Machine Learning Detection of Dust Impact Signals Observed by the Solar Orbiter

Andreas Kvammen[1], Kristoffer Wickstrøm[1], Samuel Kociscak[1], Jakub Vaverka[2], Libor Nouzak[2], Arnaud Zaslavsky[3], Kristina Rackovic Babic[3,4], Amalie Gjelsvik[1], David Pisa[5], Jan Soucek[5], and Ingrid Mann[1]

[1]Department of Physics and Technology, UiT The Arctic University of Norway, 9037, Tromsø, Norway
[2]Department of Surface and Plasma Science, Charles University Prague, 18000, Prague, Czech Republic
[3]LESIA – Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 place Jules Janssen, 92195, Meudon, France
[4]Department of Astronomy, Faculty of Mathematics, University of Belgrade, Studentski trg 16, 11000, Belgrade, Serbia
[5]Institute of Atmospheric Physics, Czech Academy of Sciences, Bocni II/1401, 141 00 Prague, Czech Republic

Correspondence: Andreas Kvammen (andreas.kvammen@uit.no)

Abstract. This article presents the results of automatic detection of dust impact signals observed by the Solar Orbiter – Radio and Plasma Waves instrument.

A sharp and characteristic electric field signal is observed by the Radio and Plasma Waves instrument when a dust particle impacts the spacecraft at high velocity. In this way, ∼5–20 dust impacts are daily detected as the Solar Orbiter travels through
5 the interplanetary medium. The dust distribution in the inner solar system is largely uncharted and statistical studies of the detected dust impacts will enhance our understanding of the role of dust in the solar system.

It is however challenging to automatically detect and separate dust signals from the plural of other signal shapes for two main reasons. Firstly, since the spacecraft charging causes variable shapes of the impact signals, and secondly because electromagnetic waves (such as solitary waves) may induce resembling electric field signals.

10 In this article, we propose a novel machine learning-based framework for detection of dust impacts. We consider two different supervised machine learning approaches: the support vector machine classifier and the convolutional neural network classifier. Furthermore, we compare the performance of the machine learning classifiers to the currently used on-board classification algorithm and analyze two years of Radio and Plasma Waves instrument data.

Overall, we conclude that detection of dust impact signals is a suitable task for supervised machine learning techniques. The
15 convolutional neural network achieves the highest performance with 96% ± 1% overall classification accuracy and 94% ± 2% dust detection precision, a significant improvement to the currently used on-board classifier with 85% overall classification accuracy and 75% dust detection precision. In addition, both the support vector machine and the convolutional neural network classifiers detect more dust particles (on average) than the on-board classification algorithm, with 16% ± 1% and 18% ± 8% detection enhancement respectively.

20 The proposed convolutional neural network classifier (or similar tools) should therefore be considered for post-processing of the electric field signals observed by the Solar Orbiter.

1 Introduction

1.1 The Dust Population in the Inner Solar System

The interplanetary dust population in the inner solar system ($\leq$ 1 AU) is formed by collisional fragmentation of asteroids, comets and meteoroids. The meteoroids and the larger dust particles are in bound orbits around the Sun and their lifetime is limited by collisions, while the smaller particles that form through collisional fragmentation are repelled from the Sun by the radiation pressure force (Mann et al., 2004). The sources and sinks of the interplanetary dust particles are well-studied at the orbit of Earth (Grün et al., 1985), while there have been few observations inside 1 AU until recent years.

Model calculations show that the number density of dust within 1 AU is diminished by collisional destruction (Ishimoto, 2000). However, there are a number of uncertainties that enter the model calculations since the dust collision rates depend both on the dust number density distribution and on the relative velocities between the dust particles. These parameters are generally unknown inside the orbit of the Earth and the estimated sizes of the fragmented dust particles are currently based on empirical relations, inferred from laboratory measurements of accelerated dust particles (Mann and Czechowski, 2005). Furthermore, there is an additional dust population with an interstellar origin that streams through the solar system. The interstellar dust distribution is largely unknown and thus complicates the analysis of the interplanetary dust population. Remote observations of the zodiacal light and the Fraunhofer corona (F-corona) provide some information of the dust population within 1 AU, but mainly of the larger ($> \mu$m) dust particles (Mann et al., 2004). For all these reasons, in-situ measurements are needed in order to better understand the role of dust in the inner solar system.

1.2 Exploration of the Inner Solar System

At present, the inner solar system is explored by the Parker Solar Probe (Szalay et al., 2020), launched August 12, 2018, and the Solar Orbiter (Müller et al., 2020), launched February 10, 2020. While systematic studies of the dust flux near 1 AU are conducted with the Solar Terrestrial Relations Observatory (STEREO) (Zaslavsky et al., 2012) and Wind (Malaspina et al., 2014). The first analyses show that a large fraction of the observed dust particles are repelled from the Sun, i.e. the dust particles are in unbound orbits (Zaslavsky et al., 2021; Szalay et al., 2020; Malaspina et al., 2020). Mann and Czechowski (2021) used model calculations to explain the impact rates observed by the Parker Solar Probe. The dust production was modeled by collisional fragmentation near the Sun and the dust trajectories were calculated with included radiation pressure and Lorentz force terms. Mann and Czechowski (2021) showed that the observed impact rates largely agree with the model calculations for dust $>$ 100 nm and proposed that the differences may be explained by the influence of smaller particles and of other dust components, such as dust in bound orbits and interstellar dust.

In this work, we analyze data acquired by the Solar Orbiter. The spacecraft orbits the Sun in an elliptic orbit with a period of approximately 6 months. At perihelion, the Solar Orbiter reaches a minimum solar distance of 0.28 AU, just within the perihelion of the Mercury orbit. The expected mission duration is 7 years, with a possible 3-year extension. The Solar Orbiter

will thus provide long-term, in-situ observations of the environment in the inner solar system with multiple instruments. One of these instruments is the Radio and Plasma Waves instrument, allowing observations of the cosmic dust flux with typical diameters ranging from ~100 nm to ~500 nm (Zaslavsky et al., 2021).

1.3 Radio and Plasma Waves Instruments for Dust Detection

Radio and plasma waves instruments (i.e. antennas) have been used for studying dust in the solar system since the Voyager mission (Gurnett et al., 1983; Aubier et al., 1983). A dust impact is observed by the spacecraft antennas as a sharp and characteristic electric field signal, produced by the impact ionization process.

The impact ionization process occurs when dust particles hit a target in space with impact speeds on the order of ~km/s or larger, impact speeds which are typical for space missions in the interplanetary medium. The kinetic energy of the impact is transferred into deformation, shattering, melting and vaporization of the dust projectile– and target material, producing a cloud of free electrons and ions on the surface of the spacecraft. Laboratory measurements (Collette et al., 2014) and model calculations (Hornung et al., 2000) indicate that the free-charge yield depends on multiple parameters, where the most important are the dust impact velocity, the dust mass, and the material of both the dust projectile and the target (the spacecraft surface) (Mann et al., 2019). The forming cloud of charged particles is partly expanding into the ambient solar wind and is partly recollected by the spacecraft. This induces the characteristic electric field signal, hereafter referred to as the dust impact signal/waveform.

Radio and plasma waves instruments allow for the entire spacecraft body to serve as a dust detector, providing a large collection area in comparison to dedicated dust detection instruments. Thus, radio and plasma waves instruments can provide dust distribution estimates based on thousands of dust impacts each year, statistical products that are difficult to acquire by dedicated dust instruments. Still, radio and plasma waves instruments have lower sensitivities than dedicated dust detectors (Zaslavsky, 2015) and the shape of the dust impact waveform is highly dependent on the potential difference between the spacecraft and the ambient plasma (Vaverka et al., 2017). This complicates the analysis of the dust distribution in the solar system since statistical studies rely on automatic dust detection with high accuracy, which is difficult to attain with the software currently in use.

1.4 Machine Learning Classification of Time Series Data

In this article, we present a machine learning-based framework as a novel method for detecting dust impact signals in radio and plasma waves instrument data. Machine learning methods, in particular neural networks in the recent decade, have been extensively used for challenging time series classification problems, such as: speech recognition (Trosten et al., 2019), heart rate monitoring (Wickstrøm et al., 2022) and human activity classification (Villar et al., 2016).

A neural network has previously been used for selecting the signals of interest observed by the WAVES instrument on board the Wind spacecraft (Bougeret et al., 1995). While an unsupervised method (self-organizing maps) was used for identifying and

categorizing plasma waves in the magnetic field data observed by the MMS_1 spacecraft (Vech and Malaspina, 2021). Still, no machine learning tools have been developed for classifying dust impacts in radio and plasma waves instrument data, although the characteristic signal produced by the impact ionization process is distinctive and could therefore be suitable for machine learning detection.

1.5 Motivation and Article Structure

The main motivation for this work was to develop a dedicated dust detection tool that can be used to automatically process the large amount of data acquired by the Radio and Plasma Waves instrument on board the Solar Orbiter. The aim was to develop a classifier with a high overall classification accuracy on a balanced data set that can make statistical studies more reliable and easier to conduct. For this project, we defined high accuracy to be ($\gtrsim 95\%$) after some initial testing. We considered ($\gtrsim 95\%$) accuracy to be satisfactory for meaningful statistical studies and a significant improvement to the currently used classification system. In order to achieve this objective, we used supervised machine learning techniques to develop the dust classifiers, trained and tested on a set of 3000 manually labeled observations.

The remaining of this article is structured as follows. Section 2 explains the Solar Orbiter – Radio and Plasma Waves observations and the on-board algorithm that is currently used for dust impact detection. Section 3 describes the procedure that was used for developing the machine learning classifiers; from the downloaded data to the training– and testing of the classifiers. Section 4 investigate the performance of the classifiers and includes the resulting dust impact rates, calculated by analyzing two years of automatically classified Solar Orbiter data. Finally, Section 5 presents the overall conclusions of this project.

2 Observations and Data Acquisition

2.1 The Radio and Plasma Waves (RPW) Instrument and the Time Domain Sampler (TDS) Receiver

This work focuses on electric field signals (i.e. waveforms) observed by the Radio and Plasma Waves (RPW) instrument on-board the Solar Orbiter (Maksimovic et al., 2020). The RPW instrument consists of 3 antennas operating synchronously and the measured electric potential is recorded by the Time Domain Sampler (TDS) receiver unit (Soucek et al., 2021).

The TDS receiver is designed to capture plasma waves (such as ion-acoustic and Langmuir waves) in the frequency range 200 Hz – 100 kHz, in addition to the dust impact signals (Soucek et al., 2021). The antenna voltages are converted to electric field values using the antenna effective lengths but are otherwise uncalibrated. We consider only signals sampled with a sampling rate of 262.1 kHz in snapshots of 16384 time-steps, acquired when the TDS receiver was operating in the XLD1 mode.

