Article

The Use of Drone Photo Material to Classify the Purity of Photovoltaic Panels Based on Statistical Classifiers

Tomasz Czarnecki * and Kacper Bloch

The Institute of Telecommunications, Faculty of Electronics and Information Technology,
Warsaw University of Technology, Nowowiejska 15/19, 00-665 Warsaw, Poland; 01112924@pw.edu.pl
* Correspondence: tomasz.czarnecki@pw.edu.pl

Abstract: The subject of this work is the analysis of methods of detecting soiling of photovoltaic panels. Environmental and weather conditions affect the efficiency of renewable energy sources. Accumulation of soil, dust, and dirt on the surface of the solar panels reduces the power generated by the panels. This paper presents several variants of the algorithm that uses various statistical classifiers to classify photovoltaic panels in terms of soiling. The base material was high-resolution photos and videos of solar panels and sets dedicated to solar farms. The classifiers were tested and analyzed in their effectiveness in detecting soiling. Based on the study results, a group of optimal classifiers was defined, and the classifier selected that gives the best results for a given problem. The results obtained in this study proved experimentally that the proposed solution provides a high rate of correct detections. The proposed innovative method is cheap and straightforward to implement, and allows use in most photovoltaic installations.

Keywords: photovoltaic panel; detection; soiling

1. Introduction

Since the 1950s, the temperature on Earth has risen on average by 0.2 C per decade. The temperature rise has a negative impact on the environment [1,2]. The consequences of climate warming include heatwaves deadly for humans, drinking water shortages, food production decline, coral reef degradation, and glacier melting [3]. To limit global warming, the European Union (EU) under the European Green Deal [4] initiated a strategy to achieve climate neutrality by 2050. The EU strategy has been endorsed by the European Parliament and the European Council, 2020 [5,6]. In order to achieve this goal, it is necessary to increase the use of renewable energy sources, which can achieve by building new renewable energy installations and increasing their energy efficiency.

Photovoltaics (PV) convert light into electricity using semiconductor materials that exhibit a photovoltaic effect. Solar cells are used to convert solar energy into electricity. The energy generated by solar cells is called "green energy". This means that it comes from a natural and renewable energy source, the sun, and its production does not release pollutants into the air [7]. This assumption is valid, but only after approximately three years of operation of the PV module because the energy expenditure needed to produce it is returned after this time [8]. In addition, the production of energy used to make the panel generates 300 kg CO_2 in the atmosphere. Advantages of photovoltaics compared to other types of RES are:

- Positive correlation between the intensity of sunlight and the daily demand for electricity,
- Increased generation in the summer period correlated with the demand for cold, and
- It enables the use of brownfield sites and poor-quality land, as well as building roofs [9].

The TÜV Rheinland Institute identified the most common problems based on its data from photovoltaic farms (Figure 1), industrial installations, and home micro-installations. We include among them:

– Dirt on PV panels,
– Incorrect installation,
– Shading,
– Discoloration of EVA foil,
– Glass breakage,
– Degradation by induced voltage,
– Path snails,
– Defective protective foil,
– Spot heating of panels [10].

The most common problem is dirty panels, which translates into huge losses in energy generated [11,12], as shown in 1. The dust accumulated on the surface of the photovoltaic panel comes mainly from the soil, rocks, construction debris, particles from car traffic, bird droppings, and pollen [13]. Dust accumulation on the surface of the panels obstructs the light, preventing it from reaching the PV cells, reducing energy production [14,15]. The energy loss depends on the amount of dust, particle size, and chemical composition of the powder. Contaminants have different effects on the light transmission process. Some dust particles can reduce the efficiency of photovoltaic devices by up to 98 [16,17]. To effectively produce electricity using PV cells, it is necessary to ensure failure-free operation of the PV installation throughout its lifetime (even up to 30 years), and a quick return on investment outlays. For this purpose, it is necessary to develop a fast, reliable, and straightforward method of checking the cleanliness of PV cells [18,19].
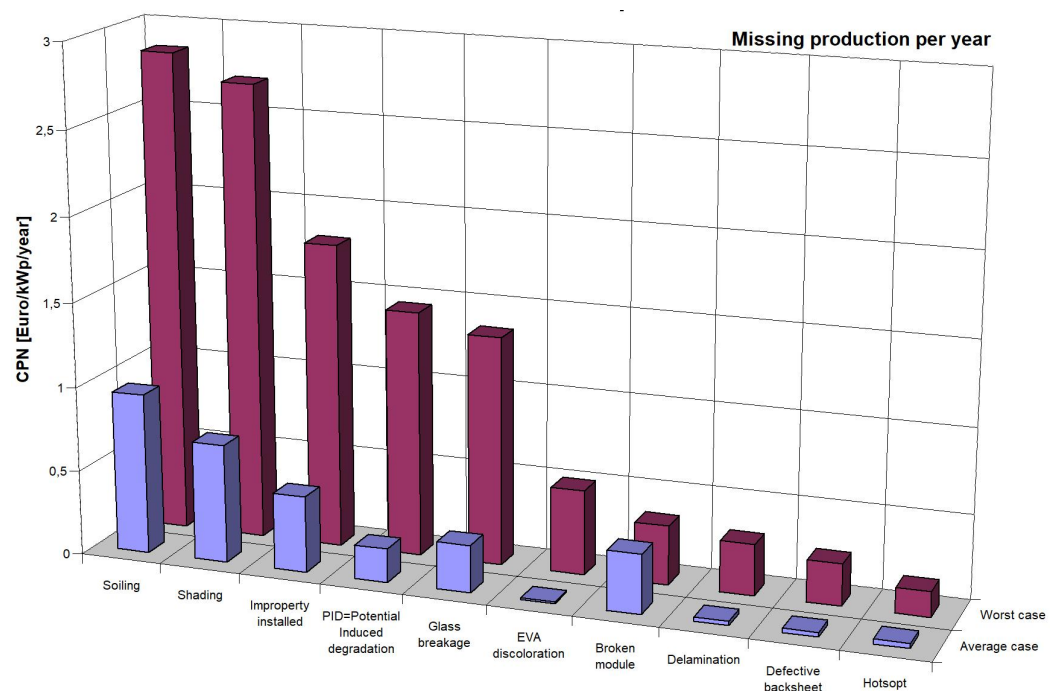

Figure 1. Missing electricity production due to various types of failure of solar panels [10].

2. Requirements

The system (Figure 2), which the algorithm is to be a part of, consists of a drone equipped with a high-resolution camera, an edge computing unit, automatic cleaning devices, and, optionally, may also consist of cloud computing. The principle of the system is as follows. The drone uses a camera to record video material showing PV panels in a given area. The recorded video material is sent to a computing unit, which extracts

individual frames from the video. It then analyzes them using the algorithm proposed, in this work, to detect dirty PV panels [20]. After detecting such a module, information about the need to clean the PV modules is sent to the cleaning devices. The photos collected can be sent to cloud computing for further analysis, e.g., to improve the algorithm's operation.
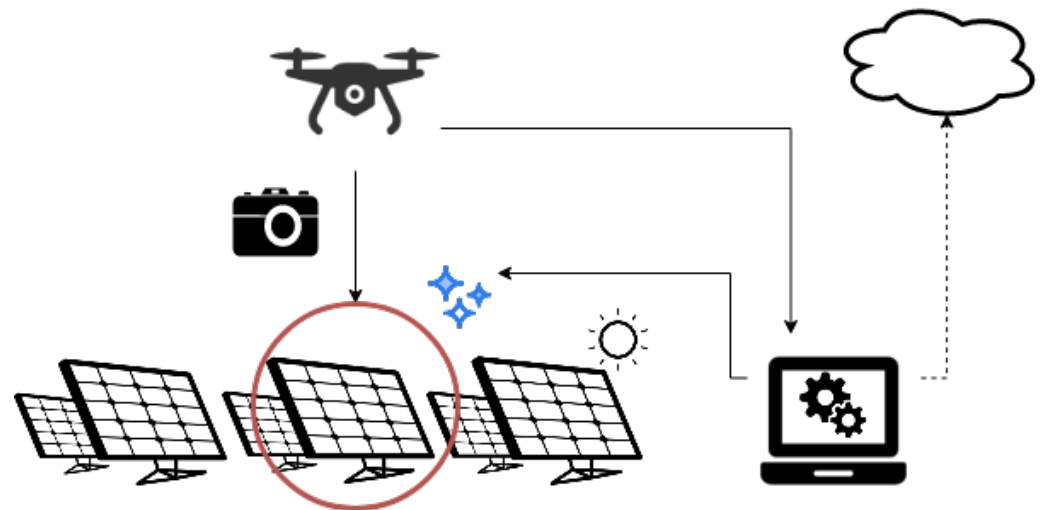

Figure 2. System.

