Use of crowdsourcing in evaluating post-classification accuracy

Ekrem Saralioglu and Oguz Gungor

Department of Geomatics, Karadeniz Technical University, Trabzon, Turkey

ABSTRACT
“Crowdsourcing” uses masses of people to solve a specific problem, usually focusing on research strategies to reduce time, cost and effort to create data. Crowdsourcing intrinsically claims that groups can make relatively smarter and better decisions than the most intelligent individual among them. We investigated to see if crowdsourcing could be used to collect control points for usage in calculating post-classification accuracy assessments. For this purpose, a test was done using class values of randomly generated 1000 control points. Its goal was to explore the accuracy of a specific class values to be entered by three different users by utilizing majority voting method. While examining 3 data sets containing 1000 points, it could be observed that the class values of only 4 points were entered incorrectly. When the support vector machine classification results were evaluated, using the same 1000 control points generated by the experts and the crowdsourcing (containing 4 faulty points), the classification accuracies were found to be 85.0487% and 84.6154%, respectively. Results show that crowdsourcing offers a quicker and more reliable post-classification accuracy assessment for high-spatial resolution multispectral images.

Introduction
Crowdsourcing is commonly described as “the practice whereby an organization enlists a variety of people, paid or unpaid, to work on a specific task or problem”. The term “crowdsourcing”, however, was first used by Jeff Howe in his article, “The Rise of Crowdsourcing”, published in WIRED (Howe, 2006) describes this method of using masses of participants to solve a problem goes back hundreds of years (Thenkabail, 2015). For example, in 1714 in Britain, a newly created commission offered a 20,000 pound sterling prize to anyone who could propose a method that accurately calculates longitude, allowing ships to pinpoint their positions at sea for the first time in history (Howe, 2008). A lesser known example of successful crowdsourcing involves Toyota. In 1936, the Japanese auto manufacturer organized a contest to design a new corporate logo. From the 27,000 entries, a winner was selected and the company’s name permanently changed from the family name, Toyoda, to its present corporate identifier, TOYOTA.

So far, crowdsourcing has been used to solve the problems in a diverse variety of disciplines. As a result, many new terms have emerged in scientific literature. Among these, “volunteered geographic information” (Goodchild, 2007) and “neo-geography” (Turner, 2006) focus on the spatial nature of the data, while “crowdsourcing” (Howe, 2006), “citizen science” (Bonney et al., 2009) and “user-generated content” (Krumm, Davies, & Narayanaswami, 2008) have a much wider range of applications. Albeit very different in nature, these terms are often used interchangeably for various activities related to geographic information science with the participation of citizens (See et al., 2016).

The biggest and fastest development in crowdsourcing was the emergence of the World Wide Web, and especially Web 2.0 technologies. No matter where you live on earth, the Internet allows you to make contributions, share resources easily and deliver results quickly. Today, the “Amazon Mechanical Turk” website is one of the most well-known crowdsourcing sites. Initially, it emerged based on the idea that there are many tasks that people can do better than computers, such as describing and listing objects in a photo. On Amazon.com’s Mechanical Turk forum, individuals or businesses specify the work to be done, give explanations on how to do it, and declare the fee to be paid per piece/hour. Persons who want the work must apply, and if they land the job, complete it within the specified time frame to get paid. (Behrend, Sharek, Meade, & Wiebe, 2011). Wikipedia, I-stock Photo, innocente.com,
OpenStreetMap are also other crowdsourcing-based applications.

On the other hand, one of the first examples of studies on crowdsourcing being used to solve geospatial problems was the “Did You Feel It” project, which was initiated in the early 1990s by United States Geological Survey (Wald, Quitoriano, Worden, Hopper, & Dewey, 2012). This project was the first significant example of crowdsourcing applied to natural disaster management (Thenkabail, 2015). The ClickWorkers project used The Global Earthquake Model as an international forum where organizations and individuals came together to develop, use and share tools and resources for the unbiased assessment of the earth-quake risk (Thenkabail, 2015) resulting in another great example of crowdsourcing.

