Article

Fine Land Cover Classification in an Open Pit Mining Area Using Optimized Support Vector Machine and WorldView-3 Imagery

Weitao Chen [1,2] **, Xianju Li** [1,2,*] **and Lizhe Wang** [1,2]

[1]   Faculty of Computer Science, China University of Geosciences, Wuhan 430074, China;
      wtchen@cug.edu.cn (W.C.); LZWang@cug.edu.cn (L.W.)

[2]   Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences,
      Wuhan 430074, China

*   Correspondence: ddwhlxj@cug.edu.cn; Tel.: +86-27-67883716

Abstract: Fine land cover classification in an open pit mining area (LCCOM) is essential in analyzing the terrestrial environment. However, researchers have been focusing on obtaining coarse LCCOM while using high spatial resolution remote sensing data and machine learning algorithms. Although support vector machines (SVM) have been successfully used in the remote sensing community, achieving a high classification accuracy of fine LCCOM using SVM remains difficult because of two factors. One is the lack of significant features for efficiently describing unique terrestrial characteristics of open pit mining areas and another is the lack of an optimized strategy to obtain suitable SVM parameters. This study attempted to address these two issues. Firstly, a novel carbonate index that was based on WorldView-3 was proposed and introduced into the used feature set. Additionally, three optimization methods—genetic algorithm (GA), k-fold cross validation (CV), and particle swarm optimization (PSO)—were used for obtaining the optimization parameters of SVM. The results show that the carbonate index was effective for distinguishing the dumping ground from other open pit mining lands. Furthermore, the three optimization methods could significantly increase the overall classification accuracy (OA) of the fine LCCOM by 8.40%. CV significantly outperformed GA and PSO, and GA performed slightly better than PSO. CV was more suitable for most of the fine land cover types of crop land, and PSO for road and open pit mining lands. The results of an independent test set revealed that the optimized SVM models achieved significant improvements, with an average of 8.29%. Overall, the proposed strategy was effective for fine LCCOM.

Keywords: open pit mining; land degradation; support vector machine (SVM); worldView-3; remote sensing

1. Introduction

Land degradation has been increasingly recognized as one of the most destructive impacts on the terrestrial environmental during the last century [1,2]. Some researchers have revealed the important effect of open pit mining on local land degradation [1–5]. Accordingly, land covers in complex open pit mining landscapes are being increasingly used as key datasets for global and local land degradation and development studies [6–12].

Currently, high resolution satellite imagery and machine learning algorithms (MLAs) have been applied to land cover classification in open pit mining areas [6,9,12–14]. MLAs can generally accept various features sets [10], which have proven to be valuable in open pit mining areas classification. Several algorithms with excellent performance have been widely used, for example, support vector

machine (SVM) [15,16], artificial neural network (ANN) [17], and random forest (RF) [15]. These three algorithms were non-parametric learning methods and no assumptions need to be made for the form of the target function. Therefore, they are becoming increasingly popular in remotely sensed classification [18].

Some researches show that SVM can successfully achieve higher classification accuracy than traditional methods [9,19–24]. Nevertheless, two challenges for SVM remain to be addressed in the remote sensing field [15,16]: (1) to select the kernel function and parameter [25]; and, (2) to optimally obtain key parameters, including kernels and penalty parameter C [22,26–28]. Our previous research show that it is challenging to achieve higher accuracy for fine land cover classification in an open pit mining area (LCCOM), because of four specific features: significant three-dimensional terrain, strong temporal-spatial variability of surface cover, spectral-spatial homogeneity of mining areas [7], and different phenological stages [9]. From this viewpoint, it is difficult to use the original SVM to improve the classification accuracy of fine LCCOM. Thus, several optimum methods have been used to determine the optimized key parameters of SVM, including several evolutionary algorithms and swarm intelligence algorithms [29,30], such as the genetic algorithm method (GA), cross-validation method (CV), and particle swarm optimization method (PSO).

It was revealed in our previous research that it is essential to obtain beneficial features from new remote sensing sensors to improve the classification accuracy in open pit mining areas [6,7,9]. WorldView-3 (WV-3) has higher spatial and spectral resolution [31], which might be further advantageous in vegetation, hydrocarbons [31], and lithology mapping. Thus, we believe that WV-3 has high potential in improving the accuracy of fine LCCOM [6], but is has not still been confirmed.

The objectives of this study are as follows: (1) to validate the ability of WV-3 in achieving higher classification accuracy of fine LCCOM; and, (2) to determine an optimized SVM method for improving the classification accuracy of fine LCCOM.

2. Materials and Methods

2.1. Study Area

An open pit mining district with an area of 65.16 km^2 located in central China (114°13'53.375"–114°20'35.538"E and 30°15'39.957"–30°18'48.352"N) was selected as the study area. Mineral resources relevant activities are very strong in the study area. The mine disturbance here with open pit mining has a history of nearly 70 years, and most of the mines are active nowadays, especially the Wulongquan mine. More detailed information regarding the study area can be found in the literature [6,7,9]. In this study, 23 sites representing different land cover types were selected as the field investigation points, as shown in Figure 1.
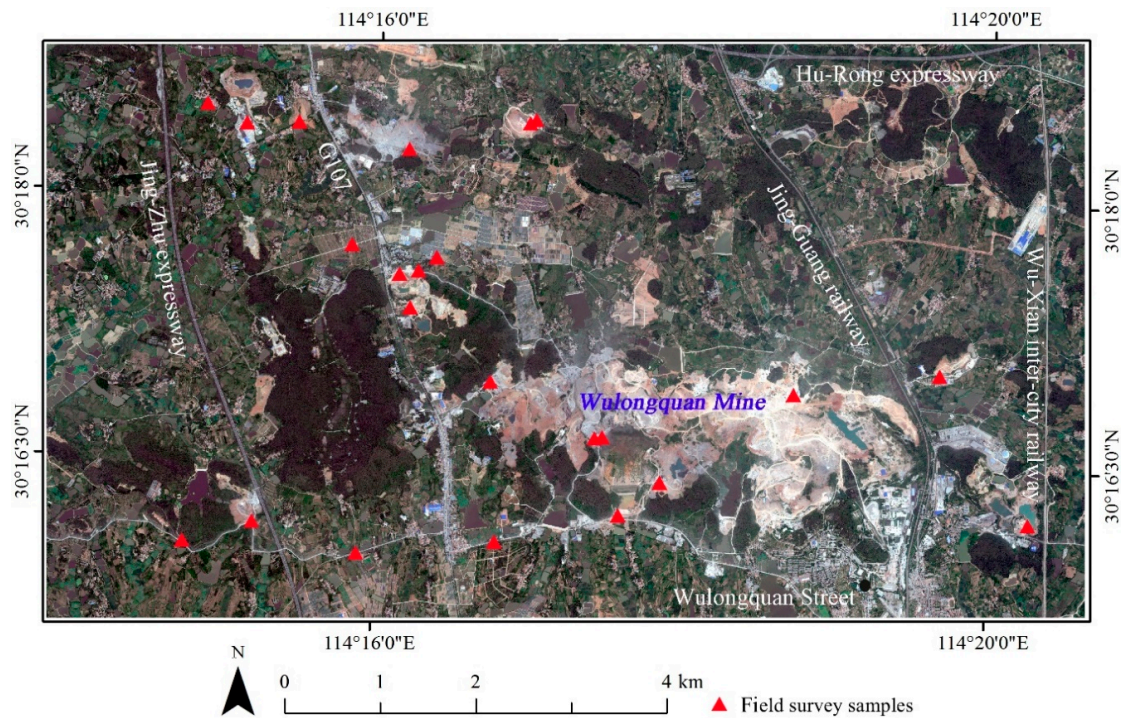
Figure 1. Study area and WV-3 fused true color image. RGB = band 5, 3, 2.

