Ice Detection on Aircraft Surface Using Machine Learning Approaches Based on Hyperspectral and Multispectral Images

drones

MDPI

Article

Ice Detection on Aircraft Surface Using Machine Learning Approaches Based on Hyperspectral and Multispectral Images

Maria Angela Musci [1,2], Luigi Mazzara [1,2,*] and Andrea Maria Lingua [1,2]

[1] DIATI, Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy; mariaangela.musci@polito.it (M.A.M.); andrea.lingua@polito.it (A.M.L.)

[2] PIC4SeR, Politecnico Interdepartmental Centre for Service Robotics, 10129 Torino, Italy

* Correspondence: luigi.mazzara@polito.it; Tel.: +39-3453183320

Received: 7 July 2020; Accepted: 14 August 2020; Published: 18 August 2020

Abstract: Aircraft ground de-icing operations play a critical role in flight safety. However, to handle the aircraft de-icing, a considerable quantity of de-icing fluids is commonly employed. Moreover, some pre-flight inspections are carried out with engines running; thus, a large amount of fuel is wasted, and CO_2 is emitted. This implies substantial economic and environmental impacts. In this context, the European project (reference call: MANUNET III 2018, project code: MNET18/ICT-3438) called SEI (Spectral Evidence of Ice) aims to provide innovative tools to identify the ice on aircraft and improve the efficiency of the de-icing process. The project includes the design of a low-cost UAV (uncrewed aerial vehicle) platform and the development of a quasi-real-time ice detection methodology to ensure a faster and semi-automatic activity with a reduction of applied operating time and de-icing fluids. The purpose of this work, developed within the activities of the project, is defining and testing the most suitable sensor using a radiometric approach and machine learning algorithms. The adopted methodology consists of classifying ice through spectral imagery collected by two different sensors: multispectral and hyperspectral camera. Since the UAV prototype is under construction, the experimental analysis was performed with a simulation dataset acquired on the ground. The comparison among the two approaches, and their related algorithms (random forest and support vector machine) for image processing, was presented: practical results show that it is possible to identify the ice in both cases. Nonetheless, the hyperspectral camera guarantees a more reliable solution reaching a higher level of accuracy of classified iced surfaces.

Keywords: hyperspectral images; multispectral data; machine learning; ice detection

1. Introduction

Human safety is one of the main concerns at airports and aircraft icing represents a significant hazard in aviation [1]. The ice formation leads the aircraft's balance in a loss of control, and de-icing and anti-icing are necessary treatments for flight safety during the winter [2,3]. However, de-icing operations require the employment of chemicals such as ethylene glycol- (EG-) or propylene glycol- (PG-) that can cause damage to the environment, in particular for the nearby surface and groundwater [3].

Ice accumulation can occur due to the supercooled droplets colliding with a hard surface forming an ice film [4] with an air temperature between 0 and −20 °C [5]. As reported by the FAA (Federal Aviation Administration), structural or in-flight ice and ground ice can be identified [6,7]. The former occurs when the aircraft is flying through visible water such as rain or cloud droplets. The latter, instead, may accumulate on parked aircraft due to precipitation and atmospheric conditions.

According to the temperature, liquid water content, speed of the formation process, aircraft surface temperature and shape, particle concentration and size, it is possible to distinguish three different types of structural ice [8] (Figure 1):

- Rime ice that is a milky-white deposit of the ice, and it is the result of rapidly freezing of small droplets at low temperature after impacting with the aircraft surfaces. It grows at low temperature (<15 °C) with low water liquid content. The rime density is lower than 0.2–0.3 g cm^{-3}, and it is composed mainly of discrete ice granules [9].
- Clear or glaze ice, glassy transparent ice, caused by large droplets that run back on the aircraft surface with slow freezing, has a density higher than 0.8 or 0.9 g cm^{-3} [10].
- Mixed-ice that has mixed features of the previous ones, because it forms when both rime and clear ice accumulate on the aircraft.

Figure 1. Ice types: rime ice, clear ice, and mixed-ice. (Photo credit: NASA, adapted from [8]).

Frost, snow (or slush), fog, drizzle, rain (and their freezing states), and ice pellets can be considered the foremost examples of ground icing.

The sublimation of deposited water vapor on the aircraft can form frost when the temperature is at or below freezing. A fog formed of supercooled water droplets which freeze upon the impact with the aircraft surface, also known as freezing fog, produces a coating of rime/clear ice.

Rain and drizzle, uniform precipitations of liquid water particles, can be distinguished by drop diameters and proximity. The rain is characterized by drops with a diameter more than 0.5 mm very separated, instead of drizzle, that has close drops with diameters less than 0.5 mm. These two kinds of precipitations, in the freezing state, can create ice deposits with a transparent appearance. Snow or slush are precipitations of ice crystals. The slush is formed by water-saturated snow.

As gathered from the Manual of Aircraft Ground De-icing/Anti-icing Operations [2], the difference between in flight and on ground icing is not referred mainly on the characteristics of the ice but the impacts on the flight and the de-icing procedures. As it is possible to notice from the definitions of freezing fog or frost and freezing rain, clear and rime ice can also occur on the ground.

Due to the physical characteristics of these types of ice, their identification is currently based on visual (e.g., for rime ice, snow) and tactile (e.g., for clear ice, frost) inspections carried out by trained and qualified ground crew or flight crew [11].

The cleaning process moreover involves the use of a considerable amount of aircraft de-icing fluids (ADFs), because targeted operations are not achievable. The contamination check shall cover all surfaces that have aerodynamic-, control-, sensing-, movement- or measuring-function such as wings, tail surfaces, engine, fuselage, antennas, and sensors. This investigation requires enough visibility of these parts. A verification of cleaned surfaces shall always be made after the de-icing/anti-icing, and this inspection can be either visual or tactile. The whole procedure is time-consuming and demanding, especially since it is crucial to maintain the flight's schedule. A time-effective strategy for ice detection is required to limit ADFs use and improve the management of the crew's operations.

In this context, UAV (uncrewed aerial vehicle) [12] imagery combined with machine learning algorithms has shown excellent potential for rapid, remote, cost-effective detection tasks. This approach allows ice identification from multiple views with an automatic check-up operation.

The SEI (Spectral Evidence of Ice) project [13,14] proposes to provide an integrated solution that can handle the automatic pre-deicing inspection, ice detection, and cleaning verification procedure. The expert crew manages the request of de-icing and sends the UAV to the parking area (or hangar) of a specific aircraft that needs the procedure. The UAV can autonomously recognize the location of the aircraft and starts the inspection. Indeed, the multi-sensor UAV platform, equipped with a hyperspectral or multispectral camera, has been designed to monitor and inspect aircraft in the specific de-icing area of the airport. The main task of the drone is the identification of the location and the extension of the ice-contaminated area. For this purpose, an automatic methodology for geometric and radiometric detection of the ice has been determined. The development of computer-oriented methods for ice detection is still challenging due to the physical characteristics of the ice, variable atmospheric condition, and lack of autonomous technology in this application field.

Several devices have been developed for ice characterization for on-ground and in-flight inspection [15,16]. Some examples are based on ultrasonic, magnetostrictive, and electromagnetic sensors [17,18]. Some researchers such as Gong et al. [19] have discussed the use of the mid-infrared sensor for ice detection. In this field, the spectral imagery, not only in the mir-infrared but also in all electromagnetic range, is an emerging technology because of its high spectral and spatial resolution [20]. Our study would fulfil the gap and present the potential of hyperspectral or multispectral imaging technique in the ice detection on aircraft.

Regardless of devices for data acquisition, machine learning approaches, such as random forest (RF) [14] and support vector machine (SVM) [15], have been utilized for material detection and their characterization [21]. These algorithms perform well in reducing the complexity of the classification task associated with spectral data because they can handle the high dimensionality input space and noisy dataset [22,23].

This work, within the activities of the SEI project, tests the feasibility to use spectral sensors, such as hyperspectral and multispectral cameras, and random forest and support vector machine, as machine learning algorithms.

Firstly, the purpose is the selection of the most suitable sensor to mount on a UAV prototype that has to respond to cost requirements. For this reason, a multispectral camera, as a low-cost sensor, was examined to reduce the system production cost. At the same time, the paper addresses the definition of the time-effective automatic methodology for ice detection using the machine learning approach. As known, the hyperspectral camera has a spectral resolution of more than 100 bands instead of the multispectral camera that has a few bands (most of the cases from three to 15). A dimensionality reduction process has been applied to accurately compare the performance of the two algorithms on images with sharply different spectral resolution.

Since the UAV prototype is under construction, the experimental analysis was performed with a simulation dataset acquired on the ground. However, the methodology can be easily transferred to a UAV application.

2. Materials and Methods

This section describes the two sensors (Section 2.1), the methodology (Section 2.2), and the algorithms and accuracy assessment (Sections 2.3–2.5).

2.1. Sensor Description

The data acquisition was performed by a hyperspectral camera (Senop Rikola) and a multispectral camera (MAPIR Survey 3N). Senop Rikola hyperspectral camera is a snapshot camera based on a Fabry–Perot Interferometer [24,25]. It includes two not aligned sensors: one sensor acquires near-infrared bands (659.2–802.6 nm) and the second captures visible bands (from 502.8 to 635.1 nm). The MAPIR survey 3N is a multispectral camera, and it records RGN (red, green, and near-infrared bands) images as red (660 nm), green (550 nm), and near-infrared (850 nm) bands [26]. An RGB camera with specifications comparable with the MAPIR (same spatial resolution, optics, and pixel size) was

used to include the blue band (475 nm). The reason for the introduction of an additional band is explained in the Methodology (Section 2.2).

The Senop and the MAPIR are lightweight UAV sensors, and they were selected because they have a similar spectral range from 500 to 950 nm. Table 1 summarizes the specifications of the two sensors.

Table 1. Sensor specifications: the Senop Rikola hyperspectral camera and the MAPIR Survey3N multispectral camera.

Camera	Senop Rikola	MAPIR Survey3N
Lens optics	H 36.5°, V 36.5°	H 41° HFOV (47 mm)
Spectral range	500–900 nm	550–850 nm
Spectral channels	380	3
Spectral resolution	1 nm	-
Shutter type	Global	Rolling
Focal length	9 mm	8.25 mm
Image resolution	1010 × 1010 pixels	4000 × 3000 pixels
Pixel size	5.5 µm	1.55 µm
Weight	720 g	75.4 g
Dimensions	172.7 × 89 × 77 mm	59 × 41.5 × 36 mm
Cost	≈60,000€	≈700€

2.2. Methodology

The overall methodology is shown in the schema below (Figure 2).

For data collection, ice samples were generated in the laboratory using molds and a real section of the aircraft wing. Since the idea was to have ice samples similar to rime (white ice) and clear ice (transparent ice), two types of ice were created. Snow or other varieties of ice cited above are not considered in this analysis because its production in our laboratory was not possible. The details of sample production and data acquisition will be explained in Section 3.

