Multi-source domain adaptation of social media data for disaster management

Anuradha Khattar [1] (ID) · S. M. K. Quadri [1]

Abstract

Labeled data scarcity at the time of an ongoing disaster has encouraged the researchers to use the labeled data from some previous disaster for training and transferring the knowledge to the current disaster task using Domain Adaptation (DA). However, often labeled data from more than one previous disaster may be available. As all deep learning models are data-hungry and perform better if fed with more annotated data, it is advisable to use data from multiple sources for training a Deep Convolutional Neural Network (DCNN). One of the easiest ways is to simply combine the data from multiple sources and use it for training. However, this arrangement is not that straightforward. The models trained on the combined data from various sources do not perform well on the target, mainly due to distribution discrepancies between multiple sources. This has motivated us to explore the challenging area of multi-source domain adaptation for disaster management. The aim is to learn the domain invariant features and representations across the domains and transfer more related knowledge to solve the target task with improved accuracy than single-source or combined-source domain adaptation. This study proposes a Multi-Source Domain Adaptation framework for Disaster Management (MSDA-DM) to classify disaster images posted on social media based on unsupervised DA with adversarial training. The empirical results obtained confirm that the proposed model MSDA-DM performs better than single-source DA by up to 10.83% and combined-source DA by up to 5.06% in terms of F1-score for different sets of source and target disaster domains. We also compare our model with current state-of-the-art models. The main challenge of multi-source DA is the choice of the relevant sources taken for training since, unlike single-source DA that handles only source-target distribution drift, the multi-source DA network has to address both source-target and source-source distribution drifts.

✉ Anuradha Khattar
anuradha.khattar@mirandahouse.ac.in

S. M. K. Quadri
quadrismk@jmi.ac.in

[1] Department of Computer Science, Faculty of Natural Sciences, Jamia Millia Islamia, New Delhi, India

1 Introduction

In the past decade, we have witnessed the increasing role of various social media platforms in communication dissemination during disasters [52]. Timely information through social media platforms enables the authorities to know the ground reality, thus, making the post-disaster relief and recovery operations more effective [22]. Moreover, the messages and pictures posted on these platforms help the authorities adapt to vulnerable situations and take necessary actions [15]. Government agencies and other humanitarian organizations have acknowledged the vast outreach of these platforms for use in managing disasters [41]. A study by Hiltz et al. [18] shows that during a disaster, the two most popular platforms used by the affected community are Facebook and Twitter. In a recent survey by Khattar et al. [28], social media platforms like WhatsApp and Instagram were excessively used by young Indian students during the countrywide shutdown due to the coronavirus pandemic. Twitter has stood the test as the most widely adopted microblogging platform during a conventional disaster. Twitter allows authorities to broadcast crucial information and warnings to a broader community and permits two-way communication between the affected community and the rescuing teams.

As soon as any disaster happens, data starts accumulating on social media platforms at a very high rate that can give valuable insights into the ongoing catastrophe if handled efficiently. Several studies have been done in recent years to explore the application of artificial intelligence techniques for managing disasters and related challenges [21, 27]. Vieweg et al. [54] were among the first team of researchers to work on social media data analysis for disaster management, after which many studies [18, 20, 22] have further explored this area of research. Natural Language Processing methods have been applied to analyze tweet text posted on Twitter to filter actionable and informative messages for matching the need-based tweets to the tweets that can fulfill the needs [1, 14, 38, 50]. On the other hand, there are many challenges to using social media data, including misinformation, rumors, fake messages, and other harmful content. Also, sometimes it is not easy to apply the standard machine learning, or natural language processing methods as the messages posted on social media sites like Twitter may be multilingual or multimodal in nature [15, 52].

Most of the studies mentioned above are based on traditional machine learning and deep learning methods that require a sufficient amount of labeled data of the related domain to solve the task. It has been observed that the messages on social media platforms start accumulating as soon as any disaster happens, but practically getting this data labeled is challenging due to time and cost constraints. This limitation has motivated the researchers to apply the concept of transfer learning [6, 37] to leverage models trained on some previous disaster (source) labeled data to solve the task for the ongoing disaster (target) with unlabeled data. However, many studies on transfer learning indicate that models trained on the source domain do not perform too well when tested on the target domain due to the discrepancy between the two domains. Under these circumstances, a particular type of transfer learning called Transductive transfer learning or Domain Adaptation (DA) [55] is used to handle the domain shifts between the source and the target domains. DA assumes that the source and the target domains have different marginal probability distributions but have the same feature space and the same task to be solved. These methods can be applied to align the source and target domain distributions

to perform the target task [11]. DA seems to be a desirable approach for disaster response as it allows relief agencies to take immediate action without any delay.

Several studies have been done based on the social media data focusing on single-source unsupervised domain adaptation, which requires annotated data from the source and unannotated data from the target disaster at the time of training [2, 25, 26, 31]. However, annotated data from more than one previous disaster may be available in many situations. As all deep learning models, including Convolution Neural Networks (CNN), are data-hungry and perform better if fed with more data, it is advisable to use data from multiple sources for training a CNN. One of the easiest ways is to combine the data of all the sources into a single source and use it for training. However, this arrangement does not take care of the diverse features extracted from the images of multiple datasets, which may be unrelated. This has motivated us to explore the area of multi-source DA for disaster response and study how to handle data from multiple sources and use this data for domain adaptation.

This study presents a multi-source DA method that not only reduces the domain shift between the source and the target domains but also handles the discrepancy that exists between multiple sources. The proposed algorithm for multi-source DA extends the single-source Domain Adversarial Neural Network (DANN) by Ganin et al. [12]. We apply adversarial training to the labeled data from multiple source disasters and unlabeled data from one target disaster embedded in a deep learning pipeline through CNN. For the present study, we would restrict the number of source domains to two for all the experiments.

The main contributions of this study are as follows:

1. A novel model, Multi-Source DA for Disaster Management (MSDA-DM), is proposed that takes input from multiple (2 in the present study) labeled source domains and an unlabeled target domain and aims to align the source and target domains to solve the target task.
2. We compare the proposed model MSDA-DM with two baseline models, (a) Single-Source DA and (b) Combined-Source DA.
3. The dataset used for this study consists of the images posted on Twitter during the disasters that happened in 2017. Three different categories of disasters are chosen - hurricanes (Hurricane Harvey, Maria, and Irma), earthquakes (Mexico and Iraq-Iran Earthquake), and wildfires (California). Each experiment involves three disasters, two source disasters, and one target disaster. A combination of source and target disasters is chosen very carefully to have complete insight into the behavior of MSDA-DM under the following four cases:

   • Case I: both the source disasters and the target disaster are of the same type
        Hurricane Harvey + **Hurricane** Irma→**Hurricane** Maria
   • Case II: both the source disasters are of the same type but different from the target disaster
        Hurricane Harvey + **Hurricane** Irma→**Earthquake** Iraq-Iran
   • Case III: one of the source disasters is the same as the target disaster
        Hurricane Harvey + **Earthquake** Mexico→**Earthquake** Iraq-Iran
   • Case IV: both the source disasters as well as the target disaster are of different types
        Hurricane Harvey + **Earthquake** Mexico→**Wildfires** California

4. MSDA-DM is also compared with the recent state-of-the-art (SOTA) models.

The empirical results obtained confirm the outperformance of the proposed MSDA-DM compared to SSDA and CSDA and the other SOTA models for different sets of source and target disasters. To the best of our knowledge, the concept of multi-source DA is used for the first time on disaster images posted on the microblogged platform Twitter.

The rest of the paper is organized as follows: Section 2 covers the related work done in DA, single-source DA in disaster management, and the recent multi-source DA methods. Section 3 covers the architecture of the proposed model to classify the disaster images. Section 4 gives the details of the experiments performed. Section 5 presents the tabulated results, analysis, visualization, and comparison with baselines and current SOTA methods. Finally, we conclude the paper in Section 6 with the limitations of the proposed model and future directions in this area.

2 Related work

We discuss the related work done in this area under three subsections. Subsection 2.1 gives the background and basic idea of domain adaptation and its recent applications. Subsection 2.2 provides a brief overview of the work done based on DA in disaster management. Most of the work mentioned in this section uses a single-source DA approach. Subsection 2.3 covers a few algorithms proposed for the novel concept of multi-source DA. We could not find any work where multi-source DA has been applied for disaster management.

2.1 Domain adaptation

As per a survey study [37], transfer learning is an approach to learn from the training data (source) and apply the knowledge to solve the target task. However, the performance of the model is deteriorated due to the discrepancy between the source and target. To handle this discrepancy, a particular type of transfer learning called Transductive transfer learning or Domain Adaptation (DA) is used where the training data comes from both the source and target domains. These two domains may be different but related to each other, and the domain tasks are the same. The DA techniques are called unsupervised DA if the target domain data is unlabeled. If the target domain has a small set of labeled and unlabeled data for training, it is called semi-supervised DA [55]. Recently DA methods have been embedded along with the training of deep neural networks to learn meaningful and domain invariant features. Deep domain adaptation methods boost the learning capacity of the network on source samples so that they can perform well on target samples. The DA techniques aim to learn a mapping that can handle the shift between the distributions of the source and target domains. We can divide various DA methods into three broad categories based on the technique used for domain alignment:

(a). **Discrepancy-based methods:** The discrepancy-based methods retrieve the domain-invariant features from the source and target domains by optimizing losses using the metrics such as the MMD metric [7], 'Wasserstein metric' [47], 'Correlation Alignment (CORAL)' [51], etc.

