SEN12TS: COMBINING SENTINEL-1 DERIVED BACKSCATTER AND INSAR TIMESERIES WITH MULTISPECTRAL SENTINEL-2 IMAGERY FOR DEEP LEARNING AND DATA FUSION

Terence Conlon
Columbia University
New York, NY
tmc2180@columbia.edu

Rose Rustowicz
Descartes Labs
San Francisco, CA
rose@descarteslabs.com

ABSTRACT

Despite widespread availability of radar-derived and optical imagery at similar timescales and spatial resolutions, a number of issues complicate the creation of combined radar and optical datasets. These challenges include problems with coregistration across satellite missions; difficulties in post-processing radar imagery to correct for ground geometry and incidence angle; and a lack of reliable labeled data to pair with imagery for class-dependent research objectives. Addressing these concerns using the Descartes Labs geospatial analytics platform, this paper presents the SEN12TS dataset, a public resource that includes radiometric terrain corrected synthetic aperture radar backscatter measurements; interferometric synthetic aperture radar coherence and phase layers; local incidence angle and ground slope values; multispectral optical imagery; and decameter-resolution land cover data. Moreover, sensed imagery is available in timeseries: Within a Sentinel-1, Sentinel-2, and labeled land-use/land-cover image triplet, radar-derived imagery is collected at four timesteps 12 days apart. For the same spatial extent, up to 16 image triplets are available across the calendar year of 2020.

The SEN12TS dataset consists of image triplets over six agro-ecologically diverse areas of interest: California, Iowa, Catalonia, Ethiopia, Uganda, and Sumatra. In total, 246,400 triplets are produced at 10m resolution over 31,398 256-by-256-pixel unique spatial tiles for a total size of 1.69 TB. Two use cases are also demonstrated for the SEN12TS dataset. The first transforms radar imagery into enhanced vegetation indices by means of a generative adversarial network, and the second tests combinations of input imagery for cropland classification. The SEN12TS dataset is hosted by the Radiant Earth Foundation, where it is available for download at `https://radiant.earth/link_to_come` under a non-commercial CC BY-NC 2.0 license.

Keywords Sentinel-1 · Sentinel-2 · data fusion · deep learning · image timeseries

1 Introduction

In the field of remote sensing, large amounts of data, better access to computational resources, and recent methodological advances have allowed researchers to develop reliable applications for a variety of tasks [Zhu et al., 2017]. However, progress is contingent upon large, high-quality datasets, which due to the high costs of labelling, are fewer and less studied than widely used datasets for classic computer vision tasks [Cheng et al., 2016]. Publicly-available labeled datasets, such as those listed in Rieke [2022] and those organized for standardized access in Yeh et al. [2021], provide significant value for researchers looking to develop new remote sensing methodologies or derive new insights from existing imagery.

As remotely-sensed data become more accessible, researchers are increasingly exploring the benefits of data fusion – the process of combining multiple satellite imagery products, often across sensing domains, to extract otherwise hidden information. In this space, there have been many efforts to link radar and optical imagery, and in a few cases, produce public datasets. Of particular note are the SEN1-2 Schmitt et al. [2018] and SEN12MS Schmitt et al. [2019] datasets,

Table 1: Existing combined SAR-optical imagery datasets

Dataset	Num. Images	Image Size	Resolution	Data Sources	Data Layers	Description	Reference
SEN12MS	541,986	256 x 256	10m	Sentinel-1, Sentinel-2, MODIS	SAR *vv* + *vh*, 13 MSI bands, land cover classification	Paired S1/S2 images are combined with MODIS land cover classifications covering all inhabited continents during all meteorological seasons	Schmitt et al. [2019]
SEN1-2	564,768	256 x 256	10m	Sentinel-1, Sentinel-2	SAR *vv*, RGB bands	S1 single polarization images are paired with S2 RGB bands over a global extent	Schmitt et al. [2018]
So2SAT LCZ42	400,673	32 x 32	10m	Sentinel-1, Sentinel-2	SAR *vv* + *vh*, 13 MSI bands	S1/S2 imagery pairs are collected over 42 urban regions (with 10 smaller areas) across the globe and are assigned local climate zone classes, verified by domain experts.	Zhu et al. [2019]
SARptical	10,108	112 x 112	$\sim$1m	TerraSAR-X, UltraCAM	SAR single polarization, RGB bands	Pairs of 1m SAR imagery and 20cm optical imagery are collected over Berlin, Germany. Images are coregistered with the 3D position of the center of the image pair.	Wang and Zhu [2018]
SEN12TS	246,400	256 x 256	10m	Sentinel-1, Sentinel-2, USDA CDL, SIGPAC, ESA	SAR *vv* + *vh*, InSAR coherence + phase, 12 MSI bands, land cover classifications	S1-derived timeseries of radiometric terrain corrected backscatter and InSAR coherence and phase layers are combined with S2 images and land cover labels. Image triplets cover portions of California, Iowa, Catalonia, Ethiopia, Uganda, and Sumatra.	This paper

which present paired Sentinel-1 (S1) synthetic aperture radar (SAR) backscatter measurements and Sentinel-2 (S2) optical images in >500,000 locations across the world; these papers provide the inspiration for this dataset and paper. A non-exhaustive overview of other large-scale (>10,000 image pairs) combined SAR-optical datasets is shown in Table 1.

A number of challenges exist in the production of combined SAR-optical datasets. First, all imagery must be tightly coregistered: registration differences will degrade the usefulness of the dataset [Zhu et al., 2022]. Second, raw SAR imagery requires substantial processing to reduce noise and amplify contained signals. Imagery that is both geometrically corrected (controlled for ground geometry) and radiometric terrain corrected (controlled for incidence angle) is more valuable to the end user. Third, there are few high-quality labeled datasets to attach to large imagery datasets. While global land-use/land-cover (LULC) maps do exist, these products often have hectometer spatial resolutions coarser than relevant on-the-ground length-scales; are informed by limited ground truth; exist for a single year or period of time; and/or significantly disagree among themselves as to predicted land cover classes. Accordingly, a combined SAR-optical dataset with precise co-registration; geometric and radiometric terrain corrected SAR imagery; and attached high-quality, decameter LULC maps will provide substantial utility to remote sensing researchers.

To this end, this paper presents the SEN12TS dataset, which addresses the above-mentioned challenges and introduces the following novel features. For each image triplet in the dataset, S1-derived radar imagery is provided at a timeseries (TS) of four timesteps: $t' = t..t-3$, where $t' = t$ corresponds to the time at which the S2 optical image is collected. These timesteps are 12 days apart, representing the time between successive passes for a single S1 satellite. Along with *vh* and *vv* SAR backscatter layers, interferometric SAR (InSAR) coherence and phase layers are included at all four timesteps. The S1-derived imagery is augmented with two layers representing the local incidence angle and the

Table 2: SEN12TS dataset overview

Area of Interest	Number of Tiles	Number of Image Triplets	Size (GB)
California	3151	47,660	326.5
Iowa	5292	47,902	327.4
Catalonia	4127	47,280	323.0
Ethiopia	5828	47,670	325.9
Uganda	7173	47,676	326.1
Sumatra	5827	8212	56.8
Total	27,698	246,400	1685.7

Shuttle Radar Topography Mission (SRTM)-derived ground slope values [NASA Jet Propulsion Laboratory, 2013]. Shadow/layover masks are appended to S1 images, while scene classification layers and cloud masks are appended to S2 images. Labeled LULC classifications specific to the area of interest (AOI) are then added. Together, these 19 S1 layers, 14 S2 layers, and 1 LULC layer constitute the imagery contained within a SEN12TS image triplet in the dataset. To capture surface characteristics across the entire 2020 calendar year, up to 16 image triplets – 4 per season – are generated for each spatial extent included in the dataset, subject to imagery availability.

Imagery is collected over 6 agro-ecologically distinct AOIs – California, Iowa, Catalonia, Ethiopia, Uganda, and Sumatra. LULC classifications from the United States Department of Agriculture (USDA) Cropland Data Layer (CDL) are added to imagery set in California and Iowa [USDA National Agricultural Statistics Service, 2020]; for imagery collected over Catalonia, LULC classifications from Sistema de Información Geográfica de Parcelas Agrícolas (SIGPAC) are included [Government of Catalonia (GENCAT), 2021]; for imagery collected over Ethiopia, Uganda, and Sumatra, LULC classifications are provided by the ESA 2020 WorldCover Map v100 [Zanaga et al., 2021].

Table 2 presents an overview of the imagery contained in the SEN12TS dataset. In general, imagery is collected over randomly selected tiles in the AOIs until approximately 320 GB of imagery is available for each region. However, so few S2 images with minimal cloud cover are available across Sumatra that collecting imagery over the entire AOI only results in 56.8 GB of imagery.