2.2. Remote Sensing Data Resources

WV-3 can provide a dataset of higher spatial and spectral resolution from panchromatic, multispectral imagery to short-wave infrared imagery (SWIR) (as listed in Table 1). The use of WV-3 imagery has been found to provide better performance on mapping surface geology targets than conventional sensors [32,33]. The 2A level WV-3 data that were used in this study were acquired on 4 August 2015 with the total cloud cover of 1%.

Table 1. Main specification of worldview-3 [34].

Sensor Bands	Band Name	Wavelength Range (nm)	Sensor Resolution (m)
	Panchromatic	450–800	0.31
Visible and near infrared (VNIR)	Coastal Blue	400–450	1.24
	Blue	450–510	
	Green	510–580	
	Yellow	585–625	
	Red	630–690	
	Red Edge	705–745	
	NIR-1	77–895	
	NIR-2	860–1040	
Short-wave infrared (SWIR)	SWIR-1	1195–1225	Commercial delivery at 7.5 m resolution
	SWIR-2	1550–1590	
	SWIR-3	1640–1680	
	SWIR-4	1710–1750	
	SWIR-5	2145–2185	
	SWIR-6	2185–2225	
	SWIR-7	2235–2285	
	SWIR-8	2285–2365	

3. Methods

The overall flowchart of the methodology can be divided into six phases, as presented in Figure 2. The high resolution features of WV-3 images and optimized SVM algorithm were used in this study to

improve the classification accuracy of the fine LCCOM. Besides, an independent test set was used to examine the classification models for LCCOM.
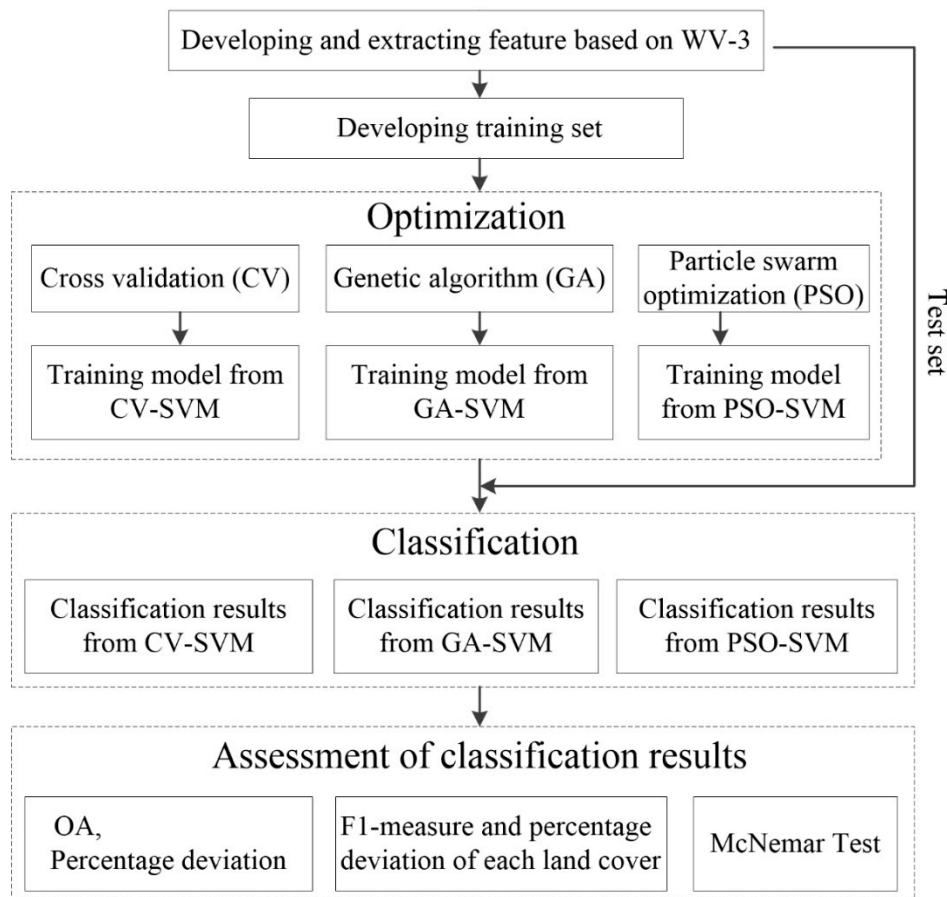

Figure 2. Overall flowchart of the methodology employed in this study.

3.1. Data Processing

Geometric registration between different bands of the used WV-3 level 2A products can meet the requirement of image fusion. According to our previous study [34], the panchromatic and multispectral data were fused based on the Gram-chmidt method because it provides higher fidelity of spatial and spectral characteristics, which is suitable for studying open pit mining areas.

3.2. Developing Features Based on WV-3

A total of 70 features divided into six types were used in this study (as listed in Table 2).

(1) Various spectral bands: eight VNIR spectral bands of the fused image and eight SWIR spectral bands.

(2) Vegetation index: one is the normalized difference vegetation index [35], which was calculated while using the first near-infrared band (NIR-1) of WV-3; the other is the soil-adjusted vegetation index [36], which was expected to improve the identification ability for bare soil and sparse vegetation in open pit areas. The formula for SAVI is as follows:

$$SAVI = [(NIR - red)/(NIR + red + L)] \times (1 + L), \tag{1}$$

According to statistic values, the value of L was determined with 0.5. Furthermore, the NIR-1 of WV-2 was used.

(3) Carbonate index (CI): a strong absorption feature at around 2.3 μm [37] characterizes carbonates. It is advantageous in distinguishing non-lithology cover types in mining areas, such as dumping grounds. Accordingly, the CI was developed based on WV-3 characteristics, as follows, in this study:

$$CI = SWIR\text{-}5/SWIR\text{-}3, \tag{2}$$

In formula (2), SWIR-5 and SWIR-3 are the fifth and third bands in the SWIR range, respectively.

(4) Principal component bands: the principal component analysis [38] was carried out to eliminate the redundancies for the eight fused multispectral bands in this study. The top three bands with a cumulative contribution rate of 99.34% were used.

