Real-time detection of clouds on high-resolution satellite videos using deep learning

Panagiotis Sidiropoulos, Eduard Vazquez and
Alistair Francis

CORTEXICA

AI SOLUTIONS FOR BUSINESS

AI for EO

- ISCF SDR draft:
  - *"…fostering meaningful collaborations within the AI and space industries at the required scale to create national and international impact that would otherwise be impossible…"*
  - *"…there has been a lack of awareness and appreciate from both the AI and EO industry of each other's maturation and capability."*
- AI: a lot of PoCs, not a lot of products
- Solution for AI professionals: (If necessary) Scale down your ambitions, scale up your effectiveness

CORTEXICA

AI SOLUTIONS FOR BUSINESS

On-board Cloud Detection on High-Resolution Videos

Desirable Characteristics

- False Positive Rate: <5%
- False Negative Rate: <10%
- Classification of clouds on cloud types
- Computational Speed: 50-200 Mpixels/second
- Consistent performance in all types of terrain
- Reliable performance across all parts of the world
- Light implementation for on-board processing

- <u>Not within the current reach of the AI SoA</u>

- Two alternatives:
  - Try to tackle the problem as is, fail by a mile, write a paper, put all the good results to a website, go to the next project
  - Scale down the problem until the development of a useful tool becomes realistic with the SoA, progressively scale it up

CORTEXICA

AI SOLUTIONS FOR BUSINESS

On-board Cloud Detection on High-Resolution Videos

Desirable Characteristics

- False Positive Rate: <5%
- False Negative Rate: <10%
- Classification of clouds on cloud types
- Computational Speed: 50-200 Mpixels/second
- Consistent performance in all types of terrain
- Reliable performance across all parts of the world
- Light implementation for on-board processing

- Not within the current reach of the AI SoA

- Two alternatives:
  - ~~Try to tackle the problem as is, fail by a mile, write a paper, put all the good results to a website, go to the next project~~
  - Scale down the problem until the development of a useful tool becomes realistic with the SoA, progressively scale it up

CORTEXICA
AI SOLUTIONS FOR BUSINESS

Our Approach

Stage 1: Cloud Coverage Estimation	
Single Number	Downlink or not image

↓

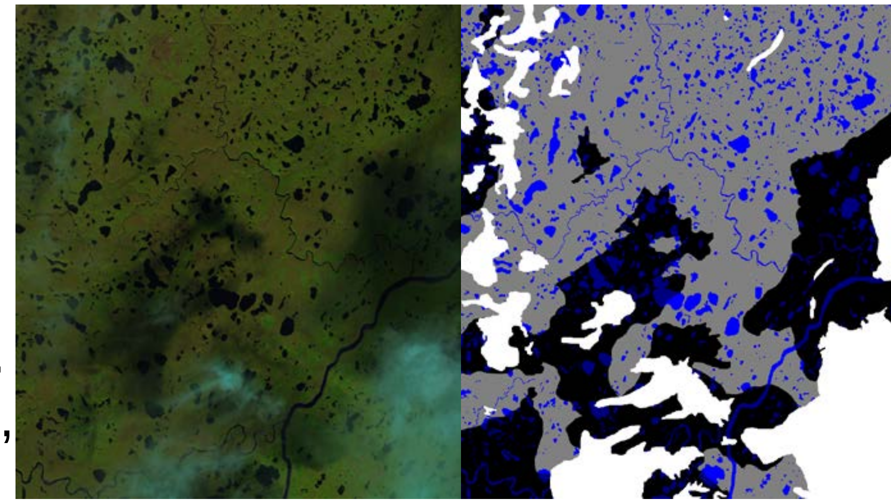
Stage 2: Visibility Masks	
Binary Mask (Visible or Not)	Partial downlink/compression

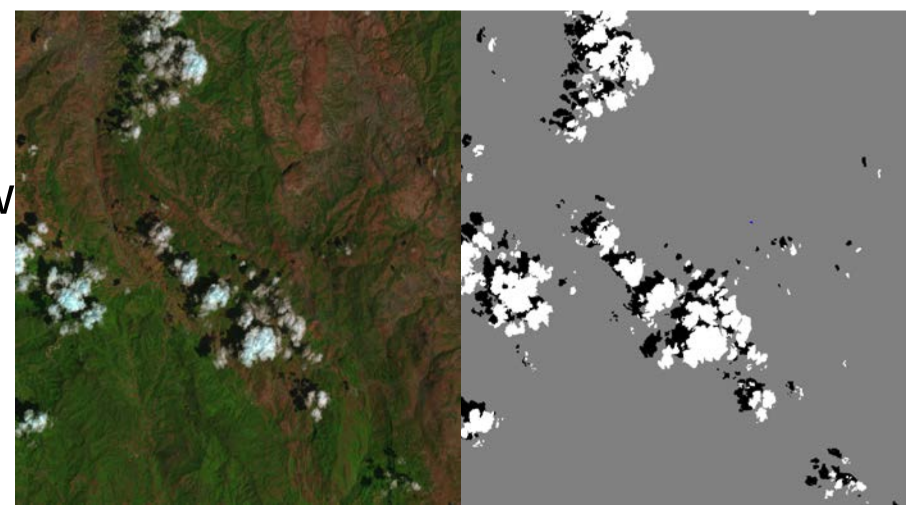
↓

Stage 3: Cloud Masks	
Mask with multiple classes	Information about cloud types, height etc.