(b). **Adversarial methods:** In the adversarial category, the widely used methods are, DANN ('Domain Adversarial Neural Network') suggested by Ganin et al. [11, 12] and Generative Adversarial Nets, GAN [16], which use synthetic data for training.

(c). **Reconstruction-based methods:** The reconstruction-based models include the Encoder-Decoder framework [13] and adversarial reconstruction [53].

Domain Adaptation plays a vital role in situations where it is challenging to get annotated data, especially for an ongoing event. Areas like natural language processing [9, 29], computer vision [10, 24, 32, 34], machine translation [42, 43], etc., are increasingly using the concept of DA without the need to have labeled data.

2.2 Single-source domain adaptation

The most common scenario under DA is using an adaptation from one source domain to the target domain where the source domain data is labeled, and target domain data is unlabeled though available at the time of training. Thus, the two terms DA and Single-source DA are generally used interchangeably. In this subsection, we will discuss the recent work done in single-source DA in the context of disaster management.

Li et al. [31] were among the first few researchers who applied transfer learning and domain adaptation methods on tweets of Hurricane Sandy as a source and the Boston Marathon bombing as the target, with 1700 and 1000 tweets, respectively. They applied the DA Naive Bayes classification for their study. They showed that unlabeled target data and labeled source data improve the network's performance only in cases where the task is more related to the target. However, their model did not perform well on similar source and target domains tasks. In that case, supervised classifiers for source showed better performance. In the absence of the availability of the labeled data for a current event, Nguyen et al. [36] simply added out-of-event data to the current-event data. They used a convolutional neural network with pre-trained Google and Crisis embeddings for the binary classification of the text tweets. They reported that although neural network-based models perform better than non-neural models, the performance of the neural model drops when out-of-event data is added to current-event data of the ongoing crisis. Alam et al. [2, 3] performed two studies on the same Nepal and Queensland earthquakes dataset to classify text tweets based on semi-supervised learning. In the first study, they used a very popular method for DA proposed by Ganin et al. [11], the unsupervised adversarial training with backpropagation, and combined it with semi-supervised graph embeddings. Their results show significant improvement compared with the baseline methods without DA. The second study extended their model to an inductive semi-supervised model. Both semi-supervised settings performed better than unsupervised, but choosing an optimal number of labeled samples from the target domain is challenging. In a recent work [10], researchers fine-tuned BERT (i.e., Bidirectional Encoder Representations from Transformers) to classify the tweets of Hurricane Harvey into eight humanitarian categories to understand the disaster evolution and identify the location from the tweet posts by proposing a hybrid pipeline. They expressed that the proposed hybrid model should be further explored for different disasters, and other DA approaches should be embedded in the hybrid pipeline.

In a recent study, Khattar et al. [26] trained six DCNNs with images of Hurricane Harvey as the source and six other disasters as targets. They proposed an unsupervised DA model that applies the maximum mean discrepancy metric with radial basis function kernel to reduce the discrepancy between the source and target disaster domains. Li et al. [33] used the concept of DA on tweet images of four disasters for damage assessment. They classified the images as *damage* and *no_damage* based on the domain adversarial neural network for DA with VGG-19 as the backbone architecture and experimented with several pairs of disasters. In a similar study by Quadri et al. [25], a semi-supervised DA model is proposed for classifying the images of Hurricane Irma and California Wildfires with Hurricane Harvey as the source disaster. They suggested that the DA approach works better on disasters of different nature. Robertson et al.

[44] extracted the images of Hurricane Harvey through Twitter's Spritzer API and are the first ones to study the authenticity of the social media images by applying deep learning models.

Most of the work done in this area is based on text data, and very few studies have used images posted on microblogs for training. Additionally, most DA methods applied for disaster management have used only single-source DA. To the best of our knowledge, we could not find any study related to multi-source DA in disaster management. We have attempted to fill this research gap through our study.

2.3 Multi-source domain adaptation

The performance of a deep learning model improves if a sufficient amount of labeled data is available for training [45], which is not always possible due to time and cost constraints. Data may be collected from multiple sources and combined to form a larger dataset to overcome this limitation. However, this arrangement is not that straightforward. It is observed that the models trained on the combined data from multiple source domains do not generalize too well on the target domain, mainly due to the additional distribution discrepancies that exist between various sources. John et al. [5] were among the first researchers who investigated if the convex combination of error bounds of different source domains can be minimized to handle the multi-domain problem. Recently much research has been done in this area, and several new techniques have been proposed to manage multiple source domains for adaptation to a target domain. Based on the transformation of the latent space, we have grouped these techniques into two broad categories (a) Discrepancy Based Techniques and (b) Adversarial Based Techniques.

2.3.1 Discrepancy based techniques

The methods belonging to this category optimize the discrepancies among various domains by minimizing their distance. Karimpour et al. [23] proposed a novel shallow 'Multi-source Domain Adaptation technique (MDA)' for classifying images by extending the Maximum Mean Discrepancy (MMD) criteria to multiple source domains to construct the domain-invariant clusters. Their experiments included single, double, triple, and quadruple domain adaptation to transfer the knowledge common in all the domains by sample reweighting to improve the model performance. They concluded that average classification accuracy increases with an increase in the number of domains. However, their model lacks generalizability and can be improved by adding weights to the source domains. Zhu et al. [58] proposed a 'Multi-Feature Spaces Adaptation Network (MFSAN)' to handle input from N sources and put a separate feature extractor for each domain with N different classifiers. The loss term of the network included three terms: MMD loss, discriminator loss, and classification loss. Zhang et al. [56] gave a new algorithm of the weighted combination of domains for the cross-entropy loss by reducing Rényi divergence. However, their algorithm performed well only in the situation when one of the training sources was the same as the test target. Another model proposed by Peng et al. [40] uses moment alignment of the feature distributions for handling domain discrepancy in a multi-domain scenario. However, their model suffers from negative transfer.

2.3.2 Adversarial based techniques

This technique learns the discriminative features of the source samples and, at the same, time tries to fool the discriminator about the origin of the sample, i.e., whether it is coming from the

source or target, by making the features of the samples indistinguishable. Zhou et al. [57] proposed a model 'Duplex Adversarial Network for Multi-source DA (DAN_ MA)' based on two adversarial networks: one for domain classification with a separate classifier for every source-target pair and a second network to bring the target domain closer to each of the source domain decision boundary. They suggested that the model can be improved by using weights for different classifiers. Sebag et al. [46] introduced MuLANN, an extension of Domain Adversarial Neural Network, DANN [12] to multi-domain learning for microscopic data. They expressed H-distance [5] as the convex combination of the distance between different pairs of domains and minimized the average H-distance to improve multi-domain learning. 'Deep Cocktail Network' (DCTN) is introduced by Chen et al. [8], in which the authors have handled category shift and the domain shift by the multi-way adversarial network.

The above models have experimented with one or more benchmark datasets for unsupervised domain adaptation Office-31, digits dataset (MNIST, MNIST-M, SVHN, Synthetic digits), and computer vision (WebcamT Vehicle counting) from the real world. Their experiments showed the superiority of adaptation for multi-source as compared to single-source. However, to the best of our knowledge, no work has been done using multi-source DA to classify disaster images posted on the microblogging platforms. Therefore, the present study extends the concept of single-source DA to multi-source DA and is based on unsupervised, homogenous DA with adversarial training.

3 Proposed work

We propose a Multi-Source Domain Adaptation framework for Disaster Management (MSDA-DM) embedded in the deep learning pipeline for multiple labeled source domains and an unlabeled target domain. The MSDA-DM architecture is an adversarial training model that extends Domain Adversarial Neural Network (DANN) [12] from single-source DA to multi-source DA. The block diagram of the model MSDA-DM is shown in Fig. 1.
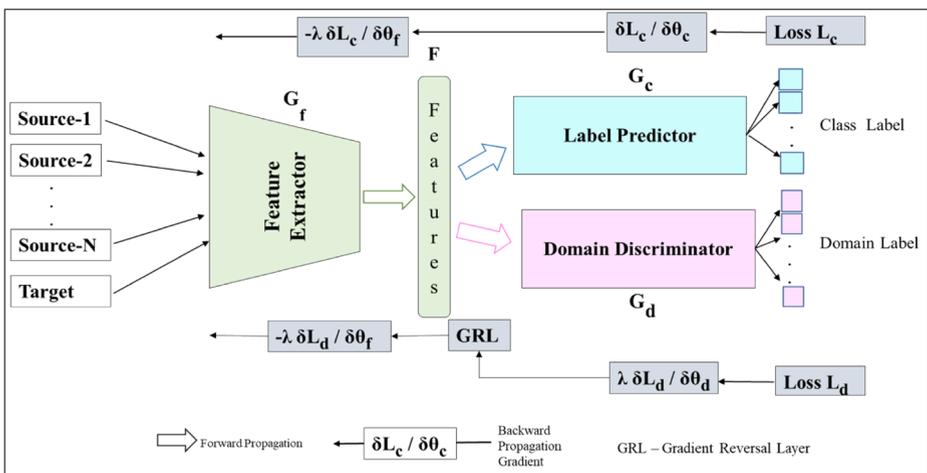

Fig. 1 Block diagram of the proposed model Multi-Source DA for Disaster Management (MSDA-DM) that takes input from multiple labeled sources and an unlabeled target

3.1 Adversarial learning for domain adaptation

The mechanism of adversarial training involves the feature extractor and domain discriminator competing with one another. The source and target features become indiscriminative by reducing the shifts between them and simultaneously making the source feature representations discriminative for the classification task of the source. The domain discriminator attempts to predict the correct domain, but the feature extractor is trying to confuse the domain discriminator about the origin of the data. When all three pieces of training are combined in one network, the network learns the domain invariant features of source and target domains while ensuring that the semantic context is preserved to classify the data. The success of this training is attributed to the addition of a gradient reversal layer (GRL). It is placed between the domain discriminator and the feature extractor and has no parameters. During backpropagation, GRL negates the gradient by a regularization parameter λ which balances the two adversarial losses. For the present study, the value of λ is taken as 1 (see Section 5.3.3).