(5) Filter images: it is helpful to improve the classification accuracy in LCCOM for applying Gaussian low-pass filters to optical images [9]. Thus, the eight fused multi-spectral bands were calculated and a kernel size of 3×3 was used.

(6) Texture measures: five features belonging to gray level co-occurrence matrix texture [39] measures were calculated with a processing window size of 9×9.

Table 2. Details of the feature set used in this study.

Image Features	Feature Names	Number
Various spectral bands	VNIR_(1, 2, … , 8), SWIR_(1, 2, … , 8)	16
Vegetation indices	Normalized difference vegetation index, Soil-adjusted vegetation index	2
Carbonate Index	CI	1
Principal component bands	P1, P2, P3	3
Filter images	Gaussian low-pass filter _VNIR_(1, 2, … , 8)	8
Texture measures	(Cor, Con, Asm, Hom, Ent)_VNIR_(1, 2, … , 8)	40
	Total	70

VNIR_(1, 2, … , 8): fused bands of eight VNIR bands; SWIR_(1, 2, … , 8): eight shortwave infrared bands; CI: carbonate Index; Cor, Con, Asm, Hom, Ent: correlation, contrast, angular second moment, homogeneity, entropy.

3.3. Land Cover Classification Schemes

As previously reported [6,9], there were seven first-level classes in the study area, i.e., open pit mining land, crop land, forest land, water, road, urban and rural residential land, and bare land. Fine classes (for details, see Table 3) should be considered in land cover classification schemes for open pit mining areas due to the big intra-class spectral and topographic differences in the first-level classes. Two new fine types were particularly added in this study (Table 3). One is green dry land with land-water resources for crops mainly from natural precipitation and with high coverage. The other is black roof, usually referring to residential land in industrial parks. All of the procedures, such as training and test set construction, SVM algorithm optimization, classification model developing and prediction, and accuracy assessment were carried out on the fine land cover classes based on the open pit mining area in this study.

Table 3. Land cover classification schemes used in this study (revised from [9]).

Fine Land Cover Types	Description
Opencast pit	Having mine pit lakes and spiral roads.
Ore processing site	Characterized by linear mineral processing facilities and highly reflective rubble.
Dumping ground	Located around stopes and may be gray in true color images.
Paddy field	Having adequate water supply and used for cultivation of rice, lotus, and other aquatic crops.

Table 3. *Cont.*

Fine Land Cover Types	Description
Vegetable and fruit greenhouse	Having white plastic film sides and roofs, and high surface albedo with regular rectangular shapes.
Green dry land	On the land water resources for crops mainly from natural precipitation and with high coverage.
Gray dry land	On the land water resources for crops mainly from natural precipitation and with low coverage.
Fallow land	No crops growing at the present stage.
Woodland	Includes timber stands, economic forests, and shelterbelts that have high chlorophyll content and are dark red in the false color image (R—NIR-1, G—Red, B—Green).
Shrub forest	Having multiple stems and shorter height, generally less than 2 m tall, and is bright red in false color images.
Forest under stress	Under the influence of surface mining development, around surface-mined land, having large amounts of deposited mineral dust, poor growth, and is grayish in true color images (R—Red, G—Green, B—Blue).
Nursery and orchard	Having a rectangular shape like cropland dotted by vegetation cover and exposed soil and is black in true color images.
Pond and stream	Including many fish ponds with regular rectangular shapes.
Mine pit pond	Lakes created during and after mining, typically with irregular shapes.
Dark road	Usually referring to asphalt highways.
Bright road	Usually referring to cement roads.
Light gray road	Usually referring to dirt roads.
Bright roof	Usually referring to urban and town areas.
Red roof	Usually referring to rural land.
Dark roof	Usually referring to residential land in industrial parks
Blue roof	Usually referring to land used for industrial parks.
Bare surface	Referring to exposed land with little vegetation.

3.4. Training Set and Test Sets

Based on our training data polygons [9], a revised version of polygons was obtained (Table 4) based on WV-3 images and the above-mentioned updated fine land cover classification scheme. Almost all of the open pit mining land was delineated in a similar manner, and the others were randomly determined across the study area. We employed a stratified random sampling method in this study. The result is that each class included 1000 samples, in which 900 samples were considered as the training set, while the other 100 were considered as the test set (Table 4). This study used spatially dependent training and test sets for fine LCCOM, with reference to the literature [6,9].

Table 4. Description of the Training Set and Test Set Used in This Study.

Fine Land Cover Types	Number of DPs	Area of DPs (km^2)	Number of Pixels in DPs (10^3)	Fraction (‰)
Open pit	17	2.22	24,635	0.04
Ore processing site	36	0.92	10,256	0.10
Dumping ground	38	0.79	8724	0.11
Paddy	36	0.10	1085	0.92
greenhouse	13	0.05	565	1.77
Green dry land	66	0.18	2021	0.49
Gray dry land	30	0.06	626	1.60
Fallow land	42	0.09	1013	0.99

Table 4. *Cont.*

Fine Land Cover Types	Number of DPs	Area of DPs (km^2)	Number of Pixels in DPs (10^3)	Fraction (‰)
Woodland	71	0.75	8298	0.12
Shrubbery	43	0.12	1295	0.77
Coerced forest	28	0.11	1228	0.81
Nursery	41	0.21	2386	0.42
Pond and stream	99	0.47	5253	0.19
Mine pit pond	12	0.06	659	1.52
Dark road	9	0.04	399	2.50
Bright road	5	0.00	32	31.63
Light gray road	63	0.06	704	1.42
Bright roof	61	0.08	932	1.07
Red roof	128	0.02	198	5.06
Dark roof	81	0.03	344	2.91
Blue roof	33	0.05	503	1.99
Bare surface	6	0.01	136	7.34

DPs: data polygons; Fraction: number of pixels in the training and test sets divided by the number of pixels in DPs.

An independent test set with the first-level land covers (i.e., crop, forest, water, road, residential, bare surface, and open pit mining) [9] was used. In this set, there were 700 samples, i.e., 100 samples for each first-level land class.

3.5. Classification Algorithm and Corresponding Parameter Optimization Methods

The penalty parameter C of SVM is key to improving the remote sensing classification accuracy. The polynomial kernels and the radial basis function (RBF) kernel were often used in the remote sensing community [40]. The used kernel parameter G can influence the complexity of the sample feature subspace distribution [16]. The classification accuracy and generalization ability of SVM will decline as G increases. When the G value is small, almost all of the training samples are support vectors. At this time, the training error is small and the test error is close to 1. However, the SVM generalization ability is poor. As the G value increases, the number of support vectors gradually decreases and the training error increases. Nevertheless, the SVM generalization ability is gradually enhanced. With the continued increase, when G reaches a certain threshold, the number of support vectors will increase again, and both training errors and test errors will also increase. At this point, SVM will start to deteriorate, regardless of classification ability or generalization ability. In addition, the penalty parameter C could influence the generalization of SVM. After the sample subspace is determined, when the C value is small, the complexity of the SVM is small, the penalty for the empirical error is small, and the empirical risk value is large.

