Computers and Electronics in Agriculture 244 (2026) 111450

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

MCDA: a novel domain adaptation with multiple classifiers for crop mapping

Changhong Xu [a,b], Maofang Gao [a,*], Yuanwei Chen [c], Jingwen Yan [d]

[a] State Key Laboratory of Efficient Utilization of Arable Land in Northern China, Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China
[b] College of Engineering, Shantou University, Shantou 515063, China
[c] Space Star Technology CO., LTD., Beijing 100086 China
[d] School of Artificial Intelligence, Guangzhou Institute of Science and Technology, 510540, China

ARTICLE INFO

Keywords:
Remote Sensing
Domain Adaptation
Crop Mapping

ABSTRACT

Accurate crop mapping is conducive to optimizing agricultural production, improving resource utilization efficiency, as well as supporting precision agriculture management and environmental monitoring. However, crop mapping often relies on a large amount of reference data as labels. Moreover, the generalization and scalability of approaches are challenging issues due to the influence of geographic locations, climatic conditions, and crop characteristics. Thereby, this study proposed the Domain Adaptation with Multiple Classifiers (MCDA) model that enables crop mapping in target domains without reference data, which consists of a feature extractor and three classifiers. Initially, the data from the source domain was trained using the feature extractor and individual classifier. Next, the feature extractor was fixed and two classifiers were trained. The decision boundaries of classifiers were maximized by calculating discrepancies between different classification results in target domain. Finally, the two classifiers were fixed and the feature extractor was trained to map the data from the source domain and target domain into a better feature space. In this study, three sets of experiments were designed to validate the stability and generalization of MCDA model by selecting different study areas in four countries to classify typical crops such as maize and winter wheat, rice, and soybeans using time series Vegetation Indices (VIs). Comparing the 1DCNN and LSTM models, as well as the four UDA models, the experimental results indicate that the MCDA model outperforms the other models in most transfer cases. With the appropriate selection of crop features, the method proposed in this study can achieve more accurate crop mapping effectively identifying crop types and field boundaries. The implementation code of our method has been publicly available at https://github.com/abbyxuchanghong-cmd/MCDA.

1. Introduction

Agriculture is one of the basic industries of human society, supporting the sustainable development of economy, environment and society worldwide. Crop classification (Zhong et al., 2019), as an important part of agricultural research, clarifies the spatial and temporal distribution of crops and planting structure (Cai et al., 2018). It is a foundation for promoting the advancement of smart agriculture. Accurate crop mapping plays a vital role in the area estimation, growth monitoring, yield evaluation, pest control, disaster warning, and other aspects related to crops (Xu et al., 2023).

With the development of remote sensing technology, the large-scale crop mapping (Li et al., 2023; Wei et al., 2021; Yang et al., 2024) has been made possible, by rapidly acquiring the satellite image over extensive areas. Besides, remote sensing imagery can also capture reflectance information in different spectral bands (Li et al., 2021), reflecting characteristics of crops such as health status, growth stages, and biophysical properties. Crop mapping using deep learning mainly adopt two mainstream research strategies, namely using spatial and temporal features of crops to classify different crop types (Cai et al., 2018). The spatial information in remote sensing images is utilized to achieve exact identification of different crops by analyzing the morphological features, colour distribution, texture structure, and more (Xu et al., 2020). On the other hand, crops are classified by extracting temporal features from remote sensing images, including growth period, growth rates, peak variations, and others (Feng et al., 2023).

* Corresponding author.
E-mail address: gaomaofang@caas.cn (M. Gao).

https://doi.org/10.1016/j.compag.2026.111450
Received 16 January 2025; Received in revised form 13 January 2026; Accepted 14 January 2026

Traditional crop mapping methods typically depend on a large amount of labels, which is not only costly to obtain, but also a very time-consuming process. Furthermore, the results of labels are significantly likely to be influenced by the subjective factors of human. Due to these limitations, it has gradually become a mainstream research interest to classify crops in study areas without labels (Wang et al., 2019). The acquisition of remote sensing images is greatly affected by regional, temporal and climatic conditions, which results in notable distribution differences between data from different domains. Such distribution discrepancies (Ma et al., 2024) make it difficult to directly apply the trained classification model in one area to another. Therefore, effectively addressing these issues and improving the generalization of models in different domains has become a crucial research topic in the field of crop classification.

Unsupervised Domain Adaptation (UDA) (Liu et al., 2022) aims to overcome the domain shift by efficient and accurate transfer of models from a labeled source domain to an unlabeled target domain. The UDA also aims to reduce the dependency on labels in the target domain, powerfully addressing the issue of insufficient labels in practical applications. The strategies of UDA are broadly divided into two main categories, namely discrepancy-based and adversarial-based methods (Ragab et al., 2023). The primary goals of discrepancy-based methods are to minimize the statistical distance between features of the source and target domains, thereby alleviating the problem of domain shift. The mainstream models in this category include Deep Domain Confusion (DDC) (Tzeng et al., 2014), Higher-order Moment Matching (HoMM) (Chen et al., 2020), Minimum Discrepancy Estimation for Deep Domain Adaptation (MMDA) (Rahman et al., 2020), Deep Subdomain Adaptation (DSAN) (Zhu et al., 2020) and more. Whereas, adversarial-based methods utilize a domain discriminator network to train a feature extractor that produces domain invariant features in different domains. The notable models in this category include Domain-Adversarial Training of Neural Networks (DANN) (Ganin et al., 2016), Conditional Adversarial Domain Adaptation (CDAN) (Long et al., 2018), Convolutional deep Domain Adaptation model for Time Series data (CoDATS) (Wilson et al., 2020), Adversarial Spectral Kernel Matching (AdvSKM) (Liu et al., 2021) and so on.

Vegetation indices (VIs) (Gao et al., 2020) that enhance characteristics or details of crops by combinatorial operations of multiple spectral reflectance from satellite data make it possible to classify crops without labels (You et al., 2023). For example, Landsat Surface Reflectance and VIs were used to classify maize, soybeans, wheat and alfalfa from nine states in the Midwest US using the random forest and two unsupervised methods in the absence of reference data (Wang et al., 2019). The high confidence pixels from crop maps in America and VIs from harmonized Landat-8 and Sentinel-2 were used to train different random forests (Hao et al., 2020). Then, the VIs of different lengths were used to classify spring wheat and canola in Alberta, cotton, maize and winter wheat in Hengshui city, as well as maize and soybeans in Nebraska. With 11 states in the United States as the source domain and three provinces in Northeast China as the target domain, the 10 spectral bands and five VIs were collected in the Sentinel-2 images to classify maize and soybeans using the multi-source UDA crop classification model (MUCCM) (Wang et al., 2022). What's more, with four states in the U.S. and two provinces in Northeast China as the study areas, the feasibility of crop mapping for maize and soybean was verified in space, time, and space–time, using the deep adaptation crop classification network (DACCN) model (Wang et al., 2023). The crops such as maize, rice, wheat, soybeans and others were classified by the Variable-dimensional Symmetric Network with Position Encoding (VPSNet) and transfer strategy based on the Inter-Regional Discrepancies in Crop Time Series (IRDCTS) to solve the problem of limited labels (Wang et al., 2024).

The supervised learning approaches for crop mapping relies on a large number of labeled samples, which are often costly and time-consuming to acquire. Moreover, remote sensing images of crops from different regions exhibit significant domain shifts due to variations in climate, soil, and other factors. To address these challenges, the UDA methods leverage the temporal characteristics of crops growth to learn domain-invariant features in the target domain without labels, thereby enabling cross-domain classification and improving the generalization ability of the model. In this study, with reference to the labels from source domain, we proposed the Domain Adaptation with Multiple Classifiers (MCDA) model to classify crops such as maize and winter wheat, rice and soybean in the target domains without labels from four countries, using the Sentinel-2 satellite images. The specific contributions of this study are as follows:

- The MCDA model was trained in three steps. In the first training step, the independent classifier was used to classify the features of source domain after the extraction by Convolutional Neural Network (CNN). It can ensure that the training process in the source domain was not affected by the target domain data, thereby allowing the feature extractor to better capture critical information from the source domain.
- Fixing the parameters of the feature extractor and training the two classifiers in the second training step, the knowledge distillation loss was used to calculate the discrepancies in classification results for the target domain. By maximizing the decision boundaries of classifiers, it is possible to effectively identify the differential samples of the same class that are within the target domain but outside the support of the source domain, thus improving the accuracy of crop classification.
- Using the idea of transfer learning to construct MCDA model for crop distribution mapping in target domains without labels, provided a feasible research solution for the crop classification in areas where it is difficult to obtain field samples. It can also provide data and technical support for crop mapping with irregular spatial distribution, complex planting structure and unbalanced proportions for crop classes.

2. Materials

2.1. Study areas

A total of six representative study areas with typical crops in four countries (the United States, the Czech Republic, China and Argentina) were selected to validate the effectiveness and generalization of the MCDA model, as shown in Fig. 1. The Lower Mississippi (Ochs et al., 2023) spans across Missouri, Kentucky, Arkansas, Tennessee, Louisiana and Mississippi, with a land area of 266,505 km^2. It has a flat topography, mild climate, and abundant rainfall, and is classified as a subtropical humid region. Plentiful precipitation, ample sunlight and fertile soil provide favourable conditions for crop growth in the Lower Mississippi, with large-scale cultivation of crops such as maize, cotton, rice, and soybeans.

Kansas is located in the central United States, covering an area of 213,096 km^2, as shown in Fig. 1 (b). It is part of the flat North American prairie, which gradually increases in topography from east to west, with an average elevation of 610 m. Kansas has a temperate continental climate, with an annual average temperature of 13°C and an annual average precipitation of about 690 mm. Kansas has vast farmlands and fertile soil, making it well-suited for crop cultivation and livestock development. The main crops include wheat, maize, soybeans, sorghum, and others.

The Czech Republic is located in central Europe, with a land area of 78,900 km^2, as shown in Fig. 1 (c). The Czech is situated in a quadrilateral basin that rises on three sides, and most of the basin is below 500 m above sea level. The Czech has a temperate continental climate, with an average temperature of about 18.5°C in summer and −3°C in winter, and an annual average precipitation of 683 mm. The main crops are wheat, barley, maize, potatoes, sugar beet and so on.

Cordoba is located in central Argentina, covering an area of 165,321

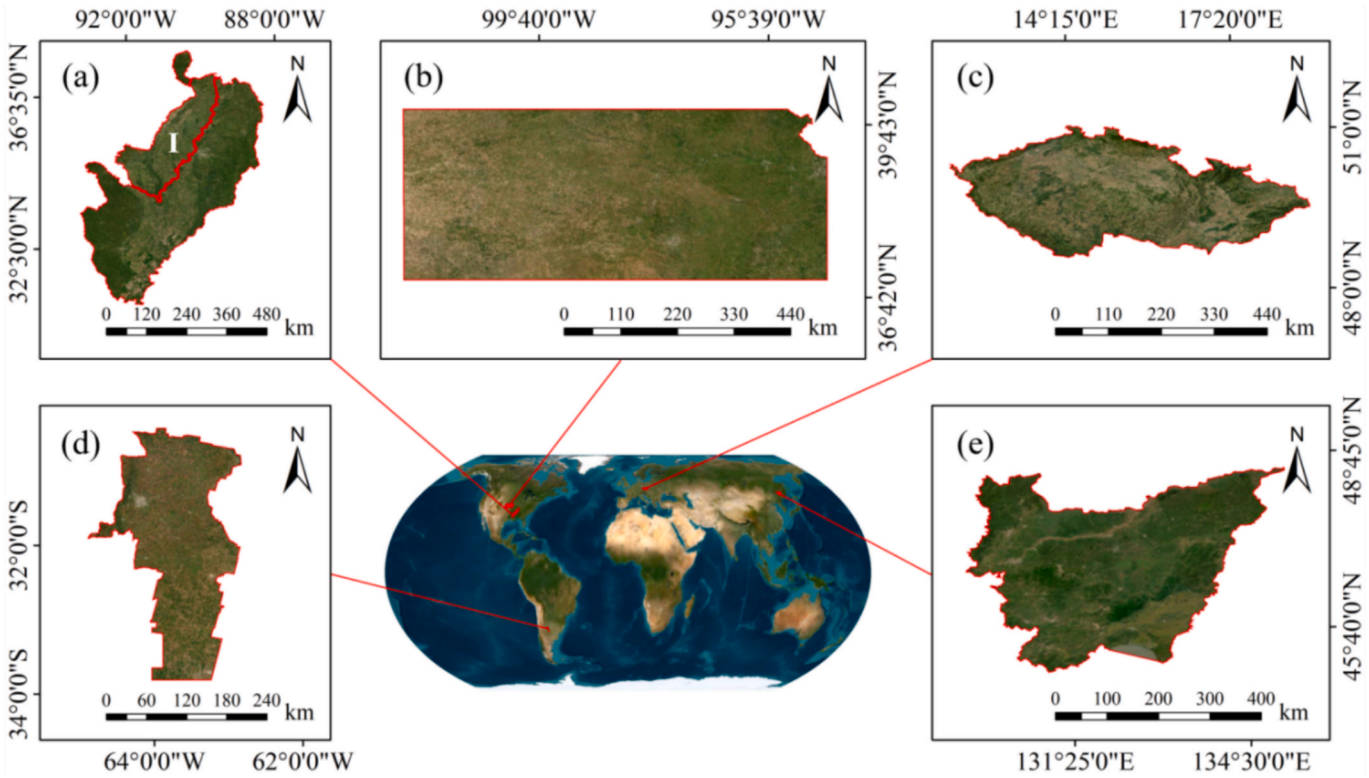

Fig. 1. The geographic locations of (a) the Lower Mississippi Region in the United States, (b) the Kansas, United States, (c) the Czech Republic, (d) the Central Region of Cordoba in Argentina and (e) the Northeast Region of Heilongjiang Province in China. The I in (a) represents a study area nested within another.

km^2, with an average elevation of 2,000 m. It has a temperate continental climate, with an average maximum temperature of 30 to 32°C and an average minimum temperature of 16 to 18°C in the east during summer. In winter, average high temperatures range from 15 to 18°C, while average low temperatures range from 2 to 4°C. Monthly precipitation ranges from 90 to 120 mm in summer, but in winter, monthly rainfall is less than 25 mm. Agriculture and animal husbandry account for 10% of Cordoba's production, with agriculture focusing on soybeans, wheat, maize, as well as other cereals.

