



Article

Automatic Identification and Statistical Analysis of Data Steps in Electric Field Measurements from CSES-01 Satellite

Jianping Huang^{1,2}, Zongyu Li^{1,2}, Zhong Li^{2,*} , Wenjing Li¹, Livio Conti^{3,4} , Hengxin Lu¹, Na Zhou¹, Ying Han² , Haijun Liu², Xinfang Chen², Zhaoyang Chen², Junjie Song² and Xuhui Shen⁵

¹ National Institute of Natural Hazards, Ministry of Emergency Management of China, Beijing 100085, China; jianpinghuang@ninhm.ac.cn (J.H.); 21661329@st.cidp.edu.cn (Z.L.); liwenjing@ninhm.ac.cn (W.L.); luhengxin@ninhm.ac.cn (H.L.); zhouna@seis.ac.cn (N.Z.)

² Institute of Disaster Prevention, Langfang 065201, China; hanying@cidp.edu.cn (Y.H.); liuhaijun@cidp.edu.cn (H.L.); chenxinfang@cidp.edu.cn (X.C.); 21661315@st.cidp.edu.cn (Z.C.); 22661327@st.cidp.edu.cn (J.S.)

³ Faculty of Engineering, Uninettuno University, 00186 Rome, Italy; livio.conti@uninettunouniversity.net

⁴ INFN—Istituto Nazionale di Fisica Nucleare, Sezione Roma Tor Vergata, 00133 Rome, Italy

⁵ National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China; shenxuhui@nssc.ac.cn

* Correspondence: lizhong@cidp.edu.cn

Abstract: The spaceborne Electric Field Detector (EFD) is one of the payloads of the China Seismo-Electromagnetic Satellite (CSES-01), which can measure electric field data at near-Earth orbit for investigating fundamental scientific topics such as the dynamics of the top-side ionosphere, lithosphere-atmosphere-ionosphere coupling, and electromagnetic field emissions possibly associated with earthquake occurrence. The Extremely Low-Frequency (ELF) waveform shows anomalous step variations, and this work proposes an automatic detection algorithm to identify steps and analyze their characteristics using a convolutional neural network. The experimental results show that the developed detection method is effective, and the identification performance reaches over 90% in terms of both accuracy and area under the curve index. We also analyze the rate of the occurrence of steps in the three components of the electric field. Finally, we discuss the stability of the statistical results on steps and their relevance to the probe's function. The research results provide a guideline for improving the quality of EFD data, and further applications in monitoring the low-Earth electromagnetic environment.

Keywords: spatial electric field; data step; convolutional neural network; data quality; detection algorithm



Citation: Huang, J.; Li, Z.; Li, Z.; Li, W.; Conti, L.; Lu, H.; Zhou, N.; Han, Y.; Liu, H.; Chen, X.; et al. Automatic Identification and Statistical Analysis of Data Steps in Electric Field Measurements from CSES-01 Satellite. *Remote Sens.* **2023**, *15*, 5745. <https://doi.org/10.3390/rs15245745>

Academic Editors: Xuemin Zhang, Chieh-Hung Chen, Yongxin Gao and Katsumi Hattori

Received: 21 October 2023
Revised: 27 November 2023
Accepted: 5 December 2023
Published: 15 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

An increasing number of electromagnetic disturbances possibly associated with earthquakes has been discovered by using satellite observations both in case studies and with statistical analyses [1]. In this framework, a key problem is the cleaning and quality insurance of data in order to reduce spurious effects such as fluctuations of measurements not induced by the investigated events. For this purpose, an automatic procedure for electromagnetic field measurements will be particularly useful in terms of controlling artificial perturbations, as well as nonphysical background and electromagnetic noise. On the basis of CSES-01 measurements, Zhang et al. [2] found a correlation between the variation in multiple parameters, such as electric field, recorded by CSES-01 and the location of the seismogenic region of a specific Indonesian earthquake. Marchetti et al. [3] integrated CSES measurements (especially Ne) with those of other anomalies from the lithosphere and atmosphere, as well as Swarm satellite measurements, for large earthquakes in the Indonesian region. Liu et al. [4] Liu evaluated the data quality of satellite measurements of ionospheric parameters by comparing them with in situ observations and ionosonde data. Yang et al. [5]

analyzed used satellite data from the China Seismo-Electromagnetic Satellite to analyze the geomagnetic polarization of 12 typical earthquakes from December 2018 to January 2023. Wang et al. [6] investigated the electromagnetic effects of earthquakes and their potential impact on space weather through analysis of data from the CSES. Guo et al. [7] focused on the seismo-ionospheric effects related to two earthquakes in Taiwan, which were detected by the China Seismo-Electromagnetic Satellite. The results suggested that earthquakes in Taiwan and the surrounding region may have affected the ionosphere through geochemical, acoustic, and electromagnetic channels.

In addition, data quality verification is an important part of studies on seismo-electromagnetic monitoring from space, and is key to the data management process in satellite missions. The main differentiating factor of geophysical investigations performed by satellite with respect to other ground measurements is that the same observation cannot be made in situ, and that the method of data quality verification is relatively more complex. Many authors have conducted in-depth studies on the extraction of features of electric fields possibly associated with earthquakes, and on the procedures for cleaning data of possible electromagnetic interferences using combinations of cutting-edge technologies. Some researchers have attempted the extraction of anomalous electromagnetic disturbances using techniques based on artificial intelligence, such as applying deep learning directly to time series, or to spectrograms, etc. Kanarachos et al. [8] proposed a signal processing algorithm that combines wavelets, neural networks, and Hilbert transformation to detect anomalies in geoelectric field signals when investigating seismic precursors. Zeren et al. [9] conducted the exploration of time-corrected cross-calibrating methods using three payloads (EFD, SCM and HPM) onboard a CSES-01. Yan et al. [10] presented two types of regular features that were observed during LAP on board the CSES-01. The first feature is characterized by a sudden drop in plasma potential and floating potential data, while the second one manifests as a spike in the dayside plasma potential and floating potential data. Wang et al. [11] employed data from five ionosonde stations and one incoherent scatter radar observatory to validate the radio occultation measurements obtained by the CSES satellite. Yuan et al. [12] applied algorithms for the automatic recognition of lightning whistler acoustic waves on search-coil magnetometer (SCM) data from CSES-01; designed fuzzy and L morphology convolution kernels to identify the characteristics of spectral and L morphological features of whistlers; and used the SVM classifier to perform feature classification. Yan et al. [13] demonstrated the reliability of in situ plasma parameters derived from the China Seismo-Electromagnetic Satellite. Chen et al. [14] showed that CSES-01 Ne data very effectively reflect solar activity, as the trend in the former is highly correlated with the trend of variations in sunspot numbers. Liu et al. [15] investigated potential seismic anomalies related to the Luding Ms6.8 earthquake that occurred on 5 September 2022 by analyzing ionospheric, infrared radiation, atmospheric electrostatic field, and hot spring ion data, the results suggested that the observed multi-sphere coupling anomalies were associated with the occurrence of the earthquake. Diego et al. [16] proposed a direct quantitative validation method based on CSES measurements of plasma parameters and the geomagnetic field. Zhao et al. [17] established the seismogenic mechanism of ELF electromagnetic waves emitted by earthquakes, using ground-based and satellite observations. Yan et al. [18] analyzed the correlation between electron Ne and Te in the topside ionosphere, utilizing in situ measurements obtained from four satellites (CSES-01, Swarm A and B and CHAMP).