CV is a time-consuming process, as it is often necessary to evaluate each parameter set at many grid points [26,41]. GAs have been used to obtain SVM kernel parameters, which were integrated into SVM algorithms to improve the accuracy of SVM classification [26,42]. PSO is selected to determine the optimum kernel function, which is particularly effective for the radial basis function kernel [43].

In this study, the GA, CV, and PSO algorithms were selected to optimize G and C parameters of SVM. The CV process was implemented in R software. The GA and PSO algorithms were processed in MATLAB R2009.

3.5.1. k-fold CV Algorithm

The k-fold CV method has been widely used in the remote sensing community [15,44–46]. More details regarding the CV could be found in the literature [44]. In this study, a five-fold CV scheme in the function "best.tune" included in the e1071 package [47] was utilized to achieve the "optimal" parameter combination. The maximum mean accuracies resulting from the k-fold CV would be considered as the optimal parameter.

3.5.2. Genetic Algorithm

GA was considered to be an adaptive optimization method according to the genetic processes of biological organisms. Further details on GA can be found in the literature [26,48]. GA can simultaneously identify the optimal SVM kernel parameters without reducing the SVM classification accuracy.

3.5.3. PSO Algorithm

PSO was proposed based on the social behavior of bird flocking, and the detailed principle on PSO could be found in the literature [49,50]. PSO has the advantages of higher efficiency, uncomplicated implementation, and significant exploration abilities for parameter optimization in SVM, both globally and locally [44,51].

3.6. Accuracy Assessment

The classification accuracies of fine LCCOM were evaluated on the basis of the test set. The overall accuracy (OA) was used to indicate the performance of the optimization classification models. The F1-measure [51] was used to describe class-specific accuracy. Moreover, the percentage deviation [52] was calculated on the basis of the three above-mentioned metrics for evaluating differences in the overall performance and the accuracy of each class among different classification models. In addition, the McNemar test was used to examine whether the parameter optimization methods could significantly improve the classification accuracy and whether there is a significant difference between the three optimization methods.

4. Results

4.1. Results of Parameter Optimization

A group of 48 parameter combinations (i.e., eight and six values for gamma and cost; for details, see Table 5) was used for the parameter optimization of SVM algorithm-based classification models. The default values for gamma and cost were $1/n$ (n is data dimension) and 1. While using CV, GA, and PSO algorithms, combinations of 2^{-7} and 2^7, 2^{-9} and 2^7, and 2^{-5} and 2^3 for G and C were obtained, respectively.

Table 5. Parameter of optimization results.

Parameter	Ranges	Results		
		CV	GA	PSO
G	$2^{-11}, 2^{-9}, \ldots, 2^3$	2^{-7}	2^{-9}	2^{-5}
C	$2^{-3}, 2^{-1}, \ldots, 2^7$	2^7	2^7	2^3

4.2. Assessment of Classification Results

Four classifications were performed in this study. Table 6 presents the F1-measure, OA, and percentage deviation of different SVM models based on default parameters and different optimization methods. Table 7 presents the results of the McNemar statistical test.

Table 6. Classification accuracy of different methods (%).

Second Level Classes	SVM	CV	GA	PSO	CV-P	GA-P	PSO-P
Open pit	45.99	51.58	50.00	52.41	12.15	8.72	13.96
Ore processing site	46.59	51.43	52.46	54.14	10.39	12.60	16.21
Dumping ground	65.95	71.70	69.61	71.22	8.72	5.55	7.99
Paddy	84.38	88.32	87.96	88.78	4.67	4.24	5.21
greenhouse	85.41	94.36	93.19	91.58	10.48	9.11	7.22
Green dry land	61.03	68.32	67.91	65.70	11.94	11.27	7.65

Table 6. *Cont.*

Second Level Classes	SVM	CV	GA	PSO	CV-P	GA-P	PSO-P
Gray dry land	73.13	87.50	84.36	81.86	19.65	15.36	11.94
Fallow land	69.84	78.67	74.15	73.63	12.64	6.17	5.43
Woodland	80.75	83.96	83.18	83.90	3.98	3.01	3.90
Shrubbery	45.03	52.48	52.31	49.50	16.54	16.17	9.93
Coerced forest	72.12	78.79	79.60	76.10	9.25	10.37	5.52
Nursery	54.08	66.29	61.54	63.44	22.58	13.79	17.31
Pond and stream	95.57	97.54	96.62	97.54	2.06	1.10	2.06
Mine pit pond	96.48	98.51	99.00	98.02	2.10	2.61	1.60
Dark road	80.00	89.86	87.44	90.38	12.33	9.30	12.98
Bright road	94.23	99.00	98.00	97.51	5.06	4.00	3.48
Light gray road	75.12	86.54	80.63	86.70	15.20	7.33	15.42
Bright roof	81.25	88.21	87.31	85.00	8.57	7.46	4.62
Red roof	93.26	95.43	96.94	94.42	2.33	3.95	1.24
Dark roof	63.21	79.26	77.78	78.14	25.39	23.05	23.62
Blue roof	95.88	97.46	98.49	95.92	1.65	2.72	0.04
Bare surface	75.13	87.88	85.43	86.15	16.97	13.71	14.67
OA	74.55	81.77	80.41	80.27	9.68	7.86	7.67

P—percentage deviation.

Table 7. McNemar test results for the mapping of open pit mining areas. fij: the numbers of cases that were wrongly classified by classifier i but correctly classified by j (i, j = 1, 2); χ2: chi-square; *p*: probability value.

Pair of Classifications	f12	f21	χ2	*p*
SVM vs. CV	229	70	84.6	<0.001
SVM vs. GA	178	49	73.3	<0.001
SVM vs. PSO	186	60	64.5	<0.001
CV vs. GA	60	90	6.0	<0.05
CV vs. PSO	46	79	8.7	<0.01
GA vs. PSO	76	79	0.1	

4.2.1. OA and Percentage Deviation

The SVM model that was optimized by CV (CV-SVM) achieved the highest OA of 81.77%, followed by GA-SVM (80.41%), PSO-SVM (80.27%), and SVM (74.55%) (Table 6). Clearly, parameter optimization greatly improved the performances of SVM models. The OA values increased by 9.68% (CV-SVM), 7.86% (GA-SVM), and 7.67% (PSO-SVM), with an average of 8.40%. Moreover, CV outperformed the GA and PSO algorithms in this case, with OA value differences of 1.69% and 1.87%.

4.2.2. F1-Measure and Percentage Deviation of Each Land Cover

Regarding the accuracy of each class (i.e., F1-measure), the three optimization methods yielded different effects. For example, only the following seven land covers complied with the above-mentioned conclusion (i.e., CV overtook GA and PSO): greenhouse, green dry land, gray dry land, fallow land, shrub, bright road, and bright roof.