Heilongjiang Province is located in northeastern China, with a total area of 473,000 km^2. The terrain is generally high in the northwest, north, and southeast, low in the northeast and southwest, comprising mountainous with an altitude of 300 to 1,000 m, tableland with an altitude of 200 to 350 m, and plains with an altitude of 50 to 200 m. Heilongjiang has a temperate continental monsoon climate, with an annual average temperature ranging between −5 and 4°C from north to south, gradually decreasing from northwest to southeast. Annual average precipitation ranges from 400 to 700 mm, more in the east and less in the west. The Mollisol in Heilongjiang is fertile and widely distributed, mainly used for cultivating crops such as rice, wheat, maize, sorghum, and soybeans.

Maize and winter wheat are the main study crops in Kansas and Czech. The growth period of maize is about 150 days, sown from April to May each year and harvested from September to October. The growth period of winter wheat is about 280 days, sown from September to November each year and harvested from May to July of the following year. Rice is the main study crop in the Lower Mississippi Region, as shown by I in Fig. 1 (a), and the Northeast Region of Heilongjiang Province, as shown in Fig. 1 (e). The growth period of rice is around 150 days, with cultivation from April to May each year and harvest from September to October. Soybeans is the main crop in the Lower Mississippi Region, as shown in Fig. 1 (a), and the Central Region of Cordoba, as shown in Fig. 1 (d). The growth period of soybeans is about 150 days, and in the Lower Mississippi, soybeans are generally sown from April to May each year and harvested from September to October. In Cordoba, however, soybeans are generally sown between October and December each year and harvested between March and May of the following year. For a more convenient description in the following, the Lower Mississippi Region I for rice is abbreviated as study area M_R, the Northeast Region of Heilongjiang Province is abbreviated as study area HLJ_R, the Lower Mississippi Region for soybeans is abbreviated as study area M_S, and the Central Region of Cordoba is abbreviated as study area S_C.

2.2. Data processing

2.2.1. Sentinel-2 data

Sentinel-2 is an Earth observation mission as part of the Copernicus programme, consisting of two satellites, Sentinel-2A and Sentinel-2B, with a revisit period of 5 days (Phiri et al., 2020). Each satellite is equipped with an identical multi-spectral instrument, which captures images covering 13 bands including visible light, near-infrared, and shortwave infrared. The spatial resolution ranges from 10 to 60 m, with the swath width of 290 km. The use of Sentinel-2 data provides reliable and effective inputs for crop mapping. Images in 6 bands of Red, Blue, Green, Near-Infrared (NIR), Short Wave Infrared (SWIR) 1, and SWIR2 were downloaded at ten-day intervals. With reference to the growth period of studied crops, Sentinel-2 images from October 2017 to September 2018 were downloaded in Kansas and Czech. Sentinel-2 images from April 2018 to October 2018 were downloaded for study areas M_R, HLJ_R and M_S. Since the study area S_C is located in the Southern Hemisphere, the seasons are opposite to other study areas. Thus, Sentinel-2 images from November 2017 to May 2018 were downloaded in study area S_C, corresponding to the growing seasons of local soybeans.

2.2.2. Data Reconstruction

Although the revisit period of Sentinel-2 is as high as 5 days, remote

sensing images for optical satellites still face the problem of corrupted or missing observation data due to various factors such as sensor failures, cloud cover, transmission errors, and so on. On Google Earth Engine (GEE) platform, Sentinel-2 data is reconstructed using the penalized least-square regression based on discrete cosine transform (DCT-PLS) method, selecting all available images with cloud coverage less than 30%.

In the DCT-PLS method, the variable x represents a series of equal interval original data, and the variable $\hat{x}$ represents a series of equal interval reconstructed data. The goal of DCT-PLS method is to find the optimal value of $\hat{x}$ that minimizes the objective function $\mathscr{F}$, which can be expressed as

$$\mathscr{F}(\hat{x}) = \boldsymbol{W} \bullet \mathrm{RSS} + \lambda \bullet P(\hat{x}) = \|\boldsymbol{W}^{1/2} \bullet (\hat{x} - x)\|^2 + \lambda \bullet \|\boldsymbol{D} \bullet \hat{x}\|^2 \quad (2\text{-}1)$$

where $\boldsymbol{W}$ is the weight matrix that w_i corresponds to each x_i and $w_i \in [0, 1]$, RSS is the Residual Sum-of-Squares, λ is a positive scalar used to control the degree of smoothing, P is the penalty function, $\boldsymbol{D}$ is the Laplace operator. According to Discrete Cosine Transform (DCT) and Inverse Discrete Cosine Transform (IDCT), $\hat{x}$ can be expressed as

$$\hat{x}_{k+1} = \mathrm{IDCT}(\Gamma \bullet \mathrm{DCT}(\boldsymbol{W}(x - \hat{x}_k) + \hat{x}_k)) \quad (2\text{-}2)$$

where k is the number of iterations and Γ is a diagonal matrix. The minimum value of the objective function $\mathscr{F}$ can be calculated by iteratively solving $\hat{x}$. The specific solution steps, experimental procedure, and open-source code can be referenced in Yang et al., (2022).

2.2.3. Reference data

The Cropland Data Layer (CDL) for the continental United States, European crop type map published by D Andrimont et al., (2021), crop type maps in Northeast China published by You et al., (2021), as well as soybean classification maps in South America published by Song et al., (2021) were used as labels in place of ground truth data. The CDL is created annually for the continental US using moderate resolution satellite imagery and extensive agricultural ground truth. It is a crop-specific land cover data layer with a resolution of 30 m, which contains information on the distribution of more than 100 crop types. According to the official CDL Metadata, the classification accuracy is generally between 85% and 95% for the major crop-specific land cover categories. The CDL is widely used as a reference in agriculture due to its advantages such as high classification accuracy and spatial resolution, timely update frequency, extensive coverage, diverse crop types and so on.

Based on the LUCAS 2018 Copernicus in-situ survey and Sentinel-1 observations, the European crop type map with a spatial resolution of 10 m was produced using the random forest algorithm to classify 19 different crop types. The overall accuracy of the European crop type map is reported as 80.3% when main crop classes are grouped. The CDL for the continental US and the European crop type map in 2018 were downloaded from the GEE platform. Using the layered mapping strategy, random forest classifiers, Sentinel-2 time series data and optimized features, crop maps of maize, soybean and rice in Northeast China from 2017 to 2019 were produced with a resolution of 10 m and a classification accuracy of over 80%. According to the satellite observations and sample field data, soybean classification maps in South America from 2000 to 2019 were generated, achieving the overall accuracies of 96%, 94% and 96% in 2017, 2018 and 2019, respectively. The CDL for the continental US and soybean classification maps in South America were resampled to a resolution of 10 m.

2.3. Data description

2.3.1. Vegetation indices

The Normalized Difference Vegetation Index (NDVI) and Barley Heading Date Index (BHDI) were selected to classify maize, winter wheat and others in Kansas and Czech. The NDVI, 2-band Enhanced Vegetation Index (EVI2), Land Surface Water Index (LSWI), Normalized Difference Water Index (NDWI) and Modified Normalized Difference Water Index (MNDWI) were selected to classify rice and others in study areas M_R and HLJ_R. The NDVI and Greenness and Water Content Composite Index (GWCCI) were selected to classify soybeans and others in study areas M_S and S_C. As mentioned above, the calculation formulas of different VIs are summarized in Table 1, where the ρ_{Red}, ρ_{Green}, ρ_{Blue}, ρ_{Nir}, ρ_{SWIR_1}, and ρ_{SWIR_2} represent the Sentinel-2 reflectance in the Red, Green, Blue, NIR, SWIR1, and SWIR2 bands, respectively. And the temporal profiles of different VIs are shown in Fig. 2, where gray, pink, green, blue, and purple indicate others, maize, winter wheat, rice, and soybeans, respectively. Fig. 2 (a) and (d) show the NDVI of maize, and winter wheat, and others in Kansas and Czech. Fig. 2 (b) and (e) show the NDVI of rice and others in study areas M_R and HLJ_R. Fig. 2 (c) and (f) show the NDVI of soybeans and others in study areas M_S and S_C. Fig. 2 (g) and (j) show the BHDI of maize, and winter wheat, and others in Kansas and Czech. Fig. 2 (h) and (k), (i) and (l), (m) and (p), as well as (n) and (q) show the EVI2, LSWI, NDWI, and MNDWI of rice and others in study areas M_R and HLJ_R. Fig. 2 (o) and (r) show the GWCCI of soybeans and others in study areas M_S and S_C.

Before applying the UDA approaches, the temporal profiles of VIs for different crops across regions were analyzed to examine the domain-invariant phenological features. The temporal profiles of NDVI are able to reflect the growth dynamics of different crops. As the crops grow and develop, more red light is absorbed while more NIR is reflected, leading to an increase in the values of NDVI. When crops enter the maturity stage, the cover of leaves reaches its maximum, resulting in the highest values of NDVI. During the later stage of maturity, leaves begin to age and turn yellow, and the reflectance of NIR band decreases, leading to a decrease in the values of NDVI. Therefore, the temporal profiles of NDVI first increase and then decrease during the growth periods of maize, rice and soybeans, as shown in Fig. 2 (a) to (f). Compared with the other crops, the winter wheat is typically planted in the autumn or winter, and as it grows, the values of NDVI begin to rise. Due to the low temperature and freezing in winter, the growth of winter wheat slows down or stagnates, causing a decrease in the values of NDVI. With the arrival of spring, when the weather warms up and the temperatures of soil rise, the winter wheat begins to resume growth and enters a rapid development stage, causing a rise in the values of NDVI again. As shown by the green lines in Fig. 2 (a) and (d), the temporal profiles of NDVI thus exhibit a clear double peak during the growth period of winter wheat. The temporal profiles of NDVI for winter wheat in Kansas and Czech show the similar trends but the peak values differ in magnitude. During the growth period of winter wheat, the temporal profiles of BHDI also exhibit a double peak in Fig. 2 (g) and (j). Whereas, the temporal profiles of BHDI exhibit a single peak during the growth period of maize.

For the rice in study areas M_R and HLJ_R, the temporal profiles of EVI2 in Fig. 2 (h) and (k) exhibit the extremely similar variation trends, while the peak values differ in magnitude. During the growth periods of rice in study areas M_R and HLJ_R, the variation trends of LSWI in Fig. 2 (i) and (l) are similar, and their peak values are highly close. The temporal profiles of NDWI are negative due to the absorption of visible light and reflection

Table 1
The calculation formulas of vegetation indices (VIs) used in this study.

VIs	Equations
NDVI	$\mathrm{NDVI} = (\rho_{Nir} - \rho_{Red})/(\rho_{Nir} + \rho_{Red})$
BHDI	$\mathrm{BHDI} = (\rho_{Green} - \rho_{Blue}) + (\rho_{Green} - \rho_{Red})$
EVI2	$\mathrm{EVI2} = 2.5^*(\rho_{Nir} - \rho_{Red})/(\rho_{Nir} + 2.4\rho_{Red} + 1)$
LSWI	$\mathrm{LSWI} = (\rho_{Nir} - \rho_{SWIR_2})/(\rho_{Nir} + \rho_{SWIR_2})$
NDWI	$\mathrm{NDWI} = (\rho_{Green} - \rho_{Nir})/(\rho_{Green} + \rho_{Nir})$
MNDWI	$\mathrm{MNDWI} = (\rho_{Green} - \rho_{SWIR_1})/(\rho_{Green} + \rho_{SWIR_1})$
GWCCI	$\mathrm{GWCCI} = \mathrm{NDVI}^*\rho_{SWIR_2}$