The hyperspectral and multispectral images were radiometrically corrected using the empirical line method (ELM) and the reference panel.

After that, the dimensionality reduction with principal component analysis (PCA) was executed on the hyperspectral data. This step allows defining two new datasets, the one composed by principal component hypercubes and the ones with a reduced number of bands. The original hyperspectral dataset and both the new ones also with multispectral images were classified to evaluate the performance of the detection and the computational time. Moreover, for the multispectral case, further analysis was made using the RGBN (red, green, blue, near-infrared) images, to understand the improvement of the additional band on the classification.

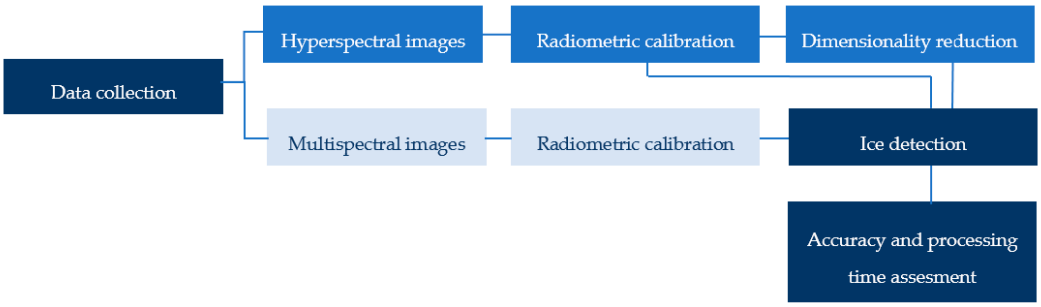

Figure 2. Methodology workflow.

2.3. Dimensionality Reduction of Hyperspectral Images: Feature Extraction and Feature Selection

The high dimensionality of hyperspectral images is a crucial problem in real-time application because it takes time both in the acquisition and in ice detection steps. Moreover, it can produce the so-called Hughes phenomenon [27]. For addressing this issue, the most popular methods for

dimensionality reduction are feature extraction and feature selection. Feature extraction refers to a linear or nonlinear transformation procedure that reduces the data redundancy in the spatial and spectral domain. Feature selection refers to a process to define a subset of the original features without a transformation [28–30]. PCA is widely used as a feature extraction method, but it can also be used for feature selection.

The PCA dimensionality reduction is based on the estimation of the eigenvalues of the covariance matrix [31–33]. For each pair of bands, the covariance is calculated as (1):

$$\sigma_{i,j} = \frac{1}{N-1}\sum_{p=1}^{N}(DN_{p,i} - \mu_i)\left(DN_{p,j} - \mu_j\right) \quad (1)$$

where $DN_{p,j}$ and $DN_{p,j}$ are digital numbers of a pixel p in the bands i and j, respectively, and the μ_i and μ_j are the averages of the DN for bands i and j. Then the covariance matrix is defined as (2):

$$C_{b,b} = \begin{pmatrix} \sigma_{1,1} & \dots & \sigma_{1,j} \\ \dots & \dots & \dots \\ \sigma_{i,1} & \dots & \sigma_{i,j} \end{pmatrix} \quad (2)$$

The roots of the characteristic equation provide the eigenvalues λ (3):

$$\det(C - \lambda I) = 0 \quad (3)$$

where C refers to the covariance matrix (2), and I is the diagonal identity matrix.

The eigenvalues indicate the quantity of original information that they compress. The variance percentage for each principal component is calculated as the ratio of each eigenvalue and the sum of all of them. Those components which contain minimum variance and, thus, the minimum number of information can be discarded. The matrix form of the principal components can be expressed as (4):

$$Y_i = \begin{pmatrix} w_{1,1} & \dots & w_{1,j} \\ \dots & \dots & \dots \\ w_{i,1} & \dots & w_{i,j} \end{pmatrix} X_j \quad (4)$$

where Y is the vector of the principal components (PC), W the transformation matrix, and X the vector of the original data, the coefficients $w_{i,j}$ are the eigenvectors, and they link the PC with the real variable providing information on their relationship. The eigenvectors can be calculated for each λ_k as (5):

$$(C - \lambda_k I)w_k = 0 \quad (5)$$

where C and I can be defined as the (3), while λ_k is the k eigenvalues and w_k is the k eigenvectors.

There are three practical criteria to select the most representative PCs [34]:

- Cumulative percentage of total variation: the representative PCs must contribute to the cumulative percentage of total variation to achieve 80% or 90%.
- Kaiser's criterion: the selected PCs must have a variance that exceeds 1, which means that if all the variables are independent, the PCs have variance equal to 1 in the correlation matrix.
- Screen graph: it is the plot of eigenvalues λ_k related to the order number of that eigenvalue k. The PC number can be selected looking at the 'elbow' in the graph, that indicates the components to retain. The number of PCs to use will be given by the smallest k characterized by the trend of λ_k. The trend of the function must be sharply decreasing on the left of 'elbow' point, while on the right, it has to be constant or weakly decreasing [35].

Once the PCs have been chosen, the interpretation of them is based on eigenvectors, derived from the (5). The meaning of PCs can be determined looking at the coefficient ($w_{i,j}$) of variables X_j. The greater $w_{i,j}$ is, the higher the correlation, and X_j is the most important for the PC [36].

2.4. Machine Learning: Random Forests and Support Vector Machine

Random forest (RF) algorithm builds multi-decision trees (forest) that operate as an ensemble trained with a bagging mechanism [37,38]. The bagging mechanism samples N random bootstraps of the training set with replacement. The number of trees characterized the forest, and the higher the number of trees, the more accurate the classification [39]. Moreover, the following parameters can affect the performance of the RF classifier: the tree depth, that is the number of splits for each tree, the split criteria, that handle the split at each node (such as GINI index), and the minimum split [40,41].

Support vector machine (SVM) is a binary algorithm and constructs an optimal hyperplane or a set of hyperplanes, that can be employed for the classification task [42]. The best hyperplane can separate data points of different classes, and it is usually the plane that has the most significant margin between the two classes [40]. SVM can be extended to the multiclass problem through two different approaches: the one-against-all or the one-against-one. In the one-against-all approach, a set of N binary classifiers is applied to the N-class problem. The second approach, one-against-one, carries out a series of binary classifiers to each pair of classes. The training sample size has a high impact on the performance of the SVM, as defined in Myburgh G. et al. [43].

2.5. Evaluation Metrics

Either for the random forest and the support vector machine, the accuracy assessment for the performance evaluation can be achieved with different parameters based on the error matrix. According to the literature, the selected parameters are the following [44,45]:

- Overall accuracy (OA) that is the ratio of the total number of correctly classified pixels with respect to the total number of reference pixels;
- User's accuracy (UA) that is the ratio of pixels correctly classified in given class with respect to all pixels classified in the specific category.

Moreover, in this specific real-time application, the computational time for the classification part was assessed. The processing time of the training procedure was not taken into account because the final goal was to use transfer learning.

3. Ice Detection: An Experimental Analysis

For the experimental analysis, it was not possible to collect the real types of ice; thus, different kinds of ice were generated in the laboratory as much as similar to the case study. Two types of ice were produced: the first one similar to the rime ice with milky-white color and the second one to the glassy clear ice.

The former was created using the water vapor condensed into the freezer at a temperature of −15 °C and its thickness reaches values between 2 and 6 cm. The latter was generated by freezing tap water within plastic molds at a temperature of −15 °C. Different plastic molds were used for producing different blocks of ice that contained from 5 to 20 mL of water with a thickness of approximately around 3 cm. The ice blocks were located on a section of an aircraft wing to simulate the typical conditions in which the ice is present.

The aluminum panel used in the tests was a section of a Socata MS.894 Rally Minerva with a dimension of 400 × 400 × 2 mm. Before icing, the panel was stored in a freezer so the icing would start with low surface temperature. Figure 3 shows the configuration of the samples.

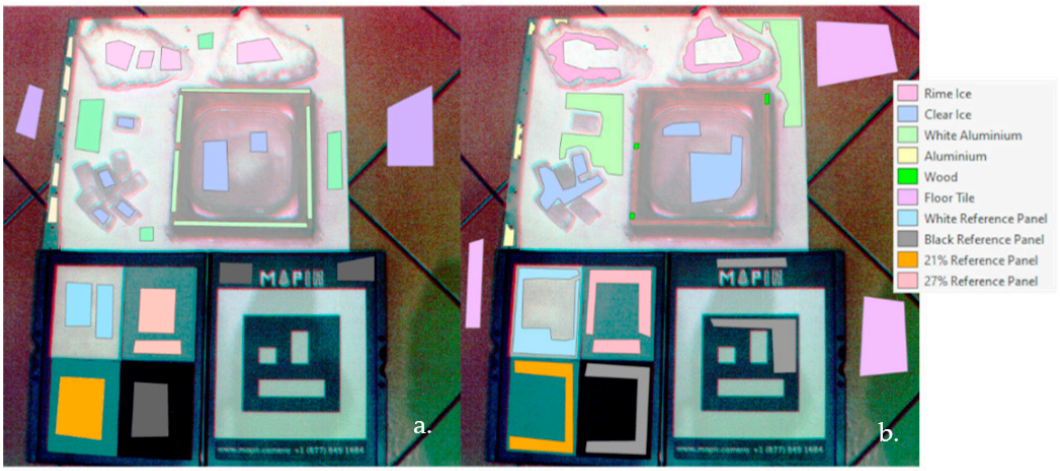

Figure 3. Example of annotated image with reference data. Training set sample (**a**) and validation set sample (**b**).

3.1. From Data Collection to Sample Annotation

The dataset was built by collecting a ground measurement. The acquisitions were performed at Photogrammetry, Geomatics and GIS Laboratory of DIATI (Department of Environment, Land and Infrastructure Engineering) at Politecnico di Torino (Italy) [46]. During this campaign, 10 images for the hyperspectral and eight for the multispectral sensor with different illumination conditions were collected (18 images in overall). The various illumination conditions were generated using a different number of lamps and a combination of lamps and natural light to simulate the real scenario in which the drones will be used in the parking area or the hangar. The term "Test" refers to each image with different environmental conditions in this paper. All data were recorded, maintaining stable positions and varying rotations of the camera slightly.

The hyperspectral camera was used in manual mode connected to the computer through a USB cable. The selected image resolution was 1010 × 1010 pixels. The images were composed by 100 bands covering the spectral range from 502 to 906 nm, with a wavelength step of 4 nm and a Full with Half Maximum resolution (FWHM, where Wide means low gap index). The integration time was set at 450 ms based on the environmental illumination condition. The sequence of the bands was automatically generated using the Rikola Hyperspectral Imager software v2.0. These parameters were chosen to cover the whole spectral range. The whole electromagnetic spectrum was also covered to identify the most characteristic bands and features of studied materials. For the MAPIR instead, the camera's sensitivity was set to ISO- 800, and the exposure time was fixed to 1/15 s.