CV provided the highest F1-measure values for greenhouse (94.36%), green dry land (68.32%), gray dry land (87.50%), fallow land (78.67%), woodland (83.96%), shrub (52.48%), nursery (66.29%), bright road (99.00%), bright roof (88.21%), dark roof (79.26%), bare surface (87.88%), and dumping ground (71.70%); GA provided the highest F1-measure values for coerced forest (79.60%), mine pit pond (99.00%), red roof (96.94%), and blue roof (98.49%); PSO provided the highest F1-measure values for paddy (88.78%), dark road (90.38%), light gray road (86.70%), open pit (52.41%), and ore processing site (54.14%). Moreover, CV and PSO were more effective in discriminating pond and stream, with the highest F1-measure value of 97.54%.

In general, all the three parameter optimization methods yielded over 90% F1-measures for the following land covers: greenhouse, pond and stream, mine pit pond, bright road, red roof, blue roof; over 80% for the following: paddy, gray dry land, woodland, dark road, light gray road, bright roof, and bare surface; over 70% for the following: fallow land, coerced forest, dark roof, and dumping ground; over 60% for the following: green dry land and nursery; and, over 50% for the following: shrub, open pit, and ore processing site.

As for the F1-measure of each class, all three parameter optimization methods showed positive effects, i.e., the percentage deviation of F1-measure ranged from 0.04% to 25.39%, with an average of 9.37%. Moreover, all of the parameter optimization methods improved 70% of the 20 land covers by over 5%: greenhouse (with an average improvement of 8.94%), green dry land (10.29%), gray dry land (15.65%), fallow land (8.08%), shrub (14.21%), coerced forest (8.38%), nursery (17.89%), dark road (11.54%), gray road (12.65%), dark roof (24.02%), bare surface (15.12%), open pit (11.61%), ore processing site (13.07%), and dumping ground (7.42%).

In particular, the CV method provided the highest F1-measure improvements for greenhouse (10.48%), green dry land (11.94%), gray dry land (19.65%), fallow land (12.64%), woodland (3.98%), shrub (16.54%), nursery (22.58%), bright road (5.06%), bright roof (8.57%), dark roof (25.39%), bare surface (16.97%), and dumping ground (8.72%); the GA method provided the highest F1-measure improvements for coerced forest (10.37%), mine pit pond (2.61%), red roof (3.95%), and blue roof (2.72%); PSO method provided the highest F1-measure improvements for paddy (5.21%), dark road (12.98%), gray road (15.42%), open pit (13.96%), and ore processing site (16.21%). CV and PSO were particularly more sensitive in discriminating pond and stream, with the highest F1-measure improvement of 2.06%.

4.2.3. McNemar Test

The McNemar test was conducted for each pair of the models based on SVM with default parameters and those optimized parameters while using the three methods. Table 7 shows pairs of classification models, the numbers of samples that one model wrongly classified and another model that correctly classified, and the corresponding chi-square and p values. The results indicate that: (1) SVM models based on CV, GA, and PSO algorithms significantly outperformed those with default parameters (chi-square values were larger than 3.84 and the p values were smaller than 0.05), i.e., the three parameter optimization methods significantly improved the classification performance; (2) there were significant differences between CV and the other two methods, i.e., CV significantly outperformed GA and PSO; and, (3) GA and PSO methods showed equivalent effects with no statistical significance.

4.3. Assessment of the Independent Test Set

A predicted map with the first-level land covers was obtained based on the classification result of the study area that derived from CV-SVM (Figure 3). The predicted map was poorer than that drawn in [9]. There were misclassifications between all of the land covers. In the west part of the study, the misclassification of crop land as forest land existed. In the whole study area, there were land covers that were wrongly classified as open pit mining land and road.

The map was then erased by the data polygons that were used to construct the dependent training and test sets. Finally, a stratified random sampling method was applied to obtain the independent test set with 700 samples. The land classes of the independent test samples were determined by a visual interpretation based on the WV-3 imagery.

Table 8 shows all of the models' OA values for the independent test set and Table 9 depicts the results of statistical test. The SVM model just achieved an OA of 57.43%. The optimized SVM models significantly outperformed the SVM, with the improvements of 10.94%, 8.71%, and 5.22% (average 8.29%). Among the three optimized SVM models, just CV-SVM overtook PSO-SVM, and no significant differences existed between other pairs of them.
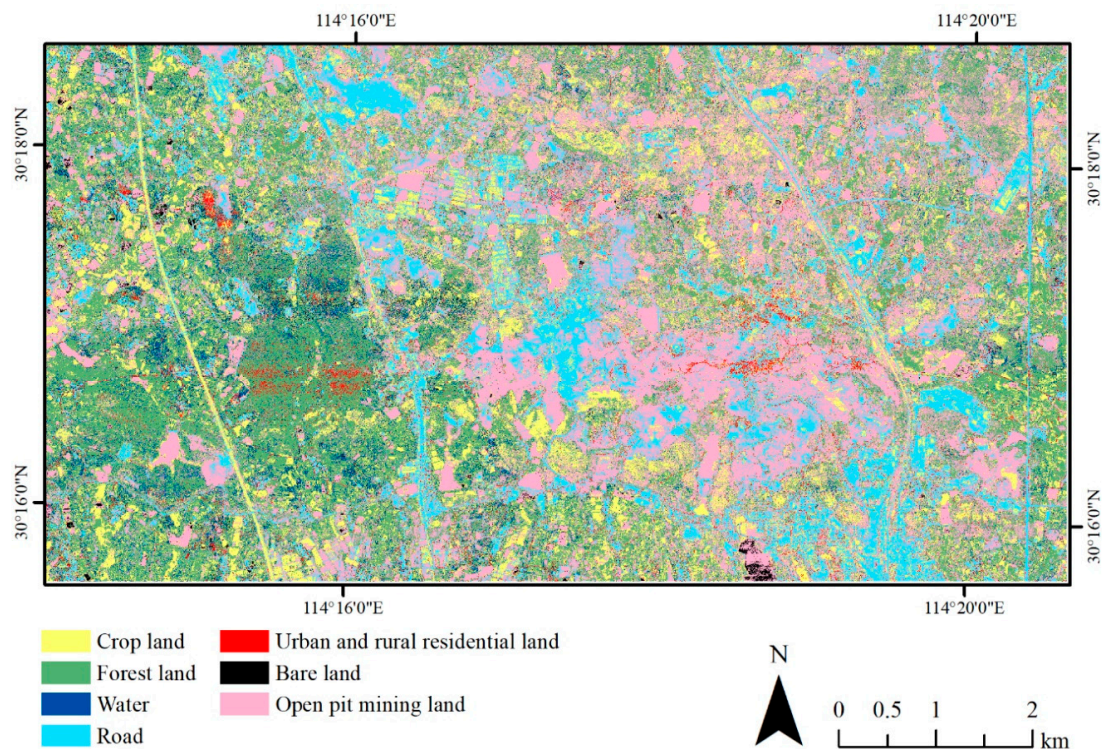
Figure 3. The predicted map of the study area with the first-level land covers.

Table 8. Overall accuracies (OAs) of different methods for the independent test set (%).

	SVM	CV	GA	PSO	CV-P	GA-P	PSO-P
OA	57.43	63.71	62.43	60.43	10.94	8.71	5.22

P—percentage deviation.