3.2 MSDA-DM

The proposed framework has three main components: a joint multi-source Feature Extractor, a multi-source Domain Discriminator, and a source-specific Label Predictor:

- The feature extractor is a DCNN that aligns the distributions of multiple source domains and a target domain and aims to learn common features among them. This way, the feature extractor minimizes the marginal probability distribution drift between various domains. In addition, the feature extractor is common for the other two components of the network - the domain discriminator and the label predictor.
- The label predictor predicts the label of the input image coming from multiple sources.
- The domain discriminator predicts the origin of the input images among multiple source domains and a target domain based on the shared domain invariant features provided by the feature extractor.

Suppose there are K labeled source domains where each kth source domain is denoted by D_{s_k} and an unlabeled target domain represented by D_t. Let G_f be the Feature Extractor, G_d be the Domain Discriminator to distinguish source and target domains, and G_c is the label predictor to predict the labels of the source domain data. Let f denotes the domain invariant features. θ_f, θ_d, and θ_c are the parameters of the feature extractor, the domain discriminator, and the label predictor. L_d and L_c are the domain discriminator loss and label predictor loss. Let n_{s_k} are the number of samples in the kth source domain D_{s_k}, n_t the number of samples in the target domain, y_i is the class label of sample x_i, d_i is the domain label for the sample x_i, and $n = n_s + n_t$.

The network aims to learn features that:

1. minimize the class label predictor loss L_c for all the sources
2. maximize the domain discriminator loss L_d for all the source domains and the target domain

The following equations represent the two adversarial losses label predictor loss and domain discriminator loss:

$$\mathbf{L_c}\left(\mathbf{G_c}\left(\mathbf{G_f}\left((\mathbf{x_i}), \mathbf{y_i}\right)\right)\right) \tag{1}$$

$$\mathbf{L_d}\left(\mathbf{G_d}\left(\mathbf{G_f}\left((\mathbf{x_i}), \mathbf{d_i}\right)\right)\right) \tag{2}$$

Summing Eqs. (1) and (2) for K number of source domains D_{s_k} (k = 1, K) and a target domain D_t.

The complete loss function for MSDA-DM is given by Eq. (3):

$$\mathbf{L}\left(\theta_f, \theta_d, \theta_c\right) = \sum_{k=1}^{K} \frac{1}{n_{s_k}} \Sigma_{x_{i \in D_{s_k}}} \mathbf{L_c}\left(\mathbf{G_c}\left(\mathbf{G_f}\left((\mathbf{x_i}), \mathbf{y_i}\right)\right)\right) - \lambda \frac{1}{n_s} \sum_{k=1}^{K} \Sigma_{x_{i \in D_{s_k} \cup D_t}} \mathbf{L_d}\left(\mathbf{G_d}\left(\mathbf{G_f}\left((\mathbf{x_i}), \mathbf{d_i}\right)\right)\right) \tag{3}$$

The two terms of Eq. (3) handle the domain shifts among multiple sources and the target and also learn the discriminative features of the target domain to solve the target task by the method of adversarial learning. This way, the proposed MSDA-DM considers the domain distribution discrepancy between the target and each of the source domains as well as between multiple source domains. Thus, it allows us to handle diverse data from multiple labeled sources and use this data for domain adaptation to improve the accuracy of the classifier. Furthermore, unlike single-source DA, which handles only source-target distribution drift, the proposed multi-source DA framework handles both source-target and source-source distribution drifts giving promising results compared with baseline methods.

4 Experimental setup

This section gives the details of performing the experiments for the proposed model. Subsection 4.1 provides the particulars of the dataset used for the experiment. The subsequent subsections cover the baseline methods, training parameters, performance metrics, and the methodology used for performing the experiments.

4.1 Datasets

We conduct experiments on the Twitter images of a publicly available multimodal dataset CrisisMMD, released by researchers Alam et al. [4] and is available on AIDR ("Artificial Intelligence for Disaster Response") platform. The dataset consists of both tweet texts and images, annotated manually by a crowdsourcing platform Figure Eight for seven disasters that happened in 2017. The two major classes for this dataset are '*informative*' and '*non-informative*'.

The image is labeled as '*informative*' if it can be used by the humanitarian agencies for assisting aid to the affected community, as shown in Fig. 2; otherwise, the image is labeled as '*non-informative*'. The images with banners, cartoons or public figures generally come under the *non-informative* category, as shown in Fig. 3. To explore the multi-source adaptation approach, the experiments are performed on different types of source and target disasters that include hurricanes (Harvey, Irma, and Maria), earthquakes (Iraq-Iran and Mexico), and wildfires (California). For conducting the experiments, we divide each of the datasets into

Fig. 2 Set of *informative* images which help humanitarian agencies to assist in disaster response

Fig. 3 Set of *non-informative* images that include banners, posters, cartoons, etc

three folders train (80%), validation (10%), and test (10%). The details of the dataset are listed in Table 1.

4.2 Baselines

The proposed model MSDA-DM is compared with two baseline models, SSDA and CSDA. In addition, we have also chosen three recent SOTA models, which have used the CrisisMMD dataset (Table 1) as used in the proposed MSDA-DM model. The selected baselines are as follows:

1. Single-Source Domain Adaptation (SSDA): We have developed an unsupervised DA network based on the popular 'Domain Adversarial Neural Network (DANN)'. The network is trained on the images from a source disaster (labeled) and a target disaster (unlabeled). SSDA classifies the images of the target disaster as *informative* and *non-informative*.
2. Combined-Source Domain Adaptation (CSDA): This is also an unsupervised domain adaptation network based on DANN. The datasets of two of the disasters are simply combined to make one more extensive dataset. The network is trained with labeled images of the combined dataset and unlabeled images of a target dataset. CSDA classifies the images of the target disaster as *informative* and *non-informative*.

Table 1 Details of CrisisMMD dataset: Type and name of six disasters, duration of data collection, the total number of images, and the number of images with labels '*informative*' and '*non-informative*'

Type of Disaster	Name of Disaster	Duration of Data Collection	No. of *informative* Images	No. of *non-informative* Images	Total Images
Hurricane	Harvey	26 Aug 2017 to 21 Sep 2017	2449	1947	4396
Hurricane	Irma	6 Sep 2017 to 21 Sep 2017	2203	2262	4465
Hurricane	Maria	20 Sep 2017 to 13 Nov 2017	2229	2289	4518
Earthquake	Mexico	20 Sep 2017 to 6 Oct 2017	839	528	1367
Wildfires	California	10 Oct 2017 to 27 Oct 2017	979	588	1567
Earthquake	Iraq-Iran	13 Nov 2017 to 19 Nov 2017	399	193	592

3. The study [10] applied several methods: CNN, RCNN, CNN GRU, BiGRU, BiLSTM, Kmax, DPCNN, and proposed a fine-tuned BERT for the classification of text tweets posted during three Hurricanes.
4. Recently, another study [49] has used text and images of all seven disasters of the CrisisMMD dataset and proposed a model based on late fusion.
5. Khattar et al. [25] suggested a semi-supervised DA method to classify the images of three disasters, Hurricane Harvey, Hurricane Irma, and California wildfires.

4.3 Training details

All the networks are trained under the open-source deep learning framework PyTorch [39], and Python is the primary language used for coding. The number of epochs for all the networks is 50, the image size is 224 × 224, and the batch size is 8. The optimizer used is mini-batch Stochastic Gradient Descent (SGD) [30] with learning rate (LR): 1e-03, momentum: 0.9, and weight decay: 5e-04. To avoid over-fitting scheduler is used to tune the learning rate at the milestones of 30 and 40 epochs by 0.1. The loss function for classification, and domain loss is Cross Entropy. The parameter λ that balances two loss terms, classification loss and domain loss, in the adversarial loss function is set to 1. The experiments are run on the Google cloud platform Colab Pro which provides Tesla P100 GPU and 27.3 GB high-RAM. More details of hyperparameter settings are listed in Table 2. Grid search is used to choose the hyperparameter values. We have performed a detailed ablation study for our choice of feature extractor and hyperparameters in Section 5.3. All the experiments are performed three times, and the average of the three is listed as the results in the study.

