

Article

Identification of the Interface in a Binary Complex Plasma Using Machine Learning

He Huang¹, Mierk Schwabe²  and Cheng-Ran Du^{1,3,*}¹ College of Science, Donghua University, Shanghai 201620, China; 2161507@mail.dhu.edu.cn² Institut für Materialphysik im Weltraum, Deutsches Zentrum für Luft- und Raumfahrt (DLR), 82234 Weßling, Germany; mierk.schwabe@dlr.de³ Magnetic Confinement Fusion Research Centre, Ministry of Education, Shanghai 201620, China

* Correspondence: chengran.du@dhu.edu.cn

Received: 18 February 2019; Accepted: 6 March 2019; Published: 12 March 2019



Abstract: A binary complex plasma consists of two different types of dust particles in an ionized gas. Due to the spinodal decomposition and force imbalance, particles of different masses and diameters are typically phase separated, resulting in an interface. Both external excitation and internal instability may cause the interface to move with time. Support vector machine (SVM) is a supervised machine learning method that can be very effective for multi-class classification. We applied an SVM classification method based on image brightness to locate the interface in a binary complex plasma. Taking the scaled mean and variance as features, three areas, namely small particles, big particles and plasma without dust particles, were distinguished, leading to the identification of the interface between small and big particles.

Keywords: complex plasma; machine learning

A complex plasma is a weakly ionized gas containing small solid particles [1,2]. The particles are highly charged by collecting ions and electrons, and interact strongly with each other. This system allows experimental studies of various physical processes occurring in liquids and solids at the kinetic level [3], such as plasma crystals [4,5], acoustic waves [6], and turbulence [7,8]. A complex plasma consisting of two differently sized microparticle types is known as binary complex plasma. Under certain conditions, two types of particles can be mixed and form a glassy system [9]. Other phenomena, such as phase separation [10,11] and lane formation [12], can also be studied in such systems. Under microgravity conditions, phase separation can occur due to the imbalance of forces despite the criterion of spinodal decomposition not being fulfilled [13]. The phase separated system then allows carrying out dedicated experiments such as wave transmission [14], and interaction of spheres with differently sized particles [15].

In complex plasmas, the particle radius usually ranges from a few to hundreds of microns. Then, the particles can be illuminated by a laser and directly recorded by a video camera [16]. The recorded image sequences can be further analyzed by using tracking algorithms [17,18] and the trajectories of individual particles can be obtained. This provides a basis to study the dynamics and interactions in a multi-particle system. However, under certain conditions, a large region of interest needs to be recorded with a high recording rate. As a result, the spatial resolution has to be sacrificed with currently affordable recording technology. Therefore, advanced image recognition techniques are desirable in the research of complex plasmas.

In recent years, machine learning has been widely applied to image recognition [19–21], such as face recognition [22], and handwriting recognition [23,24]. Machine learning methods include many different algorithms [25], such as decision trees [26], neural networks [27], Bayesian networks [28], k-Nearest Neighbor (kNN) [29] and support vector machine (SVM) [30]. Among these algorithms,

the SVM method is one of the common supervised learning models for classification and regression problems [30]. Given a training set, the aim of SVM is to find the “maximum-margin hyperplane” that divides the group of points and maximizes the distance between the hyperplane and the nearest point from either group [30]. Once the hyperplane has been established by SVM training algorithm, we can classify samples not in the training set. The SVM method has been applied in solving various practical problems, such as text and hypertext categorization [31], classification of images [32] and other scientific research projects [33,34].