Lastly, two applications are presented as initial explorations of the SEN12TS dataset. The first translates S1 radar imagery into enhanced vegetation index (EVI) predictions by means of a modified generative adversarial network (GAN), while the second compares LULC classification performance in California for different configurations of S1 and S2 imagery. The SEN12TS dataset is published with a non-commercial CC BY-NC 2.0 license, and is hosted by Radiant Earth Foundation at the following link: `https://radiant.earth/link_to_come`.

2 Data Background

The SEN12TS dataset combines timeseries of S1-derived imagery, S2 multispectral imagery, and LULC classifications in imagery triplets. Imagery is collected over unique spatial extents called tiles; these tiles have a 10m pixel resolution and size of 256-by-256 pixels. Figure 1 presents the distributions of saved imagery within the 6 AOIs contained in the SEN12TS dataset; colors of the figure dots correspond to the number of images available for the tile (see Figure 1 caption). The 6 AOIs contained in the dataset are selected to capture a degree of global diversity in land cover patterns. Among the US-based AOIs, California* offers a well-studied, highly developed agricultural region that contains a wide variety of crop types. Iowa provides a region that is likewise developed and studied with much fewer distinct crop types. Catalonia also contains similarly advanced agricultural practices, albeit with different climate and cropping patterns. These three AOIs contain high-quality, government-produced LULC maps that can be combined with imagery to explore a variety of classification-based research questions.

The following three AOIs – Ethiopia, Uganda, and Sumatra – are regions that have received comparatively less attention from remote sensing and deep learning researchers. Agricultural practices in these areas differ from those in the US and Europe: Plots are smaller, irrigation is less prevalent, and management of the land is less mechanized and more labor-intensive. In Uganda and Sumatra in particular, equatorial climates ensure persistent cloud presence throughout

*The California AOI covers a portion of the California Central Valley south of Sacramento and north of Bakersville; it does not cover the entire state and is named as such for concision.

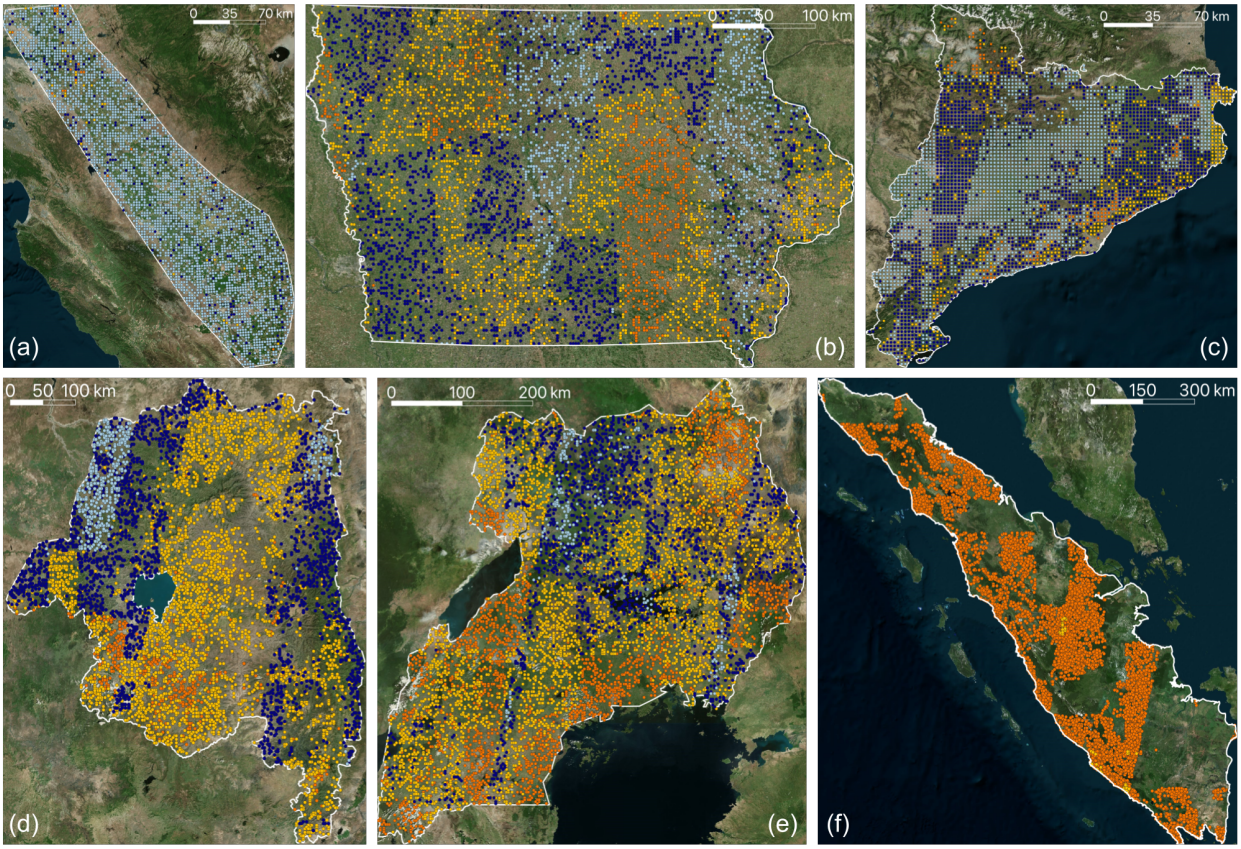

Figure 1: Locations of saved imagery in the 6 SEN12TS AOIs: (a) California, (b) Iowa, (c) Catalonia, (d) Ethiopia, (e) Uganda, and (f) Sumatra. Light blue dots indicate between 13-16 images available per tile; dark blue dots indicate between 9-12 images available per tile; yellow dots indicate between 5-8 images available per tile; and orange dots indicate between 1-4 images available per tile. All images are projected to their local Universal Transverse Mercator (UTM) zone. Background imagery is provided by Bing Aerial.

much of the year, a reality that increases the need for radar-based monitoring solutions. To focus on agriculturally intensive areas in Ethiopia, imagery is only collected in Amhara and Tigray regions, parts of the country that constitute the highly-cropped West Ethiopian Highlands.

2.1 Sentinel-1

Two polar orbiting satellites, equipped with C-band SAR (5.404 GHz) sensors, comprise the European Space Agency's (ESA) Sentinel-1 mission [Torres et al., 2012]. These satellites have global repeat periods of 6-24 days and collect cloud-free imagery both day and night regardless of weather. S1 satellites can operate in one of three imaging modes: interferometric wideswath (IW), extended wideswath, and stripmap. The main mode, IW, contains a 250km swath with decameter (10-20m, depending on the processing) pixel resolution.

Multiple products are derived from S1 sensed data, including Single Look Complex (SLC) and Ground Range Detected (GRD) backscatter measurements. SLC products are provided in zero-Doppler slant-range geometry, include a single look in each dimension, and consist of complex samples that preserve phase information [European Space Agency (ESA), 2021]. In contrast, GRD products consist of focused SAR data that have been multilooked and projected to ground range using an Earth ellipsoid model. GRD products have approximately square pixel tiling and reduced speckle at the cost of losing phase information and coarser spatial resolution. Backscatter products from the Sentinel-1 mission contain *vh* and *vv* polarized images.

To reduce the impact of ground geometry on SAR data, backscatter products can be normalized by either the ground area (σ_0) or the illuminated area projected in the look direction (γ_0) [Small et al., 2004]. Here, σ_0 images are considered to be geometric terrain corrected, while γ_0 images are considered to be radiometric terrain corrected (RTC). RTC

images control for local incidence angle, allowing for standardized information extraction on sloped terrain and across different satellite passes.

Successive satellite passes generate data that can be used to produce Interferometric Synthetic Aperture Radar (InSAR) measurements, which are effective in measuring changes in land cover over time. InSAR data are often presented as interferograms – maps that track relative ground-surface change by comparing signal coherence and phase across two SAR images covering the same spatial extent. Because of their ability to highlight changes in land cover, interferograms have significant potential for various land monitoring purposes; however, a scarcity of publicly available interferograms means that this image type is less frequently utilized compared to other S1 derived products.

2.2 Sentinel-2

Since March 2017, two identical polar-orbiting satellites have collected optical imagery as part of the ESA Sentinel-2 mission [Drusch et al., 2012]. The satellites are 180 degrees out of phase, have an orbital swath of 290km, and are equipped with a multispectral instrument (MSI) payload that samples 13 spectral bands. Of these 13 bands, four have a 10m spatial resolution (blue, green, red, near infrared), six have a 20m spatial resolution (red edge 1, red edge 2, red edge 3, red edge 4, short-wave infrared 1, short-wave infrared 2), and three have a 60m spatial resolution (coastal aerosol, water vapor, cirrus). At the equator, the S2 mission has a revisit period of approximately 5 days, ensuring frequent collection of multispectral imagery; however, clouds can reduce the availability and utility of these images, especially in tropical areas with persistent cloud cover. In correcting S2 imagery to bottom-of-atmosphere reflectance values – i.e. to the L2A processing level – the cirrus band is removed, leaving 12 spectral bands. At the L2A level, S2 images are also produced with a scene classification layer that indicates the pixelwise presence of clouds, shadows, vegetation, desert, or snow; this scene classification layer can be used to derive additional cloud, shadow, or snow masks for the corresponding optical imagery. Additional information on the L2A processing algorithm, including the corresponding classes for the scene classification raster values, is provided at European Space Agency (ESA) [2022].