The XLD1 mode is the most commonly used observational mode of the RPW-TDS system (Soucek et al., 2021). XLD1 is a hybrid mode, where channel 3 (CH_3) is operating in monopole mode while channel 1 (CH_1) and channel 2 (CH_2) are operating in dipole mode:

$$CH_1 = \left(\frac{V_1 - V_{SC}}{L_1}\right)\hat{L}_1 - \left(\frac{V_3 - V_{SC}}{L_3}\right)\hat{L}_3 \tag{1}$$

$$CH_2 = \left(\frac{V_2 - V_{SC}}{L_2}\right)\hat{L}_2 - \left(\frac{V_1 - V_{SC}}{L_1}\right)\hat{L}_1 \tag{2}$$

$$CH_3 = \left(\frac{V_2 - V_{SC}}{L_2}\right)\hat{L}_2 \tag{3}$$

where $V_i - V_{SC}$ denotes the potential difference between antenna i and the spacecraft body along the antenna boom with unit vector $\hat{L}_i$ and effective length L_i. For this work however, the 3 RPW antenna signals are all converted to monopole electric field signals ($\bar{E}_1, \bar{E}_2, \bar{E}_3$) by the following conversion:

$$\bar{E}_1 = CH_3 - CH_2 = \left(\frac{V_2 - V_{SC}}{L_2}\right)\hat{L}_2 - \left(\left(\frac{V_2 - V_{SC}}{L_2}\right)\hat{L}_2 - \left(\frac{V_1 - V_{SC}}{L_1}\right)\hat{L}_1\right) = \left(\frac{V_1 - V_{SC}}{L_1}\right)\hat{L}_1 \tag{4}$$

$$\bar{E}_2 = CH_3 = \left(\frac{V_2 - V_{SC}}{L_2}\right)\hat{L}_2 \tag{5}$$

$$\bar{E}_3 = CH_3 - CH_2 - CH_1 = \bar{E}_1 - CH_1$$

$$= \left(\frac{V_1 - V_{SC}}{L_1}\right)\hat{L}_1 - \left(\left(\frac{V_1 - V_{SC}}{L_1}\right)\hat{L}_1 - \left(\frac{V_3 - V_{SC}}{L_3}\right)\hat{L}_3\right) = \left(\frac{V_3 - V_{SC}}{L_3}\right)\hat{L}_3 \tag{6}$$

The Solar Orbiter RPW-TDS detection threshold is $\sim$5 mV, allowing dust impact identification of the cosmic dust flux with typical diameters ranging from $\sim$100 nm to $\sim$500 nm (Zaslavsky et al., 2021).

2.2 The Triggered Snapshot WaveForms (TSWF) data product and the TDS Classifier

For this project, we use the Triggered Snapshot WaveForms (TSWF) data product, processed with software version 2.1.1 and acquired over a 25 month-period, spanning between June 15, 2020, to July 14, 2022. The TSWF data product consists of signal packets (63 ms snapshots) that are down-linked only if the classification algorithm on-board the Solar Orbiter is triggered. The accuracy of the on-board classification algorithm is therefore important in order to optimize the data transfer and provide reliable data products for statistical analysis.

The input to the on-board classification algorithm, hereafter named the TDS classifier or the TDS classification algorithm, is the 63 ms signal packet, while the output is categorized into one out of three labels: *dust*, *wave* or *other*. The TDS classifier assigns the label based on 3 extracted features.

1. The snapshot peak amplitude (V_{max})

2. The ratio of the peak amplitude to the median absolute value of the signal (V_{max}/V_{med})

3. The full width half maximum (BW) of the main spectral peak, identified by analyzing the discrete Fourier transform of the signal.

The signal label is then determined by comparing the extracted feature values against configurable thresholds. The threshold criterion reflects that observations of waves are typically narrow-band (low BW) and the peak of the signal is not much larger than the median value (low V_{max}/V_{med}). In contrast, dust observations are sharp non-periodic signals (high BW) that generally have a high maximum to median amplitude ratio (high V_{max}/V_{med}). For more detailed descriptions of the TDS classifier, see Maksimovic et al. (2020) and Soucek et al. (2021). Figure 1 presents a few examples of recorded snapshots with included labels, as classified by the TDS classification algorithm.

Figure 1 illustrates that it is challenging to detect and separate dust signals from the plural of other signal shapes. In particular, the dust waveform in Sub-figure c) is classified as *other*, while the Langmuir wave and solitary wave snapshots in Sub-figures g) and h) are erroneously classified as *dust* by the TDS classification algorithm. For more information on observations of Langmuir and ion-acoustic waves in the Solar Orbiter data, see e.g. Soucek et al. (2021), and for an analysis of Wind observations of electrostatic solitary waves, see Malaspina et al. (2013).

3   Machine Learning-Based Framework for Automatic Dust Impact Detection

The goal of the machine learning classifier is to take a monopole RPW snapshot as an input and automatically output if the signal contains a dust impact or not. For this purpose, we use a supervised classifier. A supervised classifier relies on manually labeled data to learn (i.e. train) the function that maps the input observation (the electric field signal) to the output label. For this work, we focus exclusively on detecting dust impact signals, we therefore use the binary labels: *dust* or *no dust*. Additional labels, such as: *ion-acoustic waves*, *Langmuir waves* and *solitary waves*, could however be implemented in a similar machine learning-based framework.

3.1   Data Pre-Processing for Machine Learning Classification

In order to construct a balanced data set, we selected ~ 1500 waveforms classified as *dust* and ~ 1500 waveforms classified as *wave/other* by the TDS classification algorithm. The signals were randomly drawn from the TDS data archive and acquired between 15 June 2020 to 16 December 2021. The TDS signals were then pre-processed to standardize the input to the classifier and speed up the training. Standardized data further reduces bias effects and makes the manual labeling of the signals easier to conduct. For this work, a 4-step pre-processing procedure was used independently on each antenna signal, the pre-processing procedure applied on a sample signal is illustrated in Figure 2.

1. **Remove the signal offset** The electric field offset is removed by subtracting the raw signal with the median of a heavily filtered version of the raw data. A sliding median filter over 21 time-steps was selected by visual inspection of the noise characteristics. The removal of the electric field offset centers the signal around zero and reduces bias effects from offset waveforms.
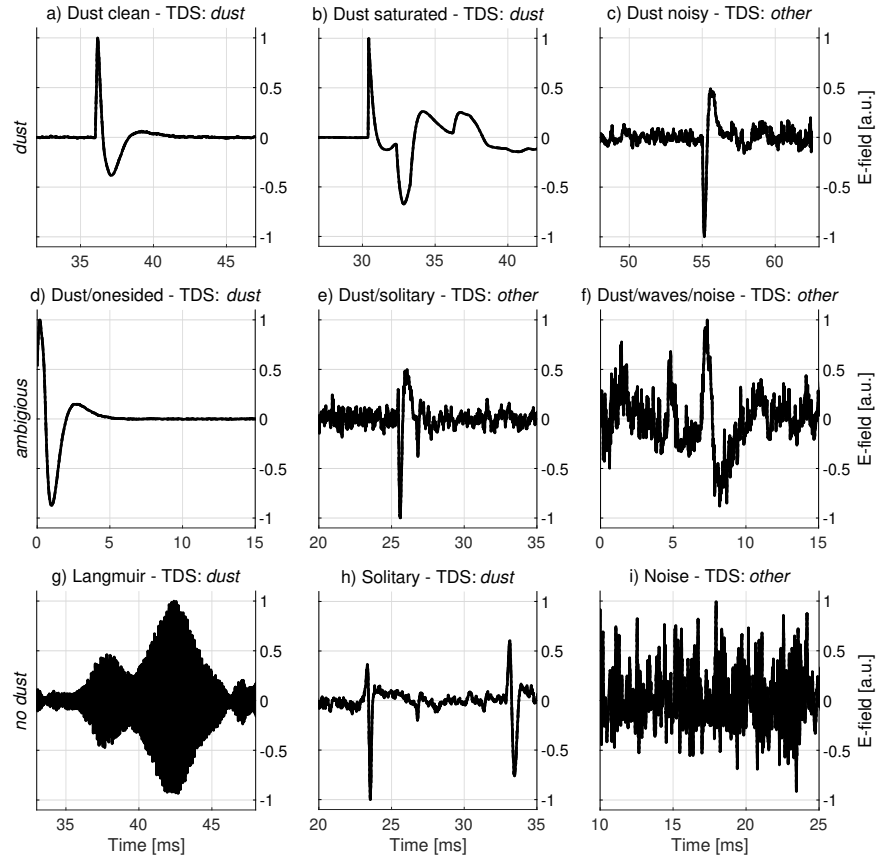
Figure 1. Waveforms recorded by the TDS receiver and measured by one of the RPW antennas. The signal label, classified by the TDS classification algorithm, is included for each snapshot in the subplot titles. The top row presents dust waveforms: a) is a clean dust impact waveform, b) shows a dust impact that saturates the receiver unit (or reaches the non-linearity limit), c) presents a weak dust impact signal that is strongly affected by noise. The middle row presents ambiguous waveforms: d) might be a dust impact, but information is limited by the signal framing, e) is likely a dust impact, but the signal shape resembles solitary waves and is strongly affected by noise, f) might be a dust impact, but noise and possible electromagnetic waves make the signal difficult to interpret. The bottom row presents waveforms without dust: g) shows Langmuir waves, characterized by the high-frequency E-field oscillations with a lower-frequency amplitude modulation, h) presents solitary waves, which sometimes resemble dust impact waveforms, i) shows a signal dominated by noise, without any clear features. Note that the full (63 ms) snapshots are zoomed to 15 ms intervals around the interesting features and that the signal amplitudes are normalized to ± 1 and centered around zero for illustrative purposes.

2. **Filter the data** The signal is filtered using a sliding median filter over 7 time-steps in order to reduce the high-frequency noise. The 7 time-step filter was selected by inspecting the power spectrum of impact signals and by noticing that most information above (f_N = 35 kHz) is buried in noise, although the TDS sampling frequency is higher (f_s = 262.1 kHz), thus making a filter length ($< f_s/f_N \approx 7.5$) appropriate without significant loss of information.

3. **Compress the data** The signal is re-sampled with a compression factor of 4 using linear 1-dimensional interpolation. The compression is done to speed up the training of the classifier, resulting in a re-sampling from 16384 to 4096 time-steps.