We applied the SVM method to achieve automatic identification of the interface in a binary complex plasma based on the brightness of the recorded images. The experiments were performed in the PK-3 Plus Laboratory on board the International Space Station (ISS). Technical details of the setup can be found in Reference [35]. An argon plasma was produced by a capacitively coupled radio-frequency (rf) generator in push–pull mode at 13.56 MHz. We prepared a binary complex plasma by injecting two types of particles. The first type was melamine formaldehyde (MF) particles of a diameter of $2.55\ \mu\text{m}$ with a mass $m_b = 1.34 \times 10^{-14}\ \text{kg}$, while the second type was SiO_2 particles of a diameter of $1.55\ \mu\text{m}$ with a mass $m_s = 3.6 \times 10^{-15}\ \text{kg}$. Using the quadrant view (QV) camera [35], a cross-section of the left half of the particle cloud (illuminated by a laser sheet) was recorded with a frame rate of 50 frames-per-second (fps).

Due to the gradient of plasma potential close to the chamber wall, both particle types were confined in the bulk plasma region, forming a three-dimensional (3D) cloud with a cylindrical symmetry. Figure 1a shows the left half of the cross section of the particle cloud. The cloud of small particles, big particles, and microparticle-free plasma can be easily distinguished with the naked eye. The two particle types were phase-separated due to the following reasons: First, the disparity of particle size ($\Delta d/\bar{d} \approx 0.5$) was much larger than the critical value of spinodal decomposition [10]. Second, both particle types were subjected to two forces under microgravity conditions, namely the ion drag force (directed outwards from the center of the plasma chamber) and the electric-field force (directed inwards to the center of the plasma chamber). The total force acting on the two particle types had a subtle difference depending on the particle diameter [13]. The synergistic effects of spinodal decomposition and force difference led to the instantaneous phase separation. Particularly the second effect drove the small particles into the inner part of the particle cloud and left big particles outside [14].

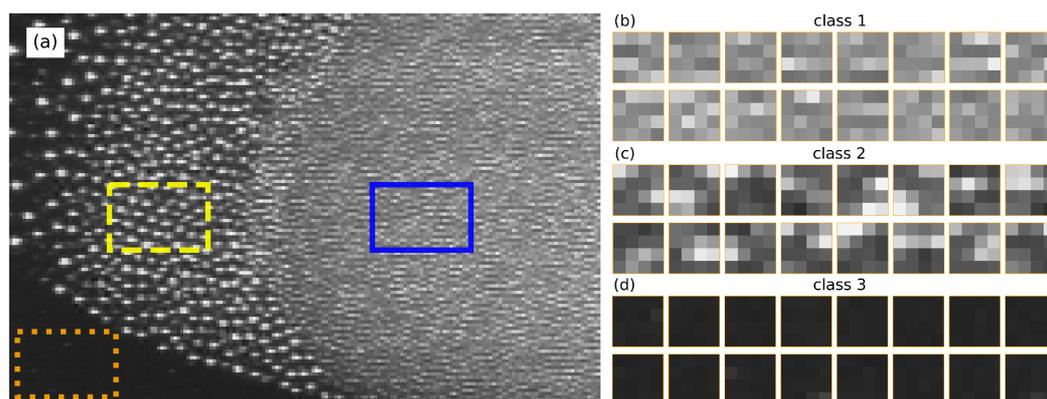


Figure 1. (a) Single image extracted from experiment recordings and the selected areas for training (highlighted by the rectangles). (b) Examples of samples of the small particle cloud corresponding to class 1 (highlighted by the blue solid rectangle). (c) Examples of samples of the big particle cloud corresponding to class 2 (highlighted by the yellow dashed rectangle). (d) Examples of samples of the background plasma without dust particles corresponding to class 3 (highlighted by the orange dotted rectangle). We randomly selected four by four pixel tiles from each area and calculated the mean and variance of the pixel brightness in each grid area as features for the SVM method. The grid size defines the spatial resolution.

