
AWS
re:Invent

WPS303

Using Amazon SageMaker for geospatial imagery with Capella Space

Kate Zimmerman
Principal Deep Learning Architect
AWS

Ganesh Yalla
Data Science Lead
Capella Space

Lloyd Hughes
Data Scientist
Capella Space

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

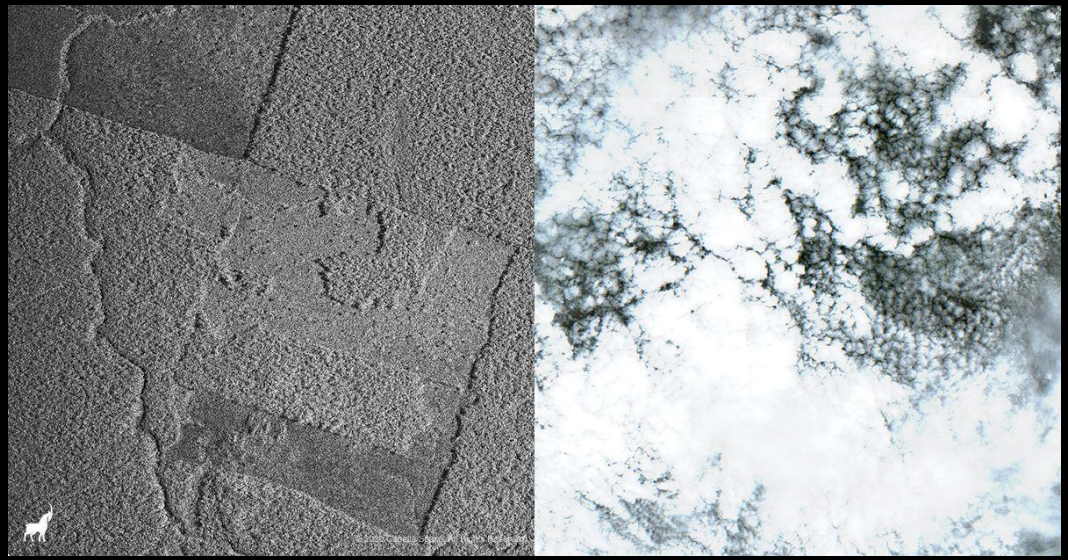
Agenda

- Introduction to Capella Space and SAR
- What makes SAR challenging for ML
- Introduction to Amazon SageMaker
- Capella’s use of Amazon SageMaker for ML on SAR

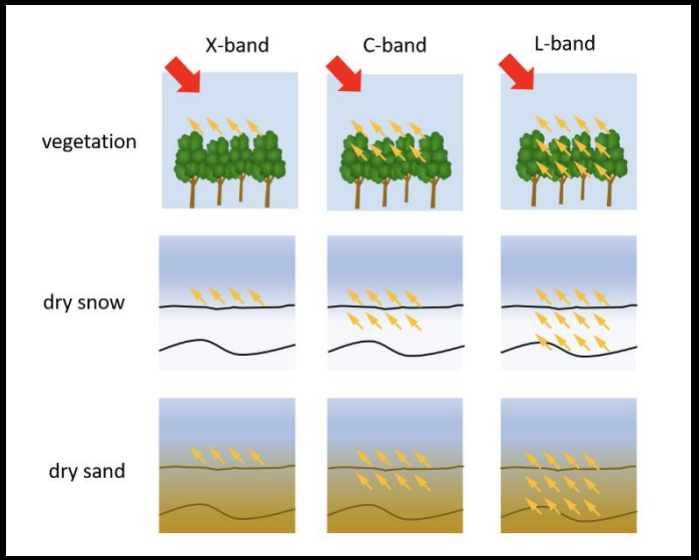


© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Synthetic Aperture Radar (SAR)



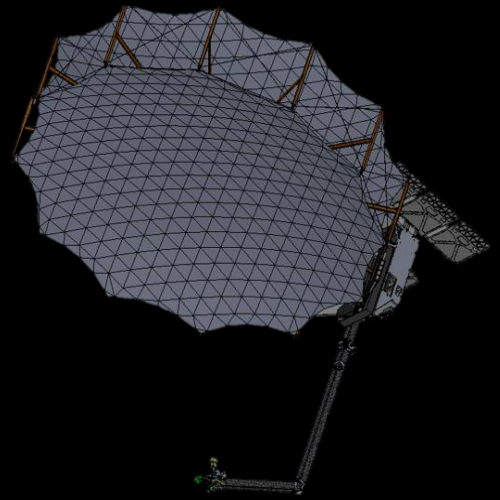
- SAR can penetrate through clouds unlike optical
- SAR can image a site day or night



- Different radar bands have different penetration capabilities
- At Capella Space we use X-band

