EARTH OBSERVATION AND GEOMATICS ENGINEERING
website: https://eoge.ut.ac.ir

Evaluating the Capability of Spatial and Spectral Fusion in Land-Cover Mapping Enhancement

Mohammad Mansourmoghaddam [1], Iman Rousta [2]*, Hamid Reza Ghafarian Malamiri [2], Mohammad Hossein Mokhtari [3]

[1]Center for Remote Sensing and GIS researches, Shahid Beheshti University, Tehran, Iran
[2]Department of Geography, Yazd University, Yazd, Iran
[3]Environmental and Desert Studies School - Management in the Arid Regions, Yazd University, Yazd, Iran

Article history:
Received: 2022-03-21 , Received in revised form: 2022-04-02, Accepted: 2022-5-04

ABSTRACT

Most remote sensing satellites, such as Landsat, Spot, and QuickBird, produce many types of images, such as Multi-spectral (MS) and Panchromatic. Through the combination of high spatial and high spectral image characteristics, image fusion may generate a high spatial and high spectral resolution MS Image. The ability of Sentinel-2, optical Landsat-8, and their two types of fused images for Yazd, Iran, to improve the accuracy of the Land Cover (LC) map using supervised Maximum Likelihood Classification (MLC) was examined and quantified in this study. The present study, using the Gram-Schmidt Pan Sharpening fusion method, first focused on the spatial fusion of Landsat-8 images with Sentinel-2 images. Then the Sentinel spectral bands were also fused with Landsat's fused spectral bands and formed a new series of data. Finally, 4 data series of Landsat-8 30-m images (IM1), Sentinel-2 10-m images (IM2), spatial fused image (FIM1), and spectral-spatial fused image (FIM2) were classified using the MLC method and were evaluated. Also, the Normalized Optimum Index Factor (NOIF) index was developed based on the Optimum Index Factor (OIF) index and the number and combination of bands with the desired amount of information were examined. Results have shown that, because Landsat-8 has a higher spectral resolution than Sentinel-2 and Sentinel-2 has a higher spatial resolution than Landsat-8, their combination has boosted the information of images used for classification. As a result, the ideal NOIF values have been defined as 0.6 to 1. Furthermore, as compared to the initial unfused image of Landsat-8, the combination might improve overall classification accuracy by 10% and the Kappa coefficient by 16.5%. Also, the list of ideal band combinations with NOIF greater than 0.6 has been reported to aid researchers in Yazd in doing their categorization more properly and quickly.

KEYWORDS

Image Fusion,
Optimal Index Factor,
Normalized Optimal Index Factor,
Classification Accuracy,
Statistical Remote Sensing

1. Introduction

Various kinds of images such as Multi-spectral (MS) and Panchromatic are derived from most remote sensing satellites including Landsat, Spot, and QuickBird. While MS images provide a better spectral resolution which is useful to recognize landcover types, Panchromatic images provide a better spatial resolution which is helpful to identify the shape and texture of objects (Huang et al., 2015; Pohl & Van Genderen, 1998; Zare Naghadehi et al., 2021). Image fusion can produce a high spatial and high spectral resolution MS Image through the combination of high spatial and high spectral image features (Li & Li, 2010). It is important why helps the classification of images by improving the diagnosis of the land cover (LC) in images. Component substitution (CS) or Multiresolution Analysis (MRA) approaches are used in traditional image fusion methods (Palsson et al., 2013). Transformations, such as Principal Component Analysis (PCA), are used in CS processes (Chavez et al., 1991; Olsen et al., 2012), or even a spectrum transformation,

* Corresponding author
E-mail addresses: irousta@yazd.ac.ir (I. Rousta)

such as the IHS (intensity-hue-saturation) transition (CARPER et al., 1990; Huang et al., 2015). After a component generated from the lower-resolution multispectral band is replaced with a component derived from the higher-resolution band, the fused bands are produced by inverse transformation (Gašparović & Jogun, 2018). In MRA, methods like high-pass filtering are frequently utilized (Chavez et al., 1991; Zheng et al., 2017). Mixed fusion techniques are based on both MRA and CS (Palsson et al., 2013), as well as geostatistical approaches like Kriging with external drift (Sales et al., 2012) and area-to-point regression Kriging (ATPRK) (Wang et al., 2016), are more recent.

Mapping land use/land cover (LU/LC) cover in large cities is a critical component in recognizing significant shifts from non-industrial to industrial regions (Mansourmoghaddam et al., 2021). Also, mapping LU/LC plays an important role in environmental analysis and assessment, as well as LU/LC preparation and decision-making (Asadi et al., 2022; Lu et al., 2012). Remote sensing is a useful tool for analyzing changes in LU/LC in features like crops, forests, and agriculture. In remote sensing and other technology fields such as medical imaging and night vision, image fusion is a hot topic of study (Keshavarzi, Sarmadian, Tirado-Corbal, et al., 2010; Li et al., 2017; Mansourmoghaddam, Rousta, et al., 2022; Pohl & Van Genderen, 2016). Combining radar and optical data increases LU/LC classification precision by combining complementary knowledge from two radically separate technologies (Keshavarzi, Sarmadian, Sadeghnejad, et al., 2010; Maleki et al., 2020; Mansourmoghaddam, Ghafarian Malamiri, Arabi Aliabad, et al., 2022; Nuthammachot & Stratoulias, 2019; Pakdaman, 2013; Sukawattanavijit & Chen, 2015). Sentinel is a new European Space Agency project to meet the Copernicus program's operating requirements. Sentinel-2 is a project for the land observation that uses polar-orbiting multispectral high-resolution images (LU/LC tracking or transition detection) and also can provide data to emergency responders (Copernicus, 2015; Gašparović & Jogun, 2018). Sentinel-2 features a 13-spectral-band optical multispectral sensor payload: 4 bands with a spatial resolution of 10 meters, 6 bands with a spatial resolution of 20 meters, and 3 bands with a spatial resolution of 60 meters (Richter et al., 2011). The Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) instruments are carried by Landsat-8, the most recently deployed Landsat satellite. OLI has 8 bands with a spatial resolution of 30 meters and 1 band with a spatial resolution of 15 meters and TIRS has 2 bands with a spatial resolution of 100 meters (Mansourmoghaddam, Ghafarian Malamiri, Rousta, et al., 2022; USGS).

Attempts to combine satellite data have been made in several experiments. Sukawattanavijit (2015) used RADARSAT 2 and Landsat-8 satellite images with wavelet-based fusion to classify ground cover in Wang Muang district in Saraburi Province, central Thailand. When comparing the single radar images to the fused image, the results revealed that the fused image had better precision. When Dimov et al. (2016) examined crop areas in Uzbekistan's Fergana Valley, they discovered that Ehlers fusion is the best method. Chen et al. (2017) investigated the efficacy of multi-source remotely sensed data fusion in the context of LC classification optimization. In comparison to the original data sources, they discovered that The fused data improved LC classification mapping accuracy by combining temporal, spectral, angular, and topographic features. Salman et al. (2017) showed that using machine vision, a fused dataset provided 99.6% precision in LU/LC classification. Puttinaovarat and Horkaew (2017) studied the detection of building structures from fused satellite images using spectral indices rather than reflectance bands. Nuthammachot and Stratoulias (2019) used Sentinel 1 and Landsat-8 fused images to classify LC in Thailand. The findings show that combining optical and SAR satellite imagery at a higher resolution will increase the understanding and classification precision of LC and land use forms in the study field.

Because of the technical limitations of satellite sensors, they cannot support all optimal spectral and spatial identification for the classification of land cover. This is especially important when working in urban areas. More variety and smaller urban components require more precision to identify the right class of each pixel. This study first introduces a new form of Optimum Index Factor (OIF), the Normalized Optimum Index Factor (NOIF) to compare these values between different images. Despite the spectral proximity of the two Landsat and Sentinel sensors, they are complementary in terms of the number of bands and spatial power. By considering the importance of LC management and assessment for Yazd because of its global and central location and also, the lack of resources to study the effect of image integration to improve LC classification in the city; this study aims to assess and quantize the capability of SAR Sentinel-2, optical Landsat-8 (Table 1), and their two types of fused images in Yazd, Iran, for improvement of accuracy LC map through Maximum likelihood supervised classification.

work.

