A Shallow Neural Network Model for Urban Land Cover Classification Using VHR Satellite Image Features

Mohamed Fawzy [1,3] *, György Szabó [2], Arpad Barsi [1]

[1] Department of Photogrammetry and Geoinformatics, Faculty of Civil Engineering, Budapest University of Technology and Economics, 3 Műegyetem rkp., K Building First Floor 31., H-1111, Budapest, Hungary - Mohamedfawzyramadanmahmoud@edu.bme.hu, barsi.arpad@emk.bme.hu

[2] Institute of Mathematics and Informatics, University of Nyíregyháza, Sóstói út 31/b A ép.112, H-4400, Nyíregyháza, Hungary - Szabo.gyorgy.dr@nye.hu

[3] Civil Eng. Dept., Faculty of Engineering, South Valley University, Qena, Egypt - Mohamedfawzy@eng.svu.edu.eg

KEYWORDS: Shallow neural networks, image classification, VHR satellite images, urban environments, Land use and land cover.

ABSTRACT:

Recently, image classification techniques using neural networks have received considerable attention in sustainable urban development, since their applications have an extreme effect on building distribution, infrastructural networks, and water resource management. In this research, a back-propagation shallow neural network model is presented for very high resolution satellite image classification in urban environments. Workflow procedures consider selecting and collecting data, preparing required study areas, extracting distinctive features, and applying the classification process. Visual interpretation is performed to identify observed land cover classes and detect distinctive features in the urban environment. Pre-processing techniques are implemented to present the used images in a more suited form for the classification techniques. A shallow neural network model (supported by MathWorks MATLAB environment) is successfully applied and results are evaluated. The proposed model is tested for classifying both WorldView-2 and WorldView-3 multispectral images with different spatial and spectral characteristics to check the model's applicability to various kinds of satellite imagery and different study areas. Model outcomes are compared to two well-known classification methods; the Nearest Neighbour object-based method and the Maximum Likelihood pixel-based classifier, to validate and check the model stability. The overall accuracy achieved by the proposed model is 86.25% and 83.25%, while the nearest neighbour approach has obtained 84.50% and 82.75%, and the maximum likelihood classifier has accomplished 82.50% and 80.25% for study area 1 and study area 2 respectively. Obtained results indicate that the developed shallow neural network model achieves a promising accuracy for urban land cover classification in comparison with the standard techniques.

1. INTRODUCTION

Remote sensing techniques offer cost-effective resources for Land Use and Land Cover (LULC) analysis using Very High Resolution (VHR) satellite images. Urban land cover interpretation is a challenging task due to the scene complexity and wide-range spatial and spectral resolution of available imagery systems. With recent development in the space industry, a variety of data are available for earth observation and sustainable LULC studies. According to the Union of Concerned Scientists (UCS), there were about one thousand Earth Observation (EO) satellites in orbit until 2022 (Geospatial World, 2023). Considerable interest and rapid progress in collecting data using EO satellites have been seen in recent years due to improved technology, relatively lower costs, and high enhancement in satellite image resolutions. Image resolutions have greatly increased; a centimetres-level of spatial resolution, spectral resolution over tens or even hundreds of different bands, radiometric resolution of 16-bit or higher, and temporal resolution with two images a day are now obtainable. The wide expansion of accessible satellite imagery rises a significant challenge to collect, arrange, and analyse data depending on standard soft-wares and manual traditional methodologies. Manual methods are time-consuming and require extensive human effort to finish the classification process. An automatic technique is needed for image classification to maximize the benefits of such available data. The 4th Industrial Revolution offers more effective solutions for image analysis using Artificial Intelligence (AI), especially

Neural Network (NN) tools. Neural networks are considered one of the most powerful techniques for satellite image studies according to their strength of using all available features and outperforming other image analysis methods in urban applications (Mäyrä, J., 2018). Several neural network algorithms have been adopted in remote sensing land cover classification involving deep and shallow neural networks with different structures based on model flexibility and required application. Classification approaches using neural networks are identified depending on the number of layers. NNs include various hidden layers using a network of interconnected units (or neurons) that learn nonlinear connections from input data (Goodfellow, et al., 2016). Deep neural networks consist of multi-convolution, pooling, and activation layers followed by fully connected layers that represent the results. Meanwhile, shallow neural network architecture is limited by a few layers only which is typical for feed-forward neural nets with the back-propagation training algorithm (Ososkov, G., and Goncharov, P., 2017). Although deep learning approaches can efficiently analyse rich spectral information contained in multispectral images to perform an accurate discrimination process, multi-dimension data poses additional challenges requiring powerful hardware and more time to train the deep network. Moreover, using plenty of features (i.e., spectral bands, indices, geometric, and spatial properties) causes more complexity to the classifier and increases the number of parameters for defining urban classes. The question of "How to choose a suitable network structure?" requires more effort to be answered (Paoletti, M., et

---

* Corresponding author

al., 2019). In comparison with deep networks, shallow neural networks take the advantage of a simple structure, few numbers of layers, and consume less time and power for the training and classification process (Gorokhovatskyi, O., and Peredrii, O., 2018). Simplicity, flexibility, fast training, and using less computational resources and memory make shallow neural networks an effective choice for VHR image classification (Lei, F., et al., 2019).