The classification of a PV panel is to assign it to one of two classes (Figure 3):

– Clean,
– Dirty.

A clean panel works with the most excellent possible efficiency under the given atmospheric conditions, or the decrease in efficiency due to impurities on its surface is less than 25%. In turn, a dirty panel is one for which we notice a reduction in performance by at least 25%.
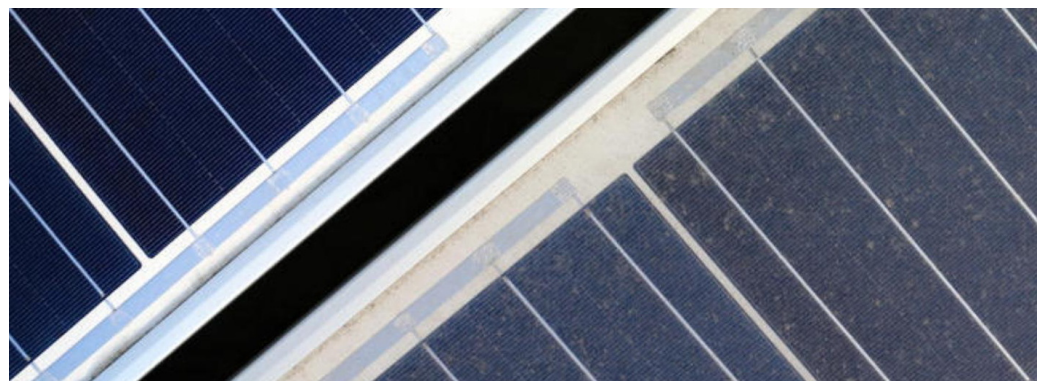

Figure 3. The classification of a PV panel—one of two classes.

2.1. Drone Photo Sourcing

In the case of video materials obtained from a drone (digital sensors), spatial resolution, the Ground Sampling Distance (GSD) [21], is of crucial importance. It measures the distance in the field between the centers of two adjacent pixels (Figure 4). For each measurement mission of the drone, the GSD is defined before the task [22]. Typical values for this type of task should be in the range of 1.5 cm–45 cm. In the case of monitoring photovoltaic panels, it is necessary to determine which defects and the amount of dirt must be detected [23]. To detect panels with mechanical problems, set the GSD within 25 cm. To detect physical damage or points more minor than the entire panel, set the GSD between 5–16 cm. In the case of dust and dirt detection missions, set the GSD > 2 cm—this will allow you to detect

even small contaminants on the panels [24]. The tested system uses RGB cameras with a resolution of 20 MP and a 1/1.7″ CMOS matrix (DJI Zenmuse H20).
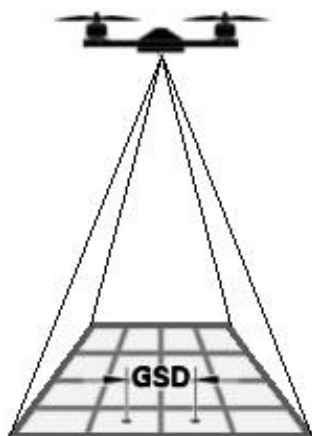

Figure 4. Ground sampling distance definition.

2.2. The Research Material

The input data will be the photographic material recorded on the test stand reflecting the actual conditions on the photovoltaic farm. The photographic material consists of 70 photos encoded in JPG format. The images show one, two, three, or four solar panels. The detailed breakdown is as follows:

– 60 photos containing one panel,
– 4 photos containing two panels,
– 4 photos containing three panels,
– 2 photos contain four panels.

The research material contained a total of four different polycrystalline PV panels, 44 clean and 44 dirty panels in total. First, pictures of clean PV panels were taken, then the dirt from the soil, rocks, construction debris, and particles from car traffic was gradually applied to them, and at a power drop of 10%, a photo was taken until archived a power drop of 50% compared to maximum power.

The images were taken:

– With adequate sunlight. The images were taken during the day with a minimum solar radiation intensity of 500 W/m^2, because below this value the PV panels are insufficiently illuminated, which means that the contrast of the photo is too low to extract the information that is important to us. The project does not assume artificial lighting of PV panels.
– Under appropriate weather conditions. Pictures cannot be taken during rainfall, as they introduce unwanted artifacts into the picture, making subsequent analysis difficult.
– At a minimum angle of 45 ° to the panel surface. Smaller values may make it impossible to extract the panel from the photo.
– At different times of the year. This approach will enable the use of classifiers throughout the year.

3. Statistical Classifiers to Classify Photovoltaic Panels

The project presents several algorithm variants that use different statistical classifiers to classify photovoltaic panels into one of two classes, clean or dirty, based on observation (feature vector). For the classification, it decided to use the characteristic feature of dirt. The image saturation decreases in the place where they occur, the image saturation decreases, and its luminance increases. Due to the two classes and two features, binary classifiers have been selected to classify observations with two parts. The algorithm consists of two stages: classifier training and classification. The photo material was divided into two sets: a

training set composed of 32 photos of clean panels and 32 photos of dirty panels. A test set consisted of 12 pictures of clean panels and 12 photos of dirty panels. Before classification, the image is pre-processed to remove unwanted background and leave only the PV module in the picture.

3.1. Extracting a Panel from a Photo

The first step in classification is to detect the PV module in the image. The PV surface is found in the image and extracted from the background at this stage. Any additional information is dropped from the input image.

This process takes place in three steps:

– Detection of all edges in the photo,
– Finding the edge of the PV panel,
– Application of a forward-looking transformation [25].

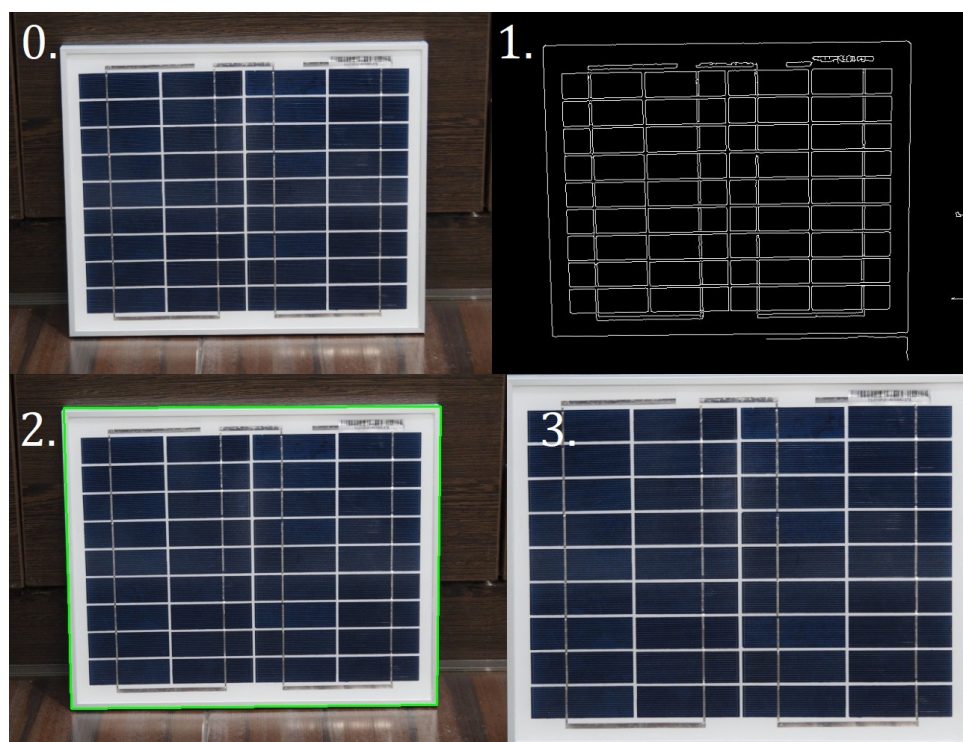The result of each of these steps is shown in Figure 5.