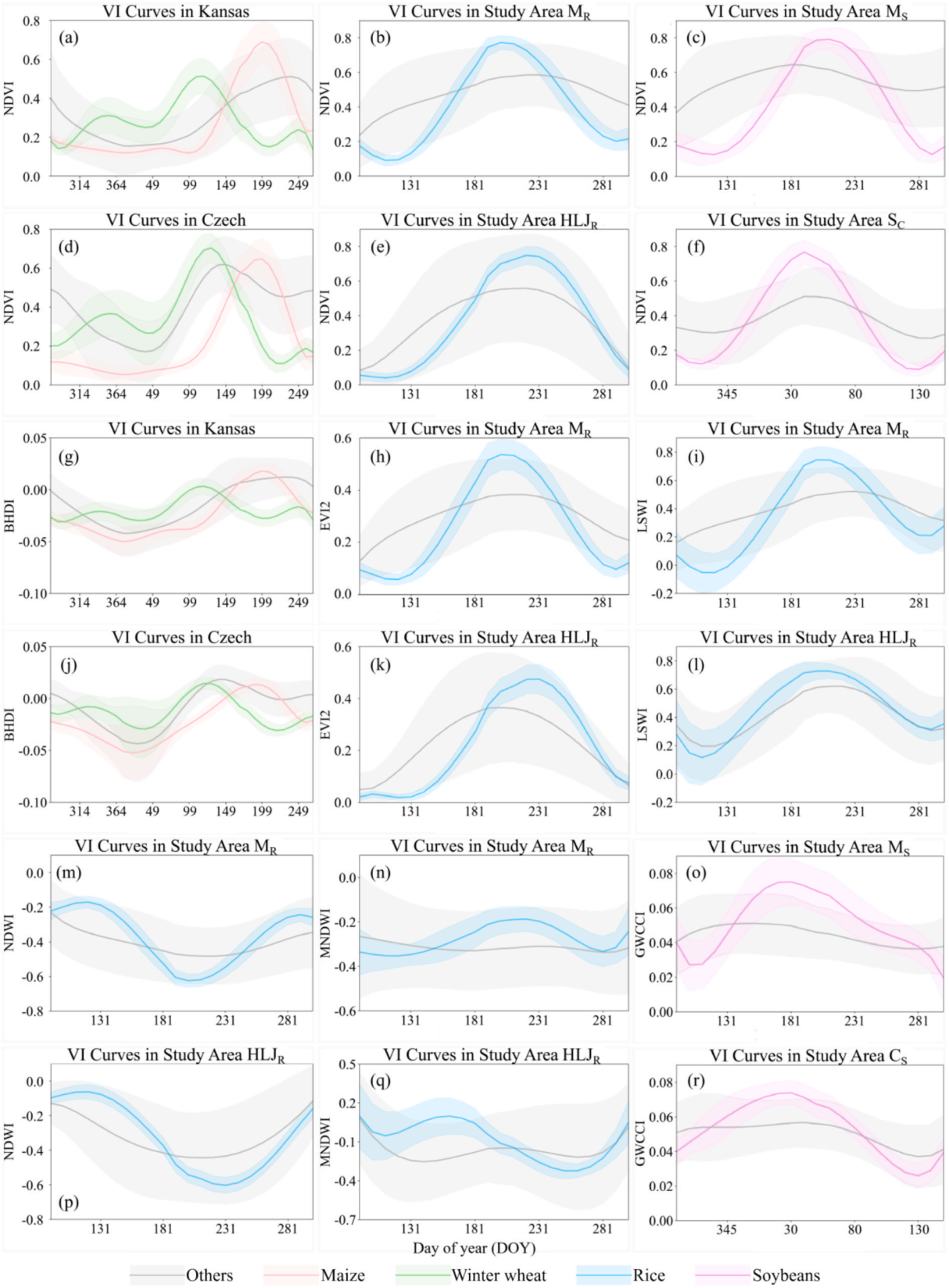

Fig. 2. The temporal profiles of different VIs. (a) to (f) show the NDVI of maize, winter wheat, rice, soybeans, and others in Kansas, Czech, and study areas M_R, HLJ_R, M_S, and S_C. (g) and (j) show the BHDI of maize, winter wheat, and others in Kansas and Czech. (h) and (k), (i) and (l), (m) and (p), as well as (n) and (q) show the EVI2, LSWI, NDWI, and MNDWI of rice and others in study areas M_R and HLJ_R. (o) and (r) show the GWCCI of soybeans and others in study areas M_S and S_C. Among them, gray, pink, green, blue, and purple represent others, maize, winter wheat, rice, and soybeans, respectively. The solid lines represent the mean values of VIs, and the shaded areas denote the standard deviations for different categories.

of near-infrared light by crops. And the temporal profiles of NDWI in Fig. 2 (m) and (p) decrease and then increase with the growth and development of rice. Although the peak values of MNDWI for rice in Fig. 2 (n) and (q) occur at different times, their temporal profiles in study areas M_R and HLJ_R exhibit a trend of first increasing and then decreasing. The GWCCI is the product of NDVI and the reflectance of SWIR1 band, where the values of reflectance are less than 1. Therefore, the values of GWCCI are generally smaller than the values of NDVI, as shown in Fig. 2 (o) and (r). The temporal profiles of GWCCI have a similar trend over the growth period of soybean in study areas M_S and S_C, and the peak values of GWCCI are considerably overlapping in Fig. 2 (o) and (r).

As illustrated in Fig. 2, despite differences in regional climate and phenological information for different study areas, the same crops exhibit similar variation trends and peak values of VIs. These similarities make it possible to map domain-invariant features. The underlying invariant growth patterns can be captured and aligned by the UDA approaches, thereby mitigating discrepancies caused by regional or phenological variations and enhancing the cross-regional transferability of models.

To verify the rationality and necessity of VIs selections, the mean values and standard deviations of the study crops and other types were presented in Fig. 2 for different study areas. In each experiment, the NDVI (Wang et al., 2023) was first selected as the primary index for distinguishing between crops and non-crops. The temporal profiles of NDVI for crops exhibit distinct seasonal variations during the crop growth periods. In contrast, the temporal profiles of NDVI for non-crops, such as bare soil, water, and buildings, remain relatively stable or consistently low throughout the seasons. The mean values of NDVI for others are significantly different from those of the study crops in Fig. 2 (a) to (f). And the standard deviations of NDVI for others are larger than those of the study crops at each time point. This is mainly because others include diverse land-cover types, such as non-crops and crops outside the focus of this study, each with their own spectral features. However, relying solely on the NDVI in the UDA approaches is insufficient to accurately distinguish different crop types due to the similar dynamics of NDVI. For example, the temporal profiles of NDVI for rice and soybeans show the similar variation trends in Fig. 2 (b) and (c).

As another major crop in Czech, barley exhibits the remarkably similar growth characteristics to winter wheat. The BHDI is primarily used to distinguish between barley and wheat, with specific details provided in Ashourloo et al., (2022). Thus, the NDVI and BHDI were selected as the key features to distinguish maize, winter wheat, and others for the UDA approaches in Kansas and Czech. The temporal profiles of NDVI for maize and soybeans are similar to those of rice, despite differences in their peak values. Especially in study area HLJ_R, the temporal profiles of NDVI for others, which include maize and soybeans as the interference crops, show the relatively large standard deviations in Fig. 2 (e), with their variation ranges covering those of rice. So the additional VIs, such as EVI2, LSWI, NDWI and MNDWI (Yang et al., 2024) were selected to distinguish rice from other crops in study areas M_R and HLJ_R. The EVI2 (Jiang et al., 2008) provides superior sensitivity in high biomass areas and minimizes the influence of soil and atmospheric effects compared with the NDVI. The LSWI (Chandrasekar et al., 2010; Zhang et al., 2023) is sensitive to the liquid water in vegetation and soil, making it useful for distinguishing rice from other crops during the early growth stage and irrigation periods. The NDWI (Mcfeeters, 1996) utilizes the Green and NIR bands to enhance the detection of open water, enabling the discrimination of rice from dryland crops or bare soil. And the MNDWI (Xu, 2006), as a modified version of NDWI, can further enhance the separability between rice and non-rice categories. Specifically in Fig. 2 (q), the temporal profiles of MNDWI for rice in study area HLJ_R exhibit significant differences from those of others. The GWCCI demonstrates a strong capability to accurately classify soybeans from other crops under the complex conditions, with the rationale described in Chen et al., (2023). And in Fig. 2 (o) and (r), the temporal profiles of GWCCI for soybeans in study areas M_S and S_C differ from those of others. Thus, the NDVI and GWCCI were used to extract soybeans from other types in this study.

2.3.2. Data statistics

The percentages of random points for different types in each study area were utilized to represent the approximate proportions of different crops within that area. Fig. 3 shows the number and proportions of points for studied categories in the training dataset. There are a total of 5,000 points in each study area. In Fig. 3, blue, pink, green, yellow, cyan, and purple indicate the number of points for different categories in Kansas, Czech, study area M_R, study area HLJ_R, study area M_S, and study area S_C, respectively. In Kansas and Czech, other accounts for 76.76% and 79.04% of the all points, while winter wheat accounts for 13.04% and 14.64% , respectively. The proportion of corn is the lowest, accounting for 10.20% and 6.32% in Fig. 3 (a). In study areas M_R and HLJ_R, the proportions of other are separately 87.30% and 72.02%. However, in study area HLJ_R, the proportion of rice is more than twice that of study area M_R, accounting for 27.98% and 12.70% respectively. In study areas M_S and S_C, the proportions of other are separately 76.22% and 57.80%. In addition, the proportions of soybeans are separately 23.78% and 42.20%, especially in study area S_C where the points of soybeans account for nearly half of the total points.

3. Methodology

3.1. Approach Overview

The workflow of this study consists of four parts as shown in Fig. 4, namely data preparation, domain adaptation, model selection and crop map.

(a) **Data Preparation**. Sentinel-2 time series images were reconstructed using the DCT-PLS method based on GEE platform at ten-day intervals, and then downloaded. Applying ArcGIS software, time-series spectral reflectance curves and corresponding crop types were extracted based on randomly distributed points on reference data. The time series VIs were smoothed using Savitzky-Golay (S-G) filter to produce source domain training dataset, source domain test dataset, target domain training dataset and target domain test dataset, denoted by X_S^{tr}, X_S^{te}, X_T^{tr} and X_T^{te}, respectively.

(b) **Domain Adaptation**. The crucial features of time-series VIs were extracted on the source and target domains using the feature extractor. The source domain training features were classified using the classifier, followed by calculating the source domain classification loss $\mathscr{L}_{C_S}$ from the classification results and corresponding true labels Y_S^{tr}. The features of the source domain and target domain were then aligned using the domain alignment algorithm.

(c) **Model Selection**. The source domain training dataset X_S^{tr}, source domain test dataset X_S^{te}, target domain training dataset X_T^{tr}, and the risk evaluation function were used to compute the risk values of different candidate models to obtain the optimal model. The optimal model was used to predict the test results for the target domain test dataset X_T^{te}.

(d) **Crop Map**. Time-series spectral reflectance of the target domain was sequentially extracted for each pixel point on the satellite images. After calculating the time series VIs, the target domain dataset that can be fed into the deep learning network was obtained through the S-G filter. The crop distribution of study area was mapped, utilizing the optimal model to predict the target domain dataset.

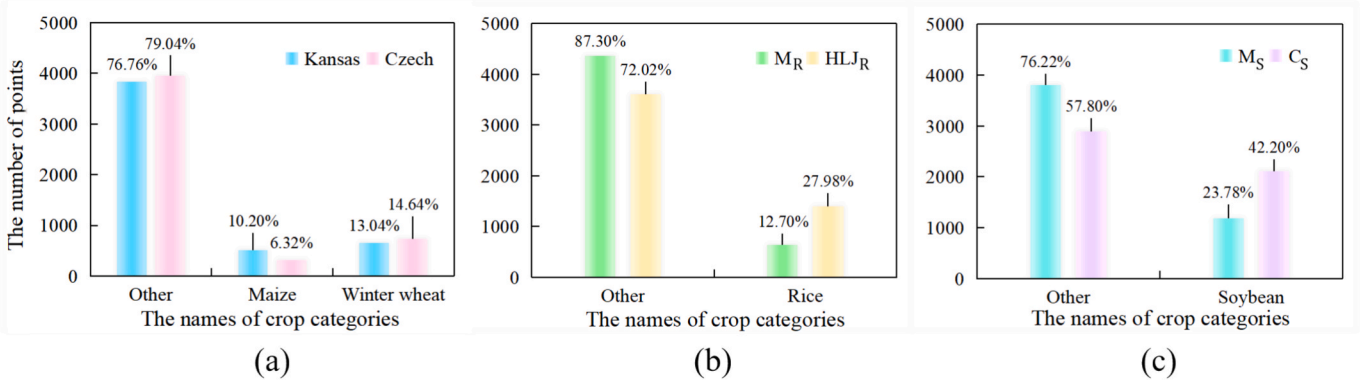

Fig. 3. The number and proportions of points for different categories in the training dataset.

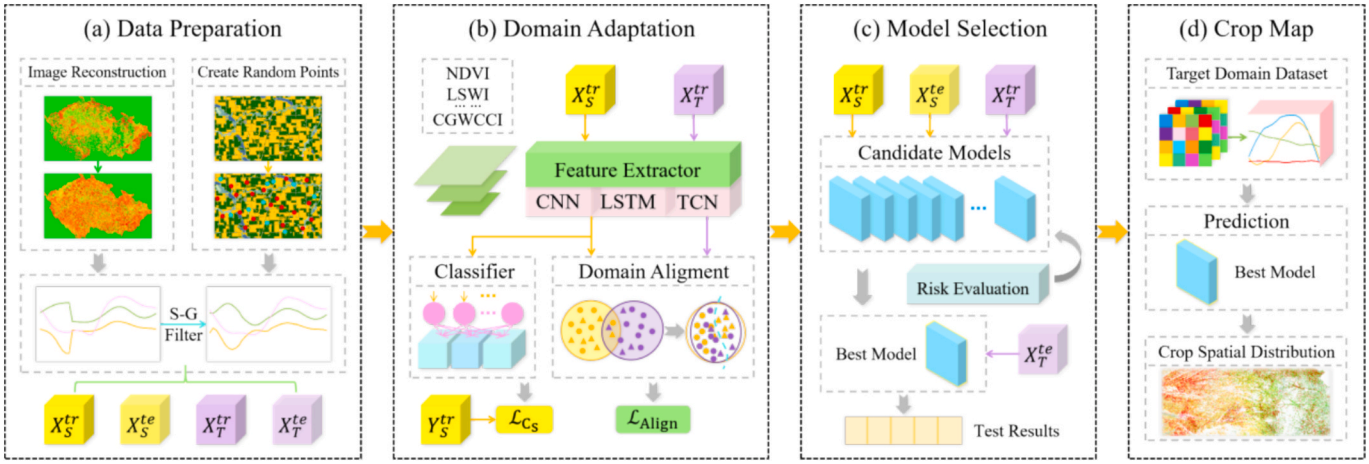

Fig. 4. The workflow of this study. (a) Data Preparation, (b) Domain Adaptation, (c) Model Selection and (d) Crop Map.