4.4 Performance metrics

The performance metrics used for the experiments are Accuracy, Precision, Recall, and F1-Score. The symbols used in the formula are TP: True Positive, TN: True Negative, FP: False Positive, and FN: False Negative.

$$\textbf{Accuracy} = (TP + TN)/(TP + TN + FP + FN)$$

$$\textbf{Precision} = TP/(TP + FP)$$

$$\textbf{Recall} = TP/(TP + FN)$$

$$\textbf{F1-Score} = 2*(Precision*Recall)/(Precision + Recall) \text{ (Harmonic Mean of precision and recall)}$$

Table 2 The setting of Hyperparameters for the experiments performed

Hyper-parameter	Value
Image Size	224×224
Batch Size	8
No. of Epochs	50
Optimizer	Stochastic Gradient Descent (SGD)
Loss Function	Cross-Entropy
Learning Rate (LR)	1e-03
Momentum	0.9
Weight Decay	5e-04
Scheduler	Multistep Decrease at milestones 30 and 40 at the rate of 0.1
Regularization Parameter (λ)	1

4.5 Methodology

We have performed four sets of experiments to classify images posted on Twitter for six disasters – including Hurricanes (Harvey, Irma, and Maria), earthquakes (Mexico and Iraq-Iran), and Wildfires in California. Each set of experiments has a unique combination of source and target disasters. Under each set of experiments, we have built three deep neural networks: SSDA, CSDA, and the proposed MSDA-DM, where SSDA and CSDA are the baseline models to evaluate the performance of the proposed model MSDA-DM. Thus, in all, twelve DNNs are trained in this study. We have used the Hurricane Harvey dataset as one of the source disasters for all the sets since it was the first one to happen and had a sufficiently large dataset for training.

The pre-trained model VGG-19 [48] is the backbone architecture to initialize the network parameters and fine-tune them. VGG-19 is a popular, frequently used model trained on a vast dataset, ImageNet, which has millions of images and classifies these images into 1000 different categories. VGG-19 model parameters are highly transferable and can generalize well on other images. This makes the model very useful for extracting features from images of other classification problems. As shown in Fig. 4, VGG-19 consists of 19 layers, including 16 convolution layers (Block 1 to Block 5) and three fully connected layers (Block 6). The hyperparameters of each convolutional layer are depicted in Fig. 4 as conv1(in-channel, out-channel, filter-size, stride, padding). Each convolution layer is followed by a batch normalization layer and a nonlinear activation function ReLU. Finally, a Maxpool layer follows each block for downsampling with filter-size:2, stride:2, and padding:0.

Following are the implementation details of the three architectures trained for this study. We have kept the number of sources limited to two for these experiments, but it can easily be extended to more than two. The combination of source and target disasters is chosen very carefully to have complete insight into the behavior of MSDA-DM.

Some short forms are used in the following explanation: **SD** for Source Disaster, **TD** for Target Disaster, **SD1** for Source Disaster-1, **SD2** for Source Disaster-2.

1. **Single-Source Domain Adaptation (SSDA)**: As discussed in Section 4.2, SSDA is an unsupervised domain adaptation network based on the popular Domain Adversarial Neural Network (DANN), which takes input from one source and a target. For all the
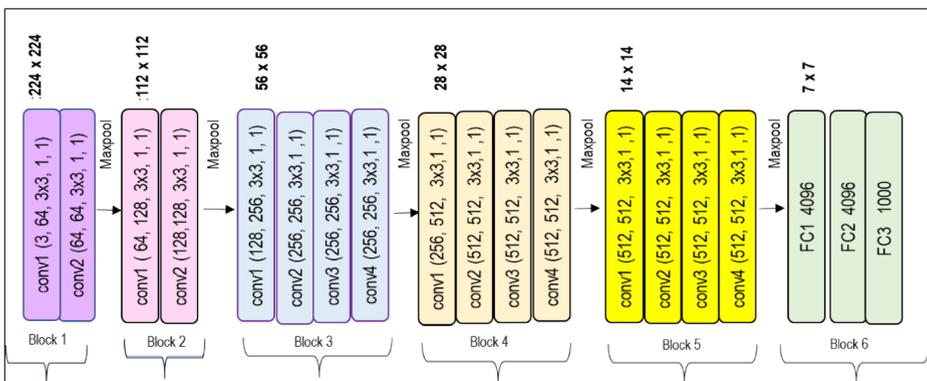

Fig. 4 VGG-19 network [48] has 16 Conv. and 3 FC layers. The hyperparameters of each Conv layer are depicted as (in-channel, out-channel, filter-size, stride, padding). There are 5 Maxpool layers with filter-size:2, stride:2, and padding:0

sets of experiments, Hurricane Harvey is the source disaster, and Hurricane Maria, Iraq-Iran Earthquake, and California Wildfires are the target disasters. The SD data is labeled while the TD data is unlabeled.

Case I: Hurricane Harvey (SD) ⟶ Hurricane Maria (TD)
Case II: Hurricane Harvey (SD) ⟶ Earthquake Iraq-Iran (TD) ⎤ Case II and Case III are
Case III: Hurricane Harvey (SD) ⟶ Earthquake Iraq-Iran (TD) ⎦ same for SSDA
Case IV: Hurricane Harvey (SD) ⟶ Wildfires California (TD)

2. **Combined-Source Domain Adaptation (CSDA):** The datasets of two source disasters are merged to form one larger dataset used for input. As discussed in Section 4.2, CSDA is an unsupervised domain adaptation network based on the popular Domain Adversarial Neural Network (DANN), which takes input from two merged sources and a target. The SDs data is labeled while the TD data is unlabeled. Networks are trained for the following four cases:

Case I: (Hurricane Harvey + Hurricane Irma) $_{combined}$ (SDs)→Hurricane Maria (TD)
Case II: (Hurricane Harvey + Hurricane Irma) $_{combined}$ (SDs)→Earthquake Iraq-Iran (TD)
Case III: (Hurricane Harvey + Earthquake Mexico) $_{combined}$ (SDs)→Earthquake Iraq-Iran (TD)
Case IV: (Hurricane Harvey + Earthquake Mexico) $_{combined}$ (SDs)→Wildfires California (TD)

3. **Multi-Source Domain Adaptation for Disaster Management (MSDA-DM):** The general architecture of the proposed model MSDA-DM with N-labeled source domains and an unlabeled target domain is discussed in Section 3.2. In the present
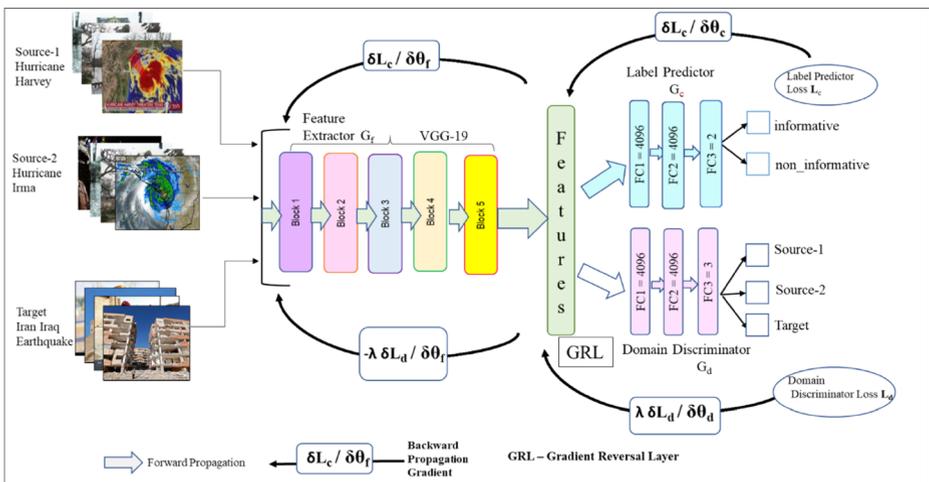
Fig. 5 Detailed architecture of the implemented model MSDA-DM with input images from two labeled source disasters and an unlabeled target disaster. Pe-trained model VGG-19 is used as Feature Extractor (G_f) depicted by five blocks. The details of each block are given in Fig. 4. G_f is common for the other two network components, Label Predictor (G_c) and Domain Discriminator (G_d). The dimension of the last dense layer FC3 for G_c is set to 2 since there are two class labels, *informative* and *non-informative*. For G_d, it is set to 3 for three domains source-1, source-2, and target. Gradient Reversal Layer (GRL) is placed between G_f and G_d, which multiplies the gradient by -λ during backpropagation, (λ = 1) for the proposed model

study, we have implemented MSDA-DM with two sources only. The pseudocode of MSDA-DM is given in 'Pseudocode-1', and a more detailed version of the proposed architecture of MSDA-DM is shown in Fig. 5 with two source disasters and a target disaster. The datasets used for source and target disasters are chosen to understand the additional impact of the domain shift introduced between the two sources in addition to the shift between the source and the target. Depending on the source/ target combination, the whole experiment is conducted under four cases. Case I: both the source disasters and the target disaster are of the same type, Case II: both the source disasters are of the same type but different from the target, Case III: one of the source disasters is the same as the target disaster, Case IV: both the source disasters, as well as the target, all three are of different types. All the source disaster datasets are labeled, while the target disaster datasets are unlabeled.

