Deep Domain Adaptation based Cloud Type Detection using Active and Passive Satellite Data

Xin Huang*, Sahara Ali*, Chenxi Wang†,§ Zeyu Ning*, Sanjay Purushotham*, Jianwu Wang* and Zhibo Zhang‡
*Department of Information Systems, University of Maryland, Baltimore County (UMBC), Baltimore, MD, USA
†Joint Center for Earth Systems Technology, University of Maryland, Baltimore County (UMBC), Baltimore, MD, USA
‡Department of Physics, University of Maryland, Baltimore County (UMBC), Baltimore, MD, USA
§Earth Science Division, NASA Goddard Space Flight Center, Greenbelt, MD, USA
Email: {xinh1, sali9, chenxi, zeyning1, psanjay, jianwu, zhibo.zhang}@umbc.edu

***Abstract*—Domain adaptation techniques have been developed to handle data from multiple sources or domains. Most existing domain adaptation models assume that source and target domains are homogeneous, i.e., they have the same feature space. Nevertheless, many real world applications often deal with data from heterogeneous domains that come from completely different feature spaces. In our remote sensing application, data in source domain (from an active spaceborne Lidar sensor CALIOP onboard CALIPSO satellite) contain 25 attributes, while data in target domain (from a passive spectroradiometer sensor VIIRS onboard Suomi-NPP satellite) contain 20 different attributes. CALIOP has better representation capability and sensitivity to aerosol types and cloud phase, while VIIRS has wide swaths and better spatial coverage but has inherent weakness in differentiating atmospheric objects on different vertical levels. To address this mismatch of features across the domains/sensors, we propose a novel end-to-end deep domain adaptation with domain mapping and correlation alignment (DAMA) to align the heterogeneous source and target domains in active and passive satellite remote sensing data. It can learn domain invariant representation from source and target domains by transferring knowledge across these domains, and achieve additional performance improvement by incorporating weak label information into the model (DAMA-WL). Our experiments on a collocated CALIOP and VIIRS dataset show that DAMA and DAMA-WL can achieve higher classification accuracy in predicting cloud types.**

***Index Terms*—domain adaptation, remote sensing, cloud type detection, deep learning, multi-sensor, weak supervision**

I. Introduction

Cloud and atmospheric aerosols are two critical components that significantly impact Earth's radiative energy balance, hydrological and biological cycles, air quality and human health [1]. For example, clouds constantly cover about two-third of Earth's surface and alter global energy distribution by reflecting solar radiation and absorbing thermal emission from the surface. Satellite-based remote sensing is the only means to monitor the global distribution of aerosols and clouds. Thus, improvements in aerosol and cloud observations are a major focus of NASA's Earth Science endeavor, and numerous satellite sensors have been developed to observe and retrieve aerosol and cloud properties. They can be largely divided into two groups: 1) active sensors such as spaceborne Lidar (e.g., CALIOP) and Radar (e.g., CloudSat) and 2) passive sensors such as MODIS, VIIRS and ABI. Active sensors collect data by providing their own source of energy to illuminate the objects they observe; while passive sensors collect different sets of data attributes by detecting natural energy (e.g., solar radiation and thermal emission) that is emitted or reflected by the objects. The advantages of active sensors, compared to passive sensors, include their capability of resolving the vertical location of aerosol/cloud layer, better sensitivity to aerosol type and cloud phase and better performance during night-time and polar region. On the other hand, passive sensors observe column integrated radiation and have inherent weaknesses in differentiating atmospheric objects on different vertical levels. However, passive sensors always have wide swaths and better spatial coverage. Classifying cloud and aerosol types from passive sensors is an important application in satellite remote sensing due to its wide spatial coverage.

Our previous study [2] has shown proper use of machine learning (ML) algorithms, such as Random Forest (RF), can have better cloud type detection accuracy than physical-based algorithms. However, the algorithms cannot directly learn from multiple active and passive sensors. For example, RF can be either developed for CALIOP or VIIRS data, but it cannot jointly learn from both sensors since there is a mismatch of features/variables among the sensors. RF and other ML algorithms do not generalize to new combinations of the learned features beyond those seen during the training process. Moreover, many ML algorithms generally cannot do joint label predictions if the labels are missing for one of these sensors during training time. To address these issues, we employ deep learning based domain adaptation which can automatically learn feature representations from multiple sensors with different features/variables in an end-to-end way. Our model will be able to predict labels for all sensors even when the labels are absent for some sensors at training time by transferring knowledge from one set of sensors (e.g., CALIOP) to another set of sensors (e.g., VIIRS).

Domain adaptation, a transductive transfer learning technique to transfer knowledge representation from labeled source domain [3], has been thoroughly studied in computer vision [4], [5] and natural language processing (NLP) applications [6], [7]. Recently, the deep learning paradigm has become popular in domain adaptation due to its ability to learn rich, flexible, non-linear domain-invariant representations [8], [9]. However, few of these approaches have been adapted for

remote sensing applications. Moreover, domain adaptation techniques using deep neural network have been mainly used to solve the distribution drifting problem in homogeneous domains [10]. The data in the homogeneous domains usually share similar feature spaces and have the same dimensionalities. Nevertheless, real world applications often deal with heterogeneous domains that come from completely different feature spaces with different dimensionalities. In our remote sensing application, two remote sensor datasets collected by active and passive sensors respectively are heterogeneous. In particular, CALIOP actively detects hundreds of features within an atmospheric column. In this study, only 25 features that are sensitive to cloud phase and aerosol types are used. The selected CALIOP data are fully labeled with 6 cloud and aerosol types (see Section V). VIIRS is an imaging radiometer, which collects radiometric measurements of the surface and atmosphere in the visible and infrared bands. Quite a few physical-based algorithms are developed for detecting cloud and aerosols using VIIRS observations. However, it has been shown that passive sensors may have difficulties in prediction cloud/aerosol types in complicated atmospheres (e.g., overlapping cloud and aerosol layers, cloud over snow/ice surface, etc.). Our goal is to adapt the good cloud representation from active sensor (source domain) to passive sensor (target domain) and build a domain invariant classifier to accurately classify different cloud types (labels) in the passive sensing dataset.