Figure 5. Successive stages of extracting the PV panel from the photo. 0—starting image; 1—detected edges in the picture; 2—outer edges of the PV panel found; 3—the result of the perspective transformation.

3.2. Observation

By visually comparing a clean and dirty panel, you can see that the surface of the dirty panel looks dull and lacks color intensity. This is because the dominant color of the dust is gray, and the shades of gray are not saturated. In addition, the surface of the pure solar panel is dark as the cell material absorbs incident light. Therefore, the amount of light reflected by the panel is limited. For a dirty PV module, its surface looks brighter because less light is absorbed by the cell, and therefore more light is reflected from the surface and scattered by the deposits [26]. Table 1 shows the average values of color saturation and luminance for the same panel in two cases, when clean and dirty. Figure 6 shows a visual comparison of this panel. These panels differ mainly in appearance. No structure is visible, no light reflection, surface heterogeneity. In turn, Figure 7 presents all the results of observations in a graphical form. These values confirm the hypothesis that for dirty panels, the image saturation decreases, and the luminance increases.

Table 1. Average values of saturation and luminance for a clean and dirty panel.

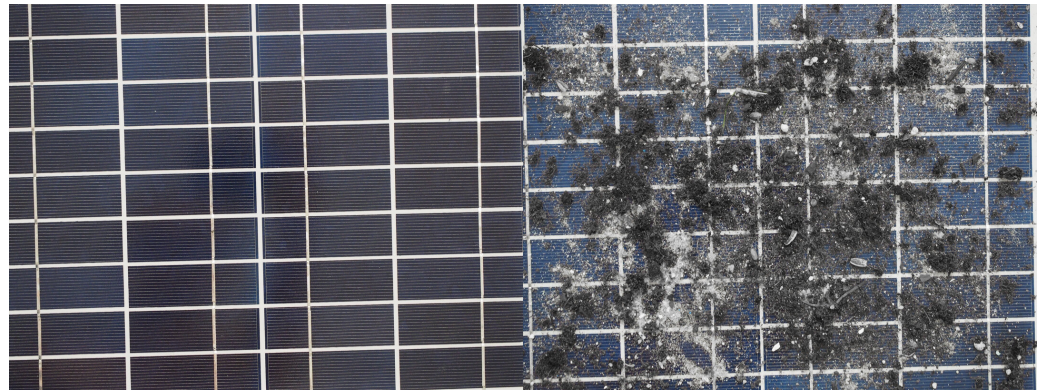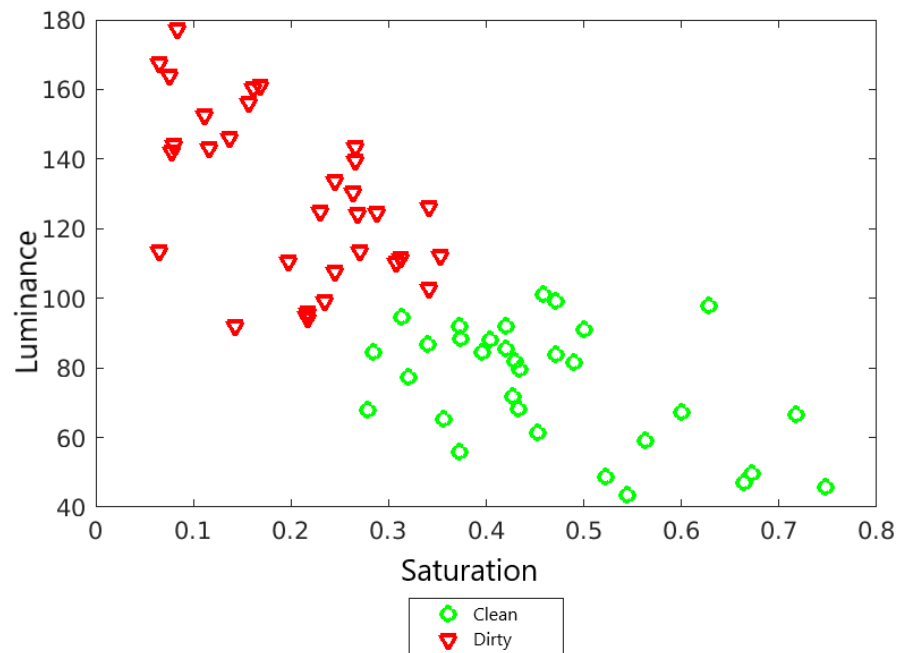|	Clean	Dirty
Saturation	0.4707	0.0891
Luminance	84.3143	143.0543

Figure 6. A clean panel (**left**) and a dirty panel (**right**) a comparison.

Figure 7. Observation results.

Image luminance is a value representing the image's brightness, calculated by ITU BT.601 [27]. For each pixel, calculate its luminance value and then calculate the arithmetic mean of these values.

To calculate the saturation value of the image, you can use the fact that the saturation is one of the components of the HSV model. So you can convert the image from RGB to HSV, extract the component corresponding to the saturation for each pixel, and then calculate the arithmetic mean of these values.

3.3. The Classifier of the k Nearest Neighbors

The classifier of the *k* nearest neighbors (kNN) classification method assigns a classified object to the class that is most frequently represented among the *k* closest neighbors from the training set [28]. In order to find the *k* nearest neighbors of the test object, the Euclidean

distance between the test object and all training objects is calculated [29]. In this case, the classifier consists of two features, so it is two-dimensional, which means that can lace of the things be placed on the Euclidean plane in the form of a point with Cartesian coordinates. Then, the Euclidean distance between two points is expressed by the formula:

$$d(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2} \tag{1}$$

where:

　　p—first point;
　　q—second point;
　　$d(p,q)$—Euclidean distance between points p and q;
　　p_1—coordinate X of point p;
　　p_2—coordinate Y of point p;
　　q_1—coordinate X of point q;
　　q_2—coordinate Y of point q.

As shown in Table 1, the values of saturation and contrast differ significantly in terms of magnitude, so the difference would dominate the calculated Euclidean distance because it affects the distance value much more. For this reason, it is necessary to normalize the value so that all dimensions for which the distance is calculated are equally relevant. Normalization consists of making the variable's values belong to the interval [0, 1]. The formula that expresses it:

$$x_j(i) = \frac{x_j(i) - min(x_j)}{max(x_j) - min(x_j)} \tag{2}$$

where:

　　i—next vector index;
　　j—index of feature;
　　$max(x_j)$—the maximum value of the variable j;
　　$min(x_j)$—the minimal value of the variable j.