3.2. MCDA model

The structure of domain adaptation model with multiple classifiers is shown in Fig. 5, which mainly consists of two parts, namely the feature extractor and classifiers. As a core component of the model, the structural design of the feature extractor directly affects learning effectiveness and classification performance. It consists of three convolutional blocks, each including a one-dimensional convolutional layer, batch normalization, a ReLU activation function, and a max-pooling operation with a stride of 2. In Fig. 5, the yellow module indicates the source domain dataset and the purple module indicates the target domain dataset. The green module represents the feature extractor Γ, which consists of three convolutional blocks of different sizes. The cyan module is the source domain classifier C_S, and the pink and blue modules represent two different classifiers C_1 and C_2. Each classifier is implemented as a linear layer, which maps the extracted feature to the corresponding class. The feature extractor maps data from different domains to a new feature space, while the classifier performs

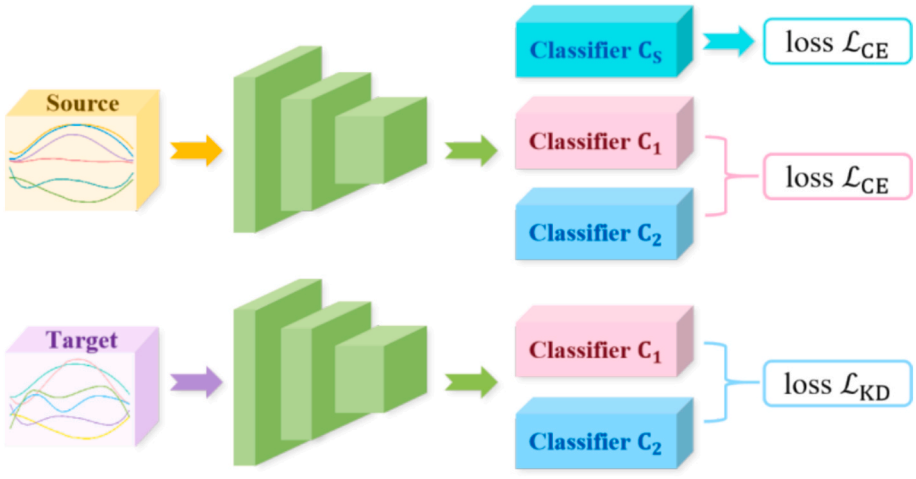

Fig. 5. The structure of domain adaptation model with multiple classifiers.

classification on the extracted features. The parameters of the model are updated and trained according to the different losses.

The cross entropy loss $\mathscr{L}_{CE}$ (Mao et al., 2023) is calculated based on the classification results and labels of source domain. The formula is as follows:

$$\mathscr{L}_{CE}(X_S, Y_S) = -\mathbb{E}_{(\boldsymbol{x_S}, \boldsymbol{y_S})\sim(X_S, Y_S)} \sum_{k=1}^{K} \mathbb{I}_{[k=y_S]} \log p(f(\boldsymbol{x_S})) \tag{3-1}$$

where k denotes different classes, p denotes the classifier, f denotes the feature extractor. The $\mathbb{I}$ is the indicator function, which is set to be 1 when the condition is satisfied and set to 0 otherwise. The knowledge distillation loss $\mathscr{L}_{KD}$ (Hou et al., 2021) in the target domain is calculated based on the results from different classifiers. The formula is as follows:

$$\mathscr{L}_{KD}(X_T) = \mathscr{D}_{KL}(p_1(f(\boldsymbol{x_T})), p_2(f(\boldsymbol{x_T}))) \tag{3-2}$$

where $\mathscr{D}_{KL}$ denotes the Kullback-Leibler (KL) divergence. And in this study, the KL divergence is employed to measure distributional discrepancies between the outputs of two classifiers C_1 and C_2 on the target domain. The greater the distributional discrepancies, the larger the KL divergence. When the distributions are identical, the KL divergence is zero. Compared to the simple Euclidean distance, the KL divergence can more accurately describe the inconsistency between the predictive distributions of two classifiers C_1 and C_2.

3.2.1. Domain adaptation

The MCDA model with three training steps belongs to an adversarial-based method. In the first step, the feature extractor Γ and source domain classifier C_S are trained, as shown in Fig. 6 (a). This step aims to learn the feature representations from the source domain by an independent classifier, providing a stable initial feature space for subsequent target alignment. The features from source domain training dataset X_S^{tr} are extracted using the feature extractor Γ, and these features are then classified by the source domain classifier C_S. The source domain classification loss $\mathscr{L}_{C_S}$ is calculated using cross entropy loss $\mathscr{L}_{CE}$, based on the classification results and corresponding true labels Y_S^{tr}. The optimization objective is as follows:

$$\min_{\Gamma, C_S} \mathscr{L}_{C_S}(X_s, Y_s) \tag{3-3}$$

The feature extractor Γ and source domain classifier C_S are trained by continuously updating the source domain classification loss $\mathscr{L}_{C_S}$. Fig. 6 (b) illustrates the final training results for the source domain in the first step. The yellow indicates the source domain data, the solid and dashed lines indicate the different categories, and the cyan line indicates the decision boundary of the source domain classifier C_S. After minimizing the source domain classification loss $\mathscr{L}_{C_S}$, the optimal feature extractor Γ for the source domain dataset and the source domain classifier C_S are obtained.

In the second step, the parameters of the feature extractor Γ is fixed and the classifiers C_1 and C_2 are trained to maximize the discrepancy in predictions for the target domain. When the outputs of the classifiers C_1 and C_2 are consistent, it indicates that the prediction results of features from the target domain are relatively accurate. On the contrary, large discrepancies between the outputs of the classifiers C_1 and C_2 suggest that the feature from the target domain has not yet been aligned with the source domain. This step helps to expose the target features that remain misaligned with the source domain, particularly those lying near the decision boundaries of the classifiers. Fig. 7 (a) illustrates the training process of the second step. The features from the source domain training dataset X_S^{tr} are extracted using the feature extractor Γ, followed by the classifiers C_1 and C_2 to get different classification results of the source domain training dataset. The two source domain cross entropy losses $\mathscr{L}_{C_1}$ and $\mathscr{L}_{C_2}$ are calculated by different classification results and labels Y_S^{tr} corresponding to source domain. Similarly, the features of the target domain dataset X_T^{tr} are extracted using the feature extractor Γ, and the different predictions of the target domain are obtained after two classifiers C_1 and C_2. The knowledge distillation loss $\mathscr{L}_{KD}$ is calculated using the two different predictions in target domain. The optimization objective is as follows:

$$\min_{C_1, C_2} \mathscr{L}_{C_1}(X_S, Y_S) + \mathscr{L}_{C_2}(X_S, Y_S) - \mathscr{L}_{KD}(X_T) \tag{3-4}$$

By training the classifiers C_1 and C_2, the optimal decision boundaries that can perfectly distinguish the different categories in the source domain are obtained. At the same time, the decision boundaries are maximized for the discrepant samples that are within the target domain but outside the support of the source domain, as shown in Fig. 7 (b). The purple indicates the target domain data, and the pink and blue lines indicate the decision boundary of the classifiers C_1 and C_2.

In the third step, the parameters of the classifiers C_1 and C_2 are fixed and the feature extractor Γ is trained to minimize the discrepancy on target samples, thereby pushing the features towards a better domain-invariant space and achieving domain adaptation. The features from the target domain dataset X_T^{tr} are extracted again using the feature extractor Γ, and the different predictions of the target domain are obtained after two classifiers C_1 and C_2. Then the knowledge distillation loss $\mathscr{L}_{KD}$ of the two predictions is calculated in target domain, as shown in Fig. 8 (a). The optimization objective is as follows:

$$\min_{\Gamma} \mathscr{L}_{KD}(X_T) \tag{3-5}$$

On the basis of obtaining the optimal decision boundary of the classifiers C_1 and C_2, the divergence of different predictions in the target domain is minimised by training the feature extractor Γ. So then, the features from the target domain are constantly close to the features from the source domain, as shown in Fig. 8 (b). After continuous updating and training of the domain adaptation model, the feature extractor Γ maps the source domain and target domain data into a better feature space, and classifiers C_S, C_1 and C_2 with optimal decision boundaries that can distinguish between different crop classes are obtained.

Unlike conventional discrepancy-based methods, the MCDA model constructs an adversarial learning mechanism through the feature extractor Γ and two classifiers C_1 and C_2 in the target domain to achieve the domain adaptation. By adopting multiple classifiers for adversarial learning instead of a domain discriminator, the MCDA model guides the

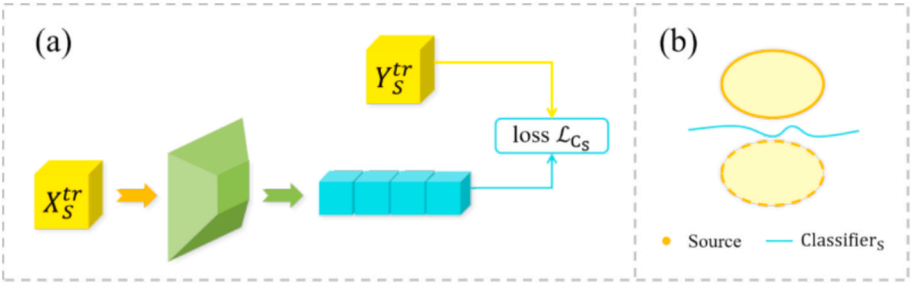

Fig. 6. The training process and results of the first step.

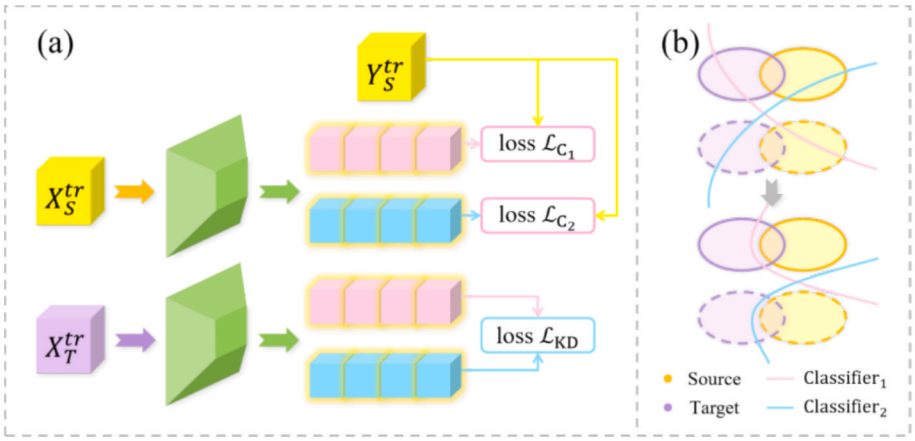

Fig. 7. The training process and results of the second step.

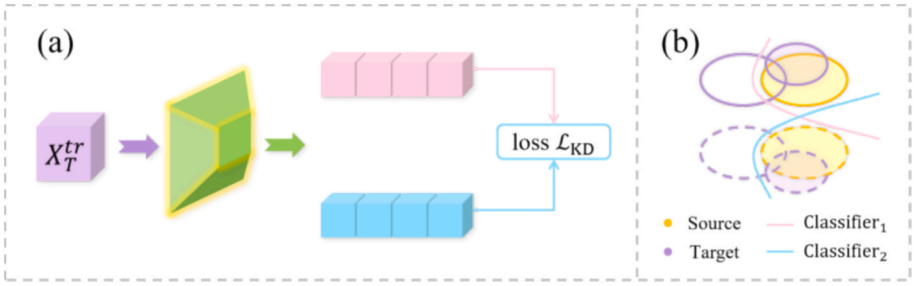

Fig. 8. The training process and results of the third step.

feature extractor to map target samples into the support space of the source domain, thereby learning more robust domain-invariant features. In the second step, two classifiers C_1 and C_2 are trained to maximize discrepancies between predictions on the target domain, thereby exposing target features that are misaligned with the source domain. In the third step, the feature extractor Γ is trained to minimize discrepancies between predictions of two classifiers C_1 and C_2 on the target domain, enabling it to learn more robust domain-invariant feature representations for different domains. Through this alternating "maximize-minimize" optimization process, the MCDA model establishes a game mechanism similar to that of adversarial networks, effectively realizing domain adaptation based on the adversarial learning.

3.2.2. Model selection

Model selection and hyper-parameter tuning are long-standing problems in the tasks of UDA due to the absence of labels from target domain to supervise the training process. Thus, the Deep Embedded Validation (DEV) risk (You et al., 2019) was used to select the optimal hyper-parameters of MCDA model. Without labels from target domain, the DEV risk evaluates the performance of the model by the features in source and target domains. To begin with, the logistic regression model r_θ is trained to distinguish whether the training features come from the source or target domain. The optimization objective is as follows:

$$\min_{r_\theta} \mathscr{L}_r = \left[log r_\theta(Z_S^{tr}) + log(1 - r_\theta(Z_T^{tr}))\right] \tag{3-6}$$

where Z_S^{tr} and Z_T^{tr} represent the training features in the source and target domains. Next, the trained logistic regression model r_θ is used to compute the importance weights w_f of the source domain test dataset, which is written as

$$w_f(X_S^{te}) = \frac{N_S^{tr}}{N_T^{tr}} \frac{1 - r_\theta(Z_S^{te})}{r_\theta(Z_S^{te})} \tag{3-7}$$

where N_S^{tr}/N_T^{tr} represents the ratio of sample sizes in the source domain and target domain training dataset, and Z_S^{te} represents the features from the source domain test dataset. Then, the weighted cross entropy loss $\mathscr{L}_S^{te}$ is computed through the importance weights w_f from source domain test dataset, which is written as

$$\mathscr{L}_S^{te} = \left\{w_f(X_S^{te}) \mathscr{L}_{CE}\left(m(X_S^{te}), Y_S^{te}\right)\right\}_i^{N_S^{te}} \tag{3-8}$$

where m is a candidate model, and N_S^{te} is the sample size of source domain test dataset. In the end, the DEV risk $\mathscr{R}_{DEV}$ is computed based on the importance weights w_f and weighted cross entropy loss $\mathscr{L}_S^{te}$ in the source domain test dataset, which is written as

$$\mathscr{R}_{DEV} = \text{mean}\left(\mathscr{L}_S^{te}\right) + \eta \text{mean}\left(w_f\left(X_S^{te}\right)\right) - \eta \tag{3-9}$$

where $\eta = -\text{Cov}\left(\mathscr{L}_S^{te}\right)/\text{Var}\left(w_f\left(X_S^{te}\right)\right)$ is the optimal coefficient.