The contributions of this paper are summarized as follows. We also open sourced our implementation in PyTorch[1].

- We develop a novel end-to-end deep domain adaptation with domain mapping and correlation alignment (DAMA) to learn domain invariant feature representation from multiple heterogeneous satellite remote sensing sensors.
- We extend the above DAMA model with weak supervision by incorporating the noisy/weak labels (DAMA-WL) from the target domain and achieve higher accuracy in cloud type detection.
- Experiments show our DAMA model achieves higher accuracy in detecting cloud types in the challenging passive satellite remote sensing data compared to other state-of-the-art ML models and our DAMA-WL can improve accuracy further by utilizing the weak labels.

The rest of the paper is organized as follows. Related work is introduced in Section II. Active and passive satellite remote sensing data and problem are described in Section III. Section IV introduces the end-to-end model of the deep domain adaptation based cloud type detection and the components of the model. Section V discusses the experiment and evaluation of our method. Section VI concludes the paper.

II. Related Work

Over the past few decades, a variety of remote sensing aerosol and cloud retrieval algorithms have been developed based on the physical principles and the radiative transfer of light scattering and absorption within aerosol and cloud fields (see review by [11]). These physical-based retrieval algorithms are the backbone of many widely used aerosol and cloud property products for weather and climate studies [12], [13]. Traditionally, many of these algorithms use a lookup table approach, in which one must prescribe aerosol and cloud types. The challenge is to ensure that the algorithm has the means to select the appropriate model.

Although highly successful, it is challenging to improve these physical-based algorithms. For example, according to [14], there is no absolute separation between "aerosol" and "cloud". Most, if not all, the physical-based retrieval algorithms that rely on threshold adjustments could have a lack of flexibility in consideration of sensor and environmental differences. Additional threshold adjustments and tests are required to apply an existing algorithm to different sensors. Thus, physical-based algorithms are expensive.

Machine learning (ML) and artificial intelligence (AI) techniques may overcome the challenges faced by physical-based algorithms. Since ML algorithms are written to autonomously find information (e.g., patterns of spectral, spatial, and/or time series data), they can learn hidden signatures of different types of objects. ML algorithms are portable and can be easily applied to active and/or passive sensor measurements. [2] introduced two Random Forest (RF) machine learning models for cloud mask and cloud thermodynamic phase detection using spectral observations from VIIRS data. [15] developed a deterministic self-organizing map (SOM) approach and applied it on satellite data based cloud type classification. Deep learning [16] is also a promising technique, already having revolutionized many fields such as computer vision [17], natural language processing [18], and is increasingly being used in remote sensing applications [19]. Those approaches can learn representations of multiple variables in a single domain.

Domain adaptation has been widely used in learning domain invariant representation from source and target domains. In unsupervised domain adaptation with unlabeled target domain, several approaches have been developed to minimize the feature distribution difference between the source domain and target domain. DCC [20] and DAN [21] have used Maximum Mean Discrepancy (MMD) loss to train a deep neural network and learn a representation that is both discriminative and domain invariant. [10] introduced a correlation alignment based method in the homogeneous domain adaptation in computer vision. Its architecture is based on CNN with a classification layer, and a correlation layer is used to minimize the difference in the second-order statistics between the source and target domains. [22] introduced an adversarial learning based domain adaptation method that combines adversarial learning with discriminative feature learning. It specifically learns a discriminative mapping of target images to the source feature space by simultaneously fooling a domain discriminator in distinguishing the encoded target images from source images. The state-of-the-art approaches are mostly applicable to homogeneous domain adaptation in image classification, in which

[1]DAMA source codes are available at: https://github.com/big-data-lab-umbc/deep-multisensor-domain-adaptation

the source and target domains are both two dimensional data and share similar feature space. To the best of our knowledge, few of the deep domain adaptation approaches have been used in the remote sensing application or in the heterogeneous domains [23], especially with the heterogeneous nature of datasets collected by active and passive sensors.

III. Data and Problem Description

In this section, we take a close look at the the heterogeneity of our source (active) and target (passive) remote sensing data, and its domain adaptation challenge.

A. Active and Passive Satellite Data

The Visible Infrared Imaging Radiometer Suite (VIIRS) [24], [25] is a passive instrument onboard polar orbiting satellites Suomi-NPP and JPSS-1, and will be JPSS-2 after 2022. VIIRS collects visible and infrared imagery and provides less than 1-km spatial resolution observations (native 750m for VIIRS moderate spatial resolutions bands) and wide spatial coverage. Passive sensors observe column-integrated radiation. Accuracy of cloud and/or other atmospheric particles detections could significantly decrease if the whole column is highly heterogeneous (e.g., multi-layered clouds with different thermodynamic phases). In contrast to passive satellite sensors that observe column-integrated, active sensors are more reliable in recognizing objects in different layers because of the high vertical resolutions. For example, the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) onboard CALIPSO satellite [26], [27] operates at wavelengths of 532nm and 1064nm, measuring lidar backscattering profiles at a 30m vertical and 333 m along-track resolution. CALIOP also measures the perpendicular and parallel signals at 532nm, along with the depolarization ratio at 532nm that is frequently used in cloud phase discrimination algorithms because of its strong particle shape dependence [27]. Although active sensors are very sensitive to cloud and aerosol layers, they have limited spatial coverage. By taking into account these strengths and weaknesses of both CALIOP and VIIRS, we intend to generate reliable label datasets based on CALIOP Level-2 (version 4) product. The VIIRS Level-1B observations and solar/satellite geometries, and the CALIOP and VIIRS Level-2 cloud mask and thermodynamic phase products will be used for training, validation, testing, and comparison.