To identify the interface of the particle cloud, we applied the SVM method to distinguish two particle types and the background microparticle-free plasma, which defined the three possible classes. First, we prepared the training sets and defined the features. Three areas (representing three classes) were selected in one frame of the experimental video (marked by rectangles in Figure 1a). Class 1 is the small particles area, class 2 is the big particles area, and class 3 is the background plasma without dust particles. The areas are far from the interface to avoid ambiguity, and their class can be easily identified with the naked eye. To define features, we randomly selected 4×4 pixel tiles from the selected area. Figure 1b–d shows a part of the tile collections. Here, each tile was one sample. The mean and the variance of each sample were selected as features. As the variance was much bigger than the mean, we rescaled both features so that their magnitudes were comparable:

$$S_i^j = \frac{x_i^j - x_{min}^j}{x_{max}^j - x_{min}^j} \tag{1}$$

where i stands for sample i , $j = 1$ represents variance and $j = 2$ represents mean. x_{min}^j and x_{max}^j are the minimum and maximum of the j feature of all samples of all classes. We labeled each sample as 1, 2, or 3 based on the area it belonged to and repeated the process for a few frames. All samples were randomly divided into two sets, namely the training set and the test set. Here, we selected 216 samples in each class for training. Table 1 shows the values of the features and labels of a few samples.

Table 1. A few samples of the scaled training set.

| No. | Variance | Mean | Label |
|-----|----------|-------|-------|
| 1 | 0.184 | 0.893 | 1 |
| 2 | 0.000 | 0.019 | 3 |
| 3 | 0.201 | 0.915 | 1 |
| 4 | 0.998 | 0.712 | 2 |
| 5 | 0.702 | 0.477 | 2 |
| ⋮ | ⋮ | ⋮ | ⋮ |

Next, we applied one of the support vector machine methods, namely the support vector classification (SVC), to the training set. The algorithm was provided by scikit-learn API [36,37]. The SVC method implemented the “one-against-one” approach for multi-class classification. The “one-against-one” approach involves constructing a machine for each pair of classes. Thus, three classifiers were constructed, each constructed by two different classes of training points, i.e., “big” or “small particles”, “big” or “particle-free”, and “small” or “particle-free”. When applied to a sample, each classifier gave one vote to the winning class, and the sample was labeled with the class having most votes [38]. The parameter C in the scikit-learn API was set to 5, which is the penalty parameter of the error term. A larger values of C allows for fewer incorrect classifications. The linear kernel type was used in the algorithm and the remaining parameters were the default parameters. More detailed instructions can be seen in scikit-learn API [36,37].

The results of the SVC classification are shown in Figure 2. Dark blue dots represent samples in the background microparticle-free plasma, light blue dots represent samples in the cloud of big particles, and dark red dots represent samples in the cloud of small particles. The classification lines are indicated by the blue and red dotted lines. When the sample is located below the blue line in Figure 2, the corresponding pixels represent the background microparticle-free plasma. When the sample is above the blue line and above the red line, the corresponding pixels belong to the cloud of big particles. When the sample is below the red line, the corresponding pixels belong to the cloud of small particles.

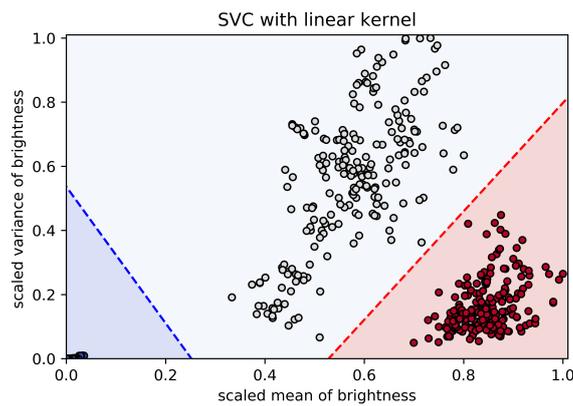


Figure 2. The results of the SVC classification. Dark blue dots represent samples in the background microparticle-free plasma, light blue dots represent samples of the big particle cloud, and dark red dots represent samples of the small particle cloud. The classification lines are indicated by the blue and red dotted lines.

