Feature deformation network with multi-range feature enhancement for agricultural machinery operation mode identification

Weixin Zhai[1,2†], Zhi Xu[1,2†], Jinming Liu[1†], Xiya Xiong[1†], Jiawen Pan[1,2*], Sun-Ok Chung[3], Dionysis Bochtis[4], Caicong Wu[1,2]

(1. *College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China*;
2. *Key Laboratory of Agricultural Machinery Monitoring and Big Data Application, Ministry of Agriculture and Rural Affairs, Beijing 100083, China*;
3. *Department of Smart Agricultural Systems, Graduate School, Chungnam National University, Daejeon 34134, Republic of Korea*;
4. *Institute for Bio-Economy and Agri-Technology (iBO), Centre for Research and Technology-Hellas (CERTH), Thessaloniki 57001, Greece*)

Abstract: Utilizing the spatiotemporal features contained in extensive trajectory data for identifying operation modes of agricultural machinery is an important basis task for subsequent agricultural machinery trajectory research. In the present study, to effectively identify agricultural machinery operation mode, a feature deformation network with multi-range feature enhancement was proposed. First, a multi-range feature enhancement module was developed to fully explore the feature distribution of agricultural machinery trajectory data. Second, to further enrich the representation of trajectories, a feature deformation module was proposed that can map trajectory points to high-dimensional space to form feature maps. Then, EfficientNet-B0 was used to extract features of different scales and depths from the feature map, select features highly relevant to the results, and finally accurately predict the mode of each trajectory point. To validate the effectiveness of the proposed method, experiments were conducted to compare the results with those of other methods on a dataset of real agricultural trajectories. On the corn and wheat harvester trajectory datasets, the model achieved accuracies of 96.88% and 96.68%, as well as F1 scores of 93.54% and 94.19%, exhibiting improvements of 8.35% and 9.08% in accuracy and 20.99% and 20.04% in F1 score compared with the current state-of-the-art method.
Keywords: road-field trajectory classification, efficientNet, feature deformation network, multi-range feature enhancement, agricultural machinery operation mode recognition
DOI: 10.25165/j.ijabe.20241704.8831

Citation: Zhai W X, Xu Z, Liu J M, Xiong X Y, Pan J W, Chung S, et al. Feature deformation network with multi-range feature enhancement for agricultural machinery operation mode identification. Int J Agric & Biol Eng, 2024; 17(4): 265–275.

1 Introduction

In recent years, the process of agricultural modernization in China has been developing rapidly, and agricultural machinery has been adopted for large-scale agricultural operations, thus generating a large amount of trajectory data[1,2]. The trajectory is a time-ordered coordinate sequence that records agricultural machinery during the process of driving, reflecting the spatial changes in the machinery over a period of time; this sequence is used to depict the movement trajectory[3]. There are many ways to collect trajectory data, and one of the main ways is to utilize the Global Navigation Satellite System (GNSS) for data collection[4]. GNSSs are navigation and positioning systems that provide users with two-dimensional coordinates and temporal and velocity information about an object at any time and place[5]. After a large amount of data has been collected, further analyzing and applying these trajectory data is more important in the actual production chain. Its wide range of application areas includes but is not limited to recognizing the cross-area operational behavior of agricultural machinery[6,7], planning the operation path of agricultural machinery[8-10], recognizing the companion behavior of agricultural machinery, automated navigation of agricultural machinery[11,12], agricultural road network construction[13,14], agricultural big data precision analysis[15-17], operation efficiency evaluation of agricultural machinery[18,19], and conducting in-depth studies on the quantitative relationship between the operational efficiency of agricultural machinery and weather[20].

To implement effective analysis and mining of agricultural machinery trajectory data, a key prerequisite task is to first understand the driving scene and activity type related to each point in the trajectory of agricultural machinery, which is also called field-road classification. Field-road classification aims at semantically segmenting trajectories by learning the trajectory data of agricultural machinery when it operates and recognizing its

Received date: 2024-01-26 **Accepted date:** 2024-06-13
Biographies: Weixin Zhai, PhD, Associate Professor, research interest: big data of spatio temporal, big data mining of agricultural machinery, cartography and geographic information system, Email: zhaiweixin@cau.edu.cn; **Zhi Xu**, Master, research interest: big data mining of agricultural machinery, Email: xuzhihh@163.com; **Jinming Liu**, Bachelor candidate, research interest: data science and big data technology, Email: 2279315755@qq.com; **Xiya Xiong**, Bachelor candidate, research interest: data science and big data technology, Email: xiongxiya_sia@163.com; **Sun-Ok Chung**, PhD, Professor, research interest: precision agriculture, Email: sochung@cnu.ac.kr; **Dionysis Bochtis**, PhD, Professor, research interest: precision agriculture, Email: d.bochtis@certh.gr; **Caicong Wu**, PhD, Professor, research interest: unmanned driving and autonomous operation of agricultural machinery, big data mining of agricultural machinery, Email: wucc@cau.edu.cn.
†The authors contributed equally to this work.
***Corresponding author: Jiawen Pan**, Doctor candidate, research interest: computational intelligence, computer vision. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China. Tel: +86-15588880367, Email: cau_panjiawen@cau.edu.cn.

operation modes. The operation modes include fields and roads. When working in the field, the trajectory points of agricultural machinery are semantically labelled “field points”, while when driving on a road, they are labelled “road points”. This semantic segmentation method helps to understand and distinguish the trajectory characteristics of agricultural machinery in different scenarios more accurately and provides a strong theoretical basis for the subsequent analysis of agricultural machinery behavior.

At present, semantic segmentation of trajectories using artificial intelligence technology has become a popular research topic. It can not only more accurately identify the behavior of agricultural machinery in field operations and road driving but also provide deeper semantic labels to the trajectory data, thus making the subsequent analysis of the data more detailed and targeted. Chen et al.[21] proposed a DBSCAN field-road classification model based on direction distribution (BDFRTS). This method introduces inference rules for speed and direction; for example, the driving direction of agricultural machinery is mostly parallel, and the driving speed is slower when the machinery is working in the field. In contrast, when the machinery is driving on the road, the classification results of the unsupervised clustering method DBSCAN are corrected. In addition, Poteko et al.[22] extracted statistical parameters based on individual points and their temporal neighbors as features that were subsequently fed into a decision tree (DT) for learning to achieve field-road classification; Xiao et al.[23] proposed an XGBoost field-road classification model with a time window feature extraction operators to extract features and a recursive feature elimination algorithm to eliminate redundant features (DR-XGBoost). The above methods achieve field-road classification by building machine-learning models or combining logical assumptions for inference. Although these methods have achieved acceptable results, they still have drawbacks. The BDFRTS methods use only latitude and longitude as inputs, which makes it difficult to comprehensively capture trajectory features, and the model depends on parameters that need to be constantly adjusted to achieve better performance. The DT models do not adequately mine the intrinsic information of the data and focus only on the features of individual trajectory points without acknowledging the connections between trajectory points. Although the DR-XGBoost model improves feature extraction by taking into account the connections between points, it still utilizes discrete features, which necessitates the discretization of continuous features. However, this discretization process can potentially result in information loss.

