Hue University Journal of Science: Agriculture and Rural Development
pISSN: 2588-1191; eISSN: 2615-9708
Vol. 132, No. 3C, 2023, P. 19–35, DOI: 10.26459/hueunijard.v132i3C.7141

USING GIS, EUROSAT DATASET, AND CONVOLUTIONAL NEURAL NETWORKS IN RESIDENTIAL LAND DETECTION IN THUA THIEN HUE PROVINCE, VIETNAM

Ho Viet Hoang[1] *, Tran Thi Phuong[1, 2], Nguyen Bich Ngoc[1]

[1] University of Agriculture and Forestry, Hue University, 102 Phung Hung St., Hue, Vietnam
[2] Centre for Climate Change Study in Central Vietnam, University of Agriculture and Forestry, Hue University, 102 Phung Hung St., Hue, Vietnam

* Correspondence to Ho Viet Hoang <hoviethoang@huaf.edu.vn>
(*Received: March 13, 2022; Accepted: July 10, 2023*)

Abstract. This paper evaluates the level of suitability of applying GIS, EuroSAT dataset, and Convolutional Neural Networks to recognize residential land in Thua Thien Hue, Vietnam. We employed EuroSAT dataset to train and validate the ResNet152 model. Then, the assessed model was used to detect residential land in the location. The classification accuracy assessment of the Resnet152 model trained with EuroSAT dataset is based on some metrics, including Overall Accuracy, Recall, Precision and F-score, and the residential land detection accuracy of trained Resnet152 was evaluated by using Overall Accuracy. The results show that the performance of Resnet152 over EuroSAT dataset acquires high quality, with the model's classification accuracy reaching 0.98 and all other metrics greater than 0.9. Besides, the model’s detection accuracy of the Resnet152 model and the ground-truth data is as high as 0.88. The most incorrect detection occurs between residential land and two other lands, namely industrial land and herbaceous vegetation. Finally, it is undeniable that this EuroSAT-based ResNet152 model has significant potential to apply in the central region of Vietnam.

Keywords: EuroSAT, ResNet152, residential land, Thua Thien Hue

1 Introduction

Remote sensing (RS) provides powerful, fast, simple, and practical techniques for precise and accurate change detection, mapping, and inventory monitoring [1]. The number of available RS images has increased significantly with the advent of a new generation of satellite missions, typically consisting of satellite constellations with short revisit times and high-resolution sensors [2]. Today, monitoring dynamic processes, estimation of biophysical parameters, and classification problems can be solved by using multiple data sources, namely WorldView-1, WorldView-2, GeoEye-1, Landsat, and Sentinel [3–5]. Consequently, analysts can use these datasets for tasks such as updates on urban areas, desertification, agriculture land monitoring, soil mapping, water resources monitoring, and land use/land cover detection [6].

The analysis of residential land is essential in analyzing human-environment interactions [7]. Compared with traditional methods, such as aerial photography and field surveys, satellite data can be considered a powerful tool for monitoring residential areas and can be used by urban and regional planners at a fraction of the cost and time [8]. Many studies have been conducted to classify residential land and separate it from other land cover types [9, 10].

One of the major breakthrough technologies of 2013 was deep learning, one of the fastest-growing trends in big data analysis [11]. It has become state-of-the-art in many fields, such as image recognition. Deep learning models can learn information representations of raw input data and achieve high performance in many computer vision tasks because they can automatically learn information representations with multiple levels of abstraction [12]. Among these deep learning methods, Convolutional Neural Networks (CNNs) have attracted considerable attention and are used for a wide range of RS applications, including detection and change detection [13].

It is crucial to have high-quality datasets with appropriate categories for CNNs in order to maximize their performance [14]. Thus, a large quantity of training data is required to feed this model. One of the available RS datasets is EuroSAT, which can be utilized for CNNs to improve geographic maps by detecting changes in land use and land cover [15]. However, there have not been any studies that use or evaluate the capacity of using EuroSAT and CNNs models in the detection of land cover in Vietnam, in general, and the central region of Vietnam, in particular. According to the mentioned facts above, we implemented the assessment of what extent EuroSAT dataset and CNNs technique can be used to detect residential land in central Vietnam.

2 Methods

2.1 Study area

The province of Thua Thien Hue covers an area of 5033 km^2 and is located approximately in the centre of Vietnam on the North Central Coast. It borders Quang Tri province to the north, Quang Nam province and Da Nang to the south, Laos to the west, and the East Sea to the east [16].

In recent times, there has been an effort by the Thua Thien Hue province to expand and upgrade residential land toward becoming one of the direct-controlled municipalities of Vietnam by 2025. This initiative is supported by the implementation of several construction projects within Thua Thien Hue, including Huong Long ward, Xuan Thuy ward, and Phong Dien town [17–19]. Consequently, there is an urgent need for rapid detection of residential land in Thua Thien Hue province to enable effective monitoring of its changes.