2. The Study Area

The city of Yazd is one of the metropolises of Iran and is the capital of the province and city of Yazd. The city is located between longitude 54° and 22' east and 31° and 53' north latitude (Figure 1). The area of this city is equal to 110 km2. The average altitude of Yazd is 1228 meters above sea level and the average annual temperature of this city is 20 degrees Celsius. The city of Yazd has two historical and new contexts that the components of the historical context of this city of Yazd, including buildings, blocks, and uses of the city

are intertwined and follow a hierarchical system (BEHZADFAR & NOURMOHAMMADZAD, 2011).

3. Materials and Methodology

In order to quantify the fusion effect of Landsat-8 images by Sentinel-2 images, an image of Landsat and an image of Sentinel were selected from the city of Yazd at close intervals (Table 1). After performing radiometric and atmospheric corrections on the image, first, the 30-meter visible bands of Landsat-8 based on Sentinel-2 images are fused to 10 meters (spatial fusion), then the 10-meter spectral bands of Sentinel-2 are added to the fused set (spectral fusion). Four image sets including the 30-meter Landsat-8 6-band image are called image-1 (IM1), the 10-meter Sentinel-2 4band image is called image-2 (IM2), the 10-meter fused Landsat-8 6-band is called fused image-1 (FIM1), and the 10-meter fused Landsat8-Sentinel-2 10-band (including 6 Landsat bands and 4 Sentinel bands) image is called fused image-2 (FIM2) then have arrived. Then the statistics of each image were calculated based on Normalized Optimum Index Factor (NOIF). Finally, after classifying each of the images in the Maximum likelihood method, it compares the LC classification accuracy of each of the images (Figure 2)

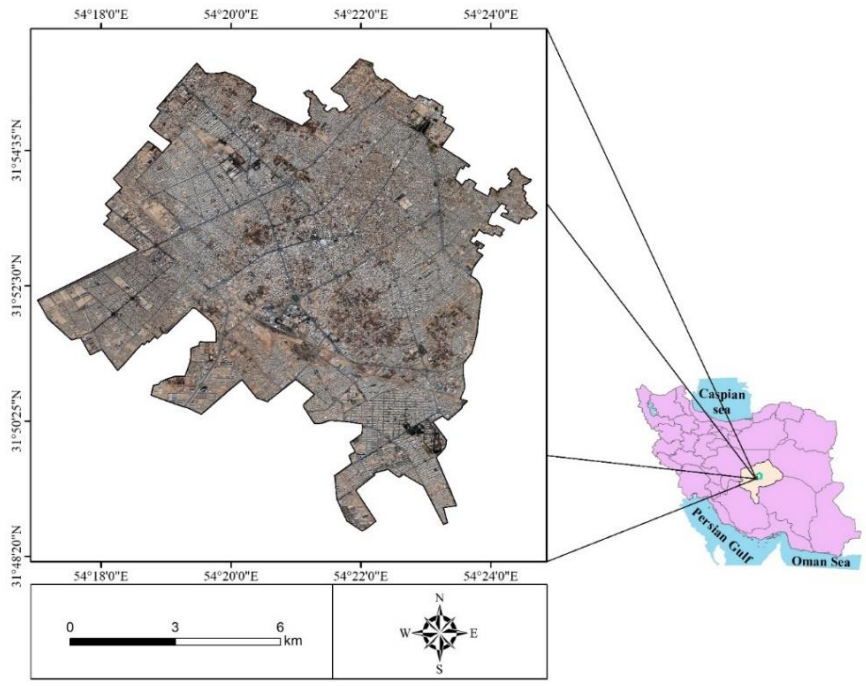


Figure 1. The study area, Yazd city, from the perspective of Bing satellite (2019) and its position in Iran

Table 1. Characteristics of images used

Satellite	Spatial resolution	Wavelength (µm.) (eesa; USGS)	Acquisition date	Acquisition time (GMT)
Landsat-8	30m bands	Blue: 0.45 Green: 0.53 Red: 0.64 NIR[1]: 0.85 SWIR[2]1: 1.57 SWIR2: 2.11	20200804	06:56:53
Sentinel-2-A	10m bands	Blue: 0.49 Green: 0.55 Red: 0.66 NIR: 0.83	20200819	06:56:31

[1] Near Infrared
[2] Short Wave Infrared

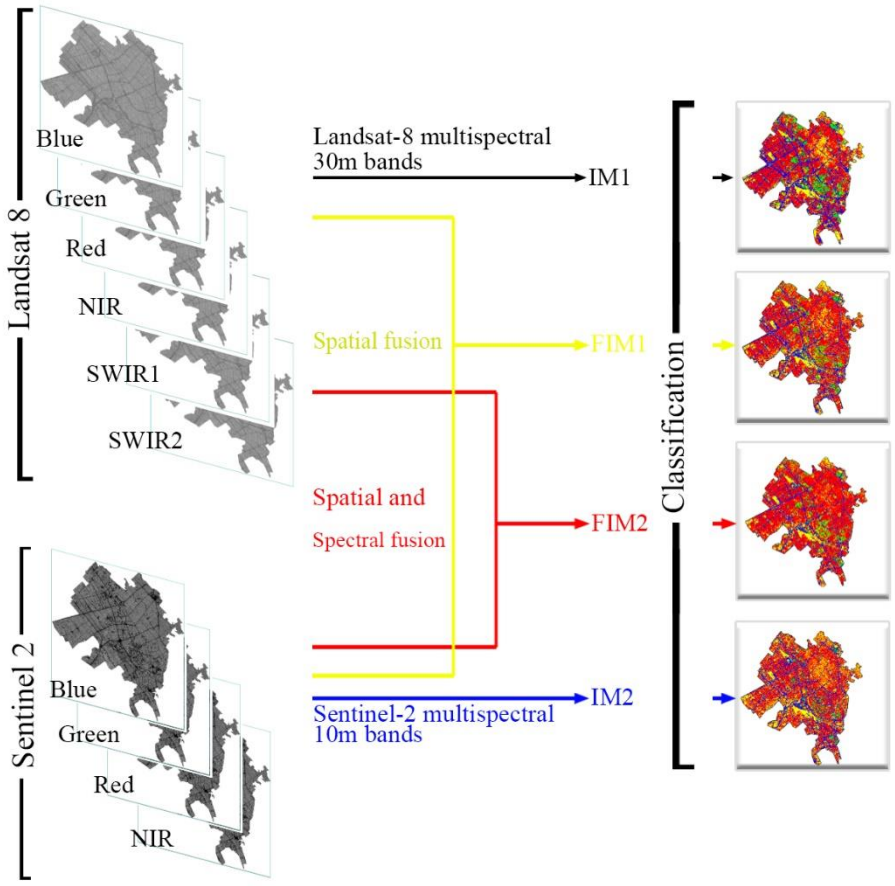


Figure 2. The flowchart of the study

3.1. Pan-Sharpening Technique

Let The Gram-Schmidt pan-sharpen procedure is one of the most widely used high-quality image sharpening processes; that's why companies including Esri, ENVI, and others use the Gram-Schmidt approach in their product sets (Maurer, 2013). The process procedure is as follows:

1- As a linear combination of the n MS (Multi Spectral) bands, create a virtual low-resolution Pan band (Equation (1)) (Maurer, 2013):

$$Pan_{sim} = \sum_{k=1}^{n} W_k MS_k \quad (1)$$

Where n is the number of MS bands, and k and W are the counter and weights of each of MS band.

2- By considering each band as a high-dimensional vector and starting with the simulated pan band as the first vector, render all bands orthogonal using the Gram-Schmidt vector orthogonalization, or a modified variation of it. Both the incoming bands and the claims of the scalar products are first rendered dc-free for the Gram-Schmidt pan sharpening (subtraction of their mean). In this method, the initial Gram-Schmidt scalar products are transformed into covariances. The iterative process remains unchanged: calculate the angle between both the Red and Pan bands, then move the Red band until it is perpendicular to the Pan band. Calculate the angles between the Green band and the Pan band and the rotated Red band in the next step, then rotate the Green band such that it is orthogonal to both the Pan band and the rotated Red band. The list goes on. The bands are de-correlated using the Gram-Schmidt forward transform (Maurer, 2013).