3.3. Evaluation Metrics

The confusion matrix was utilized to represent the data distribution of all the predicted responses, where the rows represented actual classes and the columns represented predicted classes. The overall accuracy (OA) and weighted-averaged F1 score ($F1_w$) were employed to evaluate the performance of models. The OA is the ratio of the total number of correctly predictions to the total number of samples, expressed as:

$$OA = \frac{1}{N} \sum_{k=1}^{K} c_{k,k} \tag{3-10}$$

where N is the number of samples, K is the number of classes, $c_{i,j}$ denotes the element in the i th row and j th column of the confusion matrix. Obviously, the OA reflects the performance of models as a whole. The $F1_w$ assigns the weights for different classes depending on the number of samples in each class, expressed as:

$$F1_w = 2\frac{P_w R_w}{P_w + R_w} \tag{3-11}$$

where P_w and R_w represent individually the weighted-averaged precision and recall. The P_w and R_w are calculated as follows:

$$P_w = \frac{1}{N}\sum_{k=1}^{K} N_{r,k} P_k \tag{3-12}$$

$$R_w = \frac{1}{N}\sum_{k=1}^{K} N_{r,k} R_k \tag{3-13}$$

where $N_{r,k}$ denotes the actual number of k th class, P_k denotes precision of k th class, R_k denotes recall of k th class. The P_k and R_k are calculated as follows:

$$P_k = \frac{1}{N_{c,k}} c_{k,k} \tag{3-14}$$

$$R_k = \frac{1}{N_{r,k}} c_{k,k} \tag{3-15}$$

where $N_{c,k}$ denotes the number of predictions for k th class. To sum up, the $F1_w$ is suitable for datasets with imbalanced class proportions.

3.4. Experimental Setup

All experiments in this study were carried out on the Intel (R) Xeon (R) Gold 5218R CPU @ 2.10 GHz processor and NVIDIA GeForce RTX 3090 video card. There are four datasets in the source and target domains, each with 5000 samples respectively. For the 1DCNN and LSTM models, target domain data was not used to supervise the training process. The models were trained using the source domain training dataset, and the source domain test dataset was utilized to fine-tune the hyper-parameters. The performance of the models was then evaluated by the target domain test dataset. The UDA models were trained using the source domain and the target domain training dataset. In the absence of labels in target domain, the source domain training dataset, source domain test dataset, and target domain training dataset were employed to select the optimal hyper-parameters. In the same way, the performance of UDA models was then evaluated using the target domain test dataset. The adaptive moment estimation (Adam) was selected as the optimizer of all models. The rectified linear unit (ReLU) was selected as the activation function for the convolutional layers.

3.4.1. Parameter settings

The 1DCNN with three convolutional blocks was used as the feature extractor for the MCDA model in this study. The kernel size determines the receptive field of each convolutional layer, and the number of filters in convolutional layers influences the dimensionality and expressive capacity of the extracted temporal features. Considering the heterogeneity of data characteristics across study areas, the kernel sizes and numbers of filters were selected through experiments. Based on experience and empirical rules, the convolution kernel size was typically set to 3 or 5. The number of filters in the first convolutional layer was set to 32 or 64, with the second layer containing twice as many filters as the first. The number of filters in the third convolutional layer was flexibly chosen from 32, 64, or 128 to ensure that the network captures rich and effective multi-scale features across diverse datasets. The specific parameter configurations of the feature extractor for different source and target domains are detailed in Table 2.

The hyper-parameters of the MCDA model, including the epoch, batch size, learning rate and weight decay of the optimizer, adjustment schedule and decay factor for the learning rate, loss weight from the source domain, and loss weight from the target domain, were fine-tuned using the DEV risk. Based on a suite of experience and rules, different optional values were set for the hyper-parameters. The values for the epoch and batch size were 40 and 32, respectively. The optional values for learning rate were $[10^{-2}, 5\times10^{-3}, 10^{-3}, 5\times10^{-4}]$, the optional values for weight decay were $[10^{-4}, 10^{-5}, 10^{-6}]$, the optional values for step size in the StepLR scheduler were $[5, 10, 20]$, and the optional values for decay factor of learning rate were the same as those for learning rate. The range of loss weight from the source domain was set to $[10^{-1}, 10]$, and the range of loss weight from the target domain was set to $[10^{-2}, 10]$. For each experiment in different domains, 50 sets of hyper-parameters were randomly generated to train the UDA models. The three sets of hyper-parameters with the smallest values of DEV risk were selected to train the optimal models. And the hyper-parameters corresponding to the optimal models were then saved. Additionally, the random seed 42 was selected for all experiments to facilitate the debugging and validation of models, as well as the reproduction of results. The hyper-parameters of optimal models from different study areas are summarized in Table 3.

Table 2
The specific parameter configurations of the feature extractor for different source and target domains.

Source Domain	Target Domain	Kernel Size	Number of Filters (First Layer)	Number of Filters (Last Layer)
Kansas	Czech	3	32	128
Czech	Kansas	3	32	32
M_R	HLJ_R	5	64	64
HLJ_R	M_R	3	32	32
M_S	S_C	3	64	128
S_C	M_S	3	32	128

3.4.2. Comparative experiments

Feature extractors play a vital role for UDA models, so comparative experiments were conducted using 1DCNN, LSTM, and Transformer-based structures as feature extractors to explore their impacts on the performance of the MCDA model. The detailed parameter settings for the 1DCNN-based feature extractor are described above. For the LSTM-based feature extractor, the hidden state was set to 32 or 64, and the number of recurrent layers was varied from 1 to 5 to explore the optimal performance of the MCDA model. The Transformer-based feature extractor consists of a one-dimensional convolutional layer with a kernel size of 1 and the transformer encoder layers. Among these parameters, the number of filters in the convolutional layer was set to 32 or 64. The number of attention heads was set to 2, 4 or 8, the dimension of the feedforward network was set to 64 or 128, and the dropout value as set to 0.2 or 0.4.

In addition to the models proposed in this study, the 1DCNN, LSTM, DANN, MMDA, AdvSKM, and CoTMix models were used to compare with the performance of MCDA model. The 1DCNN model is a typical deep learning method for processing one-dimensional sequential data, which captures the local features of temporal data through convolution computations. And it is widely used as a baseline model in time-series related tasks. The LSTM model is a type of Recurrent Neural Network (RNN) that addresses the issues of gradient vanishing and gradient explosion in traditional RNNs, thereby enabling more effective processing of long-sequence data. For the DANN (Ganin et al., 2016) model, the Gradient Reversal Layer (GRL) is utilized for adversarial training between the domain classifier and feature extractor. The feature extractor is trained to maximize the accuracy of label prediction and minimize the accuracy from the domain classifier. While the optimization objective of the domain classifier is to maximize the accuracy of the domain classification. The second order statistics and Maximum Mean Discrepancy (MMD) are employed to align features from source and target domains in the MMDA (Rahman et al., 2020) model. The AdvSKM (Liu et al., 2021) model is a metric-based domain matching approach that

Table 3
The hyper-parameters of optimal models from different study areas.

	Source Domain	Target Domain	Source Domain	Target Domain	Source Domain	Target Domain
	Kansas	Czech	M_R	HLJ_R	M_S	S_C
Learning rate	10^{-2}		10^{-3}		10^{-2}	
Weight decay	10^{-4}		10^{-6}		10^{-6}	
Step size	5		5		20	
Decay factor of learning rate	5×10^{-4}		10^{-2}		5×10^{-4}	
Source Loss Wt	4.64		5.59		9.86	
Domain Loss Wt	2.23		2.60		8.88	
	Source Domain	**Target Domain**	**Source Domain**	**Target Domain**	**Source Domain**	**Target Domain**
	Czech	Kansas	HLJ_R	M_R	S_C	M_S
Learning rate	5×10^{-3}		5×10^{-4}		5×10^{-4}	
Weight decay	10^{-4}		10^{-5}		10^{-6}	
Step size	5		20		20	
Decay factor of learning rate	10^{-3}		10^{-3}		5×10^{-4}	
Source Loss Wt	4.97		3.76		9.58	
Domain Loss Wt	9.93		0.20		9.74	

introduces the hybrid spectral kernel network as an inner kernel to reformulate the MMD metric. The CoTMix (Eldele et al., 2023) model generates two intermediate augmented views for the source and target domains via a cross-domain temporal mixup strategy. And it employs the contrastive learning to maximize the similarity between each domain and its corresponding augmented view, thereby progressively mitigating the distribution shift across domains.

4. Results

4.1. Training performance

The training performance of the MCDA model is illustrated in Fig. 9 with Czech, and study areas HLJ_R and S_C as target domains. The yellow, blue, and purple curves represent the experimental results when Czech, study area HLJ_R, and study area S_C are used as target domains, respectively. Fig. 9 (a) to (c) depict the variation curves of the training and test accuracy from the source domain, represented by solid and dashed lines, respectively. Fig. 9 (d) to (f) display the variation curves of the loss functions across the three training steps, corresponding to different target domains. As shown in Fig. 9, the accuracy from the source domain increases rapidly with training iterations and gradually stabilizes. And the training accuracy is consistently higher than the test accuracy. Meanwhile, the losses in all three training steps decrease progressively during the training processes, indicating that the models are well trained and gradually converged.

4.2. Backbones Comparison

The performance of the MCDA model with different feature extractors, such as 1DCNN, LSTM, and Transformer-based structures, was evaluated in three sets of experiments. The OA and $F1_w$ were used to assess classification performance, while the training time was employed to measure computational efficiency. Table 4 presents the experimental results of the MCDA model with different backbone networks as feature extractors between Kansas and Czech. In the experiments using Kansas as the source domain and Czech as the target domain, the OA and $F1_w$ of MCDA model with three feature extractors achieved over 80%. In contrast, when Czech was used as the source domain and Kansas was used as the target domain, the MCDA model based on the Transformer as

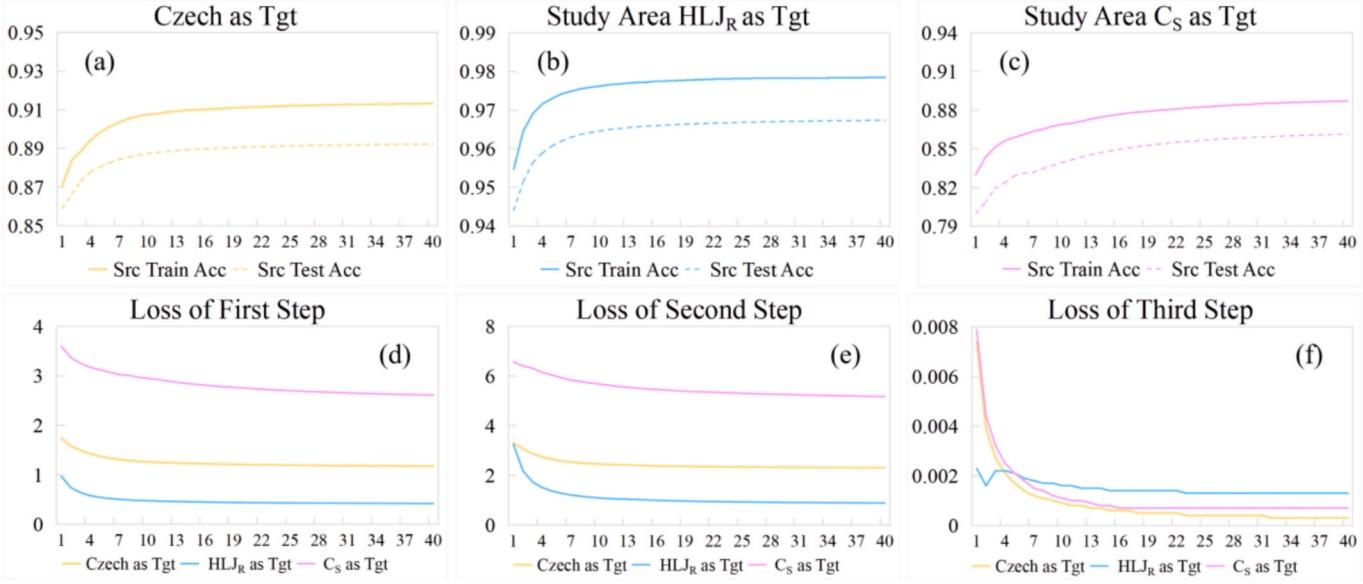

Fig. 9. The training performance of the MCDA model with Czech, and study area HLJ_R, study area S_C as target domains, represented by yellow, blue, and purple. (a) to (c) show the training and test accuracy from the source domain, represented by solid and dashed lines. (d) to (f) show the loss functions for three training steps in different target domains.

Table 4
The experimental results of the MCDA model with different backbone networks in Kansas and Czech.

	Source Domain		Target Domain	Source Domain		Target Domain
	Kansas		Czech	Czech		Kansas
Backbones Networks	**OA**	**$F1_w$**	**Training Time (s)**	**OA**	**$F1_w$**	**Training Time (s)**
1DCNN	86.36%	86.42%	58.61	85.64%	85.03%	56.36
LSTM	84.88%	84.94%	56.19	79.48%	75.92%	57.35
Transformer	84.54%	85.25%	65.83	75.56%	72.70%	65.54

the feature extractor showed the lowest performance. And in Table 4, the training time of the MCDA models using 1DCNN and LSTM as feature extractors were quite comparable.