Test data also needs to be normalized. When normalizing the test set, one should use the maximum and minimum values determined on the training set. Figure 8 shows the decision surface for the kNN classifier for $k = 7$.
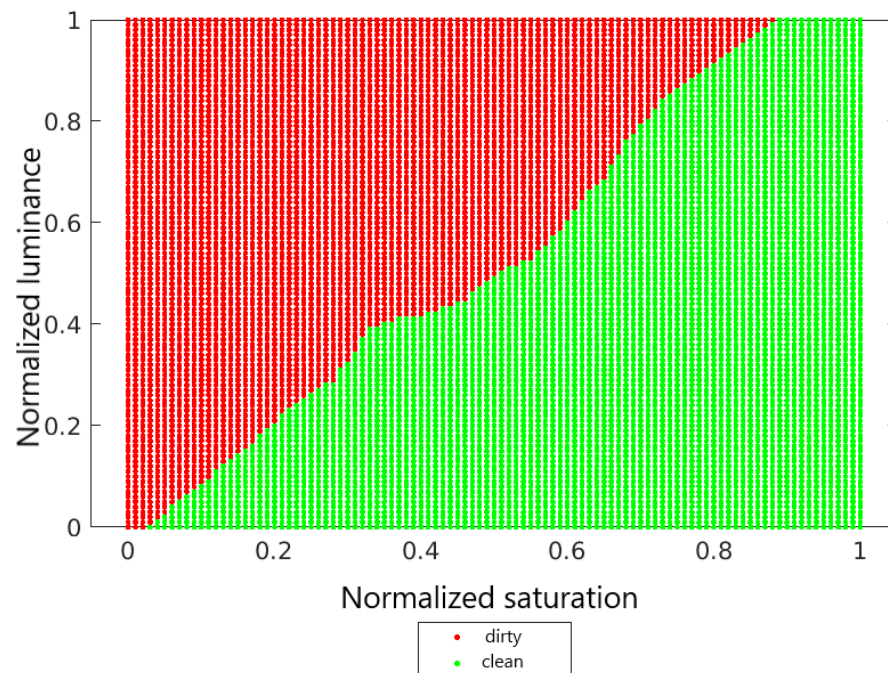

Figure 8. Decision area for the kNN classifier for $k = 7$.

3.4. Naive Bayesian Classifier

The naive Bayes classifier is a simple probabilistic classifier that assumes that all features are mutually independent, hence the so-called "naivety" of this classifier. It uses Bayes' theorem, and the classification result is based on a conditional probability comparison. The class for which the posterior probability value is the highest is selected [30].

$$P(Y_k|X) = \frac{P(X|Y_k)P(Y_k)}{P(X)} \tag{3}$$

where:

Y—vector of classes;
Y_k—class;
X—vector of features of classified object;
X_i—feature;
$P(Y_k)$—probability *a priori*;
$P(X|Y_k)$—probability of occurrence;
$P(Y_k|X)$—probability *a posteriori*;
$P(X)$—probability of occurrence of set of features.

Using the assumption of the classifier's naivety:

$$P(X|Y) = \prod_{i=1}^{n} P(X_i|Y) \tag{4}$$

where:

$P(X|Y_k)$—probability of occurrence;
$P(X_i|Y)$—conditional probability of occurrence of a given feature provided that a given class occurs.

The values of the features are continuous, so we assume that for each part X_i the distribution $P(X_i|Y_k)$ is a normal distribution:

$$P(X_i|Y_k) = N(\mu_{ik}, \sigma_{ik}^2) = \frac{1}{\sqrt{2\pi\sigma_{ik}^2}} e^{-\frac{1}{2\sigma_{ik}^2}(X_i - \mu_{ik})^2} \tag{5}$$

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{m} 1_Y(Y^j = Y_k)} \sum_{j=1}^{m} X_i^j 1_Y(Y^j = Y_k) \tag{6}$$

$$\sigma_{ik}^2 = \frac{1}{\sum_{j=1}^{m} 1_Y(Y^j = Y_k)} \sum_{j=1}^{m} (X_i^j - \mu_{ik}^2) 1_Y(Y^j = Y_k) \tag{7}$$

where:

1_Y—indicator function;
σ_{ik}^2—variance;
μ_{ik}—average value.

In the case under consideration, we have 2 features and 2 classes:

$$X = [X_{saturation}, X_{luminance}] \tag{8}$$

$$Y = [Y_{clean}, Y_{dirty}] \tag{9}$$

therefore:

$$P(Y_{clean}|X) = \frac{P(Y_{clean}) * P(X_{saturation}|Y_{clean}) * P(X_{luminance}|Y_{clean})}{P(X)} \tag{10}$$

$$P(Y_{dirty}|X) = \frac{P(Y_{dirty}) * P(X_{saturation}|Y_{dirty}) * P(X_{luminance}|Y_{dirty})}{P(X)} \quad (11)$$

To compare $P(Y_{clean}|X)$ and $P(Y_{dirty}|X)$, it is not necessary to calculate $P(X)$ because this value is constant and only serves as a scaling function. This approach reduces the computational effort of the classifier. The classifier's decision rule is as follows:

$$Y_k = \begin{cases} clean & \text{if } P(Y_{clean}|X) \geqslant P(Y_{dirty}|X) \\ dirty & \text{otherwise} \end{cases} \quad (12)$$

Figure 9 shows the decision surface of the naive Bayes classifier.
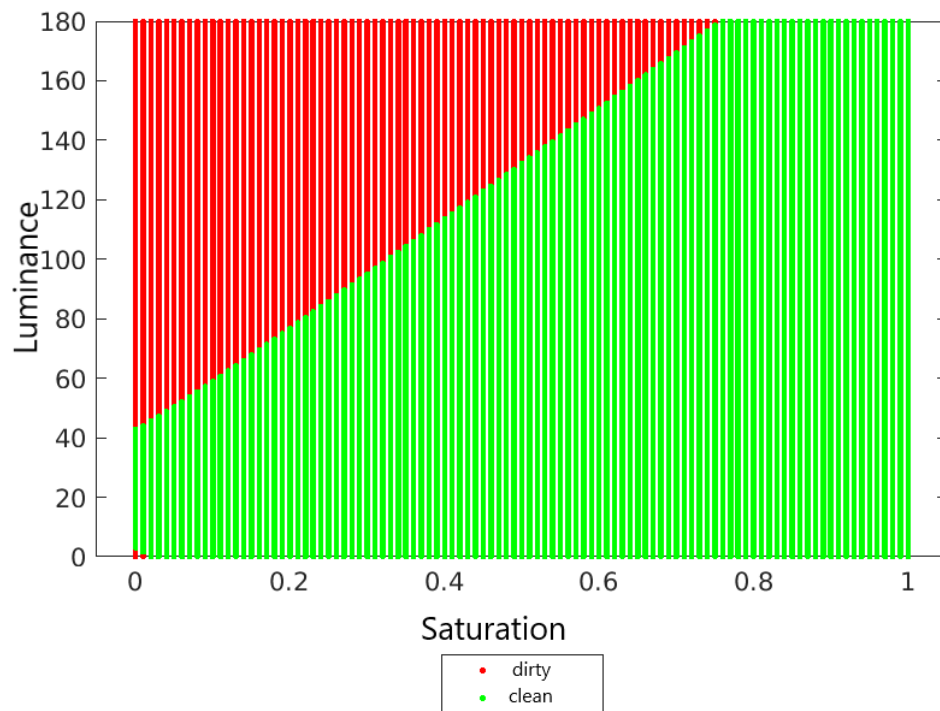

Figure 9. Decision surface for the naive Bayesian classifier.

3.5. Fisher's Linear Discriminator

The Fisher Linear Discriminator (FLD) is used for supervised classification and produces a linear discriminant rule. The task of discriminant analysis for two classes can be defined as [31]:

find the direction *a*, hat best separates the learning subgroups, and as a measure of class separation along a given direction *a* take the square of the distance between the arithmetic means of the subgroups along this direction, taking into account the variability of the intra-group observation.

$$\frac{(a'\overline{x}_2 - a'\overline{x}_1)^2}{a'Wa} \quad (13)$$

The direction of *a* best separating the classes is the direction that maximizes the expression (13):

$$argmax_a \frac{(a'\overline{x}_2 - a'\overline{x}_1)^2}{a'Wa} \quad (14)$$

The solution is:

$$a = W^{-1}(\overline{x}_2 - \overline{x}_1) \tag{15}$$

where:

W—intragroup covariance matrix;

a—direction vector of the searched line;

$\overline{x}_1$—group mean of observations included in the class *clean*;

$\overline{x}_2$—group mean of observations included in the class *dirty*;

n_1—number of observations included in the class *clean*;

n_2—number of observations included in the class *dirty*.