Recently, deep learning-based methods have been proposed to further facilitate field road classification tasks. Deep learning methods mainly involve setting up the deep structure of complex models, collecting a large amount of data, and transforming the data (including spatiotemporal relationship graphs and pixel images) so that the model can automatically extract important information from the data and thus better understand and differentiate between the essential features of different trajectory data for classification. Chen et al.[24] utilized the spatiotemporal relationship between each point and its neighboring points to construct spatiotemporal graphs, thus obtaining a rich representation of the features for each GNSS point and then applying a graph convolution neural network (GCN) to propagate the features between graph nodes, thus aggregating the point information in the trajectory to achieve classification. Furthermore, Zhang et al.[25] proposed transforming trajectory data into images and applying an object detection model to detect objects in images; additionally, applying the traditional DBSCAN clustering method (DBSCAN+object detection); finally combining two classification results using the Davis-Boldin Index (DBI). Chen et al.[26] also extracted two types of input feature vectors, statistical feature vectors, and visual feature vectors, to represent a GNSS trajectory point and then integrated the input features using a BiLSTM network to ultimately achieve classification for each GNSS point (BiLSTM). Zhai et al.[27] proposed a generative adversarial network-bidirectional long short-term memory network (GAN-BiLSTM) field-road classification model. They used Generative Adversarial Networks (GAN) for data augmentation to obtain a balanced dataset, which was subsequently fed into BiLSTM with an attention mechanism and focal loss as the loss function for classification. The deep learning methods described above further extract trajectory features and achieve better performance. However, a spatiotemporal graph needs to be constructed, which consumes a large amount of computational resources; the results of the DBSCAN+object detection model and BiLSTM model rely on the selection of the mapping model that transforms the data into images, and the mapping process introduces biases through operations, which prevent image data from accurately expressing the original trajectory information.

Although these studies have made promising progress in model design and motion information extraction and all of them have improved the accuracy of field-road classification to a certain extent, shortcomings such as an imbalanced distribution of raw data, lack of overall motion information, and high computational loss still exist.

To overcome the limitations mentioned above, this study proposed a feature deformation network with multi-range feature enhancement for agricultural machinery operation mode identification (FDRNet). First, to fully explore the feature distribution of agricultural machinery trajectory data, a multi-range feature enhancement module was introduced. The raw data were processed using short-range and long-range feature extraction methods to capture the spatial structure, movement trends, and boundary details of agricultural machinery trajectories. This approach accurately describes the movement patterns and behaviors of agricultural machinery, thereby enhancing the ability of the model to represent trajectory data. Second, the feature deformation method was proposed (FD). To further enrich feature representation, FD first trains trajectory features in a high-dimensional space, expanding the data dimensionality. FD then maps the feature elements of each trajectory point into a feature map. This approach transforms the field road trajectory classification problem into an image classification problem. This method is simple and efficient and provides a new paradigm for field road trajectory classification tasks. Finally, to extract features of different scales and depths from trajectories and filter out features highly relevant to the results, this study adopted EfficientNet-B0 as the backbone network. EfficientNets are a series of composite scaling networks designed to achieve better performance than traditional networks with lower model sizes and computational resources[28]. The method involves using a set of scaling coefficients to uniformly adjust the network’s depth, width, and resolution. EfficientNet-B0 is the baseline model in the EfficientNet series. This network effectively captures the spatial information of trajectory points and relationships between different feature channels in a short time. The weight of each channel is dynamically adjusted based on this information. Therefore, this enables the network to focus more on important features and improves the model’s perception of trajectory features. EfficientNet-B0 extracts rich features from multiple perspectives to accurately predict the activities of each trajectory point, such as "field" and "road." To validate the effectiveness of the proposed approach, experiments were conducted on datasets of wheat and corn harvester trajectories. The results demonstrate that our method achieves the best outcomes in field road trajectory classification tasks. The main contributions of this study are as follows:

1) Introducing a multi-range feature enhancement method to uncover feature information within agricultural machinery trajectories. This captures changes in trajectory features and movement trends of machinery, constructing spatiotemporal relationship features;

2) A trajectory feature deformation method is proposed. Initially, trajectory dimensionality is increased through a linear mapping layer, followed by the deformation of the trajectory temporal features into the feature maps. The field-road trajectory classification task is thereby transformed into an image classification task. This deformation overcomes the defect that traditional CNN-based network can not deal with trajectories directly, providing a universal image model solution for the field-road trajectory classification task.

3) The network EfficientNet-B0 from the field of image classification was chosen as the backbone network. This approach captures both local and global information on trajectory features, identifying crucial features. This network extracts diverse features at different scales and depths from trajectory data, allowing it to precisely comprehend the distribution of trajectory features. As a result, it can more effectively recognize agricultural machinery operation mode.

2 Materials and methods

2.1 Overview

To overcome the limitations of the current models used for field-road trajectory classification, this study developed a network suitable for trajectory feature extraction of agricultural machinery called FDRNet. This section will elaborate on the principle of the FDRNet algorithm. It is composed of three key components: feature augmentation, trajectory feature map creation, and a model for classifying field road trajectories. The specific algorithm flow chart is shown in Figure 1.

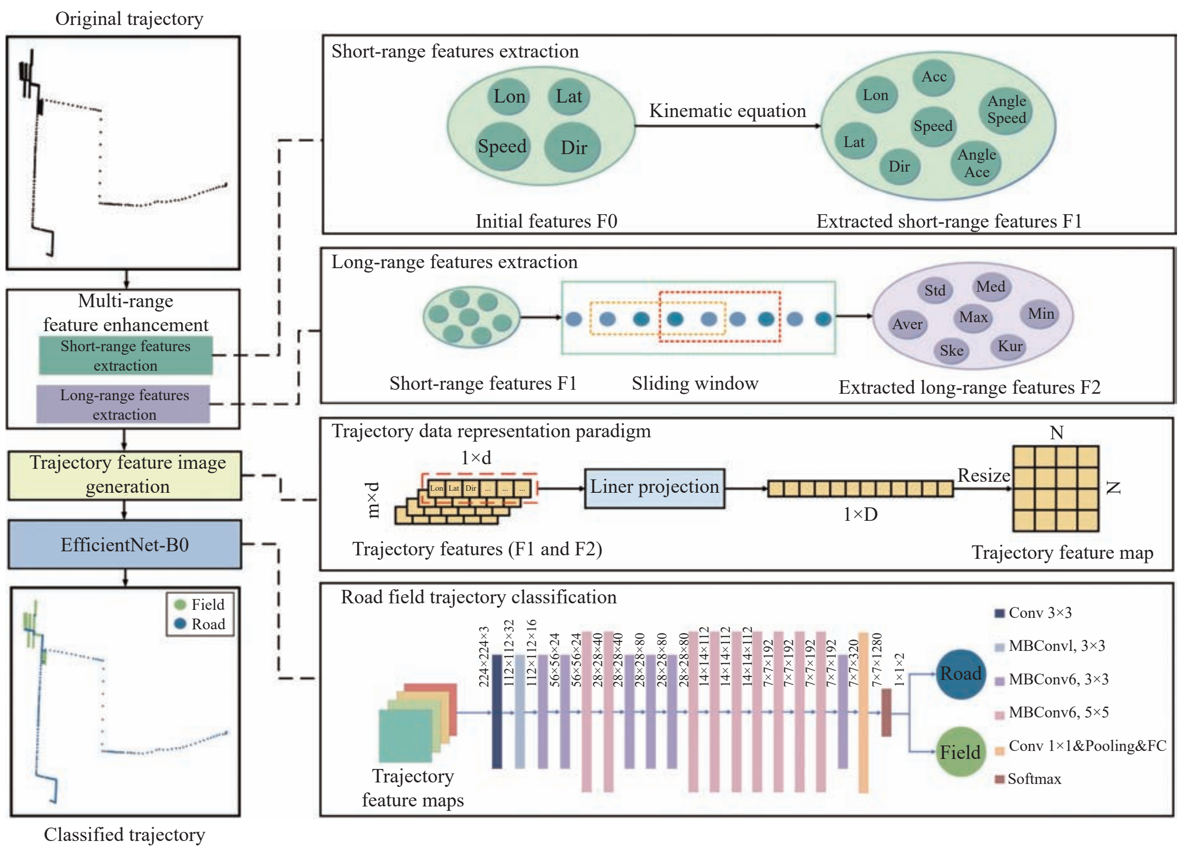

Figure 1 The pipeline of the proposed FDRNet

2.2 Multi-range feature enhancement

This section introduces a multi-range feature enhancement method. It first utilizes a short-range feature extraction method (SRF) to capture the relationships between different features of the same trajectory point. Then, a long-range feature extraction method (LRF) is used to identify changes in the feature distribution of different trajectory points.