The OA, $F1_w$, and training time of the MCDA model with 1DCNN, LSTM, and Transformer-based structures as feature extractor are reported in Table 5, between the study areas M_R and HLJ_R. The MCDA model with the LSTM as the feature extractor exhibited considerably lower performance in study area M_R as the source domain and study area HLJ_R as the target domain. Similarly, the MCDA model based on Transformer as the feature extractor showed inferior performance and required the longest training time.

Table 6 presents the experimental results of the MCDA model, where different backbone networks serve as feature extractors between study areas M_S and S_C. The performance with LSTM and Transformer-based structures was worse to that of the 1DCNN-based model. In the experiments with study area S_C as the source domain and study area M_S as the target domain, the OA and $F1_w$ of the MCDA model with different backbone networks reached above 80%. In the comparative experiments with different backbone networks, the MCDA model using 1DCNN as the feature extractor outperformed those based on LSTM and Transformer. Moreover, the MCDA model with the Transformer-based feature extractor required a longer training time.

4.3. Accuracy Comparison

The experimental results of different models, in which Kansas and Czech serve as the source and target domains for each other, are shown in Table 7. Specifically, in the experimental results with Kansas as the source domain and Czech as the target domain, the OA of LSTM model was the lowest, while the accuracy of all other models achieved more than 80%. The OA of MMDA model was close to that of MCDA model in the experimental results with Czech as the source domain and Kansas as the target domain, whereas the rest of models did not perform well in the experiments, especially the LSTM and AdvSKM models. From Table 7, it can be seen that whether Kansas or Czech is used as the target domain, the MCDA model achieves the best OA and $F1_w$. Whether Kansas or Czech was used as the source domain, the 1DCNN model achieved the shortest training time. Among the UDA methods, the DANN model had the shortest training time, while the training time of the other UDA models were very similar.

The experimental results for study areas M_R and HLJ_R, serving as the source and target domains for each other, are shown in Table 8. In the experimental results with study area M_R as the source domain and study area HLJ_R as the target domain, the OA and $F1_w$ of LSTM model were still the lowest, at 77.26% and 78.50% respectively. The accuracy of other models ranged between 80% and 90%. In particular, the MCDA model had the highest OA and $F1_w$ with 92.50% and 92.57%, respectively. However, in the experimental results with study area HLJ_R as the source domain and study area M_R as the target domain, although the MCDA model performed better, the DANN model achieved the highest OA and $F1_w$. And in the experimental results with study area M_R as the target domain, the OA and $F1_w$ of MCDA model were close to those of the CoTMix model. Additionally, the 1DCNN model still maintained the shortest training time in the experiments between study areas M_R and HLJ_R. The CoTMix model had the longest training time, nearly twice that of the 1DCNN model.

The experimental results for study areas M_S and S_C, where each serves as both the source and target domains, are shown in Table 9. In the experimental results with study area M_S as the source domain and study area S_C as the target domain, the accuracy of other models ranged between 70% and 76%. However, the OA and $F1_w$ of MCDA model were separately 81.90% and 81.93%, which was over 6% higher than the accuracy of other models. In the experimental results where study area S_C was the source domain and study area M_S was the target domain, the OA of all models exceeded 80%. Especially, the OA and $F1_w$ of the MCDA model were 82.34% and 82.40%, respectively. The training time of different models in the experiments between study areas M_S and S_C was similar to that in other areas. And the training time of the DANN model was close to that of the LSTM model.

As shown in Fig. 10, the confusion matrices for different models in different target domains were plotted to further demonstrate the performance of the models in more detail. Fig. 10 (a) to (g) show the confusion matrices of different models for maize, winter wheat, and other classes in Czech. Fig. 10 (h) to (n) display the confusion matrices of different models for rice and other classes in study area HLJ_R. Fig. 10 (o) to (u) present the confusion matrices of different models for soybeans and other classes in study area S_C. Each column of the confusion matrix represents the predicted class, and the sum of each column indicates the total number of predictions for that class. Each row represents the actual class and the sum of each row represents the real number of that class. From the confusion matrix, various metrics of models such as accuracy, precision, recall and others in the target domain test dataset can be calculated. The precision for each class can be calculated based on the columns of the confusion matrix. In the target domain test results for Czech, the precision of the AdvSKM model for maize is 0, but the precision for winter wheat is the highest. In the target domain test results for study area HLJ_R, the precision for rice of the MCDA model is the highest, at 84.28%. Similarly, in study area S_C, the precision for soybeans of the MCDA model is the highest, at 77.52%. The recall for each class can be calculated from the rows of the confusion matrix. Even if the precision for the studied crops can reach the maximum in all experiments, the recall for each class is not necessarily the best. Therefore, the

Table 5
The experimental results of the MCDA model with different backbone networks in study areas M_R and HLJ_R.

	Source Domain		Target Domain	Source Domain		Target Domain
	M_R		HLJ_R	HLJ_R		M_R
Backbones Networks	**OA**	**$F1_w$**	**Training Time (s)**	**OA**	**$F1_w$**	**Training Time (s)**
1DCNN	92.50%	92.57%	58.33	92.10%	91.38%	57.60
LSTM	80.08%	80.99%	56.10	87.00%	82.35%	63.80
Transformer	83.34%	84.09%	64.31	86.74%	81.13%	65.37

Table 6
The experimental results of the MCDA model with different backbone networks in study areas M_S and S_C.

	Source Domain		Target Domain	Source Domain		Target Domain
	M_S		C_S	C_S		M_S
Backbones Networks	**OA**	**F1$_w$**	**Training Time (s)**	**OA**	**F1$_w$**	**Training Time (s)**
1DCNN	81.90%	81.93%	57.48	82.34%	82.40%	57.95
LSTM	72.00%	72.14%	58.16	81.62%	81.50%	54.95
Transformer	73.70%	73.88%	65.94	82.10%	81.90%	65.20

Table 7
The experimental results of different models in Kansas and Czech.

	Source Domain		Target Domain	Source Domain		Target Domain
	Kansas		Czech	Czech		Kansas
Evaluation Metrics	**OA**	**F1$_w$**	**Training Time (s)**	**OA**	**F1$_w$**	**Training Time (s)**
1DCNN	83.60%	84.00%	34.21	77.00%	77.23%	34.74
LSTM	79.46%	80.14%	45.31	63.48%	68.24%	45.79
DANN	82.86%	83.42%	50.56	81.26%	76.82%	47.01
MMDA	84.14%	84.81%	59.72	85.28%	83.97%	53.12
AdvSKM	84.94%	81.47%	56.23	75.76%	65.31%	54.06
CoTMix	82.28%	81.98%	57.96	79.68%	79.18%	58.72
MCDA	**86.36%**	**86.42%**	58.61	**85.64%**	**85.03%**	56.36

Table 8
The experimental results of different models in study areas M_R and HLJ_R.

	Source Domain		Target Domain	Source Domain		Target Domain
	M_R		HLJ_R	HLJ_R		M_R
Evaluation Metrics	**OA**	**F1$_w$**	**Training Time (s)**	**OA**	**F1$_w$**	**Training Time (s)**
1DCNN	82.44%	82.84%	33.63	87.78%	83.92%	33.19
LSTM	77.26%	78.50%	45.10	86.26%	81.20%	45.52
DANN	82.54%	82.14%	49.65	**93.56%**	**93.25%**	50.15
MMDA	86.66%	86.34%	55.48	90.34%	88.47%	55.79
AdvSKM	89.40%	89.78%	53.38	86.90%	81.70%	60.19
CoTMix	85.38%	86.09%	59.08	92.58%	91.38%	60.84
MCDA	**92.50%**	**92.57%**	58.33	92.10%	91.38%	57.60

Table 9
The experimental results of different models in study areas M_S and S_C.

	Source Domain		Target Domain	Source Domain		Target Domain
	M_S		S_C	S_C		M_S
Evaluation Metrics	**OA**	**F1$_w$**	**Training Time (s)**	**OA**	**F1$_w$**	**Training Time (s)**
1DCNN	73.98%	74.14%	34.29	81.02%	79.17%	32.83
LSTM	70.42%	70.29%	42.55	80.44%	80.46%	42.22
DANN	75.46%	75.21%	46.58	81.40%	81.23%	47.15
MMDA	75.66%	75.44%	56.00	80.94%	81.29%	59.09
AdvSKM	75.20%	75.00%	54.45	80.20%	79.18%	51.47
CoTMix	75.82%	75.98%	54.11	80.22%	81.24%	53.93
MCDA	**81.90%**	**81.93%**	57.48	**82.34%**	**82.40%**	57.95

performance of models should be evaluated from multiple perspectives based on the specific situation.

4.4. Visual predictions

The regions with large-scale crop cultivation were randomly clipped from the target domains of Czech, and study areas HLJ_R and S_C for detailed visual predictions to showcase the transfer ability of different models, as shown in Fig. 11. And in Fig. 11, yellow indicates maize, green indicates winter wheat, blue indicates rice, purple indicates soybeans, and all black indicate others. Fig. 11 (1) displays the Sentinel-2 imagery in Czech, Fig. 11 (a) shows the corresponding label with the European crop type map as references, and Fig. 11 (b) to (h) depict the visual prediction maps of the different models in Czech. For the labels of fields in Fig. 11 (a), it is evident that some pixels are missing in the annotations of some crop class. In Fig. 11 (g), the prediction results of the CoTMix model almost lose the relevant information of the maize category. For the performance of maize and winter wheat, the proposed approach in this study offered more complete predictions of fields and clearer boundary delineation, especially for winter wheat. Nevertheless, compared to the reference map in Fig. 11 (a), the accuracy of crop predictions needs improvement, as the underestimation of maize occurs in all models.

Similarly, Fig. 11 (2) presents the Sentinel-2 imagery in study area HLJ_R, Fig. 11 (i) shows the corresponding label with the crop type maps in Northeast China as references, and Fig. 11 (j) to (p) depict the visual prediction maps of the different models in study area HLJ_R. In addition to rice, maize and soybeans were also extensively cultivated in Fig. 11

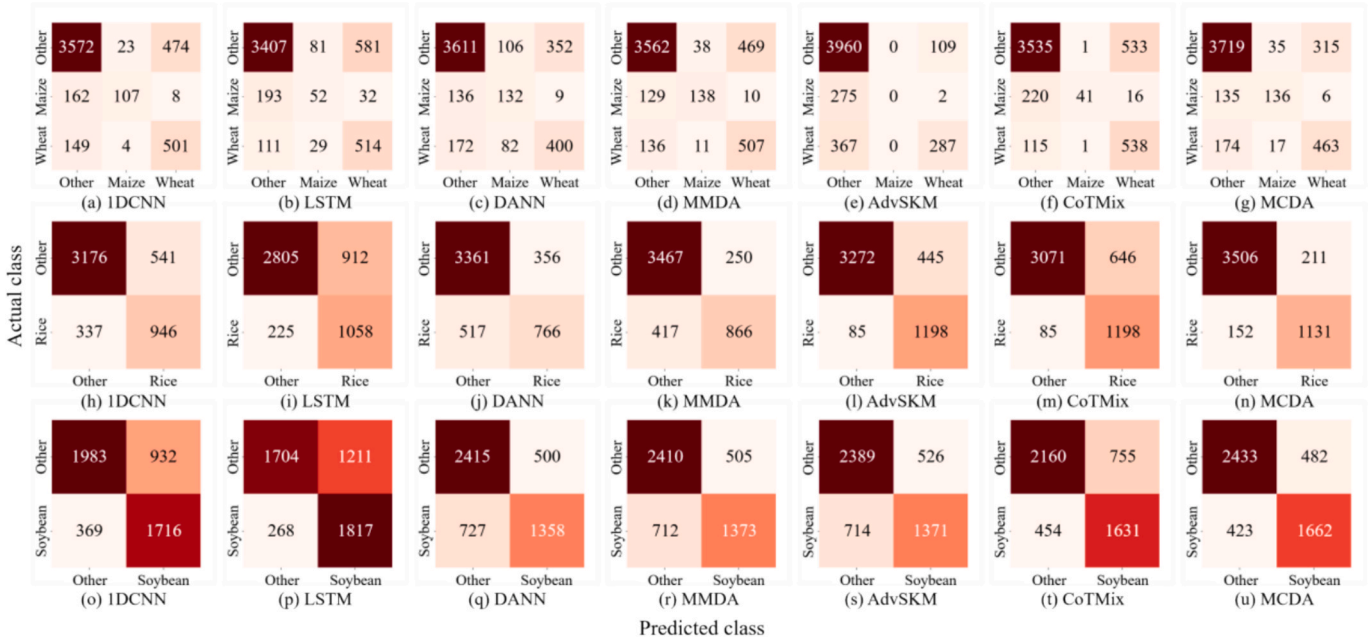

Fig. 10. The confusion matrices of different models in Czech, and study areas HLJ_R and S_C as target domains.

(2). Comparing the label in Fig. 11 (i), the predictions of MCDA model for rice in Fig. 11 (p) were more precise than those of other models. And the boundary predictions of DANN and CoTMix models were clearer than those of other models in Fig. 11 (l) and (o). But the 1DCNN, LSTM, and AdvSKM models mistakenly classified large areas of maize as rice.

In the same way, Fig. 11 (3) shows the Sentinel-2 imagery in study area S_C, Fig. 11 (q) exhibits the corresponding label with the soybean classification maps in South America as references, and Fig. 11 (r) to (x) depict the visual prediction maps of the different models in study area S_C. In the label of Fig. 11 (q), the field boundaries were inaccurately depicted and roads were not distinguished. Compared to the label in Fig. 11 (q), the MCDA model distinguished most of the roads clearly. Moreover, the predictions for soybeans of MCDA model in Fig. 11 (x) are more accurate than those of other models. The 1DCNN model overestimated soybeans while other UDA models underestimated soybeans. In Fig. 11 (s), the LSTM model shows the worst predictive performance.