2. RESEARCH PROBLEM

Urban environment represents one of the most complex study areas for remote sensing studies due to the wide-ranging spatial and spectral heterogeneity of surface materials. Also, spectral similarities between different land cover classes cause considerable confusion for more classification techniques. Urban classes have various similarities in geometry (e.g., roads, bridges, and railway routes), similarities in colour (e.g., buildings, walk lanes, and bare soil), and similarities in texture (e.g., asphalt roads, parking lots, and some building surfaces). Class resemblances make scene analysis more complicated in such urban environments. The recent development of spectral and spatial resolution leads to a significant interest in using VHR satellite imagery for urban feature extraction. VHR satellite images provide up-to-date information regarding the Earth's resources, which is vital for sustainable urban development, agricultural investigation, water resource management, and disaster control (Fawzy, et al., 2020). During the past two decades, hundreds of image classification methods were developed considering pixel and object-based techniques. Despite the increased number of classification approaches for land-use analysis, the accuracy and efficiency of these methods are still insufficient to fulfil the requirements of real-world applications such as sustainable urban planning and land management (Huang, B., et al., 2018). Image classification meets more challenges due to spectral similarities of land cover classes that pixel-based techniques fail to deal with. Object-Based Image Analysis (OBIA) techniques are effective for land use and land cover analysis. OBIA overcomes the problem of spectral similarities in pixels by dividing the satellite image into more appropriate segments according to both spectral and spatial characteristics. However, the growing number of LULC classes with various features makes OBIA a time-consuming method, requiring more human effort, in addition to utilizing a large amount of computer memory and extra computation power to finish the process (Fawzy, et al., 2022). Furthermore, OBIA requires rearranging parameters, functions, and/or algorithms for each different image and different study area, which limits the generalizability and transferability of the strategy. Essentially, an automated strategy is required to accurately identify the diversity of urban classes using VHR images. Neural network-based approaches offer multiple solutions for image classification drawbacks (Längkvist, M., et al., 2016). The key feature of NN-based algorithms is that NNs do not require prior feature extraction for each study area, thus increasing the generalization capabilities (Sertel, E., et al., 2022). Neural networks are capable of collecting, arranging, and analysing features to train, validate, and test models for classifying urban areas. In this article, shallow neural networks are investigated for VHR satellite image classification. A back-propagation shallow neural network model is developed, applied, and evaluated with the objective of maximizing image classification results in urban study areas.

3. OUTLINE OF OBJECTIVES

The main objective of this research is to study the effectiveness and competitiveness of shallow neural networks for satellite image classification in urban environments. A proposed methodology is introduced using neural networks to overcome the limitations of traditional classification strategies and maximize extracted urban land cover information from VHR satellite images. Also, the following procedures have been considered:

- Developing, applying, and evaluating a shallow neural network model for urban land cover classification using VHR multispectral satellite images.
- Testing the efficiency of neural networks for image classification in terms of image resolution, type and degree of planning for the selected study areas, and the required level of accuracy for each application.
- Applying the suggested model to different case studies for land cover classification in the urban environment.

4. STATE-OF-THE-ART

Remote sensing applications using VHR satellite images have received widespread attention in urban development due to their notable effects on building allocation, infrastructural systems, and water resource utilization. Several approaches have applied NNs for urban land cover analysis depending on VHR satellite images. Neural network-based techniques are considered a highly promising tool for remote sensing image analysis which produce considerable results in the last few years (Martins, V. S., et al., 2020). Recently, remote sensing experts have paid great attention to neural networks to achieve significant success at many image classification tasks such as feature extraction and LULC analysis (Cheng, Gong, et al., 2018). Ma, Lei, et al., (2019) introduced the major neural network concepts pertinent to remote sensing reviewing 200 published papers during the previous two years covering nearly every application of remote sensing from pre-processing to mapping. Neupane, B., et al., (2021) focused on urban remote sensing and performed a meta-analysis for papers related to current research problems, data sources, data pre-processing, and training details for various architectures. Analysis showed that neural networks outperform traditional methods in terms of accuracy, and address several faced challenges. Nguyen et al., (2013) developed a five-layer network algorithm for satellite image classification into six urban classes: airport, bridge, forest, harbour, land, and urban. Hudjakov and Tamre, (2013) applied a network with two convolutional and one linear classifier to assign land cover to four classes: roads, grasses, houses, and bushes. Chen et al. (2014) applied NNs to extract individual vehicles from satellite images. Ishii et al., (2015) compared the neural network algorithms with Support Vector Machine (SVM) for surface object detection using Landsat 8 imagery and concluded that NN results outperformed SVM.

5. METHODOLOGY

The indicated methodology applies a back-propagation shallow neural network model to deal with multiple features using VHR multispectral satellite images for land cover classification in urban environments. To meet the research objectives, the following procedures are presented (Figure 1), and the results are evaluated:

- Collecting the required data and the used satellite images for the selected study areas.

- Applying pre-processing techniques to present the used images in a better appropriate form for the classification process.
- Developing a shallow neural network model consisting of a few fully connected layers. Through the model, required features are extracted for each pixel focusing on basic characteristics (e.g., spectral bands and spectral indices) of input images.
- Applying the developed model for classifying urban case studies to validate and check the model stability in comparison with standard classification methods considering overall accuracy and kappa coefficient.
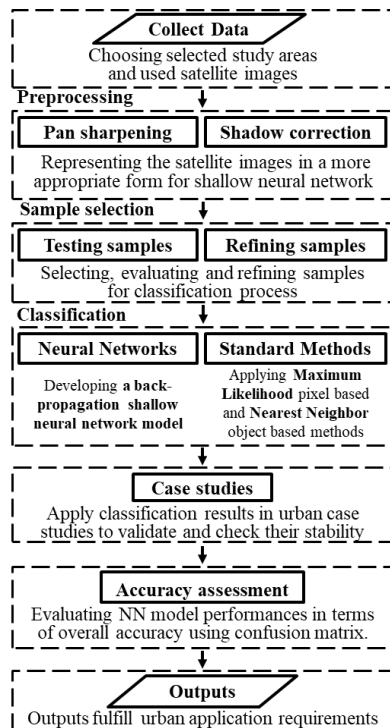

Figure 1. Flow chart of the proposed methodology.