Since 2009, the ongoing Geo-Wiki project has focused on the solution of geospatial problems utilizing contributions from “crowds”. Due to the large discrepancies in certain places in MODIS, GlobCover and GLC-2000 land cover maps, the Geo-Wiki project was initiated as an online volunteer network to increase the accuracy of global land cover maps. Another good example of crowdsourcing is the Global Earth Observation System of Systems (GEOSS) that was built by the Group on Earth Observations (GEO). Working in isolation from each other, GEOSS is a set of coordinated, independent earth observation, information and processing systems that interact and provide access to a diverse information resource for a broad range of users from both public and private sectors. As a crowdsourcing application, the “GEOSS Portal” offers a single Internet access point for users seeking data, imagery and analytical software packages relevant to all parts of the globe by connecting users to existing data bases and portals to provide reliable, up-to-date and user-friendly information for decision-makers, planners and emergency managers. Commercial remote sensing satellite company Digital Globe initiated another successful crowdsourcing application. Thanks to the Tomnod platform, more than two million Internet users have tried to help search and rescue teams to find aircraft for Malaysian airways lost in 2014 (Karaman et al., 2015).

The spatial resolutions of remote sensing images, which started at 80 m with the launch of the first Landsat satellite, have reached 30 cm with WV-3 and WV-4 satellites, and a couple of centimeters or better with the emerging of unmanned aerial vehicles (UAVs). The increase in spatial resolution indicates that the resulting image size and data processing time have also increased, creating unexpected difficulties in analyzing the data. Although most of the data collection and processing can be completed automatically, analysis and interpretation of remote sensing data is a complex task and still may not be entirely fulfilled by machine-driven algorithms.

One of the most common users of remotely sensed images can be found in the creation of thematic maps, such as land cover maps based on image classification. This is done by assigning pixels to predefined classes either by detecting groups of pixels (clusters) in image data that have similar spectral characteristics (unsupervised classification) or by selecting representative sample sites of known cover types to classify the entire image (supervised classification). In either case, the resulting classified image can be considered a thematic map, showing the land cover type found in the region. These thematic maps are used to create and update land cover maps for different applications related to monitoring the earth on a regional and global scale (Pasolli, Melgani, Alajlan, & Conci, 2013). There are two approaches for estimating thematic correctness of these maps. While the first approach uses control points directly on the ground, the second one establishes control points on aerial photography and/or high-resolution satellite imagery (Carlotto, 2009). Depending on the spatial resolution of the imagery, it is critical to detect the number of control points to be selected. It is known that for 30-m spatial resolution Landsat imagery, more than 250 control points are needed to estimate the average accuracy of a class at ±5% (Congalton, 1991). Akar & Güngör, 2015 used WorldView-2 (WV-2) image with a size of 1000 × 1000 and a spatial resolution of 2 m, and they calculated the number of necessary control point for accuracy assessment as 735 using the multinomial distribution equation given by Congalton and Green (1999). This number would be much higher in WV-3 and WV-4 images because of their higher spatial resolutions, making it difficult and time-consuming to enter the class values of large amounts of control points needed for accuracy assessment. Therefore, it may be appropriate to use collective intelligence, employing human perception and interpretation ability, to cope with such a large amounts of data (Hu et al., 2012).

This study investigates the usability of a crowdsourcing concept in a post-classification accuracy assessment of satellite images with high-spatial resolutions (1 m or better spatial resolution). For this purpose, randomly generated 1000 points scattered across a pan-sharpened WV-2 image are provided to volunteers through a web interface so they can determine and enter real class values of these points through a web interface. Users can, if they
wish, use the help document that was prepared and presented in the web interface to learn how to enter the class value of each point. The class value of each of the 1000 points is entered three times by three different users, and the class value of each point determined by a majority voting method. For a final verification, the resultant data set is also compared to the reference data set created by an expert and controlled by another expert. This verification process reveals that only 4 of the 1000 points were incorrectly entered by the users. These control points are used to determine the classification accuracy of a WV-2 image classified using support vector machines (SVM) algorithm. The results show that the crowdsourcing approach can be used with post-classification accuracy assessment of high-resolution satellite images.

Materials and methods

Data set

WV-2 multispectral (MS) image of Sürmene, a small town within the Trabzon province (located on the Black Sea coast in Northeastern Turkey) was used. The WV-2 satellite provides a panchromatic band covering the spectrum from 450 to 800 nm with 0.50 m spatial resolution. It also offers eight MS bands (i.e. four standard bands namely red, green, blue and near-infrared-1 and four additional bands namely, coastal, yellow, red edge and near-infrared 2) covering the spectral range from 400 to 1050 nm at a spatial resolution of 1.84 m (Padwick, Deskevich, Pacifici, & Smallwood, 2010). WV-2 images are used for enhanced spectral analysis, mapping and monitoring applications, land-use planning, disaster relief, exploration, defense and intelligence, and visualization and simulation environments.

With the QGIS program, randomly selected 1000 points were created in GeoJSON format on the WV-2 image over the study area. In the attribute table of each point, a “class value” field was added to enter the real class value for each control point.