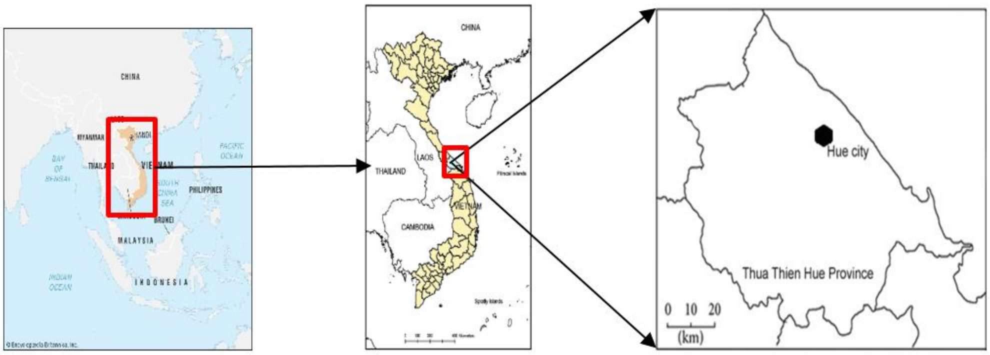


Figure 1. Location of Thua Thien Hue province

2.2 Data collection

EuroSAT

The EuroSAT dataset is a recently introduced dataset composed of 27,000 labelled and geo-referenced images. The images are derived from Sentinel-2 satellite imagery and encompass 13 spectral bands. In addition, the dataset comprises a total of 10 different classes, allowing classification and analysis of land use and land cover patterns [15].

Table 1. Spectral bands of EuroSAT data

No	Band	Spatial resolution (m)	Central wavelength (nm)
1	B01 - Aerosols	60	443
2	B02 - Blue	10	490
3	B03 - Green	10	560
4	B04 - Red	10	665
5	B05 - Red edge 1	20	705
6	B06 - Red edge 2	20	740
7	B07 - Red edge 3	20	783
8	B08 - NIR	10	842
9	B08A - Red edge 4	20	865
10	B09 - Water vapor	60	945
11	B10 - Cirrus	60	1375
12	B11 - SWIR 1	20	1610
13	B12 - SWIR 2	20	2190

Source: [15]

The images on the EuroSAT dataset are already preprocessed through two steps. First, satellite images are gathered over European cities distributed in over 34 countries. Second, a dataset comprising 27,000 labelled and georeferenced image patches is generated by using the acquired satellite images. The image patches are cropped into 64×64 pixels and have been checked manually [15].

In this study, we mainly focus on the residential class with a total number of 3,000 scenes. Some of the representative residential class scenes are illustrated below.

Sentinel-2

Sentinel-2 is a satellite system consisting of Sentinel-2A and Sentinel-2B, which offers a wide swath width of 290 km and a high revisit time. With one satellite, the revisit time is 10 days at the equator, while with two satellites, it is reduced to five days under cloud-free conditions, which results in 2–3 days at mid-latitudes [20]. It provides essential support for terrestrial observations, particularly in the areas of vegetation, soil and water cover, inland waterways and coastal areas assessment, land use and change detection mapping, disaster relief support, and climate change monitoring [21].

Figure 2. Sample patches of the residential class covered in the proposed EuroSAT dataset

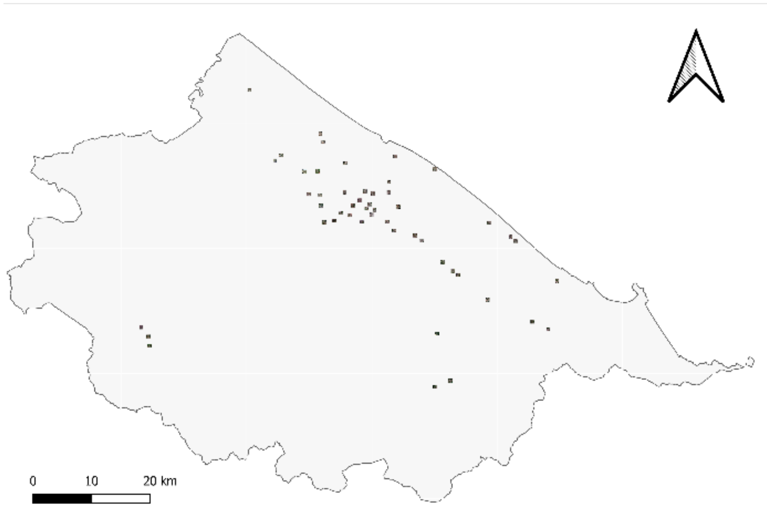


Figure 3. Location of ground-truth data extracted from Sentinel-2 image

In this research, the Sentinel-2 image was directly downloaded from https://scihub.copernicus.eu/. The collected image was a level-2A product acquired by Sentinel-2A on the 24th of July, 2022, at 03:15:31 AM. It was acquired over Tile 48QYD during Relative Orbit 118 and processed with PDGS Processing Baseline 04.00. Subsequently, a total of 50 of 64 × 64 pixel scenes of the residential land were derived randomly within Thua Thien Hue province's boundary. This kind of data was used as test data to calculate the detection accuracy of the EuroSAT-based CNNs model in the Thua Thien Hue province.