6. EXPERIMENTAL WORKS

6.1 Data Used

WorldView series provides high resolution Earth observation imagery for environment monitoring. WorldView-2 offers a panchromatic band of 0.50m spatial resolution and eight multispectral bands of 2.00m spatial resolution: blue, green, red, and first near-infrared, in addition to coastal blue, yellow, red-edge, and a second near-infrared bands (Figure 2). WorldView-3 supplies a panchromatic band of 0.30m spatial resolution and sixteen multispectral bands of 1.24m spatial resolution. As a first multi-payload, super-spectral, WorldView-3 comprises the same eight multispectral bands as WorldView-2, in addition to eight short waves infra-red bands (Figure 3).
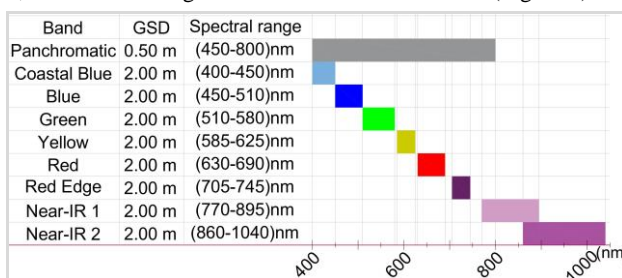

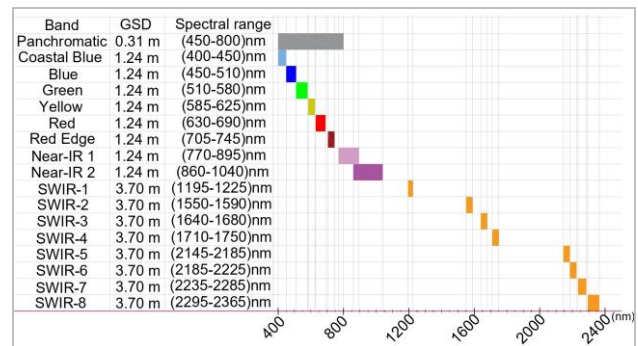**Figure 2.** WV-2 eight multispectral band specifications.

Figure 3. WV-3 sixteen multispectral band specifications.

6.2 Study Areas

Considering WorldView-2 and WorldView-3 specifications and visual analysis of used images, two study areas were chosen. study area 1 is selected, from WorldView-2 image, for Qena City, Egypt, captured in March 2017, X=472098.25:472582.25, Y=2893729.25:2893934.75 UTM. It is a semi-planned area with five land cover classes: water, vegetation, bare soil, road, and building (Figure 4-a). Study area 2 is chosen, from WorldView-3 image, for Provo City, USA, captured in June 2013, X=439271.94:439734.46, Y=4455703.74:4456328.70 UTM. It is a well-planned area with the same land cover classes (Figure 4-b).

(a) Study area 1   (b) Study area 2
Figure 4. (a) Study area 1 of Qena City, Egypt,
and (b) Study area 2 of Provo City, USA.

6.3 Image Pre-Processing

Image pre-processing is essential to improve the analysis process and maximize the extracted information from VHR satellite images. Two main tasks of image pre-processing are presented for used images: data fusion and shadow correction.

6.3.1 Data Fusion

A data fusion process is required to inject high spatial resolution from the panchromatic band into multispectral resolution bands to get a high spatial and multispectral quality pan-sharpened image. Panchromatic images contain a high spatial resolution which is advantageous for geometric studies and feature extraction. Whereas, multispectral images present various spectral resolutions that are required for spectral land use and land cover studies. A variety of techniques for image fusion have been presented in the state-of-art including Principal Component Analysis (PCA), Brovey Transform (BT), and Intensity Hue Saturation (IHS) transformation (Pohl and Genderen, 2016; Ehlers et al., 2010). PCA is applied to used satellite images as all input bands are pan-sharpened and fitted for the feature extraction process (Fawzy, 2020).

6.3.2 Shadow Correction

Shadows pose a major challenge for VHR satellite images and affect information quality. However, shadow area has weak reflectance, it still provides valuable information to apply shadow restoration. Shadow correction process is presented to compensate the brightness difference between shadow and non-shadow areas within two main steps: shadow detection and shadow compensation. Shadow detection is applied using a set of indices depending on various colour models to detect shadows in the presence of dark objects, while shadow compensation aims to recompense brightness differences between shadow and non-shadow areas in VHR satellite images. Study areas, of the presented paper, involve different shadow and dark water regions. A distinct issue is a misclassification between shadow and water classes. The Optimized Shadow Index (OSI) proposed by Mostafa and Abdelhafiz, (2017) is considered to differentiate between shadow and water pixels. Additionally, the Linear Correlation Correction (LCC) method introduced by Sarabandi et al., (2004) is applied for shadow compensation to provide meaningful information despite the weak signals in shadow areas (Fawzy, et. al, 2020).

6.4 Sample Selection

Classification is one of the most essential steps for urban feature extraction using VHR satellite images. Most image classification techniques require adequate and effective samples for training. Samples are usually identified manually and labelled by human visual inspection or field exploration. Training samples strongly affect classification results. Consequently, training samples should be meaningful, pure, representative, and cover the entire image. Selecting effective samples requires two main procedures to fit the classification process and enhance final results: sample testing and sample refining (Fawzy, et al., 2021).

6.4.1 Training Sample Testing

a) Intensity Distribution

Signature of each group of samples, for one class, requires viewing the intensity count histograms. It is significant to determine if the brightness histogram has a unimodal distribution for one type of land cover e.g., water, vegetation, and road samples, or multimodal for samples with heterogeneous materials e.g., building and bare soil (Figure 5). A distribution with more than one mode represents samples of more than one distinct class that need to be refined.
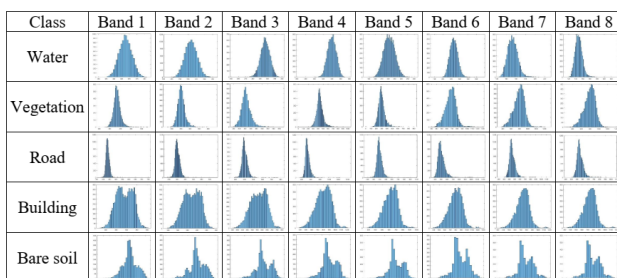

Figure 5. Samples intensity distribution for study area 1.

b) Separability of Classes

Separate and sufficient samples enable the classifier to distinguish various class signatures. Class separability could be measured and evaluated using multiple equations, for separability distance measurement, available in many remote sensing applications (e.g., Jefferies-Matusita in Erdas software

with values from 0 to 1414). High separability refers to good representative samples with fewer spectral similarities. Meanwhile, low separability represents more confusion between samples due to spectral similarities and leads to a high number of misclassified areas where samples should be improved. Signature separability values for land cover classes of study area 1 are shown in Table 1. The highest separability can be seen between both water, vegetation, and other classes, while little fewer values exist between building-road and building-bare soil classes due to wide heterogeneous materials of building and spectral overlap with road and bare soil.