SPARCS Dataset: Preliminary Development

- Created by M. J. Hughes, released by USGS in 2016
- Comprises 80 scenes from LANDSAT-8, 1MPixel each
- 7 classes in masks, including: cloud, cloud shadow, land, snow, water etc.
- Characterised by mostly rural terrain, with various cloud morphologies and illumination conditions
- LANDSAT-8 has 10 bands in NIR and visible
- We restrict data to only RGB preview images, to simulate RGB camera mode.
- Resolution of 30m per pixel

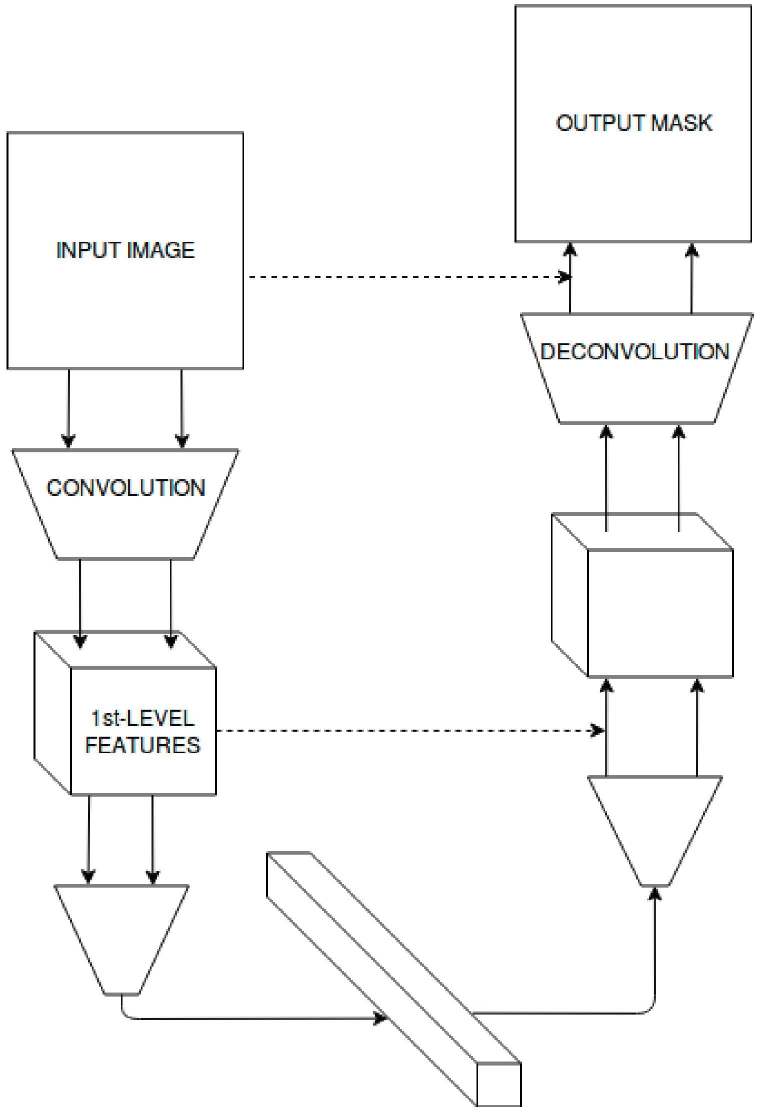




CORTEXICA

AI SOLUTIONS FOR BUSINESS

Deep-Learning Algorithm: Preliminary Development



- Transforming Autoencoder architecture:
  - Features extracted across image through convolution, and then decoded into class map
  - Inspired by Fully Convolutional Networks like U-Net
- Normalisation of all dataset images to a common range (no saturation)
- Can take arbitrary image dimensions as input
- Residual connections in network allows for low-level colour information, as well as deep features to inform each pixel's predicted class