To evaluate the accuracy of the classification, we selected 1000 samples (with knowledge of their class) from the test set, which was set aside before the training. The accuracy was defined as the percentage of correctly classified samples out of all the selected test samples. Here, we studied the dependence of the accuracy of the SVM model on the size of the training set and the spatial resolution. As we can see in Figure 3a, with a small set of training samples, the accuracy depended on which samples we selected. On the one hand, if the randomly selected training samples represented the overall properties of the class, the resulting accuracy was high. On the other hand, if the selected samples represented only a part of the properties, the accuracy was low. This randomness led to a relatively low accuracy with big standard deviations. However, the accuracy improved quickly as the number of training samples increased. When the number of training samples exceeded 40, the accuracy approached 100%. For the classification, the spatial resolution depended on the size of each tile. Bigger tiles include more information in each tile but lead to lower spatial resolution. In Figure 3b, we show the dependence of the accuracy of the classification on the side length of each tile. With a considerable size of the training set, the accuracy already exceeded 90% when each tile included only four (2×2) pixels. This shows that the mean value of the pixel brightness alone played a significant role in classification. The accuracy rose further with the tile size. It exceeded 98% when there were more than sixteen (4×4) pixels in each tile. Comparing the results using features with (blue lines) and without (orange lines) scaling in Figure 3, we found that the scaling allowed for a reduction of the size of the training set and increased the spatial resolution of the classification.

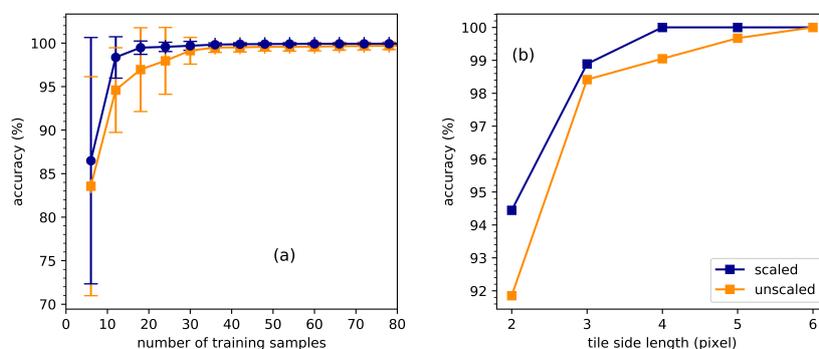


Figure 3. Dependence of the accuracy of the classification on: the number of samples in the training set (a); and the tile side length (b). The blue line denotes the accuracy with scaling, while the orange line denotes the accuracy without scaling. In (a), the tile side length is set to 4 and, in (b), the number of training samples is set to 216.

Finally, we applied the trained algorithm to distinguish the two particle types and microparticle-free plasma in another frame of the recorded video. The original image and the results after classification are shown in Figure 4a,b, respectively. As we compare these two panels, we see that the small particles (white area), big particles (gray area), and the background microparticle-free plasma (black area) are clearly distinguished. An interface can be drawn between the small and big particles (highlighted by the red line). Here, the red line is obtained by calculating each of the demarcation points of the horizontal pixels, while the green dotted line is obtained by connecting the innermost big particle cloud pixels. The discrepancy may be caused by the presence of a third type of particles with intermediate size in the experiment run.