3- Replacement of the low-resolution virtual Pan band with a high-resolution Pan band that has been gain and bias modified. All MS bands should be upsampled accordingly (Maurer, 2013).

4- Use the same transform coefficients to backward the Gram-Schmidt transform, except just on the high-resolution bands. The pan-sharpened image in high resolution is the product of this reversed forward Gram-Schmidt transform. (To read more, refer to: (Laben & Brower, 2000; Maurer, 2013)).

3.1.1 Accuracy Assessment

To assess the accuracy of the fusion results, two images FIM1 and FIM2 were checked using the mean square error (RMSE). RMSE is one of the approaches to describe spatial cross-validation. It can evaluate spatial interpolations. Because it is the most often reported and misconstrued of the

three average-error statistics, the RMSE is of particular relevance (Willmott & Matsuura, 2006)

3.2. NOIF Index

OIF is a mathematical calculation of all potential three bands in R-G-B format (Qaid & Basavarajappa, 2008). Chavez et al. are the creators (Chavez et al., 1982). OIF values were calculated to identify the most advantageous band combination (Cengiz et al., 2006) and rate the band subsets based on their detail (Beauchemin & Fung, 2001). It is dependent on the overall variance and the correlation between different bands (Jensen, 1996). The OIF is calculated using the following algorithm for any subset of three bands (Equation (2)) (Qaid & Basavarajappa, 2008):

$$OIF = \left[\frac{\sum_{i=1}^{n}\sigma(i)}{\sum_{j=1}^{n}|r(j)|}\right] \quad (2)$$

where $\sigma(i)$ is the k band's standard deviation and $r(j)$ is the value of the correlation matrix.

A high OIF value means that the bands contain a lot of detail (e.g., a high standard deviation) and less duplication (e.g., low band correlation) (Buhe et al., 2007). To compare the OIF value of each level of fusion, this study has normalized the OIF index using Equation (3):

$$NOIF = \frac{oif - oif_{min}}{oif_{max} - oif_{min}} \quad (3)$$

Where NOIF values greater than 0.6 (more than 0.5 as the mean) have been considered optimal OIF and NOIF lower than 0.1 have been ignored.

3.3. Maximum Likelihood Classification Method

The maximum likelihood classification (MLC) algorithm is one of the most widely used parametric classification methods (Colditz et al., 2006). d spectral components, which are independent Gaussian random variables, describe each land cover class, according to this procedure (Richards & Richards, 1999). Data is believed to be distributed to a predefined probability model for parametric classifiers. The training samples are used to calculate the parameters for this distribution. The likelihood is maximized for the classification norm. To measure the likelihood for each class, it is assumed that a multivariate gaussian distribution with mean vector μ_k and covariance matrix S_k exists (Colditz et al., 2006). The following classification rule is used to approximate the pixel vector S_k's class label k (Equation 4) (Colditz et al., 2006):

$$k = \ln(S_k) - [(x_i - \mu_k)^T S_k (x_i - \mu_k)] \quad (4)$$

Due to the optimum statistical representation of the class, a single training class corresponds to each training sample for the classification procedure. The training classes are recoded into four ground cover classes after classifying.

3.3.1 Accuracy Assessment

The user, the producer, as well as the Kappa coefficient, and overall accuracy were used to assess classification precision. The confusion or error matrix, which is commonly used in classification accuracy estimation, can accurately be used to approximate the kappa coefficient. The primary emphasis would be on the simplest case of a binary confusion matrix, which is commonly used in analyses of land cover transition, for example (Foody, 2020). The Kappa test is a nonparametric statistic that was used to determine how well user-assigned and predefined values were agreed (Ishtiaque et al., 2017). Kappa is a calculation efficiency index for binary characteristics that are commonly used. When the agreement is fine, meaning that the raters agree in their classification in the event, the Kappa coefficient is 1.0 (100 percent), and when the agreement is no greater than that predicted by chance, it is 0.0. It may also be negative, implying that there is less consensus than would be predicted provided the marginal distributions of ratings by chance (Thompson & Walter, 1988).

The accuracy tests, called overall accuracy, user accuracy, and producer accuracy were performed using Equations (5) to (7) respectively (Bokaie et al., 2016; Pal & Ziaul, 2017).

Divide the total number of correctly categorized pixels by the total number of pixels in the error matrix to get the overall accuracy of the classification map (Jensen, 1996) (Equation 5).

$$Overall\ Acc. = \left\{\frac{\sum CCP(Diagonal)}{\sum CRP} * 100\right\}, \quad (5)$$

where $CCP(Diagonal)$ is the corrected classified pixels (diagonals) and CRP is the corrected reference pixels.

The result is a measure of commission error, which is calculated by dividing the total number of right pixels in a class by the total number of pixels that were currently included in that group. This metric is known as the user's consistency or dependability (Jensen, 1996) (Equation 6).

$$User\ Acc. = \left\{\frac{\sum CCP(Category)}{\sum CPC\ (Row)} * 100\right\}, \quad (6)$$

where $CCP(Category)$ is the corrected classified pixels (category) and $CPC(Row)$ is the classified pixels in that category (the row total).

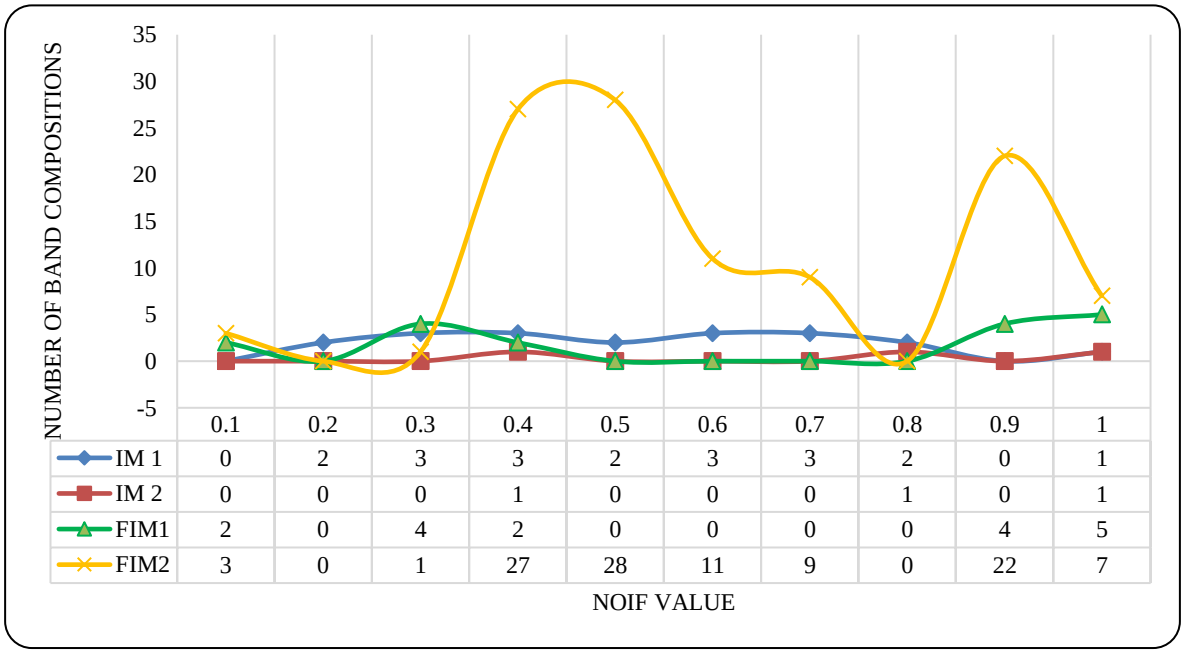
The producer's accuracy is defined as the total number of right pixels in a class divided by the total number of pixels in that class as calculated from ground reference data. This metric is an indicator of omission error which shows the likelihood of a reference pixel being accurately identified (Jensen, 1996) (Equation 7).

$$Producer\ Acc. = \left\{\frac{\sum CCP(Category)}{\sum CPC\ (Column)} * 100\right\}, \quad (7)$$

$CPC(Column)$ are the classified pixels in that category (the column total), and $CCP(Category)$ is the corrected classified pixels (diagonals).