4. **Normalize the signal** The data is normalized to be between -1 and 1 by dividing all data samples with the maximum absolute value of the signal. The normalization makes the machine learning classifier more robust to variations in the signal strength and eases the parameter optimization during training.
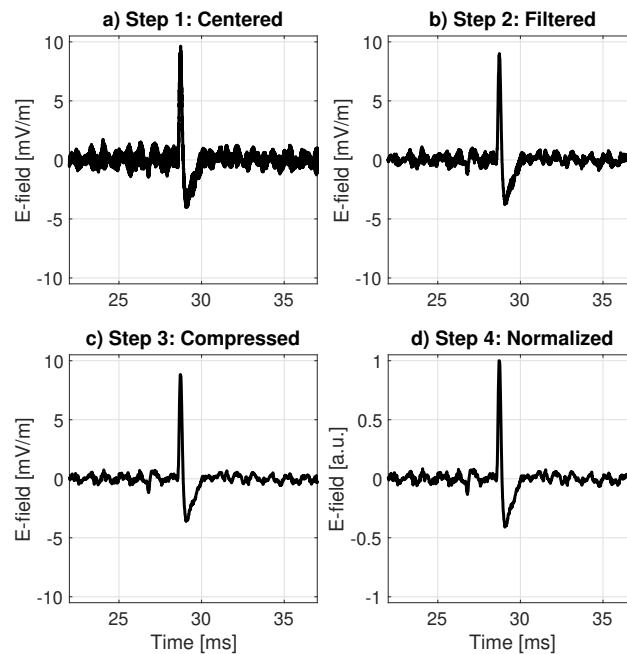
Figure 2. A dust waveform observed by antenna 2 on September 8, 2021. The sub-figures illustrate the different stages of the pre-processing procedure. a) The electric field offset is removed and the signal is centered around 0 mV/m. b) The signal is filtered by a median filter over 7 time-steps to reduce the high-frequency noise. c) The signal is compressed by a factor of 4 to reduce the data size. d) The waveform is normalized by the maximum absolute value of the signal in order to ease the parameter optimization of the machine learning classifier. Note the waveform is zoomed to a 15 ms time period around the dust impact in order to better visualize the impact shape modification by the pre-processing procedure.

3.2 Manual Waveform Labeling

Manually labeled data is used both to train the machine learning classifiers and to test the performance of the trained models. Thus, great care is needed in order to construct a high-quality labeled data set, without significant contamination of corrupted data files, biases and mislabeled signals.

We manually labeled the data into either *dust* or *no dust*. Each signal was displayed without indications of the previously assigned label by the TDS classifier in order to reduce bias effects. Furthermore, a zoom function was used to investigate the areas of interest and options were included both to correct labeling mistakes by the user and to indicate ambiguous signals that do not clearly fit into any label (*dust* or *no dust*). Appendix A presents the Graphical User Interface (GUI) that was used to label the 3000 observations.

It should be noted that 134 signals (i.e. 4.5%), out of 3000 manually labeled waveforms, were marked as ambiguous and did not clearly fit into either the *dust* or *no dust* label, see the middle row of Figure 1 for ambiguous examples. Furthermore, the manual waveform labeling was done by one scientist, although with consultations with other experts. Thus, it is to be expected that different scientists will disagree on a proportion (up to 5%) of the manual labels. The disagreement level could possibly be reduced if several experts labeled the same data set, and the labeling consensus was used as the effective waveform label.

3.3 Developing the Machine Learning Classifiers

The manually labeled data was split into a training set (containing 80% of the data) and a testing set (with the remaining 20%). The training data is used to optimize the free parameters of the machine learning classifiers with respect to the assigned labels, while the testing data is used as an independent set to evaluate the performance of the trained classifiers. The performance of a machine learning classifier is quantified by comparing the outputs of the trained model to the labels of the testing data. Figure 3 illustrates the data flow; from the TDS data sets to the machine learning performance metrics.

There are numerous machine learning techniques that are suitable for time series classification. In this work, we focus on two well-known techniques: the Support Vector Machine (SVM), described in Sub-section 3.4, and the Convolutional Neural Network (CNN), discussed in Sub-section 3.5.

3.4 The Support Vector Machine (SVM)

The support vector machine (Boser et al., 1992; Cortes and Vapnik, 1995) is a robust and versatile classification algorithm, considered to be one of the most influential approaches in supervised learning (Goodfellow et al., 2016). SVMs learn the decision hyperplane that maximizes the discriminative power between the observations categorized into two classes (in this case: *dust* or *no dust*). However, SVMs are highly dependent on the representation of the data and often achieve sub-optimal performance on high-dimensional data (when used directly). In this case, the observations from 3 antenna measurements, each
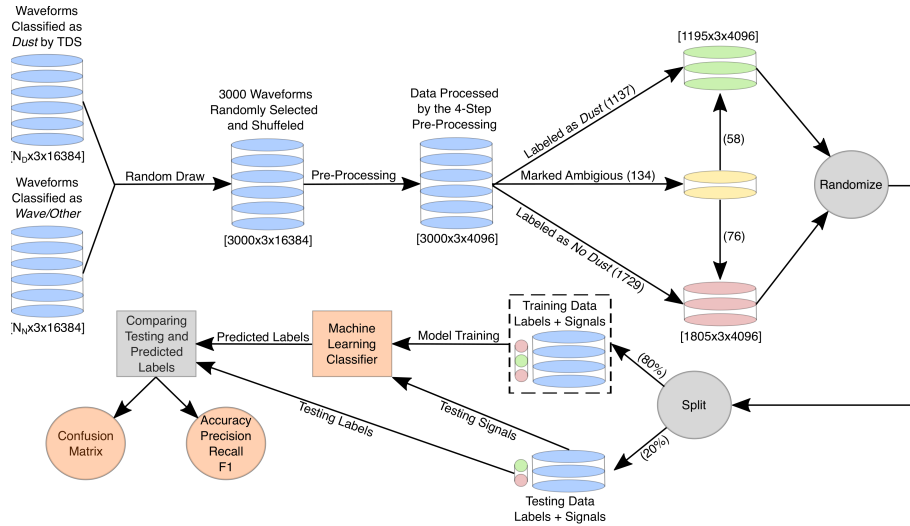
Figure 3. Data flow; from the TDS data sets to the machine learning performance metrics. The diagram illustrates the data flow by the black arrows and the applied process by the arrow label. The cylinders indicate the signal waveforms and the cylinder color indicates the associated label. The gray circles mark data transformation processes. The random draw of the TDS data and the pre-processing is explained in Sub-section 3.1, while the manual labeling is described in Sub-section 3.2. A description of the randomization and splitting of the manually labeled data into a training and a testing set is included in Sub-section 3.3. Sub-sections 3.4 and 3.5 explain the training and testing of the machine learning classifiers. Finally, the performances of the machine learning classifiers are compared and evaluated in Sub-section 4.1.

with 4096 time-steps, are both high dimensional and noisy (each time-step contain little information). It is therefore common to extract important characteristics (i.e. features) from the data to provide the SVM with compactly represented information with less noise and redundancies.

3.4.1 Feature Extraction

In order to develop a baseline machine learning classifier, comparable to the on-board TDS classification algorithm, a simple 2-dimensional SVM classifier was considered. Thus, every observation with dimension (3x4096) is represented by a 2-dimensional feature vector (1x2). After some initial testing, we selected two features that had a high discriminative power between the *dust* and *no dust* observations.

1. **The standard deviation** The mean standard deviation is calculated over the 3 antenna channels, each with 4096 time-steps. The standard deviation is an appropriate feature since normalized *dust* signals typically have a lower mean standard deviation than normalized *no dust* signals.

2. **The convolution ratio** The $\log_{10}$ value of the convolution ratio ($|\text{conv}|_{\max}/|\text{conv}|_{\text{median}}$) is calculated, where |conv| is the absolute values of the convolution of the antenna signals with a normalized Gaussian of width 0.5 ms. $|\text{conv}|_{\max}$ is the maximum value of |conv|, while $|\text{conv}|_{\text{median}}$ is the median. The convolution ratio was selected as a feature since the *dust* signals typically have a larger convolution ratio than the *no dust* signals. The Gaussian width of 0.5 ms was experimentally found to give high correlations with dust impact signals.

3.4.2 Training the Support Vector Machine

The 2 features (standard deviation and convolution ratio) were extracted from all observations in the training data. The decision hyperplane, in this 2-dimensional case a decision line, is defined by a polynomial of degree 2 that is optimized by minimizing the non-separable SVM cost function, see e.g. Theodoridis and Koutroumbas (2009) for details. The SVM classifier was trained with a slack variable factor of 1 and equal weighting between the *dust* and *no dust* observations. The 2-dimensional SVM is computationally inexpensive to optimize with a training time of ~1 second on a modern laptop. Figure 4 illustrates the training of the SVM classifier.
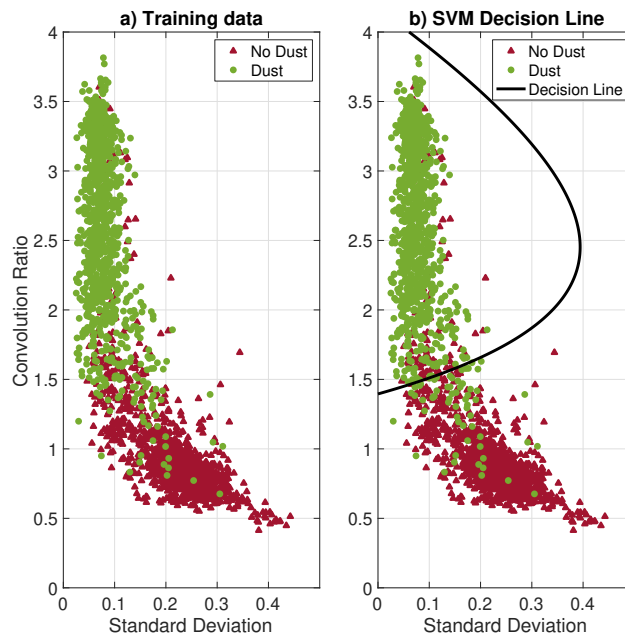
Figure 4. a) The (1x2) feature vectors extracted from all (2400) observations in the training data, the associated labels are indicated in green (*dust*) and red (*no dust*). b) The SVM decision line is defined as a second-order polynomial, obtained by minimizing the non-separable SVM cost function. The optimized SVM decision line appears to be reasonable, and most observations are separable in the training data.