Table 9. McNemar test results of different models for the independent test set. fij: the numbers of cases that were wrongly classified by classifier i but correctly classified by j (i, j = 1, 2); χ2: chi-square; *p*: probability value.

Pair of Classifications	f12	f21	χ2	*p*
SVM vs. CV	127	60	24.0	<0.001
SVM vs. GA	102	44	23.0	<0.001
SVM vs. PSO	95	51	13.3	<0.001
CV vs. GA	51	60	0.7	
CV vs. PSO	39	62	5.2	<0.025
GA vs. PSO	56	70	1.6	

5. Discussion

5.1. Effectiveness of The Used Features

The features that were used in this study were like that of previous studies [6,9]. The importance of each feature was assessed and then compared with those in other related studies [9]. Chen et al. [6] further confirmed the importance grades of each feature set and determined whether there were significant differences among them. Overall, the commonly used feature sets, such as spectral information, principal component bands, filter images, and texture measures, are effective for fine LCCOM.

Nevertheless, there are three significant differences between this study and the aforementioned two studies, which resulted in different effects on the classification accuracies of different land covers. (1) This study focused on the fine classification of land covers, which is more difficult than that of

coarse land covers (neglecting intra-class misclassification) [6,9]. (2) This study utilized more spectral information, soil adjusted vegetation index, and the proposed CI, but not topographic variables and the standard deviation filters, which were the top two important feature sets in the previous studies [6]. (3) The WV-3 data that were used in this study were of much higher resolution than that of Ziyuan-3 image in the previous two studies (about seven times), which further increased the classification difficulty at a fine scale.

Taking open pit and ore processing sites, for example, their low accuracies (about 50%) could be attributed to the inherent difficulty of classification, misclassifications between each other, and the lack of topographic variables. Although Li et al. [9] achieved the highest OA of 87.34% for three land covers of open pit, to a certain degree, it could be attributed to the use of 10% data as the training set and the effect of spatial auto-correlation [53] between the training and test sets. The proposed CI could theoretically help to distinguish different open pit mining land classes, providing relatively higher accuracies for dumping ground. However, it could not differentiate the open pit and ore processing site. Obviously, the spectral band of WV-3 should be further explored to generate more effective lithology indices for fine LCCOM. In addition, the integration of lithology indices, topographic variables, and other features might be effective, and they will be considered in the future.

Similarly, the shrub and nursery exhibited low accuracies, owing to the very high resolution of WV-3 and misclassifications between each other. The addition of more spectral information and soil adjusted vegetation index was not enough for distinguishing them. Higher-level features that are derived from spectral information and more effective spatial features, such as the standard deviation filters, have potential in improving their accuracies, and we will investigate them in the future.

The highest OA for the independent test set was just 63.71%. It was lower than the feature subset-based RF, SVM, and ANN models (with OA values of 77.57%, 72.00%, and 64.29%, respectively), and all feature-based RF and SVM models (with OA values of 74.86% and 68.00%). This revealed that the topographic features were important.

5.2. Dependency of Test and Training Sets and Sampling Scheme

When compared to [6,9], this study used spatially dependent training and test sets for fine LCCOM, through which good OA and F1-measures could be obtained for most fine land covers. In general, independent training and test sets are the prerequisites of reliable accuracy assessment. Although the acquisition method of training and test sets determines their spatial auto-correlation [53], the effect was within normal range, as reflected by the statistical data in Table 4. In addition to the white road (31.63%), the number (or area) fractions of samples in the training and test sets and those in data polygons for each land cover class had the maximum, minimum, and average values of 7.34‰, 0.04‰, and 1.53‰, respectively. The used data were only a small portion of data polygons and the spatial auto-correlation of training and test samples was very small. Therefore, it would not have affected the reliability of the accuracy assessment.

The classification in this study was different from the subclassification of three open pit mining lands in [9]. The spatial auto-correlation was very large; however, the large test set ensured the reliability of the accuracy assessment.

The following conclusions could be drawn regarding the effects of spatial auto-correlation on each land cover. (1) On the whole, spatial auto-correlation had little effect. For example, the land covers of gray dry land, gray road, dark roof, and bare surface exhibited higher spatial auto-correlation (i.e., the above-mentioned higher fraction; 1.60‰, 1.42‰, 2.91‰, and 7.34‰) and their F1-measures were moderate (73.13%, 75.12%, 63.21%, and 75.13%; obtained from the SVM model with default parameters). Moreover, high F1-measures could be obtained for some land covers with low spatial auto-correlation, i.e., wood land (0.12‰ and 80.75%) and pond and stream (0.19‰ and 95.57%). It was clear that the separability of these classes was the dominant reason, although there were some other land covers with high spatial auto-correlation and high F1-measures, such as greenhouse (1.77‰ and 85.41%), mine pit pond (1.52‰ and 96.48%), bright road (31.63‰ and 94.23%), bright roof (1.07‰ and

81.25%), red roof (5.06‰ and 93.26%), and blue roof (1.99‰ and 95.88%). (2) The low accuracies of three surface-mined lands could be partly attributed to the insufficient number of training samples, which led to the low spatial auto-correlation to some degree. It could be further concluded that more training samples are necessary and the spatial auto-correlation should be fully and reasonably exploited for open pit mining lands with relatively low separability.

The results of independent test set in this study were worse than those in [9]. These results might reflect the true predict ability of the SVM models. More training data were necessary for better predicted map of the study area, owing to the complexity and difficulty of LCCOM and fine LCCOM. Small training data were just applicable for model comparison.

5.3. Influence of Parameter Optimization

For SVM-based models, parameter optimization is indispensable and it can significantly affect the obtained results [16,24]. Furthermore, many previous studies reported that parameter optimization has positive effects. Similarly, it can be concluded in this study that the three parameter optimization methods significantly improved the classification accuracy of LCCOM. Moreover, CV significantly outperformed GA and PSO, and GA slightly overtook PSO. A statistical test was also performed to determine whether there were significant improvements after parameter optimization and whether there were significant differences among them. In contrast, most studies only used one of the common optimization algorithms or focused only on modifying specific algorithms. Few studies have compared different optimization algorithms. For example, a comprehensive assessment was carried out for different optimization methods in SVM, such as continuous ant colony optimization, GA, imperialist competitive algorithm, and PSO [54,55]. The authors concluded that the suitability of the algorithm depended on the specific application. The limited parameter space [55] and the complexity of LCCOM in this study might obstruct the performance of more complex algorithms, such as GA and PSO.

Some studies have also reported that the combinatorial optimization of feature selection and parameter optimization in SVM might be more effective than only optimizing the parameters in SVM [55,56]. In this study, multiple types of features with highly redundant and correlated information were utilized.

Different land covers showed different sensitivities regarding the three parameter optimization methods. For example, CV was more suitable for most of the fine land covers of crop land (i.e., CV achieved the most substantial accuracy improvements), and PSO for road and open pit mining lands. Schuster et al. [52] and Li et al. [9] also drew a similar conclusion that easily distinguishable land covers are less sensitive to the additional operation, which is the parameter optimization in this study and feature selection in the previous two studies.