Existing field-road trajectory classification models mostly exploit the initial features of trajectories, such as longitude, latitude, speed, and direction. However, the information embedded in trajectories is not sufficiently mined. Based on the above issue, the SRF was presented, where the initial features of the trajectory are computed by a kinematic formula to obtain the short-range features of the trajectory. Assuming a set of n trajectory points is represented as T={(longitude$_i$, latitude$_i$, time$_i$, speed$_i$, direction$_i$, lable$_i$)| I=1, 2, 3, …, n} where longitude$_i$, latitude$_i$, time$_i$, speed$_i$, direction$_i$, label$_i$ are the longitude, latitude, time, speed, direction, and label of the ith trajectory point, respectively. Here, label$_i$∈{0, 1}, where "0" represents a road point and "1" represents a field point. In this study, the initial features include the longitude, latitude, speed, and direction from the set T. Then, the speed difference was calculated, acceleration, direction difference, angular velocity, angular velocity difference, and angular acceleration as short-range features of the ith trajectory point (Equations (1)-(6)).

$$\mathrm{speedDiff}_i = \mathrm{speed}_i - \mathrm{speed}_{i-1} \tag{1}$$

$$\mathrm{Acc}_i = \frac{\mathrm{speedDiff}_i}{\mathrm{time}_i - \mathrm{time}_{i-1}} \tag{2}$$

$$\mathrm{Ang}_i = \mathrm{direction}_i - \mathrm{direction}_{i-1} \tag{3}$$

$$\mathrm{angVel}_i = \frac{\mathrm{Ang}_i}{\mathrm{time}_i - \mathrm{time}_{i-1}} \tag{4}$$

$$\mathrm{angVeldiff}_i = \mathrm{angVel}_i - \mathrm{angVel}_{i-1} \tag{5}$$

$$\mathrm{angAcc}_i = \frac{\mathrm{angVeldiff}_i}{\mathrm{time}_i - \mathrm{time}_{i-1}} \tag{6}$$

where, $\mathrm{speedDiff}_i$, Acc_i, Ang_i, angVel_i, $\mathrm{angVeldiff}_i$, and angAcc_i represent the speed difference, acceleration, direction difference, angular velocity, angular velocity difference, and angular acceleration of the ith point, respectively.

There are significant differences in the movement characteristics of agricultural machinery in the field and on the road. Normally, when working in fields, agricultural machinery tends to move at an approximately constant low speed. Consequently, during certain continuous time intervals, the average, median, and standard deviation of the acceleration of trajectory points are approximately zero. On the other hand, when agricultural machinery travels on roads, its speed is higher than that in the field, and it often needs to perform actions such as acceleration, braking, and turning. This leads to considerable variations in the speed, acceleration, angular difference, angular velocity, and angular acceleration at different time points. Therefore, according to the distinct distributions of trajectory features in fields and on roads, this paper proposes the LRF. Within a sliding window n, the LRF calculates the mean, median, standard deviation, maximum, minimum, skewness, and kurtosis of the above short-range features for the trajectory point point_i as long-range features. Here, the window size n represents the number of adjacent points used for feature calculation around the current trajectory point point_i. To capture feature changes within short and long distances, the sliding window sizes were empirically set to 5 and 20. Skewness and kurtosis were used to measure the asymmetry and steepness of the feature distribution, as shown in Equations (7) and (8).

$$S = \frac{1}{n}\sum_{i=1}^{n}\left[\left(\frac{X_i-\mu}{\sigma}\right)^3\right] \tag{7}$$

$$K = \frac{1}{n}\sum_{i=1}^{n}\left[\left(\frac{X_i-\mu}{\sigma}\right)^4\right] \tag{8}$$

where, S, K, μ, and σ represent the skewness, kurtosis, mean, and standard deviation, respectively.

On the basis of the initial features of agricultural machinery trajectories, this study initially extracts short-range features from the instantaneous state of the machinery. Long-range features were subsequently derived by exploring changes in agricultural machinery states over a period of time. This process establishes correlations between different trajectory points, enhancing the feature representation of agricultural machinery trajectories.

2.3 Feature deformation network

Current research on field-road trajectory classification can be broadly divided into four types: density clustering-based field-road trajectory classification models, decision tree-based field-road trajectory classification models, image processing-based field-road trajectory classification models, and graph convolutional network-based field-road trajectory classification models. However, each type of model has its own limitations. For instance, the density clustering-based field-road trajectory classification model requires a large amount of manual parameter tuning and exhibits limitations in generalization; moreover, this model lacks precision at the edges of trajectories (where field trajectory points meet road trajectory points). The decision tree-based field-road trajectory classification model overlooks the correlation between trajectory features and is not suitable for handling multidimensional trajectory data containing both continuous and discrete features. The image processing-based field-road trajectory classification model risks information loss during data conversion and struggles with parameters such as pixel size and image size. Finally, the GCN-based field-road trajectory classification model has a large number of parameters and requires long training times. When constructing point relationships, only local spatiotemporal relationships are considered, and the process consumes a significant amount of hardware resources. To overcome the aforementioned limitations, this paper proposes a novel pattern recognition method tailored for agricultural machinery. It employs a feature deformation network to handle the field-road trajectory classification task. Specifically, in this section, we first use a trainable linear projection to map the enhanced features to the required dimensions. Subsequently, GNSS points are automatically generated into “feature maps” via a feature deformation method. Next, the EfficientNet-B0 image classification model is utilized as the backbone network to extract features at different levels from the trajectory feature maps. Finally, based on the extracted features, a linear classifier is employed to predict the categories (“field” or “road”) of each point.

2.3.1 Trajectory feature image generation

Due to the limited number of enhanced features, the network’s depth was restricted. Therefore, the model cannot effectively learn the feature distribution of the trajectory. Moreover, sequential trajectory data cannot be directly input into an image classification network. To overcome these challenges, a feature deformation network was presented. The network first introduces a trainable linear projection layer. This layer, through a fully connected network, maps the dimensionality of the enhanced features to a more expansive space. Subsequently, a feature deformation method was defined for expressing trajectory data. This deformation maps the feature vectors of GNSS trajectories to feature images. Each element in the feature vector corresponds one-to-one with a pixel in the feature image. Notably, each point in the GNSS trajectory represents an independent feature image. This method effectively resolves the issue of inputting trajectory data into the model.