 ### 3.4.3 Testing the Support Vector Machine

The performance of the trained SVM classifier is evaluated using the independent testing data, i.e. the remaining manually labeled data (20 %) that was not used for training the classifier. Figure 5 presents the SVM classification performance on the testing data.
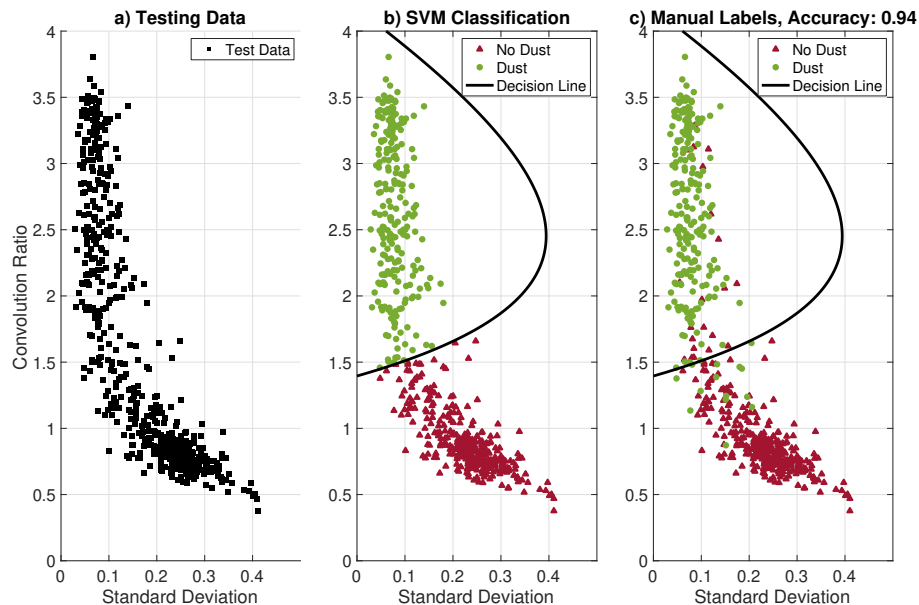
Figure 5. a) The (1x2) feature vectors extracted from the testing data (600 observations with hidden labels). b) The testing data is classified using the trained SVM decision line, where all observation within the polynomial line is classified as *dust* while all observations outside are classified as *no dust*. c) The "true" labels (from the manual labeling) are revealed. It is clear that some observations are confused, predominantly near the decision line. Still, the SVM classifier achieves an overall classification accuracy of 94%, calculated by comparing the outputs from the SVM classification (Sub-figure b) to the "true" labels (Sub-figure c).

250   Overall, the SVM classifier achieves a classification accuracy of 94% on the testing data using the 2-dimensional feature vectors. Note that the inclusion of additional extracted features can possibly enhance the SVM performance. Several additional features can be considered, such as; the mean amplitude of the signal, the range between the signal maximum and minimum values and the cross-correlation length (the time lag to the first zero crossing).

3.4.4 Explainability of the Support Vector Machine

255 Ideally, we want to develop a machine learning classifier that not only has a high accuracy, but also makes decisions that are understandable for human experts (Holzinger et al., 2019). In other words, we want to be able to explain why the machine learning classifier selected the predicted class for a given observation. In machine learning, this is often referred to as the

explainability of the trained classifier. Figure 5 presents the testing data in the 2–D feature vector space, but this plot gives no clear indications of how different signal shapes are distributed and which signatures are confused by the SVM classifier. In order to better understand the decisions made by the SVM classifier, the signal examples in Figure 1 are studied in detail. The analysis is presented in Figure 6.

It should be noted that the signal examples in Figure 6 are not representative for the general distribution of observations in the 2–D feature vector space, since most observations are clustered in distinct *dust* and *no dust* regions, as can be seen in Figure 5. Figure 6 focuses mostly on signal examples that are challenging to classify. Still, Figure 6 indicates that the SVM classifier provides mostly comprehensible outputs, but might have difficulties classifying weak dust impact signals and signals with important signatures located at the edge of the snapshot frame.

3.5 The Convolutional Neural Network (CNN)

Convolutional Neural Networks are algorithms designed for processing grid-like data and have achieved premium performance on a number of different tasks in the recent decade, such as image (He et al., 2016; Kvammen et al., 2020), video (Karpathy et al., 2014), and time series (Wang et al., 2017; Wickstrøm et al., 2021) classification.

3.5.1 Feature Extraction

Unlike the SVM, the CNN does not require pre-defined feature extraction routines. Instead, the CNN extracts the features based on a chain of convolution operations and automatically optimizes the convolution filters based on the training data and the associated labels.

For this work, we employed the 3-layer fully convolutional network architecture presented in Wang et al. (2017) and suggested for time series classification after extensive testing (Wickstrøm et al., 2022; Fawaz et al., 2020; Karim et al., 2019). The Rectified Linear Unit (ReLU) function (Glorot et al., 2011) was used as the activation function and Batch Normalization (BN) (Ioffe and Szegedy, 2015) was used at each convolutional layer in order to regularize the network and accelerate the training process. Figure 7 presents the employed CNN architecture.

3.5.2 Training the Convolutional Neural Network

The 3-layer fully convolutional network consists of 267010 free parameters (weights and biases) that need to be optimized to solve the dust impact classification task. The free parameters are randomly initialized and thereafter optimized using the ADAM gradient descent optimizer (Kingma and Ba, 2014). The CNN was trained for 225 epochs with a cross-entropy loss function using the 2400 labeled observations in the training data. CNNs are computationally expensive to optimize, as compared to the SVM classifier, and a training time of ∼20 minutes was required using TensorFlow on a MacBook Pro with a 32-core M1 Max GPU chip. For more details on neural network training and optimization, see for example (Montavon et al., 2012).
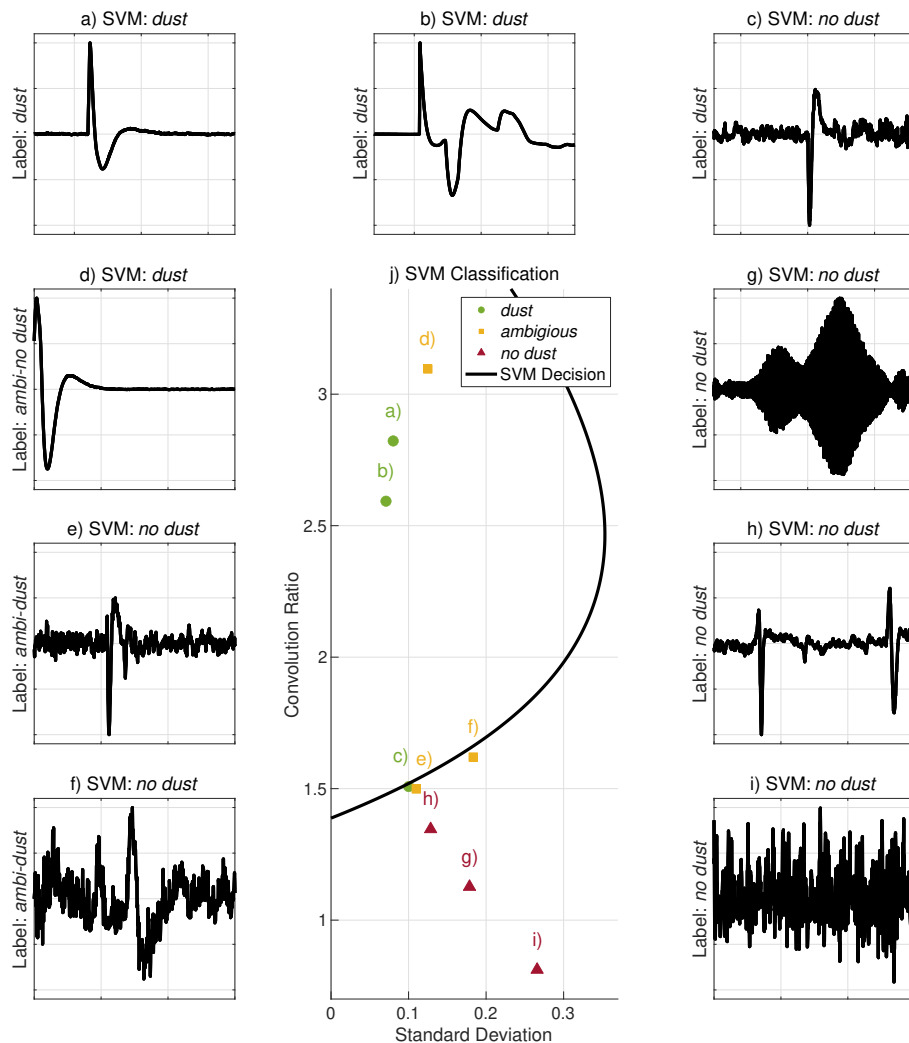
Figure 6. The signal examples are presented in Sub-figures (a-i), the manual labels are indicated along the y-axis and the predicted labels, classified by the SVM decision line, are presented in the sub-plot titles. Sub-figure j) presents the associated signal examples in the 2–D feature vector space along with the SVM decision line. The *dust* signals are illustrated in green, the *ambiguous* signals are marked in yellow and the *no dust* signals are indicated in red. The SVM classifier provides mostly explainable outputs. The clear dust signals (a-b) are located well within the SVM decision line, the *ambiguous* signals (e-f) are located near the decision line while the no dust signals (g-i) are clearly located outside. However, *dust* signal c) is erroneously located just outside the decision line, this can possibly be explained by the weak signal-to-noise ratio. In addition, signal d) is located well within the decision line, although this signal is labeled *ambiguous-no dust* due to the signal framing, this indicates that the SVM might have difficulties classifying signatures located at the edge of the snapshot frame. Note that the signals are zoomed to 15 ms intervals around the interesting features, similar to the examples in Figure 1.
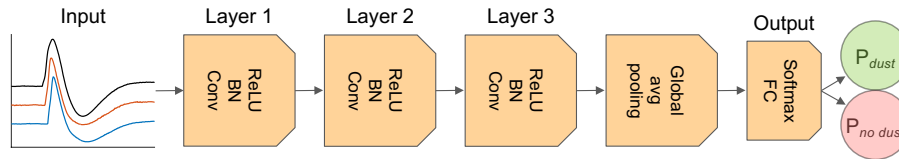
Figure 7. The 3-layer fully convolutional network used for the dust impact classification task. The input to the network is the (3x4096) waveform. The feature extraction process is defined by 3 convolutional layers, consisting of 128, 256 and 128 independent filters with kernel lengths of 8, 5 and 3 weights, respectively. Batch normalization (BN) is used at each convolutional layer to regularize the inputs and the Rectified Linear Unit (ReLU) function was used as the activation function. Finally, the output of the convolutional layers (with dimension 128x4096) is averaged in the global pooling layer to a feature vector with dimension (128x1). The class score is then determined in a Fully Connected (FC) network layer and the output label probabilities (P_{dust}, $P_{no\ dust}$) are calculated using the softmax function. The figure is adopted from Wickstrøm et al. (2021).

3.5.3 Testing the Convolutional Neural Network

In order to visualize the features extracted by the CNN, we employ the t-distributed Stochastic Neighbor Embedding (t-SNE) method (Van der Maaten and Hinton, 2008). The t-SNE method is used for visualizing high-dimensional data by assigning each observation a location in a 2–D space such that similar observations are modeled by nearby points while dissimilar observations are modeled by distant points with high probability. The (128x1) testing feature vectors, extracted in the global pooling layer, are presented in a 2–D t-SNE map in Figure 8, along with a visualization of the CNN classification performance.

Overall, the CNN obtains a high ($\gtrsim$ 95%) classification accuracy and might therefore be suitable for automatic processing of electric field signals observed by the RPW instrument on board the Solar Orbiter.