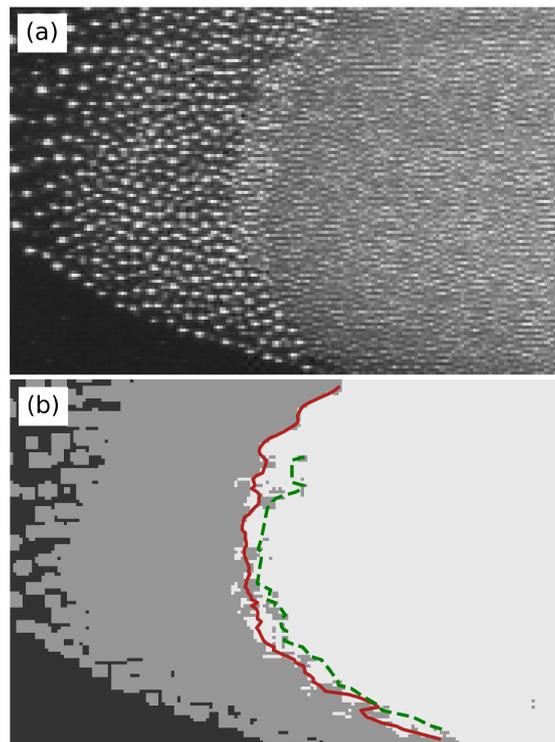


Figure 4. Original image selected from the experiment recording (a); and the results obtained by the SVC classifier (b). Area of small particles (white), big particles (gray), and the background microparticle-free plasma (black) are clearly distinguished. The red and green curves indicate the interface detected with two different methods (see text).

In summary, we applied the SVM method to achieve automatic identification of the interface in a binary complex plasma. The experiments were performed in a binary complex plasma under microgravity conditions on board the ISS, where the particle size cannot be directly deduced in the recorded images by the QV camera. The results show that this method can effectively distinguish small and big particles and the background microparticle-free plasma using the scaled mean and variance of the pixel brightness with low demand for the training samples.

Author Contributions: H.H. performed the analysis and wrote the original draft. M.S. designed the experiments and edited the manuscript. C.-R.D. conceived the idea and oversaw the project.

Funding: The research was funded by the National Natural Science Foundation of China (NSFC), Grant No. 11405030. The PK-3 Plus project was funded by the space agency of the Deutsches Zentrum für Luft- und Raumfahrt e.V. with funds from the Federal Ministry for Economy and Technology according to a resolution of the Deutscher Bundestag under grant number 50WM1203.

Acknowledgments: We thank the PK-3 Plus team for providing the experimental data and we also thank Christoph Räth for the useful comments.