Field data

The identification of the residential class for each scene was quickly obtained through visual interpretation of high resolution in Google Earth. However, to ensure that the scenes extracted from the Sentinel-2 image were correctly residential class, we carried out a field survey. The fieldwork was conducted from 10/7 to 30/7/2022, and the main work was verifying residential land through visual interpretation and taking photos.

2.3 Convolutional Neural Networks

Residual networks (Resnet) is one kind of CNNs with similar structures but different depths [22]. Resnet introduces a structure called the Residual Learning Unit to mitigate the degradation of deep neural networks. The structure of this unit is a feedforward network with a link that inserts new inputs into the network and generates new outputs. The main advantage of this unit is that it achieves better classification accuracy without increasing the complexity of the model [23].

Resnet152 obtains the best accuracy among the members of the Resnet family [22]. Figure 4 illustrates the basic architecture of Resnet152.

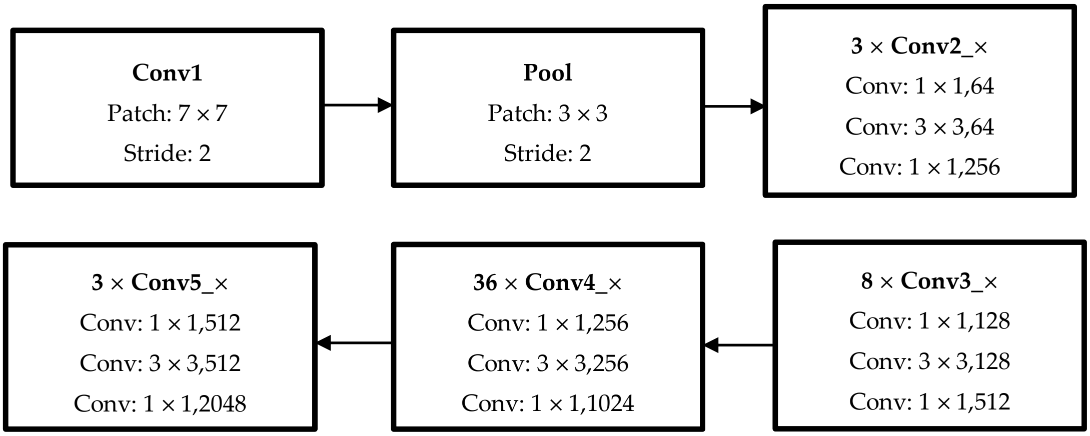


Figure 4. Resnet152 architecture [22]

In this study, we split scenes within the EuroSAT dataset into training data (80%) and validation data (20%). Then, the Resnet152 model was implemented to learn how residential land is over the training data.

2.4 Accuracy assessment

Classification accuracy of the Resnet152 trained with EuroSAT dataset

In order to evaluate the classification accuracy of the model, researchers introduced Overall Accuracy (OA) [25]. The value of OA ranges from 0 to 1. Higher values mean higher accuracy. The formula of OA is presented below.

$$OA = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} \tag{1}$$

Although accuracy is the most widely used and reliable metric in classification, it may not always be an adequate metric for evaluating the performance of a model. In cases where the class with the highest number of samples in the dataset has low accuracy, other classes with high accuracy may be overshadowed. This can lead to an incomplete evaluation of the model's performance. Therefore, it is important to consider additional metrics, namely Precision, Recall, and F-score that can provide a more nuanced assessment of the model's performance [26]. These metrics take into account the distribution of samples across different classes and can provide insights into the model's ability to correctly identify both positive and negative cases.

The Recall measures the proportion of actual positive cases that are correctly identified by the model, while the Precision measures the proportion of positive predictions made by the model that are actually correct. They are calculated as follows:

$$Recall = \frac{TP}{TP + FP} \tag{2}$$

$$Precision = \frac{TP}{TP + FN} \tag{3}$$

where TP is True Positive; FP is False Positive; FN is False Negative.

The F-score is the harmonic mean of the Precision and the Recall and provides a more balanced evaluation of the model's performance. It is calculated as follows:

$$\mathrm{F}_{score} = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{4}$$

If the F-score is low, this could indicate that the model is either unable to identify all positive cases (low Recall) or incorrectly identifies many negative cases as positive (low Precision). In other words, the model is not able to effectively capture the features of the data and does not provide accurate predictions. On the other hand, if the F-score is high, this indicates that the model correctly identifies both positive and negative cases and provides reliable predictions.

Residential land detection accuracy of trained Resnet152 in Thua Thien Hue province

To evaluate the accuracy of the residential land detection, we generated the value of OA based on comparing the predicted scenes obtained from the Resnet512 model and ground-truth data.

3 Results and discussion

3.1 Assessment of the performance of EuroSAT-based Resnet152 model

As mentioned above, the EuroSAT dataset contains 27,000 labelled image patches, divided into 10 classes, namely, annual crop, forest, herbaceous vegetation, highway, industrial, pasture, permanent crop, residential, river, and sea/lake. The specific data for each class are described in Figure 5. Among 3,000 images of the residential land in the EuroSAT dataset, there are 2,400 residential images used as training data and 600 used as validation data in the Resnet152 model.