2.3 Land-Use/Land-Cover Label Sources

The following subsections detail the sources of LULC labels contained within the SEN12TS image triplets. Relationships between the raster values and crop types are detailed in a supplementary spreadsheet provided alongside the dataset.

2.3.1 United States Department of Agriculture Cropland Data Layer

The United States Department of Agriculture (USDA) Cropland Data Layer (CDL) is a raster, geo-referenced, crop-specific land cover data layer created annually for the continental United States using moderate resolution satellite imagery and extensive agricultural ground truth [USDA National Agricultural Statistics Service, 2020]. The CDL is produced at 30m resolution and contains 116 discrete land cover classes.

Early applications of the CDL estimated crop acreage and yield; however, the product is now used more widely, including to measure cropping frequency, create cultivated layers, and monitor disasters such as hurricanes and wildfires [Sandborn et al., 2019]. Between 2008 and 2016, the number of annual peer-reviewed journal articles or conference proceedings available through Google Scholar that utilized the CDL increased by 600% [Lark et al., 2017]. An independent assessment of the CDL found that the product is highly accurate and suitable for annual land cover applications [Luman and Tweddale, 2008].

The 2020 CDL is resampled via nearest-neighbor interpolation to 10m and used as the LULC layer for all imagery triplets in California and Iowa.

2.3.2 Sistema de Información Geográfica de Parcelas Agrícolas

Sistema de Información Geográfica de Parcelas Agrícolas (SIGPAC) combines satellite imagery and cadaster data to compile a LULC dataset for the entirety of Catalonia, Spain [Government of Catalonia (GENCAT), 2021]. SIGPAC contains 170 LULC classifications at the parcel level and has been produced every year since 2015. SIGPAC LULC labels have been combined with satellite imagery for a number of research objectives, including irrigation mapping [Gao et al., 2018], land abandonment modelling [Corbelle-Rico et al., 2012], and semantic segmentation of cropland [Pedrayes et al., 2021].

The SIGPAC 2020 dataset is downloaded from Government of Catalonia (GENCAT) [2021]; plot polygons are then rasterized to a 10m resolution. The rasterized LULC layer is used for all image triplets in Catalonia.

2.3.3 European Space Agency WorldCover 2020 Map V100

The European Space Agency (ESA) recently published the WorldCover 2020 Map v100, a global land cover map [Zanaga et al., 2021]. With its 10m resolution, the map joins another recent product from ESRI and Impact Observatory [Karra et al., 2021] to offer the highest resolution of any publicly available global LULC map. To create the ESA LULC map, researchers generated labeled pixels using GEOWIKI. These training labels are paired with S2 L2A multispectral and S1 RTC backscatter timeseries. Cloudy, saturated, or shadowed S2 pixels are removed using the scene classification layer of the L2A product. A gradient boosted decision tree algorithm (CatBoost) is then used for model training [Dorogush et al., 2017]. As a final step, auxiliary datasets such as OpenStreetMaps [OpenStreetMaps Contributors, 2021], Global Human Settlement Layer [Corbane et al., 2020, 2021], and Global Surface Water Explorer [Pekel et al., 2016] are used in the determination of expert rules for improving prediction quality.

The ESA WorldCover 2020 Map V100 achieves an overall accuracy of 74.4% across the 11 classes simulated on the holdout validation tiles. Given the map's extent, spatial resolution, and accuracy, its LULC classifications are used for all image triplets in Ethiopia, Uganda, and Sumatra – AOIs without region-specific mapping products of their own.

2.4 Dataset Collection using the Descartes Labs Platform

All images contained within the SEN12TS dataset are generated and processed by Descartes Labs. Using the Descartes Labs platform, images are organized into imagery sets, defined as a grouping of imagery containing: a single timestep's (time $t' = t$) S2 spectral bands and scene classification map; a four timestep timeseries (time $t' = t..t-3$) of RTC S1 backscatter measurements; a four timestep timeseries (time $t' = t..t-3$) of InSAR interferogram coherence and phase measurements; local incidence angle and SRTM-provided ground slope information; and auxiliary land cover classifications. A maximum difference in collection time of three days is specified for S1-derived and S2 imagery at time $t' = t$. A single imagery set covers one tile, defined as a 256-by-256-pixel square at 10m resolution in the local Universal Transverse Mercator (UTM) coordinate representation. All layers composing an imagery set are precisely coregistered for the tile in question; layers with a native resolution larger than 10m are upsampled using nearest neighbor resampling.

All twelve S2 bottom-of-atmosphere corrected L2A spectral bands and the scene classification map at time $t' = t$ are included in an image triplet. To filter out cloudy scenes, images are limited to only those from 100km-by-100km S2 granules with less than 10% aggregate cloud cover. The Descartes Labs S2 cloud mask product is then used to determine cloud coverage at the pixel level. While this product prvoides an industry-competitive S2 cloud layer, it, like all cloud masks, is imperfect. Highly reflective pixels covering buildings or snow-covered landscapes are sometimes misclassified as clouds, while diffuse cirrus clouds are sometimes missed.

The SAR-derived timeseries contain four timesteps 12 days apart, with the latest timestep (time $t' = t$) falling within three days of the collection day of the S2 image; accordingly, the first timestep in the timeseries occurs approximately 36 days before the collection day of the paired optical bands. Collecting image timeseries in twelve-day increments ensures that all SAR-derived layers within an imagery set are taken from the same satellite pass.

RTC SAR backscatter layers (γ_0) are created by combining terrain corrected backscatter layers (σ_0) and RTC factors. Here, σ_0 layers are generated on a burst-by-burst basis using geocoded SLCs as an intermediate product. The correction factors are generated using SAR metadata, digital elevation maps, and orbit information; they are unique for each ground-surface pixel and satellite pass, and do not contain any radar imagery. The correction factors also include a Descartes Labs-produced mask indicating whether σ_0 measurements fall within shadow or layover areas – regions that cannot be properly imaged due to steep surrounding slopes and the sensing geometry of the radar instrument.

Interferogram coherence and phase measurements are included for each of the four timesteps in the imagery set. Here, interferograms are created by using SAR data taken 12 days apart and are associated with the date of the second SAR collection: An interferogram created from SAR data at timesteps $t' = t$ and $t' = t - 1$ is associated with time $t' = t$, where timesteps are given in 12-day increments. Interferograms are generated using a Gaussian kernel with a spatial wavelength of 80m (corresponding to a spatial resolution of $\sim$50m). The resultant coherence and phase layers have a spatial resolution of approximately 20m, which is then upsampled via nearest neighbor resampling to 10m to align with the optical and backscatter layers.

Imagery triplets are only included in the SEN12TS dataset if the S1 and S2 images each contain fewer than 1% of pixels deemed invalid, due to either cloud cover, shadow, or layover [†]. To inform users as to which pixels are valid for each image, shadow/layover and cloud masks are included as the last layers in the S1 and S2 imagery stacks, respectively. The shadow/layover masks appended to the radar-derived S1 imagery contain 0s where the pixel falls

[†]Imposing an invalid pixel limit of 0% results in too few available triplets in Uganda and Sumatra, tropical regions with persistent cloud cover.