Capella Space

CONSTELLATION OF SMALL RADAR SATELLITES



Rapid path to deployment:
less expensive to build and launch

ANY TIME, ANY WEATHER

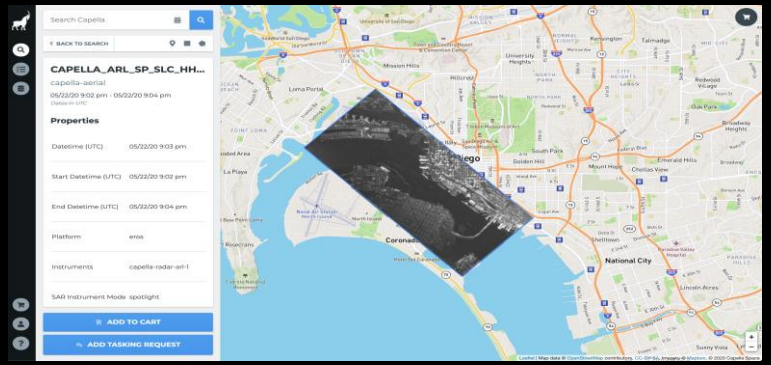
FREQUENT REVISIT

VERY HIGH RESOLUTION IMAGING

FASTEST FROM ORDER TO DELIVERY

The Capella vision

HIGH QUALITY COMMERCIAL SAR DATA AND PRODUCTS AS A SERVICE



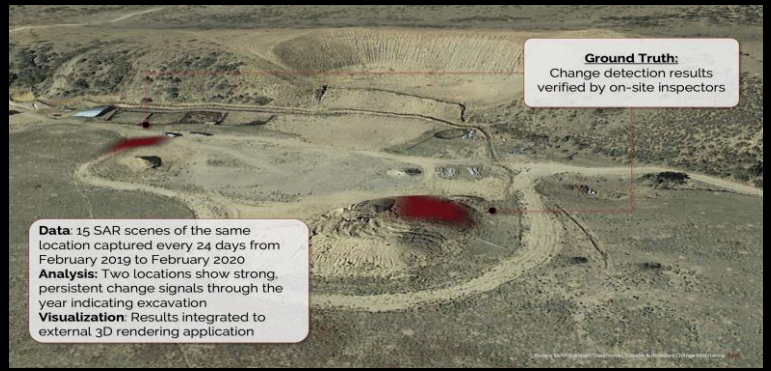
Capella Console and Tasking API



Data – High-quality, high-resolution images



Derived Products - Computer Vision and ML

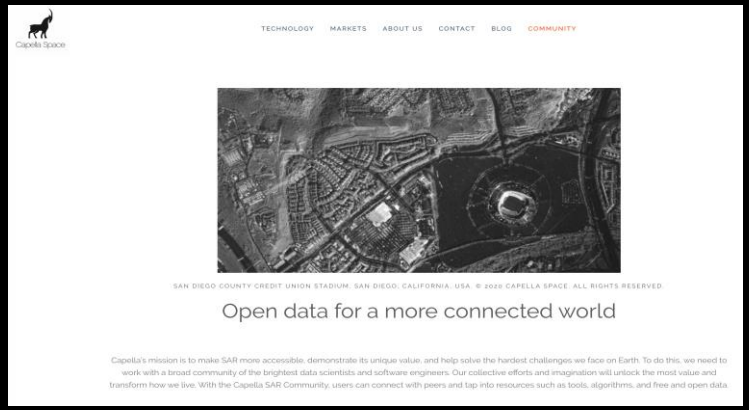


Change detection and monitoring



© 2020 Capella Space. All Rights Reserved

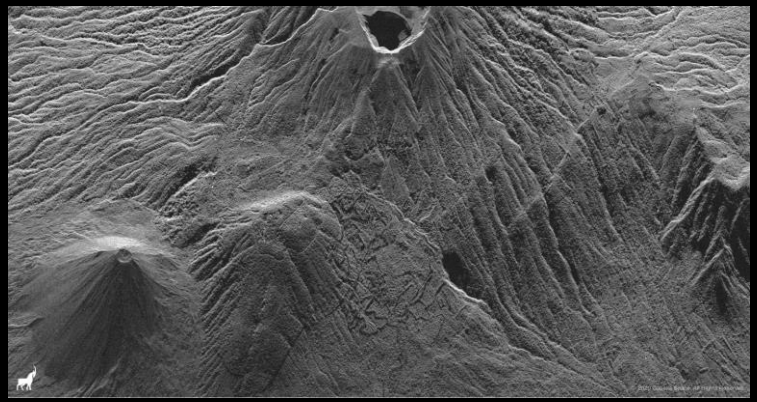
Capella community page and high-resolution images EAP