4. Results and Discussion

To evaluate the accuracy of fusion results, their RMSE statistics were compared. All three images matched each other with an error of less than 0.0000. The results of the comparison of four sets of images before and after fusing have shown that Landsat bands improved for classification based on NOIF results. In this way, while IM1 had most of its bands between 0.2 and 0.8 of NOIF value i.e 18 of 20 possible triple-band compositions (90% of all), and only 1 band composition (5% of all) with perfect NOIF (1), it has optimally changed to 2 band compositions (10% of all) with NOIF 0.1, 4 compositions (20% of all) with NOIF 0.3, 2 compositions (10% of all) with NOIF 0.4, 4 compositions (20% of all) with NOIF 0.9 and a significant number of 5 compositions (25% of all) with perfect NOIF (1) when spatially fused to 10-meters (FIM1). It is probably because the Landsat improved spatial resolution bands help thanks to Sentinel-2 10 meters bands. Although IM2 had just 4 compositions, one with NOIF 0.4, one with 0.8, and one with perfect NOIF (1) (one in 0 has been ignored), its spectral combination with FIM1, Helped in achieving a statistically significant 3 combinations (3% of all) with NOIF 0.1, 1 combination (1% of all) with NOIF 0.3, 27 combinations (23% of all) with NOIF 0.4, 28 combinations (23% of all) with NOIF 0.5, 11 combinations (9% of all) with NOIF 0.6, 9 combinations (8% of all) with NOIF 0.7, 22 combinations (18% of all) with NOIF 0.9 and an increased number of perfect NOIF (1) to 7 combinations (6 of all) in FIM2 (Figure 3). It shows although Sentinel-2 and Landsat-8 4 first bands are almost similar in terms of wavelength, minor differences in wavelength (Table 1) that improve the image radiometrically on the one hand, and the presence of 2 additional SWIR1 and SWIR2 bands in the Landsat on combination with these 10-meter images, on the other hand, have improved the information of the FIM2.



	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
IM 1	0	2	3	3	2	3	3	2	0	1
IM 2	0	0	0	1	0	0	0	1	0	1
FIM1	2	0	4	2	0	0	0	0	4	5
FIM2	3	0	1	27	28	11	9	0	22	7

Figure 3. Change in NOIF values of the 4 images' triple-band compositions before and after fusing

Since this paper considered NOIF +0.6 as the optimal OIF value, the change of the number of triple-band compositions with NOIF 0.6 from 3 compositions in IM1 to 11 compositions in FIM2, NOIF 0.7 from 3 to 9 compositions, NOIF 0.9 from 0 to 22 compositions, and finally perfect NOIF (1) from 1 to 7 compositions have derived statistically acceptable results from fusing. The increased number of optimal band compositions clearly says the increase of differences between bands and also more useful information for classification after fusion has been performed (Table 2 and Figure 4).

Table 2. Comparison of the number of the 4 images triple-band composition with NOIF +0.6 before and after fusing

	IM 1	IM 2	IM 3	IM 4
0/6	3	0	0	11
0/7	3	0	0	9
0/8	2	1	0	0
0/9	0	0	4	22
1	1	1	5	7

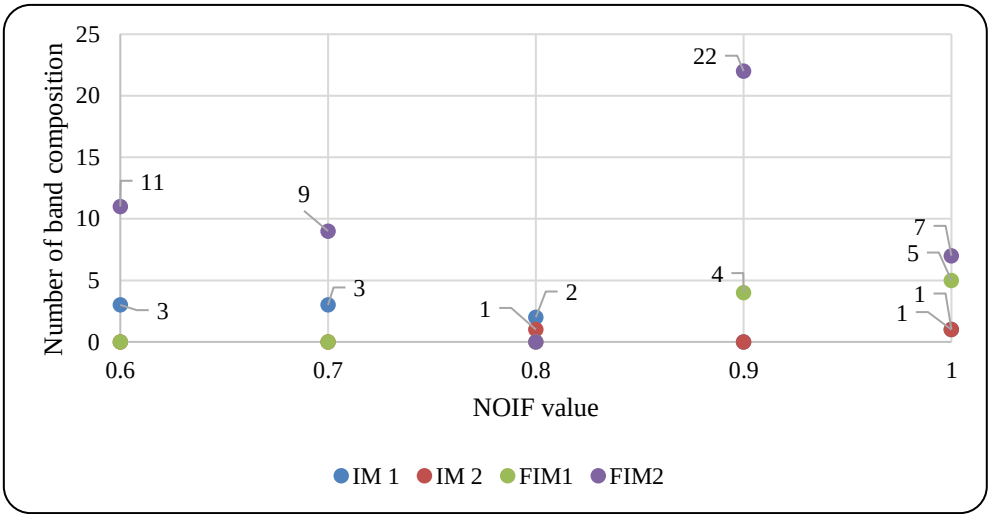


Figure 4. Schematic of the number of 4 images' triple-band compositions with NOIF +0.6 before and after fusing

Therefore, after fusing (FIM1 and FIM2), valuable band compositions with equal to/more than 0.6 (mean + 1) of NOIF for FIM1 and FIM2 have been delineated as shown in Table 3 and **Error! Reference source not found.**. As shown in Table 3, one of the infrared or thermal infrared bands is seen in all band combinations with NOIF +0.6. Furthermore, in all combinations with optimal NOIF (1), at least two of these bands are observed in the band composition, which indicates the importance and valuable information of these bands for image information extraction.

Table 3. Valuable band compositions for FIM1

Band Composition	NOIF	Band Composition	NOIF
2-3-5	0.9	2-6-7	1
2-3-6	0.9	3-5-6	1
2-3-7	0.9	3-5-7	1
2-5-7	0.9	3-6-7	1
2-5-6	1		

The tables **Error! Reference source not found.** shows the a mazing results of combining the Sentinel and Landsat bands to obtain higher NOIF values. More far bands (spectrally) have decreased the redundancy, increased the information (and consequently increase the NOIF). Thus, in band combinations with NOIF ≥ 0.6, one or both infrared and thermal infrared bands of Landsat can be seen, and in more than half of the cases, one of the Sentinel spectral bands is also observed. This is despite the fact that in all NOIFs ≥ 0.7, at least two sentinel bands are present in the band composition. But in NOIF ≥ 0.9, most of the bands that are far from each other are seen in combination with an infrared or thermal infrared band and a sentinel band; such as the blue or green band of Landsat in combination with an infrared or thermal infrared band of Landsat in combination with often one of the Sentinel green or infrared bands. At the optimal NOIF value, i.e. 1, the dominant band combinations mainly consist of a Landsat near-infrared band or close to that, a red band combined with a Landsat thermal infrared band or Sentinel blue band combined with a Sentinel green or red

band. These combinations represent different information that exists in each of these bands in order to combine and extract useful information from the images.

Table 4. Valuable band compositions for FIM2 (Note: Bands 1-6 are from Landsat-8 and Bands 7-10 are from Sentinel-2)

Band Composition	NOIF	Band Composition	NOIF
1-5-6	0/6	1-6-10	0/9
1-5-7	0/6	1-7-8	0/9
1-6-7	0/6	1-7-10	0/9
2-5-6	0/6	2-5-8	0/9
2-5-7	0/6	2-5-9	0/9
4-5-7	0/6	2-5-10	0/9
4-5-6	0/6	2-6-8	0/9
2-6-7	0/6	2-6-9	0/9
3-6-7	0/6	2-6-10	0/9
3-5-6	0/6	3-5-10	0/9
3-5-7	0/6	2-7-8	0/9
1-8-9	0/7	2-7-9	0/9
1-8-10	0/7	3-6-8	0/9
1-9-10	0/7	3-6-10	0/9
2-9-10	0/7	2-7-10	0/9
3-8-9	0/7	3-5-8	0/9
3-8-10	0/7	3-6-9	1/0
3-9-10	0/7	3-7-9	1/0
2-8-9	0/7	3-7-10	1/0
2-8-10	0/7	3-7-8	1/0
1-5-8	0/9	3-5-9	1/0
1-5-9	0/9	4-5-8	1/0
1-5-10	0/9	1-7-9	1/0
1-6-8	0/9	2-3-4	1/0
1-6-9	0/9		

The results of performing LC classification of Yazd city by maximum likelihood method (Figure 5) have shown that generally, the overall accuracy of classification has increased from 77.9% in IM1 to 88.6% in FIM2 (10.7%). Similarly, IM2 and FIM1 have resulted in 82.4% and 84.6% in overall accuracy respectively. The kappa coefficient results also have indicated a 26.5% increase from 62.2% in IM1 to 78.7% in FIM2. IM2 and FIM1 have experienced a kappa coefficient of 67.4% and 71.4% respectively (Figure 6). This confirms the role of image radiometric and spatial characteristics improvement through fusion in LC classification.