Case I: Type of SD1 = Type of SD2 = Type of TD
Hurricane Harvey + **Hurricane** Irma→**Hurricane** Maria
Case II: (Type of SD1 = Type of SD2) ≠ Type of TD
Hurricane Harvey + **Hurricane** Irma→**Earthquake** Iraq-Iran
Case III: Type of SD1 ≠ (Type of SD2 = Type of TD)
Hurricane Harvey + **Earthquake** Mexico→**Earthquake** Iraq-Iran
Case IV: Type of SD1 ≠ Type of SD2 ≠ Type of TD
Hurricane Harvey + **Earthquake** Mexico→**Wildfires** California

---

Pseudocode-1: Multi-Source Domain Adaptation for Disaster Management (MSDA-DM)

```
Input: images of SD1, SD2, and TD; labels of SD1, SD2
Output: Best_Model to binary classify labels of TD into: "informative" and "non-informative"
Preprocessing:
    resize the input images to 224x224
    convert to Tensor
    RandomFlip horizontally
    normalize
    shuffle
    divide into train_set, validation_set, test_set
Model Training:
1. My_Model = VGG-19                    #Load the pre-trained model
2. Set the dimension of the last fully connected layer(FC3) of My_Model to 2 for Label_Predictor
   and 3 for Domain_Discriminator
3. Set the hyperparameters:
       total_epochs = 50
       optimizer = SGD
       one_batch = 8
       weight_decay = 5e-04
       momentum = 0.9
       LR = 1e-03
       regularization_parameter = 1
       loss_fun = CrossEnropyLoss
4. fine-tune My_Model to get the Best_Model
   Best_Model = My_Model
   for epochs in range (0, total_epochs):
       for images in range (0, len(train_set), one_batch):
           load images od SD1,SD2,TD and labels of SD1,SD2
           compute label_predictor_loss for SD1 and for SD2          #eq (1)
           compute domain_discrepancy_loss for SD1, SD2 and TD       #eq (2)
           total_loss = label_predictor_loss + domain_dicrepency_loss  #eq (3)
           backpropagate and update the parameters of My_Model
       validate My_Model on validation_set
       compare Best_Model and My_Model
Model_Testing:
 5. Best_Model is tested on the test_set of TD
```

4.6 Complexity of MSDA-DM

Let L be the number of layers in the neural network.

Computation at each neuron is represented as: $y_i = \varphi \ (\Sigma \ w_i x_i)$ where

   i = {1, 2 …, n}, n is the number of neurons in each layer
   w is the weight matrix of size (n x n)
   x is the input vector of size (n × 1)
   y_i is the output
   φ is the non-linear activation function (ReLU in this case)

Complexity of matrix-vector multiplication ($w_i x_i$) is $O(n^2)$.

Complexity of elementwise activation φ is $O(n)$.

Complexity of $\varphi \ (\Sigma \ w_i x_i) = O \ (n + n^2) = O(n^2)$.

Complexity of the network during forward propagation with L layers $= O(L*n^2)$.

During Backward Propagation (BP), if we assume that Gradient Descent takes, I number of iterations to converge.

Complexity of the network during BP is $= O(I*L*n^2)$.

For simplicity let us assume that the network has n number of layers with each layer having n number of neurons and number of iterations for BP is also n then the complexity of BP is given by $O(I*L*n^2) = O(n*n*n^2) = \mathbf{O(n^4).}$

5 Results and discussion

In Section 5.1 we report the results obtained by training neural networks under four cases for the proposed model and two baseline models. We discuss the performance of the proposed algorithm and the baseline models and compare them with SOTA models in Section 5.2. Section 5.3 covers the ablation analysis.

5.1 Quantitative results

All the experiments were performed three times, and the average of the three is reported as results in Tables 3, 4, 5, and 6. Figure 6 shows the F1-score and accuracy of the three networks, SSDA, CSDA, and MSDA-DM, under four cases in a bar graph.

5.2 Performance analysis

All the comparisons among various models are made based on the F1-score. Based on the quantitative results obtained under four sets of experiments, we conclude that the performance of the proposed network improves in the range of 2.54% to 10.83% in comparison to SSDA and 0.87% to 5.06% in comparison to CSDA. A detailed comparison of MSDA-DM with the baseline models SSDA and CSDA and other state-of-the-art methods is discussed below. Table 7 lists the percentage increase in the performance of MSDA-DM as compared to the two baselines. A bar graph has depicted the same in Fig. 7.

Table 3 **Case I: Type of SD1 = Type of SD2 = Type of TD** (Accuracy, Precision, Recall, and F1-score (all in %) of two baseline models, SSDA and CSDA, and the proposed model MSDA-DM for the binary classification of images)

SSDA (Single-Source Domain Adaptation)

Source		Target	
Hurricane Harvey		Hurricane Maria	
Acc.	Precision	Recall	F1-score
76.43	76.61	76.47	76.41

CSDA (Combined-Source Domain Adaptation)

Source-1	Source-2	Target	
Hurricane Harvey	Hurricane Irma	Hurricane Maria	
Acc.	Precision	Recall	F1-score
77.31	77.36	77.29	77.29

MSDA-DM (Multi-Source Domain Adaptation for Disaster Management)

Source-1	Source-2	Target	
Hurricane Harvey	Hurricane Irma	Hurricane Maria	
Acc.	Precision	Recall	F1-score
79.08	**79.66**	**79.04**	**78.95**

Table 4 **Case II: (Type of SD1 = Type of SD2) ≠ Type of TD** (Accuracy, Precision, Recall, and F1-score (all in %) of two baseline models, SSDA and CSDA, and the proposed model MSDA-DM for the binary classification of images)

SSDA (Single-Source Domain Adaptation)				CSDA (Combined-Source Domain Adaptation)						MSDA-DM (Multi-Source Domain Adaptation for Disaster Management)					
Source	Target			Source-1	Source-2	Target				Source-1	Source-2	Target			
Hurricane Harvey	Earthquake Iraq-Iran			Hurricane Harvey	Earthquake Irma	Earthquake Iraq-Iran				Hurricane Harvey	Earthquake Irma	Earthquake Iraq-Iran			
Acc.	Precision	Recall	F1-score	Acc.	Precision	Recall	F1-score			Acc.	Precision	Recall	F1-score		
72.95	71.12	73.48	71.22	73.77	74.16	77.29	73.03			**79.51**	**77.43**	**80.27**	**78.09**		

Table 5 Case III: Type of SD1 ≠ (Type of SD2 = Type of TD) (Accuracy, Precision, Recall, and F1-score (all in %) of two baseline models, SSDA and CSDA, and the proposed model MSDA-DM for the binary classification of images)

SSDA (Single-Source Domain Adaptation)				CSDA (Combined-Source Domain Adaptation)				MSDA-DM (Multi-Source Domain Adaptation for Disaster Management)		
Source			Target	Source-1	Source-2		Target	Source-1	Source-2	Target
Hurricane Harvey			Earthquake Iraq-Iran	Hurricane Harvey	Earthquake Mexico		Earthquake Iraq-Iran	Hurricane Harvey	Earthquake Mexico	Earthquake Iraq-Iran
Acc.	Precision	Recall	F1-score	Acc.	Precision	Recall	F1-score	Acc.	Precision Recall	F1-score
72.95	71.12	73.48	71.22	82.79	80.61	82.71	81.18	**83.61**	**81.47** **83.32**	**82.05**

Table 6 Case IV: **Type of SD1 ≠ Type of SD2 ≠ Type of TD** (Accuracy, Precision, Recall, and F1-score (all in %) of two baseline models, SSDA and CSDA, and the proposed model MSDA-DM for the binary classification of images)

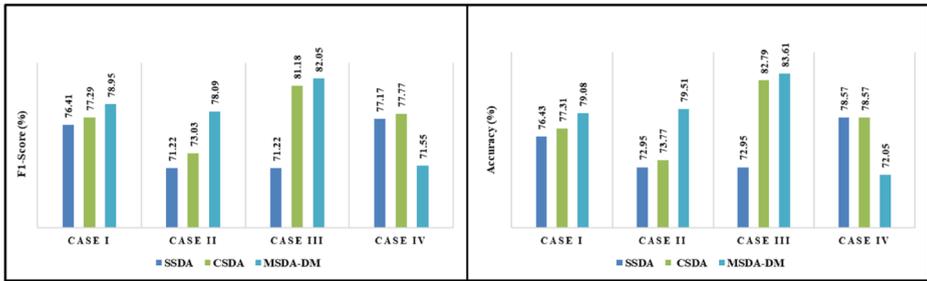| SSDA (Single-Source Domain Adaptation)					CSDA (Combined-Source Domain Adaptation)						MSDA-DM (Multi-Source Domain Adaptation for Disaster Management)					
Source		Target				Source-1	Source-2	Target				Source-1	Source-2	Target		
Hurricane Harvey		Wildfires California				Hurricane Harvey	Earthquake Mexico	Wildfires California				Hurricane Harvey	Earthquake Mexico	Wildfires California		
Acc.	Precision	Recall	F1-score			Acc.	Precision	Recall	F1-score			Acc.	Precision	Recall	F1-score	
78.57	77.70	77.00	77.17			**78.57**	**77.48**	**78.36**	77.77			72.05	70.99	71.85	71.55	

Fig. 6 F1-score and Accuracy of three models: Single-Source Domain Adaptation (SSDA), Combined-Source Domain Adaptation (CSDA), and Multi-Source Domain Adaptation for Disaster management (MSDA-DM) under four cases. Case I: both the sources and the target disaster are the same, Case II: both the source disasters are the same but different from the target, Case III: one of the source disasters is the same as the target disaster, Case IV: both the source disasters as well as the target are different. MSDA-DM outperforms in all the cases except in case IV

5.2.1 Comparison of MSDA-DM with SSDA and CSDA

Case I: Type of SD1 = Type of SD2 = Type of TD

Hurricane Harvey + Hurricane Irma→Hurricane Maria

Since both the source disasters and the target disaster are of the same type, i.e., hurricanes, they share more common features and have less discrepancy in their distributions. The performance of MSDA-DM for this case improves by 2.54% compared to SSDA and by 1.66% compared to CSDA.