The observations can be divided into a subgroup of observations classified as class *clean* and into a subgroup of observations classified as class *dirty*:

$$x_{11}, x_{12}, \ldots, x_{1n_1} \text{observations from class } clean$$
$$x_{21}, x_{22}, \ldots, x_{2n_2} \text{observations from class } dirty$$

Then, we can write the group averages as:

$$\overline{x}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} x_{ki} \text{ dla } k = 1,2. \tag{16}$$

In order to assess the intragroup variability of the covariance matrix, it is necessary to assume that both subgroups have the same covariance matrix, then:

$$W = \frac{1}{n-2} \sum_{k=1}^{2} (n_k - 1) S_k = \frac{1}{n-2} \sum_{k=1}^{2} \sum_{l=1}^{n_k} (x_{kl} - \overline{x}_k)(x_{kl} - \overline{x}_k)' \tag{17}$$

where:

$n = n_1 + n_2$

S_k—sample covariance matrix of subgroup k;

x'—vector transposition x.

Having the designated direction a, both means of classes $\overline{x}_1$ i $\overline{x}_2$ and the new observation x, we can define the classification rule:

$$X = \begin{cases} clean & \text{if } |a'x - a'\overline{x}_1| < |a'x - a'\overline{x}_2| \\ dirty & \text{otherwise} \end{cases} \tag{18}$$

where:

x—new observation;

X—the class assigned to the new observation.

What after qualifying the boundary case to class *clean* comes down to the following decision rule:

$$X = \begin{cases} clean & \text{if } (\overline{x}_2 - \overline{x}_1)'W^{-1}[x - \frac{1}{2}(\overline{x}_1 + \overline{x}_2)] \geqslant 0 \\ dirty & \text{otherwise} \end{cases} \tag{19}$$

Figure 10 shows the discriminant line. It is a straight line perpendicular to the line a and passing through the middle of the line connecting points $a\overline{x}_1$ and $a\overline{x}_2$. The discriminant line equation is as follows:

$$x_L = 170,5716401240x_N + 44,975372529300 \tag{20}$$

where:

x_N—value of saturation;

x_L—value of luminance.

Figure 11 shows the designated decision areas.
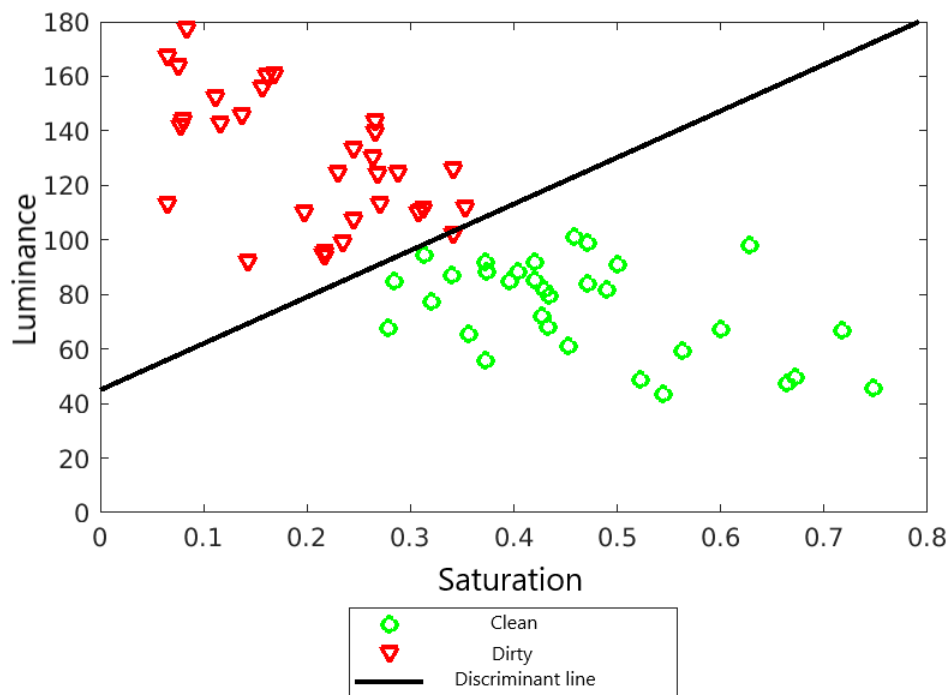
Figure 10. Decision boundary between classes.
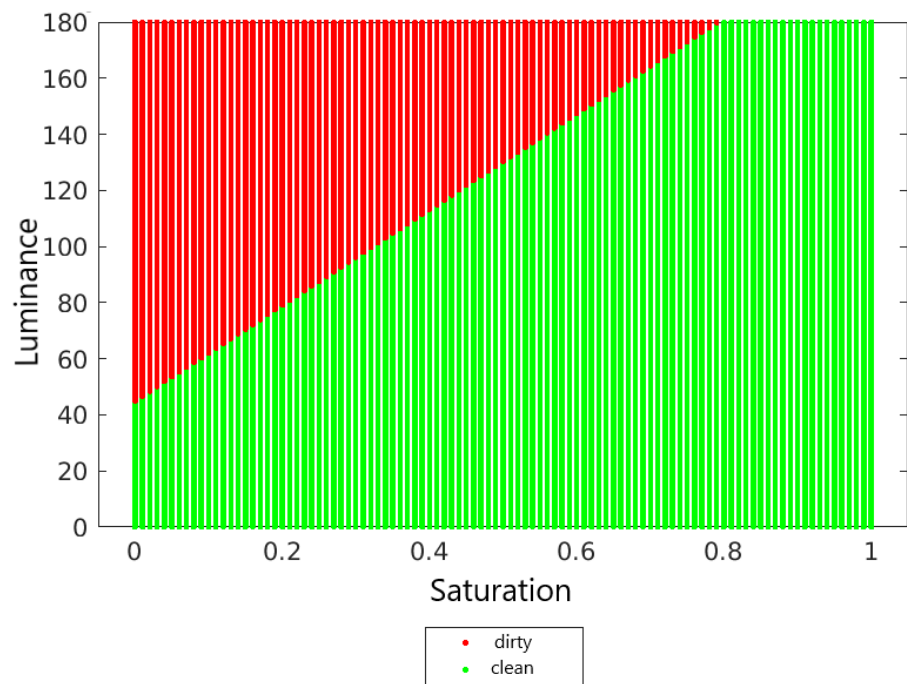
Figure 11. Decision surface for the linear Fischer discriminant.

4. Analysis and Research

The aim of the work was to test the quality and effectiveness of the proposed classifiers. For this purpose, the classic metrics of binary classification were used [32]. The decision surfaces of classifiers were also analyzed. In the case of the kNN classifier, its various variants were analyzed to select the optimal value of the *k* parameter.

4.1. kNN Classifier

TThe kNN classifier is characterized by the fact that for a new observation, the class that is most frequently represented among the k closest neighbors from the training set is selected. The correct performance of this classifier depends on the number of neighbors. The problem is the difficulty of selecting a priori the appropriate value for the k parameter, so different values of this parameter were analyzed. Only odd values of k have been considered, as such values guarantee that there will be no draw situation.

$$k \in \{1, 3, 5, 7, 9, 11, 13\} \tag{21}$$

Table 2 presents the results of the classification for different values of the k parameter. Based on these results, classic binary classification metrics have been developed, which can be found in the Table 3. The test set consisted of 12 clean panels and 12 dirty panels.

It can be concluded that the optimal value of the k parameter is 7. For this value, the kNN classifier is characterized by the highest sensitivity, specificity and precision. Additionally, the F1 metric reaches the highest value for $k = 7$, which can be seen in the Figure 12.

Table 2. Classification results for different values of parameter k.

k	P	N	TP	TN	FP	FN
1	12	12	10	8	4	2
3	12	12	10	10	2	2
5	12	12	11	10	2	1
7	12	12	11	11	1	1
9	12	12	10	11	1	2
11	12	12	10	10	2	2
13	12	12	9	8	3	2

Table 3. Classic binary classification metrics for different values of parameter k.

k	TPR	TNR	PPV	NPV	F1
1	83%	67%	71%	80%	77%
3	83%	83%	83%	83%	83%
5	92%	83%	85%	91%	88%
7	92%	92%	92%	92%	92%
9	83%	92%	91%	85%	87%
11	83%	83%	83%	83%	83%
13	75%	67%	69%	73%	72%

4.2. Decision Surfaces

Figure 13 shows the decision surfaces of the classification of the 7NN classifier, the naive Bayes classifier and the linear Fischer discriminant. In order to compare these surfaces, the luminance and saturation values have been normalized so that, for each classifier, the values are determined from the same selected range.