Capella Space community sign-up



Sunderbans National Forest, India

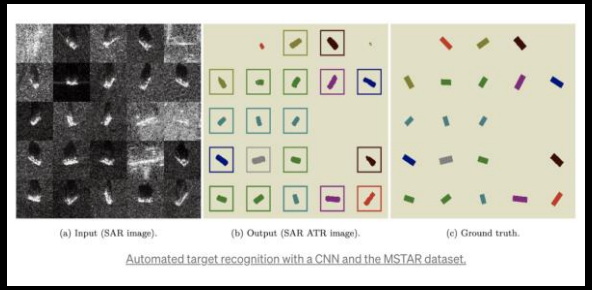


Santa Ana Volcano, El Salvador

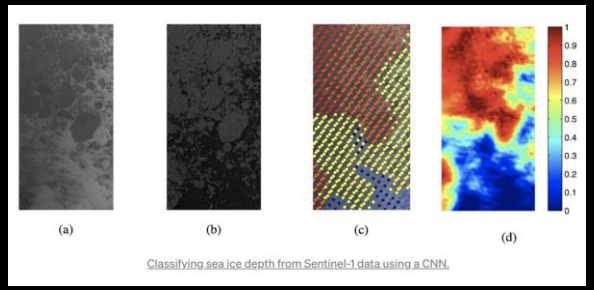


Debswana Jwaneng Diamond Mine, Botswana

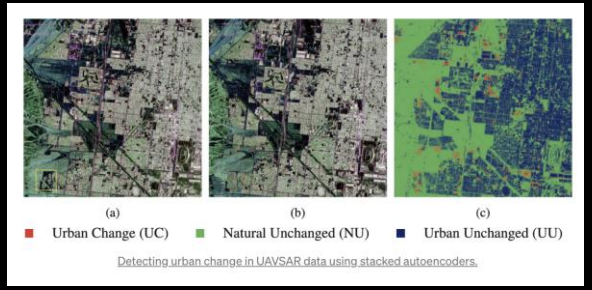
ML Use-cases for SAR imagery



Object detection



Land cover classification



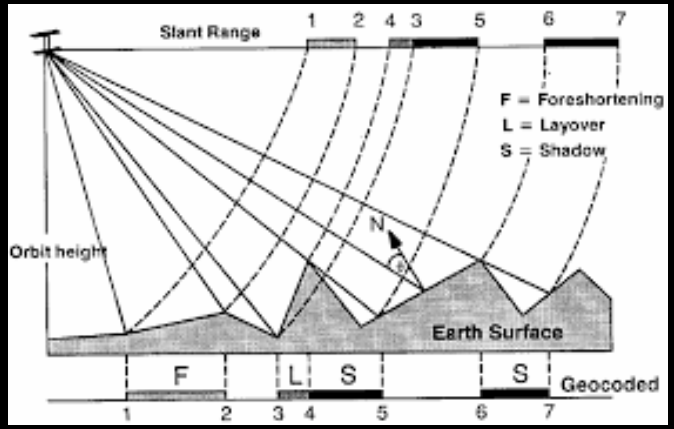
Change detection



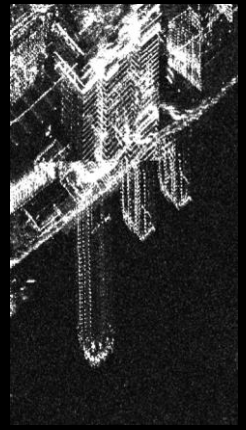
Data augmentation

- Medium Blog by Scott Soenen (VP, Product Engineering, Capella Space): https://towardsdatascience.com/deep-learning-and-sar-applications-81ba1a319def
- Capella Space Blog: https://www.capellaspace.com/blog/

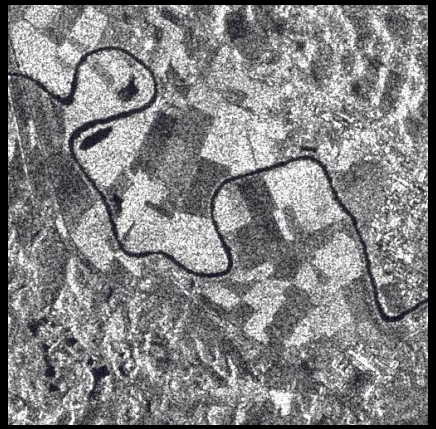
Challenges for ML on SAR



Geometric Effects (source: ESA)



Example of building layover

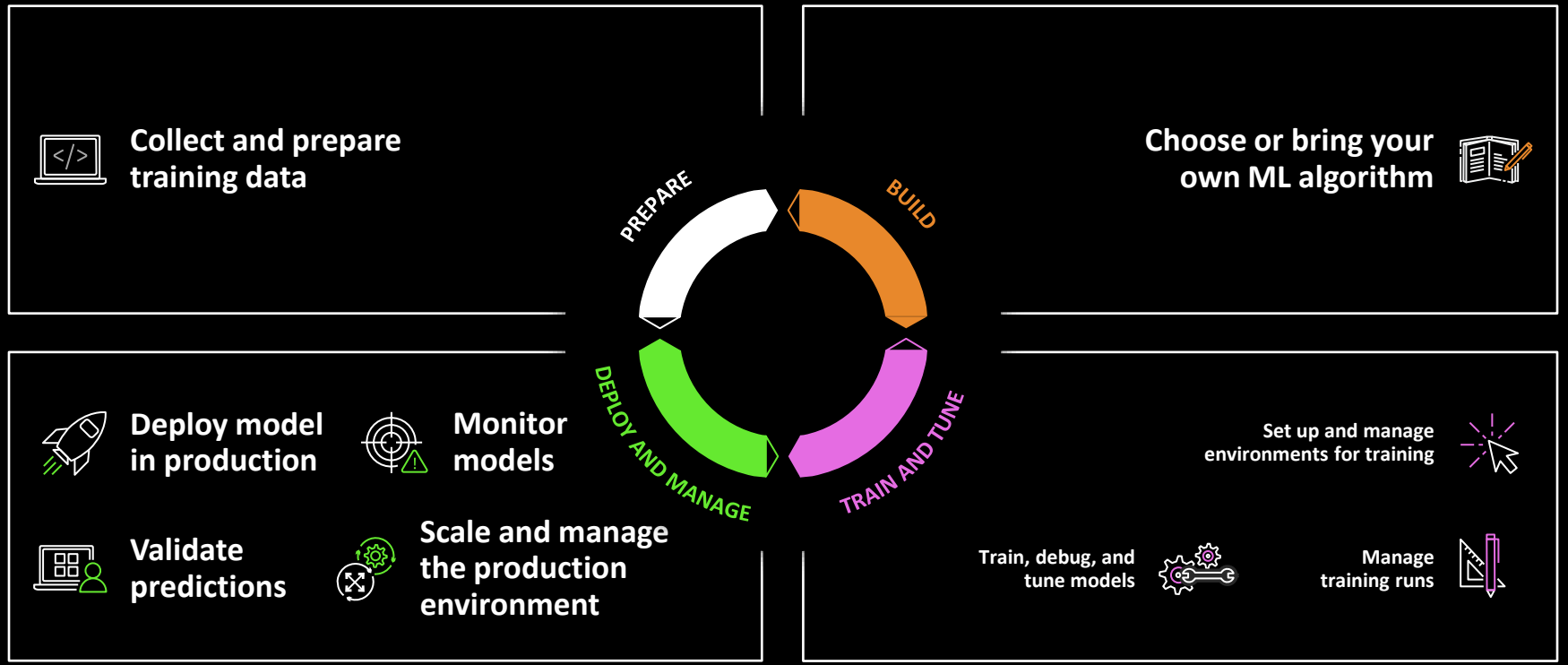


Speckle (coherent interference)

So, how do we label/train/deploy ML algorithms with SAR data?