There is a set of trajectory data $P = \vec{p_1}, \vec{p_2}, \vec{p_3}, \ldots, \vec{p_m}$, $\vec{p_i} \in R^d$, where m is the number of trajectory points and d is the input feature dimension for each trajectory point. In this paper, after the linear projection layer, a d-dimensional feature vector is mapped to D dimensions. The output of this projection was referred to as the feature embedding. The $1\times D$-dimensional feature embedding $F=\{f_1, f_2, f_3, \ldots, f_D\}$ of each trajectory point $\vec{p_i}$ is resized into a feature matrix $F' \in R^{N\times N}$, where $N\times N=D$. The overall process is illustrated in Figure 2. In this way, the trajectory points are deformed into a “feature map” that the image classification model can process directly. Here, the $d$dimensional feature vector corresponds to the features extracted in Section 2.2, and N^2 is the number of pixels in the feature image. Compared to existing methods of rasterizing image conversion, the advantage of the proposed method lies in its abundance of features and simplicity of implementation, which can effectively enhance the classification performance of the model.

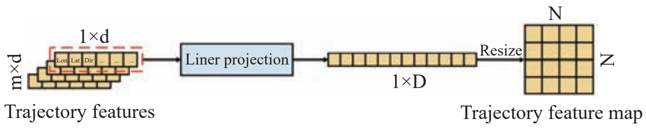

Figure 2 The deformation of the trajectory feature

2.3.2 Network structure

Through muti-range feature enhancement and feature deformation, each agricultural machinery trajectory point was represented as a feature map rich in spatial information. To further capture trajectory features at various scales and depths and to select features highly relevant to the outcome, this study employed EfficientNet-B0 as the backbone network. EfficientNet-B0 scales its depth, width, and resolution proportionally. Hence, this approach can capture agricultural machinery trajectory features at different levels of abstraction and granularity, achieving effective multiscale feature extraction. Additionally, this network can efficiently capture the spatial information of trajectory points and the relationships between different trajectory feature channels in a short period of time. By adjusting the weight of each channel, trajectory features that are highly correlated with the outcome were selected. Therefore, this model not only enhances the ability of the model to recognize patterns and details in trajectory data but also meets the real-time requirements of field and road trajectory classification tasks.

The network architecture of EfficientNet-B0 is listed in Table 1 and Figure 3. The network is divided into 9 stages, where the first stage consists of a regular convolutional layer with a 3×3 kernel size

Table 1 EfficientNet-B0 baseline network

Stage i	Operator $\widehat{F}_i$	Resolution $\widehat{H}_i \times \widehat{W}_i$	#Channels $\widehat{C}_i$	#Layers $\widehat{L}_i$
1	Conv 3×3	224×224	32	1
2	MBConv1, k3×3	112×112	16	1
3	MBConv6, k3×3	112×112	24	2
4	MBConv6, k5×5	56×56	40	2
5	MBConv6, k3×3	28×28	80	3
6	MBConv6, k5×5	14×14	112	3
7	MBConv6, k5×5	14×14	192	4
8	MBConv6, k3×3	7×7	320	1
9	Conv 1×1&Pooling&FC	7×7	1280	1

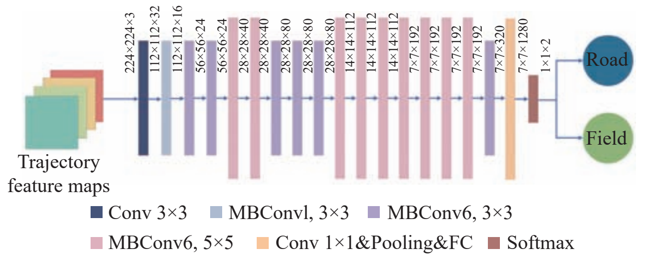

Figure 3 The network structure of EfficientNet-B0

and a stride of 2 (including batch normalization (BN) and an activation function (Swish)). Stages 2 to 8 comprise repeated stacks of MBConv structures and Layers $\widehat{L}_i$ indicates the number of times the MBConv structure is repeated in that stage. Stage 9 consists of a regular 1×1 convolutional layer (including BN and Swish), an average pooling layer, and a fully connected layer (FC). In Table 1, MBConv$n$ denotes the multiplier factor n, which determines how the first 1×1 convolutional layer within the MBConv expands the input trajectory feature matrix's channels by n times. In EfficientNet-B0, n is assigned a value of 1 or 6. Additionally, k3×3 or k5×5 represents the kernel size used in the depthwise convolution within the MBConv. The resolution indicates the size of the input trajectory feature matrix $\langle \widehat{H}_i, \widehat{W}_i \rangle$. Channels $\widehat{C}_i$ represents the number of output trajectory feature matrix channels after passing through that stage.

MBConv, originating from the InvertedResidualBlock of the MobileNetV3 network, differs in EfficientNet's MBConv, which employs the Swish activation function, along with the inclusion of the Squeeze-and-Excitation (SE) module in each MBConv. The structure of the MBConv network is illustrated in Figure 4. Each MBConv consists of a 1×1 regular convolution (including BN and Swish), a k×k Depthwise Convolution (including BN and Swish), an SE module, a 1×1 regular convolution (including BN), and a Dropout layer. The value of k is either 3 or 5. In contrast to regular convolutions, Depthwise Convolution calculates with a convolution kernel only for a single channel of input trajectory features, significantly reducing the parameter count. Additionally, MBConv employs residual connections, where the input is directly added to the output, forming a residual block. This helps prevent gradient vanishing issues, making the network easier to train. Swish introduces non-linear factors, aiding the model in learning complex non-linear relationships and trajectory feature representations. Moreover, Swish includes a learnable parameter β, allowing the activation function's shape to adapt based on the trajectory data distribution. This adaptability contributes to enhancing the network's generalization ability across different data distributions. Its formula is depicted in Equation (9). When β=0, the activation function becomes a linear function $f(x)=x/2$; when $\beta=\infty$, the activation function becomes the ReLU function. Swish has only a lower bound, ensuring no gradient saturation during training, providing stronger regularization effects. Furthermore, its continuous differentiability facilitates easier optimization of network parameters during backpropagation.

$$f(x) = x \cdot \mathrm{sigmoid}(\beta x) \tag{9}$$

The SE module is essentially a channel attention mechanism comprising Squeeze, Excitation, and Scale stages. In the Squeeze stage, the SE module reduces the dimensions of each channel through Global Average Pooling (AvgPooling). This operation averages all elements in each channel, compressing the channel information into a single value. This value represents the global importance of that channel. In the excitation stage, 2 fully connected layers are used to learn the weights of each channel.

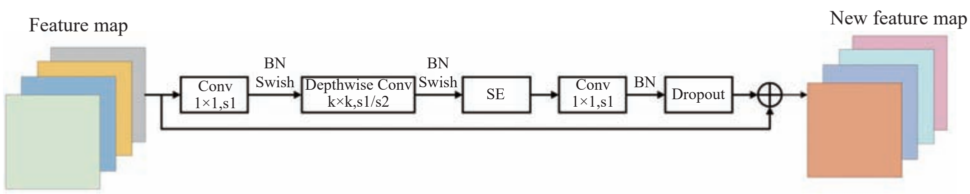

Figure 4 The structure of the MBConv

These weights, known as excitation factors, measure the importance of each channel globally. Finally, in the scale stage, the original trajectory feature map is reweighted channel-wise by multiplying it with the learned excitation factors. This operation allows the network to adaptively emphasize important channels during training, thereby improving overall performance. The network architecture is shown in Figure 5. The main goal of the SE module is to learn channel relationships in the input trajectory feature map and dynamically adjust the weight of each channel accordingly. Therefore, the SE module enables the network to focus more on crucial features, enhancing the model's perception of trajectory features.