The verification data set, containing class values of 1000 control points created by an expert and controlled by another expert, was used for an accuracy assessment. The expert class values of 77 of the 1000 points were not identified. Hence, no class value was entered for these points.

The world map with open layers was provided as the base map. Cadastral boundary data and classified image data with SVM algorithm are other layers presented on the web interface.

Processing steps

Since its aim was to implement a post-classification accuracy assessment analysis with a crowdsourcing approach rather than a single person, this study needed a classified image. For this purpose, instead of using the original WV-2 MS data, pan-sharpened WV-2 MS data (pan-sharpened using WV-2 panchromatic band) were classified by the SVM classification algorithm so that the class values of the control points can be more clearly detected by users. This pan-sharpened WV-2 MS data were used in the accuracy assessment section of the study. To evaluate post-classification accuracy, conventionally, 1000 random sample points were selected on fused WV-2 image. As an alternative to conventional post-classification accuracy assessment approach, a crowdsourcing interface was created, designed to be used by multiple online users to identify class values of each sample point on the fused MS image.

Gram–Schmidt pan-sharpening method

Gram–Schmidt pan-sharpening method is similar to Principal component analysis (Karathanassi, Kolokousis, & Ioannidou, 2007). The Gram–Schmidt method combines the high-spatial resolution panchromatic image with low-spatial resolution MS one to increase the spatial resolution of the MS image (Yuhendra & Kuze, 2011). The first step is to compute a simulated panchromatic band from combination of the lower spatial resolution MS bands. In general, this is achieved by averaging the MS bands. This simulated panchromatic band is employed as the first band of the MS image. Second, a Gram–Schmidt orthogonalization transformation is performed on the simulated panchromatic band and the MS bands to get Gram–Schmidt bands. Then, the high-spatial resolution panchromatic band is replaced with the first Gram–Schmidt band. Finally, the inverse Gram–Schmidt transform is applied to create the pan-sharpened MS bands (Laben et al. 2000).

The Gram–Schmidt method is an effective pansharpening method (Karathanassi et al., 2007; Klonus & Ehlers, 2009; Saralioglu et al., 2016, Yuhendra & Kuze, 2011). This method has two main advantages. First, there is no limit on the number of bands that can be processed in one shot using this technique. Second, the spectral properties of the original MS data are preserved in the fused image satisfactorily (Li, Liu, Wang, Zhao, & Wang, 2004). As given in Figure 1, the MS bands of the WV-2 image are fused with the panchromatic band, resulting in a pan-sharpened high-resolution image.
Since the pan-sharpened image obtained has a spatial resolution of 0.50 m, the class values are more distinguishable by the users.

**Image classification**

SVMs are a nonparametric statistical learning algorithm that originally aims at binary classification, identifying the most appropriate hyperplane that provides the maximum margin separating two classes. (Cortes & Vapnik, 1995; Huang, Davis, & Townshend, 2002; Ustuner, Sanli, & Dixon, 2015; Vapnik, 1995). This method is now widely used for the classification of the remotely sensed data since it has been enhanced to be used for also multiple classes (Abdikan, Bilgin, Sanli, Uslu, & Ustuner, 2015; Adam, Mutanga, Odindi, & Abdel-Rahman, 2014; Adelabu, Mutanga, & Adam, 2014; Schuster, Förster, & Kleinschmit, 2012; Schuster, Schmidt, Conrad, Kleinschmit, & Förster, 2015; Ustuner et al., 2015). In case studies, the SVM method has been found to provide higher classification accuracy than many commonly used classification techniques such as maximum likelihood classification (Deilami, Ahmad, & Zabihi, 2014; Huang et al., 2002; Melgani & Bruzzone, 2004; Mondal, Kundu, Chandniha, Shukla, & Mishra, 2012; Naguib, Farag, Yahia, Ramadan, & Abd Elwahab, 2009; Oommen et al., 2008; Pal & Mather, 2005; Yu, Porwal, Holden, & Dentith, 2012).