Class Pairs									
W:V	W:R	W:B	W:BS	V:R	V:B	V:BS	R:B	R:BS	B:BS
1414	1414	1414	1414	1413	1411	1412	1322	1409	1240
W: Water, **V:** Vegetation, **BS:** Bare soil, **R:** Road, and **B:** Building									

Table 1. Signature separability for classes of study area 1.

c) Homogeneity of Class

Training samples are essential to be tested according to the range of mean and standard deviation to determine the homogeneity or heterogeneity of each class. For a set of samples, a close range of mean values refers to one class, while a wide range refers to overlapped classes. Low standard deviation refers to high homogeneous samples (e.g., water, vegetation, and road), while high standard deviation indicates more heterogeneity of samples (e.g., building and bare soil classes). Table 2 illustrates the statistical properties for samples of study area 1.

Band	Water Mean	Water Std Dev	Vegetation Mean	Vegetation Std Dev	Road Mean	Road Std Dev	Building Mean	Building Std Dev	Bare soil Mean	Bare soil Std Dev
B1	67.5	4.1	384.5	17.7	431.6	16.8	504.0	52.9	468.1	23.1
B2	236.7	3.7	252.6	21.5	299.1	18.2	388.1	57.4	353.6	27.1
B3	296.2	6.7	359.0	43.4	397.3	35.8	595.9	112.7	546.6	57.8
B4	339.0	10.2	434.6	72.0	520.3	57.3	852.9	180.5	798.5	97.1
B5	221.1	8.8	298.5	64.1	388.2	48.8	663.1	141.8	634.2	82.2
B6	225.7	10.0	666.0	103.4	493.4	62.5	848.4	180.1	849.8	109.6
B7	172.9	11.3	779.6	132.0	443.9	61.4	779.8	160.4	818.0	108.2
B8	148.0	11.0	713.2	124.2	415.3	58.6	723.3	147.0	770.5	103.7

**Table 2.** Sample statistical properties of study area 1.

6.4.2 Training Samples Refining

Image processing soft-wares offer many tools for testing samples. Non-representing samples require refining, and incorrect ones should be replaced. Sample refinement focus on the purest samples of all classes. As a result, the classification process could achieve optimal results.

6.5 Feature Extraction

Neural networks have attracted a lot of interest due to their flexibility in processing all features and classifying various kinds of satellite images. Shallow learning requires manually chosen features and meaningful amounts of labelled data (Sainos-Vizuett, M., and Lopez-Nava, I. H., 2021). Image classification using neural networks depends on the prior feature extraction process to feed the model for the classification process. Selected samples are arranged to extract the required features for the shallow neural network. Each pixel represents one input with all values for spectral bands. In addition to spectral band values, a set of spectral indices are calculated and involved in the model inputs (Figure 6). Indices are important to detect classes using the differences between spectral reflectance. The used indices are calculated depending on WorldView-2 and WorldView-3 spectral bands, e.g., the

Normalized Difference Water Index (NDWI) (Gao, 1996), Normalized Difference Vegetation Index (NDVI) (Knipling, 1970), World View Soil Index (WV-SI), WorldView Building Index (WV-BI) (Wolf, 2012), and Road Extraction Index (REI) (Shahi et al., 2015). Feature extraction is the first step in developing an effective shallow neural network model.
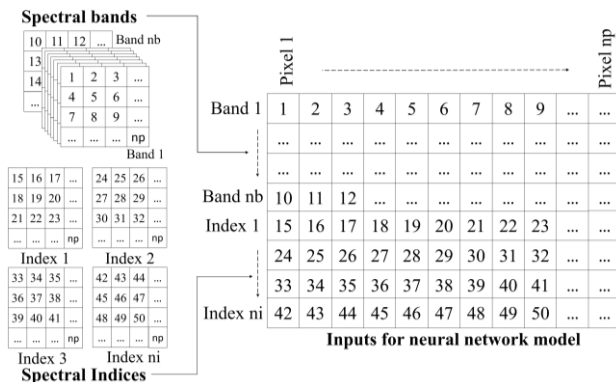
Figure 6. Extracted features including spectral bands and indices values for each pixel.

6.6 Shallow Neural Network Model

A shallow neural network requires manually selected samples and meaningful labelled data. The proposed model is built depending on a set of features and indices extracted from multispectral satellite images. The presented procedures (Figure 7) start with selecting, testing, and refining samples from the used satellite image. Selected samples are arranged as a matrix of required features focusing on the spectral value of all bands and spectral indices. Finally, a shallow neural network is applied depending on the input features and input targets to classify multispectral satellite images. To improve classification accuracy, the parameters of used shallow neural network model should be identified carefully (Lei, F., et al., 2019). The number of layers and neurons is changeable considering the required classification level for a particular application and the time of the classification process. The used model focuses on optimizing input features and the number of layers and neurons on the network to maximize classification results. After several trials, a model consisting of two hidden layers with one hundred neurons for each one is chosen, which balance between the classification accuracy and the required time for training and classifying processes. The suggested model is applied successfully in two different study areas and results are evaluated.
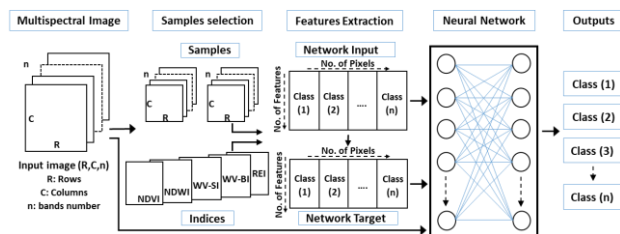
Figure 7. Shallow neural network flow chart.