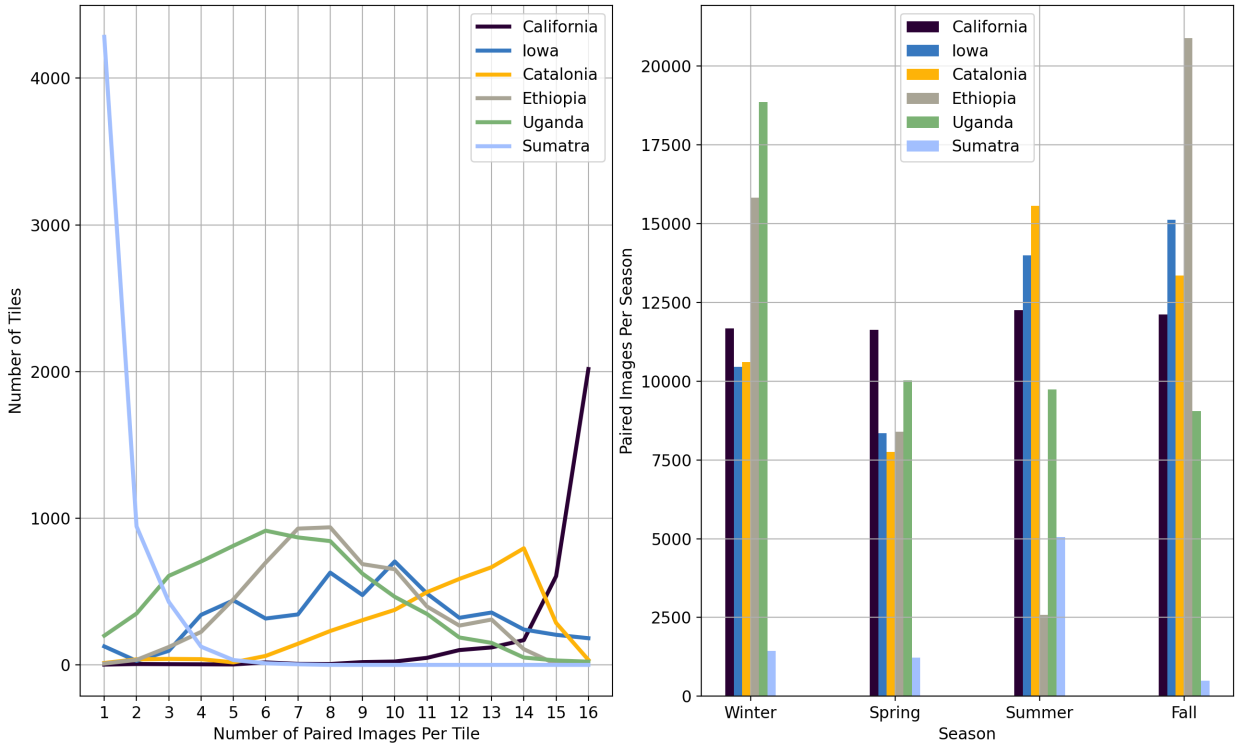

Figure 2: Statistics on imagery availability by (a) tile and (b) season for the six SEN12TS AOIs.

within the generated SAR shadow/layover mask and 1s over valid pixels; the cloud masks appended to the S2 imagery contains 0 over pixels deemed cloudy or shadowed per the utilized cloud shadow mask and 1s over valid pixels. No post-processing is performed on any of the sensed layers; users can test their own methods of post-processing – e.g. despeckling – with the SEN12TS dataset.

To evaluate land cover change over a period of time, multiple image triplets are collected for the same tile, subject to imagery availability. Here, the calendar year is divided into four seasons: winter, stretching from December 1st to February 28th; spring, stretching from March 1st to May 31st; summer, stretching from June 1st to August 31st, and fall, stretching from September 1st to November 30th. Imagery collection is limited to 4 triplets per tile per season, resulting in a maximum of 16 image triplets per tile. All imagery is collected over the 2020 calendar year.

Figure 2 presents regional statistics for the SEN12TS imagery: (a) displays the number of tiles (i.e. unique spatial extents) that contain a set quantity of image triplets; (b) shows how these images are distributed by season. Figure 2(a) demonstrates that California and Catalonia AOIs contain more tiles with a higher number of image triplets. Due to persistent cloud cover limiting the number of available optical images in Uganda and Sumatra, most tiles in these AOIs offer few valid image triplets.

3 SEN12TS Dataset

The following section discusses the structure, novelties, and caveats of the SEN12TS dataset.

3.1 Dataset Structure

The SEN12TS dataset is organized by AOI and imagery type. Each AOI contains */s1/*, */s2/*, and */labels/* folders that hold the S1, S2, or labeled LULC images respectively. S1 and S2 images are saved with the same file name, differentiated by their S1 or S2 parent folder. S1/S2 image file names are saved the following format:

lat_AAA_lon_BBB_s1date_YYYY-MM-DD_s2date_YYYY-MM-DD_s1track_CCC_s1lookpass_EF.tif

Table 3: Layer organization within S1, S2, and labeled land-use/land-cover (LULC) image triplets.

Layer	S1 Image	S2 Image	Labeled Image
0	SAR *vh*, time $t' = t$	Coastal aerosol, time $t' = t$	LULC classification
1	SAR *vv*, time $t' = t$	Blue, time $t' = t$	
2	InSAR coherence, time $t' = t$	Green, time $t' = t$	
3	InSAR phase, time $t' = t$	Red, time $t' = t$	
4	SAR *vh*, time $t' = t - 1$	Red edge, time $t' = t$	
5	SAR *vv*, time $t' = t - 1$	Red edge 2, time $t' = t$	
6	InSAR coherence, time $t' = t - 1$	Red edge 3, time $t' = t$	
7	InSAR phase, time $t' = t - 1$	Near infrared, time $t' = t$	
8	SAR *vh*, time $t' = t - 2$	Red edge 4, time $t' = t$	
9	SAR *vv*, time $t' = t - 2$	Water vapor, time $t' = t$	
10	InSAR coherence, time $t' = t - 2$	Short-wave infrared 1, time $t' = t$	
11	InSAR phase, time $t' = t - 2$	Short-wave infrared 2, time $t' = t$	
12	SAR *vh*, time $t' = t - 3$	Scene classification	
13	SAR *vv*, time $t' = t - 3$	Cloud mask	
14	InSAR coherence, time $t' = t - 3$		
15	InSAR phase, time $t' = t - 3$		
16	Local incidence angle		
17	SRTM slope		
18	Shadow/layover mask		

where *AAA* and *BBB* are the latitude and longitude, respectively, of the center of the tile; *YYYY-MM-DD* refers to the year, month, and day of the S1 and S2 imagery collection dates at time $t' = t$, which will differ by up to three days; *CCC* is the unique S1 orbit number on which the S1 image was collected; *E* refers to the look direction of the S1 satellite, either R (right) or L (left); and *F* refers to the S1 orbital pass, either A (ascending) or D (descending).

In contrast, one labeled LULC image is included for each unique tile in the dataset. As up to 16 paired S1/S2 images exist for a tile, fewer labeled LULC images exist than S1/S2 images; the same LULC image should be used for all S1/S2 images covering a specific tile. Labeled LULC images adhere to a similar file naming convention to the S1 and S2 images:

lat_AAA_lon_BBB_date_2020_lulc_label.tif

where *AAA* and *BBB* are the latitude and longitude, respectively, of the center of the tile, values that can be used to link the labeled LULC layer to S1/S2 images over the same spatial extent. All images are saved as georeferenced TIF files.

Table 3 presents the organization of layers within the image triplets. S1 image timesteps are 12 days apart, corresponding to a 36-day difference between the first (time $t' = t$) and last (time $t' = t - 3$) sets of radar imagery within an image triplet. S1 image stacks are saved as 32-bit floats. SAR backscatter layers are presented in decibels (dB) with a physical range between -40 dB and 30 dB; InSAR coherence is presented as a fraction between 0 and 1; InSAR phase is presented as a value between -π and π; and the local incidence angle and SRTM slope values are presented as degrees. In contrast, S2 and labeled LULC images are saved as 16-bit signed integers. Optical layers in the S2 stack contain values between 0 and 10,000 – representing the range between zero and full reflectance – while values in the scene classification layers range between 0 and 11. Labeled LULC images are also saved as 16-bit signed integers, containing region-specific raster values corresponding to each regions' LULC layer data source.

3.2 Dataset Novelties

The SEN12TS dataset is unique in six distinct ways from other publicly available combined SAR-optical datasets:

1. Four-timestep timeseries of SAR-derived images (RTC backscatter and InSAR phase and coherence) are collected for each imagery set. Moreover, multiple images for the same tile are contained within the SEN12TS datasets, thus providing coverage of the same spatial extent at multiple times during the year.
2. InSAR coherence and phase timeseries are included within each imagery set.
3. RTC backscatter layers (γ_0) are provided. RTC allows for reliable information extraction on sloped surfaces, along with comparison of SAR backscatter measurements from different satellite passes.
4. Before radiometric terrain correction, backscatter layers (σ_0) are derived from SLC products, instead of GRD products. Derivation of backscatter measurements from SLC data ensures finer spatial resolution (10m compared to 20m) and a larger dynamic range (12-bit instead of 8-bit). Furthermore, RTC backscatter layers that are derived from GRD products can experience pixel offsets between scenes in the same swatch and/or low backscatter at tile edges; these issues are avoided in the SLC-derived γ_0 layers in the SEN12TS dataset.
5. Shadow/layover and cloud masks specify invalid pixel locations that may interfere with applications of the dataset. Moreover, S2 scene classification layers provide an assessment of which pixels contain clouds, shadows, vegetation, desert, water, or snow; these scene classification layers can then be used to derive additional cloud, shadow, or snow masks.
6. Decameter resolution LULC maps are included within each imagery set. By adding raster layers from the USDA CDL, SIGPAC, or the ESA 2020 WorldCover Map v100 to the dataset, sensed imagery is combined with reliable, highly descriptive raster layers that can be used for a variety of classification-dependent use cases.