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

References

- Fortov, V.; Ivlev, A.; Khrapak, S.; Khrapak, A.; Morfill, G. Complex (dusty) plasmas: Current status, open issues, perspectives. *Phys. Rep.* **2005**, *421*, 1–103. [[CrossRef](#)]
- Morfill, G.E.; Ivlev, A.V. Complex plasmas: An interdisciplinary research field. *Rev. Mod. Phys.* **2009**, *81*, 1353–1404. [[CrossRef](#)]
- Chaudhuri, M.; Khrapak, S.A.; Morfill, G.E. Complex plasma the plasma state of soft matter. *AIP Conf. Proc.* **2010**, *1308*, 207–214. [[CrossRef](#)]
- Chu, J.H.; Lin, I. Direct observation of Coulomb crystals and liquids in strongly coupled rf dusty plasmas. *Phys. Rev. Lett.* **1994**, *72*, 4009–4012. [[CrossRef](#)] [[PubMed](#)]
- Zhukhovitskii, D.I.; Fortov, V.E.; Molotkov, V.I.; Lipaev, A.M.; Naumkin, V.N.; Thomas, H.M.; Ivlev, A.V.; Schwabe, M.; Morfill, G.E. Nonviscous motion of a slow particle in a dust crystal under microgravity conditions. *Phys. Rev. E* **2012**, *86*, 016401. [[CrossRef](#)]
- Teng, L.W.; Chang, M.C.; Tseng, Y.P.; Lin, I. Wave-particle dynamics of wave breaking in the self-excited dust acoustic wave. *Phys. Rev. Lett.* **2009**, *103*, 245005. [[CrossRef](#)] [[PubMed](#)]
- Zhdanov, S.; Schwabe, M.; R ath, C.; Thomas, H.M.; Morfill, G.E. Wave turbulence observed in an auto-oscillating complex (dusty) plasma. *EPL (Europhys. Lett.)* **2015**, *110*, 35001. [[CrossRef](#)]
- Schwabe, M.; Zhdanov, S.; R ath, C. Instability onset and scaling laws of an auto-oscillating turbulent flow in a complex plasma. *Phys. Rev. E* **2017**, *95*, 041201. [[CrossRef](#)] [[PubMed](#)]
- Du, C.R.; Nosenko, V.; Thomas, H.M.; Morfill, G.E.; Ivlev, A.V. Slow dynamics in a quasi-two-dimensional binary complex plasma. *arXiv* **2016**, arXiv:1609.01456.
- Ivlev, A.V.; Zhdanov, S.K.; Thomas, H.M.; Morfill, G.E. Fluid phase separation in binary complex plasmas. *EPL (Europhys. Lett.)* **2009**, *85*, 45001. [[CrossRef](#)]
- Wysocki, A.; R ath, C.; Ivlev, A.V.; S utterlin, K.R.; Thomas, H.M.; Khrapak, S.; Zhdanov, S.; Fortov, V.E.; Lipaev, A.M.; Molotkov, V.I.; et al. Kinetics of fluid demixing in complex plasmas: Role of two-scale interactions. *Phys. Rev. Lett.* **2010**, *105*, 045001. [[CrossRef](#)]
- S utterlin, K.R.; Wysocki, A.; Ivlev, A.V.; R ath, C.; Thomas, H.M.; Rubin-Zuzic, M.; Goedheer, W.J.; Fortov, V.E.; Lipaev, A.M.; Molotkov, V.I.; et al. Dynamics of lane formation in driven binary complex plasmas. *Phys. Rev. Lett.* **2009**, *102*, 085003. [[CrossRef](#)] [[PubMed](#)]
- Killer, C.; Bockwoldt, T.; Sch utt, S.; Himpel, M.; Melzer, A.; Piel, A. Phase separation of binary charged particle systems with small size disparities using a dusty plasma. *Phys. Rev. Lett.* **2016**, *116*, 115002. [[CrossRef](#)] [[PubMed](#)]
- Yang, L.; Schwabe, M.; Zhdanov, S.; Thomas, H.M.; Lipaev, A.M.; Molotkov, V.I.; Fortov, V.E.; Zhang, J.; Du, C.R. Density waves at the interface of a binary complex plasma. *EPL (Europhys. Lett.)* **2017**, *117*, 25001. [[CrossRef](#)]
- Schwabe, M.; Zhdanov, S.; Hagl, T.; Huber, P.; Lipaev, A.M.; Molotkov, V.I.; Naumkin, V.N.; Rubin-Zuzic, M.; Vinogradov, P.V.; Zaehring, E.; et al. Observation of metallic sphere-complex plasma interactions in microgravity. *New J. Phys.* **2017**, *19*, 103019. [[CrossRef](#)]
- Shukla, P.K.; Mamun, A.A. Introduction to Dusty Plasma Physics. *Plasma Phys. Control. Fusion* **2002**, *44*, 395–395. [[CrossRef](#)]
- Schwabe, M.; Du, C.R.; Huber, P.; Lipaev, A.M.; Molotkov, V.I.; Naumkin, V.N.; Zhdanov, S.K.; Zhukhovitskii, D.I.; Fortov, V.E.; Thomas, H.M. Latest results on complex plasmas with the PK-3 plus laboratory on board the international space station. *Microgravity Sci. Technol.* **2018**, *30*, 581–589. [[CrossRef](#)]
- Mohr, D.P.; Knappek, C.A.; Huber, P.; Zaehring, E. Particle detection algorithms for complex plasmas. *AIP Conf. Proc.* **2018**, *1925*, 020010. [[CrossRef](#)]
- Rawat, W.; Wang, Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Comput.* **2017**, *29*, 2352–2449. [[CrossRef](#)]
- Gao, Y.; Ma, J.; Yuille, A.L. Semi-supervised sparse representation based classification for face recognition with insufficient labeled samples. *IEEE Trans. Image Process.* **2017**, *26*, 2545–2560. [[CrossRef](#)]
- Wang, M.; Deng, W. Deep face recognition: A survey. *arXiv* **2018**, arXiv:1804.06655.
- Wright, J.; Yang, A.Y.; Ganesh, A.; Sastry, S.S.; Ma, Y. Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intel.* **2009**, *31*, 210–227. [[CrossRef](#)]