3.5.4 Explainability of the Convolutional Neural Network

Neural networks have traditionally been regarded as black boxes (Shwartz-Ziv and Tishby, 2017; Alain and Bengio, 2016), where the network carries out the desired task, but the network decisions are difficult to interpret. However, progress has been made in recent years for making the neural network decisions more accessible and easier to interpret (i.e. explainable) for human users (Samek et al., 2021). In this section, we analyze the CNN decisions by employing Class Activation Maps and the previously described t-SNE method.

Class Activation Maps (CAMs) (Zhou et al., 2016) highlights the regions of the data that are important for a considered label (c) by analyzing the features extracted in the global pooling layer and the weights in the FC layer that are associated with label (c), see e.g. (Wang et al., 2017) for a detailed description. The outcome of the CAM analysis is that we can visualize the sections of the signal that are influential for the CNN classification decision. Figure 9 presents the CAM analysis of the signal examples in Figure 1 along with an illustration of the signal features in a dimension-reduced t-SNE space. Note that the
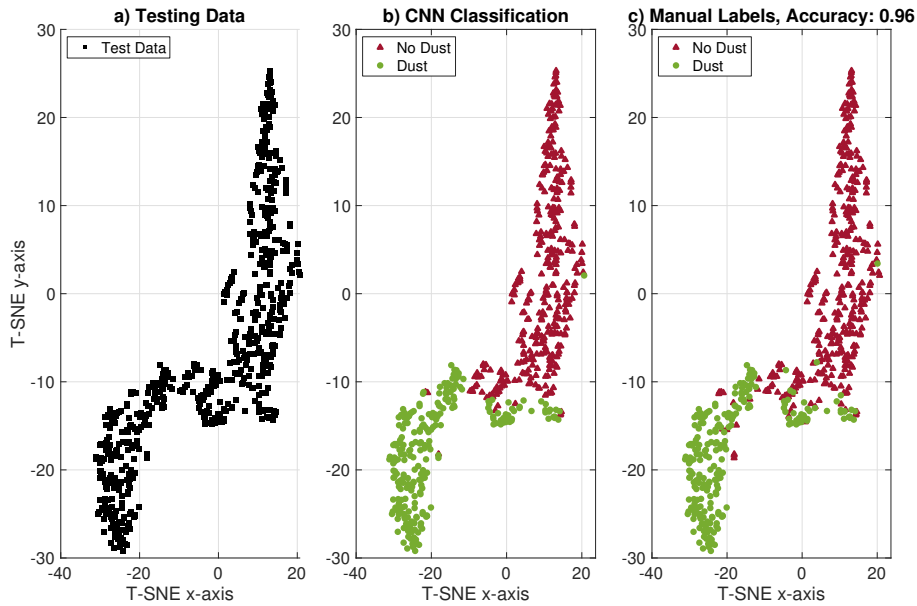
Figure 8. a) The testing data (600 observations with hidden labels) is visualized by a dimension-reduced t-SNE map where similar feature vectors are modeled by nearby points while dissimilar observations are modeled by distant points with high probability. b) The testing data classified by the trained CNN. c) The "true" manual labels are presented. Only a few observations, predominantly in the transition region between the *dust* and *no dust* observations are confused. An overall classification accuracy of 96% is calculated by comparing the labels predicted by the CNN to the manual labels. Note that the presented testing data is the same data set that was used to test the SVM classifier, illustrated in Figure 5.

t-SNE mapping in Figure 9 is different from the t-SNE mapping in Figure 8, since Figure 9 considers a different CNN where the signal examples are specifically excluded from the training data.

The CAM analysis in Figure 9 illustrates that the CNN makes classification decisions that are comprehensible (in most cases). It is however interesting to note that signal c), manually labeled as *dust*, is erroneously classified as *no dust* by the CNN, and that this decision is largely based on the tail (the relaxation period) of the impact signal. It should however be noted that it is more difficult to explain the *no dust* predictions than the *dust* predictions since the *no dust* CNN decisions are based on the lack of a signature (*dust* impact), rather than on the presence of a signature. In addition, signal d), manually labeled as *ambiguous-no dust*, is classified as *dust* by the CNN, and this decision is based on a wide region of the signal with emphasis on the tail of the (ambiguous) dust impact signal, this section might not have been highlighted as particularly important by a human expert.

In general however, the CNN achieves a high accuracy ($\gtrsim 95\%$) and makes decisions that are mostly in-line with human interpretation. It is therefore reasonable to infer that the CNN will have a performance comparable to the agreement level
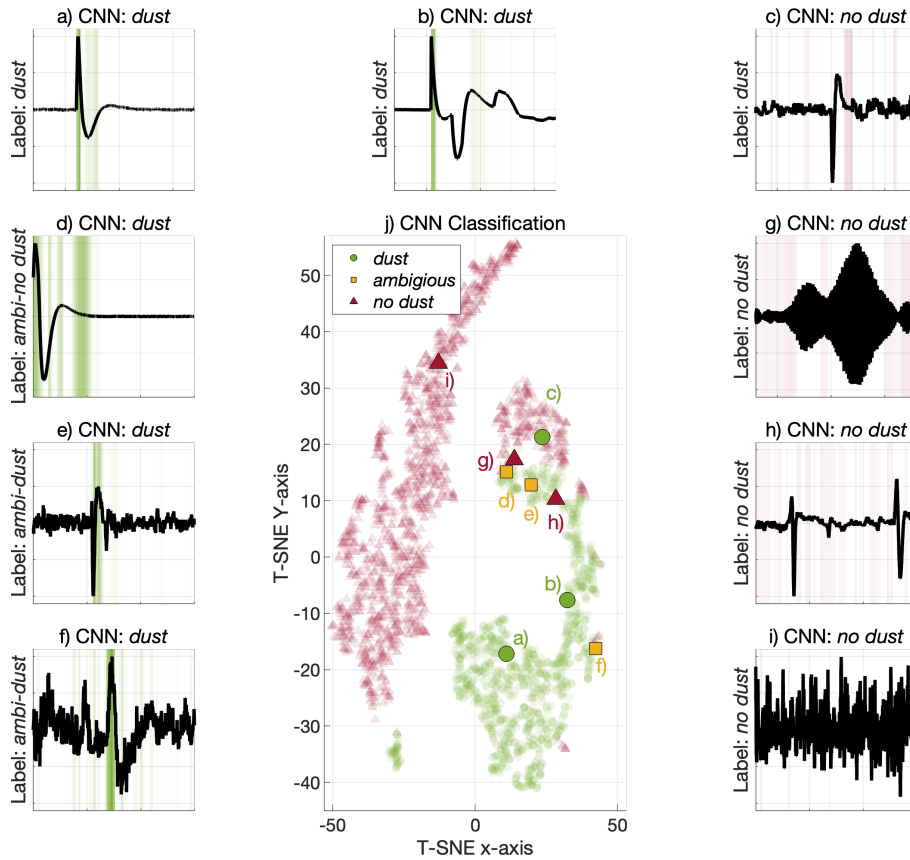
Figure 9. The signal examples and the CAM analysis are presented in Sub-figures (a-i), the manual labels are indicated along the y-axis and the predicted label, classified by the CNN, is presented in the sub-plot titles. The highlighted green color indicates the CAM values associated with the *dust* class, the green regions therefore emphasize the regions that are considered important by the CNN for detecting dust impact signatures. Similarly, the red color indicates the regions that are influential for the *no dust* class. Note that the signals are zoomed to 15 ms intervals around the interesting features, similar to Figure 1 and Figure 6. Sub-figure j) presents the associated signal examples in the t-SNE space along with the training data signals as transparent points. The *dust* signals are illustrated by the green dots, the *ambiguous* signal examples are marked in yellow and the *no dust* signals are indicated in red. The t-SNE map shows that the clear *dust* signals (a-b) are distinctly located in a green (*dust*) region whereas the clear *no dust* signal i) is distinctly located in a red (*no dust*) region. The remaining signals are located in more mixed regions. It should however be noted that the observations are represented by a 128-dimensional feature vector in the CNN and that the (2–D) t-SNE representation presented in Sub-figure j) diminishes a lot of information, meaning that even the signals located in a mixed region of the t-SNE plot might be separable in the 128-dimensional feature vector space.

between human experts, where disagreement predominantly occurs for ambiguous and noisy signals, while clear *dust* and clear *no dust* signals are classified correctly.

4 Results and Discussions

4.1 Analysis of the Classification Performance

The average classification performance is obtained by training and testing the machine learning classifiers over 10 runs, each run with different training and testing sets. The classifiers are initialized from scratch and the training and testing sets are selected independently 10 times by randomization and splitting of the manually labeled data, as indicated by the gray circles in Figure 3. The average class-wise performance of the on-board TDS classifier and the machine learning SVM and CNN classifiers are summarized as confusion matrices in Figure 10. Overall, the CNN has the highest performance for both *dust* and *no dust* classification. In addition, both the SVM and the CNN classifiers obtain stable performances with only small variations for each run.
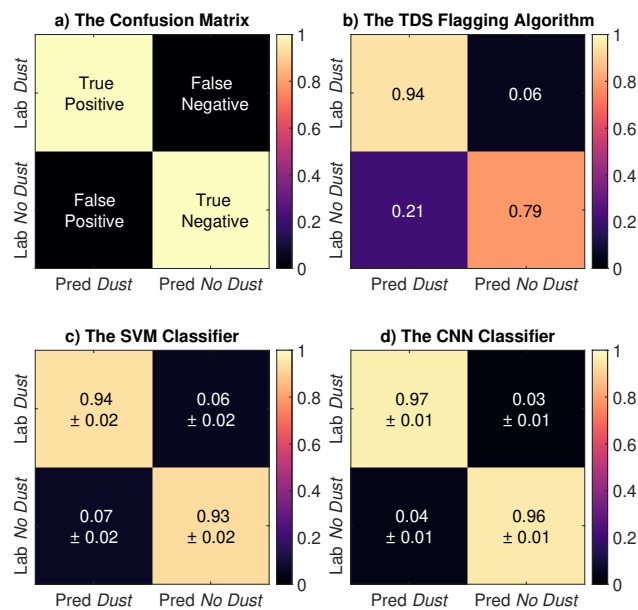
Figure 10. a) The confusion matrix entries are described by the true (correctly classified) and false (erroneously classified) values, as compared to the manual labels (Lab), positive indicate *dust* predictions (Pred) and negative indicate *no dust* predictions. b) The TDS classifier confuses *dust* and *no dust* observations, where a significant proportion (> 0.20) of *dust* predictions are manually labeled as *no dust*. c) The SVM classifier predicts both *dust* and *no dust* observations with a high (> 0.90) accuracy. d) The CNN classifier predicts a very large (> 0.95) proportion of both *dust* and *no dust* observations correctly.