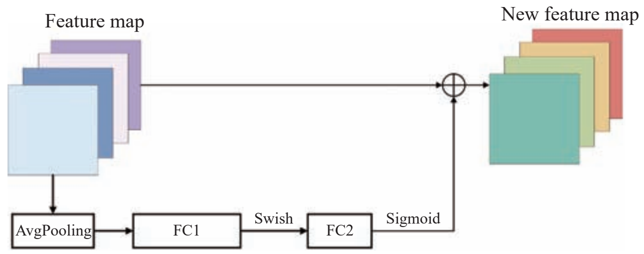

Figure 5 The structure of the SE

The EfficientNet-B0 network ultimately produces a global average feature. Next, the classification layer uses the softmax function to transform the trajectory features extracted by EfficientNet-B0 into categorical scores. This helps determine whether a trajectory point belongs to a field or a road, completing the task of classifying field-road trajectories.

After the augmented features are mapped into high-dimensional space, trajectory points are transformed into feature images using a trajectory representation paradigm as the model input. EfficientNet-B0 not only extracts local correlation information of adjacent elements in the trajectory point feature vector but also captures global correlation information of distant features within the feature vector. Additionally, based on the characteristics of MBConv, EfficientNet-B0 can effectively capture the spatial information of trajectory points and channel relationships of feature maps. This method reduces the number of parameters and lowers the computational complexity while extracting implicit features of agricultural machinery trajectories from multiple dimensions. Therefore, this approach can efficiently achieve field-road trajectory classification.

3 Results and discussion

3.1 Experimental setup

The Key Laboratory of Agricultural Machinery Operation Monitoring and Big Data Application, Ministry of Agriculture and Rural Affairs, People's Republic of China, provided the Grain Harvester Trajectory dataset for the study. To evaluate the model's performance, two sets of agricultural machinery trajectory datasets were selected with different crop types, trajectory numbers, operation times, geographical locations, sampling frequencies, and GNSS receivers. These datasets were named the "Maize Harvester Trajectory Dataset" and the "Wheat Harvester Trajectory Dataset", respectively, based on crop type. The detailed information is presented in Table 2. The initial features of each trajectory point include coordinates (longitude and latitude, WGS84), time (YYYY: MM:DD - hh:mm:ss), speed (m/s), direction (°), height (m), and labels. The labels represent manually annotated trajectory point categories ("Field" or "Road"). The data cleaning on each GNSS trajectory dataset were performed, including noise point smoothing and duplicate point removal, following the approach[21]. Afterward, the datasets were randomly divided into training and testing sets at a ratio of 8:2. All the experimental data used in this study are available at https://github.com/Agribigdata/dataset_code. For the experiments, Python, PyTorch, an NVIDIA Tesla V100 GPU, and an Intel(R) Xeon(R) Gold 6226R CPU @ 2.90 GHz were employed. The hardware environment for the experiments is detailed in Table 3.

Table 2 Information of GNSS trajectory datasets

Parameters	Maize Harvester Trajectory Dataset	Wheat Harvester Trajectory Dataset
Number of trajectories	90	117
Number of points	828 891	932 493
Acquisition period	2019.9.15-2019.10.15	2021.6.1-2021.6.30
Geographical location	Henan Province, Anhui Province, Shandong Province, Hubei Province, Jilin Province, Liaoning Province Shanxi Province, and Inner Mongolia Autonomous Region	Xinyang City, Shangqiu City, Zhumadian City, Kaifeng City, Pingdingshan City, Zhoukou City, and Xinxiang City in Henan Province
Acquisition frequencies	86% points were recorded every 20 s	93% points were recorded every 5 s
GNSS Receiver Co.	Jiangsu Langhe Control System Co., and Beijing Universal Mobile Linking Technology Co.	Jiangsu Langhe Control System Co.
Agriculture Machinery Co.	Weichai Lovol Heavy Industry Co., and Jiangsu World Agriculture Machinery Co.	Jiangsu World Agriculture Machinery Co.

Table 3 Hardware environment of the experiment

Configuration	Parameter
Programming language	Python3.7
Library and wrapper	PyTorch1.13.1
CPU	Intel(R) Xeon(R) Gold 6226R CPU @ 2.90 GHz
GPU	NVIDIA Tesla V100
Operating system	Ubuntu18.04

3.2 Performance metrics

In this study, multiple evaluation metrics were employed to comprehensively assess the performance of the proposed model. The following are the definitions of the evaluation metrics used in this study. The accuracy (Equation (10)) represents the proportion of correctly classified samples by the model out of the total samples. Here, TP represents the number of true-positive samples, TN represents true-negative samples, FP represents false-positive samples, and FN represents false-negative samples. Precision (Equation (11)) measures the model's ability to correctly identify positive samples among those it predicts as positive. Recall (Equation (12)) quantifies the model's ability to correctly recognize positive samples among all actual positive samples. The F1 score (Equation (13)) is a combined metric that balances both precision and recall to assess classification accuracy and the model's ability to identify positive samples. In this research, each of these metrics separately was calculated for the "field" and "road" categories as positive cases. Subsequently, we compute the average metric values across all categories. By taking into account all of these evaluation metrics, the aim is to provide a comprehensive assessment of the model's performance.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{11}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{12}$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{13}$$

3.3 Method comparisons and discussion

To validate the reliability and effectiveness of FDRNet, it was compared with 6 existing models for road classification: Decision Tree (DT)[22], DBSCAN+Rules[21], Graph Convolutional Network (GCN)[24], Random Forest (RF), XGBoost, and Multilayer Perceptron (MLP). Among them, GCN is the state-of-the-art method among all current field road classification models.

Tables 4 and 5 present the overall classification results of the proposed method and other models on the corn and wheat harvester trajectory datasets. On the corn harvester trajectory dataset, our method achieved an accuracy of 96.88% and an F1-score of 93.54%. The accuracy was 7.64% higher than that of the second-best model (DBSCAN+Rules), and the maximum improvement in accuracy was 38.92% over that of the DT. The F1-score was 13.66% higher than the second-best model (DBSCAN+Rules) and exhibited a maximum improvement of 38.27% over the DT. Similarly, on the wheat harvester trajectory dataset, our method also obtained the best accuracy (96.68%) and F1-score (94.19%). The accuracy was 8.73% higher than that of the second-best network (GCN) and showed a maximum improvement of 45.54% over that of the DT. The F1-score was 18.32% higher than the second-best network (GCN) and exhibited a maximum improvement of 46.39% over DT. Tables 6 and 7 detail the specific classification results of the proposed method and other methods on the maize and wheat harvester trajectory datasets. From the tables, it is evident that existing models exhibit limitations in classifying road trajectory points. However, the model of this study significantly improves upon the other models in this respect. On the maize and wheat harvester trajectory datasets, FDRNet achieved F1-scores of 88.89% and 90.40%, respectively, for road trajectory points, surpassing the best existing results by 22.73% (DBSCAN+Rules) and 31.60% (GCN), respectively.