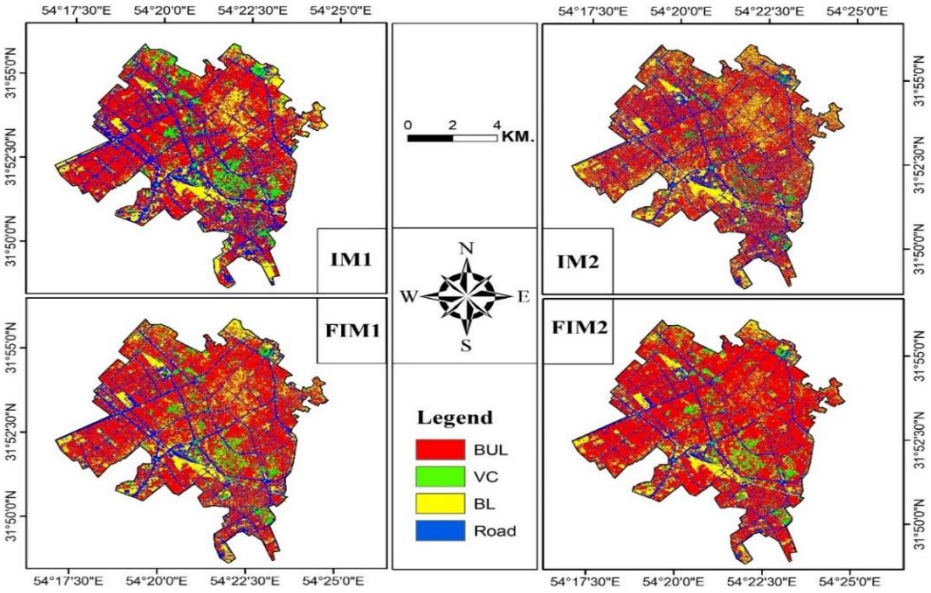


Figure 5. LC classification of Yazd city in four Built-up Land (BUL), Vegetation Cover (VC), Bare Lands (BL), and Road classes on the four types of images for 2020, Yazd city

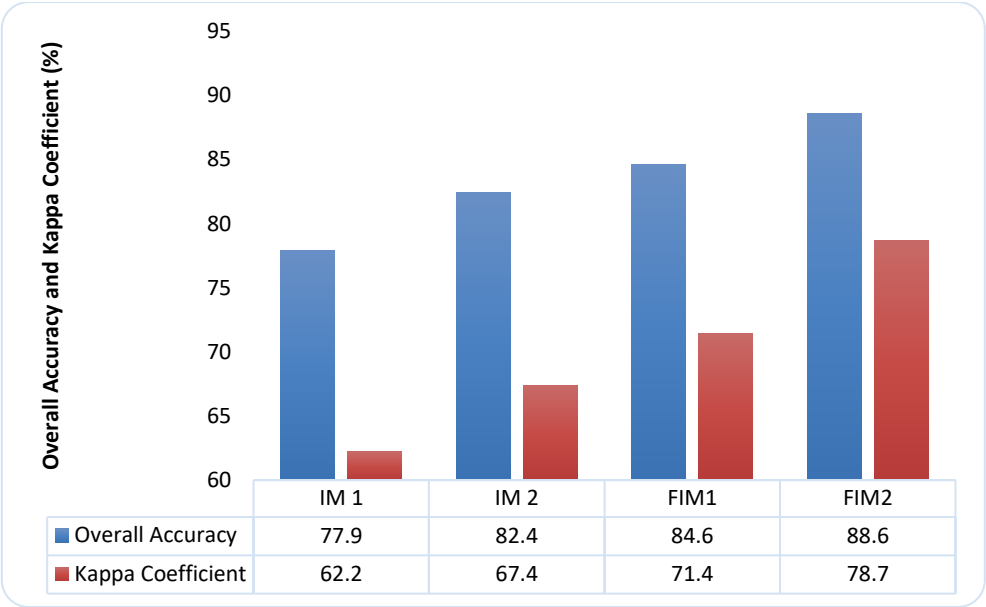


Figure 6. Statistically and schematically showed of Increase in Overall accuracy and Kappa coefficient before and after fusion.

The results of user accuracy have shown that except for BUL class, FIM2 has shown higher accuracy for all other classes. Thus, in comparison to IM1, BL user accuracy has increased from 81.3% to 91.8% (10.5% increase), VC has shown a 3.5% increase from 96.6% to 100%, and the road class user accuracy has increased significantly from 21.5% to 50% (28.5%) in FIM2. These results show the important effect of Landsat-8 and Sentinel-2 fusion on the improvement of LC classification accuracy. Although the BUL class user accuracy has increased from 83.3% in IM1 to 87.3% (4 %) in FIM2, it has shown its best user accuracy in IM2 with a little more (1.4% increase compared to FIM2) precision, 88.7% of user accuracy. By considering the high spatial power of 10-meter Sentinel-2 and the physical pattern of BUL, the slightly less accuracy of fused images can be because of the addition of some uncertainty by the fusion model to the image. Also, the user accuracy of BUL in FIM1 is a bit (0.6%) more than in FIM2. Furthermore, the results of IM2 were 9.4% better than FIM1 in the road class, also IM2, FIM1, and FIM2 were shown the same results (100% of user accuracy) for the VC class, which may be because the presence of high resolution (10-meters) red and near-infrared (NIR) bands of sentinel-2 which are the most important combination for vegetation identification due to red-edge existence in this area of wavelength (Figure 7).