The classification performance is further evaluated by the accuracy, precision, recall and F1 score. The definitions for the performance metrics are included in Appendix B. The average performance metrics, calculated over 10 runs, are summarized in Table 1. Again, the CNN has the highest performance across all metrics. The CNN obtain a significant improvement in the classification performance with a statistical significance at a level of 0.01, computed using a t-test. The t-test was computed in a pairwise manner between both the CNN and the SVM scores, and the CNN and the TDS scores. In all cases, the enhanced performance of the CNN classifier was significant.

Table 1. The TDS, SVM and CNN classification performance metrics: accuracy, precision, recall and F1-score. The SVM and CNN scores and error values are the mean and the standard deviation across 10 training runs. The bold numbers indicate statistically enhanced performance with a significance level of 0.01, computed using a t-test.

Classifier	Accuracy	Precision	Recall	F1 Score
TDS	0.850	0.746	0.944	0.833
SVM	0.936 ± 0.012	0.903 ± 0.027	0.941 ± 0.017	0.921 ± 0.015
CNN	$\mathbf{0.964 \pm 0.006}$	$\mathbf{0.939 \pm 0.020}$	$\mathbf{0.972 \pm 0.008}$	$\mathbf{0.955 \pm 0.008}$

The results from both the confusion matrices and the performance metrics strongly suggest that the SVM and CNN classifiers provide binary classification results with higher reliability than the TDS classifier, and further that the CNN is the most reliable classifier overall. We therefore propose that the CNN classifier (or similar tools) should be considered for post-processing of the TDS data product in statistical studies of dust impacts observed by the Solar Orbiter RPW instrument.

In addition, it should be noted that 134 signals (i.e. 4.5%), out of 3000 manually labeled waveforms, were marked as ambiguous, illustrated by the yellow cylinder in Figure 3, and did not clearly fit into either the *dust* or *no dust* label, see Figure 1 for label examples. It is therefore improbable to achieve a classification accuracy exceeding ∼98% for the considered data set, and an accuracy approaching ∼99% should be considered suspicious and can be an indication of over-fitting.

Both the trained SVM and CNN classifiers are computationally inexpensive to run. 1000 observations are classified in 5 seconds using the SVM model while the CNN classifier requires 50 seconds on a modern laptop, including the needed time for pre-processing and feature extraction. The proposed machine learning classifiers are therefore suitable for processing large data sets with thousands of new observations acquired every month as the Solar Orbiter continues its operation.

4.2 The Dust Impact Rate

In this section, we use the trained classifiers to automatically process a large data set, consisting of 104 032 observations. This data set contains all TSWF observations acquired over a 25-month period, spanning between June 15, 2020, to July 14, 2022,

360 that satisfies the criteria in Section 2.1 (sampling rate of 262.1 kHz, 16384 time-steps and XLD1 mode).

Figure 11 presents the TDS, SVM and CNN daily impact rates with included error estimates. The daily impact rate is calculated from the automatically detected daily dust impact number and the time-dependent TDS-RPW duty cycle. The number of dust particles detected by the Solar Orbiter on each day can be modeled as a Poisson process (Kočiščák et al., 2022),

365 where the variance in the daily count is equal to the daily count number, resulting in the standard deviation errorbars presented in Figure 11. The impact rate function curves are obtained by fitting the dust flux model from Zaslavsky et al. (2021) with an included offset:

$$R = F_{1\mathrm{AU}} S_{col} \left(\frac{r}{1\mathrm{AU}} \right)^{-2} \frac{\nu_{\mathrm{impact}}}{\nu_\beta} \left(\frac{\nu_{\mathrm{impact}}}{\nu_{\mathrm{impact}}(1\mathrm{AU})} \right)^{\alpha\delta} + C \qquad (7)$$

Where $F_{1\mathrm{AU}}$ is the unknown cumulative flux of particles above the detection threshold at 1 AU and $S_{col} = 8\mathrm{m}^2$ is the Solar

370 Orbiter collection area, as defined in Zaslavsky et al. (2021). Furthermore, r is the radial distance from the sun, ν_{impact} is the relative velocity between the spacecraft and the dust particles, assuming a constant radial and azimuthal velocity vector for the dust particles: ν_β = [50 km/s, 0 km/s], and the product $\alpha\delta$ = 1.3, as suggested in Zaslavsky et al. (2021). The assumed constant radial velocity is a good approximation for dust in hyperbolic orbits originating near the Sun that is deflected outward by the radiation pressure force. Finally, we included a constant impact rate offset: C, in order to obtain an improved fit. The

375 description of the dust flux in Equation 7 is based on the assumption that the dust– and spacecraft orbits are in the same orbital plane.

Figure 11 shows that the machine learning classifiers detected significantly more dust particles than the TDS classifier. The SVMs obtained a dust impact detection enhancement of 16% ± 1% while the CNNs had an 18% ± 8% increase. Both the SVM

380 and the CNN classifiers obtain impact rates that are notably higher around the aphelion and distinctly lower in the vicinity of the perihelion, resulting in a lower dynamic range of the impact rates than observed in the TDS data product.

Furthermore, Figure 11 illustrates that the fitted SVM and CNN impact rate function cures are in very good agreement. It is promising that two entirely different machine learning approaches provide comparable impact rates after classifying a large

385 data set (consisting of 104 032 observations) when trained– and tested on a limited data set consisting of 3000 observations. This suggests that both the SVM and CNN classifiers have obtained stable performances and can be used to classify observations outside the domain of the training and testing data.

Still, the shape of the dust impact signal is dependent on the local plasma environment, where influential parameters are:

390 the electron plasma density, the mean electron velocity and the electron temperature (Zaslavsky, 2015; Babic et al., 2022). These parameters will vary throughout the spacecraft orbit. It should therefore be noted that the machine learning classifiers were trained and tested on waveforms acquired over a one-and-a-half-year period, spanning between June 15, 2020, to December 16, 2021. During this period, the Solar Orbiter sampled the interplanetary medium at solar distances ranging from ∼0.5 AU
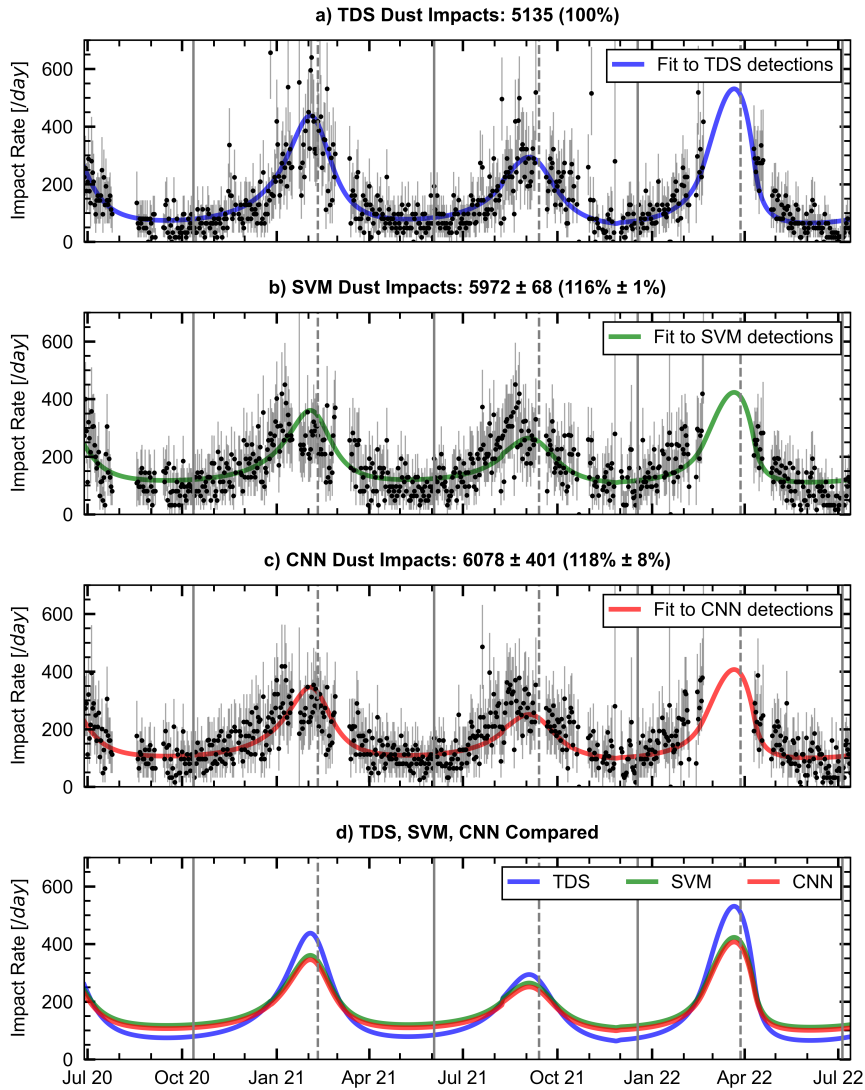
Figure 11. a) The daily dust impact rates according to the TDS classifier. The full vertical lines indicate times when the Solar Orbiter is at aphelion while the dashed lines indicate times at perihelion. b) The median of the daily impact rates classified by 10 trained SVM classifiers. c) The median of the daily impact rates from the 10 CNN classifiers. The impact rate function curves are obtained by fitting the dust flux model from Zaslavsky et al. (2021), Equation 7. d) The impact rate function cures are compared. The SVM and CNN dust impact rates are very similar, whereas the TDS provides notably smaller impact rates at aphelion and higher impact rates at perihelion. The accumulated dust impact detections for the TDS classification algorithm and the mean and standard deviation of the dust impact detections for the 10 CNN and SVM classifiers are presented in the sub-plot titles. Note that the large data gap around April 2022 (perihelion 3) is due to a different observational setup for the Solar Orbiter RPW-TDS system, where the sampling frequency was doubled. This data was excluded since it can not be reliably classified by the SVM/CNN methods without additional data processing and/or training.

to $\sim$1.0 AU. The spacecraft will however reach a minimum solar distance of 0.28 AU, and the performance of the machine learning classifiers might suffer if the observed dust impact shapes in the vicinity of $\sim$0.3 AU are significantly different from the dust impact shapes at $\sim$0.5 AU to $\sim$1.0 AU.

Finally, we note that a dip in the SVM and CNN dust impact rates can be observed in Figure 11, roughly 0.5 – 1 month before perihelia 1 and 2 (no data for perihelion 3). This dip is possibly due to a change in the relative velocity between the spacecraft and the interstellar dust particles, which is upstream at $259°$ in the Ecliptic coordinate system. Still, there is a large natural (Poisson) variation in the dust impact rates at perihelion that make visual analysis difficult with the presented data set. In addition, complicating effects will have an enhanced influence on the daily dust count number towards the sun, such as an enhancement in false detections due to increased variability in the ambient plasma and validity degradation of the dust flux model assumptions in Equation 7 close to the formation region of the hyperbolic dust particles.