### 6.7 Standard Classification Methods

Shallow neural network results are evaluated compared to traditional classification methods to check the stability and efficiency of the proposed approach. For this comparison, Nearest-Neighbour and Maximum-Likelihood are selected as most promising classification methods for object and pixel-

based strategies respectively (Fawzy, Et. Al, 2020). Nearest Neighbour is applied at the object level using eCognition software, while Maximum Likelihood is performed at the pixel level using Erdas software. Selected samples of the proposed NN model are used as training areas for both nearest neighbour and maximum likelihood methods. Finally, the accuracy assessment of classified images is implemented using error matrix depending on the same reference points.

7. RESULTS AND DISCUSSION

7.1 Accuracy Assessment

Shallow neural network performance is evaluated on both a qualitative and quantitative level. Visual analysis is applied to evaluate classification outcomes qualitatively by comparing input and classified images to roughly identify the size and location of the errors (Figure 8). While quantitative evaluation is performed considering the correlation between the classification results and reference points. A confusion matrix is derived to the classified images with 400 random points distributed over all classes. Reference points are defined at least 50 points per class. Points for each category are adjusted considering the area of class in the entire scene and the relative importance of that class for a particular application.

(a) Study area 1   (b) Study area 2
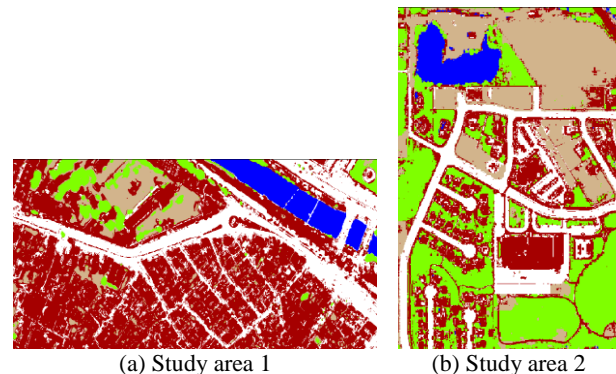
Figure 8. Neural network classification results.

Error matrices are shown in Table 3 and Table 4. It can be seen that the proposed model has achieved an overall accuracy of 86.25% and 83.25%, and kappa statistics of 0.8193 and 0.7881 for study area 1 and study area 2 respectively. For individual classes of study area 1, the producer's accuracy ranges from 81.17% for building to 98.21% for water class, and the user's accuracy ranges from 77.53% for road to 100.00% for the water class. Study area 2 achieves producer's accuracy ranges from 67.97% for bare soil to 92.65% for road, and user's accuracy ranges from 67.02% for building to 100.00% for the water class.

Classified Data	Bare soil	Building	Road	Vegetation	Water	Total Row	Producer's Accuracy	User's Accuracy
Bare soil	**48**	9	1	3	0	61	85.71%	78.69%
Building	3	**125**	12	0	0	140	81.17%	89.29%
Road	1	18	**69**	0	1	89	83.13%	77.53%
Vegetation	4	2	1	**48**	0	55	94.12%	87.27%
Water	0	0	0	0	**55**	54	98.21%	100.00%
Total column	56	154	83	51	56	**400**	-	-
Overall accuracy **86.25%** Overall kappa statistics **0.8193**								

Table 3. Confusion matrix for shallow neural network classification of study area 1.

Classified Data	Bare soil	Building	Road	Vegetation	Water	Total Row	Producer's Accuracy	User's Accuracy
Bare soil	**87**	1	0	2	3	93	67.97%	93.55%
Building	19	**63**	5	7	0	94	91.30%	67.02%
Road	7	5	**63**	2	0	77	92.65%	81.82%
Vegetation	15	0	0	**73**	1	89	86.90%	82.02%
Water	0	0	0	0	**47**	47	92.16%	100.00%
Total column	128	69	68	84	51	**400**	-	-
Overall accuracy **83.25%** Overall kappa statistics **0.7881**								

Table 4. Confusion matrix for shallow neural network classification of study area 2.

The nearest neighbour object-based method has achieved an overall accuracy of 84.50% and 82.75%, and kappa statistics of 0.7941 and 0.7833, while the maximum likelihood pixel-based strategy has got an overall accuracy of 82.50% and 80.25%, and kappa statistics of 0.7729 and 0.7507 for study area 1 and study area 2 respectively. Detailed results obtained by object-based and pixel-based methods are shown in the Appendix. It can be seen that the shallow neural network model slightly outperforms traditional methods in terms of overall accuracy and KAPPA coefficient (Figure 9). Also, it achieves convenient producer's and user's accuracy in both study areas (Figure 10).
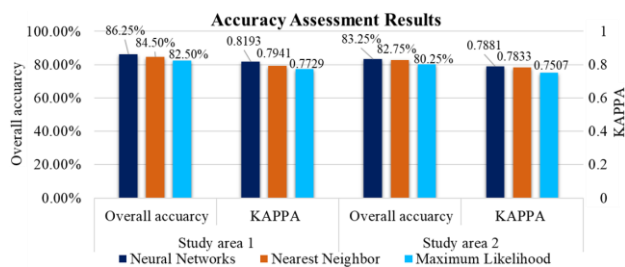
Figure 9. Accuracy assessment results.
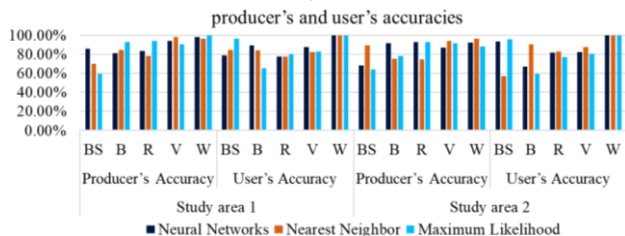
Figure 10. Producer's and user's accuracy results.