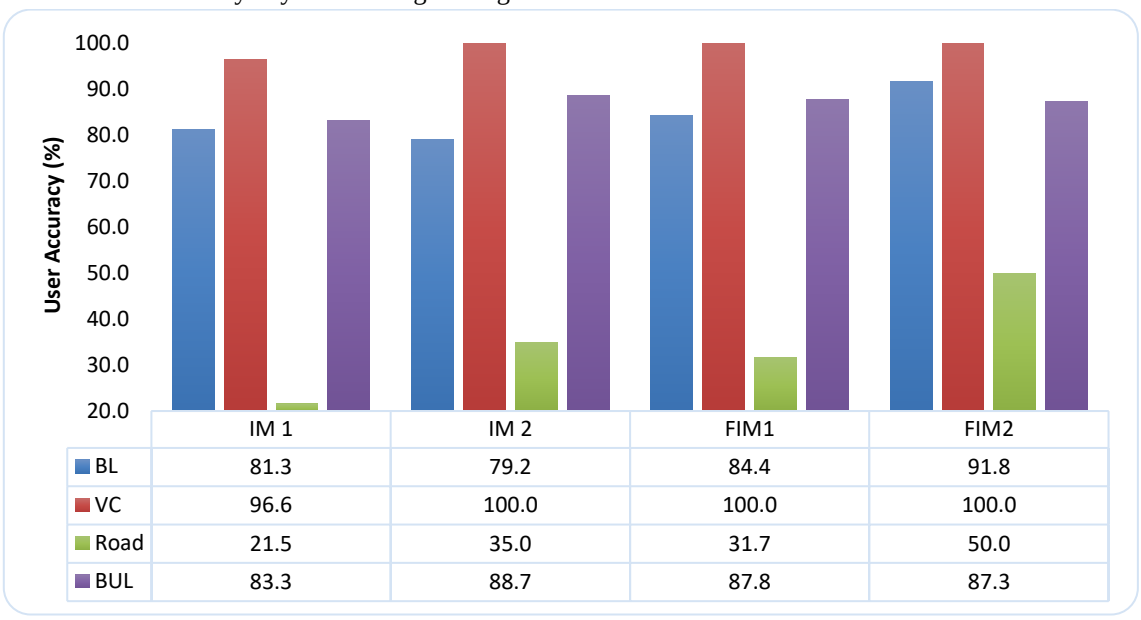
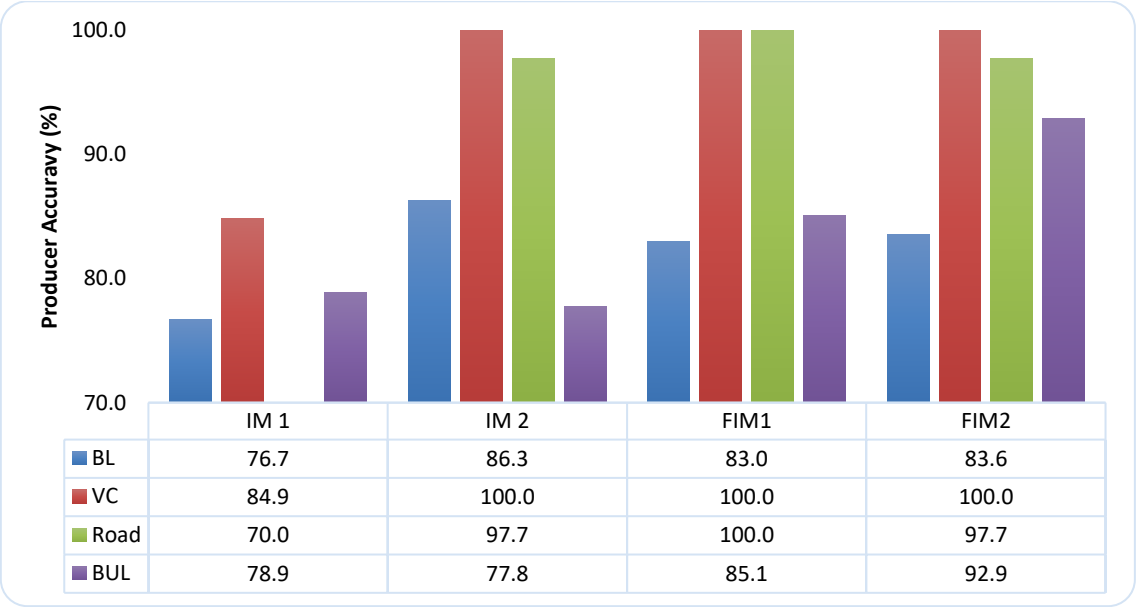


Figure 7. Comparison of user accuracy (%) before and after fusing

Producer accuracy statistics also have indicated that FIM2 was the most accurate image in the classification of VC (together with IM2 and 3) and BUL by 92.9% and 100% respectively. Although the producer accuracy of BL has risen from 76.7% in IM1 to 83.6% in FIM2 (6.9%), classification in IM2 has had the most accurate result based on producer

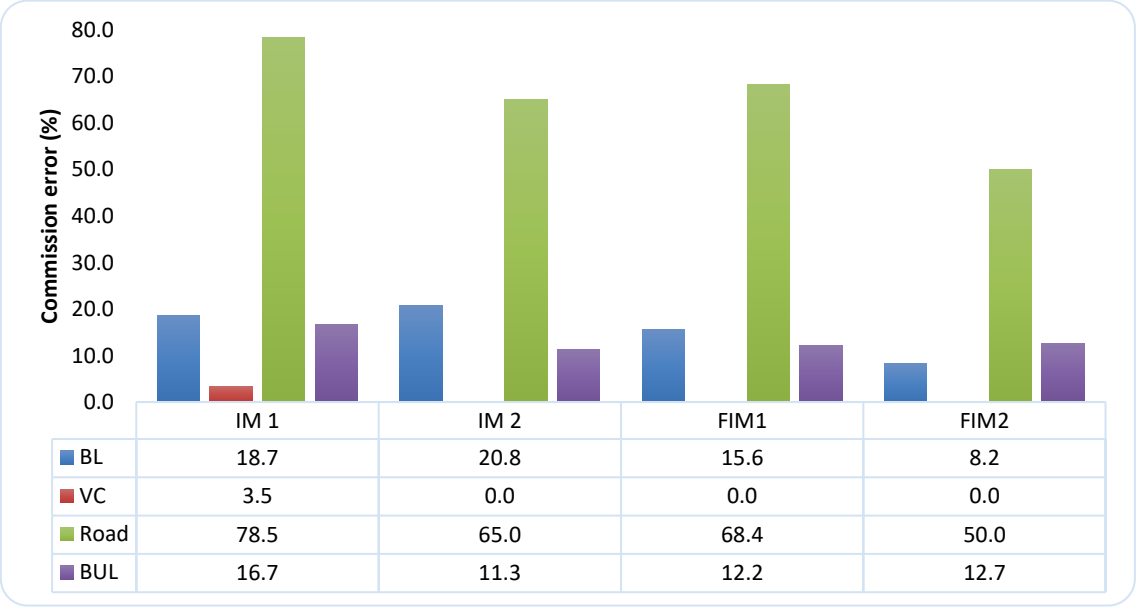
accuracy with 86.3%. Furthermore, FIM1 has had a bit (2.3%) more accurate result than FIM2 and IM2 with 100% of producer accuracy (Figure 8) in the road class.



	IM 1	IM 2	FIM1	FIM2
BL	76.7	86.3	83.0	83.6
VC	84.9	100.0	100.0	100.0
Road	70.0	97.7	100.0	97.7
BUL	78.9	77.8	85.1	92.9

Figure 8. Comparison of producer accuracy (%) before and after fusing

Overall, except for BUL, the commission error of all other LC classes had been at the minimum in FIM2 compared to other images. It confirms the effect of spatial and radiometric fusion of Sentinel-2 and Landsat-8 on reducing the error of classification. BL has had the most commission error in IM2 and the least in FIM2. This class has had a maximum error of 20.8% in IM2 and after that 18.7% in IM1. Its error has sharply decreased to 8.2% in FIM2. Also, FIM1 with a 15.6% of error, had reduced 16.6% the commission error by compared to IM1. The VC class has had just a 3.5% of error in IM1 and other images had not any error. While the road class has had a significant error in IM1 (78.5%) and FIM1 (68.4%), its error has dropped for FIM2 to 50%. In a different trend, BUL has had a minimum error of 11.3% in IM2. Its maximum error was in IM1 with 16.7% and FIM1 and FIM2 had the next lower error with 12.2% and 12.7% respectively (Figure 9). Again, as stated in the accuracy section, the physical pattern of the city has been recognized more accurately by sentinel-2 due to its 10-meter spatial resolution and non-intervention fusion uncertainty and errors.

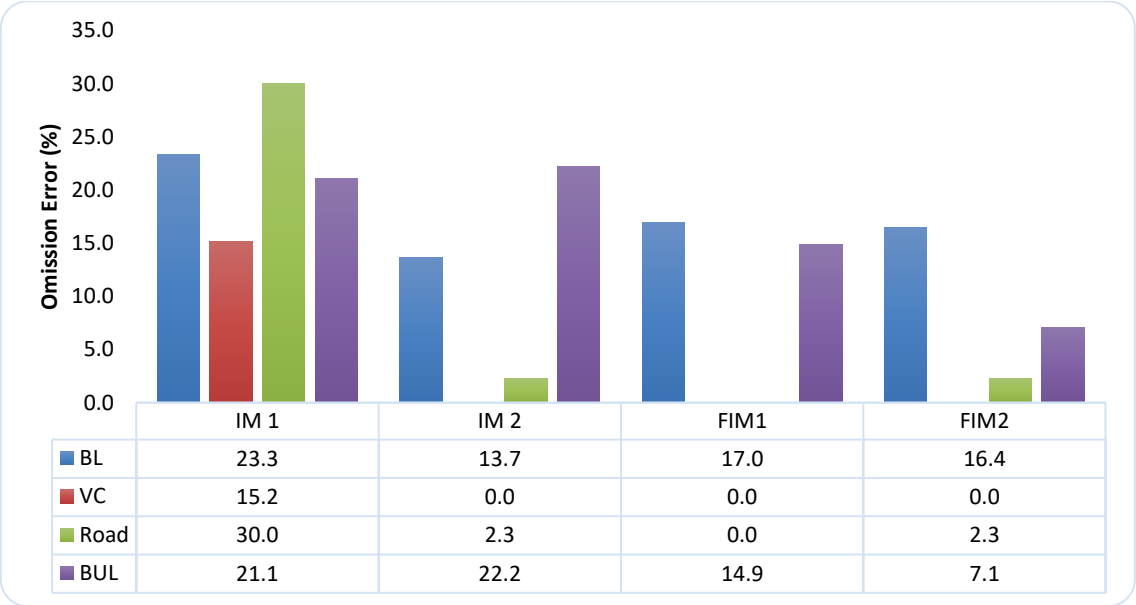


	IM 1	IM 2	FIM1	FIM2
BL	18.7	20.8	15.6	8.2
VC	3.5	0.0	0.0	0.0
Road	78.5	65.0	68.4	50.0
BUL	16.7	11.3	12.2	12.7

Figure 9. Comparison of the commission error (%) before and after fusing

The omission error results have shown the BL class has had a maximum and minimum error of 23.3% and 13.7% in IM1 and IM2 respectively. FIM2 has had the second minimum error of 16.4% and after that FIM1 with 17% of