Table 4 Overall performances of all methods on the corn harvester trajectory dataset

Method	Accuracy/%	Precision/%	Recall/%	F1-score/%
DT	57.96	58.01	60.57	55.27
DBSCAN+Rules	89.24	82.78	77.72	79.88
GCN	88.61	87.93	67.85	72.64
RF	78.13	71.67	73.78	72.52
XGBoost	68.46	64.03	67.76	64.11
MLP	84.08	88.89	69.30	72.95
FDRNet	96.88	92.56	94.58	93.54

Table 5 Overall performances of all methods on the wheat harvester trajectory dataset

Method	Accuracy/%	Precision/%	Recall/%	F1-score /%
DT	51.14	53.17	57.41	47.80
DBSCAN+Rules	70.84	63.75	70.68	63.90
GCN	87.95	87.64	71.39	75.87
RF	78.29	68.06	68.25	68.60
XGBoost	62.28	57.87	61.46	56.38
MLP	85.70	90.46	66.82	70.95
FDRNet	96.68	95.65	92.88	94.19

Table 6 Comparison of seven segmentation models on "field" and "road" categories on the corn harvester trajectory dataset

Method	Field			Road		
	Precision/%	Recall/%	F1-score/%	Precision/%	Recall/%	F1-score/%
DT	82.65	55.27	66.25	33.36	65.87	44.29
DBSCAN+Rules	91.81	95.45	93.60	73.75	59.98	66.16
GCN	88.72	98.92	93.54	87.15	36.79	51.73
RF	87.39	82.62	84.94	55.95	64.94	60.11
XGBoost	85.81	69.19	76.61	42.25	66.33	51.62
MLP	82.80	99.29	90.30	94.99	39.31	55.60
FDRNet	98.63	97.74	98.18	86.49	91.43	88.89

Table 7 Comparison of seven segmentation models on "field" and "road" categories on the wheat harvester trajectory dataset

Method	Field			Road		
	Precision/%	Recall/%	F1-score/%	Precision/%	Recall/%	F1-score/%
DT	82.14	48.51	61.00	24.17	60.92	34.61
DBSCAN+Rules	91.00	70.94	79.73	36.50	70.42	48.08
GCN	88.03	98.44	92.94	87.25	44.34	58.80
RF	87.16	84.97	86.05	48.96	53.53	51.14
XGBoost	85.38	62.89	72.43	30.35	60.04	40.32
MLP	84.86	99.62	91.65	96.06	34.03	50.25
FDRNet	97.19	98.81	97.99	94.12	86.96	90.40

The varied results in Tables 4-7 can be attributed to differences in their information extraction capabilities. DBSCAN+Rules is a density-based clustering algorithm that uses only longitude and latitude as features and fails to fully extract the feature distribution among trajectory points. The DT, RF, and XGBoost methods rely solely on individual kinematic features without effectively integrating information from different features, resulting in insufficient feature extraction. Additionally, these methods make independent judgments based on each feature, overlooking the correlation between features. An MLP, a common feedforward neural network model with fully connected layers, tends to learn shallow-level features and lacks the ability to extract complex, abstract high-level features effectively. A GCN constructs a spatiotemporal relationship graph using trajectory points, storing the connectivity between corresponding nodes in an adjacency matrix. However, this approach computes only the local spatiotemporal relationships between trajectory points, neglecting the extraction of global motion information. Subsequent networks heavily depend on the features extracted from the spatiotemporal graph. In contrast, the approach of this study extracts rich trajectory feature information from multiple perspectives, capturing the correlated information among different trajectory points comprehensively. Subsequently, EfficientNet-B0 is employed to extract associated features between different features within an individual trajectory point.

Moreover, the weights of the feature map channels are dynamically adjusted, and important features are automatically selected to enhance the model's perception of trajectory features. Therefore, this model can more effectively identify trajectories. Four sets of maize and wheat harvester trajectory data were randomly selected from outside of the training and testing sets as the observation examples. Figures 6 and 7, respectively present the classification results of the seven algorithms on these examples. The images use remote sensing satellite images as the base. The green dots represent trajectory points predicted as fields, and the red dots represent trajectory points predicted as roads. From the figures, it is evident that the proposed model can predict the trajectory point categories more accurately, demonstrating that this method can efficiently and accurately perform field-road trajectory classification.

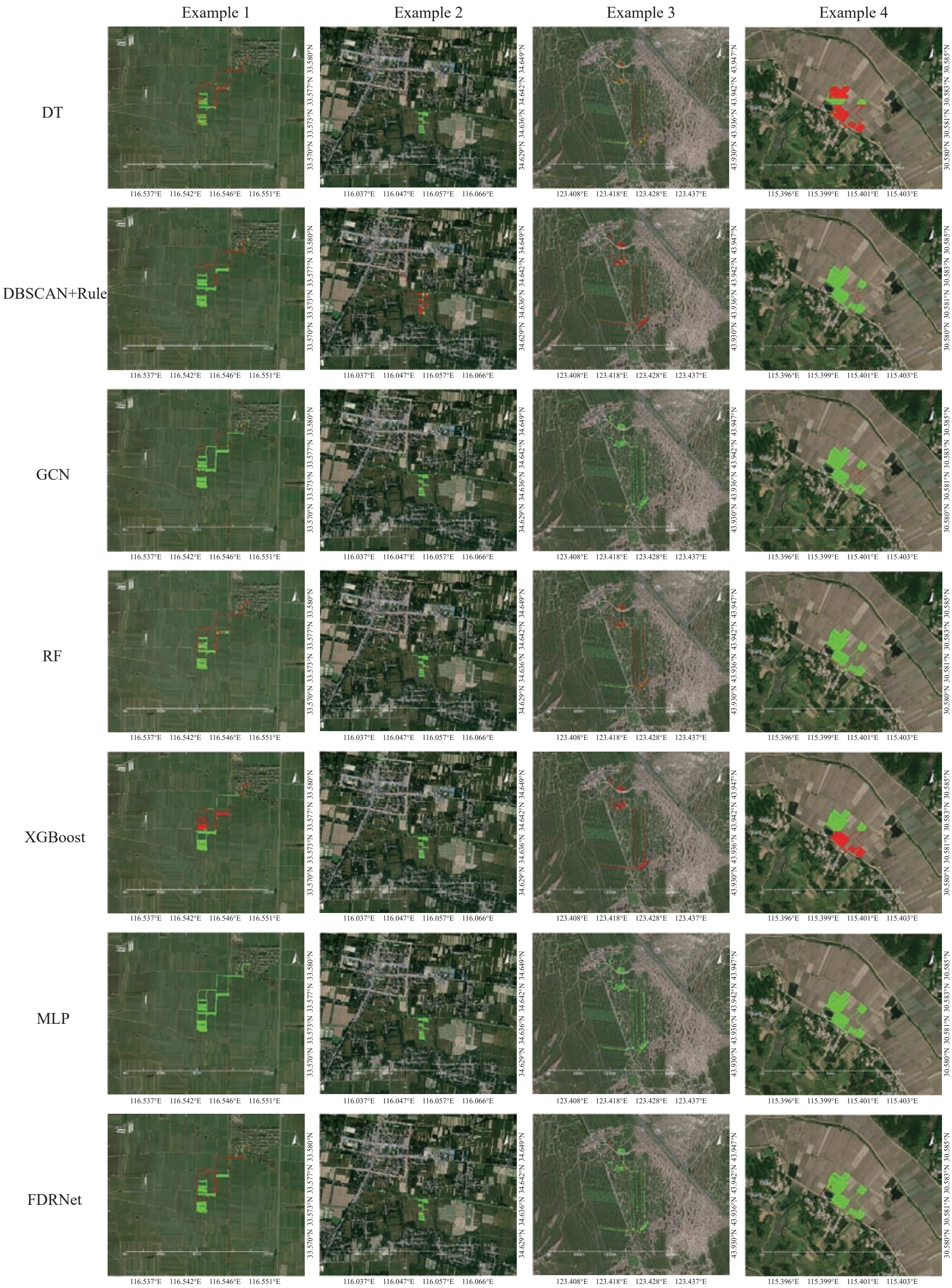

Figure 6 Field-road trajectory classification results for all models on the corn harvester trajectory dataset

3.4 Ablation experiments

The deformation feature module is the core part of FDRNet. It provides a new paradigm for CNN-based networks to directly handle time-series agricultural machinery trajectories. If this module is removed, FDRNet cannot work properly. Therefore, for the purpose of validating the effectiveness of FDRNet, we only excluded the multi-range feature enhancement module. To evaluate the impact of multi-range feature enhancement on the performance of the FDRNet model, ablation experiments were conducted on corn and wheat harvester trajectory datasets to assess the influence of multi-range feature enhancement (FE). The results are presented in Tables 8-11.