The SVM method was applied to the fused WV-2 MS image with a 0.50 m spatial resolution. Kernel functions commonly used in SVMs can be generally aggregated into four groups; namely, linear, polynomial, Radial Basis Function (RBF) and sigmoid kernels (Huang et al., 2002; Kavzoglu & Colkesen, 2009; Keerthi & Lin, 2003; Pal & Mather, 2005). RBF has some advantages compared to the other kernels. In contrast to the linear kernel, the RBF moves samples from non-linear maps to a higher dimensional space to deal with situations in which the relationship between class labels and attributes is not linear. The number of hyper parameters that affect the complexity of model selection is less in RBF when compared to the polynomial kernel. Sigmoid kernel is not valid (i.e. not the inner product of two vectors) under some parameters (Hsu, Chang, & Lin, 2003; Vapnik, 1995). RBF also requires defining a small number of parameters that are widely known to generally produce good results (Keerthi & Lin, 2003; Pal & Mather, 2005; Petropoulos, Kalaitzidis, & Vadrevu, 2012). Therefore, the RBF kernel was chosen for the study.

In order to run SVM classifier in ENVI (Environment for Visualizing Images), it is necessary to enter the penalty parameter, the number of pyramid levels to be used, the classification probability threshold value and the “gamma (g)” value in the kernel function. Generally, very little guidance exists in traditional visualization data, concerning the criteria to be used in selecting the kernel-specific parameters (Carrao et al., 2008; Li & Liu, 2010). In our case, parameterization of the above RBF kernel function was based on performing a number of permutations and combinations using classification accuracy as a measure of quality. A similar approach was followed by the researchers (Fauvel, Chanussot, & Benediktsson, 2006; Kueemmerle, Chaskovskyy, Knorn, Radeloff, & Kruhlav, 2009; Pal & Mather, 2006). In addition, suggestions provided in the ENVI User’s Guide (2008) were also taken into account when doing kernel function parameterization. As a result, the gamma (g) parameter was set to a value equal to the inverse of the number of the spectral bands of the WV-2 (0.125), whereas the penalty parameter was set to its maximum value (100), forcing no misclassification during the training process. The pyramid parameter was set to a value of zero, forcing the WV-2 imagery to be processed at full resolution. Whereas, a classification probability with a threshold of zero was used, requiring that all image pixels had to be sorted into one of the classes.

To initiate the training process in the ENVI program, at least 240 sample pixels were collected for each class by selecting more than 1 polygon per class, resulting in a total of 1755 pixels collected to train 7 classes namely, “forest”, “hazelnut”, “shade”, “soil”, “tea”, “building” and “road” classes (see Figure 2).

**Generating reference points**

The Accuracy Assessment utility allows analysts to compare certain pixels in the thematic raster layer with reference pixels that already have known class values. To determine the post-classification accuracy in a non-biased way, a sufficient number of reference pixels must be randomly selected in the image to be classified. (Congalton, R., 1991). Congalton and Green (1999) offered the following multinominal distribution equity pixels to calculate the accuracy of the classification (Akar & Güngör, 2015).
\[ n = \frac{(\prod_{i} (1 - \pi_{\alpha}^i))}{(b_{i}^2)}; \]

where "n" represents the number of reference pixels, \( B = (a/k) \times 100 \), "\( a \)" is the confidence interval. "\( k \)" is the number of class intervals, "\( i \)" is the ratio of the area of the \( i \)th value of the entire area and "\( b_{i} \)” is the required accuracy. According to the above equation, a minimum of 735 reference pixels are needed to estimate the mean accuracy of a class within a plus or minus five confidence intervals.

A 1000 points were randomly generated by the QGIS program to exceed the required minimum number and a column-labeled "class" was added to the attribute table, allowing users to enter the correct class values for each reference point. The result was ultimately recorded in GeoJSON format.

**Crowdsourcing web application**

In the crowdsourcing web interface that uses MySQL database for database operations, vector layers are generated in a GeoJSON format. In the interface, users are asked to fill an empty "class" field for each reference point. Site domain, www.cremotesensing.app, is purchased and a web site designed, using a Visual Studio environment. ISS server is used as the web server. On front-hand side, HTML, CSS, Bootstrap and JavaScript are used, while C#, jQuery and various libraries are used on back-hand side. The architecture of crowdsourcing web interface is illustrated in Figure 3.

When a user visits the web page www.cremotesensing.app.com, registration is mandatory so that the user can participate in the crowdsourcing application. Their names, surnames and contact information are requested on the registration form. If the administrator approves the request, users can access the interface and carry out their assigned task. The available layers in the interface include World map created with OpenLayers, a fused WV-2 MS image, 1000 randomly selected reference points, cadastral boundaries and a classified image.

The following rules and restrictions apply to the web application:

- Multiple users can access the web page simultaneously.
- 1000 randomly selected points generated at the beginning are used throughout the whole process.
- When the first user logs into the system, 20 out of 1000 randomly selected points are displayed on the screen. By clicking on any of the 20 points on the WV-2 image, the user needs to enter the class value corresponding to each point. If the users want, they can also use the help document created at the interface.
- A user sometimes may not be able to decide what the class value of a reference point should be; however, the user can enter and save as
many points as he/she wants, or leave the entries blank.