SAR 101 blog by Jason Brown (Capella Space):
https://www.capellaspace.com/sar-101-an-introduction-to-synthetic-aperture-radar/

The Machine Learning workflow is iterative and complex



Amazon SageMaker helps you build, train, and deploy models

Collect and prepare training data
Fully managed data processing jobs/ data labeling workflows

Choose or bring your own ML algorithm
Collaborative notebooks, built-in algorithms/models



Deploy model in production
One-click deployment and auto-scaling

Monitor models
Automatically spot concept drift

Validate predictions
Add human review of predictions

Scale and manage the production environment
Fully managed with auto-scaling for 75% less

Set up and manage environments for training
One-click training

Train, debug, and tune models
Debugging and optimization

Manage training runs
Easily track and compare experiments

Amazon SageMaker Deep Dive: https://www.youtube.com/watch?v=Ac5m4-xWtCU

Amazon SageMaker

End-to-end solution for ML



Makes ML easy

One-click train/deploy, built-in algorithms, automatic tuning



Scalable performance

On-demand access to powerful GPUs, auto-scaling, distributed training



Cost-effective

Up to 70% cost reduction for data labeling, managed spot training, lowest cost inference in the cloud



Secure

Complete set of security features: network isolation, access controls, encryption, compliance



DevOps-ready

Fully managed ML ops with the option to run on Kubernetes

SINGLE IDE

perform all ML steps with a web-based visual interface

Scale to

100s

of GPUs efficiently

At least

54% **LOWER TCO**

than other cloud-based options

KUBERNETES

and Kubeflow integration

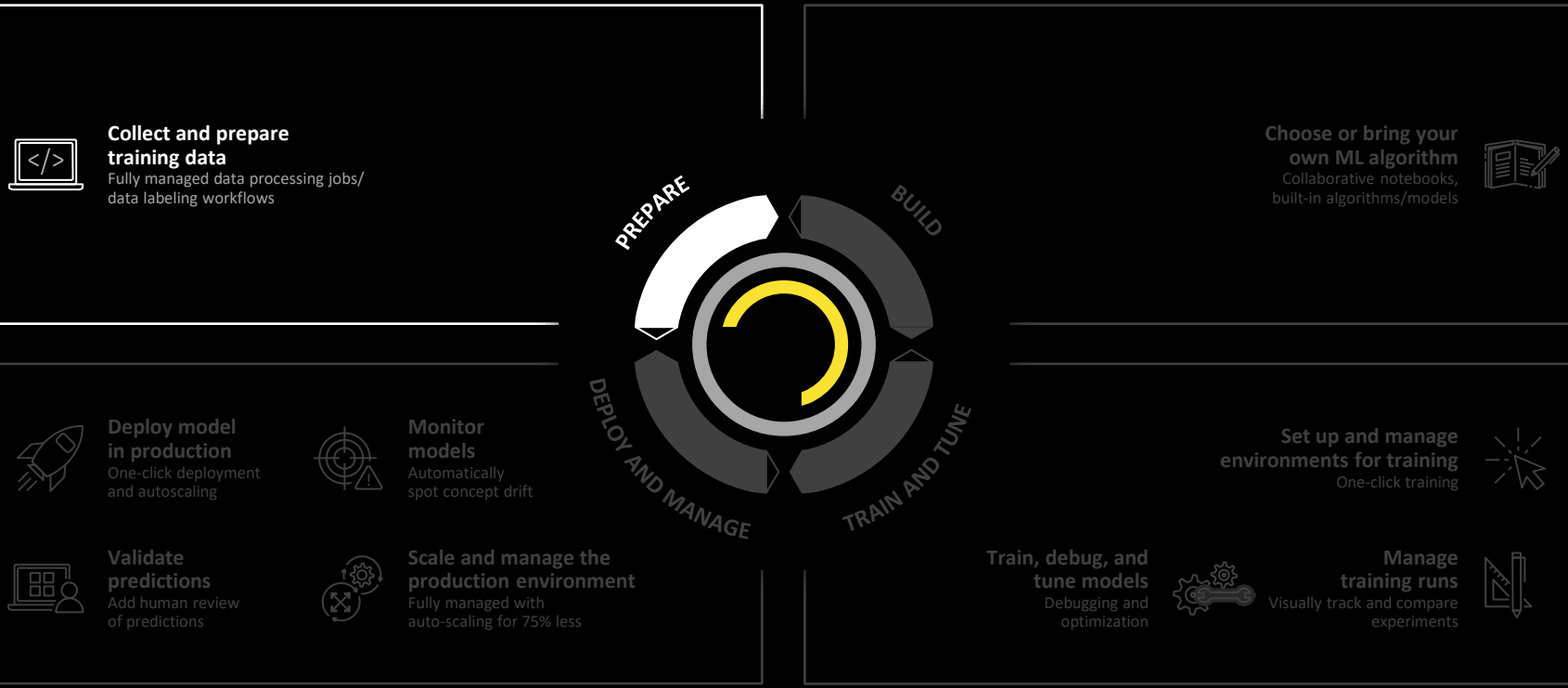
Key Amazon SageMaker features for geospatial data

Features used to address common challenges with geospatial imagery



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Collect and prepare training data

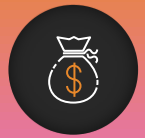


WEB-BASED IDE FOR ML

AUTOMATICALLY BUILD AND TRAIN MODELS

Amazon SageMaker Ground Truth

Build highly accurate training datasets using Machine Learning



Reduce data labeling costs
by up to 70%



Access labelers
through Amazon Mechanical Turk, Amazon approved vendors, or use private human labelers



