.zip/ .tar having Report , results and Jupyter notebook .

What you will get on completion of Challenge:

1. Participation Certificate for all team members who submit the project.
2. Award certificates and Mementos to winners .
3. Chance to present to ISRO experts.

About Input dataset:

1. 20 Data sets for model generation - Resourcesat-2/2A- Liss-4 Georeferenced Level-2 data with 5.8 m – 3 bands.
   "Training Data "
   Eg : RAF25JAN20250422200097000 55SSANSTUC00GTDB
   "Browse "having browse jpegs of all the sets for quick view.
2. Help document (TOA_help.zip) about the parameters required for generation of TOA reflectance images (Solar Elevation, Solar Exo-Atmospheric Irradiance and Earth-Sun Distance)
3. Additional data from Resourcesat-2/2A-Liss4-FMX can be downloaded from Bhoonidhi portal (bhoonidhi.nrsc.gov.in) which are Free. Also, Cartosat-1 30m DEM is available from Bhoonidhi for free. Any issues mail to coding_challenge@nrsc.gov.in .
4. 10 test datasets will be shared in the last week of the challenge to be used for evaluation only for submission of final output.

Software Stack to be used: Python , GDAL / Roserio, OpenCV / Scikit-image, PyTorch,Tensorflow ,(any other open sources packages to be mentioned in resource.txt) , QGIS (optional for visualization)

Evaluation Method:

- NRSC team shall evaluate the submissions made .
- Evaluation of model on test dataset shall be carried out based on evaluation metrices (IoU (Intersection over Union) , Precision(how many predicted positives are actual positives) , Recall(how many of the actual positives were correctly predicted) , F-1 Score (Harmonic mean of precision and recall) and accuracy(Overall correctness of the model)), expected accuracy is greater than 90% .
- NRSC team shall run inference code on another set of test data to check model performance with evaluation metrices .
- Shortlisting of teams for next round shall be based on model performance on test data shared, model inference on test data at NRSC along with inference code execution time.
- Shortlisted teams shall have to give presentation to NRSC in WebEx mode .
- Also, may be asked to demonstrate the execution of the code at NRSC which will be informed to the team .

- The entire solution should be automated without any manual interventions. It will be executed sequentially for each scene . Processing time will be calculated from execution of first line of the code till the generation of outputs .
- Visual inspections shall also be done to verify .

Submissions Description:

1. Output.zip
   - I. Report.pdf giving the design , methodology, training and validation graphs on evaluation metrices and details of Training Datasets. Training Model Structure:
     Input-Layer1-Layer2-Layer3…. Output
   - II. Model Learning Curves and Optimization parameters used (depending on the model chosen): Training.csv , Train/Test ratio=80/20 or ? , Time taken= (minutes) , Input Patch Size= , No of Epochs= , Soft max function used= , Learning rate=
   - III. Requirements.txt with script to install the packages along with versions .
   - IV. Training_Labeled_data.zip
   - V. Inference Code (Inference_Code.zip) along with help document to run on any test dataset.
     DatasetId= , CloudScore= (% of Cloud pixels), ShadowScore=(%shadow pixels) , IOU= (%) , Precision= (%), Recall= (%), F-1 Score= (%) and Accuracy=(%)
   - VI. Model code (Model.zip) along with help document.
   - VII. For each Test Evaluation Data a folder (DatsetId) to be created containing
     1. Georeferenced tiff(8bit) file(mask.tiff) for all test datasets with class value NO CLOUD:0, CLOUD:1 and SHADOW :2 for each input image .
     2. Shape files for CLOUD and SHADOW for all test dataset as cloudshapes.zip and shadowshapes.zip

     eg RAF25JAN20250422200097000 55SSANSTUC00GTDB
     mask.tiff
     cloudshapes.zip
     shadowshapes.zip

No	Test Evaluation Dataset Id	% Cloud pixels	% shadow pixels	IOU	Precision	Recall	F-1 score	Accuracy
1	RAF25JAN20250422200097000 55SSANSTUC00GTDB							
2								
…								
10								

Output.zip

I. Report.pdf
II. Training.csv
III. Requirements.txt
IV. Training_Labeled_data.zip
V. Inference_Code.zip
VI. Model.zip
VII.

DatasetId-1
1. mask.tiff
2. cloudshapes.zip
3. shadowshapes.zip

,…. DatasetId-10

References

1. https://spatialthoughts.com/2023/12/25/liss4-processing-xarray/
2. Zhiwei Li, Huanfeng Shen, Qihao Weng, Yuzhuo Zhang, Peng Dou, Liangpei Zhang, Cloud and cloud shadow detection for optical satellite imagery: Features, algorithms, validation, and prospects, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 188, 2022, Pages 89-108, ISSN 0924-2716, https://doi.org/10.1016/j.isprsjprs.2022.03.020.
3. Cloud Detection from RGB Color Remote Sensing Images with Deep Pyramid Networks, July 2018,DOI:10.1109/IGARSS.2018.8519570, Conference: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium.
4. S. Singhal, L. James, A. R. V. G, S. C. V, M. K. S and R. R. Nidamanuri, Cloud Detection from AWiFS Imagery using Deep Learning," 2023 International Conference on Machine Intelligence for GeoAnalytics and Remote Sensing (MIGARS), Hyderabad, India, 2023, pp. 1-4, doi: 10.1109/MIGARS57353.2023.10064610.
5. https://bhoonidhi.nrsc.gov.in/bhoonidhi_resources/help/docs/User-Guide-Surface_reflectance_from_Resourcesat-2A.pdf
6. https://bhoonidhi.nrsc.gov.in/bhoonidhi_resources/help/docs/R2_data_user_handbook.pdf

About Resourcesat-2/2A

- Payloads/ Sensors: Optical: Awifs , Liss-3 , Liss-4
- Imaging Modes : Multispectral - AWifs & Liss3 Nadir viewing , Liss-4 Tiltable (+/- 26 deg across track)
- Spatial Resolution(m): 56 m , 23.5 m , 5.8 m
- Spectral band (µm) : AWifs/Liss-3:VNIR (3 bands)+ SWIR (1 band), (G-B2 0.52 - 0.59 ,R-B3 0.62 - 0.68 ,NIR-0.77 - 0.86 ,SWIR-B5 1.55 - 1.70), Liss-4: VNIR (3 bands) Liss-4 (MX (G- 0.52 - 0.59 ,R-B3 0.62 - 0.68 ,NIR-0.77 - 0.86) / Mono (R-B3 0.62 - 0.68))
- Swath(km) : AWifs(A:370 + B:370)km, Liss3: 140 km , Liss-4: 70 km, Fixed Referencing (Path/Row)
- Temporal Resolution: 5 days , 24 days , 58 days
- Coverage : India and Surroundings
- Data availability: Resourcesat-2 : since May 2011 , Resourcesat-2A : since Dec 2016

  https://bhoonidhi.nrsc.gov.in/bhoonidhi/index.html
- Data Accesses: Free & Open, Within 4 - 8 hours of acquisition /Request placing
- Image Acquisition Plan : 10.30 am Equatorial node crossing , Automatic Systematic planning , Request based for CAL-VAL , IGS , Global regions
- Data products : (Geotiff) OrthoProducts (without GCP)
- Similar missions : IRS-1A/1B /1C/1D , Sentinel -2, Landsat-8/9 , Upcoming : Resourcesat-3/3A
- Major applications: Land, Agriculture , Water , urban , coastal studies

---