The data gathered by CSES-01 are widely used, and the electric field data are full of information, including not only possible earthquake precursors but also indications of other disturbances of the near-Earth electromagnetic environment. Yang et al. [19] developed a complete geomagnetic model for equatorial regions using CSES-01 data, which provided valuable insights into the Earth's magnetic field in these areas. In another study, Ghamry et al. [20] reported the first detection of Pi2 pulsation using CSES-01, which has implications for understanding plasma wave propagation in the Earth's magnetosphere. Furthermore, Gou et al. [21] carried out an examination of plasma bubbles using the

multiparametric payloads of CSES, which revealed new information about their formation and dynamics. Marchetti et al. [22] explored the effects of volcanic eruptions on the ionosphere using CSES-01 data, shedding light on the complex interactions between the Earth's atmosphere and its magnetic field. In addition to these examples, Spogli et al. [23] investigated geomagnetic storms using CSES-01 data, contributing to our understanding of space weather phenomena and their impact on Earth's magnetic environment. Compared with classical machine learning algorithms, deep learning achieves higher recognition accuracy and a greater generalization ability. Among them, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are the two most commonly used and relevant [24]. CNN is a deep learning neural network, simulating the learning mechanism of the human brain in order to ensure high accuracy of training samples and test samples, improve the efficiency of image labeling, and facilitate the management and updating of the image classification system. Therefore, in the present work, a CNN is proposed to identify and classify the steps that appear in waveform data of the electric field measured by the EFD payload of the CSES-01 in order to improve the quality of the data and increase the applicability of electric field measurements, thus providing reliable observations for further analyses including those studying seismo-electromagnetic emissions.

2. Overview of EFD

On 2 February 2018, CSES-01 was launched successfully. Its main objectives are to monitor the near-Earth space environment and to investigate possible EM perturbations. The satellite undertakes a circular Sun-synchronous orbit at an altitude of approximately 500 km, LTDN at 14:00, with a designed lifetime of 5 years [25].

The Electric Field Detector (EFD) is designed for the measuring electric field in the space plasma environment, and is one of the main payloads of CSES-01. EFD adopts the active double-probe detection principle and is conceived for worldwide coverage, thus providing accurate and complete monitoring when studying electromagnetic perturbations possibly associated with the occurrence of earthquakes [26]. The electric field is calculated from the difference in measured potential between two EFD probes. The EFD instrument is composed of four deployable booms, each of them carrying on their tip a 60 mm-diameter spherical probe (labeled "A", "B", "C", and "D", respectively) forming a tetrahedral structure. The {X, Y, and Z} satellite coordinates system is defined as follows: the X axis is parallel to the velocity vector of the satellite, Z points in the direction of the nadir, and Y points towards the point of intersection of Z and Y [27], as shown in Figure 1.

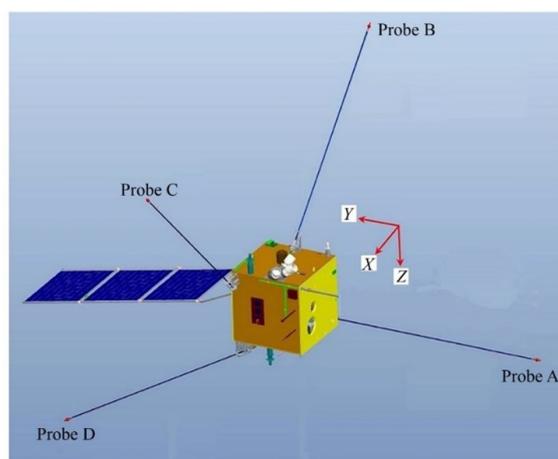


Figure 1. Spatial distribution of EFD spherical probes and extension rods.

The intensity of the electric field vector is obtained by measuring—for each of the four probes—the difference of electric potential between the surface of a probe and the circuit ground, calculating the three differences of electric potential (dop) between pairs of the four probes and then dividing the dop between couples of probes for the distance between

them. The EFD is designed to measure the intensity of the electric field components in the broad frequency range from DC to 3.5 MHz, subdivided into 4 channels, which are frequency bands, named: ULF (DC–16 Hz), ELF (6 Hz–2.2 kHz), VLF (1.8 kHz–20 kHz) and HF (18 kHz–3.5 MHz). According to the different data sampling rates at ULF, ELF, VLF and HF, the working period (WP) normally lasts 247.808 s. It is equally divided into 121 sub-working periods, each one lasting 2.048 s. The first of the 121 sub-working periods is called the bias current-corrected period (BCP). The other 120 2.048 s sub-working periods are called sampling periods (SP), and each SP is divided equally into 50 sub-sampling periods (SSP) of 40.96 ms. The data products from DPU are different in each of the different frequency bands. For the ELF frequency band, the sampling rate is 5 kHz and the outputs are each probe's voltage values [27].

3. Convolutional Neural Network Model

Convolutional Neural Network is the one of most common and interesting multilayer architectures used in deep learning. Compared with traditional models, CNNs are suitable for extracting feature information from images and audio recordings with high accuracy in recognition and classification [28,29] for image classification, target detection and behavior recognition, semantic segmentation, etc. Specific applications often require networks with different structures such as AlexNet, LeNet, VGGNet and ResNet [30–33].

In a CNN, the convolutional layer is the layer one with the role of extracting various features from the input, increasing the dimensional features by enhancing the useful ones, and reducing the effect of noise. The process of convolution is implemented through a convolution kernel, also known as a filter, which represents a feature that can be computed by sliding it over the input image by means of a sliding window, multiplying the kernel elements by the elements in the input region, and integrating in order to obtain the features of the input data [34,35], as follows:

$$x_{m_i, n_i} = \sum_{a=1}^k \sum_{b=1}^q \left[\left(x_{m_{i-1}+a, n_{i-1}+b} \right) \cdot w_{a,b} \right] \quad (1)$$

where x_{m_i, n_i} is the element of the m -row, n -column of the i -layer whose elements are being convolved, $x_{m_{i-1}+a, n_{i-1}+b}$ denotes the element of the row $m + a$, column $n + b$ of the $i - 1$ input layer, and $w_{a,b}$ presents the weight of the a -row, b -column of the kernel.

A convolutional layer usually contains multiple kernels. By computing the input matrix with different kernels, different features of the input data can be extracted. To increase the representational power of the model, bias terms are usually also added after the convolutional computation and the results are fed into a nonlinear activation function f for nonlinear mapping in order to extract the features X_i of the data, with the following equation:

$$X_i = f \left(b_i + \sum_{k=1}^{c_i} X_{i-1}^k \cdot W_i^k \right) \quad (2)$$

where b_i is the bias term of the i layer, c_i is the number of channels of the i layer, W_i^k is the weight matrix of the i layer of the k channel, which is the convolution kernel, and X_{i-1}^k is the output matrix of the $i - 1$ layer of the k channel.

To improve the accuracy of image feature acquisition and learning, the number of layers of convolutional neural networks usually needs to be deepened. Generally, the pooling layer is placed after the convolutional layer to reduce the amount of data processing and retain feature information by dimensioning down the convolutional output results to improve the generalization ability of the model.