7.2 Discussion

The suggested methodology introduces pre-processing tasks including pan sharpening and shadow correction processes to enhance the input raw data. Thanks to pre-processing, the results become more valuable. Also, the developed model depends on a designed neural networks function (code) considering not only the intensity value of each pixel, but also additional features like indices, neighbourhood, texture, and digital elevation models. Result analysis shows that shallow neural networks are an effective technique for classifying water and vegetation classes. Also, it offers appropriate results for road detection, while it needs more intensive work to enhance building and bare soil extraction. Error matrices illustrate some vegetation-bare soil misclassifications; 7% (4/56) of bare soil points in study area 1 and 12% (15/128) in study area 2 were classified as vegetation, against 6% (3/51) of vegetation points in study area 1 and 6% (3/48) in study area 2 were predicted as bare soil. Misclassifications between vegetation and bare soil

classes occur since some patches among vegetation areas are bare of plants which are misclassified as bare soil. Additionally, a weak vegetation effect appears in some bare soil areas that are misclassified as vegetation (Figure 11).

(a): Reference points in VHR images   (b): Field observation images

Figure 11. Confusion between vegetation and bare soil classes.

In addition, building-road and building-bare soil confusions affect the classification outcomes. In study area 1, 12% (18/154) of building points were assigned to road and 6% (9/154) were seen as bare soil, against 14% (12/83) of road points and 5% (3/56) of bare soil points were seen as building class. Meanwhile, in study area 2, 15% (19/128) of bare soil points are classified as building and 5% (7/128) are seen as road, also 7% (5/69) of building points are misclassified as road against 7% (5/68) of road points are estimated as building class. Confusion of building, road, and bare soil occurs due to the spectral similarities between their materials and wide heterogeneous surfaces of buildings in urban environments. Added to that, a few inaccurate points are noticed in water classification results because of the shallow water effect.

The proposed model has effectively performed the classification process for both well-planned and semi-planned study areas using WorldView-2 and WorldView-3 satellite images. Shallow neural networks have presented a fast response for the training and testing process, and are easy to be involved in automatic strategies. Results fulfil the scientific knowledge gaps in the performance of neural network algorithms for urban land cover classification, built-up area maps, infrastructural networks, and water resource distribution. However, shallow neural network outcomes are changed after every iteration of re-training or re-applying the model, which makes the neural network technique a black box. Also, the used strategy has a major drawback as it depends on classifying each pixel separately, which occasionally adds a "Salt-and-Pepper" effect to classified images. "Salt-and-Pepper" impact is a result of high heterogeneity between nearby pixels. Each pixel is handled independently of its surrounding ones in pixel-based techniques. Despite their neighbourhood, the surrounding pixels frequently belong to different classes (Kelly, M., et al., 2011). To overcome this pixel-based "Salt-and-Pepper" impact, it is recommended to apply an image segmentation process first and apply neural network classification on the object level instead of pixels. However, the acceptable performance of shallow neural networks for VHR image classification in urban environments, new strategies are needed to optimize the outcomes and maximize classification results. Segmenting the entire image into a set of objects is suggested to enhance the classification process depending on the object level. Integration between shallow neural networks and object-based features is predicted to add valuable information to the designed model and improve the shallow neural network performance.

8. CONCLUSION AND FUTURE WORK

Neural networks achieve significant progress in a variety of image analysis tasks including image classification and feature extraction for sustainable urban management using VHR satellite images. In this article, a shallow neural network model has been developed for image classification consisting of two hidden layers with one hundred neurons structure and applied to different study areas. Shallow neural networks are effective for different spatial and various spectral resolution images. The

proposed model is compared to two standard classification methods and achieves promising results with an overall accuracy of 86.25% and 83.25% for study area 1 and study area 2 respectively. It achieves efficient outcomes for water, vegetation, and road, meanwhile building and bare soil classes still require more enhancement. It is concluded that shallow neural networks can be effective enough to produce classification results that outperform state-of-the-art object and pixel-based classification techniques, with some limitations that require new strategies to enhance classification results.

ACKNOWLEDGEMENTS

## REFERENCES

Chen, X., Xiang, S., Liu, C. L., & Pan, C. H. (2014), Vehicle detection in satellite images by hybrid deep convolutional neural networks. IEEE Geoscience and remote sensing letters, 11(10), 1797-1801.

Cheng, G., Yang, C., Yao, X., Guo, L., & Han, J. (2018), When Deep Learning Meets Metric Learning: Remote Sensing Image Scene Classification Via Learning Discriminative CNNs. IEEE transactions on geoscience and remote sensing, 56(5), 2811-2821.

Ehlers, M., S. Klonus, P. Johan Åstrand, and P. Rosso (2010), Multi-sensor image fusion for pansharpening in remote sensing, International Journal of Image and Data Fusion, 1(1), 25-45.

Fawzy, M. (2020), Urban Feature Extraction from High Resolution Satellite Images, M.SC Thesis, Civil Engineering Dept., South Valley University, Egypt.

Fawzy, M., Czinkoczky, A., Szabó, G. (2022), Object-based Image Analysis Assessment for Urban Applications Using VHR Satellite Images, paper presented at Geospatial Information Conference and Exhibition "The meeting of theory and practice in geospatial information", Debrecen-Hungary, November 2-3, pages 143-150.

Fawzy, M., Khodary, F., & Mostafa, Y. G. (2021), Selecting Training Samples Automatically from VHR Satellite Images for Image Classification. Sohag Engineering Journal, 1(1), 71-84.

Fawzy, M., Mostafa, Y. G., & Khodary, F. (2020), Automatic Indices Based Classification Method for Map Updating Using VHR Satellite Images. JES. Journal of Engineering Sciences, 48(5), 845-868.

Gao, B.-C. (1996), NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space, Remote sensing of environment, 58(3), 257-266.

