Automatic Feature Extraction from UAV Imagery using Deep Learning

Ranjan Negi*

Additional Directorate General Military Survey, Indian Army
*Corresponding Author's email: ranjannegi.377y@gov.in

Abstract

One of the most used technologies for mapping is UAV based mapping. Very accurate and feature rich large-scale maps can be prepared using the high-resolution images acquired through UAV platform. These images after ortho-rectification are processed in a GIS software for extracting various topographical features like roads, buildings, land parcels, water bodies, etc. to prepare the map. Extraction of map features from the images is usually done by digitizers by creating point, polyline, or polygon features to depict various topographical features. This task is often too cumbersome and time consuming. With the advent of AI and Deep Learning, cartographers can automate this process using machines to perform the feature extraction task to quickly convert an image to a map. This work is aimed at developing a deep learning model for extracting features from high-resolution UAV acquired imagery to generate quality feature extracted and classified map products. For this, a model based on U-Net architecture is developed to automate map generation from locally acquired UAV images. The model was employed to produce semantically segmented classes from the images for identifying and extracting features in study area selected. The evaluation of model and segmented images exhibited the potential of using neural networks in automatic feature extraction resulting in automating the map making process thereby reducing the project turnaround time from data acquisition to map delivery. Though there is a need felt for further research, the work undertaken indicates great prospect of using AI/ML for cartography and geospatial applications and may provide a potential new paradigm in digital cartography by opening new possibilities for map-making.

Keywords UAV imagery, digital cartography, feature extraction, deep learning.

Introduction

With the advent of technology, data acquisition for map making has become more rapid with use UAV for mapping. High resolution images captured by UAVs are popularly used for large-scale mapping. The benefits of UAV-based mapping include low cost, high spatial and temporal resolutions, fast acquisition and quicker re-flying. Extraction of topographical features from the images is normally done by digitizers manually which is time consuming and need a lot of manpower. To overcome this, automatic extraction method provides faster production of vector data. AI trained algorithms can be developed for automatic feature extraction however they are required to be customized for specific imagery dataset and purpose. There are various methods of automatic feature extraction that have been developed for example pixel-based algorithms, object-based algorithms, land-base, tensor flow, neural network, shadow and cognition (O Marena et al., 2021). Image segmentation is one such method which can be used for feature extraction from a high-resolution image.

Unlike satellite based low-resolution images, high-resolution UAV based images allow much more freedom in image segmentation for extracting region of interest (ROI). With the very high resolution, it becomes possible to classify pixels from a UAV-based image which shall aid in automated feature extraction (Zhao et al., 2018). Numerous researches have been done for utilizing deep learning-based image segmentation techniques, however this work focuses on using U-Net architecture based deep learning method for segmentation to automate map making process using locally acquired UAV images.

Materials and Methods

Study Area: The dataset used in this study has been locally acquired using a 42 MP optical sensor payload mounted on a UAV platform operated at about 120m altitude. The dataset selected has high resolution (3cm GSD) orthorectified images mosaiced together covering area about 4 sq. km. The area selected for the research is a village in Assam where predominantly the buildings or houses are having corrugated tin shed roofs. As the village has moderate to heavy foliage cover, many houses are partially covered under the trees making it challenging for extraction of actual footprint of such houses.

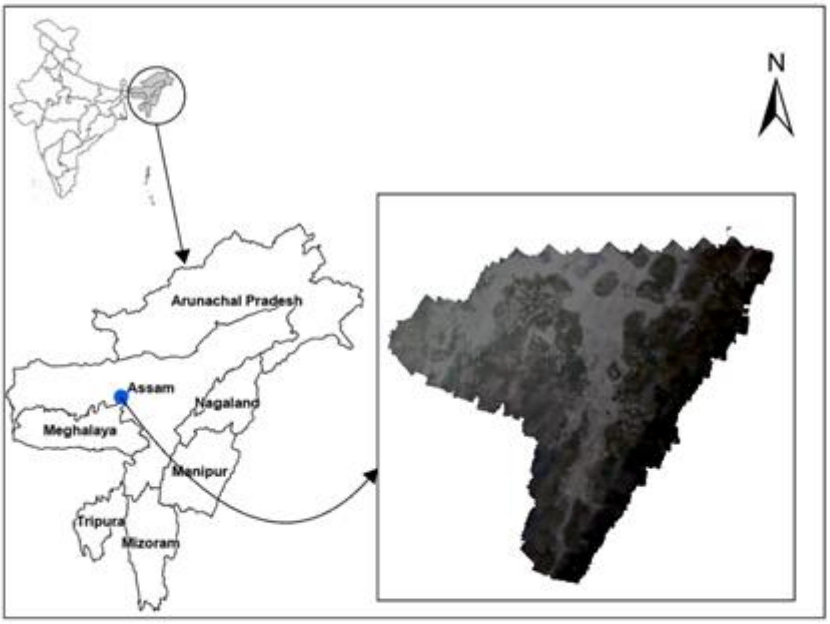

Fig. 1 Study Area Map.