The fully-connected layer acts as a “classifier” in a CNN, integrating features from the convolutional and pooling outputs and outputting classification results through the output layer. In the image classification task, the fully connected layer determines the class of the image based on the features extracted from the previous convolutional layers and pooling

layers. The fully connected layer reshapes and integrates the two-dimensional feature map to produce the final output results. As for the output layer, a logistic function or a normalized exponential function is used to activate and release the output predicted values.

4. Experiment Design

4.1. Dataset Selection and Pre-Processing

In this article, we analyze the waveform of an electric field in the ELF band. In order to study the steps occurring in the electric field data, we consider 50 samples of an SSP data packet together with 50 samples of the previous SSP to obtain the image of the plot to be analyzed with the CNN algorithm. There are two types of steps: a straight-down step (that is a sudden and sharp decrease) and a smooth or gradual step, as shown in Figure 2.

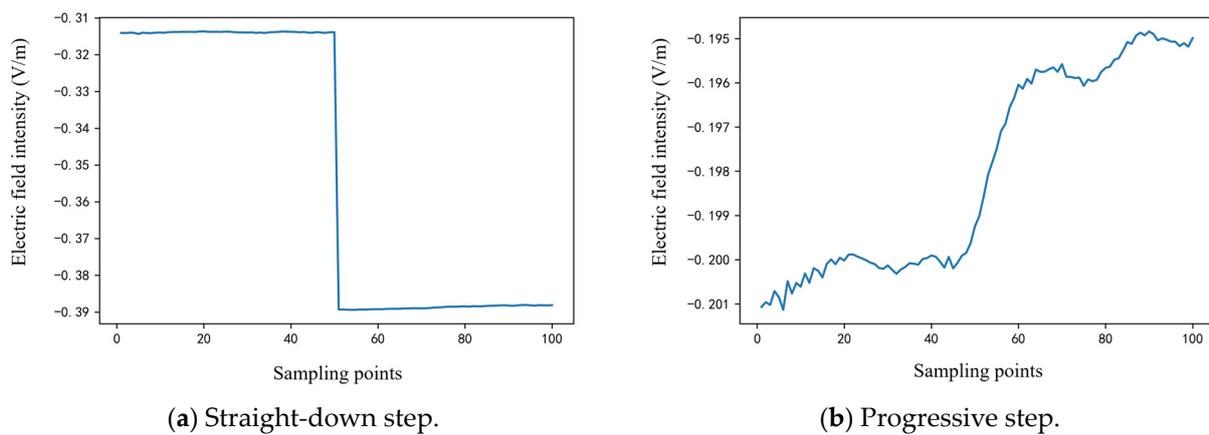


Figure 2. Examples of straight-down step, with a sharp decrease, (a) and progressive step, with a smooth increase, (b) in the waveform of one of three electric field components observed by EFD.

In signal processing data analysis, it is often necessary to segment data into different sections. For two adjacent data packets, A and B, we can use the following mathematical definitions to determine the types of connection between them:

Let M_A and M_B be the mean values of data packets A and B, respectively, while S_A and S_B are their variances, and K_A and K_B are their respective slopes.

- (1) If $|M_A - M_B| > 3S_A$, then the type of step connection between A and B is “straight-down”.
- (2) If $|M_A - M_B| \leq 3S_A$:
 - (i) If $K_A \times K_B > 0$, then the type of step connection between A and B is “progressive”;
 - (ii) If $K_A \times K_B < 0$, then the type of step connection between A and B is “straight-down”.

4.2. CNN-Based Method for Detecting Steps Anomalies

The flow chart of the recognition and classification algorithms for steps, based on CNN, is shown in Figure 3. The main blocks of the process are the loading of the electric field waveform data; the initialization and hyperparameter tuning optimization of CNN; model training and testing visualization; algorithm evaluation; the automatic recognition of electric field steps, and statistical analysis.

The images fed to the CNN algorithm are plots of waveform obtained from the EFD experiment, which are divided into two categories: straight-down and progressive. Each class includes 1000 samples with 900 training images and 100 test images. These samples were manually screened and identified from the ELF band during the period of November 2019 to November 2020. Specifically, these samples include both straight-down and progressive types of data steps.

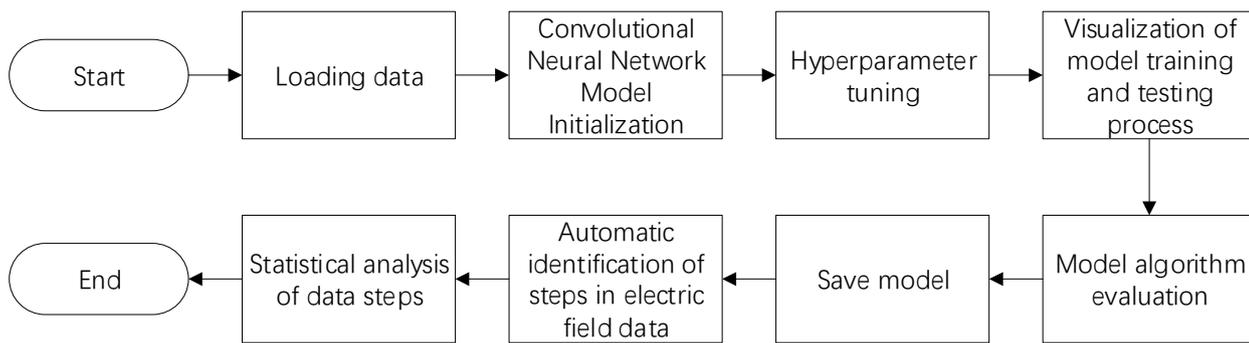


Figure 3. Flowchart of the recognition, classification and statistical analysis algorithms based on CNN.

The training process of the CNN model is as follows:

- (1) The fully connected layer of the model is applied to the dataset, and the output layer is replaced with a Softmax layer and an additional Dropout layer to avoid overfitting;
- (2) Set appropriate values of hyperparameters according to the experimental conditions and dataset size. For example, set the value of batch_size to 32, the value of epoch to 50, and the value of initial learning rate to 0.001;
- (3) Input the modified network model into the dataset and retrain it to get the new weights;
- (4) Save the trained model and weights.

The testing process of the CNN model is as follows:

- (1) Use the trained network model weights;
- (2) Start the program to run the test sample through the CNN model layer by layer and output the results;
- (3) Compare the output of the CNN model with the labels of the test samples, determine whether the output category of each image is correct, and perform statistics on the classification results;
- (4) Repeat steps (2)~(3) until all images in the test set have been tested and classified.

In our study, we employed standard kernels with dimensions of 5×5 and 3×3 . Specifically, we utilized 96 5×5 kernels in the initial and second layer, and finally incorporated 256 3×3 kernels in the third layer. To preserve the spatial resolution of the input images while extracting hierarchical features through convolutional layers, we employed zero-padding to ensure consistent spatial dimensions of the input and output feature maps throughout the network. This allowed us to maintain the original dimensions of the input images as they progressed through the network, while still enabling the extraction of increasingly abstract representations at each subsequent layer. In order to optimize the model, we utilized cross-validation during training to determine the optimal combination of hyperparameters. We calculated the average accuracy of the model on the test set and selected the hyperparameter combination that yielded the highest accuracy as the final training parameters for the model. Table 1 shows the hyperparameters and optimization methods.

Table 1. The hyperparameters and optimization methods.