6. Conclusions

In this study, SVM models with three parameter optimization methods were investigated for improving the higher accuracy of fine LCCOM based on WV-3 images. Overall, the accuracy was significantly improved. In particular, the fine land covers from mining activities could be identified based on our proposed strategy. This is different from that of our previous study, which focused only on the coarse land cover classes in open pit mining areas. Some important conclusions can be drawn. Firstly, the proposed CI based on WV-3 was useful in distinguishing the dumping ground from other open pit mining lands. Secondly, parameter optimization methods can significantly improve the classification accuracy of fine LCCOM. CV significantly outperformed GA and PSO, and GA slightly overtook PSO. Finally, CV was more suitable for most of fine land covers of crop land, and PSO for road and open pit mining lands. The three optimized SVM models also achieved significant improvements based on the independent test set. In general, the fine land covers in an open pit area could be classified with higher accuracy while using WV-3 and SVM algorithms based on parameter optimization. In the future, combinatorial optimization should be investigated [57], and we will focus on the generalization ability of the proposed strategy in different mining regions.

References

1. Saviour, M.N. Environmental impact of soil and sand mining: A review. *Int. J. Sci. Environ. Technol.* **2012**, *1*, 125–134.

2. Mbaya, R.P. Land degradation due to mining: The gunda scenario. *Int. J. Geogr. Geol.* **2013**, *2*, 144–158.

3. Qian, T.; Bagan, H.; Kinoshita, T.; Yamagata, Y. Spatial-temporal analyses of surface coal mining dominated land degradation in holingol, Inner Mongolia. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 1675–1687. [CrossRef]

4. Sonter, L.J.; Herrera, D.; Barrett, D.J.; Galford, G.L.; Moran, C.J.; Soares-Filho, B.S. Mining drives extensive deforestation in the Brazilian Amazon. *Nat. Commun.* **2017**, *8*, 1013. [CrossRef]

5. Johansen, K.; Erskine, P.D.; McCabe, M.F. *Using Unmanned Aerial Vehicles to Assess the Rehabilitation Performance of Open Cut Coal Mines*; Elsevier: Amsterdam, The Netherlands, 2019; p. 209.

6. Chen, W.; Li, X.; He, H.; Wang, L. Assessing different feature sets' effects on land cover classification in complex surface-mined landscapes by ZiYuan-3 satellite imagery. *Remote Sens.* **2018**, *10*, 23. [CrossRef]

7. Chen, W.; Li, X.; He, H.; Wang, L. A review of fine-scale land use and land cover classification in open-pit mining areas by remote sensing techniques. *Remote Sens.* **2018**, *10*, 15. [CrossRef]

8. Gong, P.; Wang, J.; Yu, L.; Zhao, Y.; Zhao, Y.; Liang, L.; Niu, Z.; Huang, X.; Fu, H.; Liu, S.; et al. Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *Int. J. Remote Sens.* **2013**, *34*, 2607–2654. [CrossRef]

9. Li, X.; Chen, W.; Cheng, X.; Wang, L. A comparison of machine learning algorithms for mapping of complex surface-mined and agricultural landscapes using ZiYuan-3 stereo satellite imagery. *Remote Sens.* **2016**, *8*, 514. [CrossRef]

10. De L. Lobo, F.; Souza-Filho, P.W.M.; de Novo, E.M.L.m.; Carlos, F.M.; Barbosa, C.C.F. Mapping mining areas in the Brazilian Amazon using MSI/sentinel-2 imagery (2017). *Remote Sens.* **2018**, *10*, 1178. [CrossRef]

11. Sellers, P.J.; Meeson, B.W.; Hall, F.G.; Asrar, G.; Murphy, R.E.; Schiffer, R.A.; Huemmrich, K.F. Remote sensing of the land surface for studies of global change: Models algorithms experiments. *Remote Sens. Environ.* **1995**, *4257*, 3–26. [CrossRef]

12. Yu, L.; Xu, Y.; Xue, Y.; Li, X.; Cheng, Y.; Liu, X.; Porwal, A.; Holden, E.J.; Yang, J.; Gong, P. Monitoring surface mining belts using multiple remote sensing datasets: A global perspective. *Ore Geol. Rev.* **2018**, *101*, 675–687. [CrossRef]

13. Maxwell, A.E.; Warner, T.A.; Strager, M.P.; Pal, M. Combining rapideye satellite imagery and lidar for mapping of mining and mine reclamation. *Photogramm. Eng. Remote Sens.* **2014**, *80*, 179–189. [CrossRef]

14. Maxwell, A.E.; Warner, T.A.; Strager, M.P.; Conley, J.F.; Sharp, A.L. Assessing machine-learning algorithms and image- and lidar-derived variables for GEOBIA classification of mining and mine reclamation. *Int. J. Remote Sens.* **2015**, *36*, 954–978. [CrossRef]

15. Goldblatt, R.; You, W.; Hanson, G.; Khandelwal, A.K. Detecting the boundaries of urban areas in India: A dataset for pixel-based image classification in google earth engine. *Remote Sens.* **2016**, *8*, 634. [CrossRef]

16. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 247–259. [CrossRef]

17. Schalkoff, R.J. *Artificial Neural Networks (International ed)*; McGraw-Hill Higher Education: New York, NY, USA, 1997.

18. Belgiu, M.; Drăgu, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]

19. De Boissieu, F.; Sevin, B.; Cudahy, T.; Mangeas, M.; Chevrel, S.; Ong, C.; Rodger, A.; Maurizot, P.; Laukamp, C.; Lau, I.; et al. Regolith-geology mapping with support vector machine: A case study over weathered Ni-bearing peridotites, New Caledonia. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *64*, 377–385. [CrossRef]

20. Ishida, H.; Oishi, Y.; Morita, K.; Moriwaki, K.; Nakajima, T.Y. Development of a support vector machine based cloud detection method for MODIS with the adjustability to various conditions. *Remote Sens. Environ.* **2018**, *205*, 390–407. [CrossRef]

21. Karan, S.K.; Samadder, S.R. Improving accuracy of long-term land-use change in coal mining areas using wavelets and support vector machines. *Int. J. Remote Sens.* **2018**, *39*, 84–100. [CrossRef]

22. Morales-Alvarez, P.; Perez-Suay, A.; Molina, R.; Camps-Valls, G. Remote sensing image classification with large-scale gaussian processes. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 1103–1114. [CrossRef]

23. Negri, R.G.; Da Silva, E.A.; Casaca, W. Inducing contextual classifications with kernel functions into support vector machines. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 962–966. [CrossRef]

24. Xiao, X.; Zhang, T.; Zhong, X.; Shao, W.; Li, X. Support vector regression snow-depth retrieval algorithm using passive microwave remote sensing data. *Remote Sens. Environ.* **2018**, *210*, 48–64. [CrossRef]

25. Maulik, U.; Chakraborty, D. Remote sensing image classification: A survey of support-vector-machine-based advanced techniques. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 33–52. [CrossRef]

26. Liu, P.; Choo, K.K.R.; Wang, L.; Huang, F. SVM or deep learning? A comparative study on remote sensing image classification. *Soft Comput.* **2017**, *21*, 7053–7065. [CrossRef]

27. Sukawattanavijit, C.; Chen, J.; Zhang, H. GA-SVM algorithm for improving land-cover classification using SAR and optical remote sensing data. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 284–288. [CrossRef]

28. Zhuo, L.; Zheng, J.; Li, X.; Wang, F.; Ai, B.; Qian, J. A genetic algorithm based wrapper feature selection method for classification of hyperspectral images using support vector machine. In Proceedings of the Geoinformatics 2008 and Joint Conference on GIS and Built Environment: Classification of Remote Sensing Images, Guangzhou, China, 28–29 June 2008; Volume 7147, p. 71471J.