CORTEXICA

AI SOLUTIONS FOR BUSINESS

SPARCS Dataset Results

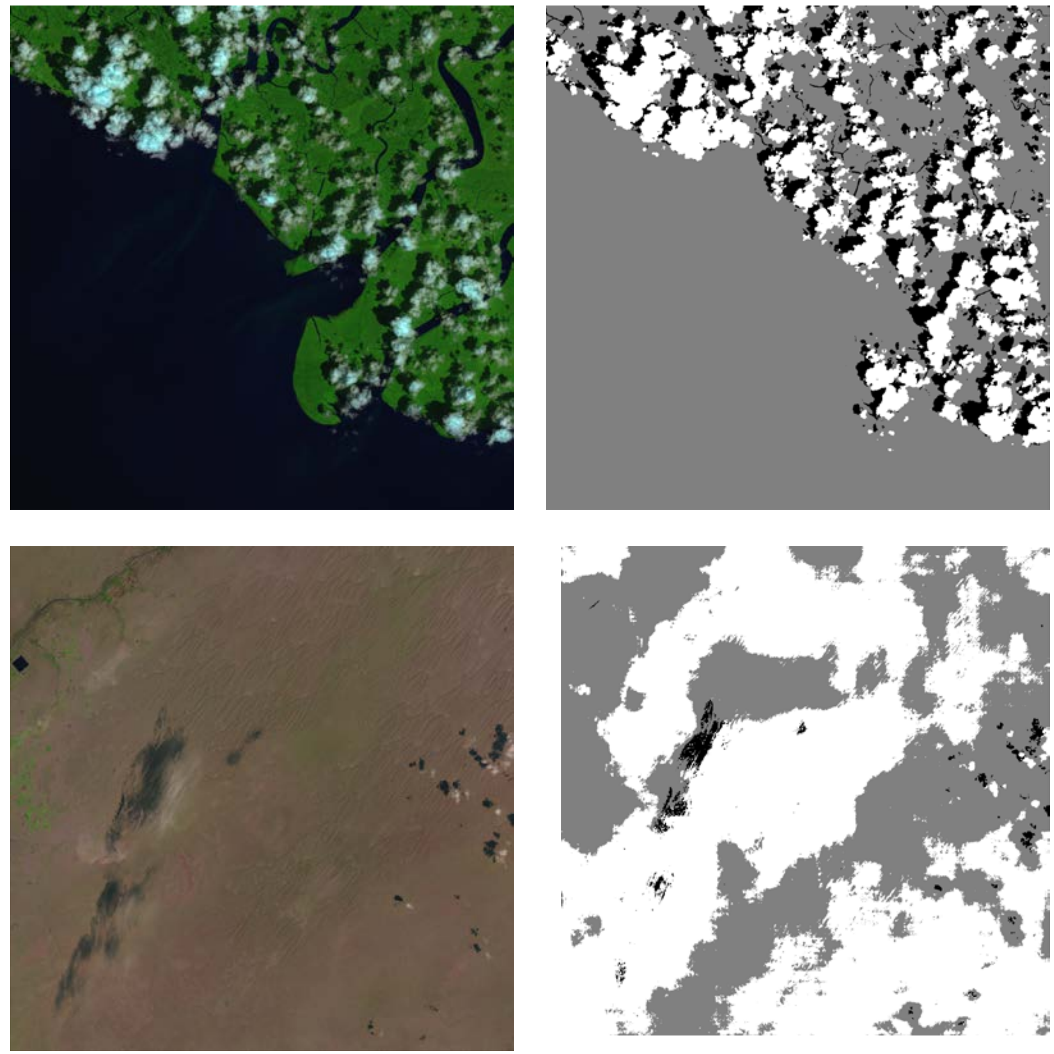
CORTEXICA
AI SOLUTIONS FOR BUSINESS

- Cloud Coverage Estimation
  - Mean: 3.21%
  - Median: 1.04%
  - 75%-Percentile: 2.47%
  - 90%-Percentile: 8.02%
  - Image Error >5%: 12.5%
  - Image Error >20%: 2.5%
  - Wrong Decision (50% Threshold): 2.5%

- Cloud Mask Estimation
  - Mean Land Recall: 94.1%
  - Mean Land Precision: 94.1%
  - Land Recall <90%: 15%
  - Land Recall <80%: 5%
  - Mean Cloud Recall: 89.6%

Speed: 30 Mpixels/second

SPARCS Images | Cloud Masks



Discussion

- Accuracy
  - Cloud coverage estimation is consistent and accurate
    - But, more research needed to model outliers: (1) how frequently they occur (2) are they correlated to a certain type of terrain or/and specific location (3) can they be avoided
  - Cloud masks are rather accurate: On average, 89.6% of the clouds can be masked out by losing only 5.9% of the land
    - But, no urban areas in SPARCS dataset
  - Overall, very good preliminary results but not statistically significant
- Speed
  - 30 MPixels/second on desktop computer with GPU
    - Real-time 1MPixel video at 30fps
  - Spatial and temporal downsampling offers greater efficiency
  - Mobile GPU platforms are slower by an order of magnitude

Current Work

- Annotation: In-house manual annotation is now underway on video frames from over 500 Carbonite-2 videos.
  - RGB, 80cm/pixel video, up to 25Mpixels frame size
  - 3 Carbonite-2 frames ≈ SPARCS dataset
  - Annotated dataset planned to be 400-500 times larger than SPARCS dataset (in pixel count)
  - 3 visibility classes: thick cloud (surface is not visible), thin cloud, no cloud
- New Tricks
  - Fast white-colour estimation to identify cloud candidates
  - Use parallax to eliminate falsely detected clouds in videos (falsely detected clouds wouldn't move relative to the ground)
- Optimization
  - Optimise model architecture for higher computational efficiency and real-time bandwidth
  - Implement on mobile GPU for real-time on-board use