Hyperparameter	Value	Optimization Method
Learning rate	0.001	Stochastic Gradient Descent (SGD)
Batch size	32	SGD
Momentum	0.5	SGD
Weight decay	0.0005	L2 regularization
Dropout rate	0.5	Dropout regularization

Due to the unbalanced number of samples in the categories of the dataset, which include many normal samples and few anomalous ones, the evaluation metrics of anomaly

detection are generally more complex and cannot be calculated only for accuracy and loss values. The most commonly used evaluation index is the Receiver Operating Characteristic curve (ROC curve), which is more objective than accuracy and can characterize the results more comprehensively.

4.3. Analysis of Experimental Data

In training the CNN, the images in the dataset are first of all divided and labeled in categories. In the training process, images are used as input to the CNN for iterative training to shorten the training time. The accuracy and loss values of the training and test sets are shown in Figure 4.

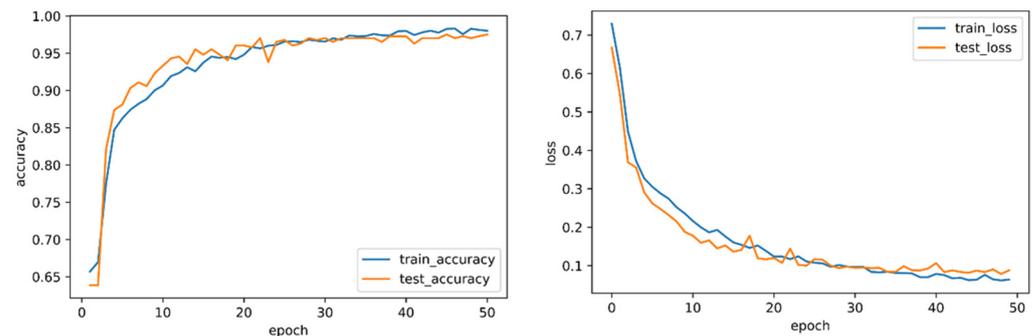


Figure 4. Example of accuracy (**left panel**) and loss values (**right panel**) for the training and test sets.

With the augmentation of iterations, there is a noticeable enhancement in the accuracy of both the training and test sets. With a number of iterations of approximately 50, the rate of improvement in the accuracy for both sets gradually stabilizes. When the training phase is complete, the accuracy of the model reaches 95.2% on the training set and 91.1% on the test set, while the value of the loss function is less than 0.1 on both training and test sets.

To verify the effectiveness of the developed classification algorithm, under different parameters, we have adopted the ROC curve to calculate the accuracy of each category. The abscissa of the ROC curve is False Positive Rate (FPR), and the ordinate is True Positive Rate (TPR), such that the ROC curve describes the equilibrium state of the classifier between TPR and FPR. The ROC curve has the important property that when the distribution of positive and negative samples in the test set changes, the ROC curve can remain unchanged, i.e., it is insensitive to the positive and negative sample imbalance problem. The closer the ROC curve is to the upper left corner, the better the performance of the classifier is. If the classifier's performance is evaluated by the ROC curve, we can use the area under curve (AUC) metric, which is the area under the ROC curve, and takes values no more than 1. If a positive sample and a negative sample are randomly selected, the AUC characterizes the probability that these two samples are correctly distinguished. An example of ROC curve is shown in Figure 5.

The ROC curves of Figure 5 show that the CNN algorithm achieved high accuracy in the classification with AUC metrics higher than 90%. Therefore, the trained model can be applied to identify the steps of the electric field waveform in EFD data.

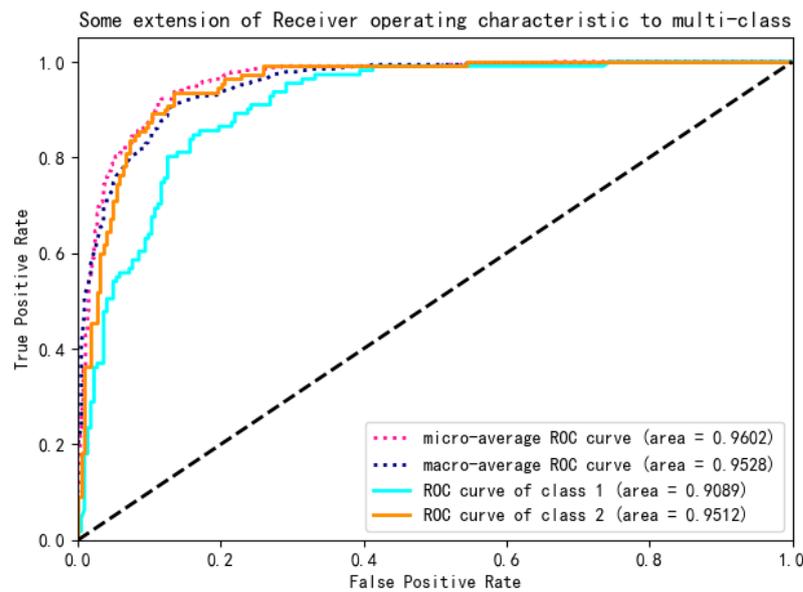


Figure 5. Example of ROC curves.

5. Experimental Results

CSES-01 adopts a near-circular Sun-synchronous orbit, and a complete orbit is divided into ascending (nighttime acquisition, at about 02:00 LT) and descending (daytime acquisition, at about 14:00 LT) semi-orbits for the satellite path from south to north latitudes, and vice versa. The CSES-01 revisit time is 5 days. In this article, we study the data of CSES-01 during a seismically quiet period. The orbit number ends with the digit 0 for descending and 1 for ascending, as per our analysis of the ELF waveform data collected from CSES-01.

To evaluate the performance of our model, we first analyzed four complete orbits during periods unaffected by external disturbances. Next, we extended our analysis to cover the revisit cycle of the CSES, which includes approximately 76 orbits, providing global-scale coverage. This allowed us to understand the regional characteristics of the data and further validate our model's accuracy during these quiet periods. The results are presented in detail in Sections 5.1 and 5.2.

5.1. Statistical Analysis of Steps in Ascending Semi-Orbits

The electric field data in the ELF band analyzed in this article are constituted by the waveforms of the three components (E_x , E_y , and E_z) evaluated on the basis of the differences of potential measured between the three pairs of probes ab, cd, and ad, respectively.

We selected the four ascending semi-orbits 26688, 27143, 27163, and 27180 in the period 16–20 December 2022 and conducted a further analysis on the ascending data within this revisit cycle. The results are shown in Figure 6, where we provide the distribution of steps in each single semi-orbit (without distinguishing the step type) (Figure 6a) and the distribution of straight-down and progressive steps for each electric field component, accumulated over the revisit cycle (Figure 6b).

According to the statistical results shown in Figure 6a, it can be found that the four selected ascending semi-orbits show similar distribution between the E_x , E_y , and E_z components, while Figure 6b shows that in all components the number of progressive steps is higher than that of straight-down steps. However, there are some differences: in the 26688 orbit, the total number of both types of steps is 108 (with 25.93% of the straight-down type and 74.07% of the progressive type); in the 27143 orbit, the total number of both types of steps is 86 (with 39.53% of the straight-down type and 60.47% of the progressive type); in the 27163 orbit, the total number of both types of steps is 80 (with 45% of the straight-down type and 55% of the progressive type); finally, in the 27180 orbit, the total

number of both types of steps is 86 (with 35.29% of the straight-down type and 64.71% of the progressive type).