error. In the same trend with commission error, the VC class has had just a 15.2% of omission error in IM1 and all other images have had no error as said in the accuracy section, the 10-m red and NIR have worked well to recognize the vegetation in Sentinel-2 images. The Road class has had a relatively significant error of 30% in IM1 as the highest error. On the other hand, it has had no error in FIM1 and it confirms the effect of spatial fusing on reducing the error. Also, IM2 and FIM2 with the same error of 2.3% are in second place. These cases indicate the higher spatial resolution required to improve the separation of this class. BUL as the only class in which FIM2 (with 7.1% error) could reduce a notable (14%) amount of error compared to IM1 (with 21.1% error) and 2 (with 21.1% error) has emerged as optimal. FIM1 with a 14.9% of error has placed second (Figure 10).



	IM 1	IM 2	FIM1	FIM2
BL	23.3	13.7	17.0	16.4
VC	15.2	0.0	0.0	0.0
Road	30.0	2.3	0.0	2.3
BUL	21.1	22.2	14.9	7.1

Figure 10. Comparison of the omission error (%) before and after fusing

5. Conclusion

This study aimed to evaluate and quantify the effects of image fusing of Landsat-8 and Sentinel-2 on the LC classification of Yazd city. to aim this goal the pan-sharpening fusion algorithm and maximum likelihood classification algorithm have been used. Results have shown that overall, because of the higher spectral resolution of Landsat-8 than sentinel-2 and the higher spatial resolution of Sentinel-2 than Landsat-8, their combination has increased the information of image uses for classification. These results are consistent with the results of Nuthammachot and Stratoulias (2019). In this manner, optimal NOIF values have been delineated from 0.6 to 1. Therefore, using the method of combining images of these two sensors to extract more information from the image by the classifier in the city of Yazd city is recommended. Results have shown that the number of bands that have values of 0.6 has increased from 3 to 11, 0.7 from 3 to 9, 0.9 from 0 to 22, and 1.0 (as perfect NOIF) has risen from 1 to 7. Furthermore, this study showed an increase in the accuracy of post-fusion classification results. These results can show the application of fusion methods for Landsat-8 and Sentinel-2 in increasing the accuracy of the classifier in Yazd city. The results of this section correspond to the results of Chen et al.'s (2017) research. Thus, the combination could increase the overall accuracy of classification by 10%, and also the Kappa coefficient by 16.5% compared to the first unfused image of Landsat-8. User and producer accuracies of classification also have increased for the BL: 10.4% and 6.9%, the VC: 3.5% and 15.2%, the road class: 28.5% and 27.7%, and for the BUL: 4% and 14% respectively after the spatial and spectral fusion has performed. Each classification contains some error, Therefore, reducing these errors is one of the necessary and appropriate measures in order to achieve a more accurate map. The commission, and Omission errors of the LC classification of Yazd after spatial and spectral fusing have dropped for BL: 10.4% and 6.9%, the VC: 3.5% and 15.2%, the road class: 28.5%, and 27.7%, and for the BUL: 4% and 14%. Therefore, the image fusion of the two sensors Sentile-2 and Landsat-8 was able to reduce classification errors well and increase the accuracy of the output map. The results are similar to Salman et al.'s (2017) and Sukawattanavijit's (2015) research. In order to further guide the use of the results of this study, a list of optimal band combinations which have NOIF of more than 0.6 has also been noted to help studies that work on Yazd to perform their classification more accurately and fastly. Due to the ease of access and free access of Sentinel and Landsat images, the authors suggest that in order to obtain newer and more complementary results, future research will compare the effect of different fusion methods on images of these two sensors and provide a desirable method for fusion to obtain

land-cover maps as accurately as possible. Like any work, fusion also has its limitations, such as pixel sizes cannot be fused more than a certain amount due to mathematical and spectral limitations. Also, in general, the more homogeneous the surface under the pixel, the better the accuracy of the fusion performance, and the more heterogeneous, the lower the accuracy. It is suggested that future studies perform the spatial fusion of Landsat-8 images with high-resolution images (such as airborne or UAV images) in combination with spectral fusion with more spectral bands such as Sentinel 2, and compare the results with the results of the present research in the field of increasing the accuracy of the land cover classification and the number of bands with high NOIF.

References

Asadi, M., Oshnooei-Nooshabadi, A., Saleh, S.-a., Habibnezhad, F., Sarafraz-Asbagh, S., & Van Genderen, J. L. (2022). Simulation of Urban Sprawl by Comparison Cellular Automata-Markov and ANN.

Beauchemin, M., & Fung, K. B. (2001). On statistical band selection for image visualization. PE & RS-Photogrammetric Engineering and Remote Sensing, 67(5), 571-574.

BEHZADFAR, M., & NOURMOHAMMADZAD, H. (2011). COGNITION OF STRUCTURE OF YAZD CITY HISTORIC PHYSIC TEXTURE.

Bokaie, M., Zarkesh, M. K., Arasteh, P. D., & Hosseini, A. (2016). Assessment of urban heat island based on the relationship between land surface temperature and land use/land cover in Tehran. Sustainable Cities and Society, 23, 94-104.

Buhe, A., Tsuchiya, K., Kaneko, M., Ohtaishi, N., & Halik, M. (2007). Land cover of oases and forest in XinJiang, China retrieved from ASTER data. Advances in Space Research, 39(1), 39-45.

CARPER, W., LILLESAND, T., & KIEFER, R. (1990). The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data. Photogrammetric Engineering and remote sensing, 56(4), 459-467.

Cengiz, O., Sener, E., & Yagmurlu, F. (2006). A satellite image approach to the study of lineaments, circular structures and regional geology in the Golcuk Crater district and its environs (Isparta, SW Turkey). Journal of Asian Earth Sciences, 27(2), 155-163.

Chavez, P., Berlin, G. L., & Sowers, L. B. (1982). Statistical method for selecting landsat MSS. J. Appl. Photogr. Eng, 8(1), 23-30.

Chavez, P., Sides, S. C., & Anderson, J. A. (1991). Comparison of three different methods to merge multiresolution and multispectral data- Landsat TM and SPOT panchromatic. Photogrammetric Engineering and remote sensing, 57(3), 295-303.

Chen, B., Huang, B., & Xu, B. (2017). Multi-source remotely sensed data fusion for improving land cover classification. ISPRS Journal of Photogrammetry and Remote Sensing, 124, 27-39.

Colditz, R. R., Wehrmann, T., Bachmann, M., Steinnocher, K., Schmidt, M., Strunz, G., & Dech, S. (2006). Influence of image fusion approaches on classification accuracy: a case study. International journal of remote sensing, 27(15), 3311-3335.

Copernicus. (2015). Sentinel-2A Launch. http://www.copernicus.eu/sites/default/files/documents/Sentinel_2A.pdf

Dimov, D., Kuhn, J., & Conrad, C. (2016). Assessment of cropping system diversity in the fergana valley through image fusion of landsat 8 and sentinel-1. ISPRS annals of the photogrammetry, remote sensing and spatial information sciences, 3, 173.

eesa. MultiSpectral Instrument (MSI) Overview. https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument

Foody, G. M. (2020). Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. Remote sensing of environment, 239, 111630.