The two datasets of images were radiometrically calibrated using the Empirical Line Calibration tool of ENVI 4.7 [47]. Then, the images of each sensor were manually annotated. In both cases, the same 10 classes were considered: rime ice, clear ice, white aluminum, aluminum, floor tile, wood and reference panel (white, black, grey 21%, and grey 27%). The representative classes were only the rime ice, the clear ice, and the white aluminum (Figure 3). These classes were chosen according to the materials that it was possible to distinguish in the real case at aircraft scale. The selected materials were related to the object (in our case, the aircraft) and the ice. Other materials were included in the background in different classes to improve the performance of the classification. The option of a single class for background materials would alter the accuracy of outcomes. The number of samples per class for each dataset are reported in Table 2.

Table 2. Hyperspectral and multispectral reference samples per class.

Class	Hyperspectral Samples (Pixels)	Multispectral Samples (Pixels)
Rime ice	10,156	101,797
Clear ice	10,457	104,819
White aluminum	11,632	116,588
Aluminum	3580	35,912
Wood	10,896	93,624
Floor tile	19,812	198,572
White reference panel	9545	95,672
Black reference panel	16,024	160,617
21% grey reference panel	12,980	130,089
27% grey reference panel	10,670	106,949
Average pixels per class	10,523	104,058

The training and test samples were collected based on visual interpretation. ArcGIS Pro 2.5.0 toolbox was used to create polygons as reference data for each class.

3.2. Dimensionality Reduction of Hyperspectral Data: Results

To reduce the hyperspectral data dimensionality, PCA was carried out using the "Principal components tools" of ArcGIS Pro 2.5.0 [48]. As described in Section 2.3, it was possible to adopt the PCA as feature extraction and band selection method.

As the first step, the feature extraction was performed to define the principal components. In the second step, the selected PCs were used for significative band selection. Both feature extraction and band selection methods were applied for understanding the best solution for ice detection.

Therefore, for the feature extraction, the eigenvalue and cumulative variance were obtained to identify the number of principal components (PCs), which means the new dimensionality. The outcomes of the first image only were reported as an example because the selection process and conclusions were the same for the other pictures. Table 3 shows the percentages of the primary five components of the sample image. As can be seen from Table 3, three PCs reach 90.31% of the total variance in original data for the first criterion and pass the 1% for Kaiser's rule (Section 2.3). As a consequence, the dimensionality of the new representation is three, and the rest of the component can be discarded. Moreover, also the screen plot in Figure 4 illustrates that it is possible to identify three as the number of PCs (third rule described in Section 2.3).

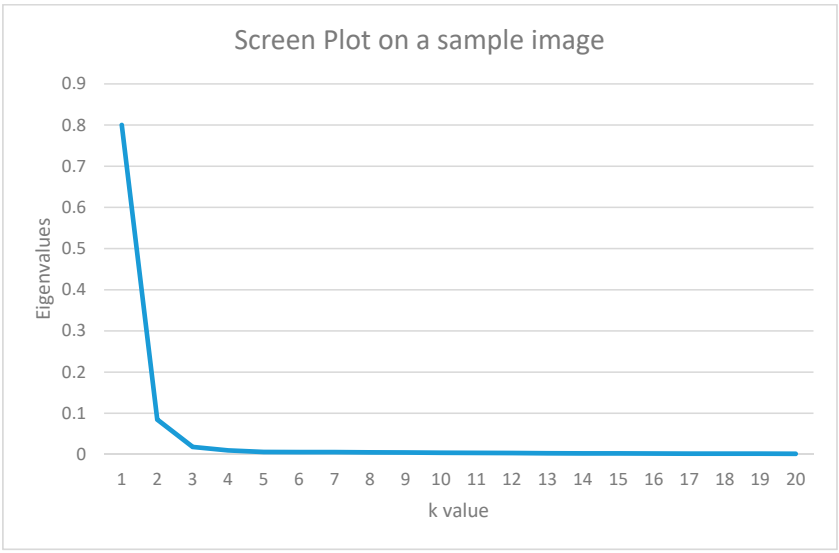

Figure 4. Screen graph on a sample image. Zoom of the 'elbow'.

Table 3. Principal component analysis (PCA): example of eigenvalue and cumulative variance in percentage on a single sample image.

Component	Percent %	Cumulative%
1	80.0010	80.0010
2	8.4818	88.4828
3	1.8341	90.3169
4	0.9935	91.3104
5	0.5888	91.8992

After the identification of the PC number, the first three principal components were used to select a reduced number of original bands for the classification task. The band selection process was carried out using the eigenvectors for each PC. The higher the absolute value of the band eigenvector, the higher the importance of that band for the specific principal component. According to this criterium, considering that the number of significative bands is strictly related to the application, a threshold of eigenvector values, defined for each component, allows to identify the significative bands. The plot of eigenvector values reports the correlation between spikes of the function and the representative bands. Figures 5–7 represent the eigenvector values with respect to the band number for the three selected principal components in four representative images (Test1, Test2, Test6, Test10) among 10 hyperspectral images. Different illuminations and the change of the state of ice characterize these four tests, and this comparison was made to check the recurrence of the most significant bands, that can be selected. The presence of spikes in the eigenvectors function allows to recognize the bands, for all the images for the three selected PCs (Table 4):

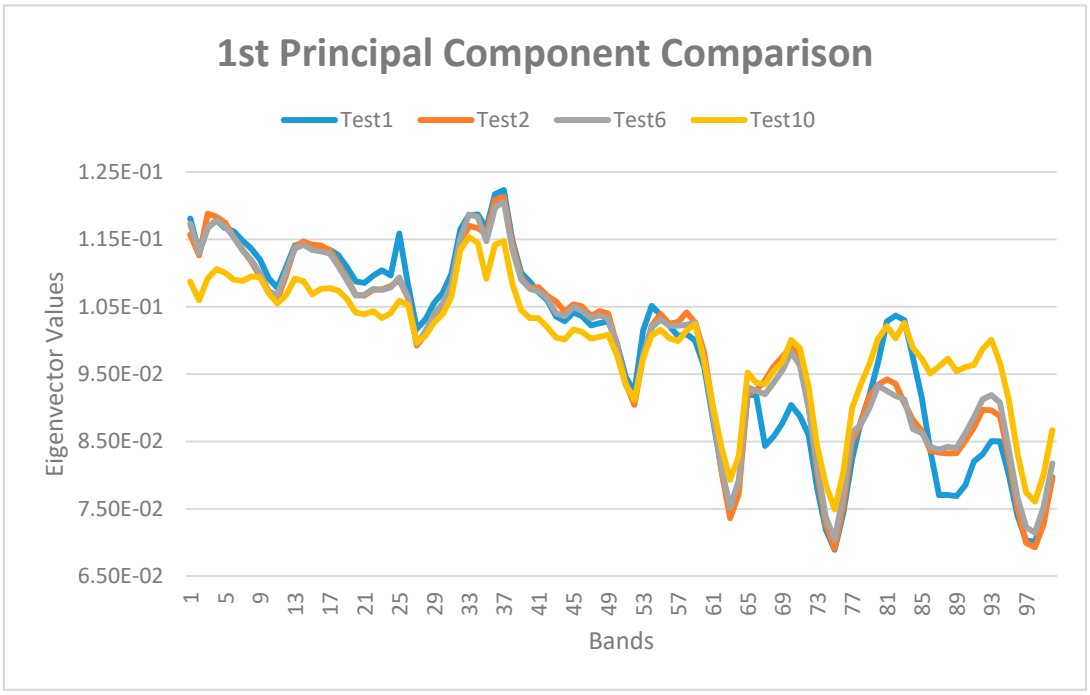

Figure 5. First principal component eigenvalues plot.

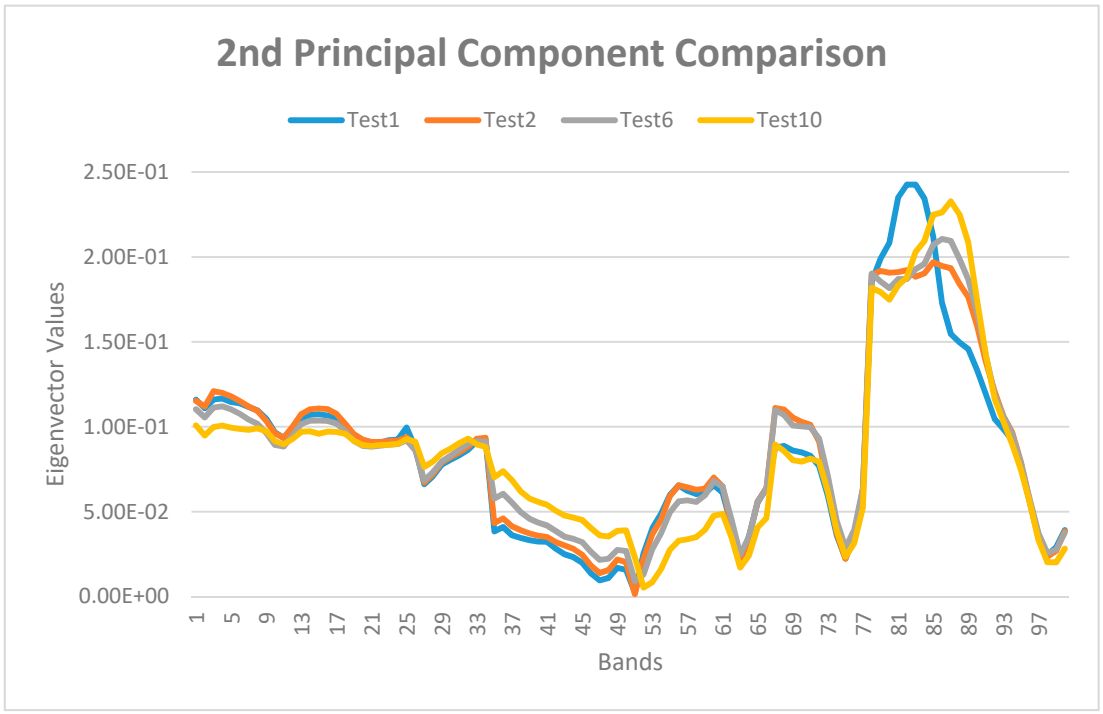

Figure 6. Second principal component eigenvalues plot.