Preparing Training Sample Dataset: Data pre-processing was carried out to prepare the data for image segmentation. For preparation of training data, a subset of the imagery was taken and polygons layer defining building/houses footprint, vegetation and background were created. The subset image with classified polygons was then rasterize to create the labelled image which will be used as ground truth for model. The subset image and the labelled image were use as training sample dataset.

In order to feed the training data to our model, the training datasets were converted into small patches of 256 x 256-pixel size. The patches of labelled image had pixel values corresponding to the colors assigned to each label class. Each patch of the labelled image was then categorized into classes viz, huts, vegetation, and background. After this, it was ensured that training images and labels are correctly aligned by visualizing both by using matplotlib plots.

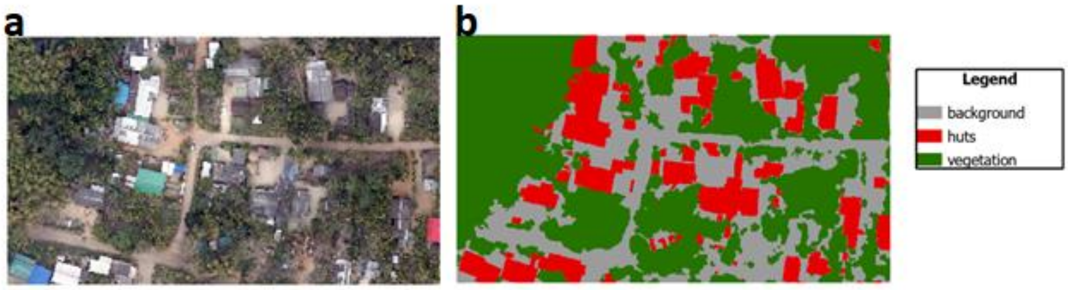

Fig. 2 Training Dataset. **(a)** Subset image. **(b)** Corresponding labelled image.