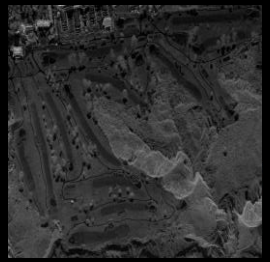
Achieve accurate results quickly



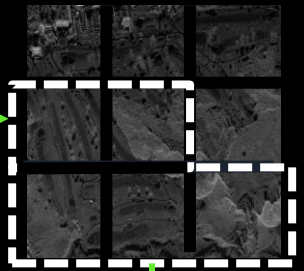
Built-in data labeling workflows
support 3D point clouds, videos, images, and text

Data labeling workflow

Raw image to label

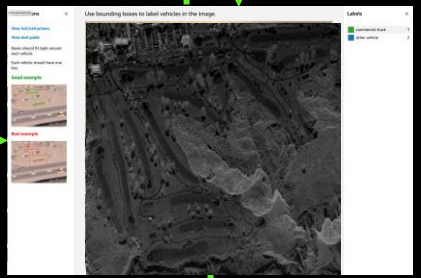


Tile pre-filtering



Data Labelers
Label Data

Multiple labelers per images, label consolidation



Chips to be Labeled

Consolidated Training Data

Amazon SageMaker Build/Train/Deploy Model

Expert Reviewers

Manual review of low confidence labels

Train and tune models



Collect and prepare training data
Fully managed data processing jobs/ data labeling workflows



Choose or bring your own ML algorithm
Collaborative notebooks, built-in algorithms/models



Deploy model in production
One-click deployment and autoscaling



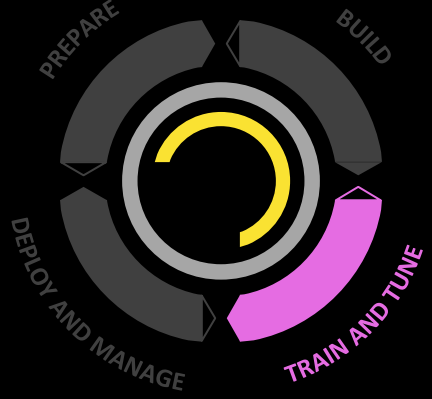
Monitor models
Automatically spot concept drift



Validate predictions
Add human review of predictions



Scale and manage the production environment
Fully managed with auto-scaling for 75% less



Set up and manage environments for training
One-click training



Train, debug, and tune models
Debugging and optimization



Manage training runs
Visually track and compare experiments

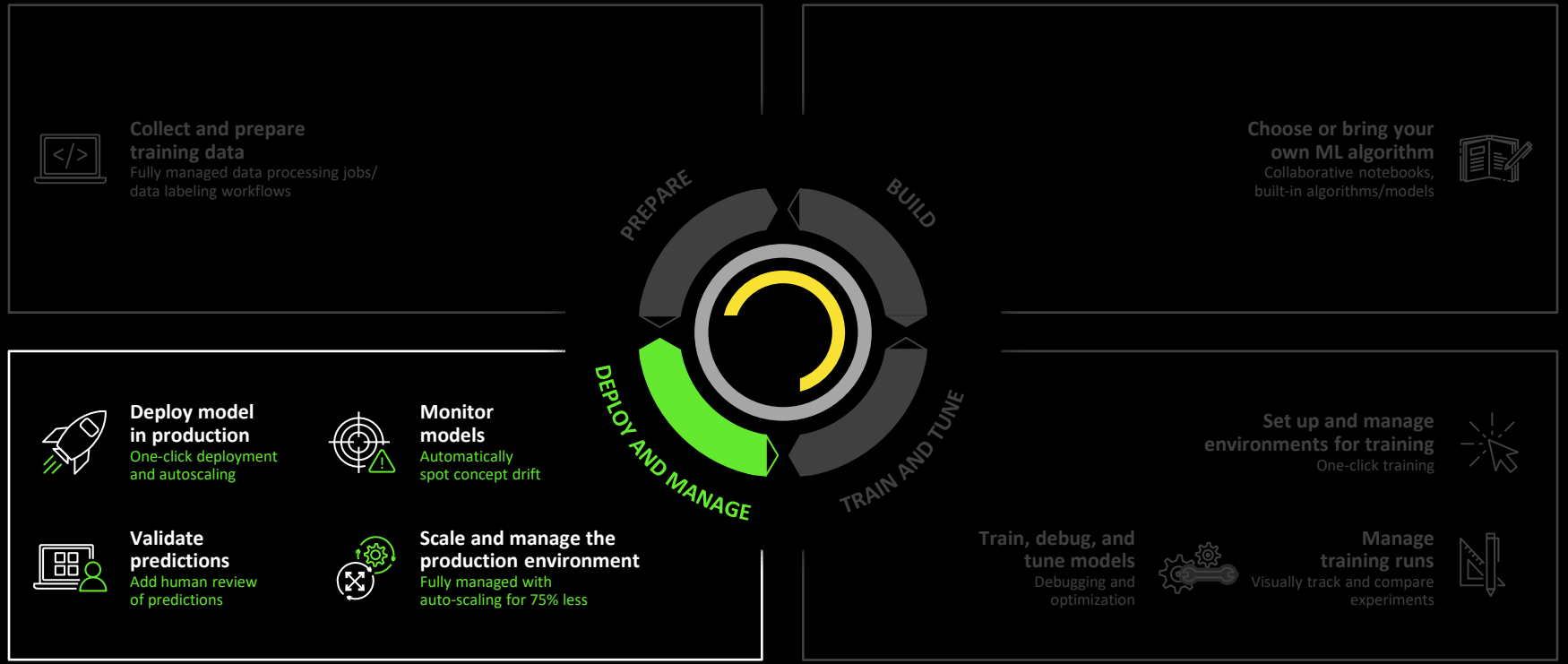


WEB-BASED IDE FOR ML

AUTOMATICALLY BUILD AND TRAIN MODELS


Train with your own algorithms
Single pass training
Distributed by default
Not memory bound
Train on a data stream
Checkpoint for re-training
Train your model with
ONE CLICK using
Amazon SageMaker