5 Conclusions

5.1 Summary and Scientific Implications

We have presented a machine learning-based framework for fully automated detection of dust impacts observed by the Solar Orbiter – Radio and Plasma Waves (RPW) instrument. Two different supervised machine learning approaches were considered: the Support Vector Machine (SVM) and the Convolutional Neural Network (CNN). The CNN classifier obtained the highest performance across all evaluation metrics and achieved 96% $\pm$ 1% overall classification accuracy and 94% $\pm$ 2% dust detection precision, a significant improvement to the currently used on-board TDS classification algorithm with 85% overall classification accuracy and 75% dust detection precision. We therefore conclude that the CNN classifier (or similar tools) should be considered for post-processing of the TDS data product for statistical studies of dust impacts observed by the Solar Orbiter.

The SVM and CNN classifiers were used to analyze 104 032 observations acquired over a two-year period, spanning between June 15, 2020, to July 14, 2022. On average, the machine learning classifiers detected more dust particles than the currently used TDS algorithm, the SVMs had a 16% $\pm$ 1% detection enhancement and the CNNs had an 18% $\pm$ 8% increase. Furthermore, the SVM and CNN classifiers were in very good agreement and both classifiers obtained a notably higher dust impact rate in the vicinity of aphelion and a distinctly lower impact rate at perihelion, as compared to the dynamic range of the TDS impact rates. This might indicate a higher ambient dust distribution than previously observed. This result is significant since it suggests the presence dust populations other than the hyperbolic dust particles in the data. Possible other populations are interstellar dust and interplanetary dust in bound orbits.

The labeled data and the trained SVM and CNN classifiers are available online with included user instructions. The proposed method and the presented classifiers can thus provide the interplanetary dust community with thoroughly tested and more

reliable data products than those currently in use. The daily dust count numbers from the CNN classification were employed by Kočiščák et al. (2022) to infer meaningful physical properties of the dust population by modeling the number of dust detections within a day as a Poisson-distributed random variable. Kočiščák et al. (2022) further demonstrated that the same procedure did not provide dust parameters that were in-line with prior knowledge when using the daily dust detections from the TDS classification. This result is independent of the manually labeled testing data, which might be prone to biases, and further suggests that the CNN approach provides more reliable data products than the currently used TDS algorithm.

5.2 Outlook and Method Constraints

The presented machine learning classifiers may be considered for on-board processing of the observed electric field signals. However, the trained SVM and CNN classifiers presented in this article are trained on Triggered Snapshot WaveForms (TSWF) data, and should not be used for processing 'untriggered' signals without additional training and testing on 'untriggered' data. Additional training can also be used to further enhance the performance of the machine learning classifiers. In particular, adding labeled data acquired near the sun ($\sim$0.3 AU) and during periods of strong solar activity will likely improve the overall accuracy and make the machine learning classifiers more robust to challenging conditions.

It should also be noted that the classifiers presented in this work are trained and tested on data labeled by one scientist, although with consultations with other experts. Labeled data from several experts can provide machine learning classifiers that are more in-line with the labeling consensus in the interplanetary dust community. Additional labeling can also be used to extend the machine learning classifiers to include automatic detection of other characteristic signatures, such as: ion-acoustic, Langmuir and solitary waves (Soucek et al., 2021).

Finally, we would like to highlight that a machine learning-based framework can be developed for automatic post-processing of data acquired by radio and plasma waves instruments on-board other spacecrafts, such as: the Solar Terrestrial Relations Observatory (STEREO) (Zaslavsky et al., 2012), Wind (Malaspina et al., 2014), and the Parker Solar Probe (Szalay et al., 2020). Automatic and reliable detection of dust impact signals observed by multiple instruments at several locations and over several years will likely facilitate statistical studies that will enhance our understanding of the role of dust in the inner solar system, beyond what is attainable with the data products that are currently in use.

Code and data availability. The code used for this work, the trained classifiers and the training and testing data sets are available at: https: //github.com/AndreasKvammen/ML_dust_detection. The Triggered Snapshot WaveForms (TSWF) data files can be downloaded at: https: //rpw.lesia.obspm.fr/roc/data/pub/solo/rpw/data/L2/tds_wf_e/

Appendix A: Graphical User Interface for Manual Labeling

Figure A1 presents the Graphical User Interface (GUI) that was used to manually label all considered (3000) signals into either *dust* or *no dust*. In addition, efforts were made to use a similar setup (with the same monitor and figure resolution) throughout the manual labeling in order to reduce bias effects.
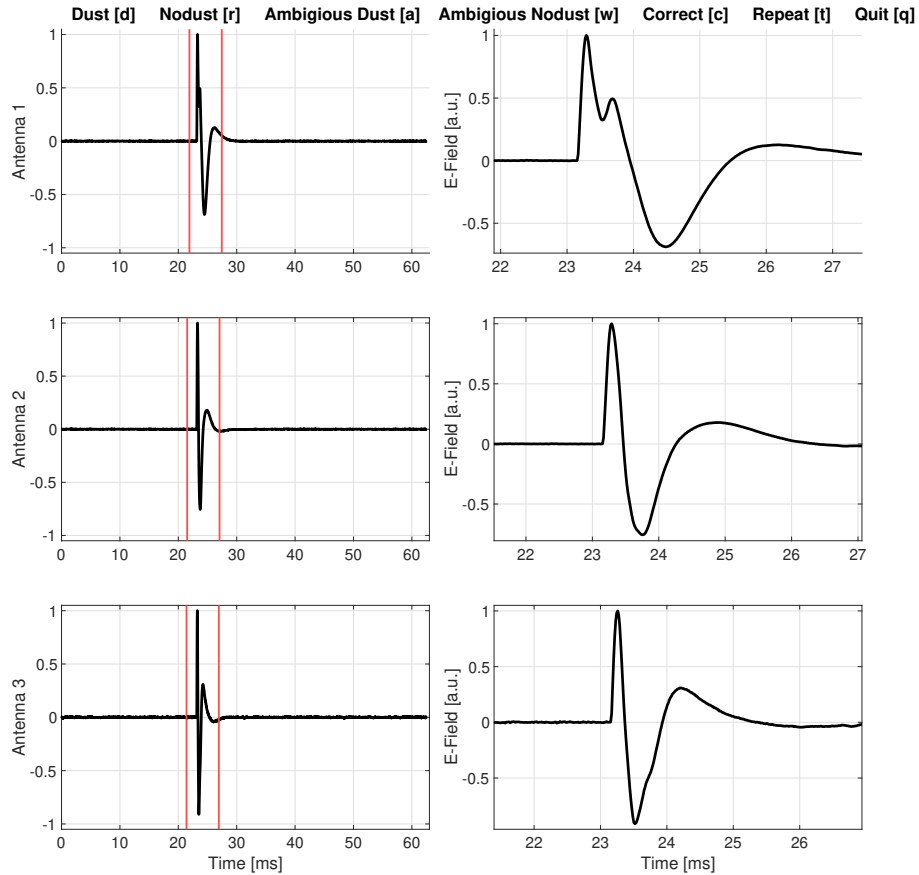
Figure A1. The manual labeling user interface showing a signal observed July 28, 2021. The left column displays the full snapshot (from 0 to ∼63 ms) at all antennas. An area of interest is selected by adjusting the red vertical lines. The right column displays the signal within the area of interest. The signal can be labeled as *dust* by pressing the [d] key on the keyboard and *no dust* by pressing the [r] key. The signal is indicated to be ambiguous if the waveform does not fit clearly into either of the two labels, note however that signals indicated to be ambiguous were also labeled into either *dust* or *no dust* using the [a] and [w] keys. There is also an option to correct [c] the previously labeled signal (in case of an error), repeat [t] the area of interest selection and quit [q] the manual labeling user interface.

Appendix B: The Classification Performance Metrics

The classification performance metrics are calculated using the True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) values, defined by comparing the predicted classes and the manually labeled classes, illustrated in Figure 10.

The overall accuracy of the classifier is the proportion of observations that were correctly predicted by the classifier. The accuracy is mathematically defined as:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{B1}$$

Precision (in this case) is defined as the proportion of data points predicted by the classifier as *dust*, whose "true" label is indeed *dust*. Precision is therefore calculated as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{B2}$$

Recall (in this case) is the proportion of observations manually labeled as *dust*, that were correctly predicted as *dust* by the classifier. Recall is defined as:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{B3}$$

The F1 score acts as a weighted average of precision and recall and is calculated as:

$$\text{F1} = 2 \left(\frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \right) \tag{B4}$$

Author contributions. AK wrote the article text, trained and tested the machine learning classifiers and manually labeled the waveforms. KW aided the development of the machine learning classifiers, analyzed the machine learning performance metrics and commented/edited the article. SK performed the dust impact rate analysis, aided with the theoretical background and commented/edited the article. JV and LN contributed to the analysis of the TDS waveforms and the theoretical background. AZ and KRB contributed with the theoretical background and helpful discussions. AG contributed with knowledge of the Solar Orbiter data availability and discussions on the dust waveform shapes. DP and JS provided the data used for this work and explained the data content. IM is the main contributor to the theoretical background, aided the article with numerous comments/suggestions/discussions and shared knowledge that was crucial for this work.

Competing interests. The authors declare that there are no competing interests.

Acknowledgements. AK thanks Audun Theodorsen for aiding the motivation and objective of the article. In addition, AK thanks Juha Vierienen, Björn Gustavsson and Patrick Guio for helpful discussions. AK acknowledges support from the Research Council of Norway

485   (grant numbers 262941 and 326039). SK is supported by the Tromsø Research Foundation under grant 19_SG_AT. JV, DP and JS acknowledge the support of Czech Science Foundation grant 22-10775S. KRB acknowledges the support from the Ministry of Education, Science and Technological Development of the Republic of Serbia through contract no. 451-03-9/2021-14/200104. The authors would like to thank the reviewers for their valuable suggestions and appropriate comments.

References

490 Alain, G. and Bengio, Y.: Understanding intermediate layers using linear classifier probes, ArXiv, 2016.

Aubier, M., Meyer-Vernet, N., and Pedersen, B.: Shot noise from grain and particle impacts in Saturn's ring plane, Geophysical Research Letters, 10, 5–8, 1983.

Babic, K. R., Zaslavsky, A., Issautier, K., Meyer-Vernet, N., and Onic, D.: An analytical model for dust impact voltage signals and its application to STEREO/WAVES data, Astronomy & Astrophysics, 659, A15, 2022.

495 Boser, B. E., Guyon, I. M., and Vapnik, V. N.: A training algorithm for optimal margin classifiers, in: Proceedings of the fifth annual workshop on Computational learning theory, pp. 144–152, 1992.

Bougeret, J.-L., Kaiser, M. L., Kellogg, P. J., Manning, R., Goetz, K., Monson, S., Monge, N., Friel, L., Meetre, C., Perche, C., et al.: Waves: The radio and plasma wave investigation on the Wind spacecraft, Space Science Reviews, 71, 231–263, 1995.

Collette, A., Grün, E., Malaspina, D., and Sternovsky, Z.: Micrometeoroid impact charge yield for common spacecraft materials, Journal of Geophysical Research: Space Physics, 119, 6019–6026, 2014.