Case II: (Type of SD1 = Type of SD2) ≠ Type of TD

Hurricane Harvey + Hurricane Irma→Earthquake Iraq-Iran

The source domains have data from disasters of the same type, i.e., hurricane but adapted to a different disaster type, i.e., earthquake. The F1-Score for the MSDA-DM improves by 6.87% compared to SSDA and 5.08% compared to CSDA. MSDA-DM performs better in Case II than Case I, as the DA approach is more useful when the source and target disasters are not the same and have more discrepancies before adaptation [32, 55]. In case I, the distributions of

Table 7 % increase in the F1-Score and Accuracy from single-source (SSDA) to Multi-Source (MSDA-DM) and from CSDA to MSDA-DM under four cases. *Case I* both the sources and the target disaster are the same, *Case II* both the source disasters are the same but different from the target, *Case III* one of the source disasters is the same as the target disaster, *Case IV* both the source disasters as well as the target are different, *H* Hurricane, *E* Earthquake, *W* Wildfires

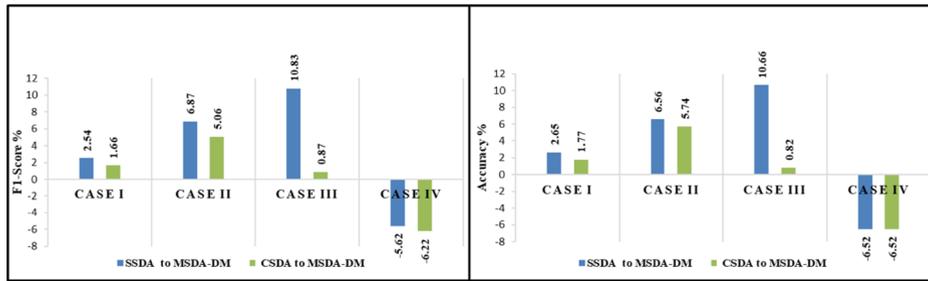| % increase in F1-Score from SSDA & CSDA to MSDA-DM			% increase in Accuracy from SSDA & CSDA to MSDA-DM		
SD1 + SD2→TD	SSDA to MSDA-DM	CSDA to MSDA-DM	SD1 + SD2→TD	SSDA to MSDA-DM	CSDA to MSDA-DM
H + H→H Case I	2.54	1.66	H + H→H Case I	2.65	1.77
H + H→E Case II	6.87	5.06	H + H→E Case II	6.56	5.74
H + H→E Case III	10.83	0.87	H + H→E Case III	10.66	0.82
H + H→W Case IV	-5.62	-6.22	H + H→W Case IV	-6.52	-6.52

Fig. 7 % increase in the F1-Score and Accuracy from single-source (SSDA) to Multi-Source (MSDA-DM) and from CSDA to MSDA-DM under four cases. Case I: both the sources and the target disaster are the same, Case II: both the source disasters are the same but different from the target, Case III: one of the source disasters is the same as the target disaster, Case IV: both the source disasters as well as the target are different

SD1 and the TD disaster are already similar before adaptation. Therefore, adding SD2, which is again of the same type, results in minor improvement compared to Case II, where there is more scope for adaptation.

Case III: Type of SD1 ≠ (Type of SD2 = Type of TD)

Hurricane Harvey + Earthquake Mexico→Earthquake Iraq-Iran

In this case, the type of SD1 is a hurricane, and that of TD is an earthquake, i.e., source and target disasters have different distributions before adaptation. Therefore, we expect performance improvement when the DA method is applied even with a single source, SD1. Furthermore, adding SD2, which is different from SD1 but the same as TD, has given a tremendous boost to the performance of the model by 10.83% as compared to SSDA. This confirms that the proposed model is not only bringing the distributions of SD1 and TD closer but also taking care of the distributions of the two source disasters, SD1 and SD2. Hence the proposed model MSDA-DM can handle both source-target and source-source discrepancy, which is the limitation of other MSDA models.

Case IV: Type of SD1 ≠ Type of SD2 ≠ Type of TD

Hurricane Harvey + Earthquake Mexico→Wildfires California

Both the source disasters are different. They are also different from the target disaster leading to the situation where it is more challenging for the model to handle the vast difference in the feature spaces of three entirely different distributions. This combination of all three different domains decreases the performance of MSDA-DM compared to SSDA by 5.62% and by 6.22% compared to CSDA. Thus, the network's performance deteriorates when the categories of both the source disasters and the target disaster differ. We conclude that the performance of the proposed model MSDA-DM depends on the choice of the source and target disasters.

5.2.2 CSDA vs. SSDA

The Accuracy and F1-score of all combined-source domain adaptation, CSDA networks are better than single-source domain adaptation, SSDA networks. This demonstrates that simply combining images of two source disaster datasets into one larger dataset improves the F1-score in the range from 0.6% to 9.96% due to an increase in the size of the source dataset for training a DCNN.

The above analysis of the experimental results demonstrates that DA from more than one source improves the network's performance and should be preferred whenever data is available from multiple sources. At the same time, just randomly adding sources into the network pipeline may not always return an improved model. The performance of the model also depends on the nature of the source domains and the target domain, i.e. their similarity in their marginal probability distributions. Thus, the type of disasters chosen for training and testing as sources and target do affect the performance of the network.

5.2.3 Performance of MSDA-DM compared to other state-of-the-art studies done in recent times in the field of disaster management

Table 8 compares the proposed model MSDA-DM with the recent studies performed on various text and image disaster datasets. The parameters used for comparison are Accuracy (Acc.), Precision (P), Recall (R), and F1-score, all in %. It is clear from the results that the proposed model MSDA-DM outperforms the state-of-the-art models presented in the recent past. In this section, we would like to do a detailed discussion of the performance of our model compared to three of the latest studies : [10, 25, 49]. The reason for choosing these models is that these studies have used the same dataset as ours, i.e., the CrisisMMD dataset, making the comparison more logical.

- Wu et al. [10] built several text-based models, BiGRU, BiLSTM, CNN GRU, CNN LSTM, DPCNN, KMax CNN, RCNN, for the classification of tweets of three hurricanes (Harvey, Irma, Maria) for situational awareness with test accuracy ranging from 58.09% to 71.77%. Their proposed model, a fine-tuned BERT, showed a test accuracy of 75.37%. The MSDA-DM model of Case I of our study used images of hurricanes Harvey and Irma as source disasters and Maria as the target disaster, resulting in an accuracy of 79.08%, confirming that DA with multiple source disasters outperforms the models trained on the single disaster.
- The study by Sreenivasulu et al. [49] proposed a multimodal classification model based on late fusion using text and image tweets of seven disasters. Their results showed an F1-score of 72.84% for Hurricane Maria, 64% for California wildfires, and 67% for the Iraq-Iran earthquake. The proposed model MSDA-DM performs better in all the cases.
- Another study [25] based on semi-supervised DA with Hurricane Harvey as the source and Hurricane Irma as the target disaster shows an F1-score of 77.90%. The proposed unsupervised model MSDA-DM with Hurricane Harvey and Irma as the source and Hurricane Maria as the target achieves an F1-score of 78.95%, confirming the multi-source network's superiority over single-source. On the other hand, in the case of Harvey as the source and California wildfires as the target, the single-source model's performance is better than Case IV of MSDA-DM, where the two source disasters, Hurricane Harvey and the Mexico Earthquake, and the target disaster California wildfires three are different leading to the poor performance of the model.

5.3 Ablation studies

For the proposed method based on unsupervised domain adaptation, the most critical hyperparameters to be set are backbone architecture, learning rate (LR) and regularization

Table 8 Comparison of the proposed model MSDA-DM with the recent studies on disaster datasets. The parameters used for comparison are Accuracy (Acc.), Precision (P), Recall (R), and F1-score (all in %)

Ref.	Approach	Source Disaster(s)	Target Disaster	Acc.	P	R	F1-Score
[10]	Fine-tuned BERT	The combined dataset of Hurricane Harvey, Hurricane Maria and Hurricane Irma		75.37	-	-	-
	CNN GRU			67.63	-	-	-
	RCNN			71.77	-	-	-
	BiGRU			67.88	-	-	-
	CNN LSTM			58.09	-	-	-
	KMax CNN			66.50	-	-	-
	DPCNN			63.74	-	-	-
	BiLSTM			71.02	-	-	-
[49]	Binary classification of text and image tweets based on fusion	Hurricane Harvey		77.70	-	-	77.60
		Hurricane Maria		72.96	-	-	72.84
		Hurricane Irma		73.82	-	-	73.55
		Mexico Earthquake		74.29	-	-	74.25
		California Wildfires		65.00	-	-	64.00
		Iraq-Iran Earthquake		68.18	-	-	67.00
[35]	ANN and CNN	Hurricane Harvey		75.90	76.00	76.00	76.00
[2]	Semi-Supervised DA with Graph Embeddings and adversarial training to classify the text tweets into two classes: relevant and non-relevant	Nepal Earthquake	Queensland Earthquake	-	67.48	65.90	65.92
		Queensland Earthquake	Nepal Earthquake	-	58.63	59.00	59.05
[3]	Semi-Supervised inductive Learning to classify the text tweets into two classes relevant and irrelevant	Nepal Earthquake		-	-	-	52.32
		Queensland Earthquake		-	-	-	75.08
[25]	Semi-supervised DA for the binary classification of image tweets	Hurricane Harvey	Hurricane Irma	77.90	77.88	77.97	77.90
		Hurricane Harvey	California Wildfires	79.87	79.11	77.27	77.90
[33]	Unsupervised Domain Adversarial Neural network to classify images as damage and no-damage	Ecuador Earthquake	Mathew Hurricane	68.7	79.1	68.3	72.6