Deploy and manage models

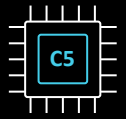


WEB-BASED IDE FOR ML

AUTOMATICALLY BUILD AND TRAIN MODELS

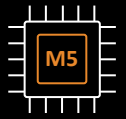
Wide selection of options for cost-effective compute

CPU INSTANCES

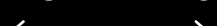


Small models,
low throughput

ELASTIC INFERENCE



NETWORK ATTACHED



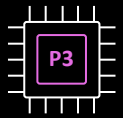
INFERENCE ACCELERATOR

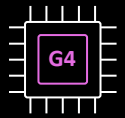


eia1.medium

Mid-sized models, low-latency
budget with tolerance limits

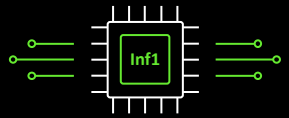
GPU INSTANCES





Large models, high throughput,
and low-latency access to CUDA

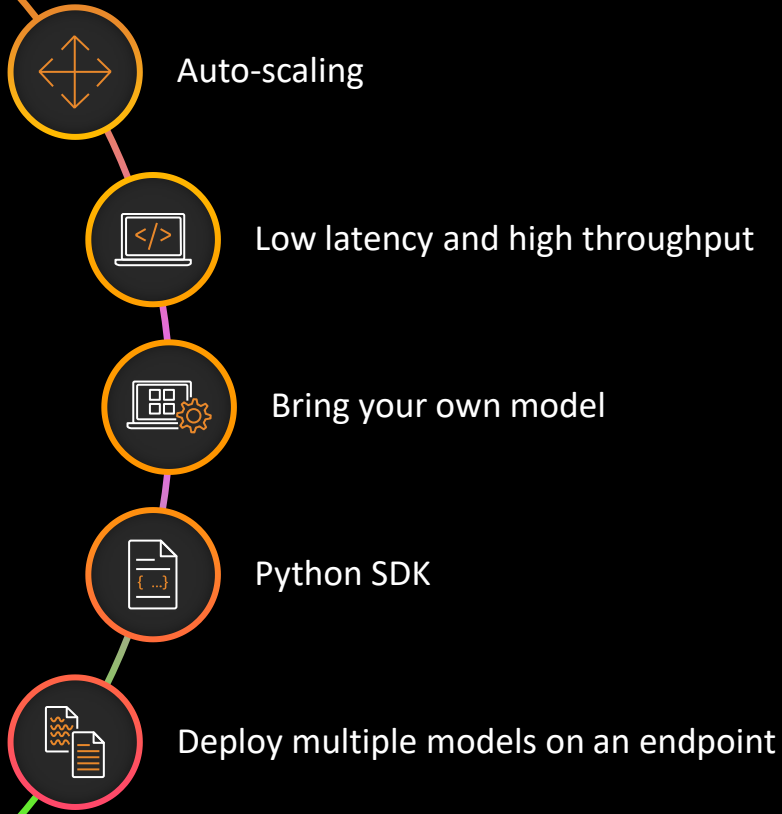
CUSTOM CHIP



Inf1: High throughput, high performance,
and lowest cost in the cloud

Amazon SageMaker
is fully managed

One click model deployment



SAR Despeckling

What is Speckle?

- Deterministic, multiplicative, non-Gaussian

Existing Despeckling approaches:

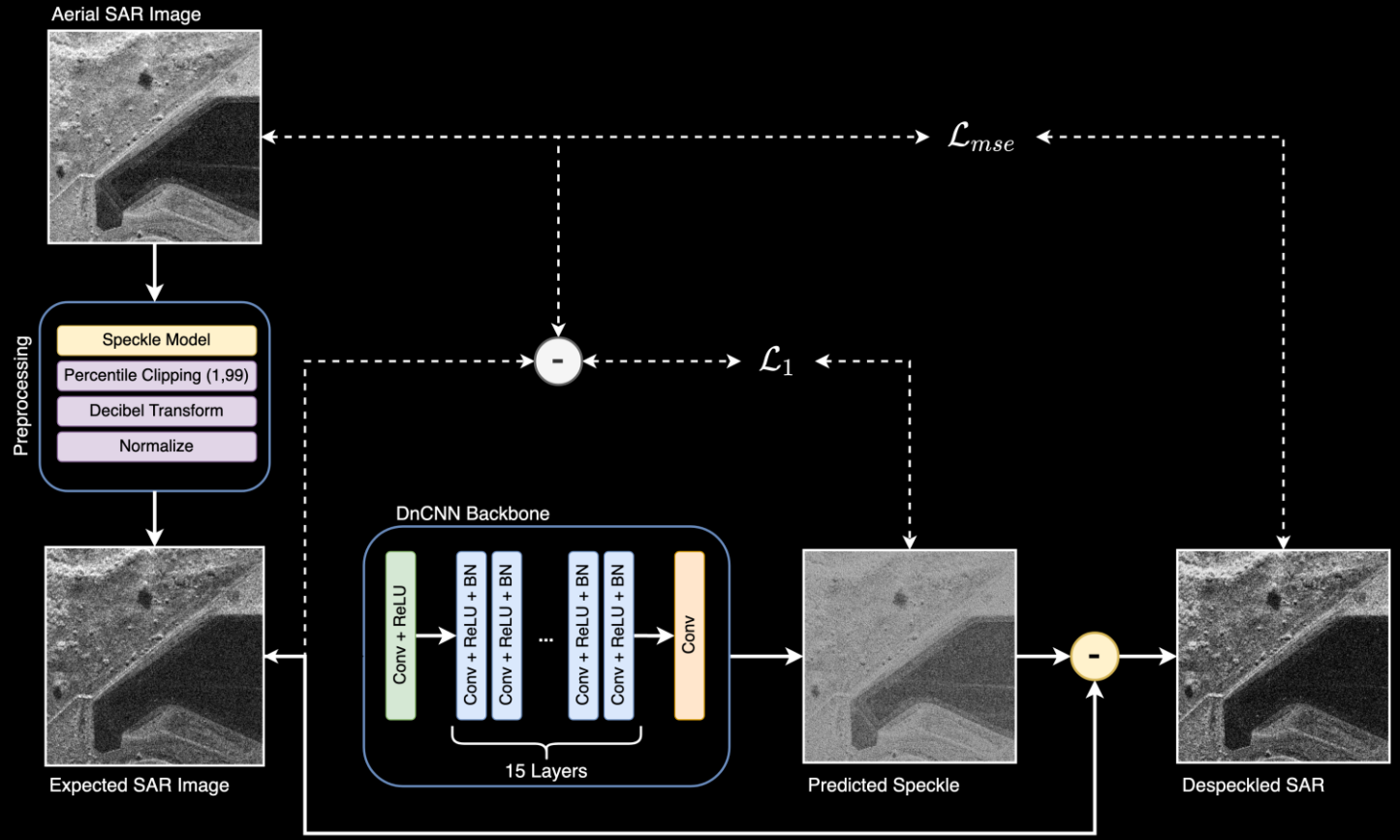
- Statistical filtering: Lee, Median, Non-local
- Decreased spatial resolution