A major difference between VIIRS and CALIOP sensor data is their spatial coverage. Figure 1 shows a coverage difference using a full day data collection using NASA Earth Data World View website [28]. VIIRS has nearly full coverage of the Earth while CALIOP only covers yellow line area which is much smaller than the coverage of VIIRS.

B. Domain Adaptation Challenge for Active and Passive Satellite Data

In heterogeneous domain adaptation, the feature spaces between the source and target domains are nonequivalent and the dimensions may also generally differs [23]. In our satellite remote sensing application, the source (active) domain

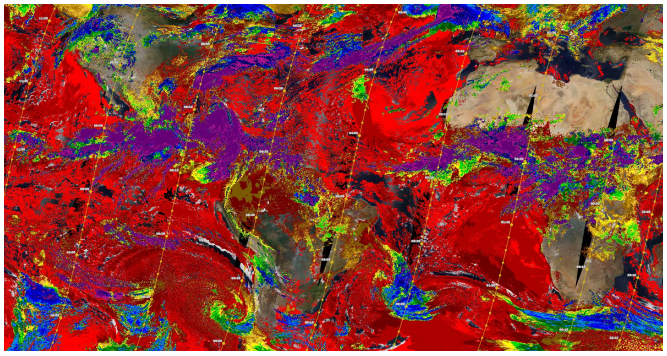

Fig. 1. An example showing the spatial coverage differences between VIIRS (global coverage) and CALIOP (yellow lines) data (Credits: NASA).

data, contains sensing data with 25 attributes collected by the CALIOP active spaceborne Lidar sensor while the target (passive) domain data contains another group of sensing data with 20 attributes collected by the VIIRS passive spectroradiometer sensor. The attribute names, descriptions and pairwise data distribution exploratory analysis can be found in [29]. Two remote sensing datasets have completely different feature spaces due to the nature of the data they collect, CALIOP data has better separations for the cloud types as the data are more evenly distributed compared to VIIRS in which the majority data are mixed together in the distribution.

Following our previous work at [2], we filter nighttime data records and choose the daytime records with $0 <$ Solar Zenith angle (SZA) < 80. Four auxiliary attributes shared in both CALIOP and VIIRS datasets are surface temperatures, surface emissivity, surface type and snow ice index. The latitude and longitude of the pixel are also provided in both CALIOP and VIIRS datasets. In total, there are 6 auxiliary features supplemented to both the source and target domains to train the domain adaptation model. The CALIOP cloud labels are used as reference label information in collocated CALIOP and VIIRS datasets. A collocation algorithm [30] that fully considers the spatial differences between the two instruments and parallax effects is used to generate our collocated datasets. Since cloud types derived from CALIOP dataset capture good representation of cloud, we use the CALIOP cloud types (label) as the ground truth for the collocated CALIOP and VIIRS data records.

As the VIIRS satellite has wider spatial coverage, by learning a domain invariant classifier from the collocated CALIOP and VIIRS data, we will be able to use it to classify different cloud types for the un-collocated VIIRS/Suomi-NPP satellite data that is off the track of CALIOP/CALIPSO satellite. The problem can be formulated as follows.

Given CALIOP (source domain) training examples $D_s = \{x_i\}, x_i \in R_s^{d_s}$ with labels $L_s = \{y_i\}, y_i \in \{1, ..., L\}, i \in \{1, .., N\}$ and unlabeled VIIRS (target domain) dataset $D_t = \{u_i\}, u_i \in R_t^{d_t}$, our model is to learn a domain invariant labeling function (classifier) $f : R_s^{d_s} \longrightarrow L_s$ with $f(D_s) = f(D_t)$.

IV. Deep Domain Adaptation based Cloud Type Detection using Active and Passive Satellite Data

The remote satellite sensing data raises more challenges as the data captured by passive sensor and active Lidar are high dimensional, globally covered and heterogeneous in nature. In this paper, we propose an end-to-end deep domain adaptation with domain mapping and correlation alignment (DAMA) and apply it to classify the heterogeneous remote satellite cloud and aerosol types. In training phase, there are two branches of inputs that are source domain data features and target domain data features that have different dimensionalities and heterogeneous feature spaces. As shown in Figure 2, our model introduces a heterogeneous domain mapping to transform the feature space of target domain into the feature space of source domain, and uses feature extraction layer to train the shared representative features between the source and target domain. At last, it adds a correlation layer to the end of the shared layers, inspired by the idea of correlation alignment introduced in [10]. By incorporating the correlation loss and classification loss in training the domain adaptation network, we find the network can maximize the classification accuracy on the target domain by minimizing the difference in the second-order statistics between the source and target domains. In the testing phase, only VIIRS (target domain) data is fed into the deep neural network by going through the deep domain mapping layer and feature extraction layer. The trained source classifier can then be applied to classify the output of the feature extraction layer as the domain invariant feature representation has been generated from the flow. Figure 2 demonstrates our end-to-end deep domain adaptation with domain mapping and correlation alignment which will be further explained in detail in the rest of the section.