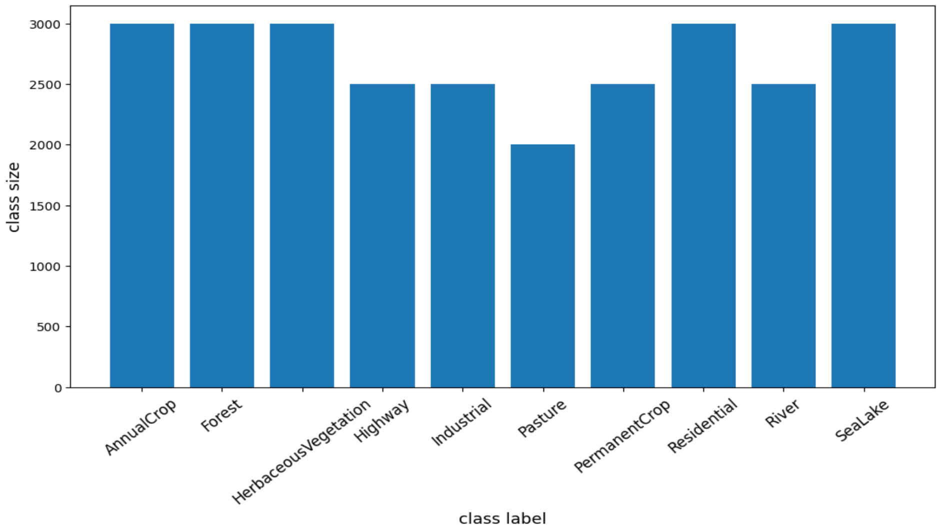
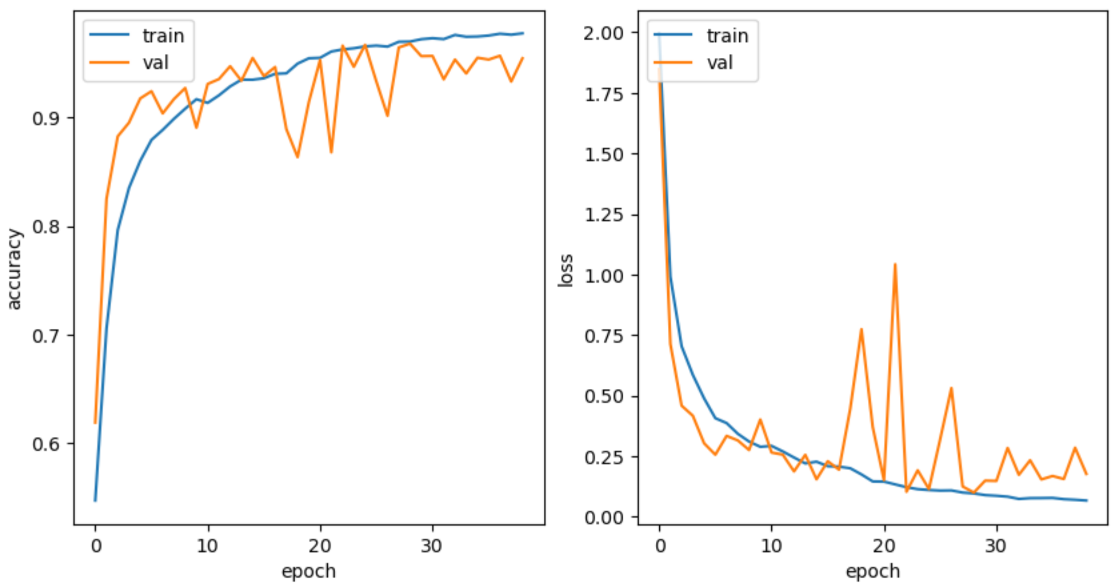


Figure 5. EuroSAT class distribution

After 39 epochs of the training process, the Resnet152 model was completed with an accuracy of 0.98 and a loss of under 0.03, with a stable learning process, as can be seen from the accuracy graph (Figure 6a) and the loss graph (Figure 6b). Importantly, both accuracy scores grow during training, and both loss scores get smaller.



a) b)

Figure 6. Training accuracy graph (a) and training loss graph (b) for the Resnet152 model

The ability of the Resnet152 model to achieve high accuracy and reduce loss to very low figures shows that the underfitting of the CNNs model does not happen. In other words, the training set has enough observation while compared with variables. Consequently, the Resnet152 model can find out the relationship between input data and the response variables because the model is not too complex to model the EuroSAT data. In addition, this result also indicates that the Resnet152 algorithm is able to discover the specific pattern between training and validation set variables in the high-dimensional dataset or significant input variables. Therefore, the graph shows that the Resnet152 model has enough potential if appropriately implemented.

Table 2 illustrates the metrics for assessing the performance of the Resnet152 model with the EuroSAT dataset. It can be seen that the values of Precision, Recall and F-score of all land cover categories are higher than 0.9. This indicates that the model is performing well in terms of precision and recall, striking a balance between these metrics. Precision, Recall, and F-score for the residential land are 0.933022, 0.998333, and 0.964573, respectively. These figures are more remarkable than those of some of the land use classes in the EuroSAT dataset, such as annual crop, herbaceous vegetation, and permanent crop. It means that the model provides more reliable predictions for residential land than others.