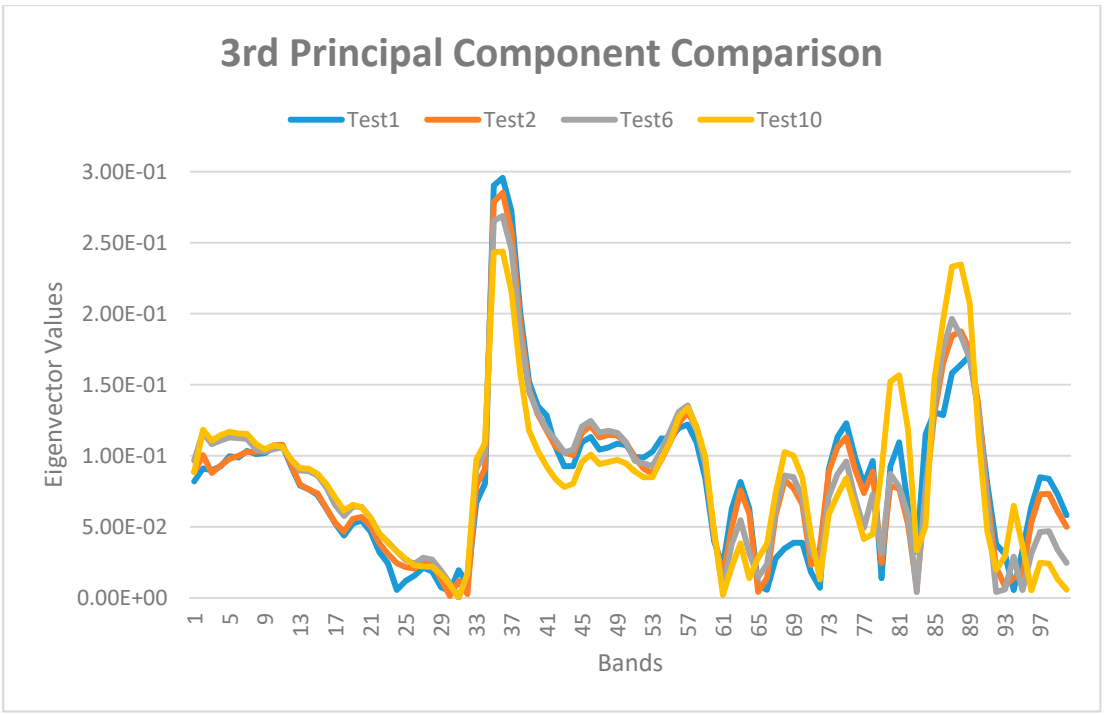

Figure 7. Third principal component eigenvalues plot.

As it is possible to notice in Table 4, there are recurring bands in each test. Taking into account all the identified bands, a new hypercube with 27 bands, that are 1 (506.31 nm), 3–7 (from 514.48 to 530.11 nm), 14 (558.28 nm), 25 (602.47), 32–38 (from 630.2 to 654.19 nm), 78–89 (from 817.58 to 861.65 nm) can be generated. However, considering only the popular bands in each principal component of all images, the number of significative bands can be further reduced to 10. The significative bands, in this latter case, are 4–5 (from 518.12 to 522.48 nm), 33–37 (from 634.36 to 650.38 nm), 83–85 (from 837.98 to 846.21 nm).

Table 4. Band selection for the three principal components (PC) in the representative tests (Test1, Test2, Test6, Test10).

Principal Component	Test1	Test2	Test6	Test10
1	1, 3–7, 25, 32–38	1, 3–6, 14, 33–37	1, 3–6, 32–37	4, 5, 32–34, 36, 37
2	80–85	80–85	78–89	83–89
3	35–37	35–37	35–37	35–37, 86–89

Three new datasets came from the dimensionality reduction process: 10 new images composed by the three PCs, 10 new hypercubes with 10 bands and 10 hypercubes with 27 bands. The first set of modified hypercubes was created through the toolbox "Principal Component Analysis" of ArcGIS Pro 2.5.0 selecting three as a maximum number of principal components. The two remaining datasets with reduced hypercubes were generated using a customized routine of Matlab for hyperspectral data decomposition and the "Composite Bands" tool of ArcGIS Pro 2.5.0 [49] for the selected band composition.

3.3. Hyperparameter Tuning for Random Forest and Support Vector Machine

The hyperparameter tuning process plays a crucial role in improving the accuracy of RF and SVM algorithms.

Before starting with the hyperparameters adjustment, data were split in 80% for training and 20% for testing. The tuning of hyperparameters was made on the training set for defining a model. Accuracy assessment was carried out for either the training and testing set to verify the performances of the model in the classification task. The validation curve allows to visualize the values of the model hyperparameters, and it shows different values of the single hyperparameter related to the accuracy trend.

The optimized hyperparameters were chosen according to two criteria. The first one is the minimum difference between the overall accuracy of training and validation models, and the second one is the best user's accuracy only for the validation. It is necessary to notice that for the evaluation of the accuracy the random choice of the samples has to be taken into account; thus, tolerance has to be considered. Accuracy analysis of training and validation are presented for both the algorithms.

Moreover, the accuracy assessment of the rime ice, clear ice, and white aluminum is under the attention among the other classes. These three classes are distinctive in the real de-icing application. The clear ice, as explained in the Introduction (Section 1), is critical to identify by visual inspection. Thus, it has a relevant weight in this analysis.

The tuning was implemented either for RF or SVM on a single image (Test_1) of both datasets (hyperspectral and multispectral images) and the "Segmentation and Classification tools" of ArcGIS Pro 2.5.0 [50]. The tests were made on a window workstation (Windows 10) with an Intel® Core™ I7-6500U CPU at 2.50 Ghz, GPU AMD Radeon™ R7 M360 (Iceland) (six compute units at 980 MHz, 2048 MB) and 16 GB of RAM.

Since SEI project application requires a near real-time approach (Section 1), in this section, the computational time was evaluated because one of the aims of the optimization is the definition of the trade-off between accuracy and processing time.

As described in Section 2.4, the hyperparameters for each classifier have to be tuned. They are the same either for hyperspectral and multispectral. The optimization was executed in manual mode. For the RF algorithm, the maximum number of samples for each class was fixed; we tuned two hyperparameters: the tree depth and the number of trees. For what concerns the SVM instead, only the maximum number of samples per class was tuned.

3.3.1. Hyperparameters Tuning for the Hyperspectral Dataset

Starting for the RF, the sample size was set to 2000 for each class for tuning the tree depth and the number of the trees. The tree depth optimization was done varying its value from 5 to 30. Instead,

the number of trees was fixed to 50. As reported in Table 5, the difference between the overall accuracy (OA) is comparable in all training and validation configuration. The case with a depth equal to 5 was kept out because it was not reasonable with a low number of trees. Therefore, looking just to the validation results, the tree depth equal to 30 produces the best value of OA and clear ice accuracy. For these reasons, the selected tree depth was equal to 30.

Table 5. Training accuracy (on the left) and validation accuracy (on the right) for random forest (RF) tree depth optimization. R_i stands for rime ice, C_i for clear ice, and W_a for white aluminum. In yellow, the selected optimized hyperparameter.

D_Trees	Training Accuracy (%)				Validation Accuracy (%)			
	R_i	C_i	W_a	Overall	R_i	C_i	W_a	Overall
5	73.6842	91.0615	74.7475	92.5425	42.4802	92.9577	64.8855	85.5371
10	80.1887	98.8636	92.9412	97.0971	44.4992	94.9346	69.5067	87.7175
15	78.8018	98.8506	93.4524	96.8969	43.128	95.0083	68.9602	87.0574
20	84.6535	93.956	93.8547	97.2472	41.2747	94.108	68.7908	86.6773
25	79.3427	97.1429	93.4911	96.8468	42.2256	92.5159	68.932	86.5173
30	81.6038	99.422	95.3757	97.2472	43.6863	95.7816	70.8692	87.83

Instead, for the number of tree selection, the tree depth value was fixed to 30, according to Table 5. The number of trees was varied from 5 to 50 (Table 6). Following the same reasoning defined for the selection of tree depth, the cases with a lower number of trees were excluded. Indeed, the differences among the OA is comparable in the other configurations. Concerning Table 6, the cases 15, 30, and 50 were characterized by a similar OA value that is also the highest one (87.9% on average). However, the accuracy of the C_i class leads with a gap in the case 50. As a consequence, the number of trees equal to 50 was the optimized value.

Table 6. Training accuracy (on the left) and validation accuracy (on the right) optimization of RF number of trees. R_i stands for rime ice, C_i for clear ice, and W_a for white aluminum. In yellow the selected optimized hyperparameter is shown.

N_Trees	Training Accuracy (%)				Validation Accuracy (%)			
	R_i	C_i	W_a	Overall	R_i	C_i	W_a	Overall
5	85,3403	96.5909	91.0053	97.1471	39.0855	93.9542	66.8908	86.0172
10	83.3333	99.422	95.4802	97.5475	41.1854	93.9344	67.7368	86.6573
15	**82.381**	**99.422**	**94.2529**	**97.3974**	**45.4699**	**95.1604**	**71.3681**	**87.9587**
20	80.7512	98.3051	95.2663	97.2472	43.2602	94.2122	71.205	87.2575
25	85.8586	97.2067	95.0549	97.6977	42.8571	93.4959	72.0339	87.3253
30	**82.6923**	**97.7011**	**93.2203**	**97.1972**	**44.0476**	**94.3548**	**71.1538**	**87.8743**
35	86.8687	99.4253	95.6522	97.998	44.081	93.75	70.892	87.3375
40	83.0846	96.4912	90.3226	96.9469	42.1217	94.2997	68.7988	86.5973
45	81.6425	98.8506	92.1348	97.0971	42.6404	95.098	69.9681	87.4375
50	**81.6038**	**99.422**	**95.3757**	**97.2472**	**43.6863**	**95.7816**	**70.8692**	**87.83**

In Figure 8, the comparison between training and validation overall trends among all the considered depths and numbers of trees can be appreciated. The validation curves confirm the previous observations and the criteria used for the optimized hyperparameters selection.

Table 7 presents the RF processing time for the training in the analyzed configurations considering the tuning of both the hyperparameters. In the case of D_Trees, the processing times are not excessively influenced by the increase of the depth number. While in the case of N_trees, the higher the number of trees, the higher the computational time. However, the computational time is stable after 15 N_tree because the selected sample size does not affect the number of trees. It is possible, thus, to choose 50 as N_trees value.

Figure 8. RF validation curve for maximum depth (on the **left**) and maximum number of trees (on the **right**).

Table 7. Processing time for the training in RF using different values of tree depth (on the left) and the number of trees (on the right). In yellow the time for training the model with the selected optimized hyperparameter is shown.

D_Trees	Processing Time	N_Trees	Processing Time
5	34″	5	14″
10	31″	10	26″
15	26″	15	32″
20	26″	20	31″
25	27″	25	32″
30	33″	30	29″
-	-	35	33″
-	-	40	30″
-	-	45	33″
-	-	50	33″

For what concerns the SVM instead, the maximum number of samples per class ranged between 100 and 5000 samples. The differences between OA training and validation remains 8% on average in all configurations (Table 8). The highest value of the OA (validation) occurs in the case of 5000 samples, but this configuration was excluded because the related computational time is too long (3h4′25″) (Table 9). Thus, the cases of 100, 750, and 1000 were taken into account. The configuration 100 was discarded because the sample size was small, and the random choice of the sample hugely affects the overall accuracy in the images.

Table 8. Training accuracy (on the left) and validation accuracy (on the right) optimization of support vector machine (SVM) number of the sample. R_i stands for rime ice, C_i for clear ice, and W_a for white aluminum. In yellow the selected optimized hyperparameter is shown.