A. Deep Domain Mapping (DDM)

To adapt to the completely different feature spaces, i.e., the heterogeneity of the source and target domain, we introduce a deep learning based approach to learn a transformation to map the target feature space into the source feature space. It equalizes the number of features in source and target domains, and also transforms both domains into the same feature space.

In our remote sensing dataset, the target domain (VIIRS) has wider spatial coverage but with no label information. The source domain (CALIOP) has better representation for cloud types and is fully labeled, so mapping the target domain to source domain can preserve the discriminating power of the source domain and can also transfer it into the down-streaming learner.

We design a deep neural network to perform the deep domain mapping (DDM) between the source and target domain. The input of the DDM network is the target domain data and the output of the network is the transformed target domain data in the source domain feature space. Because the source domain data and target domain data are collocated remote sensing data with the same longitude and latitude coordinates, mean squared error (MSE) loss function is used to measure the error of the DDM network. Specifically, given source domain training examples $D_s = \{x_i\}, x \in R_s^{d_s}, i = 1, ..., n_s$ and unlabeled target data set $D_t = \{u_i\}, u \in R_t^{d_t}, i = 1, ..., n_t$, with $d_s \neq d_t$ and $R_s^{d_s} \neq R_t^{d_t}$. Because the source domain and target domain are collocated data so we have $n_s = n_t$. The DDM is learnt to transform the target domain into source domain feature space by minimizing l_2 loss function:

$$l_2 = \frac{1}{n_t} \sum_{(i=1)}^{n_t} (DDM(u_i) - x_i)^2 \tag{1}$$

By minimizing the l_2 error we aim to map the features of the target domain into the feature space of the source domain that has better feature representation. The l_2 loss is co-trained with the correlation alignment and classifier losses in an end to end fashion by retaining the computation graph while training the deep domain mapping.

Our multiple domain experiments in Table II show DDM can significantly improve the classification accuracy, demonstrate that domain adaptation and correlation alignment (to be introduced in next section) work well on the multiple domain data from the same feature space. The proposed heterogeneous deep domain mapping network is also generic and flexible. It can be plugged into other domain adaptation methods and used in areas other than climate data analytics.

B. Domain Adaptation with Correlation Alignment

The domain adaptation in our DAMA model consists of a set of shared multilayer perceptron (MLP) feature layers (shown in Figure 2) that is used to extract the domain invariant representation between source and target domain, and a correlation layer that is used to minimize domain shift by aligning the second order statistics of source and target data distributions.

After transforming the target domain into the feature space of source domain via DDM, the dimension of the transformed target domain is identical to dimension of the source domain and the source domain and target domain become homogeneous. The correlation alignment can be formulated as follows.

Given source domain training examples $D_s = \{x_i\}, x_i \in R_s^{d_s}$ with labels $L_s = \{y_i\}, y_i \in \{1, ..., L\}, i \in \{1, ..., n_s\}$ and unlabeled transformed target data set $D_t = \{x_j^*\}, x_j^* \in R_s^{d_s}, j \in \{1, ..., n_t\}$, we can compute the covariance matrix of source domain and target domain, represented as C_s and C_t respectively:

$$C_s = \frac{1}{n_s - 1}(D_s^T D_s - \frac{1}{n_s}(\mathbf{1}^T D_s)^T(\mathbf{1}^T D_s)) \tag{2}$$

$$C_t = \frac{1}{n_t - 1}(D_t^T D_t - \frac{1}{n_t}(\mathbf{1}^T D_t)^T(\mathbf{1}^T D_t)) \tag{3}$$

where $\mathbf{1}$ is a column vector with all elements equals to 1.

We use the correlation loss proposed in [10] to measure the distance between the second order statistics (covariances) of the source and target data:

$$l_{coral} = \frac{1}{4d^2}||C_s - C_t||_F^2 \tag{4}$$

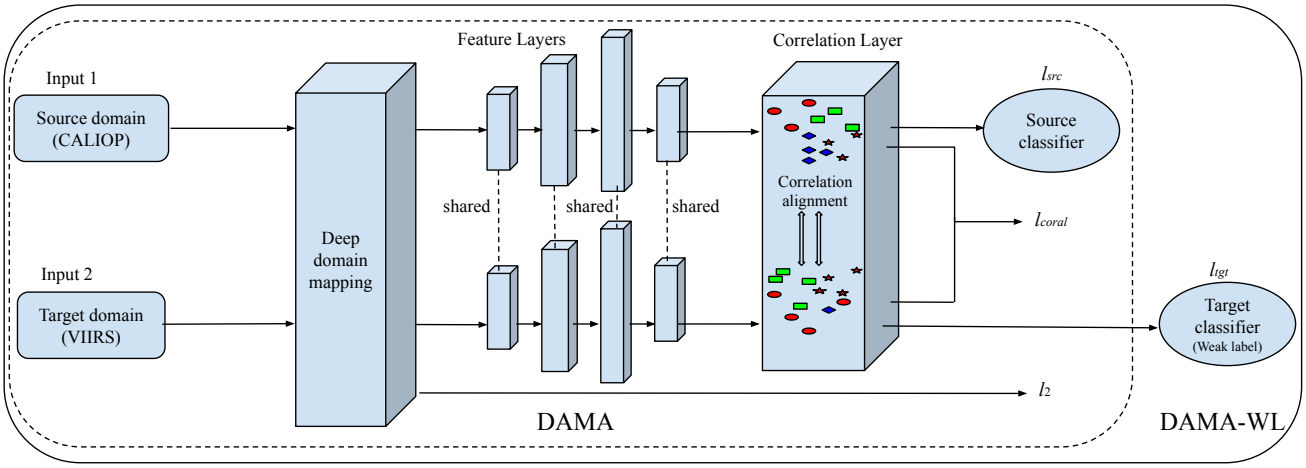

Fig. 2. DAMA: Network architecture of deep domain adaptation with domain mapping and correlation alignment. Deep domain mapping is used to map the target domain into the feature space of source domain. The model uses several multilayer perceptron (MLP) layers to learn the shared representative features between the source and target domain. A correlation layer is added to output of the feature extractor layer. At the end of the network is the source classifier that classifies the source domain in training phase. DAMA-WL adds a target classifier trained with weak label of the target domain in addition to DAMA.