Table 2. Results of Resnet152 model

No	Class	Precision	Recall	F-Score
1	Annual Crop	0.960067	0.961667	0.960866
2	Forest	0.988468	1.000000	0.994200
3	Herbaceous Vegetation	0.950820	0.966667	0.958678
4	Highway	0.976744	0.924000	0.949640
5	Industrial	0.979210	0.942000	0.960245
6	Pasture	0.982051	0.957500	0.969620
7	Permanent Crop	0.943320	0.932000	0.937626
8	Residential	0.933022	0.998333	0.964573
9	River	0.976190	0.984000	0.980080
10	Sea/Lake	0.998328	0.995000	0.996661

Figure 7. Examples for predictions and ground truth labelling images (inside parentheses) in validation data

In Figure 7, we can observe multiple outcomes resulting from the application of the trained Resnet152 model on the validation data. These outcomes are the result of the model's ability to accurately predict the target variable.

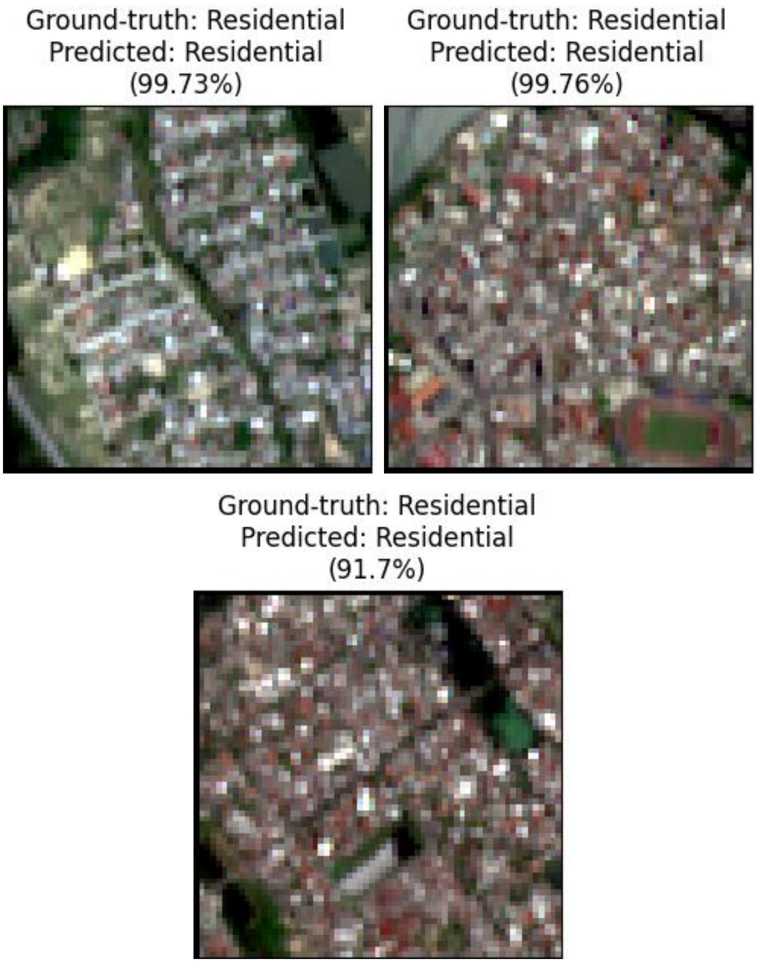
3.2 Assessment of detection accuracy of Resnet152 model in Thua Thien Hue province

After assessing the performance of the Resnet152 model based on the EuroSAT dataset, we evaluated the capacity of applying this trained model in Thua Thien Hue province, Vietnam. The confusion is depicted in Table 3.

Table 3. Confusion about residential class detection of trained Resnet152 model in Thua Thien Hue province

		Ground-truth data		OA
		Residential	**Others**	
Prediction	**Residential**	44	6	0.88

Table 3 presents the results of residential class detection of the Resnet152 model over Thua Thien Hue province. It can be clearly seen that the overall accuracy obtained after testing data is quite high at 0.88. In 50 testing images, 44 images predict correctly, while 6 images do not.