N_Samples	Training Accuracy (%)				Validation Accuracy (%)			
	R_i	C_i	W_a	Overall	R_i	C_i	W_a	Overall
100	77.6744	97.7401	92.1687	96.3964	56.6879	99.1667	72.4719	91.4671
250	86.8687	97.7401	96.7568	97.9980	46.8354	96.4567	74.5247	89.7206
500	86.6667	1.0000	94.6237	97.9479	44.7059	95.2756	73.1707	88.6228
750	93.5135	99.4382	98.4615	99.0490	51.8644	**93.9623**	90.9091	90.9182
1000	91.8919	98.895	97.3822	98.7487	52.8571	**92.6740**	89.0000	91.0679
2000	91.5344	1.0000	97.8836	98.9990	49.6815	93.8697	89.4444	89.8204
5000	97.1910	1.0000	98.4925	99.5996	52.9825	92.8058	91.1917	91.3673

Table 9. Processing time for the training in SVM using different values of sample size. In yellow the time for training the model with the selected optimized hyperparameter is shown.

N_Samples	Processing Time
100	2′8″
250	12′26″
500	24′51″
750	47′21″
1000	1h18′2″
2000	1h33′36″
5000	3h4′25″

For the remaining cases, since OA is comparable, thus the hyperparameter selection was based on the accuracy of the clear ice. In the validation, the C_i accuracy is 94.0% in the 750 sample case, instead of 92.7% in the 1000 sample case. According to this consideration, the selected number of samples was 750. The validation curves confirm that this parameter is the best fit (Figure 9).

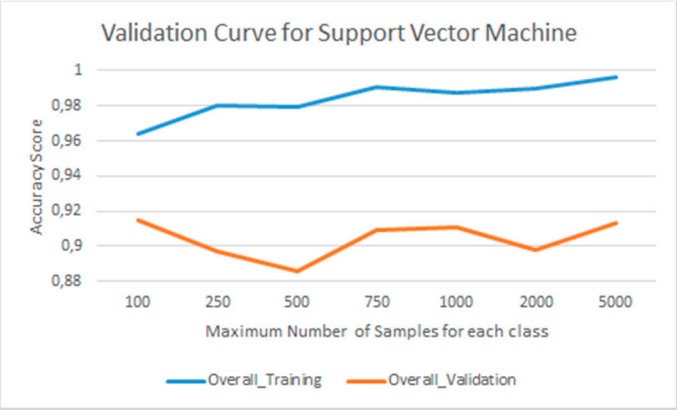

Figure 9. SVM validation curve for the maximum number of samples.

Table 9 reports the processing time for each configuration. It is possible to notice that, as expected, the computational time increases according to the increase of the sample size.

3.3.2. Hyperparameter Tuning for the Multispectral Dataset

With the RF, the samples size was set to 10,000 for each class. Tree depth and the number of trees ranges were chosen according to the number of samples and the image resolution. For multispectral images, the training sample size is five times greater than the hyperspectral, and the resolution is 4000 × 3000 pixels instead of 1010 × 1010 pixels. The case 50_30 was chosen as starting point according to the previous tuning on hyperspectral data (Section 3.3.1). The tree depth optimization was done, varying its value from 30 to 60. The number of trees ranged from 50 to 125.

Table 10 presents the training and validation accuracies considering all the combinations of the number of trees (xx in the test code) and the depth tree (yy in the test code).

The OA in training and validation is constant in all configurations, 81% and 77%, respectively. As a consequence, the best configuration can be defined, looking only to the validation accuracy. The OA accuracy is not strongly affected by the different hyperparameters. However, the test 100_30 presents the highest value of OA (77,8%). Looking at the clear ice UA, the best case should be 125_40 with a value of 69.8% instead of 69.7% in the case 100_30. From these observations, it is not possible to recognize this case as the best fit without the computational time analysis because the UA for C_i is quite similar (Table 11).

Table 10. Training accuracy (on the left) and validation accuracy (on the right) RF tree depth and the number of trees optimization. R_i stands for rime ice, C_i for clear ice, and W_a for white aluminum. The test name is defined as xx_yy, where xx is the number of trees, yy is the depth. In yellow the selected optimized hyperparameter is shown.

Test (xx_yy)	Training Accuracy (%)				Validation Accuracy (%)			
	R_i	C_i	W_a	Overall	R_i	C_i	W_a	Overall
50_30	70.5078	66.9162	56.3398	81.6832	55.3701	68.7615	60.2531	77.1311
50_40	69.8004	67.3592	56.2977	816582	54.1104	67.8014	60.6952	77.0461
50_50	69.5568	67.5865	55.8317	81.5282	53.0954	68.4173	60.8929	77.2861
50_60	70.0849	67.6036	54.2184	81.4681	53.5284	68.5125	60.6882	77.6211
75_30	68.8278	67.2374	55.5764	81.4281	53.9481	67.9496	60.4317	77.1311
75_40	69.9799	67.5952	56.2234	81.7932	53.4521	68.7778	59.3454	76.9912
75_50	69.5681	66.3261	55.0188	81.5732	53.0235	67.9594	59.4336	77.0561
75_60	68.8259	66.6204	53.4672	81.2381	54.9031	68.4853	61.2722	77.3711
100_30	67.2472	67.9566	54.3531	81.4881	56.2536	**69.6554**	61.6554	77.8161
100_40	71.1968	67.5524	56.8807	81.8782	54.2343	69.0955	60.4542	77.6161
100_50	70.6303	67.463	57.5832	81.3081	54.5098	68.7132	61.0991	77.3811
100_60	69.8595	67.5313	55.8442	81.3281	54.0376	69.2891	60.8715	77.5611
125_30	68.8588	66.3749	54.7497	81.2081	53.224	69.2393	59.2529	77.4511
125_40	69.0385	67.5875	55.7325	81.7182	52.5876	**69.8535**	60.451	77.6211
125_50	69.5795	67.8422	56.7398	81.7282	53.6558	69.1431	59.2657	77.3261
125_60	691205	66.6516	55.627	81.4181	53.7237	69.7354	60.792	77.7861

Table 11. Processing time for the training in RF using different values of tree depth and the number of trees. In yellow the time for training the model with the selected optimized hyperparameter is shown.

Test	Processing Time
50_30	2′59″
50_40	2′54″
50_50	2′57″
50_60	3′4″
75_30	4′50″
75_40	5′8″
75_50	5′4″
75_60	5′13″
100_30	6′55″
100_40	7′23″
100_50	7′25″
100_60	7′24″
125_30	9′23″
125_40	8′58″
125_50	9′1″
125_60	9′25″

For what concerns the processing time, in all configurations, the trend increases according to the number of trees increasing. The case 100_30 was selected because it has a good trade-off between processing time (6′55″ instead of 8′58″ for the case 125_40), overall and glassy ice accuracy.

For the SVM classifier, only the sample size per class ranged between 500 and 2000 samples. In Table 12, it can be noticed that cases with 1500 and 2000 samples have the best OA and the latter also has the highest value for the C_i accuracy (68.6%). Nonetheless, taking into account the computational time (Table 13), the test with 2000 samples lasts around 20 min more than the test with 1500 samples (1h5′43″). The best fit can be considered the configuration with 1500 samples.

Table 12. Training accuracy (on the left) and validation accuracy (on the right) optimization of SVM number of the sample. R_i stands for rime ice, C_i for clear ice, and W_a for white aluminum. In yellow the selected optimized hyperparameter is shown.

N_Samples	Training Accuracy (%)				Validation Accuracy (%)			
	R_i	C_i	W_a	Overall	R_i	C_i	W_a	Overall
500	67.8201	62.1644	54.3074	80.9231	51.924	64.772	60.339	75.9512
1000	67.1179	64.2766	57.4959	80.463	51.8227	66.7797	62.0635	76.6062
1500	66.7683	64.8352	54.526	80.213	52.5362	67.6934	61.0333	77.0711
2000	68.616	66.6042	54.0258	80.283	53.9737	68.603	61.3453	77.0461

Table 13. Processing time for the training in SVM using different values of the number of samples. In yellow the time for training the model with the selected optimized hyperparameter is shown.

N_Samples	Processing Time
500	24′47″
1000	50′47″
1500	1h5′43″
2000	1h23′20″

Table 13 provides the computational time in all configurations and displays that the increase of sample size defines the increasing of the processing time trend in a proportional way.

3.4. Ice Detection using Hyperspectral Data: Results

The ice detection was performed on three types of hypercubes:

1. the original hypercubes collected with the Senop Rikola (Section 3.1);
2. the reduced hypercubes composed by the selected bands (27 bands and 10 bands) (Section 3.2);
3. the images formed by the three principal components (PC images) (Section 3.2).

For the classification, the "Classify Raster" tool of ArcGIS Pro 2.5.0 [51] and the Test_1 was employed for the training. The analysis, in this section, is focused on two main parameters: the accuracy and the computational time for the classification only.

In general, as explained in Section 2.5 both overall accuracy and user's accuracy were used for assessing the classification. As mentioned in Section 3.1, some materials are included in the background, but at the same time annotated as different classes to check the performance on different materials. Since these classes were not included in the real scenario, because in that case, the background will be different (e.g., asphalt instead of floor tile), the overall accuracy was included just to show the general performance of the algorithms. However, the primary parameter is the user's accuracy, because the object of this study is the detection of the ice and in detail the clear ice due to its transparent property.

For each dataset (original hypercubes, reduced hypercubes, and PC images), random forest and SVM with the optimized hyperparameters derived from Section 3.3.1 were used. For the RF, the hyperparameters selected for the classification are the number of trees equal to 50, tree depth equal to 30 and 2000 samples. For the SVM, the classification with 750 samples was performed.

As in the case of PCA (Section 3.2), the classification evaluation is shown only in four representative images (Test_1, Test_2, Test_6, and Test_10). Test_1, Test_2 and Test_6 present varied environmental conditions and Test 10 was included to display the behavior of the model in the presence of ice phase transition.

For what concerns the original dataset classified with the RF, the overall accuracy reaches a maximum value of 88%, and the computational time is 14 min on average (Table 14). The classification performs better on the clear ice than the other classes, reaching a maximum value of ca. 96%. Regarding the rime ice, its accuracy is on average under 50% since the radiometric response is similar to that related to the white aluminum (67% on average).

Table 14. Accuracy and processing time on the original dataset with random forest. R_i stands for rime ice, C_i for clear ice, and W_a for white aluminum.

Test	Accuracy (%)				Processing Time
	R_i	C_i	W_a	Overall	
Test_1	41.7311	96.3272	69.0852	87.3775	13′22″
Test_2	46.3158	96.3636	70.9497	88.8378	14′38″
Test_6	53.2446	81.6364	69.1238	86.9974	12′37″
Test_10	43.3428	52.7938	59.5405	81.5763	16′9″

With SVM classifier (Table 15), the overall accuracy reaches a maximum value of 92%, and the computational time is 17 min on average. Additionally, in this case, C_i user's accuracy is higher than the other significant classes, reaching a maximum value of ca. 97%. Rime ice accuracy is on average under 60% since the radiometric response is similar to that related to the white aluminum (78% on average).