In order to further illustrate in more detail the visual prediction performance of different models in Czech, and study areas HLJ_R and S_C, the difference maps were created and the number of misclassification points was counted for the corresponding regions of Fig. 11. Among them, Fig. 12 (a) to (g) show the difference maps of models in the corresponding region of Fig. 12 (1), where gray represents correctly classified points, and orange represents misclassified points. It can be clearly seen that there are some misclassified fields in the predictions of different models. Fig. 12 (h) summarizes the number of misclassification points for different models in the corresponding region of Fig. 12 (1). The CoTMix model has the highest number of misclassification points in Fig. 12 (h). The MCDA model has the fewest misclassification points, with a classification accuracy of 75.91%.

Likewise, Fig. 12 (i) to (o) display the difference maps of models in the corresponding region of Fig. 12 (2), where gray represents correctly classified points, and cyan represents the misclassified points. Fig. 12 (p) counts the number of misclassification points for different models. Obviously, the MCDA model successfully predicts almost all rice fields in the difference map of Fig. 12 (o). Corresponding to the region in Fig. 12 (2), the MCDA model has the fewest misclassification points in Fig. 12 (p), with a classification accuracy of 92.12%.

Fig. 12 (q) to (w) show the difference maps of models for the corresponding region of Fig. 12 (3), where gray represents correctly classified points and pink represents misclassified points. It is clear in Fig. 12 (r) that the LSTM model performs the worst, with the highest number of misclassification points. In contrast, Fig. 12 (w) shows that the MCDA model performs the best, with the fewest misclassified fields. Fig. 12 (x) counts the number of misclassification points for each model in the corresponding region of Fig. 12 (3). And the MCDA model has the least number of misclassification points and achieves a classification accuracy of 85.01%.

To further verify the computational efficiency of different models, the reasoning time during the visualization-based prediction processes are summarized in Table 10 with Czech, and study areas HLJ_R and S_C as the target domains. Among the three target domains, the 1DCNN model achieved the fastest reasoning time, followed by the LSTM model. The reasoning time of the different UDA models were very similar. And among the UDA approaches, the CoTMix model exhibited the optimal inference speed.

5. Discussion

5.1. Feature Analysis

The premise of classification is that there are commonalities among the same type of crops, while there are significant differences among different crops. Therefore, the selection of source domains and features is crucial for crop classification. For the same study area, factors such as temperature, precipitation, sunshine, soil and so on, may affect the growth of crops. Furthermore in different study areas, factors such as geographic location, climatic conditions and so on, may also affect the growth of the same type of crop, leading to greater differences in the trends of VIs in different domains. The other temporal profiles of VIs in Fig. 2 illustrate the commonalities of the same type of crop in different regions.

On the one hand, for study areas without reference data as label, the selection of source domains with similar crop cultivation conditions and accurate crop mapping is key for crop transfer and prediction. On the other hand, since study areas often cultivate multiple crops, it is very important to select appropriate data features for reducing interference from other crops. The NDVI was able to differentiate vegetation from other land types. In Kansas and Czech, the NDVI was used to distinguish between maize and winter wheat. However, in addition to maize and winter wheat, other crops such as soybeans, sorghum and others were

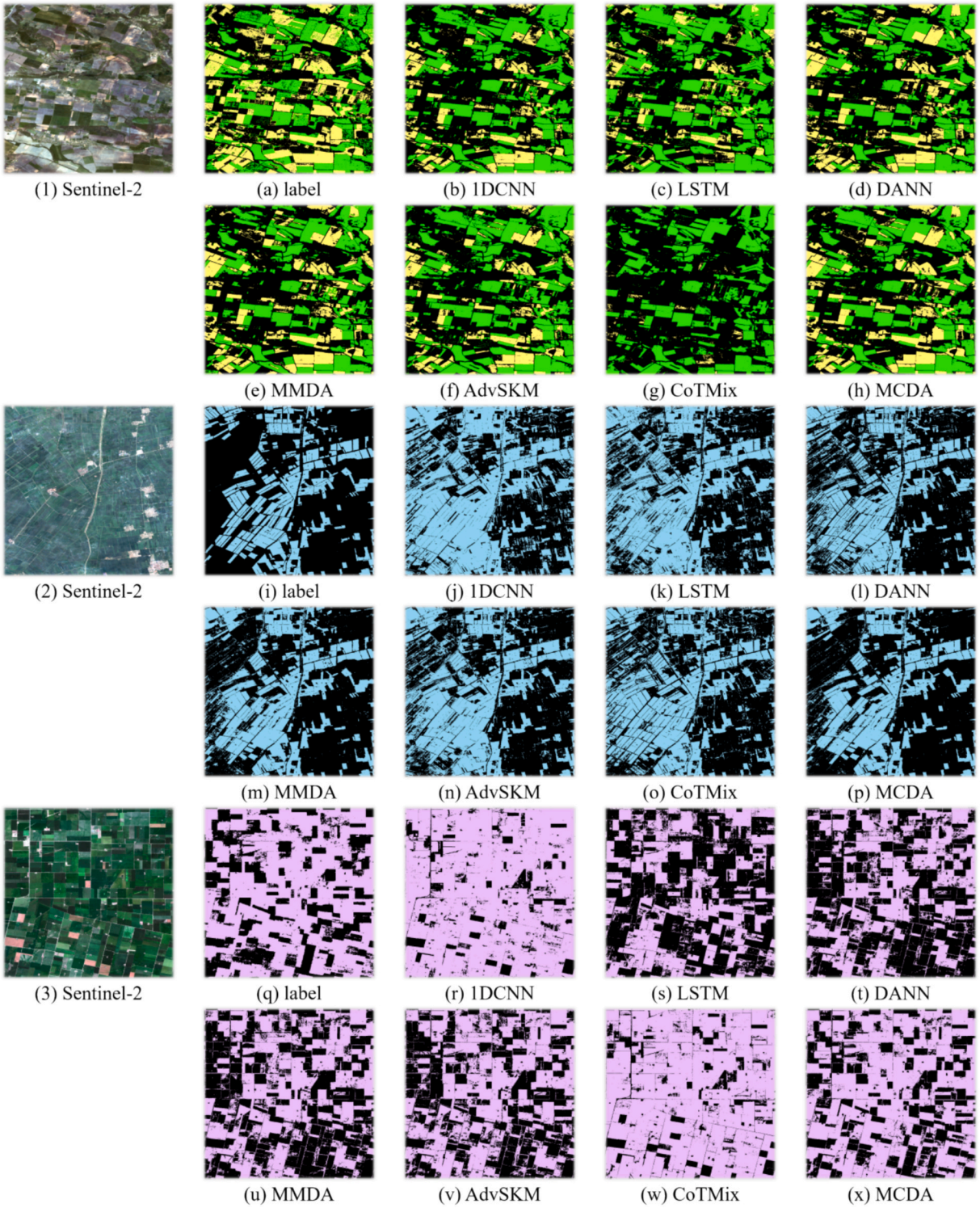

Fig. 11. The Sentinel-2 images, labels, and visual prediction maps of different models in Czech, and study areas HLJ_R and S_C as target domains. Among them, yellow, green, blue, purple, and black indicate maize, winter wheat, rice, soybeans, and other, respectively.

cultivated in Kansas, while there were extensive plantings of winter barley, sugar beets, rapeseed and turnip rapeseed, and so on in Czech. Particularly in Czech, the profiles of NDVI for winter barley, rapeseed and turnip rapeseed were very similar, so the BHDI was used to distinguish winter wheat from other crops. In addition to rice, crops such as cotton, soybean and others were grown in study area M_R, and crops such as maize and others were grown in study area HLJ_R. Since rice is a crop with high water requirement, the water spectral indices such as LSWI, NDWI and MDNWI were used to differentiate rice from other crops. During the peak growing season, the values of NDVI and SWIR1 for soybeans maintain exceptionally high, making it easy to distinguish from other land cover types. So the NDVI and GWCCI were used to distinguish between soybean and other crops in study areas M_S and S_C.

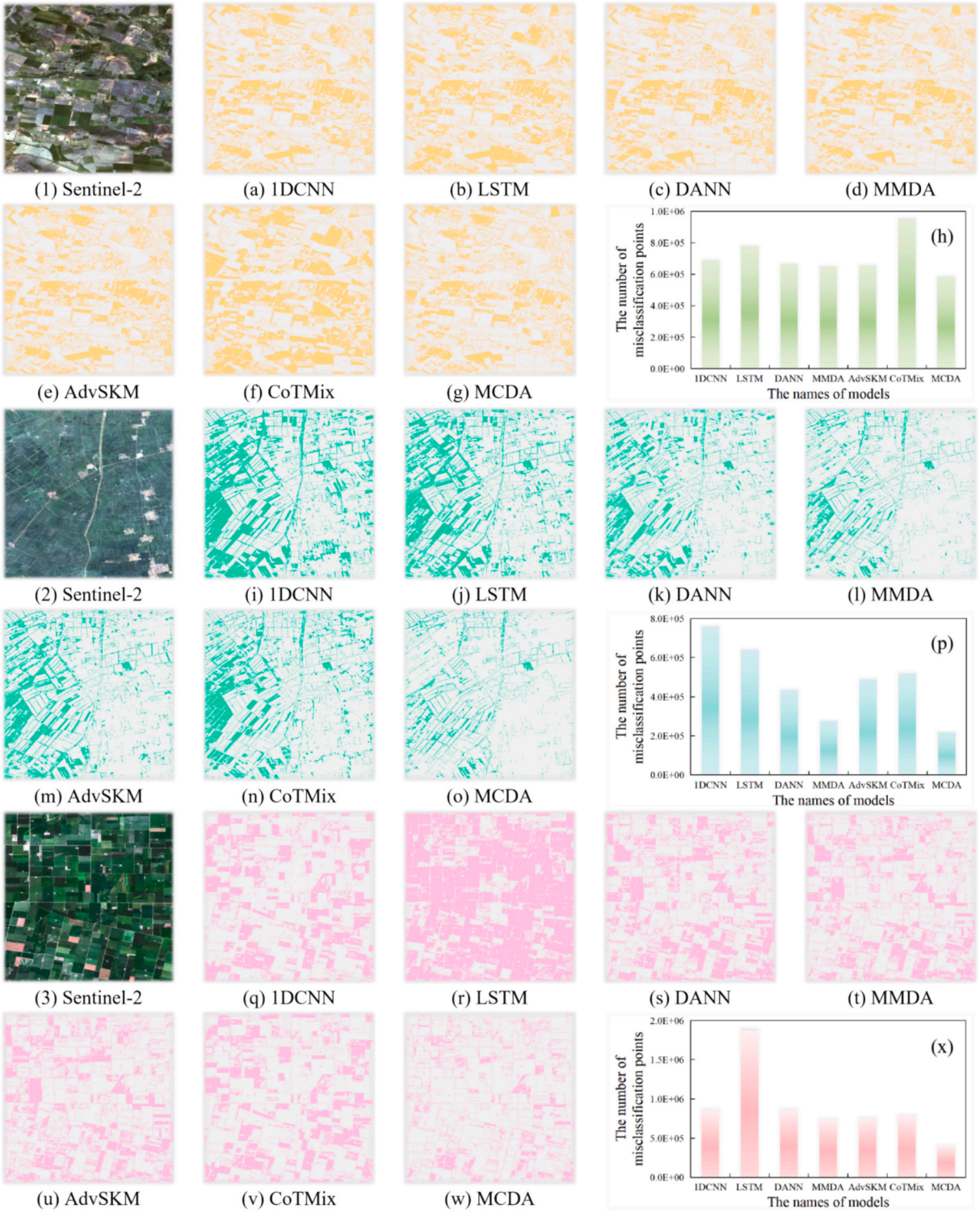

Fig. 12. The difference maps and the number of misclassification points of different models in the corresponding regions of Fig. 11. Among them, gray represents correctly classified points. Orange, cyan, and pink represent misclassified points in Czech, and study areas HLJ_R and S_C, respectively.

5.2. Model Discussion

In the comparative experiments with different backbone networks, the MCDA model using 1DCNN as the feature extractor achieved the best overall performance, outperforming those using LSTM and Transformer-based structures in both the OA and $F1_w$. The training time of the MCDA model with 1DCNN and LSTM as feature extractors were relatively close, whereas the MCDA model based on the Transformer required the longest training time. Notably, due to the relatively high complexity of the Transformer architecture, the MCDA model employing it as the feature extractor tended to overfit easily when some hyper-parameters of the model were not properly configured. In summary, the MCDA model with 1DCNN as the feature extractor can efficiently and accurately extract temporal features, exhibiting superior stability and generalization capability.

In the prediction results of the three sets of experiments, the accuracy

Table 10
The reasoning time of different models during the visualization-based prediction processes with Czech, and study areas HLJ_R and S_C as the target domains.

	Source Domain	Target Domain	Source Domain	Target Domain	Source Domain	Target Domain
	Kansas	Czech	M_R	HLJ_R	M_S	C_S
Evaluation Metrics	**Reasoning Time (s)**		**Reasoning Time (s)**		**Reasoning Time (s)**	
1DCNN	19.03		23.64		22.03	
LSTM	23.97		28.31		27.20	
DANN	32.70		39.96		34.62	
MMDA	32.68		39.45		34.69	
AdvSKM	32.50		38.69		35.02	
CoTMix	29.34		38.31		31.39	
MCDA	32.61		38.51		34.43	

of 1DCNN model is better than that of LSTM model. Accordingly, using the 1DCNN as a feature extractor is more effective than using the LSTM for this purpose. In the experiment with Czech as the target domain for classifying maize and winter wheat, the training accuracy was 91.11%, and the test accuracy was 88.81% in the source domain. In the experiment with study area HLJ_R as the target domain for classifying rice, the training accuracy could reach 97.47%, and the test accuracy could reach 96.40% in the source domain. In the experiment with study area S_C as the target domain for classifying soybeans, the training accuracy was 86.73%, and the test accuracy was 84.71% in the source domain. Based on the above, the data features from source domain were extracted using the 1DCNN and the extracted features were classified employing an independent classifier C_S. It can ensure that the training process of the source domain data is not affected by the target domain data, which is conducive to the feature extractor Γ to better capture the key information of the source domain data.