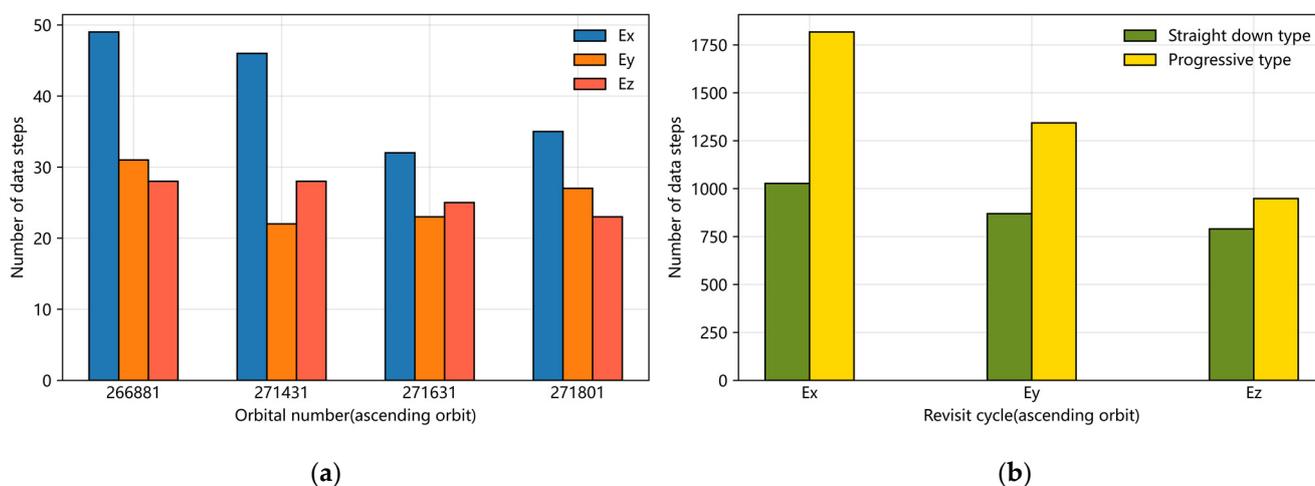


Figure 6. We show the distribution of steps detected during the four selected ascending tracks (26688, 27143, 27163 and 27180). (a) In the left panel the distribution of steps (without distinguishing the type) is provided for each orbit as a function of the Ex, Ey, and Ez components. (b) In the right panel, we show the distribution of the straight-down and the progressive steps for each electric field component accumulated over the revisit cycle of 16–20 December 2022.

Figure 6b shows that (i) the numbers of steps over Ex, Ey, and Ez components are similar, and (ii) the numbers of steps of both the straight-down and progressive types in the Ex component are higher than those in the Ey and Ez components. The total number of steps (of both types) in the ascending orbit of the revisit cycle is 6794, and the fraction of steps in the Ex component is 41.86%, in the Ey component it is 32.56%, and in the Ez component it is 25.58%.

5.2. Statistical Analysis of Steps in Descending Semi-Orbits

Based on the descending semi-orbits 26688, 27143, 27163, and 27180 in the period 16–20 December 2022, we conducted a further analysis on the descending data within this revisit cycle. The results are shown in Figure 7, where we provide the distribution of steps in each single semi-orbit (without distinguishing the step type) (Figure 7a) and the distribution of straight-down and progressive steps for each electric field component, accumulated over the revisit cycle of 16–20 December 2022 (Figure 7b).

Figure 7a shows that, as in the ascending semi-orbits, in the descending ones, the number of steps shows a similar distribution in each of the Ex, Ey, and Ez components, and the number of steps in the Ex component is again the highest. In addition, the number of progressive steps is higher than that of straight-down ones, but there are some differences between different orbits. For example, in the 26688 orbit, the total number of the two types of steps is 156 (with 27.56% of the straight-down type and 72.44% of the progressive type); in the 27143 orbit, the total number of two types of steps is 128 (with 41.41% of the straight-down type and 58.59% of the progressive type); in the 27163 orbit, the total number of both types of steps is 157 (of which the percentage of straight-down type is 41.40%, while the percentage of progressive type is 58.60%); finally, in the 27180 orbit, the total number of both types of steps is 125 (of which the percentage of straight-down type is 42.40%, while the percentage of progressive type is 57.60%).

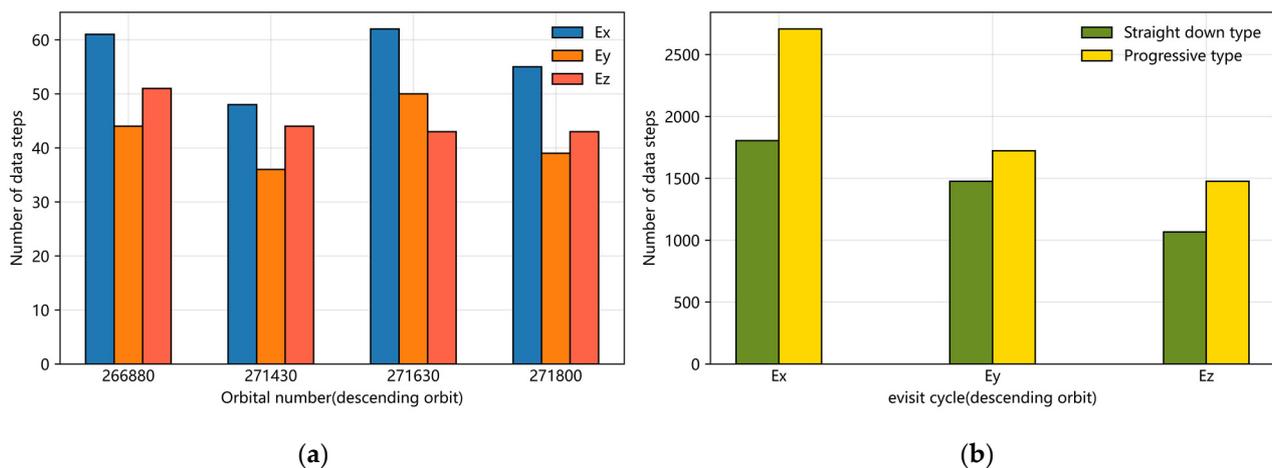


Figure 7. We show the distribution of steps detected during the four selected descending tracks (26688, 27143, 27163 and 27180). (a) In the left panel the distribution of steps (without distinguishing the type) is provided for each orbit as a function of the Ex, Ey, and Ez components. (b) In the right panel, we show the distribution of the straight-down and the progressive steps for each electric field component accumulated over the revisit cycle of 16–20 December 2022.

The statistical distribution of steps in descending semi-orbits is shown in Figure 7b: the numbers of steps in the three components have similar distributions, with the Ex component showing the greatest difference between the numbers of the two step types and the highest absolute numbers of steps of both types with respect to the Ey and Ez components. The total number of steps of both types in the descending semi-orbits of the revisit cycle is 10,250, with 44.00% occurring in the Ex component, 31.20% in the Ey component, and 24.80% in the Ez component.

In summary, the number of step anomalies in the Ex component of the ELF electric field waveform is the highest. The statistical distributions of steps in different components show similar trends. These results need to be further explored to better understand the nature of this phenomenon. These findings are relevant to developing a deeper understanding of the operational state of the CSES-01 satellite, as well as of the data acquisition and data quality insurance processes.