Table 15. Accuracy and processing time on the original dataset with SVM. R_i stands for rime ice, C_i for clear ice, and W_a for white aluminum.

Test	Accuracy (%)				Processing Time
	R_i	C_i	W_a	Overall	
Test_1	50.4886	95.1975	88.0297	90.39	15′2″
Test_2	56.6781	97.0843	82.5337	91.91	16′48″
Test_6	56.5801	80.2994	72.5888	88.39	15′4″
Test_10	63.1236	55.4054	64.4209	84.54	21′29″

As presented in Tables 14 and 15, the SVM reaches better accuracy on average than the RF. Indeed, the overall accuracy is 86.2% for the RF and 88.8% for SVM. The drawback of the SVM is the processing time. The average computational time is 14′22″ for the RF and the 17′5″ for the SVM. Therefore, the RF is faster than the SVM in the classification process.

For both algorithms, Test_10 reports low user's accuracy values compared with the other tests, because in this case, ice was starting to melt. This evidence also recurs in the reduced hypercube datasets and the PCA dataset. There is only one test in which the ice is starting to melt. Thus, it is predictable that the algorithm, in this case, works worst and the detection of the ice in other physical states was out of this preliminary study. It is well known that the ice changes its features according to its state. Thus, for the real case application, further acquisitions will be made for training the algorithm and improving the detection of the ice while changing its state to the liquid one.

The same analysis was carried out for the reduced hypercube datasets (27 bands and 10 bands) (Table 16). The reached OA with RF classifier has a maximum value of 83.8% for the hypercubes with 27 bands and 80.6% for the hypercubes with 10 bands. The computational time varies from 28.5″ for the 27 bands to 26.5″ for the 10 bands. These observations demonstrate that the two cases are comparable, and the dataset with 10 bands can be considered reliable. Moreover, the C_i user's accuracy stands that the model has in any way a good performance. As a consequence, it verifies that the set of bands selected using the PCA is adequate for the classification task.

Table 16. Accuracy and processing time on the reduced hypercube (27 and 10 bands) with RF. R_i stands for rime ice, C_i for clear ice, and W_a for white aluminum.

Bands N°	Test	Accuracy (%)				Processing Time
		R_i	C_i	W_a	Overall	
27	Test_1	41.7047	92.7769	69.1438	82.5365	25″
	Test_2	52.2158	93.28	74.6367	**86.0572**	28″
	Test_6	40.4255	70.997	67.0165	80.3561	31″
	Test_10	34.1797	54.2955	56.0549	76.3753	30″
10	Test_1	43.5185	95.4392	69.9213	80.8962	24″
	Test_2	50.1825	94.8074	74.7238	**83.2366**	27″
	Test_6	38.9886	80.3754	59.7285	77.0154	26″
	Test_10	30.7018	57.7818	56.6897	73.0546	29″

Table 17 presents the outcomes of the SVM. The overall accuracy reaches a maximum value of 87.3% for the 27 bands-hypercubes and 80.8% for the 10 bands-hypercubes. The processing time varies from 1′50″ to 1′40″, respectively, for the 27 bands and the 10 bands. Regarding the clear ice, the dimensionality reduction does not affect accuracy. Even if the SVM accuracy has the same trend of RF for the two datasets, there are still slight differences (Tables 16 and 17). Referring to the average OA for the 27 bands, the SVM performs better than the RF, while for the 10 bands is the opposite. For both reduced datasets, the SVM has a higher user's accuracy of C_i class, and it is slower compared with the RF.

Table 17. Accuracy and processing time on the reduced hypercube (27 and 10 bands) with SVM. R_i stands for rime ice, C_i for clear ice, and W_a for white aluminum.

Bands N°	Test	Accuracy (%)				Processing Time
		R_i	C_i	W_a	Overall	
27	Test_1	45.2203	94.4217	70.1794	82.84	1′33″
	Test_2	62.4833	96.7882	72.6236	**87.32**	1′52″
	Test_6	53.3607	84.0215	68.2892	82.18	1′58″
	Test_10	45.7447	61.0082	59.3465	78.02	1′57″
10	Test_1	44.1153	94.9749	70.038	79.37	1′35″
	Test_2	48.8522	96.144	72.9124	**80.84**	1′43″
	Test_6	49.3736	84.3806	67.8459	78.89	1′42″
	Test_10	33.6364	61.4594	55.3836	72.78	1′40″

Turning now to the analysis on the PC dataset, for the RF algorithm, Table 18 indicates that there is a perceivable reduction of the accuracy rate concerning original and reduced hypercubes. The average OA does not overcome the value of 72%. Regarding the computational time, it is around 38″ on average.

Table 18. Accuracy and processing time on the PC images with RF. R_i stands for rime ice, C_i for clear ice, and W_a for white aluminum.

Test	Accuracy (%)				Processing Time
	R_i	C_i	W_a	Overall	
Test_1	37.8215	94.8214	65.92	75.8552	34″
Test_2	41.1435	75.6241	78.6325	73.5947	39″
Test_6	33.5897	68.3121	57.9747	69.4139	39″
Test_10	46.1314	54.9505	72.1607	70.9542	41″

For the SVM algorithm, Table 19 presents findings comparable to the RF one. The average OA does not overcome the value of 76%. Moreover, the computational time is around 1'17" on average.

Table 19. Accuracy and processing time on the PC images with SVM. R_i stands for rime ice, C_i for clear ice, and W_a for white aluminum.

Test	Accuracy (%)				Processing Time
	R_i	C_i	W_a	Overall	
Test_1	42.4956	95.3888	68.2451	78.16	59"
Test_2	42.7742	81.6558	80.7923	78.24	1'5"
Test_6	45.1613	77.8899	64.8072	76.05	1'16"
Test_10	45.9184	56.5625	71.4886	72.29	1'48"

These observations demonstrate that either SVM and RF produce ambiguous and inaccurate outcomes in some cases and the resulting average accuracy for clear ice (77.8% for SVM and 73.4% for RF) could be not acceptable for our applications. Finally, the comparison between the two algorithms proves that SVM behaves better than RF. Overall and user's accuracy for all classes always have a higher value in SVM, but its computational time is twice as much as RF's one. Hence the classification with the SVM is still slower than the RF's.

According to the above analysis, for the ice detection in all datasets (original, reduced 27 bands, reduced 10 bands and PC), it is possible to make some general considerations. SVM and RF accuracy are comparable in all cases. Indeed, the difference between the accuracies is from 0.1% to 2%. Despite these minimal differences, the SVM presents the higher values of the user's and overall accuracy on average than the RF classifier. The reduction of dimensionality affects the overall accuracy slightly. Considering the difference between the original hypercube and the PC hypercube, the OA decreasing on average is 12% for the SVM classifier, instead of 14% for the RF one.

For the user's accuracy of the C_i, a descending trend based on the dimensionality of the feature space cannot be defined. However, the differences of clear ice accuracy between the best case (10 bands hypercube) and the worst case (PC images) is 7% for the SVM and 9% for the RF.

The processing time is strictly related to the size of the feature space. The dimensionality reduction helps to contract the processing time. In general, the analysis on the computational time shows that the RF is faster than the SVM, both in training and classification parts.

Figure 10 illustrates the results of the classification for each dataset. It can be appreciated graphically the discrepancy related to the reduction of the number of bands. Specifically, Figure 10 refers to the Test_1 hypercube classified with SVM. The RF graphical results are not included because it is impossible to detect the visual differences compared with the SVM accuracy. As can be seen, the clear ice is well detected in all the cases; while the rime ice identification gets worse according to the reduction of the bands.

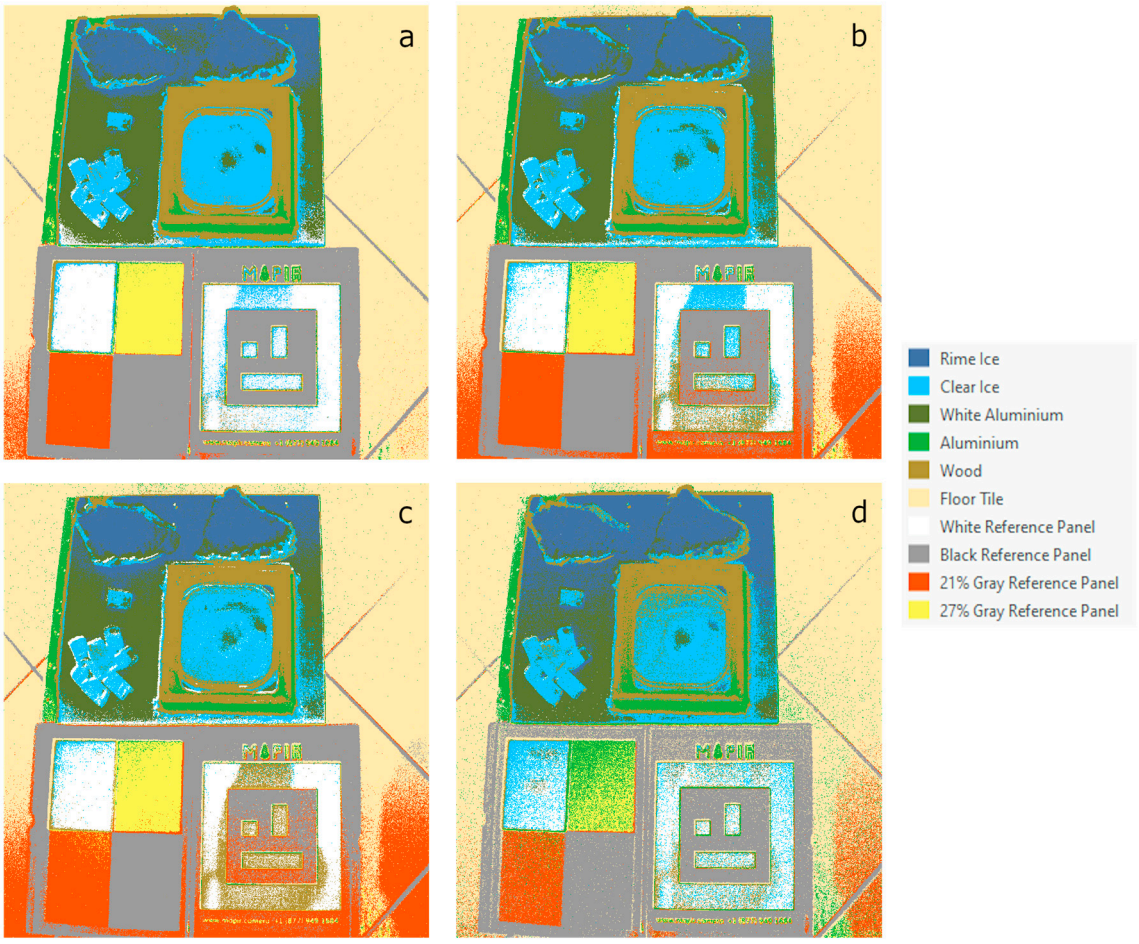

Figure 10. The classification results on Test_1 with SVM. (**a**) Original classified hypercube, (**b**) reduced classified hypercube-27 bands, (**c**) reduced classified hypercube-10 bands, and (**d**) PC classified hypercube.