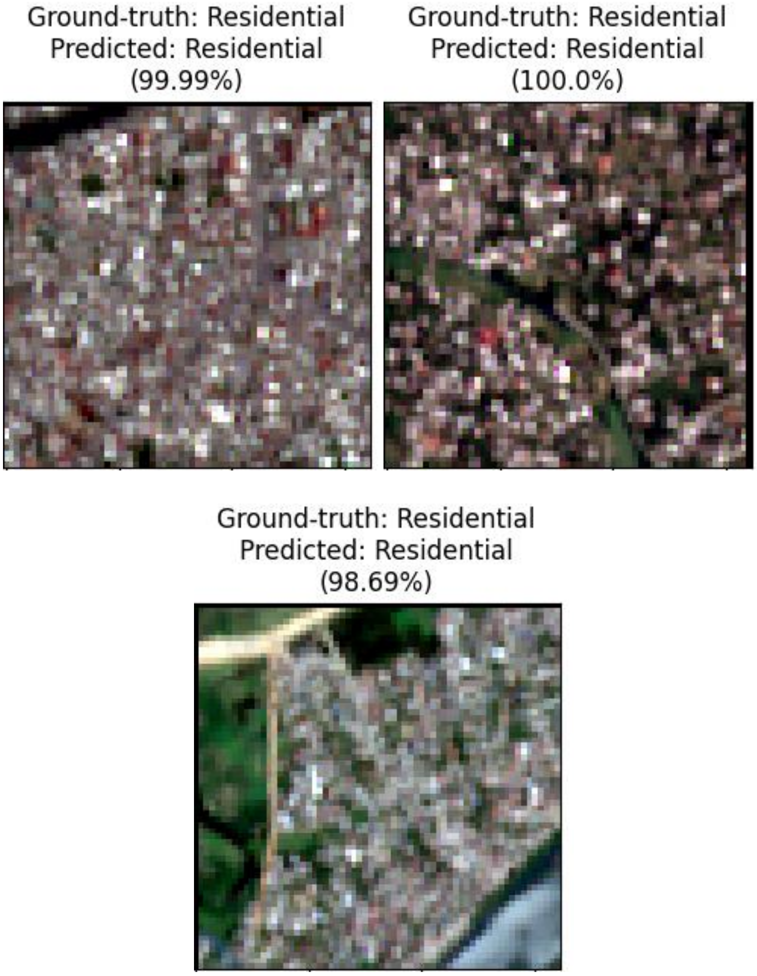
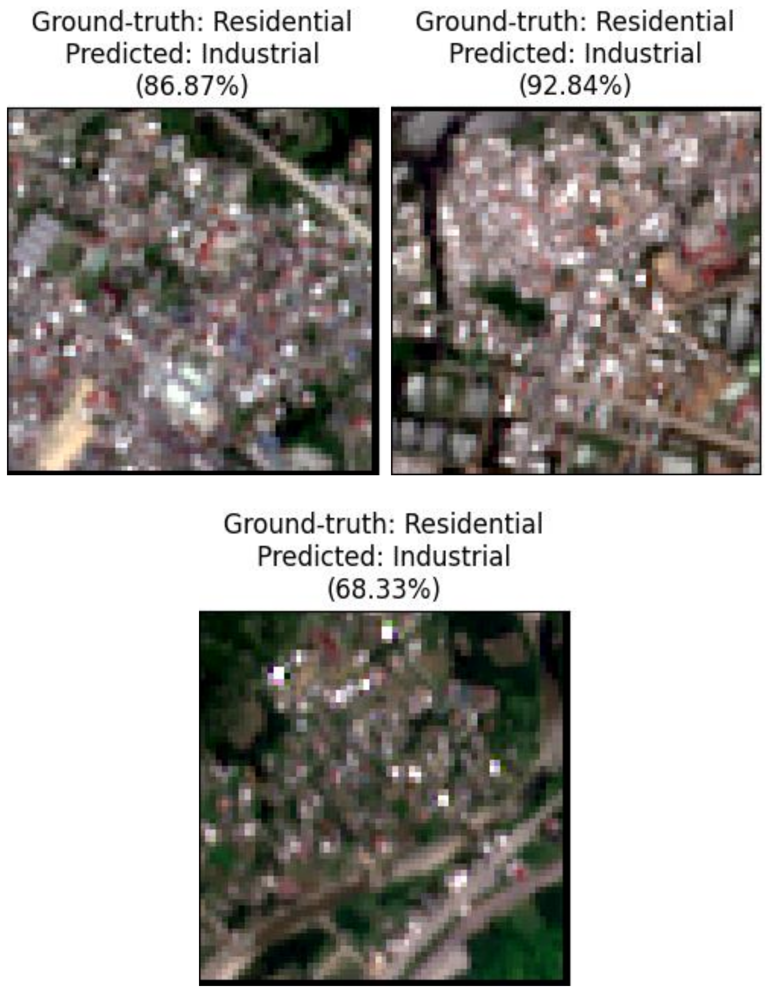


Figure 8. Correct prediction of resiential land in Thua Thien Hue province, Vietnam

Figure 8 shows a number of the predictions that are accurate for residential land within Thua Thien Hue province, Vietnam. Besides, it can be seen that the probability of these predictions is mostly high.

On the other hand, there are six wrong predictions. Among these, four predictions are assigned to the industrial class, while two others predict herbaceous vegetation class. The misleading prediction between the residential class and the industrial class is understandable when the structure and construction materials of these classes are quite similar. Meanwhile, the cause of the incorrection between the residential class and herbaceous vegetation may relate to the small-scale local residential areas in remote or mountainous regions that are normally covered with woody trees.