29. Wu, C.H.; Tzeng, G.H.; Goo, Y.J.; Fang, W.C. A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy. *Expert Syst. Appl.* **2007**, *32*, 397–408. [CrossRef]

30. Zhai, S.; Jiang, T. A novel particle swarm optimization trained support vector machine for automatic sense-through-foliage target recognition system. *Knowl. Based Syst.* **2014**, *65*, 50–59. [CrossRef]

31. Kruse, F.A.; Baugh, W.M.; Perry, S.L. Validation of digitalglobe Worldview-3 Earth imaging satellite shortwave infrared bands for mineral mapping. *J. Appl. Remote Sens.* **2015**, *9*, 096044. [CrossRef]

32. Asadzadeh, S.; de Souza Filho, C.R. Investigating the capability of Worldview-3 superspectral data for direct hydrocarbon detection. *Remote Sens. Environ.* **2016**, *173*, 162–173. [CrossRef]

33. Salehi, T.; Tangestani, H.M. Large-scale mapping of iron oxide and hydroxide minerals of zefreh porphyry copper deposit, using Worldview-3 VNIR data in the Northeastern Isfahan, Iran. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 156–169. [CrossRef]

34. Chen, W.; Zhang, L.; Zhang, Z.; Wang, Y. Comparison of fusion imageries with spectral fidelity using SPOT5. In *MIPPR 2009: Multispectral Image Acquisition and Processing*; International Society for Optics and Photonics: Orlando, FL, USA, 2009; Volume 7494, p. 74941K.

35. Goward, S.N.; Markham, B.; Dye, D.G.; Dulaney, W.; Yang, J. Normalized difference vegetation index measurements from the advanced very high resolution radiometer. *Remote Sens. Environ.* **1991**, *35*, 257–277. [CrossRef]

36. Huete, A. A soil–adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1998**, *25*, 295–309. [CrossRef]

37. Clark, R.N.; Roush, T.L. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *J. Geophys. Res.* **1984**, *89*, 6329–6340. [CrossRef]

38. Pearson, K., LIII. On lines and planes of closest fit to systems of points in space. London, Edinburgh, Dublin Philos. *Mag. J. Sci.* **1901**, *2*, 559–572. [CrossRef]

39. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural Features for Image Classification. *IEEE Trans. Syst. Man. Cybern.* **1973**, *6*, 610–621. [CrossRef]

40. Huang, C.; Davis, L.S.; Townshend, J.R.G. An assessment of support vector machines for land cover classification. *Int. J. Remote Sens.* **2002**, *23*, 725–749. [CrossRef]

41. Bazi, Y.; Melgani, F. Toward an optimal SVM classification system for hyperspectral remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 3374–3385. [CrossRef]

42. Ghoggali, N.; Melgani, F. Genetic SVM approach to semisupervised multitemporal classification. *IEEE Geosci. Remote Sens. Lett.* **2008**, *5*, 212–216. [CrossRef]

43. Bazi, Y.; Melgani, F. Semisupervised PSO-SVM regression for biophysical parameter estimation. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1887–1895. [CrossRef]

44. Cracknell, M.J.; Reading, A.M. Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Comput. Geosci.* **2014**, *63*, 22–33. [CrossRef]

45. Fassnacht, F.E.; Hartig, F.; Latifi, H.; Berger, C.; Hernández, J.; Corvalán, P.; Koch, B. Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sens. Environ.* **2014**, *154*, 102–114. [CrossRef]

46. Ghamisi, P.; Benediktsson, J.A. Feature selection based on hybridization of genetic algorithm and particle swarm optimization. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 309–313. [CrossRef]

47. Meyer, D.; Dimitriadou, E.; Hornik, K.; Weingessel, A.; Leisch, F. *Misc Functions of the Department of Statistics (e1071)*; R Package Version 1.6–4; TU Wien: Vienna, Austria, 2014; p. 1071.

48. Beasley, D.; Bull, D.R.; Martin, R.R. An overview of genetic algorithms: Part 1, fundamentals. *Univ. Comput.* **1993**, *15*, 56–69.

49. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the IEEE International Conference on Neural Networks, Perth, Western Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948.

50. Wang, L.; Geng, H.; Liu, P.; Lu, K.; Kolodziej, J.; Ranjan, R.; Zomaya, A.Y. Particle swarm optimization based dictionary learning for remote sensing big data. *Knowledge-Based Syst.* **2015**, *79*, 43–50. [CrossRef]

51. Daskalaki, S.; Kopanas, I.; Avouris, N. Evaluation of classifiers for an uneven class distribution problem. *Appl. Artif. Intell.* **2006**, *20*, 381–417. [CrossRef]

52. Schuster, C.; Förster, M.; Kleinschmit, B. Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data. *Int. J. Remote Sens.* **2012**, *33*, 5583–5599. [CrossRef]

53. Chen, W.; Li, X.; Wang, Y.; Chen, G.; Liu, S. Forested landslide detection using LiDAR data and the random forest algorithm: A case study of the Three Gorges, China. *Remote Sens. Environ.* **2014**, *152*, 291–301. [CrossRef]

54. Liu, T.; Abd-Elrahman, A.; Morton, J.; Wilhelm, V.L. Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. *GISci. Remote Sens.* **2018**, *55*, 243–264. [CrossRef]

55. Tamimi, E.; Ebadi, H.; Kiani, A. Evaluation of different metaheuristic optimization algorithms in feature selection and parameter determination in SVM classification. *Arab. J. Geosci.* **2017**, *10*, 478. [CrossRef]

56. Wang, M.; Wan, Y.; Ye, Z.; Lai, X. Remote sensing image classification based on the optimal support vector machine and modified binary coded ant colony optimization algorithm. *Inf. Sci.* **2017**, *402*, 50–68. [CrossRef]

57. Li, X.; Tang, Z.; Chen, W.; Wang, L. Multimodal and Multi-Model Deep Fusion for Fine Classification of Regional Complex Landscape Areas Using ZiYuan-3 Imagery. *Remote Sens.* **2019**, *11*, 2716. [CrossRef]