500

Cortes, C. and Vapnik, V.: Support-vector networks, Machine learning, 20, 273–297, 1995.

Fawaz, H. I., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D. F., Weber, J., Webb, G. I., Idoumghar, L., Muller, P.-A., and Petitjean, F.: InceptionTime: Finding AlexNet for time series classification, Data Mining and Knowledge Discovery, 34, 1936–1962, https://doi.org/10.1007/s10618-020-00710-y, 2020.

505 Glorot, X., Bordes, A., and Bengio, Y.: Deep Sparse Rectifier Neural Networks, in: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, edited by Gordon, G., Dunson, D., and Dudík, M., vol. 15 of *Proceedings of Machine Learning Research*, pp. 315–323, PMLR, Fort Lauderdale, FL, USA, https://proceedings.mlr.press/v15/glorot11a.html, 2011.

Goodfellow, I., Bengio, Y., and Courville, A.: Deep learning, MIT press, 2016.

Grün, E., Zook, H. A., Fechtig, H., and Giese, R.: Collisional balance of the meteoritic complex, Icarus, 62, 244–272, 1985.

510 Gurnett, D. A., Grün, E., Gallagher, D., Kurth, W., and Scarf, F.: Micron-sized particles detected near Saturn by the Voyager plasma wave instrument, Icarus, 53, 236–254, 1983.

He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, in: IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778, IEEE Computer Society, https://doi.org/10.1109/CVPR.2016.90, 2016.

Holzinger, A., Langs, G., Denk, H., Zatloukal, K., and Müller, H.: Causability and explainability of artificial intelligence in medicine, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9, e1312, 2019.

515

Hornung, K., Malama, Y. G., and Kestenboim, K. S.: Impact vaporization and ionization of cosmic dust particles, Astrophysics and Space Science, 274, 355–363, 2000.

Ioffe, S. and Szegedy, C.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, in: nternational Conference on Machine Learning, edited by Bach, F. and Blei, D., vol. 37 of *Proceedings of Machine Learning Research*, pp. 448–456, PMLR, Lille, France, https://proceedings.mlr.press/v37/ioffe15.html, 2015.

520

Ishimoto, H.: Modeling the number density distribution of interplanetary dust on the ecliptic plane within 5AU of the Sun, Astronomy and Astrophysics, 362, 1158–1173, 2000.

Karim, F., Majumdar, S., Darabi, H., and Harford, S.: Multivariate LSTM-FCNs for time series classification, Neural Networks, 116, 237–245, https://doi.org/10.1016/j.neunet.2019.04.014, 2019.

525 Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., and Fei-Fei, L.: Large-Scale Video Classification with Convolutional Neural Networks, in: 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1725–1732, https://doi.org/10.1109/CVPR.2014.223, 2014.

Kingma, D. and Ba, J.: Adam: A Method for Stochastic Optimization, in: International Conference on Learning Representations, 2014.

Kočiščák, S., Kvammen, A., Mann, I., Sørbye, S. H., Theodorsen, A., and Zaslavsky, A.: Modelling Solar Orbiter Dust Detection Rates in
530 Inner Heliosphere as a Poisson Process, arXiv preprint arXiv:2210.03562, 2022.

Kvammen, A., Wickstrøm, K., McKay, D., and Partamies, N.: Auroral image classification with deep neural networks, Journal of Geophysical Research: Space Physics, 125, e2020JA027 808, 2020.

Maksimovic, M., Bale, S., Chust, T., Khotyaintsev, Y., Krasnoselskikh, V., Kretzschmar, M., Plettemeier, D., Rucker, H., Souček, J., Steller, M., et al.: The solar orbiter radio and plasma waves (rpw) instrument, Astronomy & Astrophysics, 642, A12, 2020.

535 Malaspina, D., Horányi, M., Zaslavsky, A., Goetz, K., Wilson III, L., and Kersten, K.: Interplanetary and interstellar dust observed by the Wind/WAVES electric field instrument, Geophysical Research Letters, 41, 266–272, 2014.

Malaspina, D. M., Newman, D. L., Willson III, L. B., Goetz, K., Kellogg, P. J., and Kerstin, K.: Electrostatic solitary waves in the solar wind: Evidence for instability at solar wind current sheets, Journal of Geophysical Research: Space Physics, 118, 591–599, 2013.

Malaspina, D. M., Szalay, J. R., Pokornỳ, P., Page, B., Bale, S. D., Bonnell, J. W., de Wit, T. D., Goetz, K., Goodrich, K., Harvey, P. R., et al.:
540 In situ observations of interplanetary dust variability in the inner heliosphere, The Astrophysical Journal, 892, 115, 2020.

Mann, I. and Czechowski, A.: Dust destruction and ion formation in the inner solar system, The Astrophysical Journal, 621, L73, 2005.

Mann, I. and Czechowski, A.: Dust observations from Parker Solar Probe: dust ejection from the inner Solar System, Astronomy & Astrophysics, 650, A29, 2021.

Mann, I., Kimura, H., Biesecker, D. A., Tsurutani, B. T., Grün, E., McKibben, R. B., Liou, J.-C., MacQueen, R. M., Mukai, T., Guhathakurta,
545 M., et al.: Dust near the Sun, Space science reviews, 110, 269–305, 2004.

Mann, I., Nouzák, L., Vaverka, J., Antonsen, T., Fredriksen, Å., Issautier, K., Malaspina, D., Meyer-Vernet, N., Pavlů, J., Sternovsky, Z., Stude, J., Ye, S., and Zaslavsky, A.: Dust observations with antenna measurements and its prospects for observations with Parker Solar Probe and Solar Orbiter, Annales Geophysicae, 37, 1121–1140, https://doi.org/10.5194/angeo-37-1121-2019, 2019.

Montavon, G., Orr, G. B., and Müller, K.-R., eds.: Neural Networks: Tricks of the Trade, Springer Berlin Heidelberg,
550 https://doi.org/10.1007/978-3-642-35289-8, 2012.

Müller, D., Cyr, O. S., Zouganelis, I., Gilbert, H. R., Marsden, R., Nieves-Chinchilla, T., Antonucci, E., Auchère, F., Berghmans, D., Horbury, T., et al.: The solar orbiter mission-science overview, Astronomy & Astrophysics, 642, A1, 2020.

Samek, W. et al.: Explaining Deep Neural Networks and Beyond: A Review of Methods and Applications, Proceedings of the IEEE, pp. 247–278, 2021.

555 Shwartz-Ziv, R. and Tishby, N.: Opening the Black Box of Deep Neural Networks via Information, ArXiv, abs/1703.00810, 2017.

Soucek, J., Píša, D., Kolmasova, I., Uhlir, L., Lan, R., Santolík, O., Krupar, V., Kruparova, O., Baše, J., Maksimovic, M., et al.: Solar Orbiter Radio and Plasma Waves–Time Domain Sampler: In-flight performance and first results, Astronomy & Astrophysics, 656, A26, 2021.

Szalay, J., Pokornỳ, P., Bale, S., Christian, E., Goetz, K., Goodrich, K., Hill, M., Kuchner, M., Larsen, R., Malaspina, D., et al.: The near-sun dust environment: initial observations from parker solar probe, The Astrophysical Journal Supplement Series, 246, 27, 2020.

560 Theodoridis, S. and Koutroumbas, K.: Chapter 3 - Linear Classifiers, in: Pattern Recognition (Fourth Edition), edited by Theodoridis, S. and Koutroumbas, K., pp. 91 – 150, Academic Press, Boston, fourth edition edn., https://doi.org/https://doi.org/10.1016/B978-1-59749-272-0.50004-9, 2009.

Trosten, D. J., Strauman, A. S., Kampffmeyer, M., and Jenssen, R.: Recurrent Deep Divergence-based Clustering for Simultaneous Feature Learning and Clustering of Variable Length Time Series, in: ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 3257–3261, https://doi.org/10.1109/ICASSP.2019.8682365, 2019.

Van der Maaten, L. and Hinton, G.: Visualizing data using t-SNE., Journal of machine learning research, 9, 2008.

Vaverka, J., Pellinen-Wannberg, A., Kero, J., Mann, I., De Spiegeleer, A., Hamrin, M., Norberg, C., and Pitkänen, T.: Potential of earth orbiting spacecraft influenced by meteoroid hypervelocity impacts, IEEE Transactions on Plasma Science, 45, 2048–2055, 2017.

Vech, D. and Malaspina, D. M.: A novel machine learning technique to identify and categorize plasma waves in spacecraft measurements, Journal of Geophysical Research: Space Physics, 126, e2021JA029 567, 2021.

Villar, J. R., Vergara, P., Menéndez, M., de la Cal, E., González, V. M., and Sedano, J.: Generalized Models for the Classification of Abnormal Movements in Daily Life and its Applicability to Epilepsy Convulsion Recognition, International Journal of Neural Systems, 26, 1650 037, https://doi.org/10.1142/s0129065716500374, 2016.

Wang, Z., Yan, W., and Oates, T.: Time series classification from scratch with deep neural networks: A strong baseline, in: 2017 International joint conference on neural networks (IJCNN), pp. 1578–1585, IEEE, 2017.

Wickstrøm, K., Mikalsen, K. Ø., Kampffmeyer, M., Revhaug, A., and Jenssen, R.: Uncertainty-Aware Deep Ensembles for Reliable and Explainable Predictions of Clinical Time Series, IEEE Journal of Biomedical and Health Informatics, 25, 2435–2444, https://doi.org/10.1109/JBHI.2020.3042637, 2021.

Wickstrøm, K., Kampffmeyer, M., Mikalsen, K. Ø., and Jenssen, R.: Mixing up contrastive learning: Self-supervised representation learning for time series, Pattern Recognition Letters, 155, 54–61, https://doi.org/10.1016/j.patrec.2022.02.007, 2022.

Zaslavsky, A.: Floating potential perturbations due to micrometeoroid impacts: Theory and application to S/WAVES data, Journal of Geophysical Research: Space Physics, 120, 855–867, 2015.

Zaslavsky, A., Meyer-Vernet, N., Mann, I., Czechowski, A., Issautier, K., Le Chat, G., Pantellini, F., Goetz, K., Maksimovic, M., Bale, S., et al.: Interplanetary dust detection by radio antennas: Mass calibration and fluxes measured by STEREO/WAVES, Journal of Geophysical Research: Space Physics, 117, 2012.

Zaslavsky, A., Mann, I., Soucek, J., Czechowski, A., Píša, D., Vaverka, J., Meyer-Vernet, N., Maksimovic, M., Lorfèvre, E., Issautier, K., et al.: First dust measurements with the Solar Orbiter Radio and Plasma Wave instrument, Astronomy & Astrophysics, 656, A30, 2021.

Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., and Torralba, A.: Learning deep features for discriminative localization, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2921–2929, 2016.