Ground-truth: Residential
Predicted: Industrial
(86.87%)
Ground-truth: Residential
Predicted: Industrial
(92.84%)
Ground-truth: Residential
Predicted: Industrial
(68.33%)

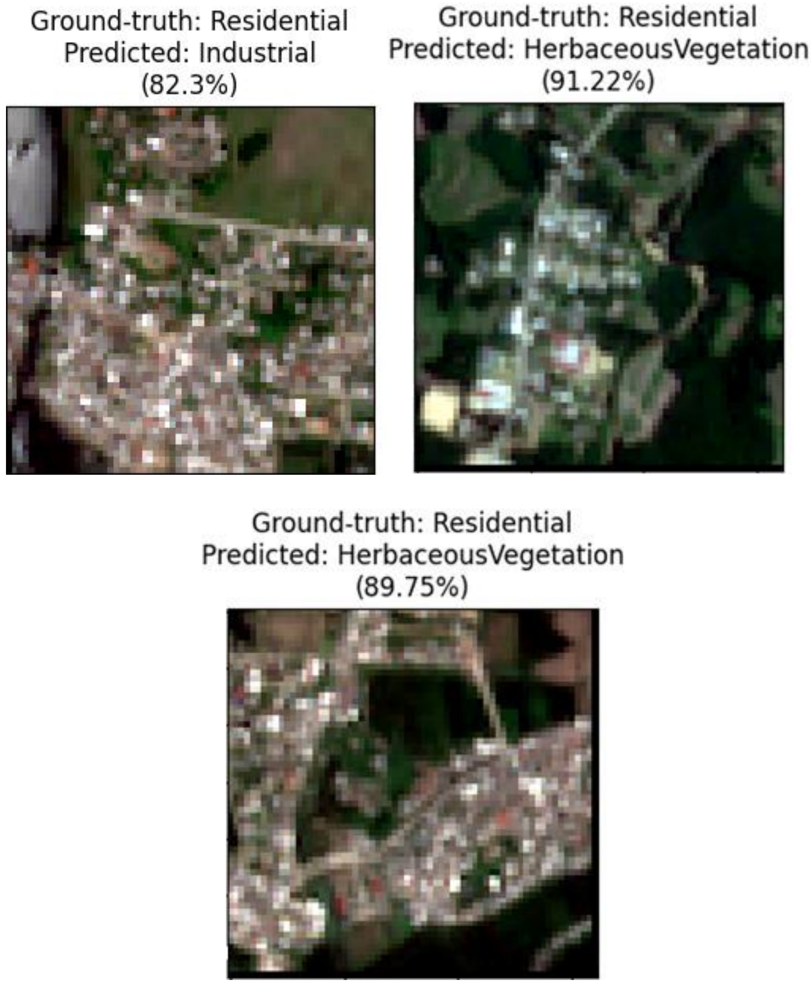


Figure 9. Incorrect prediction of resiential land in Thua Thien Hue province, Vietnam

4 Conclusion

With this paper, we successfully performed Resnet152-based CNNs fed by the EuroSAT dataset. The Resnet152 model demonstrates high accuracy, achieving a score of 0.98, a low loss rate (under 0.03), and high Recall, Precision and F-score values when applied to a qualified subset of the EuroSAT dataset. Furthermore, the combination of CNNs and the EuroSAT dataset yields a relatively high accuracy score of 0.88 in detecting residential areas in Thua Thien Hue province. However, some misclassification errors persist between the residential, industrial, and herbaceous vegetation classes. Overall, the results of this study indicate a promising potential for using the EuroSAT dataset and the CNNs model in the detection of residential land in the central region of Vietnam.

References

1. Rinaldi, Laura, Vincenzo Musella, Annibale Biggeri, and Giuseppe Cringoli (2006), New insights into the application of geographical information systems and remote sensing in veterinary parasitology, *Journal of Geospatial health*, 1(1), 33–47.
2. Tuia, Devis, Claudio Persello, and Lorenzo Bruzzone (2016), Domain Adaptation for the Classification of Remote Sensing Data: An Overview of Recent Advances, *IEEE Geoscience and Remote Sensing Magazine*, 4(2), 41–57. doi: 10.1109/mgrs.2016.2548504.
3. Amorós-López, Julia, Luis Gómez-Chova, Luis Alonso, Luis Guanter, Raúl Zurita-Milla, José Moreno, and Gustavo Camps-Valls (2013), Multitemporal fusion of Landsat/TM and ENVISAT/MERIS for crop monitoring, *International Journal of Applied Earth Observation and Geoinformation*, 23, 132–41. doi: 10.1016/j.jag.2012.12.004.
4. Pacifici, F., Longbotham, N., and Emery, W. J. (2014), The Importance of Physical Quantities for the Analysis of Multitemporal and Multiangular Optical Very High Spatial Resolution Images, *IEEE Transactions on Geoscience and Remote Sensing*, 52(10), 6241–56. doi: 10.1109/TGRS.2013.2295819.
5. Walker, J. J., de Beurs K. M., Wynne, R. H., and Gao, F. (2012), Evaluation of Landsat and MODIS data fusion products for analysis of dryland forest phenology, *Remote Sensing of Environment*, 117, 381–93. doi: 10.1016/j.rse.2011.10.014.
6. Yassine, H., K. Tout, and M. Jaber (2021), Improving Lulc Classification from Satellite Imagery Using Deep Learning – Eurosat Dataset, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLIII-B3-2021*, 369–76. doi: 10.5194/isprs-archives-XLIII-B3-2021-369-2021.
7. Meng, X. L., Nate Currit, Le Wang, Yang, X. J. (2012), Detect residential buildings from lidar and aerial photographs through object-oriented land-use classification, *Photogrammetric Engineering RS and Remote Sensing*, 78(1), 35–44.
8. Al-Dail, M. A. (1998), Change Detection in Urban Areas using Satellite Data, *Journal of King Saud University - Engineering Sciences*, 10(2), 217–27. doi: 10.1016/s1018-3639(18)30697–4.
9. Estoque, Ronald C, and Yuji (2015), Classification and change detection of built-up lands from Landsat-7 ETM+ and Landsat-8 OLI/TIRS imageries: A comparative assessment of various spectral indices, *Journal of Ecological indicators Murayama*, 56, 205–17.