Geospatial World (2023), How Many Satellites are Orbiting Around Earth in 2022? https://www.geospatialworld.net/prime/how-many-satellites-orbiting-earth/ - latest visit April 2023.

Goodfellow, I., Bengio, Y., & Courville, A. (2016), Deep Learning. MIT press.

Gorokhovatskyi, O., & Peredrii, O. (2018), Shallow convolutional neural networks for pattern recognition problems. In 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP) (pp. 459-463). IEEE.

Huang, B., Zhao, B., & Song, Y. (2018), Urban Land-Use Mapping Using A Deep Convolutional Neural Network with High Spatial Resolution Multispectral Remote Sensing Imagery. Remote Sensing of Environment, 214, 73-86.

Hudjakov, R., & Tamre, M. (2011), Ortophoto analysis for UGV long-range autonomous navigation. Estonian Journal of Engineering, 17(1), 17.

Ishii, T., Nakamura, R., Nakada, H., Mochizuki, Y., & Ishikawa, H. (2015), Surface object recognition with CNN and SVM in Landsat 8 images. In 2015 14th IAPR International Conference on Machine Vision Applications (MVA) (pp. 341-344). IEEE.

Kelly, M., Blanchard, S. D., Kersten, E., & Koy, K. (2011). Terrestrial remotely sensed imagery in support of public health: New avenues of research using object-based image analysis. Remote Sensing, 3(11), 2321-2345.

Knipling, E. B. (1970), Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation, Remote sensing of environment, 1(3), 155-159.

Längkvist, M., Kiselev, A., Alirezaie, M., & Loutfi, A. (2016), Classification and segmentation of satellite orthoimagery using convolutional neural networks. Remote Sensing, 8(4), 329.

Lei, F., Liu, X., Dai, Q., & Ling, B. W. K. (2020). Shallow convolutional neural network for image classification. SN Applied Sciences, 2, 1-8.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019), Deep Learning in Remote Sensing Applications: A Meta-Analysis and Review. ISPRS journal of photogrammetry and remote sensing, 152, 166-177.

Martins, V. S., Kaleita, A. L., Gelder, B. K., da Silveira, H. L., & Abe, C. A. (2020). Exploring Multiscale Object-Based Convolutional Neural Network (Multi-OCNN) for Remote Sensing Image Classification at High Spatial Resolution. ISPRS Journal of Photogrammetry and Remote Sensing, 168, 56-73.

Mäyrä, J. (2018), Land Cover Classification from Multispectral Data Using Convolutional Auto Encoder Networks, Master's Thesis, Faculty of Information Technology, University of Jyväskylä.

Mostafa, Y., and Abdelhafiz (2017), Shadow Identification in High Resolution Satellite Images in The Presence of Water Regions, Photogrammetric Engineering & Remote Sensing, 83(2), 87-94.

Nguyen, T., Han, J., & Park, D. C. (2013), Satellite image classification using convolutional learning. In AIP Conference Proceedings (Vol. 1558, No. 1, pp. 2237-2240). American Institute of Physics.

Neupane, B., Horanont, T., & Aryal, J. (2021), Deep Learning-Based Semantic Segmentation of Urban Features in Satellite Images: A Review and Meta-Analysis. Remote Sensing, 13(4), 808.

Ososkov, G., & Goncharov, P. (2017), Shallow and deep learning for image classification. Optical Memory and Neural Networks, 26, 221-248.

Paoletti, M. E., Haut, J. M., Plaza, J., & Plaza, A. (2019), Deep learning classifiers for hyperspectral imaging: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 158, 279-317.

Patel, H., & Upla, K. P. (2022), A shallow network for hyperspectral image classification using an autoencoder with convolutional neural network. Multimedia Tools and Applications, 1-20.

Pohl, and Genderen (2016), Remote sensing image fusion: A practical guide, Crc Press, Boca Raton, Florida, USA (Taylor & Francis).

Sainos-Vizuett, M., & Lopez-Nava, I. H. (2021). Satellite Imagery Classification Using Shallow and Deep Learning Approaches. In Pattern Recognition: 13th Mexican Conference, MCPR 2021, Mexico City, Mexico, June 23–26, 2021, Proceedings (pp. 163-172). Cham: Springer International Publishing.

Sarabandi, P., F. Yamazaki, M. Matsuoka, and A. Kiremidjian (2004), Shadow Detection and Radiometric Restoration in Satellite High Resolution Images, paper presented at IGARSS 2004. IEEE International Geoscience and Remote Sensing Symposium, IEEE, Anchorage, AK, USA, Sept 20-24.

Sertel, E., Ekim, B., Ettehadi Osgouei, P., & Kabadayi, M. E. (2022). Land Use and Land Cover Mapping Using Deep Learning Based Segmentation Approaches and VHR Worldview-3 Images. Remote Sensing, 14(18), 4558.

Shahi, K., H. Z. Shafri, E. Taherzadeh, S. Mansor, and R. Muniandy (2015), A novel spectral index to automatically extract road networks from WorldView-2 satellite imagery, The Egyptian Journal of Remote Sensing and Space Science, 18(1), 27-33.

Vitor Souza Martins (2020), Deep Learning for Land Cover Classification and Environmental Analysis Using High Resolution Remote Sensing Data, Ph.D. Thesis, Agricultural and Biosystems Engineering Dept., The Graduate College, Iowa State University, USA, https://lib.dr.iastate.edu/etd/18232.

Wolf, A. F. (2012), Using WorldView-2 Vis-NIR multispectral imagery to support land mapping and feature extraction using normalized difference index ratios, paper presented at Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII, International Society for Optics and Photonics, Baltimore, Maryland, USA, April 23- 27.