During the second step, where the parameters of the feature extractor Γ was fixed and the two classifiers C_1 and C_2 were trained, the losses of the optimization objective gradually decreased in the experiments of MCDA model. It indicates that maximizing the knowledge distillation loss $\mathscr{L}_{KD}$ between the outputs of classifiers C_1 and C_2 enables the classifiers to distinguish data that is within the target domain but outside the support of the source domain for each class. In the third step, when the classifiers C_1 and C_2 were fixed and the feature extractor Γ was trained under the optimal parameter settings, the loss functions progressively decreased with the number of iterations. It indicates that the feature extractor Γ is able to map both source domain and target domain data into a more effective feature space.

This study innovatively proposes the use of multiple classifiers in three steps to train the MCDA model, which improves the accuracy of crop classification in the target domain without reference data. From Tables 7 to 9, the accuracy of MCDA model within the same set of experiments shows minimal variation, regardless of which study areas are used as the source domain. The stability and generalization of the model proposed in this study are verified. In the confusion matrix of Fig. 10, the classification results for maize using the AdvSKM model are 0 with Czech as the source domain and Kansas as the target domain. Compared to other UDA models, the AdvSKM model is not suitable for training and prediction of maize and winter wheat in Kansas and Czech. In the predictions of Fig. 11, the method proposed in this study has better prediction performance despite inaccurate boundary and fields in the labels. Hence, when crop features are selected appropriately, it is entirely possible to achieve better crop mapping than the reference data, to accurately identify crop types and field boundaries for the temporal and spatial transfer of crops. During the training process, this study does not rely on reference data of the target domain, offering a feasible study approach for crop mapping in regions where obtaining field samples is challenging. Additionally, the proportion of other land types is often much higher than that of studied crops, as shown in Fig. 3. So the model proposed in this study can be further applied to provide data and technical support for crop mapping with irregular spatial distribution, complex planting structures, unbalanced proportions for crop classes, and so on.

5.3. Future work

Data is crucial for deep learning networks, and high-quality data can greatly improve the accuracy of crop mapping. Therefore, more complete satellite data and processing flow can be further selected to reduce the interference of clouds and outliers. Secondly, suitable source domains and accurate data features are highly effective for the classification of specific crops. Thus, more efficient spectral features or VIs are selected to classify other crops in different target domains, such as cotton, potato, sorghum, and so on. In addition, multi-source validation is also important for evaluating the generalization ability of models. Only VIs calculated by the spectral reflectance were used to classify crops in this study, so the utilization of multi-source data from different satellites is a worthwhile study direction.

Due to manual annotation errors, limitations in image quality, and the inherent constraints of classification algorithms, the reference data inevitably contain errors. As a result, the inaccurate labels may lead to imperfect supervisory signals during the training processes, which can compromise the learning performance of models on the source domain and may further undermine the accuracy and stability of cross-domain predictions. The learning strategies with improved robustness will be investigated to mitigate the impact of inaccurate labels on model training. Furthermore, incorporating high-precision ground data and performing cross-validation with multi-source information can further enhance the reliability of labels and improve the predictive performance of models.

The focus of the model proposed in this study is on the cooperative use of multiple classifiers, but the ingenious application of feature extractors can also be an interesting research direction for future work. At the same time, the choices of loss functions are also crucial, as their results directly affect the learning process of model training. Subsequently, the development of UDA models that can successfully achieve domain alignment with incomplete and noisy remote sensing data is also an entirely new challenge, especially in the face of cloud cover or other occlusions. Besides, the computational efficiency of models is also critically important. It is meaningful to explore more lightweight and robust deep learning models that are suitable for low-resource environments while still maintaining effective domain adaptation capabilities.

6. Conclusions

This study proposed the MCDA model based on UDA for crop mapping in unlabeled study areas. Six study areas were selected in four countries and three sets of experiments were designed to classify maize and winter wheat, rice and soybeans, respectively. The MCDA model outperformed the others in most transfer cases, compared to the 1DCNN and LSTM models, as well as three UDA models. When classifying maize and winter wheat, the OA and $F1_w$ were separately 86.36% and 86.42% in Czech, and 85.64% and 85.03% in Kansas. For the classification of rice, the OA and $F1_w$ reached 92.50% and 92.57% in study area HLJ_R, and 92.10% and 91.38% in study area M_R. For the classification of soybeans, the OA and $F1_w$ were 81.90% and 81.93% in study area S_C, and 82.34% and 82.40% in study area M_S. In the visual predictions, the MCDA model outperformed the other models, and the MCDA model predicted better than the reference data in terms of the field completeness in Czech and the boundary accuracy in study area S_C. Moreover, the method proposed in this study accomplished more demanding transfer scenario for time and space, particularly the crop mapping across the northern and southern hemispheres.

CRediT authorship contribution statement

Changhong Xu: Writing – review & editing, Writing – original draft, Software, Formal analysis, Conceptualization. **Maofang Gao:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization. **Yuanwei Chen:** Resources, Investigation. **Jingwen Yan:** Writing – review & editing, Supervision, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by National Key Research and Development Program of China (No. 2022YFD1500204) and State key laboratory major special projects of Jilin Province Science and Technology Development Plan(Grant No. SKL202402024)

Data availability

Data will be made available on request.

References

Zhong, L., Hu, L., Zhou, H., 2019. Deep learning based multi-temporal crop classification [J]. Remote Sens. Environ. 221, 430–443.
Cai, Y., Guan, K., Peng, J., et al., 2018. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach[J]. Remote Sens. Environ. 210, 35–47.
Xu, C., Gao, M., Yan, J., et al., 2023. MP-Net: an efficient and precise multi-layer pyramid crop classification network for remote sensing images[J]. Comput. Electron. Agric. 212, 108065.
Li, H., Song, X., Hansen, M.C., et al., 2023. Development of a 10-m resolution maize and soybean map over China: Matching satellite-based crop classification with sample-based area estimation[J]. Remote Sens. Environ. 294, 113623.
Wei, P., Chai, D., Lin, T., et al., 2021. Large-scale rice mapping under different years based on time-series Sentinel-1 images using deep semantic segmentation model[J]. ISPRS J. Photogramm. Remote Sens. 174, 198–214.
Yang, M., Guo, B., Wang, J., 2024. A novel and robust method for large-scale single-season rice mapping based on phenology and statistical data[J]. ISPRS J. Photogramm. Remote Sens. 213, 14–32.
Li, S., Li, F., Gao, M., et al., 2021. A new method for winter wheat mapping based on spectral reconstruction technology[J]. Remote Sens. (Basel) 13 (9), 1810.
Xu, J., Zhu, Y., Zhong, R., et al., 2020. DeepCropMapping: a multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping[J]. Remote Sens. Environ. 247, 111946.
Feng, F., Gao, M., Liu, R., et al., 2023. A deep learning framework for crop mapping with reconstructed Sentinel-2 time series images[J]. Comput. Electron. Agric. 213, 108227.
Wang, S., Azzari, G., Lobell, D.B., 2019. Crop type mapping without field-level labels: Random forest transfer and unsupervised clustering techniques[J]. Remote Sens. Environ. 222, 303–317.
Ma, Y., Chen, S., Ermon, S., et al., 2024. Transfer learning in environmental remote sensing[J]. Remote Sens. Environ. 301, 113924.
Liu, X., Yoo, C., Xing, F., et al., 2022. Deep unsupervised domain adaptation: a review of recent advances and perspectives[J]. APSIPA Transactions on Signal and Information Processing 11 (1).
Ragab, M., Eldele, E., Tan, W.L., et al., 2023. Adatime: a benchmarking suite for domain adaptation on time series data[J]. ACM Trans. Knowl. Discov. Data 17 (8), 1–18.
Tzeng E, Hoffman J, Zhang N, et al. Deep domain confusion: Maximizing for domain invariance[J]. arXiv preprint arXiv:1412.3474, 2014.
Chen, C., Fu, Z., Chen, Z., et al., 2020. Homm: Higher-order moment matching for unsupervised domain adaptation[C]. AAAI.
Rahman, M.M., Fookes, C., Baktashmotlagh, M., et al., 2020. On minimum discrepancy estimation for deep domain adaptation[J]. Domain Adaptation for Visual Understanding 81–94.
Zhu, Y., Zhuang, F., Wang, J., et al., 2020. Deep subdomain adaptation network for image classification[J]. IEEE Trans. Neural Networks Learn. Syst. 32 (4), 1713–1722.
Ganin, Y., Ustinova, E., Ajakan, H., et al., 2016. Domain-adversarial training of neural networks[J]. J. Mach. Learn. Res. 17 (59), 1–35.
Long, M., Cao, Z., Wang, J., et al., 2018. Conditional adversarial domain adaptation[J]. Adv. Neural Inf. Proces. Syst. 31.
Wilson, G., Doppa, J.R., Cook, D.J., 2020. Multi-source deep domain adaptation with weak supervision for time-series sensor data[C]. In SIGKDD.
Liu, Q., Xue, H., 2021. Adversarial spectral kernel matching for unsupervised time series domain adaptation[C]. In Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence.
Gao, L., Wang, X., Johnson, B.A., et al., 2020. Remote sensing algorithms for estimation of fractional vegetation cover using pure vegetation index values: a review[J]. ISPRS J. Photogramm. Remote Sens. 159, 364–377.
You, N., Dong, J., Li, J., et al., 2023. Rapid early-season maize mapping without crop labels[J]. Remote Sens. Environ. 290, 113496.
Hao, P., Di, L., Zhang, C., et al., 2020. Transfer Learning for Crop classification with Cropland Data Layer data (CDL) as training samples[J]. Sci. Total Environ. 733, 138869.
Wang, Y., Feng, L., Sun, W., et al., 2022. Exploring the potential of multi-source unsupervised domain adaptation in crop mapping using Sentinel-2 images[J]. Giscience & Remote Sensing 59 (1), 2247–2265.
Wang, Y., Feng, L., Zhang, Z., et al., 2023. An unsupervised domain adaptation deep learning method for spatial and temporal transferable crop type mapping using Sentinel-2 imagery[J]. ISPRS J. Photogramm. Remote Sens. 199, 102–117.
Wang, H., Yao, Y., Ye, Z., et al., 2024. Solution for crop classification in regions with limited labeled samples: deep learning and transfer learning[J]. Giscience & Remote Sensing 61 (1).
Ochs, C.A., Baustian, J., Harrison, A.B., et al., 2023. Rivers of the lower Mississippi Basin [M]//Rivers of North America. Elsevier 226–271.
Phiri, D., Simwanda, M., Salekin, S., et al., 2020. Sentinel-2 data for land cover/use mapping: a review[J]. Remote Sens. (Basel) 12 (14), 2291.
Yang, K., Luo, Y., Li, M., et al., 2022. Reconstruction of sentinel-2 image time series using google earth engine[J]. Remote Sens. (Basel) 14 (17), 4395.
D Andrimont R E L, Verhegghen A, Lemoine G, et al. From parcel to continental scale–A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations[J]. Remote sensing of environment, 2021,266:112708.
You, N., Dong, J., Huang, J., et al., 2021. The 10-m crop type maps in Northeast China during 2017–2019[J]. Sci. Data 8 (1), 41.
Song, X., Hansen, M.C., Potapov, P., et al., 2021. Massive soybean expansion in South America since 2000 and implications for conservation[J]. Nat. Sustainability 4 (9), 784–792.
Ashourloo, D., Nematollahi, H., Huete, A., et al., 2022. A new phenology-based method for mapping wheat and barley using time-series of Sentinel-2 images[J]. Remote Sens. Environ. 280, 113206.
Jiang, Z., Huete, A.R., Didan, K., et al., 2008. Development of a two-band enhanced vegetation index without a blue band[J]. Remote Sens. Environ. 112 (10), 3833–3845.
Chandrasekar, K., Sesha Sai, M., Roy, P.S., et al., 2010. Land Surface Water Index (LSWI) response to rainfall and NDVI using the MODIS Vegetation Index product[J]. Int. J. Remote Sens. 31 (15), 3987–4005.
Zhang, C., Zhang, H., Tian, S., 2023. Phenology-assisted supervised paddy rice mapping with the Landsat imagery on Google Earth Engine: Experiments in Heilongjiang Province of China from 1990 to 2020[J]. Comput. Electron. Agric. 212, 108105.
McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features[J]. Int. J. Remote Sens. 17 (7), 1425–1432.
Xu, H., 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery[J]. Int. J. Remote Sens. 27 (14), 3025–3033.
Chen, H., Li, H., Liu, Z., et al., 2023. A novel Greenness and Water Content Composite Index (GWCCI) for soybean mapping from single remotely sensed multispectral images[J]. Remote Sens. Environ. 295, 113679.
Mao, A., Mohri, M., Zhong, Y., 2023. Cross-entropy loss functions: Theoretical analysis and applications[C]. International conference on Machine learning. PMLR.

Hou, Y., Zheng, L., 2021. Visualizing adapted knowledge in domain transfer[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.

You, K., Wang, X., Long, M., et al., 2019. Towards accurate model selection in deep unsupervised domain adaptation[C]. International Conference on Machine Learning.

Eldele, E., Ragab, M., Chen, Z., et al., 2023. Contrastive domain adaptation for time-series via temporal mixup[J]. IEEE Trans. Artif. Intell. 5 (3), 1185–1194.