10. Ezimand, Keyvan, AA Kakroodi, and Majid (2018), The development of spectral indices for detecting built-up land areas and their relationship with land-surface temperature, *Journal of International journal of remote sensing Kiavarz*, 39(23), 8428–49.

11. Zhu, X. X., Tuia, D., Mou, L., Xia, G. S., Zhang, L., Xu, F. and Fraundorfer, F. (2017), Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources, *IEEE Geoscience and Remote Sensing Magazine*, 5(4), 8–36. doi: 10.1109/MGRS.2017.2762307.

12. Li, Ying, Haokui Zhang, Xizhe Xue, Yenan Jiang, and Qiang Shen (2018), Deep learning for remote sensing image classification: A survey, *WIREs Data Mining and Knowledge Discovery*, 8(6), e1264. doi: 10.1002/widm.1264.

13. Liu, F., L. Jiao, Tang, X., Yang, S., Ma, W. and Hou, B. (2019), Local Restricted Convolutional Neural Network for Change Detection in Polarimetric SAR Images, *IEEE Transactions on Neural Networks and Learning Systems*, 30(3), 818–33, doi: 10.1109/TNNLS.2018.2847309.

14. Russakovsky, Olga, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, et al. (2015), ImageNet Large Scale Visual Recognition Challenge, *International Journal of Computer Vision*, 115(3), 211–52. doi: 10.1007/s11263-015-0816-y.

15. Helber, Patrick, Benjamin Bischke, Andreas Dengel, and Damian Borth (2017), *EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification*, doi: 10.1109/JSTARS.2019.2918242.

16. Wikipedia (2023), Thừa Thiên Huế province. Accessed 1/26. https://en.wikipedia.org/wiki/Th%E1%BB%ABa_Thi%C3%AAn_Hu%E1%BA%BF_provine.

17. Thua Thien Hue People's Committee (2016), *Decision No. 950/QD-UBND approving the detailed construction planning (1/2000 scale) of Huong Long area, Hue city, Thua Thien Hue province*, Thua Thien Hue.

18. Thua Thien Hue People's Committee (2016), *Detailed construction planning (1/2000 scale) of Thuy Xuan area, Hue City, Thua Thien Hue province,* Thua Thien Hue.

19. Thua Thien Hue People's Committee (2016), *General planning of Phong An new urban area, Phong Dien district, Thua Thien Hue province to 2030,* Thua Thien Hue.

20. ESA (2023), *Overview of Sentinel-2*, https://sentinel.esa.int/web/sentinel/missions/sentinel-2.

21. Grivei, A. C., Neagoe, I. C., Georgescu, F. A., Griparis, A., Vaduva, C., Bartalis, Z. and Datcu, M. (2020), Multispectral Data Analysis for Semantic Assessment—A SNAP Framework for Sentinel-2 Use Case Scenarios, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 4429–42. doi: 10.1109/JSTARS.2020.3013091.

22. Nguyen, Long D., Dongyun Lin, Zhiping Lin, and Jiuwen Cao (2018), Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation, *IEEE International Symposium on Circuits and Systems (ISCAS)*, 1–5.
23. Szegedy, Christian, Sergey Ioffe, Vincent Vanhoucke, and Alexander Alemi (2017), *Inception-v4, inception-resnet and the impact of residual connections on learning*, Paper presented at the Proceedings of the AAAI conference on artificial intelligence.
24. Fathololoumi, S., Firozjaei, M. K., Li, H. and Biswas, A. (2022), *Surface biophysical features fusion in remote sensing for improving land crop/cover classification accuracy*, *Sci Total Environ*, 838 (Pt 3), 156520. doi: 10.1016/j.scitotenv.2022.156520.
25. Russell G. Congalton (1991), A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data, *Journal of Remote. Sens. Environ*, 37, 35–46.
26. Mehmet Akif Günen (2021), Performance comparison of deep learning and machine learning methods in determining wetland water areas using EuroSAT dataset, *Environmental Science and Pollution Research* (2022), 29, 21092–21106, doi: 10.1007/s11356-021-17177-z.