APPENDIX (A)

(a)

(b)

(c)

(d)
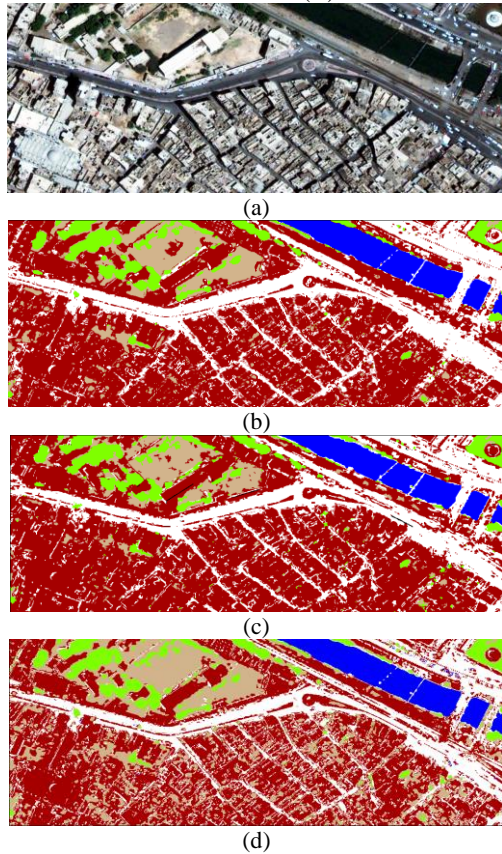
Figure A-1. Study area 1 classification results
(a) RGB image, (b) Neural Networks, (c) Nearest Neighbour
Object-based, and (d) Maximum Likelihood Pixel-based.

(a)

(a)

(c)

(d)
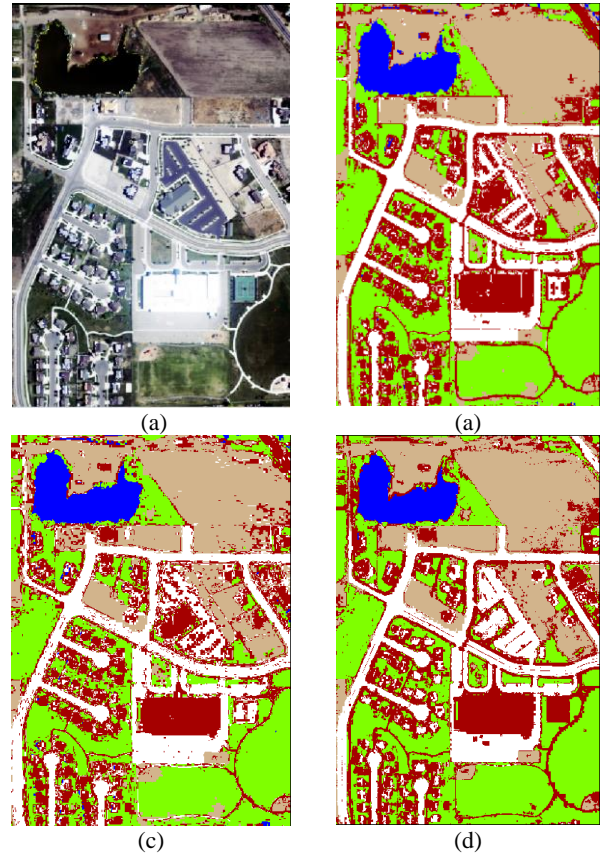
Figure A-2. Study area 2 classification results
(a) RGB image, (b) Neural Networks, (c) Nearest Neighbour
Object-based, and (d) Maximum Likelihood Pixel-based.

Classified Data	Bare soil	Building	Road	Vegetation	Water	Total Row	Producers Accuracy	Users Accuracy
Bare soil	**39**	5	1	1	0	46	69.64%	84.78%
Building	9	**130**	16	0	0	155	84.42%	83.87%
Road	2	17	**65**	0	0	84	78.31%	77.38%
Vegetation	6	2	1	**50**	2	61	98.04%	81.97%
Water	0	0	0	0	**54**	54	96.43%	100.00%
Total column	56	154	83	51	56	**400**	-	-
Overall accuracy **84.50%** Overall kappa statistics **0.7941**								

Table A-1. Confusion matrix for study area 1 using Nearest
Neighbour object-based classification.

Classified Data	Bare soil	Building	Road	Vegetation	Water	Total Row	Producers Accuracy	Users Accuracy
Bare soil	**50**	26	9	3	0	88	89.29%	56.82%
Building	3	**116**	9	0	0	128	75.32%	90.63%
Road	1	10	**62**	0	2	75	74.70%	82.67%
Vegetation	2	2	3	**48**	0	55	94.12%	87.27%
Water	0	0	0	0	**54**	54	96.43%	100.00%
Total column	56	154	83	51	56	**400**	-	-
Overall accuracy **82.50%** Overall kappa statistics **0.7729**								

Table A-2. Confusion matrix for study area 1 using Maximum
likelihood pixel-based classification.

Classified Data	Bare soil	Building	Road	Vegetation	Water	Total Row	Producers Accuracy	Users Accuracy
Bare soil	**76**	1	0	2	0	79	59.38%	96.20%
Building	25	**64**	4	5	0	98	92.75%	65.31%
Road	12	3	**64**	1	0	80	94.12%	80.00%
Vegetation	15	1	0	**76**	0	92	90.48%	82.61%
Water	0	0	0	0	**51**	51	100.00%	100.00%
Total column	128	69	68	84	51	**400**	-	-
Overall accuracy **82.75%** Overall kappa statistics **0.7833**								

Table A-3. Confusion matrix for study area 2 using Nearest
Neighbour object-based classification.

Classified Data	Bare soil	Building	Road	Vegetation	Water	Total Row	Producers Accuracy	Users Accuracy
Bare soil	**82**	2	1	1	0	86	64.06%	95.35%
Building	22	**54**	4	5	6	91	78.26%	59.34%
Road	5	13	**63**	1	0	82	92.65%	76.83%
Vegetation	19	0	0	**77**	0	96	91.67%	80.21%
Water	0	0	0	0	**45**	45	88.24%	100.00%
Total column	128	69	68	84	51	**400**	-	-
Overall accuracy **80.25%** Overall kappa statistics **0.7507**								

Table A-4. Confusion matrix for study area 2 using Maximum
likelihood pixel-based classification.