Table 8 (continued)

Ref.	Approach	Source Disaster(s)	Target Disaster	Acc.	P	R	F1-Score
		Mathew Hurricane Nepal Earthquake	Ruby Typhoon Mathew Hurricane	68.1 70.6	66.9 86.0	77.4 63.4	71.3 72.4
[26]	Unsupervised DA using Maximum Mean Discrepancy metric to classify image tweets	Hurricane Harvey	Six disasters (Average of all Six)	79.33	78.23	78.09	77.75
-	**Baseline Model (SSDA):** Single-Source DA for image classification	Hurricane Harvey Hurricane Harvey	Hurricane Maria Iraq-Iran Earthquake	76.43 72.95	76.61 71.12	76.47 73.48	76.41 71.22
		Hurricane Harvey	California Wildfires	78.57	77.10	77.00	77.17
-	**Baseline Model (CSDA):** Combined-Source DA for the classification of image tweets	Hurricane Harvey + Hurricane Irma	Hurricane Maria	77.31	77.36	77.29	77.29
		Hurricane Harvey + Hurricane Irma	Iraq-Iran Earthquake	73.77	74.16	77.29	73.03
		Hurricane Harvey + Mexico Earthquake	Iraq-Iran Earthquake	82.79	80.61	82.71	81.18
		Hurricane Harvey + Mexico Earthquake	California Wildfires	78.57	77.48	78.36	77.77
-	**Proposed Model (MSDA-DM):** Multi-Source DA for the binary classification of image tweets: 'informative' and 'non-informative'	Hurricane Harvey + Hurricane Irma	Hurricane Maria	**79.08**	**79.66**	**79.04**	**78.95**
		Hurricane Harvey + Hurricane Irma	Iraq-Iran Earthquake	**79.51**	**77.43**	**80.27**	**78.09**
		Hurricane Harvey + Mexico Earthquake	Iraq-Iran Earthquake	**83.61**	**81.47**	**83.32**	**82.05**
		Hurricane Harvey + Mexico Earthquake	California Wildfires	72.05	70.99	71.85	71.55

Table 9 Performance of SSDA with Harvey as the source and Maria as the target disaster for five backbone architectures

Backbone Architecture	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ResNet-50	66.3	66.88	66.41	66.1
DenseNet-201	67.62	72.21	68.5	66.93
EfficientNet-B3	70.04	71.22	70.77	70.54
VGG-16	75.99	77.24	76.76	76.58
VGG-19	**77.53**	**78.28**	**78.19**	**78.19**

parameter (λ), which are done by conducting the following experiments. All the experiments are performed three times, and their average is listed.

5.3.1 Backbone architecture

Five pre-trained models ResNet-50 [17], DenseNet-201 [19], EfficientNet-B3, VGG-16, and VGG-19 [48] are used as backbone architectures for the model SSDA with Hurricane Harvey as the source and Hurricane Maria as target disaster. The results summarized in Table 9 confirm that the network gives the highest accuracy (77.53%) when VGG-19 is used as the backbone architecture. Thus, we choose VGG-19 as the backbone network for all the experiments performed in this study.

5.3.2 Learning rate

To optimize the hyperparameter learning rate (LR) of the network, a set of values is taken, and the performance of the proposed architecture MSDA-DM is noted for each of these values.

Table 10 Influence of the hyperparameter LR on the performance of MSDA-DM for Hurricane Harvey + Irma as the source and Hurricane Maria as the target. The model with LR=1e-03 gives the best accuracy

Learning Rate`(LR)	Training Accuracy (%)	Test Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
1e-01	50.5	49.34	24.67	50	33.04
1e-02	77.7	71.15	71.15	71.15	71.15
1e-03	**82.44**	**78.63**	**78.63**	**78.63**	**78.63**
1e-04	81.11	78.19	78.29	78.15	78.16

Table 11 Influence of different weights for domain loss and source classification loss terms on the performance of the MSDA-DM model with Hurricane Harvey and Irma as source disasters and Hurricane Maria as the target

Weight Domain Loss	Weight Source Classification Loss	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
10	1	70.70	70.74	70.73	70.70
0.1	1	77.31	77.51	77.25	77.24
1	10	71.15	71.27	71.19	71.13
0.1	10	72.69	72.69	72.69	72.69
10	0.1	54.63	58.53	55.05	49.71
1	0.1	77.91	72.91	72.91	72.91
1	**1**	**79.08**	**79.66**	**79.04**	**78.95**

The results of Hurricane Harvey and Irma as source disasters and Hurricane Maria as the target are listed in Table 10. Based on the results, we have chosen the learning rate as 1e-03 for all our experiments as the model gives the best accuracy as 78.63% and the best F1-score of 78.63%.

5.3.3 Loss function weights

We study the effect of different weights λ for the loss terms in our loss function (Eq. 1). Three values of λ are taken 0.1,1,10 for the loss terms, and the results for the proposed model MSDA-DM for Hurricane Harvey and Irma as source disasters and Hurricane Maria as the target are listed in Table 11. We can infer from the results that giving a higher weight to either of the loss terms does not lead to a network that can generalize well on the source and target domains, which is our goal in domain adaptation. Thus, the value of λ is taken as one for both loss terms.

6 Conclusions

We have introduced a novel deep domain adaptation model, Multi-Source Domain Adaptation for Disaster Management (MSDA-DM), to classify the images posted on microblogs at the time of disaster into two classes *informative* and *non-informative*. MSDA-DM is based on adversarial learning where it receives input from multiple source disasters with labeled images and one target disaster with unlabeled images and trains a network that aims to learn the domain invariant features from these disaster domains to solve the target task. We demonstrate that MSDA-DM can outperform the baselines and current state-of-the-art methods through multiple experiments, leading to better generalization in the target domain. The main challenge of multi-source DA is the choice of the sources taken for training since, unlike single-source DA that handles only source-target distribution drift, the multi-source DA network has to address both source-target and source-source distribution drifts. In the future, we would like to extend our work to more than two source disasters and also explore a semi-supervised setting to improve the performance.

Declarations

References

1. Alam F, Ofli F, Imran M (2018) Processing social media images by combining human and machine computing during crises. Int J Human–Computer Interact 34:311–327. https://doi.org/10.1080/10447318.2018.1427831
2. Alam F, Joty S, Imran M (2018) Domain adaptation with adversarial training and graph Embeddings. In: Proceedings of the 56th annual meeting of the Association for Computational Linguistics (volume 1: long papers). Association for Computational Linguistics, Stroudsburg, pp 1077–1087

3. Alam F, Joty S, Imran M (2018) Graph based semi-supervised learning with convolution neural networks to classify crisis related tweets. 12th Int AAAI Conf web Soc media. ICWSM 2018:556–559

4. Alam F, Ofli F, Imran M (2018) CrisisMMD: multimodal twitter datasets from natural disasters. In: 12th Int. AAAI Conf. Web Soc. Media, ICWSM 2018. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM18/paper/view/17038

5. Ben-David S, Blitzer J, Crammer K, Kulesza A, Pereira F, Vaughan JW (2010) A theory of learning from different domains. Mach Learn 79:151–175. https://doi.org/10.1007/s10994-009-5152-4

6. Bengio Y (2011) Deep learning of representations for unsupervised and transfer learning. JMLR Work Conf Proc 7:1–20

7. Borgwardt KM, Gretton A, Rasch MJ, Kriegel HP, Scholkopf B, Smola AJ (2006) Integrating structured biological data by kernel maximum mean discrepancy. Bioinformatics 22:e49–e57. https://doi.org/10.1093/bioinformatics/btl242

8. Chen Z, Wei P, Zhuang J, Li G, Lin L (2021) Deep CockTail networks a universal framework for visual multi-source domain adaptation. Int J Comput Vis 129:2328–2351. https://doi.org/10.1007/s11263-021-01463-x

9. Dai Y, Liu J, Ren X, Xu Z (2020) Adversarial training based multi-source unsupervised domain adaptation for sentiment analysis. Proc AAAI Conf Artif Intell 34:7618–7625. https://doi.org/10.1609/aaai.v34i05.6262

10. Fan C, Wu F, Mostafavi A (2020) A hybrid machine learning pipeline for automated mapping of events and locations from social media in disasters. IEEE Access 8:10478–10490. https://doi.org/10.1109/ACCESS.2020.2965550

11. Ganin Y, Lempitsky V (2014) Unsupervised domain adaptation by backpropagation. 32nd Int Conf Mach learn ICML 2015 2:1180–1189

12. Ganin Y, Ustinova E, Ajakan H, et al (2017) Domain-adversarial training of neural networks. In: Advances in Computer Vision and Pattern Recognition. pp. 189–209

13. Ghifary M, Kleijn WB, Zhang M, et al (2016) Deep reconstruction-classification networks for unsupervised domain adaptation. 1–21

14. Ghosh S, Desarkar MS (2018) Class specific TF-IDF boosting for short-text classification: Application to short-texts generated during disasters. In: Companion of the the web conference 2018 on the web conference 2018 - WWW '18. ACM Press, New York, pp 1629–1637. https://doi.org/10.1145/3184558.3191621

15. Ghosh S, Ghosh K, Ganguly D, Chakraborty T, Jones GJF, Moens MF, Imran M (2018) Exploitation of social Media for Emergency Relief and Preparedness: recent research and trends. Inf Syst Front 20:901–907. https://doi.org/10.1007/s10796-018-9878-z