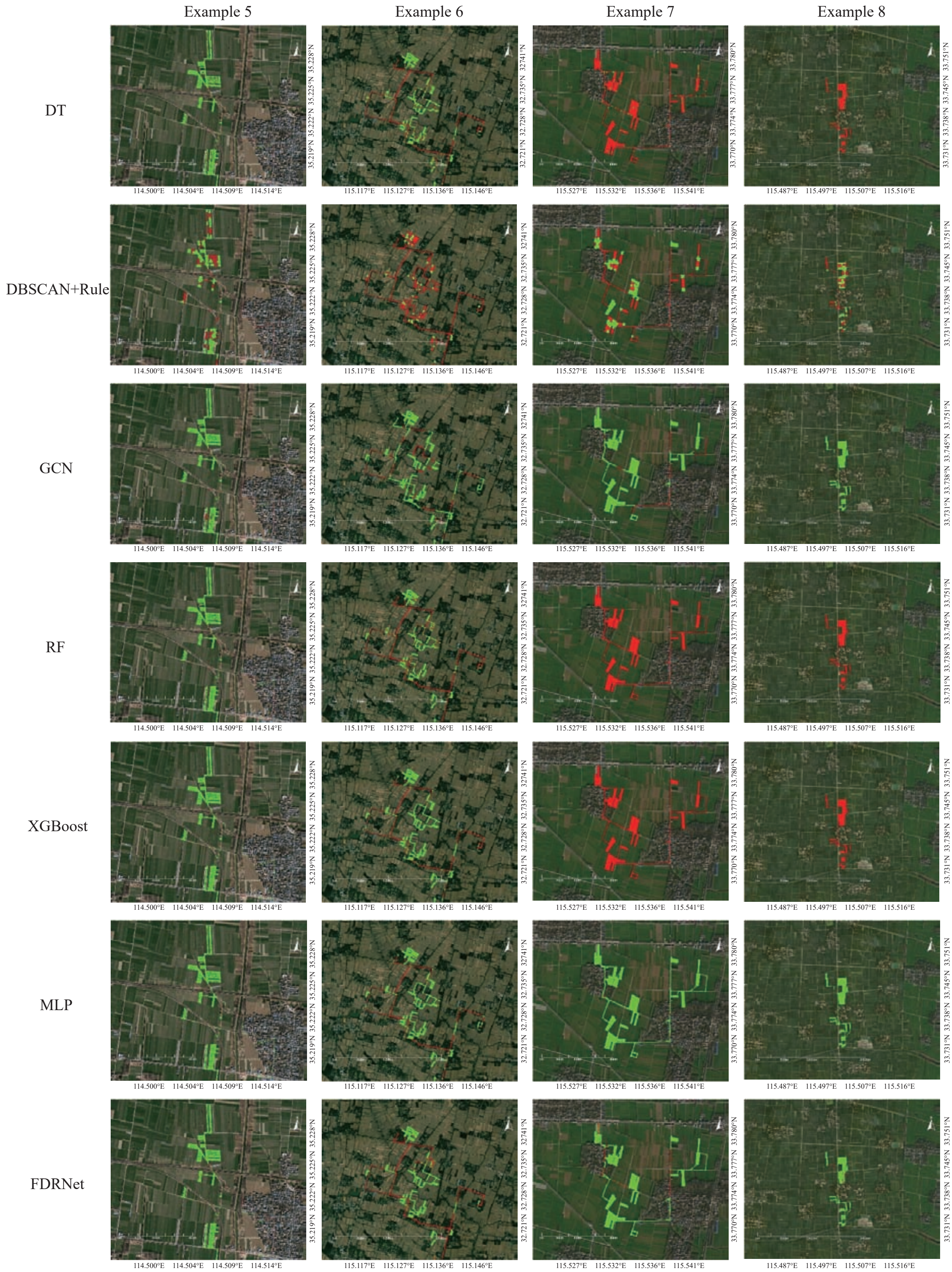

Figure 7 Field-road trajectory classification results for all models on the wheat harvester trajectory dataset

Table 8 Results of ablation experiments on feature-enhanced and balanced datasets on the corn harvester trajectory dataset

Method	Accuracy/%	Precision/%	Recall/%	F1-score/%
FDRNet	96.88	92.56	94.58	93.54
FDRNet(W/O)FE	79.69	63.72	68.47	65.25

Table 9 Results of ablation experiments on feature-enhanced and balanced datasets on the wheat harvester trajectory dataset

Method	Accuracy/%	Precision/%	Recall/%	F1-score/%
FDRNet	96.68	95.65	92.88	94.19
FDRNet(W/O)FE	84.57	73.28	70.97	72.01

The multi-range feature enhancement module enriches the number of trajectory features by incorporating short-range and long-range features. To verify its effectiveness, we removed the feature enhancement module, resulting in FDRNet(W/O)FE, which represents the model without the feature enhancement module. Table 8 shows that the addition of the feature enhancement module led to a 17.19% improvement in accuracy and a 28.29% improvement in the F1 score for the corn harvester trajectory dataset. Similarly, on the wheat harvester trajectory dataset (Table 9), the accuracy increased by 12.11%, and the F1-score improved by 22.18%. Tables 10 and 11 present the performance comparisons for the "field" and "road" categories. The addition of the feature enhancement module resulted in a more significant improvement in the F1-score for the "road" category. Specifically, for the corn harvester trajectory dataset (Table 10), there was a 46.03% improvement, and for the wheat harvester trajectory dataset (Table 11), there was a 37.15% improvement. This indicates that with a larger feature dimension, the model extracts richer motion information from agricultural machinery trajectories, enabling more effective identification of different trajectory categories, especially enhancing the performance of minority categories.

Table 10 The performances of ablation experiment on "field" and "road" categories on the corn harvester trajectory dataset

Method	Field			Road		
	Precision/%	Recall/%	F1-score/%	Precision/%	Recall/%	F1-score/%
FDRNet	98.63	97.74	98.18	86.49	91.43	88.89
FDRNet(W/O)FE	91.34	84.25	87.65	36.11	52.70	42.86

Table 11 The performances of ablation experiment on "field" and "road" categories on the wheat harvester trajectory dataset

Method	Field			Road		
	Precision/%	Recall/%	F1-score/%	Precision/%	Recall/%	F1-score/%
FDRNet	97.19	98.81	97.99	94.12	86.96	90.40
FDRNet(W/O)FE	89.61	91.94	90.76	56.96	50.00	53.25

In summary, the results of ablation experiments demonstrate the crucial role of feature enhancement in improving the performance of field-road classification models. By enhancing the model's feature dimensions from multiple perspectives, we are able to more accurately represent the feature distribution. This effectively addresses the challenge of achieving low classification performance in minority classes. As a result, the model's performance improved overall.