6. Discussions

6.1. Stability of Statistical Results of Detected Steps

In this analysis, we selected ELF waveform data collected from EFD during a quiet period. The purpose was to avoid the influence of anomalous steps caused by seismic activity. A seismically quiet period refers to a specific time frame characterized by minimal to no earthquake events. More specifically, we define a seismically quiet period as that without any recorded earthquake events with a magnitude greater than or equal to 5.0 on the Richter scale. The data were collected in four CSES-01 descending and four ascending semi-orbits, as shown in the top- and bottom-left panels, respectively, of Figure 8.

In the descending orbits, it can be seen from the ratio of the number of progressive types to the number of straight-down types that the steps in the Ex and Ez components show similar behaviors—decreasing first and then increasing, while the ratio of Ey components is continuously decreasing. In addition, the ratio of the Ex component is the highest among the three components, indicating that the Ex component has the clearest trend in the descending semi-orbits. A similar trend is observed for the ascending semi-orbits, where all the three components show a consistent trend of decreasing and then increasing, with the Ex component having the highest ratio. By statistically analyzing the data steps during the quiet period, we find that the results of the data step analysis for individual tracks show consistency with those of the experimental part of this paper, further verifying the stability of the experimental results.

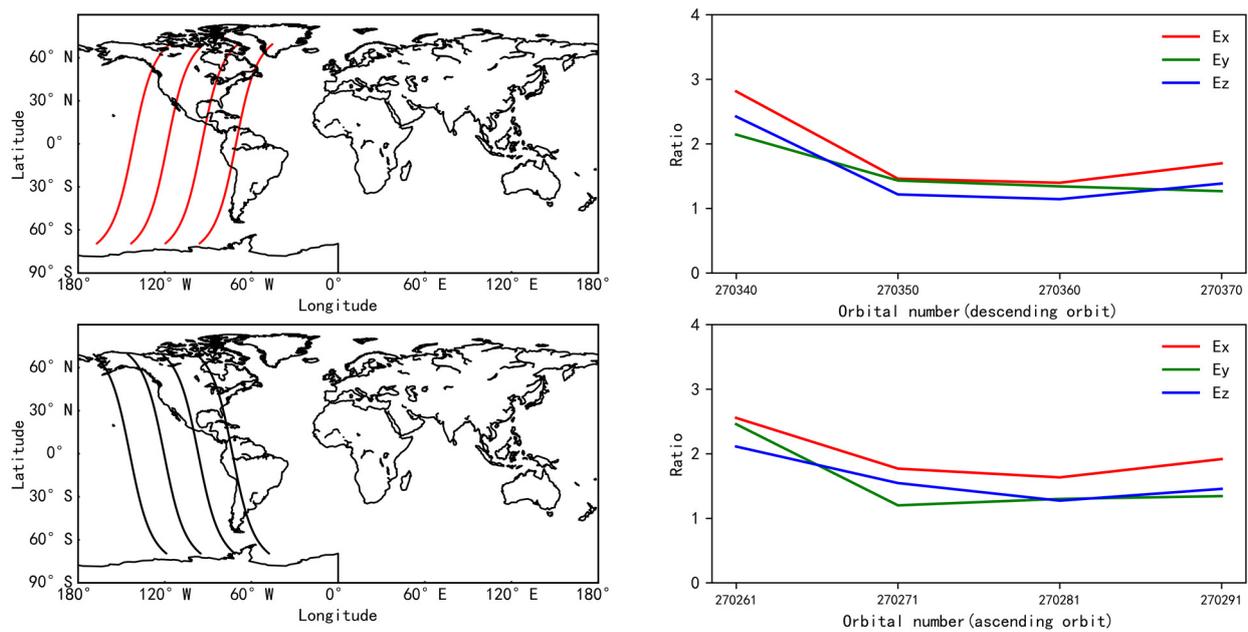


Figure 8. In the left column we show the spatial position of the analyzed descending (**top panel**) and ascending (**bottom panel**) semi-orbits. In the right column, the statistics for steps detected in descending (**top panel**) and ascending (**bottom panel**) semi-orbits are provided. The ratio refers to the proportion of progressive types to straight-down types in the data; the figure shows a graph of the trend and characteristics of three components.

6.2. Steps in Single Probe Signal and in Electric Field Components

The EFD measurements of electric field are obtained by dividing the potential difference between pairs of probes (a, b, c and d) by the distance between the probes. We analyzed the steps detected in the waveforms of each of the four probes a, b, c, and d in order to quantify their performance. The number of steps detected during the revisit period from 16 to 20 December 2022 is summarized in Figure 9. The results show that the steps occurred more frequently in probe b, which has a higher noise level than the other sensors. By analyzing the statistical distribution of the steps data, we obtained a more refined analysis of the EFD performance, which is useful for studying the noise characteristics of the probes, and for their optimization. It is worth noting that, even though a difference in the performance of probe b was found, the exact reason for this behavior is not clear and needs to be further investigated.

Since probe b is at the end of the boom in the satellite flight direction, that is, the X direction in the satellite coordinates system, and the step anomalies are more frequent in the Ex component waveform data, we can hypothesize that the effect on Ex is due to the measurements from the b sensor that enter into the calculation of Ex. The interaction between spacecraft and space environment is important for all space missions because the spacecraft perturbs the plasma around it during its motion, causing a spacecraft charging and wake effect. When the motion of the plasma relative to the spacecraft is greater than the ion thermal velocity and less than the electron thermal velocity, it is difficult for the ions to reach the spacecraft directly downstream, but relatively easy for the electrons to enter the tail, and a thin trailing region of ions will form downstream of the spacecraft, i.e., the trailing phenomenon. The trails are negatively charged, and a negative potential appears in the trails thus forming an electric field [36]. The spacecraft wake effect can cause perturbations in the surrounding plasma, which can affect the accuracy of the measurement devices. Therefore, future research work can be addressed at removing the impact of satellite motion in the plasma from the measurements taken by sensor b.

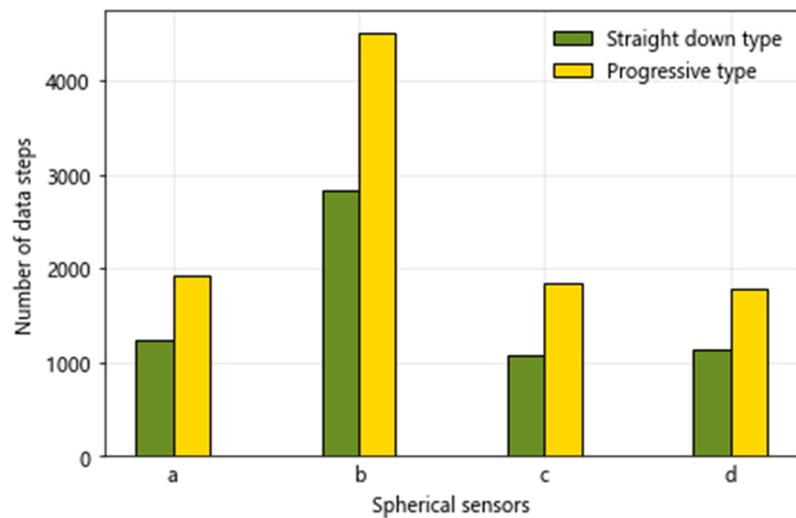


Figure 9. Spherical sensor data step statistics chart.