16. Goodfellow IJ, Pouget-Abadie J, Mirza M et al (2014) Generative adversarial nets. Adv Neural Inf Process Syst 3:2672–2680

17. He K, Sun J (2016) Deep residual learning for image recognition. In: 2017 IEEE Conf. Comput. Vis. Pattern recognition(CVPR). https://doi.org/10.1109/CVPR.2016.90

18. Hiltz SR, Lee A, Imran M et al (2020) International journal of disaster risk reduction exploring the usefulness and feasibility of software requirements for social media use in emergency management. Int J Disaster Risk Reduct 42:101367. https://doi.org/10.1016/j.ijdrr.2019.101367

19. Huang G, Liu Z, Maaten L Van Der, Weinberger KQ (2017) Densely Connected Convolutional Networks. In: 2017 IEEE Conf. Comput Vis Pattern Recognit https://doi.org/10.1109/CVPR.2017.243

20. Imran M, Castillo C, Diaz F, Vieweg S (2015) Processing social media messages in mass emergency: A survey. ACM Comput Surv 47:1–38. https://doi.org/10.1145/2771588

21. Imran M, Ofli F, Caragea D, Torralba A (2020) Using AI and social media multimodal content for disaster response and management: opportunities, challenges, and future directions. Inf Process Manag 57:102261. https://doi.org/10.1016/j.ipm.2020.102261

22. Joseph JK, Dev KA, Pradeepkumar AP, Mohan M (2018) Big data analytics and social media in disaster management. Integr Disaster Sci Manag Glob Case Stud Mitig Recover. https://doi.org/10.1016/B978-0-12-812056-9.00016-6

23. Karimpour M, Noori Saray S, Tahmoresnezhad J, Pourmahmood Aghababa M (2020) Multi-source domain adaptation for image classification. Mach Vis Appl 31(6):44. https://doi.org/10.1007/s00138-020-01093-2

24. Khattar A, Ouadri SMK (2021) A semi-supervised domain adaptation approach for diagnosing SARS-CoV-2 and its variants of concern (VOC). In: 2021 9th international conference on reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO). IEEE, pp. 1–9

25. Khattar A, Quadri SMK (2021) Deep domain adaptation approach for classification of disaster images. In: Intelligent Data Communication Technologies and Internet of Things. Springer, pp 245–259. http://link.springer.com/10.1007/978-981-15-9509-7_2

26. Khattar A, Quadri SMK (2022) Generalization of convolutional network to domain adaptation network for classification of disaster images on twitter. Multimed Tools Appl 1–28. https://doi.org/10.1007/s11042-022-12869-1

27. Khattar A, Quadri SMK (2020) emerging role of artificial intelligence for disaster management based on microblogged communication. Int Conf Innov Comput Commun. https://doi.org/10.2139/ssrn.3562973

28. Khattar A, Jain PR, Qaudri S (2020) Effects of the disastrous pandemic COVID 19 on learning styles, activities and mental health of young Indian students-a machine learning approach. In: 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, pp 1190–1195. https://ieeexplore.ieee.org/document/9120955/

29. Laparra E, Bethard S, Miller TA (2020) Rethinking domain adaptation for machine learning over clinical language. JAMIA Open 3:146–150. https://doi.org/10.1093/jamiaopen/ooaa010

30. Li M, Zhang T, Chen Y, Smola AJ (2014) Efficient mini-batch training for stochastic optimization. In: Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining. ACM, New York, pp 661–670

31. Li H, Guevara N, Herndon N, et al (2015) Twitter mining for disaster response: a domain adaptation approach. ISCRAM 2015 Conf proc - 12th Int Conf Inf Syst Cris response Manag 2015-Janua

32. Li X, Caragea D, Zhang H, Imran M (2019) Localizing and quantifying infrastructure damage using class activation mapping approaches. Soc Netw Anal Min 9(1):1–15. https://doi.org/10.1007/s13278-019-0588-4

33. Li X, Caragea C, Caragea D, et al (2019) Identifying disaster damage images using a domain adaptation approach. Proc Int ISCRAM Conf 2019-May:633–645

34. Liu H, Guo F, Xia D (2021) Domain adaptation with structural knowledge transfer learning for person re-identification. Multimed Tools Appl 80:29321–29337. https://doi.org/10.1007/s11042-021-11139-w

35. Madichetty S, Sridevi M (2019) Detecting informative tweets during disaster using deep neural networks. 2019 11th Int Conf Commun Syst Networks 2061:709–713. https://doi.org/10.1109/COMSNETS.2019.8711095

36. Nguyen DT, Al MKA, Joty S et al (2016) Rapid classification of crisis-related data on social networks using convolutional neural networks. Proc 11th Int Conf web Soc media. ICWSM 2017:632–635

37. Pan SJ, Yang Q (2010) A survey on transfer learning. IEEE Trans Knowl Data Eng 22:1345–1359. https://doi.org/10.1109/TKDE.2009.191

38. Pandey N, Natarajan S (2016) How social media can contribute during disaster events? Case study of Chennai floods 2015. In: 2016 international conference on advances in computing, Communications and Informatics (ICACCI). IEEE, pp. 1352–1356

39. Paszke A, Gross S, Massa F, et al (2019) PyTorch: an imperative style, High-Performance Deep Learning Library Adv Neural Inf Process Syst 32

40. Peng X, Bai Q, Xia X, Huang Z, Saenko K, Wang B (2019) Moment matching for multi-source domain adaptation. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 1406–1415. https://ieeexplore.ieee.org/document/9010750/

41. Phengsuwan J, Shah T, Thekkummal NB, Wen Z, Sun R, Pullarkatt D, Thirugnanam H, Ramesh MV, Morgan G, James P, Ranjan R (2021) Use of social media data in disaster management: a survey. Futur Internet 13:1–24. https://doi.org/10.3390/fi13020046

42. Pizzati F, De Charette R, Zaccaria M, Cerri P (2020) Domain bridge for unpaired image-to-image translation and unsupervised domain adaptation. Proc - 2020 IEEE winter Conf Appl Comput vision. WACV 2020:2979–2987. https://doi.org/10.1109/WACV45572.2020.9093540

43. Popel M, Tomkova M, Tomek J, Kaiser Ł, Uszkoreit J, Bojar O, Žabokrtský Z (2020) Transforming machine translation: a deep learning system reaches news translation quality comparable to human professionals. Nat Commun 11:1–15. https://doi.org/10.1038/s41467-020-18073-9

44. Robertson BW, Johnson M, Murthy D, Smith WR, Stephens KK (2019) Using a combination of human insights and 'deep learning' for real-time disaster communication. Prog Disaster Sci 2:100030. https://doi.org/10.1016/j.pdisas.2019.100030

45. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L (2015) ImageNet large scale visual recognition challenge. Int J Comput Vis 115:211–252. https://doi.org/10.1007/s11263-015-0816-y

46. Schoenauer-Sebag A, Heinrich L, Schoenauer M, et al (2019) Multi-domain adversarial learning. 7th Int Conf learn represent ICLR 2019 1–24

47. Shen J, Qu Y, Zhang W, Yu Y (2018) Wasserstein distance guided representation learning for domain adaptation. In: Proceedings of the AAAI Conference on Artificial Intelligence. https://dl.acm.org/doi/abs/10.5555/3504035.350453

48. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. 3rd Int Conf learn represent ICLR 2015 - Conf track proc 1–14

49. Sreenivasulu M et al (2020) Classifying informative and non-informative tweets from the twitter by adapting image features during disaster. Multimed Tools Appl 79(39):28901–28923. https://doi.org/10.1007/s11042-020-09343-1
50. SMS M (2021) A stacked convolutional neural network for detecting the resource tweets during a disaster. Multimed Tools Appl 80:3927–3949. https://doi.org/10.1007/s11042-020-09873-8
51. Sun B, Saenko K (2016) Return of frustratingly easy domain adaptation. In: proceedings of the thirtieth AAAI conference on artificial intelligence (AAAI-16) conference. Pp 2058–2065
52. Tang J, Yang S, Wang W (2021) Social media-based disaster research: development, trends, and obstacles. Int J Disaster Risk Reduct 55:102095. https://doi.org/10.1016/j.ijdrr.2021.102095
53. Tzeng E, Saenko K, Hoffman J, Darrell T (2017) Adversarial discriminative domain adaptation. Proc IEEE Conf Comput Vis Pattern Recognit 2017:7167–7176
54. Vieweg S, Hughes AL, Starbird K, Palen L (2010) Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 1079–1088. https://doi.org/10.1145/1753326.1753486
55. Wang M, Deng W (2018) Deep visual domain adaptation: a survey. Neurocomputing 312:135–153. https://doi.org/10.1016/j.neucom.2018.05.083
56. Zhang N, Mohri M, Hoffman J (2021) Multiple-source adaptation theory and algorithms. Ann Math Artif Intell 89(3):237–270. https://doi.org/10.1007/s10472-020-09716-0
57. Zhou Q, Zhou W, Wang S, Xing Y (2021) Duplex adversarial networks for multiple-source domain adaptation. Knowledge-Based Syst 211:106569. https://doi.org/10.1016/j.knosys.2020.106569
58. Zhu Y, Zhuang F, Wang D (2019) Aligning domain-specific distribution and classifier for cross-domain classification from multiple sources. Proceedings of the AAAI Conference on Artificial Intelligence p 33: 5989–5996