3.3 Dataset Caveats

While multiple steps are taken to minimize the amount of cloud cover in the SEN12TS dataset – limiting optical imagery collection to S2 imagery with less than 10% aggregate cloud cover; using Descartes Labs cloud masks to remove clouded and shadowed pixels; and only saving an image triplet if its S2 image contains less than 1% cloud cover – some S2 images still contain cloud artifacts. These artifacts are most frequent in imagery collected over Uganda and Sumatra, as these are tropical regions with persistent cloud cover. To quantify the likelihood of a cloud artifact, 100 S2 images are randomly selected from each region and then inspected to determine the number that contain minimal and non-minimal cloud artifacts; these results are presented in Table 4. Generally, minimal cloud cover is deemed to be cloud cover/cloud shadow that affects approximately 1% of the image or less; non-minimal cloud cover refers to cloud artifacts that occupy more than 1% of the image, or cirrus cloud cover that distorts the natural image coloration. For reference, Figure 5 in Appendix A contains images showing what qualifies as minimal and non-minimal cloud cover. While these cloud artifacts most frequently occur in small clusters of pixels, they nevertheless could pose problems for end-users if unaccounted for.

Table 4: Number of S2 images containing cloud artifacts out of a random sample of 100 images per region.

Area of Interest	Zero Cloud Artifacts	Minimal Cloud Artifacts	Non-Minimal Cloud Artifacts
California	100	0	0
Iowa	94	4	2
Catalonia	98	2	0
Ethiopia	97	2	1
Uganda	83	14	3
Sumatra	77	19	4

The cloud mask products used for the SEN12TS dataset creation also contain errors of commission. In these instances, highly reflective buildings and highly absorptive surfaces (typically standing water) are sometimes mistaken for cloud and cloud shadow, respectively. While these incorrectly masked pixels seem occur quite rarely within the dataset, it is an issue worth mentioning.

Additionally, users may encounter issues when using S1 backscatter measurements from different orbital tracks for change detection. Applying the radiometric terrain correction to backscatter measurements from different tracks should align γ_0 values to within 0.1-0.2 dB if there have been no changes to the ground geometry; however, comparing backscatter measurements from a single orbital track will provide the most sensitive detector of change.

Lastly, Figure 1 reveals that the availability of imagery is highly dependent on the collection geometries of the S1 and S2 satellites. If both S1 and S2 satellites pass regularly over a sub-area of an AOI, tiles within that sub-area will contain more images; conversely, tiles in sub-areas with less frequent satellite coverage will contain fewer images. Separate from issues of image availability, certain highly sloped (see north-central Ethiopia, Figure 1(a)) and highly urban sub-areas (see Barcelona area, Figure 1(c)) of the AOIs contain very little imagery. Here, steep surfaces create shadow and layover effects in the radar imagery, while highly reflective urban setting are misidentified as clouds. As these sub-areas often contain images with more than 1% of pixels that reside within the shadow-layover mask or images with more than 1% of pixels mislabeled as clouds, they are underrepresented in the final SEN12TS dataset.

4 Dataset Applications

The following section contains two deep-learning applications of the SEN12TS dataset.

4.1 Enhanced Vegetation Index Predictions Using Pix2Pix

Since their introduction in 2014, researchers have deployed many types of Generative Adversarial Networks (GANs) for remote sensing tasks [Goodfellow et al., 2019]. Among GANs that perform paired image transformations, the Pix2Pix architecture [Isola et al., 2017] has proven robust in applications such as LULC classification [Lin et al., 2017, Wang et al., 2018], change detection [Lebedev et al., 2018], and radar-to-optical translation [Enomoto et al., 2018, Zhang et al., 2020, Reyes et al., 2019]. To explore the applicability of the SEN12TS dataset, a GAN based on the Pix2Pix architecture is developed to translate SAR input imagery into an EVI predictions. This GAN is termed the SAR2VI. Using the near infrared (NIR), red, and blue spectral bands provided in the SEN12TS dataset, EVI is derived per Eq. (1):

$$EVI = \frac{2.5 * (NIR - RED)}{NIR + 6 * RED - 7.5 * BLUE + 10000} \tag{1}$$

Image triplets from the SEN12TS dataset are used to train the SAR2VI in the following manner: The 16 S1 bands associated with an image (*vh* and *vv* backscatter, and InSAR coherence and phase for 4 timesteps, $t' = t..t-3$) are given as inputs to the SAR2VI, while EVI layers are derived from the image triplet's S2 bands. Generated images are then compared with derived EVI layers to produce the losses that drive model training. EVI is predicted instead of the normalized difference in vegetation index (NDVI), as NDVI tends to saturate over dense vegetation and the SAR2VI generates EVI predictions with lower average pixel errors.

The SAR2VI is implemented as introduced in Isola et al. [2017], with one important modification: An L1 crop loss is added during training, proportional to the absolute difference between target values and predictions over cropland. This loss incentivizes accurate predictions over cropland, as is required in agricultural monitoring applications. For model training and testing, a portion of the SEN12TS imagery in California is selected; the USDA CDL included in these image triplets is used to determine the L1 crop loss. Random selections of 3004 and 752 image triplets constitute the training and testing datasets, respectively.

Figure 3 displays relevant layers showing the radar-optical translation process. Column A presents *vh* backscatter in the red and green bands of the images and *vv* backscatter in the blue band, both at time $t' = t$; accordingly, pixels with high *vh* values look yellow, while pixels over land cover types that do not change the polarization of the radar signal appear blue. Column B shows InSAR coherence – a measure of the consistency of the scattering processes across two SAR collects – also at time $t' = t$, with low coherence pixels appearing darker than their high coherence counterparts. In this column, surface land cover types that demonstrate little change over time, such as building or pavement, will display high coherence; land cover types with changing surface geometry, such as vegetation with leaves or branches that can move by many C-band wavelengths between passes, will induce different scattering patterns and exhibit lower coherence.

In the four displayed scene settings, alignment exists between the *vh* backscatter layers in Column A, the pixels with low coherence values in Column B, and the S2 derived EVI layer, presented in Column D. For reference, Column E presents the settings in true color RGB. Column C shows the SAR2VI-produced EVI predictions. Comparing between Column C and D, general agreement is observed between the predictions and target EVI values: high and low EVI regions are predicted as such, albeit with a loss of edge sharpness.

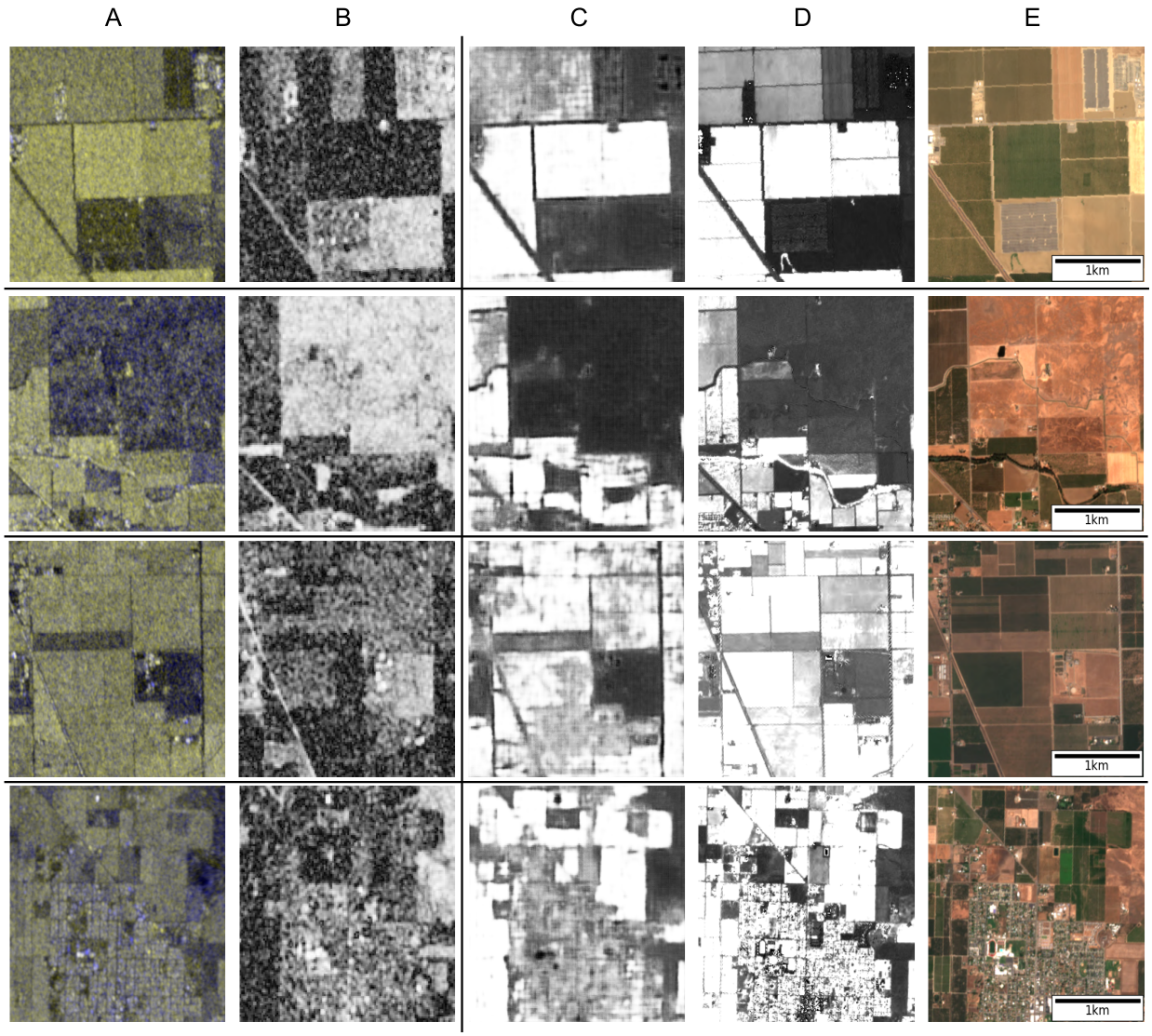

Figure 3: Column A presents SAR backscatter at time $t' = t$, with the *vh* layer in both red and green bands and the *vv* layer in the blue band; Column B presents single band images containing InSAR coherence at time $t' = t$; Column C presents SAR2VI-generated EVI predictions; Column D presents Sentinel 2-derived EVI measurements; Column E presents true-color RGB images of the landscapes. Rows 1-4 present images collected around time t= September 16, 2020; August 14, 2020; April 1, 2020; and October 18, 2020, respectively. All image settings are in California.