7. Conclusions

This paper presents the first attempt at developing a method to automatically identify the occurrence of steps in electric field waveforms detected by EFD in ELF bands using CNN. Firstly, the EFD waveform data have been selected and pre-processed; secondly, the dataset is modeled and the model parameters are adjusted according to the test set results to obtain the best-performing model; finally, the data step anomalies are detected and identified automatically based on the model's results. The experimental results show that, when using the CNN-based step anomaly detection algorithm, the accuracy and AUC index are higher than 90%.

The statistical analysis of the step anomalies in the waveform data shows that the most significant step anomalies are found in the Ex component. This study can provide an effective method for the automatic processing of waveform data and anomaly detection in an electric field. In order to avoid the influence of anomalous steps caused by seismic activity, by statistically analyzing the data steps during the quiet period, we find that the results of the data step analysis for individual tracks show consistency with those in the experimental part of this paper, further verifying the stability of the experimental results. Since the spherical sensor b—that is, the one most affected by step anomalies—is at the end of the boom in the direction of the satellite's flight, and the step anomalies are more frequent in the Ex component waveform's electric field component, we can hypothesize that the step anomaly is a trailing effect.

This study further demonstrates the importance of data quality assessment for increasing the effectiveness and optimizing the applicability of electric field data from EFD. However, it should be noted that the network model needs to be updated regularly according to the improvement of the classification system due to its relatively weak generalization ability. Meanwhile, it is also necessary to further explore how to improve the image recognition effect by refining the image sample design. The application of artificial intelligence methods to analyze EFD waveform data when detecting anomalies and classification problems has also been implemented, and this study provides support in exploring and applying artificial intelligence data analysis techniques in training algorithms for the identification and classification of typical phenomena. It also provides useful tools and methods for subsequent research on space electric field data derived via in earthquake monitoring from space environments.

Author Contributions: Conceptualization, J.H.; methodology, Z.L. (Zhong Li); software, Z.L. (Zongyu Li), Z.C. and J.S.; validation, Z.L. (Zongyu Li); formal analysis, H.L. (Haijun Liu), Y.H. and X.C.; investigation, W.L. and H.L. (Hengxin Lu); resources, X.S. and J.H.; data curation, N.Z. and H.L. (Hengxin Lu); writing—original draft, J.H. and Z.L. (Zongyu Li); writing—review and editing, J.H. and L.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the International Space Science Institute (ISSI in Bern, Switzerland and ISSI-BJ in Beijing, China), supporting International Team 23-583 led by Dedalo Marchetti and Essam Ghamry.

Data Availability Statement: CSES data are freely available upon registration and approval on the CSES web portal: www.leos.ac.cn (accessed on 1 June 2023).

Acknowledgments: This work made use of data from observations of the China Seismo-Electromagnetic Satellite supported by the China National Space Administration and the China Earthquake Administration (<https://www.leos.ac.cn>) (accessed on 1 May 2023). We thank Jungang Lei of the Lanzhou Institute of Physics, China, for his guidance and comments during the research process of this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chen, H.; Han, P.; Hattori, K. Recent Advances and Challenges in the Seismo-Electromagnetic Study: A Brief Review. *Remote Sens.* **2022**, *14*, 5893. [[CrossRef](#)]
2. Zhang, X.; Wang, Y.; Boudjada, M.Y.; Liu, J.; Magnes, W.; Zhou, Y.; Du, X. Multi-Experiment Observations of Ionospheric Disturbances as Precursory Effects of the Indonesian Ms6.9 Earthquake on 5 August 2018. *Remote Sens.* **2020**, *12*, 4050. [[CrossRef](#)]
3. Marchetti, D.; De Santis, A.; Shen, X.; Campuzano, S.A.; Perrone, L.; Piscini, A.; Di Giovambattista, R.; Jin, S.; Ippolito, A.; Cianchini, G.; et al. Possible Lithosphere-Atmosphere-Ionosphere Coupling effects prior to the 2018 Mw = 7.5 Indonesia earthquake from seismic, atmospheric and ionospheric data. *J. Southeast Asian Earth Sci.* **2019**, *188*, 104097. [[CrossRef](#)]
4. Liu, J.; Xu, T.; Ding, Z.; Zhang, X. The Comparison of Electron Density between CSES In Situ and Ground-Based Observations in China. *Remote Sens.* **2022**, *14*, 4498. [[CrossRef](#)]
5. Yang, M.; Zhang, X.; Ouyang, X.; Liu, J.; Qian, G.; Li, T.; Shen, X. Polarization Method-Based Research on Magnetic Field Data Associated with Earthquakes in Northeast Asia Recorded by the China Seismo-Electromagnetic Satellite. *Atmosphere* **2023**, *14*, 1555. [[CrossRef](#)]
6. Wang, L.; Zhang, Z.; Zhima, Z.; Shen, X.; Chu, W.; Yan, R.; Guo, F.; Zhou, N.; Chen, H.; Wei, D. Statistical Analysis of High-Energy Particle Perturbations in the Radiation Belts Related to Strong Earthquakes Based on the CSES Observations. *Remote Sens.* **2023**, *15*, 5030. [[CrossRef](#)]
7. Guo, Y.; Zhang, X.; Liu, J.; Yang, M.; Yang, X.; Du, X.; Lü, J.; Xiao, J. Seismo-Ionospheric Effects Prior to Two Earthquakes in Taiwan Detected by the China Seismo-Electromagnetic Satellite. *Atmosphere* **2022**, *13*, 1523. [[CrossRef](#)]
8. Kanarachos, S.; Christopoulos, S.; Chronos, A.; Fitzpatrick, M. Detecting anomalies in time series data via a deep learning algorithm combining wavelets, neural networks and Hilbert transform. *Expert Syst. Appl.* **2017**, *85*, 292–304. [[CrossRef](#)]
9. Zhima, Z.; Zhou, B.; Zhao, S.; Wang, Q.; Huang, J.; Li, Z.; Lei, J.; Chen, Y.; Yang, D.; Shen, X.; et al. Cross-calibration on the electromagnetic field detection payloads of the China Seismo-Electromagnetic Satellite. *Sci. China Technol. Sci.* **2022**, *65*, 1415–1426. [[CrossRef](#)]
10. Yan, R.; Guan, Y.; Miao, Y.; Zhima, Z.; Xiong, C.; Zhu, X.; Liu, C.; Shen, X.; Yuan, S.; Liu, D.; et al. The regular features recorded by the Langmuir probe onboard the low earth polar orbit satellite CSES. *J. Geophys. Res. Space Phys.* **2022**, *127*, e2021JA029289. [[CrossRef](#)]
11. Wang, X.; Yang, D.; Zhou, Z.; Cheng, W.; Xu, S.; Guo, F. Validation of CSES RO measurements using ionosonde and ISR observations. *Adv. Space Res.* **2020**, *66*, 2275–2288. [[CrossRef](#)]
12. Yuan, J.; Wang, Q.; Yang, D.; Liu, Q.; Zhima, Z.; Shen, X. Automatic recognition algorithm of lightning whistlers observed by the Search Coil Magnetometer onboard the Zhangheng-1 Satellite. *Chin. J. Geophys.* **2021**, *64*, 3905–3924. (In Chinese)
13. Yan, R.; Zhima, Z.; Xiong, C.; Shen, X.; Huang, J.; Guan, Y.; Zhu, X.; Liu, C. Comparison of electron density and temperature measurements of CSES satellite with other space-borne and ground-based observations. *J. Geophys. Res. Space Phys.* **2020**, *125*, e2019JA027747. [[CrossRef](#)]
14. Chen, W.; Marchetti, D.; Zhu, K.; Sabbagh, D.; Yan, R.; Zhima, Z.; Shen, X.; Cheng, Y.; Fan, M.; Wang, S.; et al. CSES-01 Electron Density Background Characterisation and Preliminary Investigation of Possible Ne Increase before Global Seismicity. *Atmosphere* **2023**, *14*, 1527. [[CrossRef](#)]
15. Liu, J.; Zhang, X.; Yang, X.; Yang, M.; Zhang, T.; Bao, Z.; Wu, W.; Qiu, G.; Yang, X.; Lu, Q. The Analysis of Lithosphere-Atmosphere-Ionosphere Coupling Associated with the 2022 Luding Ms6.8 Earthquake. *Remote Sens.* **2023**, *15*, 4042. [[CrossRef](#)]
16. Diego, P.; Huang, J.; Piersanti, M.; Badoni, D.; Zhima, Z.; Yan, R.; Rebutini, G.; Ammendola, R.; Candidi, M.; Guan, Y.; et al. The Electric Field Detector on Board the China Seismo Electromagnetic Satellite—In-Orbit Results and Validation. *Instruments* **2021**, *5*, 1. [[CrossRef](#)]