4 Conclusions

In this study, to effectively classify "field" or "road" points in agricultural machinery trajectories, a solution called the feature deformation network with multi-range feature enhancement (FDRNet) was proposed. FDRNet includes three key components: multi-range feature enhancement, feature deformation methods, and the EfficientNet-B0 network from the field of image classification. These elements work together to comprehensively extract features from agricultural machine trajectories. FDRNet tackles issues related to agricultural machine trajectories, including the insufficient exploration of spatiotemporal features, insufficient exploration of trajectory features, and the challenge of directly using convolutional neural networks to extract these features. This approach achieves accurate classification of agricultural machinery trajectory points into field points and road points. Experimental evaluations demonstrate that the model achieves 96.88% accuracy and 93.54% F1-score on the corn harvester trajectory dataset and 96.68% accuracy and 94.19% F1-score on the wheat harvester trajectory dataset, outperforming all the existing methods. This strongly validates the feasibility and superiority of the proposed approach for field-road trajectory classification tasks.

Due to various factors, such as the type of harvested crops, geographical location, time, weather, and agricultural machinery terminal manufacturers, the data exhibit various frequencies and feature distributions. Because of significant differences in the distribution of data features, existing models struggle to perform well across all data types. To achieve better results, we trained separate models for each distinct data distribution. However, the multitude and diversity of data types increase training costs and reduce the real-time processing of data. In future research, we plan to focus on developing a universal model for field and road classification. This model aims to learn the feature distributions of various data types, enabling accurate predictions of trajectory categories.

Acknowledgements

This research was financially supported by the National Natural Science Foundation of China (Grant No. 32301691), the National Key R&D Program of China and Shandong Province, China (Grant No. 2021YFB3901300), and the National Precision Agriculture Application Project (Grant/Contract number: JZNYYY001).

[References]

[1] Luan X D. Research on the application of artificial intelligence and computer technology in agricultural modernization. Computer Science, Agricultrual and Food Sciences, 2022; Corpus ID: 254579215.

[2] Guo T T, Wang Y F. Big data application issues in the agricultural modernization of China. Ekoloji Dergisi, 2019; 107: 3677–3688.

[3] Řezník T, Herman L, Klocová M, Leitner F, Pavelka T, Leitgeb Š, et al. Towards the development and verification of a 3D-based advanced optimized farm machinery trajectory algorithm. Sensors, 2021; 21(9): 2980.

[4] Sun Q C, Xia J H C, Foster J, Falkmer T, Lee H. Pursuing precise vehicle movement trajectory in urban residential area using multi-GNSS RTK tracking. Transportation Research Procedia, 2017; 25: 2356–2372.

[5] Quan Y M, Lau L. Development of a trajectory constrained rotating arm rig for testing GNSS kinematic positioning. Measurement, 2019; 140: 479–485.

[6] Liu L, Chen T, Yang S G, Wang X. Analysis on the mode of trans-regional allocation of agricultural machinery. American Journal of Plant Sciences, 2020; 7: 1049–1056.

[7] Li D, Liu X, Zhou K, Sun R Z, Wang C T, Zhai W X, et al. Discovering spatiotemporal characteristics of the trans-regional harvesting operation using big data of GNSS trajectories in China. Computers and Electronics in Agriculture, 2023; 211: 108003.

[8] Rodias E, Berruto R, Busato P, Bochtis D, Sørensen C G, Zhou K. Energy savings from optimised in-field route planning for agricultural machinery. Sustainability, 2017; 9(11): 1956.

[9] Liu Y, Shi L, Gao Y, Kou C Y, Yang S G, Liu L. Research on the optimized management of agricultural machinery allocation path based on teaching and learning optimization algorithm. Technical Gazette, 2022; 29(2): 456–463.

[10] Li H C, Gao F, Zuo G C. Research on the agricultural machinery path tracking method based on deep reinforcement learning. Sci Program, 2022; 2022: 6385972.

[11] Li S C, Xu H Z, Ji Y H, Cao R Y, Zhang M, Li H. Development of a following agricultural machinery automatic navigation system. Computers and Electronics in Agriculture, 2019; 158: 335–344.

[12] Yang Y, Li Y K, Wen X, Zhang G, Ma Q L, Cheng S K, et al. An optimal goal point determination algorithm for automatic navigation of agricultural machinery: Improving the tracking accuracy of the Pure Pursuit algorithm. Computers and Electronics in Agriculture, 2022; 194: 106760.

[13] Spooner P G. Minor rural road networks: values, challenges, and opportunities for biodiversity conservation. Nature Conservation, 2015; 11: 129–142.

[14] Kearney S P, Coops N C, Sethi S, Stenhouse G B. Maintaining accurate, current, rural road network data: An extraction and updating routine using RapidEye, participatory GIS and deep learning. International Journal of

Applied Earth Observation and Geoinformation, 2020; 87: 102031.

[15] Bhat S A, Huang N-F. Big data and AI revolution in precision agriculture: Survey and challenges. IEEE Access, 2021; 9: 110209–110222.

[16] Kamilaris A, Kartakoullis A, Prenafeta-Boldú F X. A review on the practice of big data analysis in agriculture. Computers and Electronics in Agriculture, 2017; 143: 23–37.

[17] Liakos K G, Busato P, Moshou D, Pearson S, Bochtis D. Machine learning in agriculture: A review. Sensors, 2018; 18(8): 2674.

[18] Lacour S, Burgun C, Perilhon C, Descombes G, Doyen V. A model to assess tractor operational efficiency from bench test data. Journal of Terramechanics, 2014; 54: 1–18.

[19] Lee J W, Kim J S, Kim K U. Computer simulations to maximise fuel efficiency and work performance of agricultural tractors in rotovating and ploughing operations. Biosystems Engineering, 2016; 142: 1–11.

[20] Chen B W, Dennis E J, Featherstone A. Weather impacts the agricultural production efficiency of wheat: The emerging role of precipitation shocks. Journal of Agricultural and Resource Economics, 2022; 47: 544–562.

[21] Chen Y, Zhang X Q, Wu C C, Li G Y. Field-road trajectory segmentation for agricultural machinery based on direction distribution. Computers and Electronics in Agriculture, 2021; 186: 106180.

[22] Poteko J, Eder D, Noack P O. Identifying operation modes of agricultural vehicles based on GNSS measurements. Computers and Electronics in Agriculture, 2021; 185: 106105.

[23] Xiao Y Z, Mo G Z, Xiong X Y, Pan J W, Hu B B, Wu C C, et al. DR-XGBoost: An XGBoost model for field-road segmentation based on dual feature extraction and recursive feature elimination. Int J Agric & Biol Eng, 2023; 16(3): 169–179.

[24] Chen Y, Li G Y, Zhang X Q, Jia J P, Zhou K, Wu C C. Identifying field and road modes of agricultural Machinery based on GNSS Recordings: A graph convolutional neural network approach. Computers and Electronics in Agriculture, 2022; 198: 107082.

[25] Zhang X Q, Chen Y, Jia J P, Kuang K M, Lan Y B, Wu C C. Multi-view density-based field-road classification for agricultural machinery: DBSCAN and object detection. Computers and Electronics in Agriculture, 2022; 200: 107263.

[26] Chen Y, Quan L, Zhang X Q, Zhou K, Wu C C. Field-road classification for GNSS recordings of agricultural machinery using pixel-level visual features. Computers and Electronics in Agriculture, 2023; 210: 107937.

[27] Zhai W X, Mo G Z, Xiao Y Z, Xiong X Y, Wu C C, Zhang X Q, et al. GAN-BiLSTM network for field-road classification on imbalanced GNSS recordings. Computers and Electronics in Agriculture, 2023; 216: 108457.

[28] Xu Y J, Liu X, Cao X, Huang C P, Liu E K, Qian S, et al. Artificial intelligence: A powerful paradigm for scientific research. The Innovation, 2021; 2(4): 100179.