, where $||.||_F^2$ denotes the squared matrix Frobenius norm and d is the number of the features.

By combining the correlation loss with the classification loss, the joint loss function is minimized to learn the latent features that can work well on the target domain:

$$l = l_{src} + \sum_{(i=1)}^{t} \lambda_i l_{coral} \tag{5}$$

Here, l_{src} is the source classifier loss calculated with the multi-class cross entropy, t is the number of the correlation layers in the deep network and λ_i is a weight that balances on the adaptation with the classification accuracy on the source domains. The classification loss l_{src} and the correlation loss l_{coral} play counterparts and reach an equilibrium at the end of training so that the representative capacities of the source domain can be adapted to the target domain, so the final classifier performs well on the target domain with higher accuracy.

C. Domain Adaptation with Weak Supervision

In the DAMA model, cloud information (i.e., the labels) in VIIRS dataset (target domain) is not used because the information is not accurate enough. In the remote satellite sensing application, inaccurate VIIRS labels are also generated from the physical-based retrieval algorithm because it could provide some information for the off-track pixels. We further extend our model to incorporate the weak label information from the target domain to help train the domain invariant feature, our overall learning task will change from unsupervised learning to weakly supervised learning [31] task. The extended deep domain adaptation model with weak label is called DAMA-WL, with the underlying principle to use additional supervision information as constraints to efficiently build the domain invariant classifier. Specifically, a target domain classifier with weak VIIRS label is added into the training phase in addition to the source domain classifier, and it will be co-trained with correlation alignment and the deep domain mapping, with target (VIIRS) classifier loss l_{tgt}, source (CALIOP) classifier loss l_{src} and correlation loss l_{coral} playing counterparts and converging, depicted in right part of Figure 2. In DAMA-WL, the joint loss is composed of the loss of source classifier, the loss of correlation alignment and the loss of target classifier:

$$l^* = l_{src} + \sum_{(i=1)}^{t} \lambda_i l_{coral} + l_{tgt} \tag{6}$$

D. End-to-End Training

The domain mapping and correlation alignment modules are trained jointly in an end-to-end fashion in order to align the heterogeneous source and target domains and build the domain invariant classifier. In each training epoch, the parameters of the correlation alignment and the domain mapping module are updated alternatively using back-propagation algorithm. That is, in every epoch, first, the model parameters are updated by using the gradients calculated from minimizing the joint loss function Eq. 5 of DAMA (or Eq. 6 of DAMA-WL), and then model parameters associated with the domain mapping part are updated by minimizing the l_2 loss of domain mapping layer. In this way, the end-to-end training of DAMA is performed.

V. Experiments

We conduct several experiments on real world remote sensing datasets to compare the performance of our proposed model with the state-of-the-art ML models.

A. Datasets and Evaluation Metrics

Our experiments use CALIOP active sensor (source) and VIIRS passive sensor (target) remote satellite sensing datasets [2] with two label settings. In the first setting, we consider ground-truth labels from CALIOP with six categories: 1) Clear

TABLE I
ACCURACY ON PREDICTING THE CLOUD TYPES ON VIIRS (TARGET) DATASET.

Models - Single Domain	**Source**	**Target**	**Day-005**	**Day-013**	**Day-019**	**Day-024**	**Day-030**	**Jan. 2017**
Random Forest	VIIRS	VIIRS	0.778	0.738	0.731	0.726	0.710	0.739
MLP-VIIRS	VIIRS	VIIRS	0.770	0.750	0.724	0.729	0.718	0.743
MLP-CALIOP	CALIOP	CALIOP	1.000	1.000	1.000	1.000	1.000	1.000
Models - Multiple Domains								
Domain Mapping Only	CALIOP	VIIRS	0.728	0.705	0.696	0.695	0.674	0.695
Correlation Align. Only	CALIOP	VIIRS	0.355	0.333	0.283	0.282	0.251	0.302
DAMA	CALIOP	VIIRS	**0.780**	**0.759**	**0.745**	**0.745**	**0.721**	**0.752**

and Clean (no cloud, no aerosol), 2) Pure Liquid Cloud (liquid cloud only, no aerosol), 3) Pure Ice Cloud (ice cloud only, no aerosol), 4) Ice and Liquid Cloud (cloud only, no aerosol), 5) Aerosol (aerosol only, no cloud) and 6) Aerosol and Cloud. In the second setting, we consider the aerosol-free pixels from CALIOP, i.e., categories 1, 2 and 3, as ground truth for source domain; we also consider aerosol-free pixels from VIIRS Cloud Top and Optical Properties Product [24] as weak labels in the target domain. The weak labels correspond to the following three categories: 1) Clear Sky (no cloud), 2) Pure Liquid Cloud, and 3) Pure Ice Cloud.