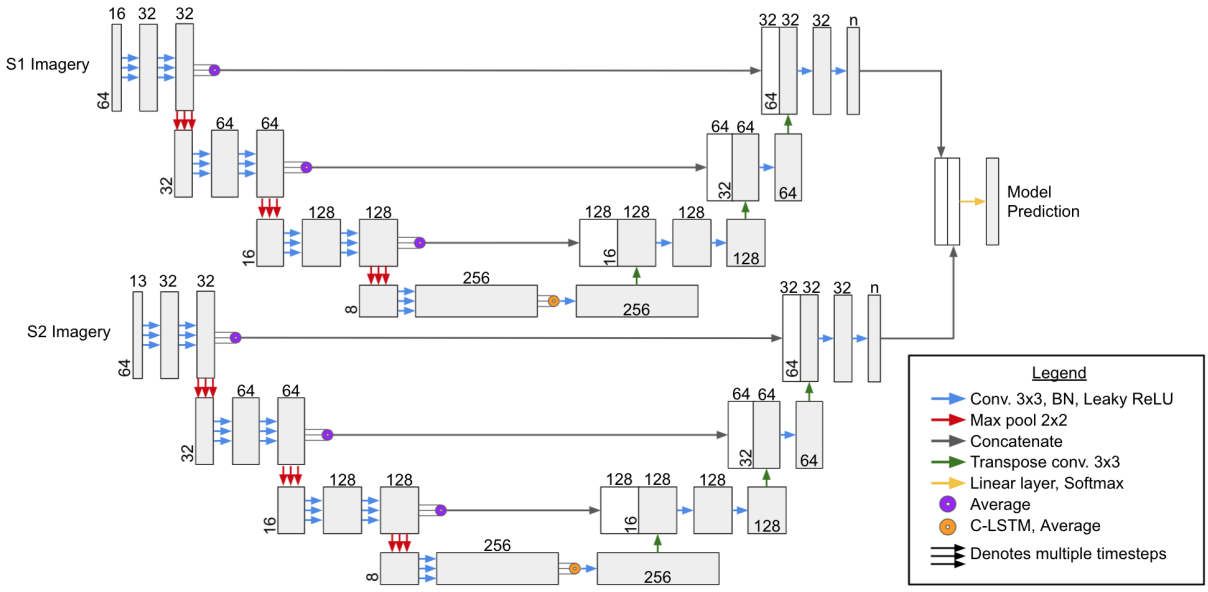

Figure 4: Network architecture adapted from M Rustowicz et al. [2019] for LULC classification.

4.2 Crop Classification

The SEN12TS dataset is also used to test crop classification models in California with the USDA CDL providing target labels. Here, input data consist of only imagery from tiles that have 16 available image triplets. The classification model is then trained on these S1 and S2 timeseries to predict a single land cover class for each pixel in the tile.

A random selection of 160 tiles with all 16 image triplets compile the training dataset. The unique land cover classes present across these images are then assessed: Labels from the 10 most frequently occurring classes are kept; all labels not contained in these 10 classes are ignored. The images are then split into into 64 pixel sub-images – each 256-by-256 image in the SEN12TS dataset yields 16 of these sub-images. All sub-images without any labeled pixels in the 10 most frequent USDA CDL classes are removed, leaving a total of 1817 sub-tiles in the training set. This process is repeated for 40 separate, randomly selected tiles to create the test dataset, which after filtering contains 454 sub-tiles worth of imagery.

Classification is performed with a model first introduced in M Rustowicz et al. [2019], adjusted for input imagery with 16 timesteps and spatial dimensions of 64x64 pixels; this adjusted network architecture is shown in Figure 4. Here, input imagery was cropped to 64x64 pixels to accommodate GPU memory constraints. The classification model contains separate modified UNets for processing the S1 and S2 imagery; when both S1 and S2-derived imagery are present during training, the separate network outputs are concatenated before the multilevel perceptron (MLP) and prediction layers. Three configurations of input data are tested: The first uses only S1 input imagery (4 bands per timestep), the second uses only S2 input imagery (12 bands per timestep), and the third uses both inputs as is seen in Figure 3. The following parameters are applied for all training: a batch size of 4; an Adam optimizer with a learning rate of 1e-3, a gradient *clipnorm* of 3 and a gradient *clipvalue* of 0.5 [Kingma and Ba, 2015]; and a categorical cross-entropy loss that is only applied over pixels classified as one of the 10 predicted classes. This loss is weighted for class balance per [King and Zeng, 2001], so that the models do not learn to overpredict the more frequently occurring classes. All input imagery is standardized bandwise to a mean of 0 and standard deviation of 1, per the training dataset.

Table 5 presents top-K accuracies for the three tested model configurations for the 10-class LULC classification task. Here, results indicate that the model trained on S2 imagery outperforms the model training on S1 imagery, and that model trained on combined S1 and S2 imagery produces the highest test set accuracies. This combined model achieves a top-1 accuracy of 64.9%, a top-3 accuracy of 85.9%, and a top-5 accuracy of 94.4%.

Table 6 presents the top-1 prediction confusion matrix for the combined S1 and S2 model. This table contains the predicted and actual classes for the entire test dataset; the crop types associated with the 10 most frequently occurring classes constitute the row and column names.

Table 5: Top-K accuracy for crop classification models.

	Top K Accuracy (%)		
	K=1	K=3	K=5
S1 imagery only	58.2	81.8	90.6
S2 imagery only	62.4	84.4	94.1
Combined S1 + S2	64.9	85.9	94.4

Table 6: Top-1 prediction confusion matrix for combined S1 and S2 classification model.

		Predicted										
		Almonds	Pistachios	Grapes	Grassland	Alfalfa	Fallow	Walnuts	Winter Wheat	Tomatoes	Developed	Total
Actual	Almonds	381,335	10,131	36,637†	18,488	5370	34,327	11,965	7908	8301	21,361	535,823*
	Pistachios	8970	147,155	8667	3718	204	46,686‡	2339	1731	2649	6816	228,935*
	Grapes	43,465†	12,524	116,983	3969	4173	7924	12,250	1354	4532	7948	215,122*
	Grassland	2502	1146	9921	49,530	2046	13,496	935	4519	220	6925	91,240
	Alfalfa	2106	1439	1416	4362	59,741	2575	924	4353	4574	2498	83,988
	Fallow	771	1721	1328	1,5271	183	84,625	141	3173	7647	5658	120,518**
	Walnuts	2260	429	5636	574	370	1099	12,553	608	427	1435	25,391**
	Winter Wheat	719	524	954	6811	1148	14,504‡	111	24,573	6473	3266	59,083
	Tomatoes	252	139	829	142	568	3727	64	1347	44,398	988	52,454
	Developed	5200	988	1494	2663	1155	5315	1032	2241	1221	26,435	47,744**
	Total	447,580*	176,196*	183,865*	105,528	74,958	214,278**	42,314**	51,807	80,442	83,330**	1,460,298

Table 6 shows that the combined S1 and S2 model underpredicts the top-3 most frequently occurring classes – almonds, pistachios, and grapes (highlighted with *) – while generally overpredicting the more underrepresented classes, such as fallow, walnuts, and developed (highlighted with **). Certain pairs of classes are also more frequently confused than others: the model mistakes grapes for almonds (and vice versa) in approximately 10% of predictions for those classes (see † highlights). Similarly, approximately 20% of actual pistachio and winter wheat pixels are each predicted as fallow (see ‡ highlights). These results are presented as an initial investigation of how the SEN12TS dataset can be used for LULC classification; a full exploration of classifier behavior on the SEN12TS dataset is left for future work.