- When points are saved by the user, only the points entered with class values are stored in the database.
- If the first user enters the class values of all 20 given points and the same/or another user makes a new request, the system displays 20 new points out of the minimum 980 points needed.
- The system will continue to display 20 points for each new request until 1000 reference points are completed.
- Suppose there are 14 points left to complete a total of 1000 points but more than 1 visitor is requesting work. In response, the system will display 14 points to each of the requesters. Note: Only the results of the person who finished the work first are stored in the database.
- It is only after all 1000 points are processed that the first data set is created.

The same number – 1000 reference points – is used for the second data set, following the same process. These steps are repeated until desired number of data sets are created.

In this study, a total of three data sets were created by users, i.e. three values are entered by users for a single reference point. (Figure 4 shows the crowdsourcing interface.)

**Help document**

A document has been provided in the HELP section, explaining clearly how users can input the reference point values correctly. This document also shows exactly how the classes should appear in the WV-2 image and which values should be entered into the reference points for each class. The Help document is shown in Table 1.

**Results and discussions**

As explained above, 3 sets of data for 1000 reference points were generated. The reason for entering more than one value for a reference point is to detect incorrectly entered values and get rid of them. If only two of the three users enter the same class value for the same reference point, then the majority of the three values are considered the class value given by the users. If all three users entered different values for a reference point, it means that point is not fully distinguishable. So, the point is left unclassified. In this study, 73 points were labeled as "unclassified". (Figure 5 illustrates a reference point that is distinguishable, and an indistinguishable one labeled as an "unclassified reference point").

In order to see how accurate the resultant reference data sets are, after being generated according to the users’ answers, a fourth data set was created for verification by determining the class values of the same 1000 reference points. During this process, it was found that it was somewhat difficult to determine the class values of 77 points. The reason for this was that the resolution of the image was too degraded after zooming-in and that the points were located on the line, separating the two classes. The generated verification set was then compared to the resultant set generated by individual users. As a result of this comparison, it was found that the class values of two points were incorrectly
entered in the verification data set. This may have been due to the user fatigue (combined with examining too many points), carelessness, mistyping or not zooming-in enough to clearly see the class values. It can be interpreted that when a single user has to deal with too many reference points, such errors often occur. This result also proves the theory that the decisions derived from a crowd are more accurate than the decisions of a single expert. In addition, when the validation data set was compared with the user-generated data set, researchers discovered the values of the four reference points were incorrectly entered by the users. If it is desired to further reduce such errors, more than three data sets must be created with different users in order for the most appropriate results to be obtained, using statistics.

In this particular study, the verification data set was generated by an expert and controlled by another expert. The resulting data set generated by crowdsourcing was used for an accuracy assessment for the SVM classification product as shown in Table 2 below.

As shown in Table 2, the effect of four points misinterpreted by the crowdsourcing on the classification accuracy is seen to be less than 1% compared to the higher percentage of errors revealed in the accuracy assessment created by the expert.

Additionally, double and triple combinations of the class values were examined in order to see which reference points users could not distinguish from the reference points seen in the WV-2 image. As explained in the help document, class values are: forest = 1, hazelnut = 2, shadow = 3, soil = 4, tea = 5, building = 6 and road = 7. Thus, reference points whose class values are entered differently in all three data sets generated by user responses were more easily evaluated. Triple combinations of the class values are given in Figure 6.

In the Figure 6, the horizontal axis represents the three different class values given by three different users for each reference point. The vertical axis represents the total number of points. Here, 125 means that users entered 1 (forest), 2 (shadow) and 5 (tea) for the same reference point in 9 out of 1000 reference points. It’s clear that users could not differentiate between forest, shadow and hazelnut classes, 19 times. This figure indicates that when determining the class values of reference points, users had the most difficulty with forest–hazel–shadow, which is not surprising since hazelnut and other trees
At first, user should tick Points WV-2 image layers on the upper left corner of the displayed image. The user can also enable other layers by clicking on them. When user clicks on a point, a class value of 1–7 should be entered for that point. When determining the class value of each point, the original fused MS data should be considered, not the classified image.

| Class        | Image                  | Color      |
|--------------|------------------------|------------|
| Forest       | ![Forest image]        | Forest area| Class value = 1
| Hazelnut     | ![Hazelnut image]      | Hazelnut area| Class value = 2
| Shadow       | ![Shadow image]        | Shadow area| Class value = 3
| Soil         | ![Soil image]          | Soil area  | Class value = 4
| Tea          | ![Tea image]           | Tea area   | Class value = 5
| Buildings    | ![Buildings image]     | Buildings area| Class value = 6
| Road         | ![Road image]          | Road area  | Class value = 7

Each class and its corresponding color.
have similar spectral characteristics. Moreover, it’s even more difficult to distinguish dark green or blackish forest areas when they’re close to black or shaded areas.