Comparing the decision surfaces of the naive Bayes classifier and the linear Fischer discriminant, it can be concluded that these surfaces are similar to each other. Although the naive Bayes classifier is not a linear classifier, for this classifier, the boundary between two classes is similar to a straight line and resembles the boundary of a linear Fischer

discriminant. In the case of the kNN classifier, its decision boundary is more wavy than for the other classifiers.
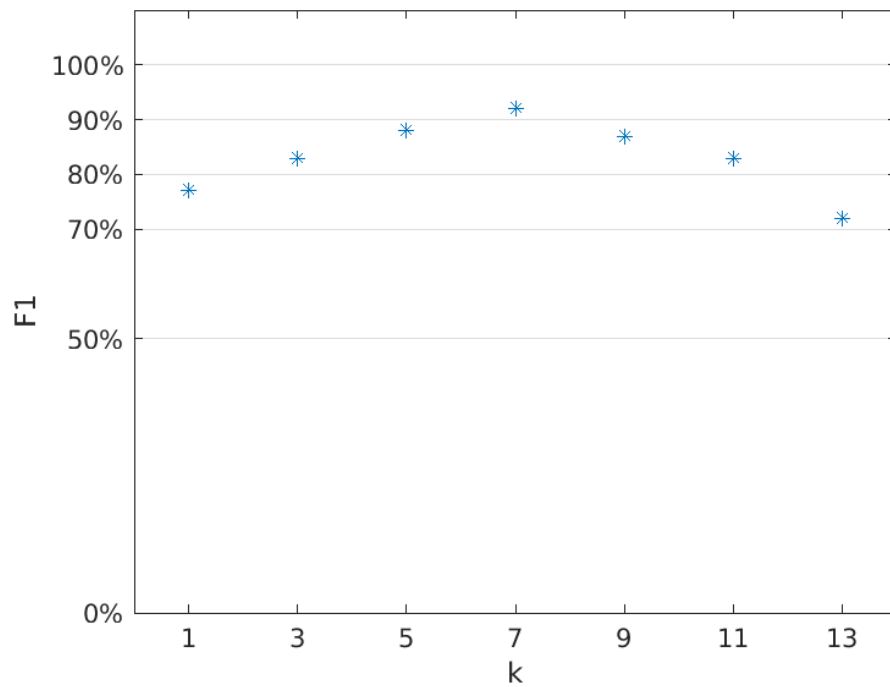

Figure 12. The value of the F1 metric depending on the parameter *k*.

4.3. Metrics and Results

To test the quality and effectiveness of the proposed classifiers, traditional binary classification metrics were used. The following classic binary classification metrics were used to assess the correctness of the type: TPR, TNR, PPV, NPV, and F1 [33]. Table 4 presents the results of panel classification into classes clean and dirty. The individual lines contain the results for each of the tested classifiers. The columns provide information on the number of clean and dirty panel samples tested, the number of true positives and true negatives detected, and the number of false positives and false negatives seen.
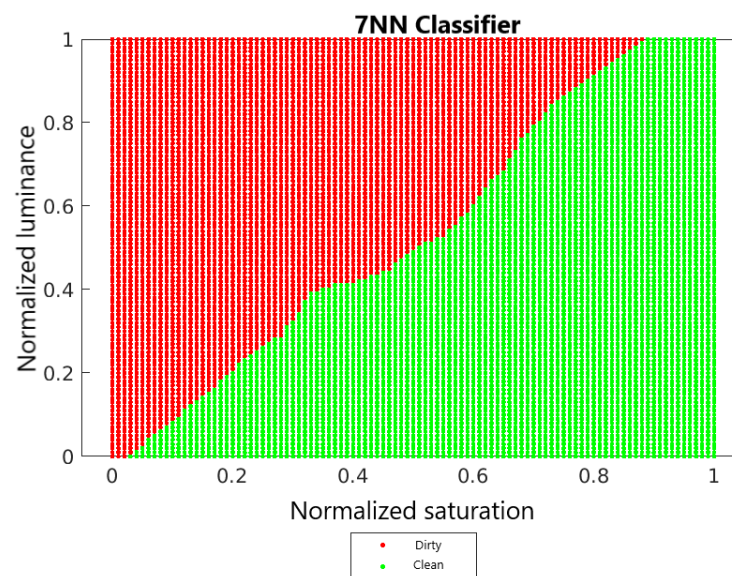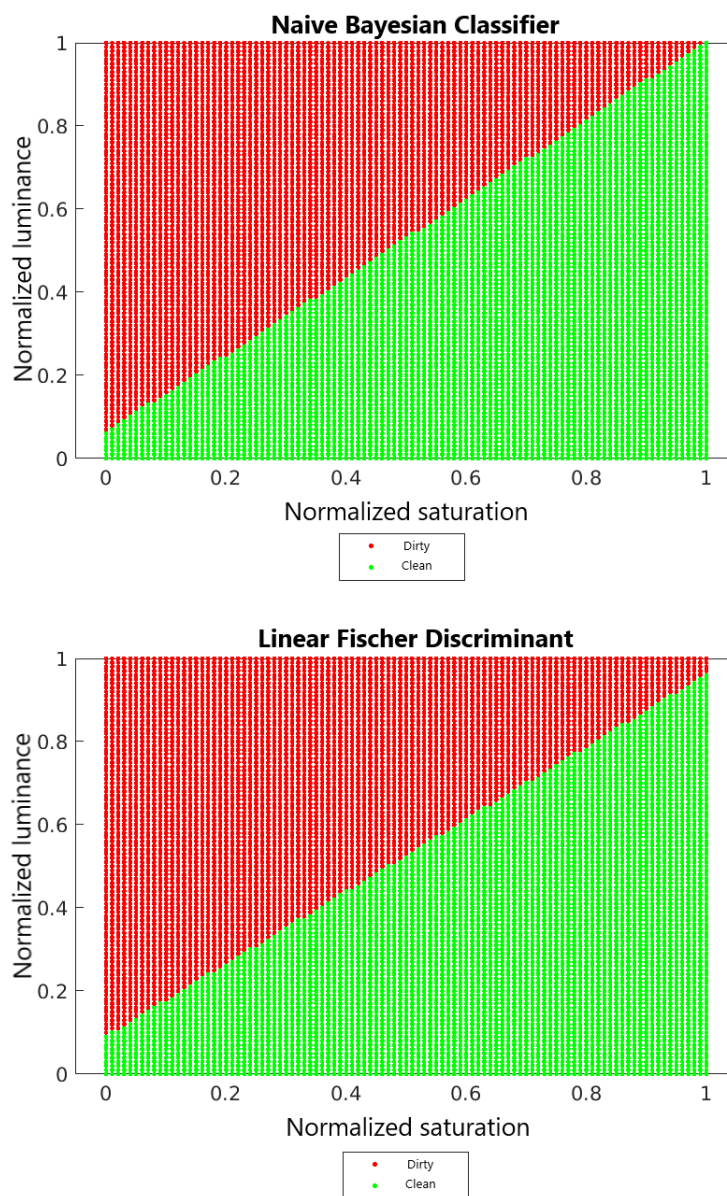

Figure 13. *Cont.*

Figure 13. Decision surfaces of each classifier.

Table 4. Classification results.

| Classifier	P	N	TP	TN	FP	FN
The classifier of the *k* nearest neighbors	12	12	10	11	1	2
Naive Bayesian classifier	12	12	11	11	1	1
Fisher's linear discriminator	12	12	11	10	2	1

Based on the results from Table 4, the classic binary classification metrics have been developed, which are presented in Table 5. The lines of this table present the metrics of each classifier, and the individual columns contain information about TPR, TNR, PPV, NPV, and F1.

Table 5. Classic metrics for binary classification.