For the two label settings, training dataset is collocated 4 months (January 2013, January 2014, January 2015 and January 2016) CALIOP and VIIRS datasets with 5,633,322 records in first label setting and 4,711,554 records in second label setting. Each built model is evaluated by predicting the labels for one month, i.e., January of the subsequent year 2017. We further perform the testing on five individual days of January 2017 that include Day 5, 13, 19, 24 and 30. Figure 3 shows the class distribution against each cloud/aerosol type (class) for the training and test VIIRS datasets with CALIOP labels (first label setting). Similarly, Figure 4 shows class distribution against each cloud type (class) for the training and test VIIRS datasets with weak VIIRS labels (second label setting). Analyzing the class distribution in the training dataset for the first label setting, as illustrated in Figure 3, we can see some class imbalance with highest class label data available for "Pure Liquid" and lowest class label data available for "Pure Cloud". Similarly, for VIIRS weak label of the second label setting, we can see highest class label data available for "Pure Liquid" and lowest class label data available for "Clear Sky", as illustrated in Figure 4.

We used Accuracy as the evaluation metric to compare all the models:

$$Accuracy = \frac{\text{Total number of correct predictions}}{\text{Total number of data points}} \quad (7)$$

B. Performance Comparison using Data from Single Domain

For non-domain adaptation model comparison, we conducted experiments on three baseline models which were trained on data from a single domain. These baseline models include 1) RF model: Random Forest trained on VIIRS data, 2) MLP-VIIRS: A deep learning based MLP model trained on VIIRS data, 3) MLP-CALIOP: A deep learning based MLP model trained on CALIOP data.

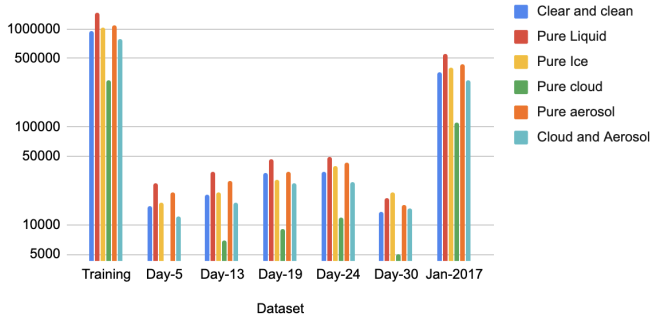

Fig. 3. Data distribution (data point count for each of 6 cloud/aerosol types) for training and test VIIRS datasets.

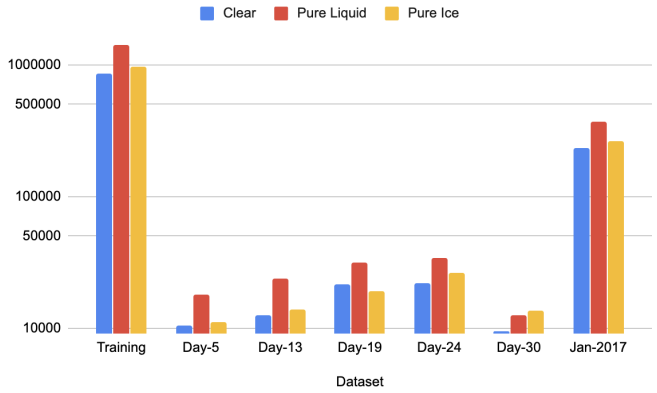

Fig. 4. Data distribution (data point count for each of 3 cloud types) for training and test VIIRS datasets.

In order to make fair comparison to our proposed model, we apply the same neural network used in the shared layer of our DAMA network to build the neural network for baseline models (MLP-CALIOP and MLP-VIIRS), with the same type and number of layers. In our experiments, the MLP (shared) layers are 4 dense layers with 128, 256, 128, 64 neurons respectively, each layer is followed with a ReLU activation function and Dropout (0.5). To train the RF model, we specify 100 as the number of trees and 15 as the maximum depth of the trees in the forest, chosen after hyperparameter tuning.

As shown in Table I, as an ML-based baseline result, RF achieves around 77.8% training, 77.0% validation and around 73.9% test accuracy (for January 2017). For the single domain experiments, we can see MLP-CALIOP achieves

TABLE II
ACCURACY ON PREDICTING THE CLOUD TYPES ON VIIRS (TARGET) DATASET WITH WEAK LABEL.

Models - Single Domain	**Label**	**Source**	**Target**	**Day-005**	**Day-013**	**Day-019**	**Day-024**	**Day-030**	**Jan. 2017**
Random Forest	CALIOP	VIIRS	VIIRS	0.957	0.947	0.934	0.933	0.917	0.939
Random Forest-WL	VIIRS	VIIRS	VIIRS	0.905	0.911	0.883	0.878	0.854	0.889
MLP-VIIRS	CALIOP	VIIRS	VIIRS	0.896	0.907	0.878	0.877	0.865	0.885
MLP-CALIOP	CALIOP	CALIOP	CALIOP	1.000	1.000	1.000	1.000	1.000	1.000
Models - Multiple Domains									
Domain Mapping Only	CALIOP	CALIOP	VIIRS	0.910	0.913	0.890	0.896	0.885	0.899
Correlation Align. Only	CALIOP	CALIOP	VIIRS	0.428	0.473	0.394	0.378	0.321	0.408
DAMA	CALIOP	CALIOP	VIIRS	0.956	0.948	0.934	0.936	0.926	0.941
DAMA-WL	CALIOP + VIIRS	CALIOP	VIIRS	**0.963**	**0.964**	**0.958**	**0.958**	**0.949**	**0.960**

100% accuracy in predicting the active sensing dataset, which is expected as the data distribution of each cloud type is very discriminative in the CALIOP. Our ultimate goal is to transfer the discriminative representation from this active sensor CALIOP to passive sensor VIIRS in order to accurately classify the cloud types in the passive dataset. In comparison, MLP-VIIRS model has lower accuracy around 74.3%, as VIIRS is a passive dataset collected by detecting the reflection of natural radiation and their feature discrimination power is weak. This observation highlights the importance of using multiple sensors data to better understand and classify the unlabeled passive sensing data that has wider spatial coverage. Our proposed deep domain adaptation model DAMA aims to achieve higher accuracy than using single domain data by transferring the discriminating power from the source domain to target domain.