The double combinations of such class values are also examined to determine which two classes are most difficult for users to distinguish (see Figure 6).

In fact, Figure 7 is derived from Figure 5. The value 46 (4 (land) and 6 (building) classes) on x-axis means 4 and 6 together appear once in 146, 346, 456 and 467 in Figure 5, which makes this double combination appear four times. As seen in Figure 6 that users mostly have difficulty (34 times) distinguishing between classes 1 (forest) and 2 (hazelnut). The second biggest problem is seen between Class 2 (hazelnut) and Class 3 (shadow) where users had trouble 26 times in making consistent decisions between them. The third biggest confusion lies between Class 1 (hazelnut) and Class 3 (shadow) where it appeared 24 times. The results show that the least problem occurs between class 4 and class 7 (only 2 times).

### Conclusion

Fifty non-trained people or users participated in this study. It concluded that non-professionals can assign class values to reference points on high-resolution satellite images with the use of a Help document. This process had to be repeated three times due to recurring issues with incorrect data entry. The resulting data set was created by taking the majority of the users’ votes. Each user spent about 10 min during this process, taking a total of 9 h. When post-classification accuracy analysis is performed by a single person, the whole process becomes prolonged and ripe with errors due to human factors, such as fatigue and distraction. Nevertheless, this study indicates that the post-classification accuracy analysis can be successfully implemented utilizing a crowdsourcing approach. In this application, the class value of each reference point was assigned by three different users. It should be noted that by increasing the number of users

---

Table 2. Accuracy assessment of classification with result and verification data set.

| Result dataset | Verification dataset |
|----------------|----------------------|
| Prod. Acc. (Per cent) | User Acc. (Per cent) | Prod. Acc. (Per cent) | User Acc. (Per cent) |
| 87.31 | 77.48 | 87.41 | 78.15 |
| 77.55 | 92.68 | 77.55 | 92.68 |
| 97.14 | 91.89 | 97.14 | 91.89 |
| 66.67 | 68.78 | 67.69 | 69.84 |
| 90.09 | 90.73 | 90.54 | 90.97 |
| 94.44 | 89.47 | 94.44 | 89.47 |
| 92 | 100 | 92 | 100 |
| Overall Acc = 84.6154 | Overall Acc = 85.0487 |

---

Figure 5. (a) Distinguishable reference point, (b) Indistinguishable reference point.

Figure 6. Triple examination of different value classes in each three databases.
not only increases accuracy but makes the process easier for anyone, including non-professionals, to use statistics to determine the most fundamentally, correct choices.

Disclosure statement
No potential conflict of interest was reported by the authors.