Classifier	TPR	TNR	PPV	NPV	F1
Classifier 7NN	83%	92%	91%	85%	87%
Naive Bayesian classifier	92%	92%	92%	92%	92%
Fisher's linear discriminator	92%	83%	85%	91%	88%

Based on the research, it can be concluded that the Naive Bayesian classifier is characterized by the highest efficiency of detection of contaminated panels. All metrics for this classifier are the highest and amounted to 92%. For the 7NN classifier, the TPR, PPV, NPV, and F1 values are lower than for the Naive Bayes classifier, and the TNR value is the same. As for the linear Fischer discriminator, the TNR, PPV, NPV, and F1 values are lower than for the Naive Bayes classifier, and the TPR value is the same. Figure 14 shows.

The conducted research shows that the naive Bayes classifier is the optimal classifier for a given problem. A very high sensitivity characterizes the Bayes classifier. This means that it identifies clean panels with high efficiency. It also identifies dirty panels with high efficiency, as the specificity of this classifier is also very high. This classifier is also very precise. As for the Fischer discriminator, it is characterized by high sensitivity but low specificity, which means that it identifies clean panels more effectively than dirty panels. It is also exact in detecting clean panels, but has low precision in detecting dirty panels. In the case of the 7NN classifier, the situation is the opposite. This classifier has a high specificity and low sensitivity, identifying clean panels more efficiently than dirty panels. It is characterized by low precision in detecting clean panels and high precision in the detection of dirty panels.
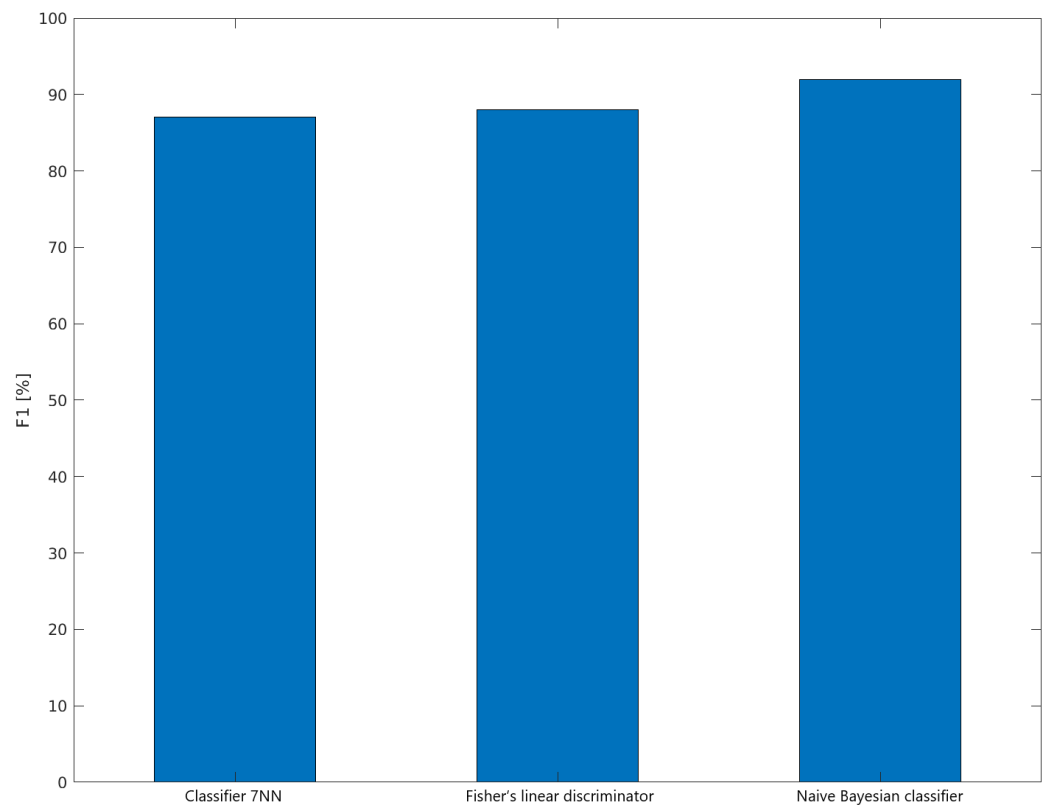
Figure 14. Value of the F1 metric for each of the classifiers.

5. Summary

Monitoring the cleanliness of photovoltaic panels is very important. In the first three years, the drop in efficiency can be as much as 15%. In places with high industrialization or dusty environments, the reductions in inefficiency are even more significant. Scientific research conducted by H. Haberlin and C. Renken from Berne University of Applied Sciences show that regular cleaning of PV modules improves their efficiency by up to 13.8% [34]. This solution is in line with the global trend of optimizing the use of photovoltaic panels. The results obtained in this study showed experimentally that the proposed solution provides a high rate of correct detections. The proposed innovative method is cheap and straightforward to implement, which allows it to be used in most photovoltaic installations and is suitable for use in an intelligent system for monitoring the cleanliness of photovoltaic panels. The presented methods of classifying the cleanliness of photovoltaic panels work well in areas with the highest concentration of dust and pollution. These are mostly suburban areas, proximity to highways, industrial plants, areas with a strong dusting of plants. There is no restriction in use. You can use photo material from a photovoltaic farm drone and photos with an appropriate resolution of home panels, e.g., on the roof, taken with a camera or telephone.

Author Contributions: Conceptualization, T.C. and K.B.; Methodology, T.C. and K.B.; Software, K.B.; Validation, T.C. and K.B.; Formal analysis, T.C.; Writing—original draft preparation, T.C. and K.B.; Writing—review and editing, T.C. and K.B.; Visualization, K.B.; Supervision, T.C.; Project administration, T.C. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

Abbreviations

The following abbreviations are used in this manuscript:

PV	photovoltaics
EVA	ethylene-vinyl acetate
TPR	true positive rate
TNR	true negative rate
PPV	positive predictive value
NPV	negative predictive value
P	positive
N	negative
TP	true positive
TN	true negative
FP	false positive
FN	false negative
RES	renewable energy sources
EU	European Union

References

1. GISTEMP Team. GISS Surface Temperature Analysis (GISTEMP), Version 4. NASA Goddard Institute for Space Studies. 2021. Available online: Available online: https://data.giss.nasa.gov/gistemp/ (accessed on 8 May 2012).
2. Lenssen, N.; Schmidt, G.; Hansen, J.; Menne, M.; Persin, A.; Ruedy, R.; Zyss, D. Improvements in the GISTEMP uncertainty model. *J. Geophys. Res. Atmos.* **2019**, *124*, 6307–6326. [CrossRef]

3. IPCC. *Global Warming of 1.5 °C. An IPCC Special Report on the Impacts of Global Warming of 1.5 °C above Pre-Industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty*; Masson-Delmotte, V., Zhai, P., Pörtner, H.O., Roberts, D., Skea, J., Shukla, P.R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Eds.; IPCC, 2019. In Press. Available online: https://www.ipcc.ch/site/assets/uploads/sites/2/2019/05/SR15_SPM_version_report_LR.pdf (accessed on 8 May 2012).

4. Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions the European Green Deal, Brussels. 11 December 2019, COM/2019/640 Final. Available online: https://www.eumonitor.eu/9353000/1/j4nvke1fm2yd1u0_j9vvik7m1c3gyxp/vl4cn7a3nez7/v=s7z/f=/com(2019)640_en.pdf (accessed on 8 May 2012).

5. Proposal for a Regulation of the European Parliament and of the Council Establishing the Framework for Achieving Climate Neutrality and Amending REGULATION (EU) 2018/1999 (European Climate Law), Brussels, 4 March 2020. COM/2020/80 Final. Available online: https://ec.europa.eu/info/sites/default/files/commission-proposal-regulation-european-climate-law-march-2020_en.pdf (accessed on 8 May 2012).

6. Amended Proposal for a Regulation of the European Parliament and of the Council on Establishing the Framework for Achieving Climate Neutrality and Amending REGULATION (EU) 2018/1999 (European Climate Law, Brussels, 17 September 2020. COM(2020) 563 Final, 2020/0036 (COD). Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0563&rid=9 (accessed on 8 May 2012).