C. Performance Comparison of using Data from Multiple Domains

For domain adaptation model comparisons, we conducted experiments on two more baseline models that use our heterogeneous domain mapping and correlation alignment respectively, using both source and target datasets. These baseline models include the following: 1) Domain Mapping Only: This model uses the deep domain mapping but no correlation alignment, 2) Correlation Alignment Only: This model uses the Correlation alignment but no deep domain mapping strategy. Comparing these baseline models with our proposed DAMA model can help understand the importance of each module in our model.

From the result of multiple sources based models in Table I, our proposed DAMA model outperforms the two domain adaptation baselines significantly. DAMA improves the accuracy by 5.7% on average of all the predictions from Day-005 to Day-030 when compared to using the Domain Mapping Only model. It also shows our approach improves more than 40% accuracy in comparison to the Correlation Alignment Only model with domain adaptation that uses the raw source and target features.

D. Performance of Domain Adaptation with Weak Supervision

We also evaluate our models' performance with the weak label from VIIRS. In this experimental setting, both training and testing of target domain are performed using the weak labels from VIIRS. Comparing VIIRS' weak labels with corresponding CALIOP category 2 and category 3 labels (further explained in Section V-A), we get around 87% label matchrate. The CALIOP label is considered as ground truth for the experiment, so we can consider the weak label is noisy of only 87% accuracy.

Table II introduces two more experiments Random Forest-WL and DAMA-WL that utilize weak label from VIIRS dataset. Random Forest-WL trains the model with weak label from VIIRS dataset of the second label setting, while DAMA-WL adds a target classifier trained with weak label of VIIRS dataset to DAMA. Other models shown in Table II use the CALIOP label from collocated CALIOP and VIIRS pixels. From Table II we can see DAMA-WL achieves highest accuracy 96.0% compared to the random forest models and other baseline models. We also see DAMA-WL brings additional 1.9% accuracy improvement compared to the DAMA method, which shows that the weak label does help train a better domain adaptation model in weak supervision on target domain.

E. Impact of Domain Mapping

The very low accuracy (around 40%) in predicting cloud satellite data with Correlation Alignment Only model exemplifies inherent complexities in heterogeneous data representation and the challenge of directly applying existing domain adaptation methods in heterogeneous domains. Our proposed deep domain mapping can mitigate the gap between the heterogeneous source and target domains and extract the domain invariant representation by integrating with the domain adaptation technique.

Our DAMA model's prediction accuracy also outperform Random Forest Model that is widely used in climate data. Supervised learning model such as Random Forest in the single domain assumes the label information on the target domain is available, in comparison, DAMA is unsupervised domain adaptation that does not require label information in the target domain, and solely rely on the label information of source domain and correlation between the source and target domain to build the model and make the prediction.

Furthermore, in domain adaptation, training on target domain with target labels is the gold standard in many domain adaptation applications as that is the best a model can achieve. So domain adaptation's performance is upper bounded by the performance on the target domain dataset (i.e., trained on the

target data and labels). However, typically in real applications target domain labels are unreliable or unavailable. Our problem setup is slightly different as the target domain labels are obtained from a different satellite (source domain) with co-located latitudes and longitudes.

VI. Conclusions

With the advances in remote sensing, we are seeing more and more satellites orbiting the Earth. By utilizing data from multiple satellites jointly, we could achieve better information retrieval for the targeted geophysics variables. Towards this goal, we presented a deep domain adaptation method with heterogeneous domain mapping and correlation alignment to employ both active and passive sensing data in cloud type detection. We further extended the domain adaptation model with weak supervision by using the noisy label from the target domain. Our experiments show that our proposed models can achieve higher accuracy in classifying the challenging passive remote sensing dataset by transferring the knowledge from the active sensing to the passive sensing dataset.

For future work, we plan to investigate taking into account the information of neighboring pixels and use of deep learning models that can capture spatial information (e.g., CNN and graph neural networks) could improve cloud type prediction.

Acknowledgement

This work is supported by grants OAC–1730250, OAC–1942714, IIS–1948399 from the National Science Foundation (NSF) and grant 80NSSC21M0027 from the National Aeronautics and Space Administration (NASA).

References

[1] O. Boucher, D. Randall, P. Artaxo, C. Bretherton, G. Feingold, P. Forster, V.-M. Kerminen, Y. Kondo, H. Liao, U. Lohmann, P. Rasch, S. K. Satheesh, S. Sherwood, B. Stevens, and X. Y. Zhang, "Clouds and aerosols," in *Climate Change 2013*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2013.

[2] C. Wang, S. Platnick, K. Meyer, Z. Zhang, and Y. Zhou, "A machine-learning-based cloud detection and thermodynamic-phase classification algorithm using passive spectral observations." *Atmospheric Measurement Techniques*, vol. 13, no. 5, pp. 2257–2277, 2020.

[3] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

[4] B. Gong, Y. Shi, F. Sha, and K. Grauman, "Geodesic flow kernel for unsupervised domain adaptation," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2012, pp. 2066–2073.