References
Abdikan, S., Bilgin, G., Sanli, F.B., Uslu, E., & Ustuner, M. (2015). Enhancing land use classification with fusing dual-polarized TerraSAR-X and multispectral RapidEye data. Journal of Applied Remote Sensing, 9, 096054. doi:10.1117/1.JRS.9.096054
Adam, E., Mutanga, O., Odinji, J., & Abdel-Rahman, E.M. (2014). Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: Evaluating the performance of random forest and support vector machines classifiers. International Journal of Remote Sensing, 35, 3440–3458. doi:10.1080/01431161.2014.903435
Adelabu, S., Mutanga, O., & Adam, E. (2014). Evaluating the impact of red-edge band from RapidEye image for classifying insect defoliation levels. ISPRS Journal of Photogrammetry and Remote Sensing, 95, 34–41. doi:10.1016/j.isprsjprs.2014.05.013
Akar, Ö., & Gungör, O. (2015). Integrating multiple texture methods and NDVI to the random forest classification algorithm to detect tea and hazelnut plantation areas in northeast Turkey. International Journal of Remote Sensing, 36(2), 442–464.
Behrend, T.S., Sharek, D.J., Meade, A.W., & Wiebe, E.N. (2011). The viability of crowdsourcing for survey research. Behavior Research Methods, 43(3), 800.
Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K.V., & Shirk, J. (2009). Citizen science: A developing tool for expanding science knowledge and scientific literacy. BioScience, 59(11), 977–984.
Carlotto, M.J. (2009). Effect of errors in ground truth on classification accuracy. International Journal of Remote Sensing, 30(18), 4831–4849.
Carrão, H., Gonçalves, P., & Caetano, M. (2008). Contribution of multispectral and multitemporal information from MODIS images to land cover classification. Remote Sensing of Environment, 112(3), 986–997.
Congalton, R.G., & Green, K. (1999). Assessing the accuracy of remotely sensed data: principles and applications. Boca Raton, FL: Lewis Publishers.
Congalton, R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, 37(1), 35–46.
Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20, 273–297. doi:10.1007/BF00994018
Deilmai, B.R., Ahmad, B.B., & Zabihi, H. (2014). Comparison of two classification methods (MLC and SVM) to extract land use and land cover in Johor Malaysia. In IOP Conference Series: Earth and Environmental Science (Vols. 20, No. 1, pp. 012052). https://doi.org/10.1088/1755-1315/20/1/012052.
ENVI (2008). Online Help, Applying Support Vector Machine Classification. Available online at: https://www.harrisgeospatial.com/docs/backgroundsvm.html (last accessed: 03.01.2019)
Fauvel, M., Chanussot, J., & Benediktsson, J.A. (2006, May). Evaluation of kernels for multiclass classification of hyperspectral remote sensing data. In ICASSP 2006, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 14–19 May 2006, Toulouse, France. IEEE, New York. Vol. 2, pp. 813–816.
Goodchild, M.F. (2007). Citizens as sensors: The world of volunteered geography. GeoJournal, 69(4), 211–221.
Howe, J. (2006). The rise of crowdsourcing. Wired Magazine, 14(6), 1–4.
Howe, J. (2008). CROWDSOURCING Kalabalıkların gıcır, bir kişinin geleceğine nasıl şekil verebilir? İstanbul: “pp. 121-122” KoşSistem Yayınları.
Hsu, C.W., Chang, C.C., & Lin, C.J. (2003). A practical guide to support vector classification. Available at: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf.
Hu, Z., & Wu, W. (2012, October). A satellite data portal developed for crowdsourcing data analysis and interpretation. In E-Science (e-Science), 2012 IEEE 8th international conference on e-Science (pp. 1–8). Chicago, Illinois: IEEE.

Huang, C., Davis, L.S., & Townshend, J.R.G. (2002). An assessment of support vector machines for land cover classification. International Journal of Remote Sensing, 23, 725–749. doi:10.1080/01431160110040323

Kahraman, F., Huroglu, C., Imamoglu, M., Ozcan, B.Y., Kalkan, M.I., Hocaoglu, M.A., & Kurt, B. (2015, May). Enterprise crowdsourcing platform for geospatial image analysis. In Signal processing and communications applications conference (SIU), 2015 23th (pp. 1813–1816). Malatya, TR: IEEE.

Karathanassi, V., Kolokoussis, P., & Ioannidou, S. (2007). A comparison study on fusion methods using evaluation indicators. International Journal of Remote Sensing, 28(10), 2309–2341.

Kavzoglu, T., & Colkesen, I. (2009). A kernel functions analysis for support vector machines for land cover classification. International Journal of Applied Earth Observation and Geoinformation, 11(5), 352–359.

Keerthi, S., & Lin, C.J. (2003). Asymptotic behaviors of support vector machines with Gaussian kernel. Neural Computation, 15, 1667–1689.

Kloumas, S., & Ehlers, M. (2009, July). Performance of evaluation methods in image fusion. In Information Fusion, 2009. FUSION’09. 12th international conference on (pp. 1409–1416). Seattle, WA: IEEE.

Krumm, J., Davies, N., & Narayanawasami, C. (2008). User-generated content. IEEE Pervasive Computing, 7(4), 10–11.

Kuemmerle, T., Chaskovsky, O., Knorn, J., Rodeloff, V.C., & Kruhlov, I. (2009). Forest cover change and illegal logging in the Ukrainian Karpathians in the transition period from 1988 to 2007. Remote Sensing of Environment, 113, 1194–1207.

Laben, C.A., Bernard, V., & Brower, W. (2000). Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. US Patent 6011875 A.

Li, C., Liu, L., Wang, J., Zhao, C., & Wang, R. (2004, September). Comparison of two methods of the fusion of remote sensing images with fidelity of spectral information. In Geoscience and remote sensing symposium, 2004. IGARSS’04. Proceedings. 2004 IEEE International (Vol. 4, pp. 2561–2564). Anchorage, AK: IEEE.