23. Hull, J.J. A database for handwritten text recognition research. *IEEE Trans. Pattern Anal. Mach. Intel.* **1994**, *16*, 550–554. [[CrossRef](#)]
24. Liu, C.L.; Nakashima, K.; Sako, H.; Fujisawa, H. Handwritten digit recognition: benchmarking of state-of-the-art techniques. *Pattern Recognit.* **2003**, *36*, 2271–2285. [[CrossRef](#)]
25. Kotsiantis, S.B. Supervised machine learning: A review of classification techniques. In *Emerging Artificial Intelligence Applications in Computer Engineering*; IOS Press: Amsterdam, The Netherlands, 2007; pp. 3–24.
26. Murthy, S.K. Automatic construction of decision trees from data: A multi-disciplinary survey. *Data Min. Knowl. Discov.* **1998**, *2*, 345–389. [[CrossRef](#)]
27. Zhang, G.P. Neural networks for classification: a survey. *IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.)* **2000**, *30*, 451–462.
28. Cheng, J.; Greiner, R. Learning bayesian belief network classifiers: Algorithms and system. In *Advances in Artificial Intelligence*; Stroulia, E., Matwin, S., Eds.; Springer Berlin Heidelberg: Berlin/Heidelberg, Germany, 2001; pp. 141–151.
29. Cover, T.; Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [[CrossRef](#)]
30. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
31. Yang, Y.; Liu, X. A Re-examination of text categorization methods. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Berkeley, CA, USA, 15–19 August 1999*; ACM: New York, NY, USA, 1999; pp. 42–49. [[CrossRef](#)]
32. Decoste, D.; Schölkopf, B. Training invariant support vector machines. *Mach. Learn.* **2002**, *46*, 161–190. [[CrossRef](#)]
33. Gaonkar, B.; Davatzikos, C. Analytic estimation of statistical significance maps for support vector machine based multi-variate image analysis and classification. *NeuroImage* **2013**, *78*, 270–283. [[CrossRef](#)]
34. Hardin, D.; Tsamardinos, I.; Aliferis, C. A theoretical characterization of linear SVM-based feature selection. In *Proceedings of the Twenty-First International Conference on Machine Learning, ICML 2004, Banff, AB, Canada, 4–8 July 2004*; pp. 377–384.
35. Thomas, H.M.; Morfill, G.E.; Fortov, V.E.; Ivlev, A.V.; Molotkov, V.I.; Lipaev, A.M.; Hagl, T.; Rothmel, H.; Khrapak, S.A.; Suetterlin, R.K.; et al. Complex plasma laboratory PK-3 Plus on the international space station. *New J. Phys.* **2008**, *10*, 033036. [[CrossRef](#)]
36. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
37. Buitinck, L.; Louppe, G.; Blondel, M.; Pedregosa, F.; Mueller, A.; Grisel, O.; Niculae, V.; Prettenhofer, P.; Gramfort, A.; Grobler, J.; et al. API design for machine learning software: experiences from the scikit-learn project. In *Proceedings of the ECML PKDD Workshop: Languages for Data Mining and Machine Learning, Prauge, Czech Republic, 23–27 September 2013*; pp. 108–122.
38. Gidudu, A.; Hulley, G.; Marwala, T. Image classification using SVMs: One-against-one vs. one-against-all. *arXiv* **2007**, arXiv:0711.2914.