Deep learning to the rescue?

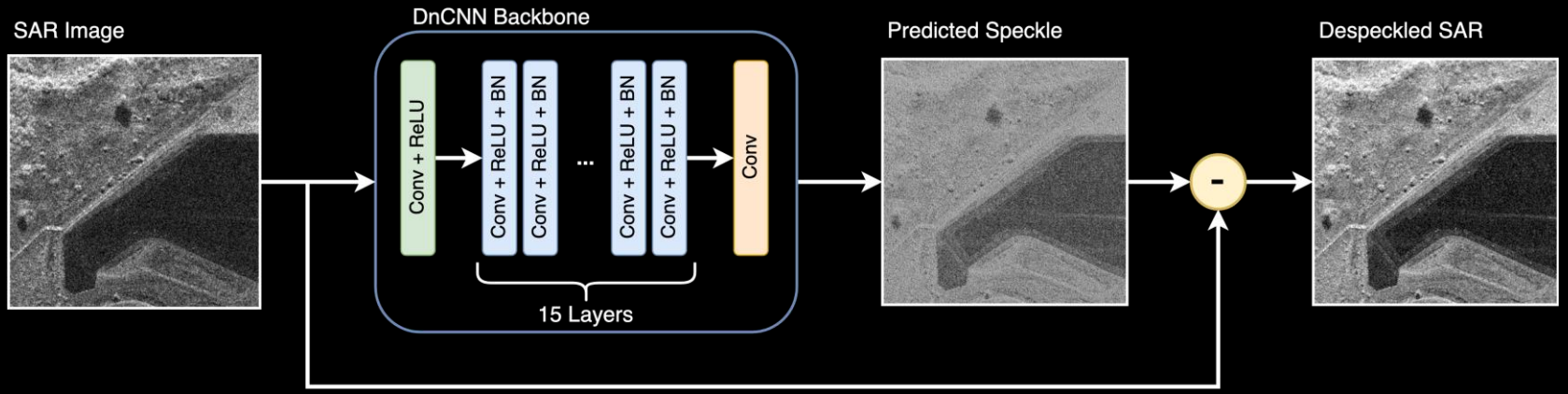
- Numerous existing works
- Capella is in a unique situation