5 Conclusion

This paper presents a novel dataset of combined Sentinel 1-derived (S1) radar imagery, Sentinel-2 (S2) optical imagery, and land-use land cover (LULC) labels over six unique global settings (California, Iowa, Catalonia, Ethiopia, Uganda, and Sumatra), all collected during 2020. The dataset is named SEN12TS, due to 1) the inclusion of radar imagery timeseries in every image triplet, and 2) inclusion of up to 16 images covering the same spatial extent in the dataset. The radar imagery in the SEN12TS dataset consists of radiometric terrain corrected (RTC) synthetic aperture radar (SAR) backscatter layers and interferometric synthetic aperture radar (InSAR) coherence and phase layers; for an image triplet, radar images are collected at four timesteps 12 days apart ($t' = t..t-3$). Radar imagery timeseries are matched with 12 S2 spectral bands and a scene classification layer collected at time $t' = t$ (defined as within 3 days of the S1 pass at time $t' = t$). For all settings, local incidence angle measurements, ground slope information are appended to the S1 images. All image triplets also contain labeled LULC raster layers. In the United States, the United States Department of Agriculture (USDA) Cropland Data Layer (CDL) provides the land cover labels; in Catalonia, labels are derived from Sistema de Información Geográfica de Parcelas Agrícolas (SIGPAC); in Ethiopia, Uganda, and Sumatra, labels are extracted from the European Space Agency (ESA) 2020 WorldCover Map v100. The entire SEN12TS dataset is collected, processed, and downloaded using the Descartes Labs platform.

Two applications of the SEN12TS dataset are also presented. The first predicts enhanced vegetation indices (EVI) given radar imagery inputs; here, a modified Pix2Pix generative adversarial network (GAN) [Isola et al., 2017] is trained and deployed over a subset of the image triplets collected in California. For the second application, a land cover classifier

is developed using combinations of S1 imagery, S2 imagery, and the USDA CDL, also for a subset of the SEN12TS data collected over California. For the classification model that intakes both Sentinel-1 and Sentinel-2 imagery, a top-1 accuracy of 64.9% is achieved.

The SEN12TS dataset is hosted by the Radiant Earth Foundation, where it can be explored and accessed at the following link: `https://radiant.earth/link_to_come`. The dataset is publicly available with a non-commercial CC BY-NC 2.0 license.

Acknowledgments

The authors would like to acknowledge and thank Mike Warren, Piyush Agram, and Matt Calef for their invaluable work in developing radiometric terrain corrections and InSAR interferograms, along with their continuous help throughout the dataset collection and paper writing processes. The authors would also like to thank the team at Radiant Earth for hosting the dataset; and Simone Fobi for her advice, support, and expertise.

Appendix A: Examples of Cloud Cover in the SEN12TS Dataset

For reference, Figure 5 presents examples of minimal and non-minimal cloud cover of S2 images in the SEN12TS dataset. The approximate prevalence of these types of cloud cover in each area of interest can be found in Table 4.

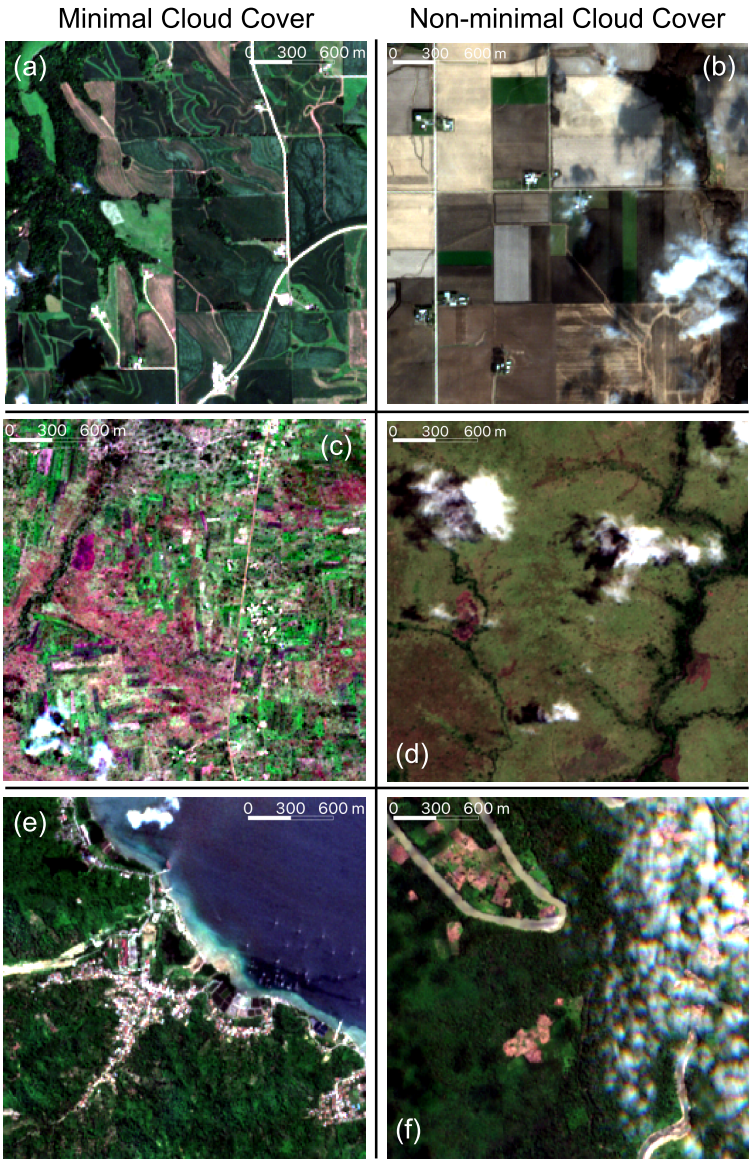

Figure 5: Example Sentinel-2 images in the SEN12TS dataset containing minimal and non-minimal cloud cover.

References

Xiao Xiang Zhu, Devis Tuia, Lichao Mou, Gui Song Xia, Liangpei Zhang, Feng Xu, and Friedrich Fraundorfer. Deep learning in remote sensing: a review. *IEEE Geoscience and Remote Sensing Magazine*, (December), 2017. ISSN 23318422. doi: 10.1109/MGRS.2017.2762307.

Guangliang Cheng, Feiyun Zhu, Shiming Xiang, Ying Wang, and Chunhong Pan. Semisupervised Hyperspectral Image Classification via Discriminant Analysis and Robust Regression. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(2):595–608, 2016. ISSN 21511535. doi: 10.1109/JSTARS.2015.2471176.

Christoph Rieke. Awesome Satellite Imagery Datasets, 2022. URL `https://github.com/chrieke/awesome-satellite-imagery-datasets`.

Christopher Yeh, Chenlin Meng, Sherrie Wang, Anne Driscoll, Erik Rozi, Patrick Liu, Jihyeon Lee, Marshall Burke, David B. Lobell, and Stefano Ermon. SUSTAINBENCH: Benchmarks for Monitoring the Sustainable Development Goals with Machine Learning. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*, number NeurIPS, 2021. URL https://openreview.net/forum?id=5HR3vCylqD.

Michael Schmitt, Lloyd Haydn Hughes, and Xiao Xiang Zhu. The SEN1-2 dataset for deep learning in sar-optical data fusion. *CoRR*, abs/1807.01569, 2018. URL http://arxiv.org/abs/1807.01569.

Michael Schmitt, Lloyd Haydn Hughes, Chunping Qiu, and Xiao Xiang Zhu. SEN12MS - A curated dataset of georeferenced multi-spectral sentinel-1/2 imagery for deep learning and data fusion. *CoRR*, abs/1906.07789, 2019. URL http://arxiv.org/abs/1906.07789.

Xiao Xiang Zhu, Jingliang Hu, Chunping Qiu, Yilei Shi, Jian Kang, Lichao Mou, Hossein Bagheri, Matthias Häberle, Yuansheng Hua, Rong Huang, Lloyd H. Hughes, Hao Li, Yao Sun, Guichen Zhang, Shiyao Han, Michael Schmitt, and Yuanyuan Wang. So2sat LCZ42: A benchmark dataset for global local climate zones classification. *CoRR*, abs/1912.12171, 2019. URL http://arxiv.org/abs/1912.12171.

Yuanyuan Wang and Xiao Xiang Zhu. The sarptical dataset for joint analysis of sar and optical image in dense urban area. In *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 6840–6843, 2018. doi: 10.1109/IGARSS.2018.8518298.

Xiaoxiang Zhu, Sina Montazeri, Mohsin Ali, Yuansheng Hua, Yuanyuan Wang, Lichao Mou, Yilei Shi, Feng Xu, and Richard Bamler. Deep learning meets sar: Concepts, models, pitfalls, and perspectives. *IEEE Geoscience and Remote Sensing Magazine (GRSM)*, 2022. URL https://elib.dlr.de/141023/.

NASA Jet Propulsion Laboratory. NASA Shuttle Radar Topogrpahy Mission Global 1 arc second [Data set], 2013. URL https://lpdaac.usgs.gov/products/srtmgl1v003/.