[5] B. Fernando, A. Habrard, M. Sebban, and T. Tuytelaars, "Unsupervised visual domain adaptation using subspace alignment," in *2013 IEEE international conference on computer vision*, 2013, pp. 2960–2967.

[6] J. Blitzer, M. Dredze, and F. Pereira, "Biographies, bollywood, boomboxes and blenders: Domain adaptation for sentiment classification," in *Proceedings of the 45th annual meeting of the association of computational linguistics*, 2007, pp. 440–447.

[7] G. Foster, C. Goutte, and R. Kuhn, "Discriminative instance weighting for domain adaptation in statistical machine translation," in *2010 conference on empirical methods in natural language processing*. Association for Computational Linguistics, 2010, pp. 451–459.

[8] E. Tzeng, J. Hoffman, T. Darrell, and K. Saenko, "Simultaneous deep transfer across domains and tasks," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 4068–4076.

[9] S. Purushotham, W. Carvalho, T. Nilanon, and Y. Liu, "Variational recurrent adversarial deep domain adaptation," in *International Conference on Learning Representations*, 2017.

[10] B. Sun and K. Saenko, "Deep coral: Correlation alignment for deep domain adaptation," in *European conference on computer vision*. Springer, 2016, pp. 443–450.

[11] S. Ackerman, S. Platnick, P. Bhartia, B. Duncan, T. L'Ecuyer, A. Heidinger, G. Skofronick-Jackson, N. Loeb, T. Schmit, and N. Smith, "Satellites see the world's atmosphere," *Meteorological Monographs*, vol. 59, pp. 4–1, 2018.

[12] S. Ackerman, R. Holz, R. Frey, E. Eloranta, B. Maddux, and M. McGill, "Cloud detection with modis. part ii: validation," *Journal of Atmospheric and Oceanic Technology*, vol. 25, no. 7, pp. 1073–1086, 2008.

[13] N. Hsu, M.-J. Jeong, C. Bettenhausen, A. Sayer, R. Hansell, C. Seftor, J. Huang, and S.-C. Tsay, "Enhanced deep blue aerosol retrieval algorithm: The second generation," *Journal of Geophysical Research: Atmospheres*, vol. 118, no. 16, pp. 9296–9315, 2013.

[14] J. V. Martins, D. Tanré, L. Remer, Y. Kaufman, S. Mattoo, and R. Levy, "Modis cloud screening for remote sensing of aerosols over oceans using spatial variability," *Geophysical Research Letters*, vol. 29, no. 12, pp. MOD4–1, 2002.

[15] W. Zhang, J. Wang, D. Jin, L. Oreopoulos, and Z. Zhang, "A deterministic self-organizing map approach and its application on satellite data based cloud type classification," in *2018 IEEE International Conference on Big Data (Big Data)*. IEEE, 2018, pp. 2027–2034.

[16] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[18] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, 2013, pp. 3111–3119.

[19] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 2, pp. 22–40, 2016.

[20] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, and T. Darrell, "Deep domain confusion: Maximizing for domain invariance," *arXiv preprint arXiv:1412.3474*, 2014.

[21] M. Long, Y. Cao, J. Wang, and M. I. Jordan, "Learning transferable features with deep adaptation networks," *arXiv preprint arXiv:1502.02791*, 2015.

[22] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 7167–7176.

[23] M. Wang and W. Deng, "Deep visual domain adaptation: A survey," *Neurocomputing*, vol. 312, pp. 135–153, 2018.

[24] S. Ackerman, R. Frey, A. Heidinger, Y. Li, A. Walther, S. Platnick, K. Meyer, G. Wind, N. Amarasinghe, C. Wang *et al.*, "Eos modis and snpp viirs cloud properties: User guide for climate data record continuity level-2 cloud top and optical properties product (cldprop), version 1," *NASA MODIS Adaptive Processing System, Goddard Space Flight Center, USA*, 2019.

[25] "Homepage — NOAA/NASA VIIRS/Suomi-NPP, JPSS," accessed: 2020-08-20. [Online]. Available: https://www.jpss.noaa.gov/viirs.html

[26] "Homepage — NASA CALIOP/CALIPSO Data," accessed: 2020-08-20. [Online]. Available: https://www-calipso.larc.nasa.gov/

[27] D. Winker, J. Tackett, B. Getzewich, Z. Liu, M. Vaughan, and R. Rogers, "The global 3-d distribution of tropospheric aerosols as characterized by caliop." *Atmospheric Chemistry & Physics*, vol. 13, no. 6, 2013.

[28] "Homepage — NASA EOSDIS Worldview Portal," accessed: 2020-08-20. [Online]. Available: https://worldview.earthdata.nasa.gov/

[29] X. Huang, S. Ali, S. Purushotham, J. Wang, C. Wang, and Z. Zhang, "Deep multi-sensor domain adaptation on active and passive satellite remote sensing data," in *1st KDD Workshop on Deep Learning for Spatiotemporal Data, Applications, and Systems (DeepSpatial 2020*. ACM, 2020.

[30] R. Holz, S. Ackerman, F. Nagle, R. Frey, S. Dutcher, R. Kuehn, M. Vaughan, and B. Baum, "Global moderate resolution imaging spectroradiometer (modis) cloud detection and height evaluation using caliop," *Journal of Geophysical Research: Atmospheres*, vol. 113, no. D8, 2008.

[31] Z.-H. Zhou, "A brief introduction to weakly supervised learning," *National Science Review*, vol. 5, no. 1, pp. 44–53, 2018.