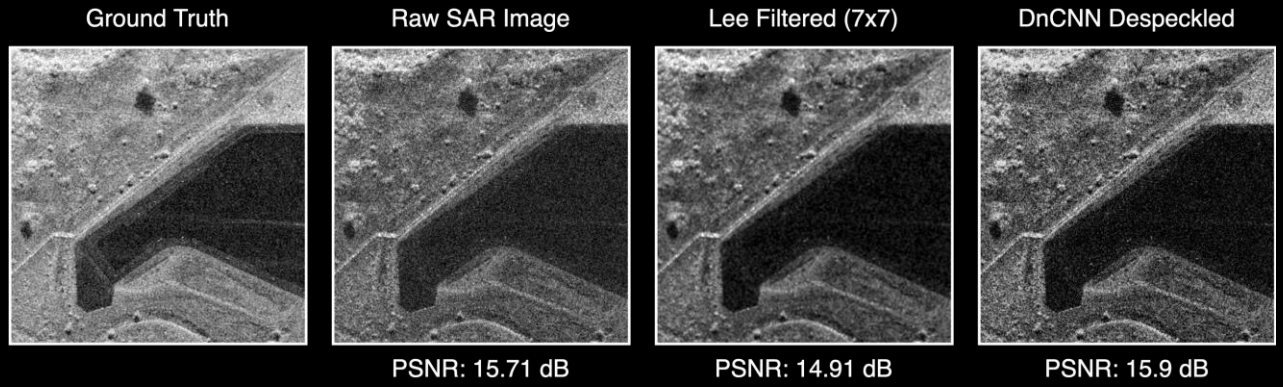


Training: Residual Despeckling Network



Inference: Residual Despeckling Network



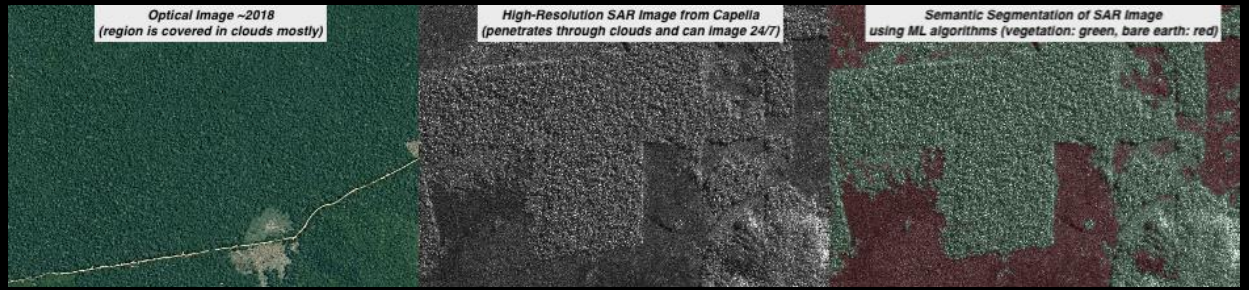


Where are we going from here?

CHANGE MONITORING RESEARCH: EARLY POCS USING CAPELLA'S OWN VERY HIGH-RESOLUTION SAR DATA

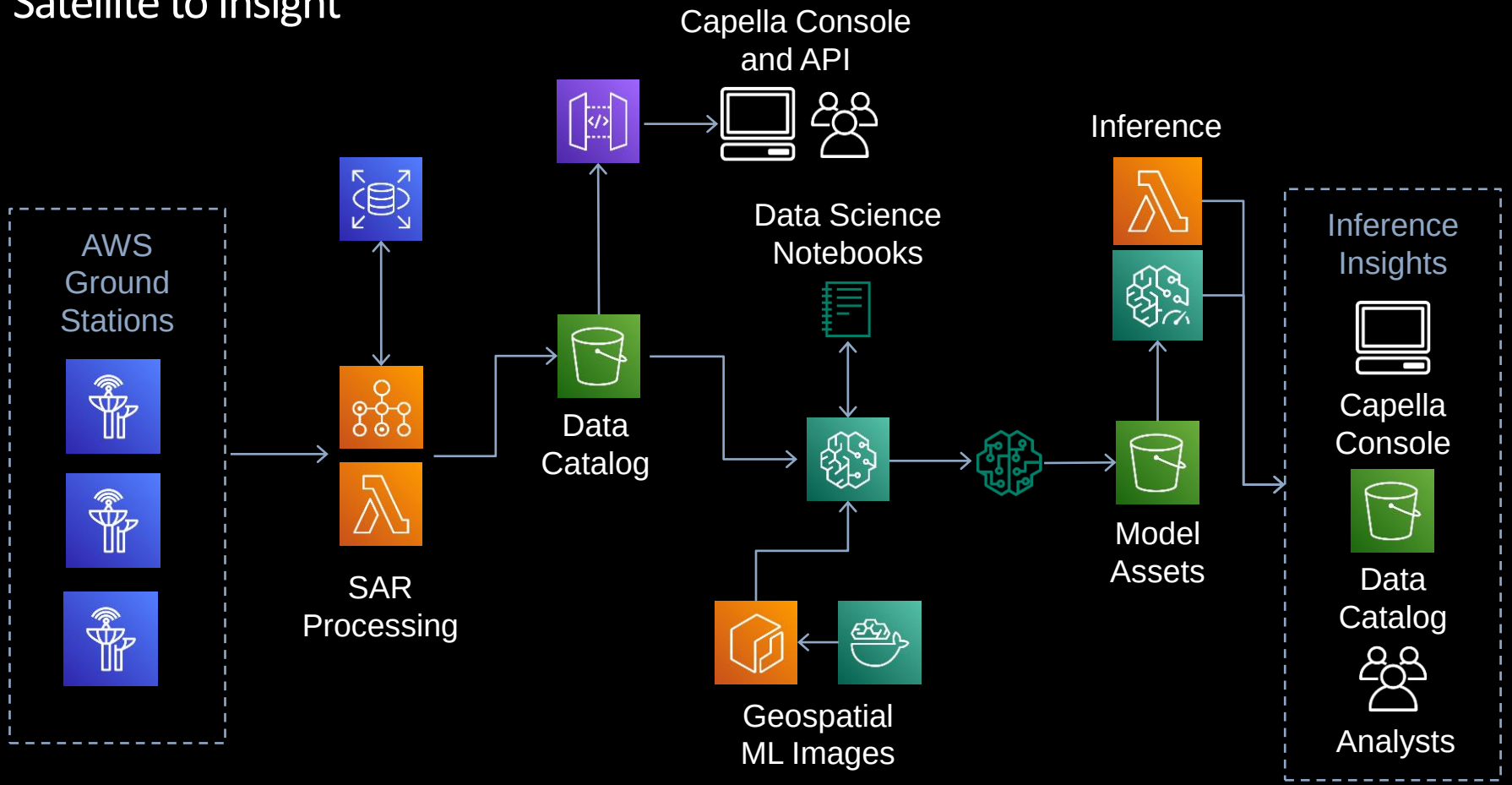


Object detection (aircrafts at an airbase)



Semantic segmentation (deforestation)

Satellite to Insight



Thank you!

Kate Zimmerman

Principal Deep Learning Architect
Amazon Web Services
werlingk@amazon.com

Ganesh Yalla

Data Science Lead
Capella Space
ganesh.yalla@capellaspace.com

Lloyd Hughes

Data Scientist
Capella Space
lloyd.hughes@capellaspace.com

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Please complete the session survey

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.