USDA National Agricultural Statistics Service. Cropland Data Layer (2020), 2020. URL https://nassgeodata.gmu.edu/CropScape/.

Government of Catalonia (GENCAT). SIGPAC Download, 2021. URL http://agricultura.gencat.cat/ca/serveis/cartografia-sig/aplicatius-tematics-geoinformacio/sigpac/descarregues/.

D. Zanaga, R. Van De Kerchove, W. De Keersmaecker, N. Souverijns, C. Brockmann, R. Quast, J. Wevers, A. Grosu, A. Paccini, S. Vergnaud, O. Cartus, M. Santoro, S. Fritz, I. Georgieva, M. Lesiv, S. Carter, M. Herold, Linlin Li, N.E. Tsendbazar, F. Ramoino, and O. Arino. ESA World Cover 10m 2020 v100, 2021.

Ramon Torres, Paul Snoeij, Dirk Geudtner, David Bibby, Malcolm Davidson, Evert Attema, Pierre Potin, Bj Örn Rommen, Nicolas Floury, Mike Brown, Ignacio Navas Traver, Patrick Deghaye, Berthyl Duesmann, Betlem Rosich, Nuno Miranda, Claudio Bruno, Michelangelo L'Abbate, Renato Croci, Andrea Pietropaolo, Markus Huchler, and Friedhelm Rostan. GMES Sentinel-1 mission. *Remote Sensing of Environment*, 120:9–24, 2012. ISSN 00344257. doi: 10.1016/j.rse.2011.05.028.

European Space Agency (ESA). Sentinel-1 Data Products, 2021. URL https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1/data-products.

David Small, Erich Meier, and Daniel Nüesch. Robust radiometric terrain correction for SAR image comparisons. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 3(C):1730–1733, 2004. doi: 10.1109/igarss.2004.1370666.

M. Drusch, U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, C. Isola, P. Laberinti, P. Martimort, A. Meygret, F. Spoto, O. Sy, F. Marchese, and P. Bargellini. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120:25–36, 2012. ISSN 00344257. doi: 10.1016/j.rse.2011.11.026.

European Space Agency (ESA). Level-2A Algorithm Overview, 2022. URL https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/level-2a/algorithm.

Avery Sandborn, Rick Mueller, Claire Boryan, Dave Johnson, Zhengwei Yang, Lee Ebinger, Arthur Rosales, Patrick Willis, Robert Seffrin, Rachel Jennings, Matt Deaton, and Hamer Hubert. NASS Geospatial Applications from the Cropland Data Layer. *ISI World Statistics Conference*, 2019. ISSN 9788578110796. doi: 10.1017/CBO9781107415324.004.

Tyler J. Lark, Richard M. Mueller, David M. Johnson, and Holly K. Gibbs. Measuring land-use and land-cover change using the U.S. department of agriculture's cropland data layer: Cautions and recommendations. *International Journal of Applied Earth Observation and Geoinformation*, 62(May 2016):224–235, 2017. ISSN 1872826X. doi: 10.1016/j.jag.2017.06.007.

Don Luman and Tari Tweddale. Assessment and Potential of the 2007 USDA-NASS Cropland Data Layer for Statewide Annual Land Cover Applications. 2008(49), 2008. URL https://www.ideals.illinois.edu/handle/2142/18134.

Qi Gao, Mehrez Zribi, Maria Jose Escorihuela, Nicolas Baghdadi, and Pere Quintana Segui. Irrigation mapping using Sentinel-1 time series at field scale. *Remote Sensing*, 10(9):1–18, 2018. ISSN 20724292. doi: 10.3390/rs10091495.

E. Corbelle-Rico, R. Crecente-Maseda, and I. Santé-Riveira. Multi-scale assessment and spatial modelling of agricultural land abandonment in a European peripheral region: Galicia (Spain), 1956-2004. *Land Use Policy*, 29(3):493–501, 2012. ISSN 02648377. doi: 10.1016/j.landusepol.2011.08.008.

Oscar D. Pedrayes, Darío G. Lema, Daniel F. García, Rubén Usamentiaga, and Ángela Alonso. Evaluation of semantic segmentation methods for land use with spectral imaging using Sentinel-2 and PNOA imagery. *Remote Sensing*, 13 (12):1–33, 2021. ISSN 20724292. doi: 10.3390/rs13122292.

Krishna Karra, Caitlin Kontgis, Zoe Statman-Weil, Joseph C. Mazzariello, Mark Mathis, and Steven P. Brumby. Global land use / land cover with sentinel 2 and deep learning. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pages 4704–4707, 2021. doi: 10.1109/IGARSS47720.2021.9553499.

Anna Veronika Dorogush, Andrey Gulin, Gleb Gusev, Nikita Kazeev, Liudmila Ostroumova Prokhorenkova, and Aleksandr Vorobev. Fighting biases with dynamic boosting. *CoRR*, abs/1706.09516, 2017. URL http://arxiv.org/abs/1706.09516.

OpenStreetMaps Contributors. OpenStreetMaps, 2021. URL https://planet.openstreetmap.org/.

Christina Corbane, Filip Sabo, Panagiotis Politis, and Vasileos Syrris. Ghs-built-s2 r2020a - ghs built-up grid, derived from sentinel-2 global image composite for reference year 2018 using convolutional neural networks (ghs-s2net), 2020.

Christina Corbane, Vasileios Syrris, Filip Sabo, Panagiotis Politis, Michele Melchiorri, Martino Pesaresi, Pierre Soille, and Thomas Kemper. Convolutional neural networks for global human settlements mapping from Sentinel-2 satellite imagery. *Neural Computing and Applications*, 33(12):6697–6720, 2021. ISSN 14333058. doi: 10.1007/s00521-020-05449-7. URL https://doi.org/10.1007/s00521-020-05449-7.

Jean François Pekel, Andrew Cottam, Noel Gorelick, and Alan S. Belward. High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633):418–422, 2016. ISSN 14764687. doi: 10.1038/nature20584.

Ian J. Goodfellow, Jean Pouget-Abadiey, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozairz, Aaron Courville, and Yoshua Bengio. Generative Adversarial Nets. *Proceedings - 2019 International Conference on Computer Vision Workshop, ICCVW 2019*, pages 3063–3071, 2019. doi: 10.1109/ICCVW.2019.00369.

Phillip Isola, Jun Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-Janua:5967–5976, 2017. doi: 10.1109/CVPR.2017.632.

Dao Yu Lin, Yang Wang, Guang Luan Xu, and Kun Fu. Synthesizing remote sensing images by conditional adversarial networks. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017-July:48–50, 2017. doi: 10.1109/IGARSS.2017.8126890.

Xiaoye Wang, Hongping Yan, Chunlei Huo, Jiayuan Yu, and Chunhong Pant. Enhancing Pix2Pix for Remote Sensing Image Classification. *Proceedings - International Conference on Pattern Recognition*, 2018-Augus(1):2332–2336, 2018. ISSN 10514651. doi: 10.1109/ICPR.2018.8545870.

M. A. Lebedev, Y. V. Vizilter, O. V. Vygolov, V. A. Knyaz, and A. Y. Rubis. Change detection in remote sensing images using conditional adversarial networks. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2:565–571, 2018. doi: 10.5194/isprs-archives-XLII-2-565-2018. URL https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-2/565/2018/.

Kenji Enomoto, Ken Sakurada, Weiming Wang, Nobuo Kawaguchi, Masashi Matsuoka, and Ryosuke Nakamura. Image translation between SAR and optical imagery with generative adversarial nets. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2018-July:1752–1755, 2018. doi: 10.1109/IGARSS.2018.8518719.

Jiexin Zhang, Jianjiang Zhou, and Xiwen Lu. Feature-guided SAR-to-optical image translation. *IEEE Access*, 8: 70925–70937, 2020. ISSN 21693536. doi: 10.1109/ACCESS.2020.2987105.

Mario Fuentes Reyes, Stefan Auer, Nina Merkle, Corentin Henry, and Michael Schmitt. SAR-to-optical image translation based on conditional generative adversarial networks-optimization, opportunities and limits. *Remote Sensing*, 11(17):1–19, 2019. ISSN 20724292. doi: 10.3390/rs11172067.

Rose M Rustowicz, Robin Cheong, Lijing Wang, Stefano Ermon, Marshall Burke, and David Lobell. Semantic segmentation of crop type in africa: A novel dataset and analysis of deep learning methods. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2019. URL https://openaccess.thecvf.com/content_CVPRW_2019/html/cv4gc/Rustowicz_Semantic_Segmentation_of_Crop_Type_in_Africa_A_Novel_Dataset_CVPRW_2019_paper.html.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In Yoshua Bengio and Yann LeCun, editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015. URL http://arxiv.org/abs/1412.6980.

Gary King and Langche Zeng. Logistic regression in rare events data. *Political Analysis*, 9(2):137–163, 2001. ISSN 15487660. doi: 10.18637/jss.v008.i02.