17. Zhao, S.; Shen, X.; Li, L.; Zhima, Z. A lithosphere-atmosphere-ionosphere coupling model for ELF electromagnetic waves radiated from seismic sources and its possibility observed by the CSES. *Sci. China Tech. Sci.* **2021**, *64*, 2551–2559. [[CrossRef](#)]
18. Yan, R.; Xiong, C.; Zhima, Z.; Shen, X.; Liu, D.; Liu, C.; Guan, Y.; Zhu, K.; Zheng, L.; Lv, F. Correlation between Ne and Te around 14:00 LT in the topside ionosphere observed by CSES, Swarm and CHAMP satellites. *Front. Earth Sci.* **2022**, *10*, 860234. [[CrossRef](#)]
19. Yang, Y.; Hulot, G.; Vigneron, P.; Shen, X.; Zhima, Z.; Zhou, B.; Magnes, W.; Olsen, N.; Tøffner-Clausen, L.; Huang, J.; et al. The CSES Global Geomagnetic Field Model (CGGM): An IGRF-Type Global Geomagnetic Field Model Based on Data from the China Seismo-Electromagnetic Satellite. *Earth Planets Space* **2021**, *73*, 45. [[CrossRef](#)]
20. Ghamry, E.; Marchetti, D.; Yoshikawa, A.; Uozumi, T.; De Santis, A.; Perrone, L.; Shen, X.; Fathy, A. The First Pi2 Pulsation Observed by China Seismo-Electromagnetic Satellite. *Remote Sens.* **2020**, *12*, 2300. [[CrossRef](#)]
21. Gou, X.; Li, L.; Zhou, B.; Zhang, Y.; Xie, L.; Cheng, B.; Feng, Y.; Wang, J.; Miao, Y.; Zhima, Z.; et al. Electrostatic Ion Cyclotron Waves Observed by CSES in the Equatorial Plasma Bubble. *Geophys. Res. Lett.* **2023**, *50*, e2022GL101791. [[CrossRef](#)]
22. Marchetti, D.; Zhu, K.; Zhang, H.; Zhima, Z.; Yan, R.; Shen, X.; Chen, W.; Cheng, Y.; He, X.; Wang, T.; et al. Clues of Lithosphere, Atmosphere and Ionosphere Variations Possibly Related to the Preparation of La Palma 19 September 2021 Volcano Eruption. *Remote Sens.* **2022**, *14*, 5001. [[CrossRef](#)]
23. Spogli, L.; Sabbagh, D.; Regi, M.; Cesaroni, C.; Perrone, L.; Alfonsi, L.; Di Mauro, D.; Lepidi, S.; Campuzano, S.A.; Marchetti, D.; et al. Ionospheric Response Over Brazil to the August 2018 Geomagnetic Storm as Probed by CSES-01 and Swarm Satellites and by Local Ground-Based Observations. *JGR Space Phys.* **2021**, *126*, e2020JA028368. [[CrossRef](#)]
24. Li, Q.; Shi, Y.; Lin, R.; Qiao, W.; Ba, W. A novel oil pipeline leakage detection method based on the sparrow search algorithm and CNN. *Measurement* **2022**, *204*, 112122. [[CrossRef](#)]
25. Shen, X.; Zhang, X.; Yuan, S.; Wang, L.; Cao, J.; Huang, J.; Zhu, X.; Piergiorgio, P.; Dai, J. The state-of-the-art of the China Seismo-Electromagnetic Satellite mission. *Sci. China Tech. Sci.* **2018**, *61*, 634–642. [[CrossRef](#)]
26. Huang, J.; Shen, X.; Zhang, X.; Lu, H.; Tan, Q.; Wang, Q.; Yan, R.; Chu, W.; Yang, Y.; Liu, D. Application system and data description of the China Seismo-Electromagnetic Satellite. *Earth Planet. Phys.* **2018**, *2*, 444–454. [[CrossRef](#)]
27. Huang, J.; Lei, J.; Li, S.; Zhima, Z.; Li, C.; Zhu, X.; Yu, W. The Electric Field Detector (EFD) onboard the ZH-1 satellite and first observational results. *Earth Planet. Phys.* **2018**, *2*, 469–478. [[CrossRef](#)]
28. Bai, Y.; Liu, J.; Huang, L.; Bai, J.; Che, X. Analysis of two convolutional neural networks for map image recognition. *Sci. Surv. Mapp.* **2021**, *46*, 126–134.
29. Kavitha, M.; Gayathri, R.; Alhudhaif, A.; Alenezi, F. Performance evaluation of deep e-CNN with integrated spatial-spectral features in hyperspectral image classification. *Measurement* **2022**, *191*, 110760. [[CrossRef](#)]
30. Krizhevsky, A.; Sutskever, I.; Hinton, G. Imagenet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [[CrossRef](#)]
31. Zeiler, M.; Fergus, R. Visualizing and Understanding Convolutional Networks. In Proceedings of the European Conference on Computer Vision 2014, Zurich, Switzerland, 6–12 September 2014; pp. 818–833.
32. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:091556.
33. He, K.; Zhang, X.; Ren, S. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2016, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
34. Shan, S.; Liu, J.; Wu, S.; Shao, Y.; Li, H. A motor bearing fault voiceprint recognition method based on Mel-CNN model. *Measurement* **2023**, *207*, 112408. [[CrossRef](#)]
35. Zhang, X.; Shi, B.; Feng, B.; Liu, L.; Gao, Z. A hybrid method for cutting tool RUL prediction based on CNN and multistage Wiener process using small sample data. *Measurement* **2023**, *213*, 112739. [[CrossRef](#)]
36. Bi, J.; Li, L. Simulations of the spacecraft charging and wake effects in the solar wind. *Chin. J. Space Sci.* **2018**, *38*, 909–914. (In Chinese) [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.