3.5. Ice Detection Using Multispectral Data: Results

The ice detection was performed in the case of multispectral images on two datasets:

1. the original data acquired with the MAPIR (RGN images);
2. the multispectral data composed with the blue band of the RGB (RGBN images).

The optimized hyperparameters identified after the tuning (Section 3.3.2) were used. For the RF, the hyperparameters selected for the classification are the number of trees equal to 100, tree depth equal to 30 and 10,000 samples. For the SVM, the number of samples is 1500.

Considering the RGN dataset, the classification assessment is described in four representative images (Test_1, Test_2, Test_6, and Test_10) that have the same characteristics of the hypercubes described in Section 3.4. For the RGBN dataset, only the Test_1 integrated with the blue band is cited to demonstrate the improvement related to the blue presence in terms of accuracy.

With the RF, the evaluation of the outcomes on RGN images shows that OA is on average 49.5%. Instead, the computational time on average is 6'43" (Table 20). As expected, the lack of blue band alters deeply also values of UA compared with hyperspectral hypercubes. For example, the maximum value for the glassy ice accuracy is lower than the 80%. Thus, it is not sufficient to be considered correct and accurate.

Table 20. Accuracy and processing time on the RGN (red, green, and near infrared) images with RF. R_i stands for rime ice, C_i for clear ice, and W_a for white aluminum.

Test	Accuracy (%)				Processing Time
	R_i	C_i	W_a	Overall	
Test_1	54.9565	69.7936	60.7158	77.6311	6′40″
Test_2	36.1789	42.5788	35.2851	47.9226	6′36″
Test_6	25.2031	33.7942	8.2353	30.9185	6′30″
Test_10	34.7062	33.3841	44.8087	41.34	7′5″

With the SVM, the accuracy assessment of the RGN dataset shows that OA is on average 49.2%. Instead, the computational time on average is 23′40″ (Table 21). The missing blue band problem is still visible. Indeed, the clear ice accuracy does not surpass the value of 67% in the best configuration.

Table 21. Accuracy and processing time on the RGN images with SVM. R_i stands for rime ice, C_i for clear ice, and W_a for white aluminum.

Test	Accuracy (%)				Processing Time
	R_i	C_i	W_a	Overall	
Test_1	52.2795	67.3587	63.1305	76.9362	22′13″
Test_2	34.8404	41.2713	32.8609	46.8527	23′9″
Test_6	25.8493	35.6347	10.6874	30.6135	23′55″
Test_10	40.3123	37.1409	42.4566	42.375	25′21″

Nonetheless, the comparison among the two algorithms shows that RF performs better classification on average and its computational time is lower than SVM's one.

Considering that the accuracy assessment for the RGN dataset is not comparable with the ones obtained using the hypercube with similar feature space size (e.g., PC images), the blue band was added to create RGBN images. Indeed, looking to the results of the band selection (Section 3.2) it is possible to notice that the blue band has essential weight.

Table 22 illustrates an overview of the accuracy and the processing time in the RGBN case. As can be seen, the OA reported a significant increase compared with the RGN images. Indeed, it overcomes the value of 80% for both the algorithms. The UA values are higher than the average of the respective values in the RGN case. At the same time, the computational time decreases for the RGBN images.

Table 22. Accuracy and processing time on the RGBN(red, green, blue, near-infrared) image with RF and SVM. R_i stands for rime ice, C_i for clear ice, and W_a for white aluminum.

Classifier	Accuracy (%)				Processing Time
	R_i	C_i	W_a	Overall	
RF	57.9867	54.2284	83.7163	84.495	4′49″
SVM	62.3726	57.3843	88.1515	86.695	8′18″

Regarding the comparison between the two algorithms in the RGBN case, the SVM produces a better classification, but its computational time is longer than the RF's one.

Figure 11 highlights the results of the classification in the two datasets using the SVM. The improvement related to the introduction of the additional band is evident. The blue band allows to reduce the classification noise and at the same time a better identification of all materials. Moreover, the enhancement of the distinction between rime ice or clear ice and white aluminum is clear.

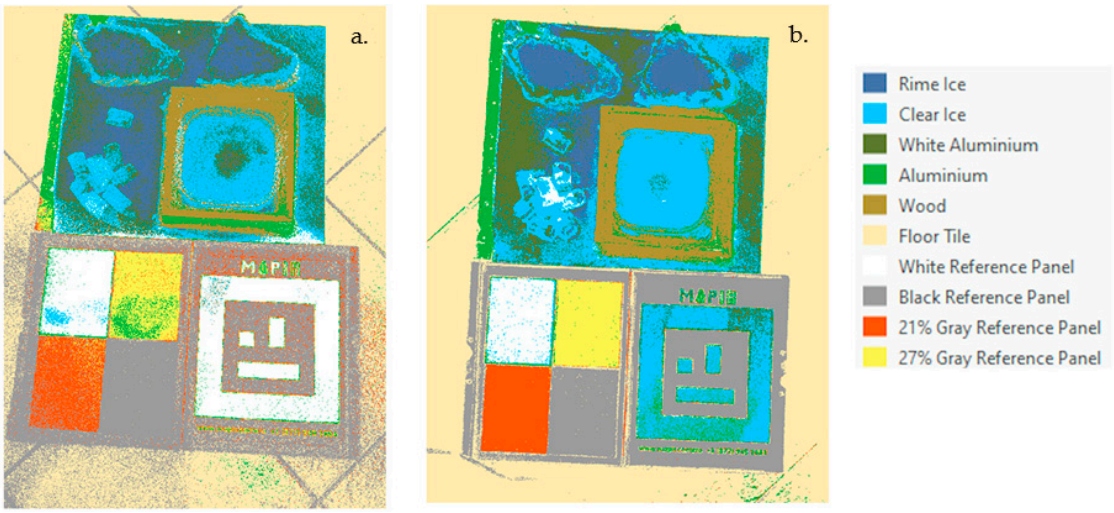

Figure 11. The classification results on Test_1 with SVM. (**a**) RGN classified images and (**b**) RGBN classified images.

4. Discussion

Since the previous section describes already the outcomes related to the dimensionality reduction of hyperspectral data and the classification with the two algorithms on the different datasets, the discussion focuses on:

- the consideration related to the use of hyperspectral or multispectral images for the ice detection task;
- the comparison between the two classifiers in terms of computation time and accuracy;
- the comparison among the detection of the different material.

For what concern the first point, this study confirms that for the ice detection task, the use of hyperspectral images is more reliable. However, on the other hand, highlights that the advantage of operating with multispectral data (with the same spatial resolution of the hyperspectral images) is related to the computational time. The latter, indeed, is one of the crucial problems in real-time application.

The outcomes demonstrate that the hyperspectral data are suitable for real-time applications after a priori analysis. A dimensionality reduction process can easily compress the size of hyperspectral data preparing the data for the classification task. This step leads to breaking down the limits related to the computational time of the original hypercubes. The experimental analysis shows that the processing time can be improved, downscaling the spectral resolution. The case of reduced hypercube with 10 bands can be considered a trade-off between accuracy and computational time, regardless of the employed algorithm.

Moreover, starting with the analysis of significative bands, a multispectral sensor can be defined for facilitating the acquisition and classification operations. It is crucial to take into account, in the case of multispectral images, that the spatial resolution is four times greater than the hyperspectral. The results obtained with the introduction of a significative band (in this case, the blue band) shows the effectiveness of predefined-band knowledge in the classification.

Regarding the classifiers, it is possible to state that the SVM performs better in terms of accuracy; on the opposite, the RF classifier is faster than the SVM. This observation is valid for both datasets: the multispectral and the hyperspectral dataset. However, the accuracy reached with the multispectral data is not comparable with ones of the hyperspectral camera regardless of the selected algorithm.

Finally, turning now to material detection, this study focuses the attention on the classification of clear ice. As explained in the Introduction (Section 1), the clear ice is not visible with the naked eye and requires tactile inspection. The ice, in the real case, could have different features (e.g., density, shape) than the ice samples that were generated in this study. At this stage of our work, other types of ice

were examined to understand if the classification was able to distinguish, as two classes, the different condition of the same material, such as transparent ice and rime ice. Due to its characteristics, the rime ice is more visible; thus, the detection is more straightforward. Concerning the ice detection reliability, it is possible to underline that both algorithms conservatively recognize both forms of ice. For instance, if an area is contaminated by ice, conservative way means that it is much more unlikely that the algorithms recognize that area as aluminum or as other material instead of ice (this is the case of a false negative). This example implies that even if the algorithms recognize an area of aluminum as ice, it is very improbable that the contrary occurs. It is possible to notice in Figures 10 and 11 that the main false negative can be associated with the white aluminum, that is identified as ice.

In the real case, not to apply de-icing fluids to some areas with ice is more dangerous than applying the de-icing procedure on some other areas that do not need it. Therefore, it could be possible that some de-icing fluids will be wasted for areas in which actually ice is not present; however, the safety, that is a primary goal, is not compromised.

Moreover, although the radiometric classification is noisy for the rime ice, the results can be improved with the use of geometric features. The previous activities of the SEI project [14] demonstrated that the rime ice is accurately identifiable in this application, using the RGB sensor and a photogrammetric approach.

5. Conclusions

In this paper, the feasibility to use spectral images for ice detection was studied, testing different sensors with different spectral resolutions such as the Senop Rikola and the MAPIR. To address this purpose, two different types of ice samples were created to understand if it was possible to distinguish clear ice (transparent ice) from rime ice (white ice). Then, images were collected in different illumination conditions, because there is no open-source and ready to use dataset to face this specific task. Moreover, semantic segmentation algorithms (such as RF and SVM) were defined, also evaluating the accuracy and the processing time.

The main challenges of this work were the definition of the efficient use of hyperspectral data in the near real-time application and the research of spectral resolution and algorithms capable of providing higher accuracy and limited computational time.

This experimental analysis demonstrates the possibility to use the reduced hypercubes and both the RF and SVM as a classifier with an OA higher than 80% on average.

As feature work, we plan to transfer the knowledge and the promising outcomes acquired through this simulation in a drone application. Moreover, the drone's application can help to consider also other kinds of ice that are not possible to reproduce in the laboratory, such as snow and freezing rain.

Author Contributions: Conceptualization, A.M.L., L.M. and M.A.M.; methodology and validation, A.M.L., L.M. and M.A.M.; writing—original draft preparation, L.M. and M.A.M.; writing—review and editing, A.M.L., L.M. and M.A.M.; visualization, L.M. and M.A.M.; supervision, A.M.L.; project administration, M.A.M.; funding acquisition, A.M.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Union and the Piedmont Region funds within the framework of the Action "MANUNET III—POR FESR 2014–2020 (project code: MNET18/ICT-3438)".