Gašparović, M., & Jogun, T. (2018). The effect of fusing Sentinel-2 bands on land-cover classification. International journal of remote sensing, 39(3), 822-841.

Huang, W., Xiao, L., Wei, Z., Liu, H., & Tang, S. (2015). A new pan-sharpening method with deep neural networks. IEEE Geoscience and Remote Sensing Letters, 12(5), 1037-1041.

Ishtiaque, A., Shrestha, M., & Chhetri, N. (2017). Rapid urban growth in the Kathmandu Valley, Nepal: Monitoring land use land cover dynamics of a himalayan city with landsat imageries. Environments, 4(4), 72.

Jensen, J. R. (1996). Introductory digital image processing: a remote sensing perspective. Prentice-Hall Inc.

Keshavarzi, A., Sarmadian, F., Sadeghnejad, M., & Pezeshki, P. (2010). Developing pedotransfer functions for estimating some soil properties using artificial neural network and multivariate regression approaches. Proenvironment Promediu, 3(6).

Keshavarzi, A., Sarmadian, F., Tirado-Corbal, R., & Sadeghnejad, M. (2010). A Sensitivity Analysis of ANN Pedotransfer Functions for spatial modeling of Soil Cation Exchange Capacity. Proenvironment Promediu, 3(6).

Laben, C. A., & Brower, B. V. (2000). Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. In: Google Patents.

Li, S., Kang, X., Fang, L., Hu, J., & Yin, H. (2017). Pixel-level image fusion: A survey of the state of the art. information Fusion, 33, 100-112.

Li, S., & Li, Z. (2010). Effects of image fusion algorithms on classification accuracy. 2010 18th International Conference on Geoinformatics,

Lu, D., Chen, Q., Wang, G., Moran, E., Batistella, M., Zhang, M., . . . Saah, D. (2012). Aboveground forest biomass estimation with Landsat and LiDAR data

and uncertainty analysis of the estimates. International Journal of Forestry Research, 2012.

Maleki, M., Van Genderen, J. L., Tavakkoli-Sabour, S. M., Saleh, S. S., & Babaee, E. (2020). Land use/cover change in Dinevar rural area of West Iran during 2000–2018 and its prediction for 2024 and 2030. Geogr. Tech, 15, 93-105.

Mansourmoghaddam, M., Ghafarian Malamiri, H. R., Arabi Aliabad, F., Fallah Tafti, M., Haghani, M., & Shojaei, S. (2022). The Separation of the Unpaved Roads and Prioritization of Paving These Roads Using UAV Images. Air, Soil and Water Research, 15, 11786221221086285.

Mansourmoghaddam, M., Ghafarian Malamiri, H. R., Rousta, I., Olafsson, H., & Zhang, H. (2022). Assessment of Palm Jumeirah Island's Construction Effects on the Surrounding Water Quality and Surface Temperatures during 2001–2020. Water, 14(4), 634.

Mansourmoghaddam, M., Rousta, I., Zamani, M., Mokhtari, M. H., Karimi Firozjaei, M., & Alavipanah, S. K. (2021). Study and prediction of land surface temperature changes of Yazd city: assessing the proximity and changes of land cover. Journal of RS and GIS for Natural Resources, 12(4), 1-27.

Mansourmoghaddam, M., Rousta, I., Zamani, M. S., Mokhtari, M. H., Karimi Firozjaei, M., & Alavipanah, S. K. (2022). Investigating And Modeling the Effect of The Composition and Arrangement of The Landscapes of Yazd City on The Land Surface Temperature Using Machine Learning and Landsat-8 and Sentinel-2 Data. Iranian Journal of Remote Sensing & GIS.

Maurer, T. (2013). How to pan-sharpen images using the gram-schmidt pan-sharpen method–A recipe. International archives of the photogrammetry, remote sensing and spatial information sciences, 1, W1.

Nuthammachot, N., & Stratoulias, D. (2019). Fusion of Sentinel-1A and Landsat-8 images for improving land use/land cover classification in Songkla province, Thailand. APPLIED ECOLOGY AND ENVIRONMENTAL RESEARCH, 17(2), 3123-3135.

Olsen, R. L., Chappell, R. W., & Loftis, J. C. (2012). Water quality sample collection, data treatment and results presentation for principal components analysis–literature review and Illinois River watershed case study. Water research, 46(9), 3110-3122.

Pakdaman, M. (2013). Using MCSST Method for Measuring Sea Surface Temperature with Modis Imagery and Modeling and Prediction of Regional Variations with Least Squares Method (case Study: Persian Gulf, Iran).

Pal, S., & Ziaul, S. (2017). Detection of land use and land cover change and land surface temperature in English Bazar urban centre. The Egyptian Journal of Remote Sensing and Space Science, 20(1), 125-145.

Palsson, F., Sveinsson, J. R., & Ulfarsson, M. O. (2013). A new pansharpening algorithm based on total variation. IEEE Geoscience and Remote Sensing Letters, 11(1), 318-322.

Pohl, C., & Van Genderen, J. (1998). Multisensor image fusion in remote sensing: concepts, methods and applications" International journal of remote sensing–Vol. 19.

Pohl, C., & Van Genderen, J. (2016). Remote sensing image fusion: A practical guide. Crc Press.

Puttinaovarat, S., & Horkaew, P. (2017). Urban areas extraction from multi sensor data based on machine learning and data fusion. Pattern Recognition and Image Analysis, 27(2), 326-337.

Qadri, S., Khan, D. M., Qadri, S. F., Razzaq, A., Ahmad, N., Jamil, M., . . . Awan, S. A. (2017). Multisource data fusion framework for land use/land cover classification using machine vision. Journal of Sensors, 2017.

Qaid, A. M., & Basavarajappa, H. (2008). Application of optimum index factor technique to Landsat-7 data for geological mapping of north east of Hajjah, Yemen. American-Eurasian Journal of Scientific Research, 3(1), 84-91.

Richards, J. A., & Richards, J. (1999). Remote sensing digital image analysis (Vol. 3). Springer.

Richter, R., Wang, X., Bachmann, M., & Schläpfer, D. (2011). Correction of cirrus effects in Sentinel-2 type of imagery. International journal of remote sensing, 32(10), 2931-2941.

Sales, M. H. R., Souza, C. M., & Kyriakidis, P. C. (2012). Fusion of MODIS images using kriging with external drift. ieee Transactions on geoscience and remote sensing, 51(4), 2250-2259.

Sukawattanavijit, C., & Chen, J. (2015). Fusion of RADARSAT-2 imagery with LANDSAT-8 multispectral data for improving land cover classification performance using SVM. 2015 IEEE 5th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR),

Thompson, W. D., & Walter, S. D. (1988). A reappraisal of the kappa coefficient. Journal of clinical epidemiology, 41(10), 949-958.

USGS. Landsat Missions: Landsat 8. https://www.usgs.gov/core-science-systems/nli/landsat/landsat-8?qt-science_support_page_related_con=0#qt-science_support_page_related_con

USGS. What are the band designations for the Landsat satellites? https://www.usgs.gov/faqs/what-are-band-designations-landsat-satellites?qt-news_science_products=0#qt-news_science_products

Wang, Q., Shi, W., Li, Z., & Atkinson, P. M. (2016). Fusion of Sentinel-2 images. Remote sensing of environment, 187, 241-252.

Willmott, C. J., & Matsuura, K. (2006). On the use of dimensioned measures of error to evaluate the performance of spatial interpolators. International Journal of Geographical Information Science, 20(1), 89-102.

Zare Naghadehi, S., Asadi, M., Maleki, M., Tavakkoli-Sabour, S.-M., Van Genderen, J. L., & Saleh, S.-S.

(2021). Prediction of Urban Area Expansion with Implementation of MLC, SAM and SVMs' Classifiers Incorporating Artificial Neural Network Using Landsat Data. ISPRS International Journal of Geo-Information, 10(8), 513.

Zheng, H., Du, P., Chen, J., Xia, J., Li, E., Xu, Z., . . . Yokoya, N. (2017). Performance evaluation of downscaling Sentinel-2 imagery for land use and land cover classification by spectral-spatial features. Remote Sensing, 9(12), 1274.