Li, D.C., & Liu, C.W. (2010). A class possibility based kernel to increase classification accuracy for small data sets using support vector machines. Expert Systems with Applications, 37(4), 3104–3110.

Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. IEEE Transactions on Geoscience and Remote Sensing, 42(8), 1778–1790.

Mondal, A., Kundu, S., Chandniha, S.K., Shukla, R., & Mishra, P. (2012). Comparison of support vector machine and maximum likelihood classification technique using satellite imagery. International Journal of Remote Sensing and GIS 1 (2): 116–123.

Nagub, A., Farag, M., Yahia, M., Ramadan, H., & Abd Elwahab, M. (2009). Comparative study between support vector machines and neural networks for lithological discrimination using hyperspectral data. Egypt Journal of Remote Sensing and Space Science, 12, 27–42.

Oommen, T., Misra, D., Twarakavi, N.K., Prakash, A., Sahoo, B., & Bandopadhyay, S. (2008). An objective analysis of support vector machine based classification for remote sensing. Mathematical Geosciences, 40, 409–424.

Padwick, C., Deskervich, M., Pacifici, F., & Smallwood, S. (2010April). WorldView-2 pan-sharpening. In Proceedings of the ASPRS 2010 annual conference (Vol. 2630). San Diego, CA, USA.

Pal, M., & Mather, P. (2005). Support vector machines for classification in remote sensing. International Journal of Remote Sensing, 26, 1007–1011.

Pal, M., & Mather, P.M. (2006). Some issues in the classification of DAIS hyperspectral data. International Journal of Remote Sensing, 27(14), 2895–2916.

Pasolli, E., Melgani, F., Alajlan, N., & Conci, N. (2013). Optical image classification: A ground-truth design framework. IEEE Transactions on Geoscience and Remote Sensing, 51(6), 3580–3597.

Petroopoulos, G.P., Kalaizidizis, C., & Vadrevu, K.P. (2012). Support vector machines and object-based classification for obtaining land-use/cover cartography from Hyperion hyperspectral imagery. Computers & Geosciences, 41, 99–107.

Saraoğlu, E., Görümş, E.T., & Gümgor, O. (2016, May). Mineral exploration with hyperspectral image fusion. In Signal processing and communication application conference (SIU), 2016 24th (pp. 1281–1284). Zonguldak, TURKEY: IEEE.

Schuster, C., Förster, M., & Kleinschmit, B. (2012). Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data. International Journal of Remote Sensing, 33, 5583–5599. doi:10.1080/01431161.2012.666812

Schuster, C., Schmoldt, T., Conrad, C., Kleinschmit, B., & Förster, M. (2015). Grassland habitat mapping by intraannual time series analysis – Comparison of RapidEye and TerraSAR-X satellite data. International Journal of Applied Earth Observation and Geoinformation, 34, 25–34. doi:10.1016/j.jag.2014.06.004

See, L., Mooney, P., Foody, G., Bastin, L., Comber, A., Estima, J., ... Liu, H.Y. (2016). Crowdsourcing, citizen science or volunteered geographic information? The current state of crowdsourced geographic information. ISPRS International Journal of Geo-Information, 5(5), 55.

Thenkabail, P.S. (Ed.). (2015). Remotely sensed data characterization, classification, and accuracies (Remote Sensing Handbook). Boca Raton, US: CRC Press.

Turner, A. (2006). Introduction to neogeography. Sebastol, CA: O’Reilly Media, Inc.

Ustunar, M., Sanli, F.B., & Dixon, B. (2015). Application of support vector machines for landuse classification using high-resolution RapidEye images: A sensitivity analysis. European Journal of Remote Sensing, 48(1), 403–422.

Vapnik, V.N. (1995). The nature of statistical learning theory. New York: Springer. doi:10.1007/978-1-4757-2440-0

Wald, D.J., Quattoriano, V., Worden, C.B., Hopper, M., & Dewey, J.W. (2012). USGS “Did You Feel It?” internet-based macroseismic intensity maps. Annals of Geophysics, 54(6), 688–707, http://dx.doi.org/10.4401/ag-5354.

Yu, L., Porwal, A., Holden, E.-J., & Dentith, M.C. (2012). Towards automatic lithological classification from remote sensing data using support vector machines. Computers & Geosciences, 45, 229–239.

Yuhendra, J., & Kuze, H. (2011). Performance analyzing of high resolution pan-sharpening techniques: Increasing image Quality for Classification using supervised kernel support vector machine. Research Journal of Information Technology, 8(1), 12–28.