Acknowledgments: This project was carried out within the activities of the PoliTO Interdepartmental Centre for Service Robotics (PIC4SeR).

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

References

1. European Organisation for the Safety of Air Navigation. The Flight Safety Foundation Aircraft Ground De/Anti-Icing. Available online: https://www.skybrary.aero/index.php/Aircraft_Ground_De/Anti-Icing (accessed on 28 June 2020).

2. International Civil Aviation Organization (ICAO) Aircraft Ground De-Icing/Anti-Icing Operations. Available online: https://www.icao.int/safety/airnavigation/OPS/Pages/Aircraft-Ground-De-IcingAnti-Icing-Operations.aspx (accessed on 25 June 2020).
3. Castro, S.; Davis, L.C.; Erickson, L.E. Plant-Enhanced Remediation of Glycol-Based Aircraft Deicing Fluids. *Pract. Period. Hazard. Toxic Radioact. Waste Manag.* **2001**, *5*, 141–152. [CrossRef]
4. Hoover, G.A. *Aircraft Ice Detectors and Related Technologies for Onground and Inflight Applications*; U.S. Department of Transportation, Fedaral Aviation Administration Technical Center, Atlantic City International Airport: Atlantic County, NJ, USA, 1993; p. 54.
5. Mingione, G.; Barocco, M.; Denti, E.; Bindi, F.G. *Flight in Icing Conditions*; Direction Générale de l'Aviation civile (DGAC): Paris, France, 1997; p. 188.
6. Federal Aviation Administration, United States Department of Transportation. *AC 00-6A—Aviation Weather for Pilots and Flight Operations Personnel*; Fedaral Aviation Administration: Washington, DC, USA, 1975.
7. Federal Aviation Administration, United States Department of Transportation. *AC 00-6B—Aviation Weather 2016*; Fedaral Aviation Administration: Washington, DC, USA, 2016.
8. Icing Hazards. Available online: https://www.weather.gov/source/zhu/ZHU_Training_Page/icing_stuff/icing/icing.htm (accessed on 28 June 2020).
9. American Meteorological Society Rime. Available online: http://glossary.ametsoc.org/wiki/Rime (accessed on 29 July 2020).
10. American Meteorological Society Glaze. Available online: http://glossary.ametsoc.org/wiki/Glaze (accessed on 29 July 2020).
11. Federal Aviation Administration, United States Department of Transportation Standardized International Aircraft Ground Deice Program (SIAGDP). Available online: https://www.faa.gov/other_visit/aviation_industry/airline_operators/airline_safety/deicing/media/standardized_international_ground_deice_program.pdf (accessed on 25 June 2020).
12. Garr-Schultz, A.; Gardner, W. Strategic Self-Presentation of Women in STEM. *Soc. Sci.* **2018**, *7*, 20. [CrossRef]
13. Falcone, A.; Miccone, D.; Eytan, A.; Theoharatos, C.; Stavropoulos, P.; Aicardi, I.; Musci, M.A.; Maria, A. Spectral Evidence of Ice on Aircrafts—The SEI Project. *Inf. Intell. Syst. Appl.* **2020**, *1*, 69–73.
14. SEI Project Consortium SEI Project. Available online: https://www.sei-project.eu/ (accessed on 7 July 2020).
15. Rizk, P.; Al Saleh, N.; Younes, R.; Ilinca, A.; Khoder, J. Hyperspectral imaging applied for the detection of wind turbine blade damage and icing. *Remote Sens. Appl. Soc. Environ.* **2020**, *18*, 100291. [CrossRef]
16. Bassey, C.E.; Simpson, G.R. Aircraft Ice Detection using Time Domain Reflectometry with Coplanar Sensors. In Proceedings of the 2007 IEEE Aerospace Conference, Big Sky, MT, USA, 3–10 March 2007; pp. 1–6.
17. Roy, C. Remote Sensing in Precision Agriculture: An Educational Primer. Master's Thesis, Iowa State University, Ames, IA, USA, July 1999.
18. Breda, J.-M.; Choisnet, J.; Naim, M. Optical Device for the Detection of Icing Conditions on Aircraft 2000. U.S. Patent US6091335A, 18 July 2000.
19. Gong, X.; Bansmer, S. 3-D ice shape measurements using mid-infrared laser scanning. *Opt. Express* **2015**, *23*, 4908. [CrossRef] [PubMed]
20. Nakauchi, S.; Nishino, K.; Yamashita, T. Selection of optimal combinations of band-pass filters for ice detection by hyperspectral imaging. *Opt. Express* **2012**, *20*, 986. [CrossRef] [PubMed]
21. Huber, N.; Kalidindi, S.R.; Klusemann, B.; Cyron, C.J. *Machine Learning and Data Mining in Materials Science*; Frontiers Media SA: Lausanne, Switzerland, 2020; ISBN 978-2-88963-651-8.
22. Camps-Valls, G.; Bruzzone, L. Kernel-based methods for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 1351–1362. [CrossRef]
23. Adam, E.M.; Mutanga, O.; Rugege, D.; Ismail, R. Discriminating the papyrus vegetation (Cyperus papyrus L.) and its co-existent species using random forest and hyperspectral data resampled to HYMAP. *Int. J. Remote Sens.* **2012**, *33*, 552–569. [CrossRef]
24. Tommaselli, A.M.G.; Santos, L.D.; de Oliveira, R.A.; Honkavaara, E. Refining the Geometric Calibration of a Hiperspectral Frame Camera with Preliminary Bands Coregistration. In Proceedings of the IGARSS 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 6468–6471.

25. Saari, H.; Aallos, V.-V.; Holmlund, C.; Mäkynen, J.; Delauré, B.; Nackaerts, K.; Michiels, B. Novel hyperspectral imager for lightweight UAVs. In *Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications VII*; Henry, D.J., Ed.; International Society for Optics and Photonics: Bellingham, WA, USA, 2010; p. 766805.
26. MAPIR Survey3 Cameras. Available online: https://www.mapir.camera/collections/survey3 (accessed on 28 June 2020).
27. Landgrebe, D.A. *Signal. Theory Methods in Multispectral Remote Sensing*; John Wiley & Sons: Hoboken, NJ, USA, 2005; ISBN 978-0-471-72125-3.
28. Preet, P.; Batra, S.S. Jayadeva Feature Selection for classification of hyperspectral data by minimizing a tight bound on the VC dimension. *arXiv* **2015**, arXiv:1509.08112.
29. Chang, C.I.; Du, Q.; Sun, T.L.; Althouse, M.L. A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 2631–2641. [CrossRef]
30. Li, F.; Lu, H. Hyperspectral images band selection using multi-dictionary based sparse representation. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 2769–2772.
31. Estornell, J.; Martı-Gavila, J.M.; Sebastia, M.T.; Mengual, J. Principal component analysis applied to remote sensing. *Model. Sci. Educ. Learn.* **2013**, *6*, 83–89. [CrossRef]
32. Fauvel, M.; Chanussot, J.; Benediktsson, J.A. Kernel Principal Component Analysis for the Classification of Hyperspectral Remote Sensing Data over Urban Areas. *EURASIP J. Adv. Signal. Process.* **2009**, *2009*, 783194. [CrossRef]
33. Torbick, N.; Becker, B. Evaluating Principal Components Analysis for Identifying Optimal Bands Using Wetland Hyperspectral Measurements from the Great Lakes, USA. *Remote Sens.* **2009**, *1*, 408–417. [CrossRef]
34. Jolliffe, I.T. *Principal Component Analysis*, 2nd ed.; Spinger Sieries in Statistics; Springer: New York, NY, USA, 2002; ISBN 978-0-387-22440-4.
35. Cattell, R.B. The Scree Test for the Number of Factors. *Multivar. Behav. Res.* **1966**, *1*, 245–276. [CrossRef] [PubMed]
36. Csillag, F.; Pásztor, L.; Biehl, L.L. Spectral band selection for the characterization of salinity status of soils. *Remote Sens. Environ.* **1993**, *43*, 231–242. [CrossRef]
37. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
38. Richards, J.A.; Jia, X. *Remote Sensing Digital Image Analysis: An Introduction*, 4th ed.; Springer: Berlin, Germany, 2006; ISBN 978-3-540-25128-6.
39. Zhang, C.; Kovacs, J.M. The application of small unmanned aerial systems for precision agriculture: A review. *Precis. Agric.* **2012**, *13*, 693–712. [CrossRef]
40. Khan, M.N.; Ahmed, M.M. Snow Detection using In-Vehicle Video Camera with Texture-Based Image Features Utilizing K-Nearest Neighbor, Support Vector Machine, and Random Forest. *Transp. Res. Rec.* **2019**, *2673*, 221–232. [CrossRef]
41. Hanselmann, M.; Köthe, U.; Kirchner, M.; Renard, B.Y.; Amstalden, E.R.; Glunde, K.; Heeren, R.M.A.; Hamprecht, F.A. Toward Digital Staining using Imaging Mass Spectrometry and Random Forests. *J. Proteome Res.* **2009**, *8*, 3558–3567. [CrossRef]
42. Bruzzone, L.; Persello, C. Approaches Based on Support Vector Machine to Classification of Remote Sensing Data. In *Handbook of Pattern Recognition and Computer Vision*; World Scientific: Singapore, 2009; pp. 329–352. ISBN 978-981-4273-38-1.
43. Myburgh, G.; van Niekerk, A. Impact of Training Set Size on Object-Based Land Cover Classification: A Comparison of Three Classifiers. *Int. J. Appl. Geosp. Res.* **2014**, *5*, 49–67. [CrossRef]
44. Lillesand, T.M.; Kiefer, R.W.; Chipman, J. *Remote Sensing and Image Interpretation*, 7th ed.; Wiley: Hoboken, NJ, USA, 2015; ISBN 1-118-34328-X.
45. Thanh Noi, P.; Kappas, M. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors* **2017**, *18*, 18. [CrossRef] [PubMed]
46. Geomatics Group of Politecnico di Torino Our Labs- Geomatics Labs. Available online: https://areeweb.polito.it/geomatics_lab/en/geomatic-lab/ (accessed on 4 July 2020).
47. L3 Harris Geospatial Docs Center—Using ENVI- Atmospheric Correction. Available online: https://www.harrisgeospatial.com/docs/AtmosphericCorrection.html#empirical_line_calibration (accessed on 4 July 2020).

48. ESRI. Principal Components. Available online: https://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/principal-components.htm (accessed on 3 July 2020).
49. ESRI. Composite Bands—Data Management Toolbox. Available online: https://pro.arcgis.com/en/pro-app/tool-reference/data-management/composite-bands.htm (accessed on 5 July 2020).
50. ESRI. An Overview of the Segmentation and Classification Toolset. Available online: https://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/an-overview-of-the-segmentation-and-classification-tools.htm (accessed on 30 June 2020).
51. ESRI. Classify Raster. Available online: https://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/classify-raster.htm (accessed on 5 July 2020).