7. What Is Green Power? Available online: https://19january2017snapshot.epa.gov/greenpower/what-green-power_.html (accessed on 8 May 2012).

8. Moser, D.; Del Buono, M.; Jahn, U.; Herz, M.; Richter, M.; De Brabandere, K. Identification of technical risks in the photovoltaic value chain and quantification of the economic impact: Economic impact of technical risks in photovoltaic plants. *Prog. Photovoltaics Res. Appl.* **2017**, *25*, 592–604. [CrossRef]

9. Ministry of National Assets. Poland's National Energy and Climate Plan for the Years 2021–2030; Version 4.1; Ministry of National Assets, 2019. Available online: https://www.gov.pl/attachment/e64830c9-440f-4f17-b3e7-8abd7667d406 (accessed on 8 May 2012).

10. David Moser, Giorgio Belluardo, Matteo Del Buono, Walter Bresciani, Elisa Veronese (EURAC) Ulrike Jahn, Magnus Herz, Eckart Janknecht, Erin Ndrio (TUV-RH) Karel de Brabandere, Mauricio Richter (3E); 1/3/2016. Available online: https://www.tuv.com/content-media-files/master-content/services/products/p06-solar/solar-downloadpage/solar-bankability_d1.1_d2.1_technical-risks-in-pv-projects.pdf (accessed on 8 May 2012).

11. John, J.J.; Tatapudi, S.; Tamizhmani, G. Influence of soiling layer on quantum efficiency and spectral reflectance on crystalline silicon PV modules. In Proceedings of the 2014 IEEE 40th Photovoltaic Specialist Conference (PVSC), Denver, Colorado, USA, 8–13 June 2014; pp. 2595–2599. [CrossRef]

12. Maghami, M.R.; Hizam, H.; Gomes, C.; Radzi, M.A.; Rezadad, M.I.; Hajighorbani, S. Power loss due to soiling on solar panel: A review. *Renew. Sustain. Energy Rev.* **2016**, *59*, 1307–1316. [CrossRef]

13. Cai, S.; Bao, G.; Ma, X.; Wu, W.; Bian, G.B.; Rodrigues, J.J.; de Albuquerque, V.H.C. Parameters optimization of the dust absorbing structure for photovoltaic panel cleaning robot based on orthogonal experiment method. *J. Clean. Prod.* **2019**, *217*, 724–731. [CrossRef]

14. Bessa, J.G.; Micheli, L.; Almonacid, F.; Fernández, E.F. Monitoring photovoltaic soiling: Assessment, challenges, and perspectives of current and potential strategies. *iScience* **2021**, *24*, 102165. [CrossRef] [PubMed]

15. Deb, D.; Brahmbhatt, N.L. Review of yield increase of solar panels through soiling prevention, and a proposed water-free automated cleaning solution. *Renew. Sustain. Energy Rev.* **2018**, *82* Part 3, 3306–3313. [CrossRef]

16. Jiang, Y.; Lu, L.; Ferro, A.R.; Ahmadi, G., Analyzing wind cleaning process on the accumulated dust on solar photovoltaic (PV) modules on flat surfaces. *Sol. Energy* **2018**, *159*, 1031–1036. [CrossRef]

17. Chanchangi, Y.N.; Ghosh, A.; Sundaram, S.; Mallick, T.K. An analytical indoor experimental study on the effect of soiling on PV, focusing on dust properties and PV surface material. *Sol. Energy* **2020**, *203*, 46–68. [CrossRef]

18. Kim, D.; Youn, J.; Kim, C. Automatic Photovoltaic Panel Area Extraction from UAV Thermal Infrared Images. *J. Korean Soc. Surv. Geod. Photogramm. Cartogr.* **2016**, *34*, 559–568. [CrossRef]

19. Ilse, K.; Micheli, L.; Figgis, B.W.; Lange, K.; Daßler, D.; Hanifi, H.; Wolfertstetter, F.; Naumann, V.; Hagendorf, C.; Gottschalg, R.; et al. Techno-Economic Assessment of Soiling Losses and Mitigation Strategies for Solar Power Generation. *Joule* **2019**, *3*, 2303–2321. [CrossRef]

20. Sovetkin, E.; Steland, A. Automatic processing and solar cell detection in photovoltaic electroluminescence images. *Integr. Comput.-Aided Eng.* **2018**, *26*, 1–15. [CrossRef]

21. Felipe-García, B.; Hernández-López, D.; Lerma, J.L. Analysis of the ground sample distance on large photogrammetric surveys. *Appl. Geomat.* **2012**, *4*, 231–244. [CrossRef]

22. Seifert, E.; Seifert, S.; Vogt, H.; Drew, D.; Van Aardt, J.; Kunneke, A.; Seifert, T. Influence of Drone Altitude, Image Overlap, and Optical Sensor Resolution on Multi-View Reconstruction of Forest Images. *Remote Sens.* **2019**, *11*, 1252. [CrossRef]

23. Jiang, H.; Yao, L.; Lu, N.; Qin, J.; Liu, T.; Liu, Y.; Zhou, C. Multi-resolution dataset for photovoltaic panel segmentation from satellite and aerial imagery. *Earth Syst. Sci. Data* **2021**, *13*, 5389–5401. [CrossRef]

24. Hwang, Y.S.; Schlüter, S.; Park, S.I.; Um, J.S. Comparative Evaluation of Mapping Accuracy between UAV Video versus Photo Mosaic for the Scattered Urban Photovoltaic Panel. *Remote Sens.* **2021**, *13*, 2745. [CrossRef]

25.   Dharampal, V.M. Methods of Image Edge Detection: A Review. *J. Electr. Electron. Syst.* **2015**, *4*, 2. [CrossRef]

26.   Cao, D.H.; Stoumpos, C.C.; Farha, O.K.; Hupp, J.T.; Kanatzidis, M.G. 2D Homologous Perovskites as Light-Absorbing Materials for Solar Cell Applications. *J. Am. Chem. Soc.* **2015**, *137*, 7843–7850. [CrossRef] [PubMed]

27.   BT.601: Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide Screen 16:9 Aspect Ratios. Available online: https://www.itu.int/rec/R-REC-BT.601 (accessed on 12 October 2021).

28.   Cheko, B. *Introduction to Scene Text and Object Detection and Recognition Understanding A Scene Image*; Scholars' Press: Chisinau, Moldo, 2020; ISBN-13: 978-613-8-93317-5.

29.   Cunningham, P.; Delany, S. k-Nearest neighbour classifiers. *Mult. Classif. Syst.* **2007**, *54*. [CrossRef]

30.   Bayindir, R.; Yesilbudak, M.; Colak, M.; Genc, N. A Novel Application of Naive Bayes Classifier in Photovoltaic Energy Prediction. In Proceedings of the 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), Cancun, Mexico, 18–21 December 2017; pp. 523–527. [CrossRef]

31.   Mello, R.F. *Moacir Antonelli Ponti, Machine Learning: A Practical Approach on the Statistical Learning Theory*, 1st ed.; Springer: Berlin/Heidelberg, Germany, 2018; ISBN-13: 978-3319949888.

32.   Kohl, M. Performance Measures in Binary Classification. *Int. J. Stat. Med. Res.* **2012**, *1*, 79–81. [CrossRef]

33.   Canbek, G.; Temizel, T.T.; Sagiroglu, S.; Baykal, N. Binary Classification Performance Measures/Metrics: A Comprehensive Visualized Roadmap to Gain New Insights. In Proceedings of the 2017 International Conference on Computer Science and Engineering (UBMK), Antalya, Turkey, 5–8 October 2017; pp. 821–826. [CrossRef]

34.   Haeberlin, H.; Schaerf, P. Long-term Behaviour of Grid-Connected PV Systems over more than 15 Years. In Proceedings of the 5th European Photovoltaic Solar Energy Conference/5th World Conference on Photovoltaic Energy Conversion, Valencia, Spain, 6–10 September 2010. Available online: http://www.pvtest.ch/Dokumente/Publikationen/144_LZPV4-Valencia-K_F_02.pdf (accessed on 8 May 2012).