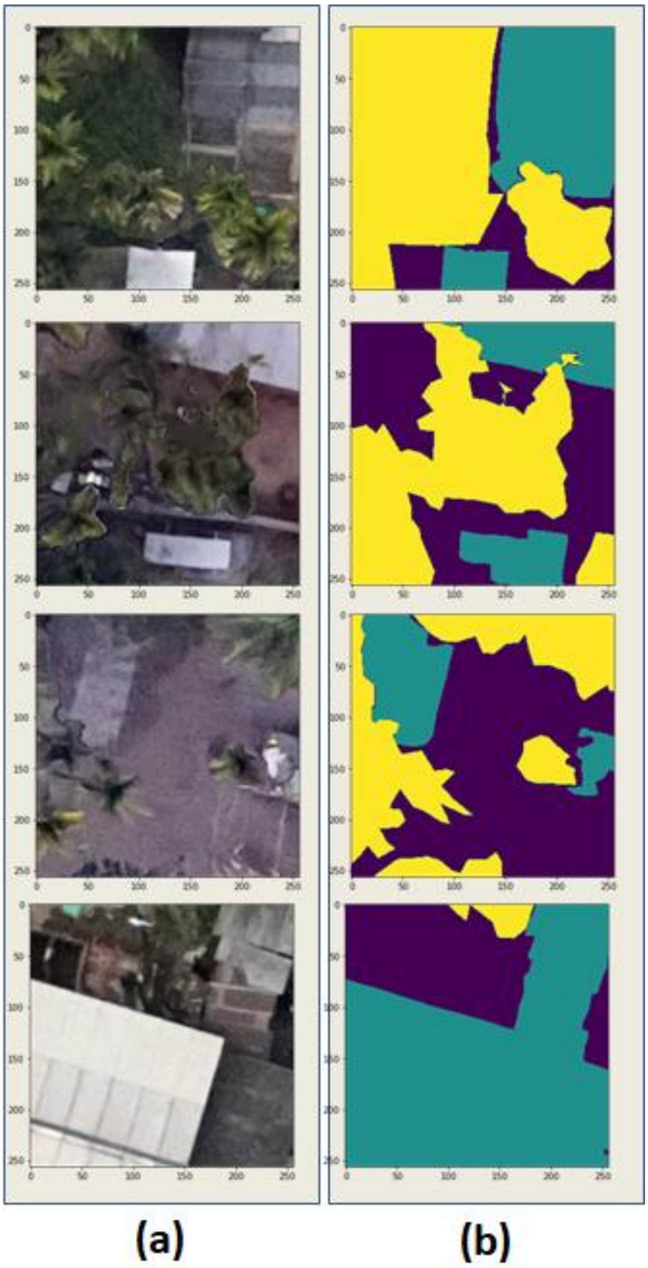

Fig. 3 Plot showing sample patches of (a) Original image and (b) Labelled image.

Model Architecture: The typical use of convolutional networks is for classification tasks, where the output to an image is a single class label (O Ronneberger et al., 2015). For model training popular convolutional neural network (CNN) architecture U-Net was used, which works with very few training images and yields more precise segmentations. U-Net consists of two critical paths; Contraction and Expansion. Contraction corresponds to general convolution Conv2D operations, where a 3 x 3 filter slide across the input image extracting features. MaxPooling2D down samples the input image size while maintaining feature information. On the other hand, Expansion corresponds to up sampling the down sampled image using transpose convolution Conv2DTranspose, to recover the input image spatial information. While up sampling, skip connections concatenate the features between the symmetric contraction and expansion layers. At the end of the decoder, one CNN layer is used with softmax as activation for getting the output image from the model.

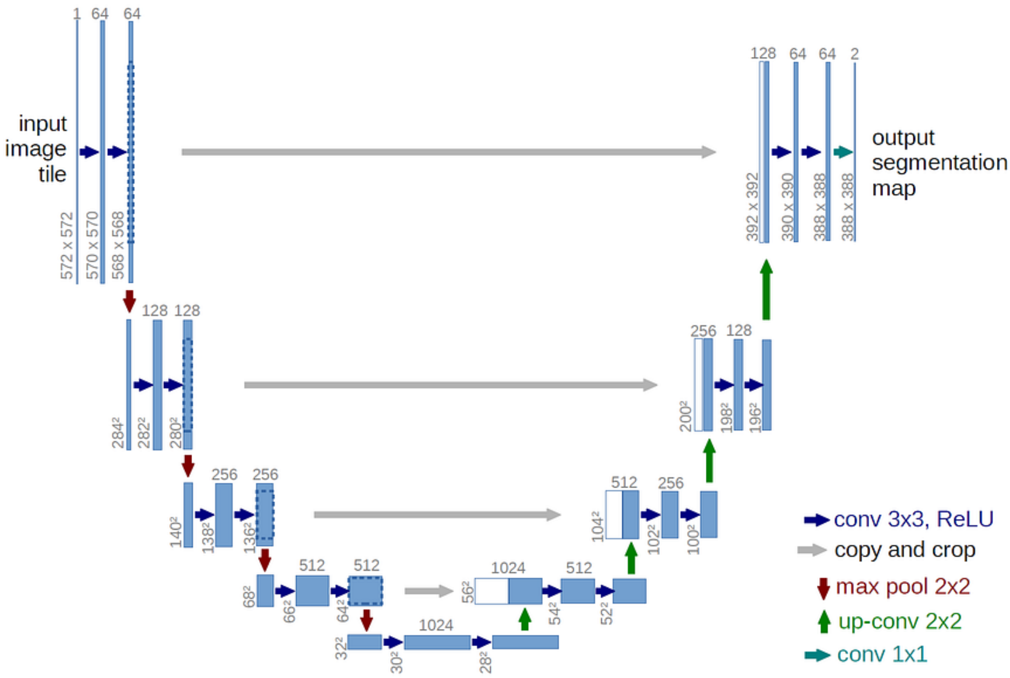

Fig. 4 U-Net Architecture as described in the reference paper by Ronneberger, et al. (2015).

Results

In the experiments the model was trained on the dataset created using the subset image. A total of 252 training sample images of size 256 x 256 and their corresponding labelled images were used for model training. The training was carried out in a total of 50 epochs and their results analyzed.

Prediction:

Accuracy of the model was also carried out by running the model over the test images and prediction made by the model in connection with segmentation of images into different classes were analyzed by plotting the predicted images.

Fig. 5 Training Vs. Validation Plot.

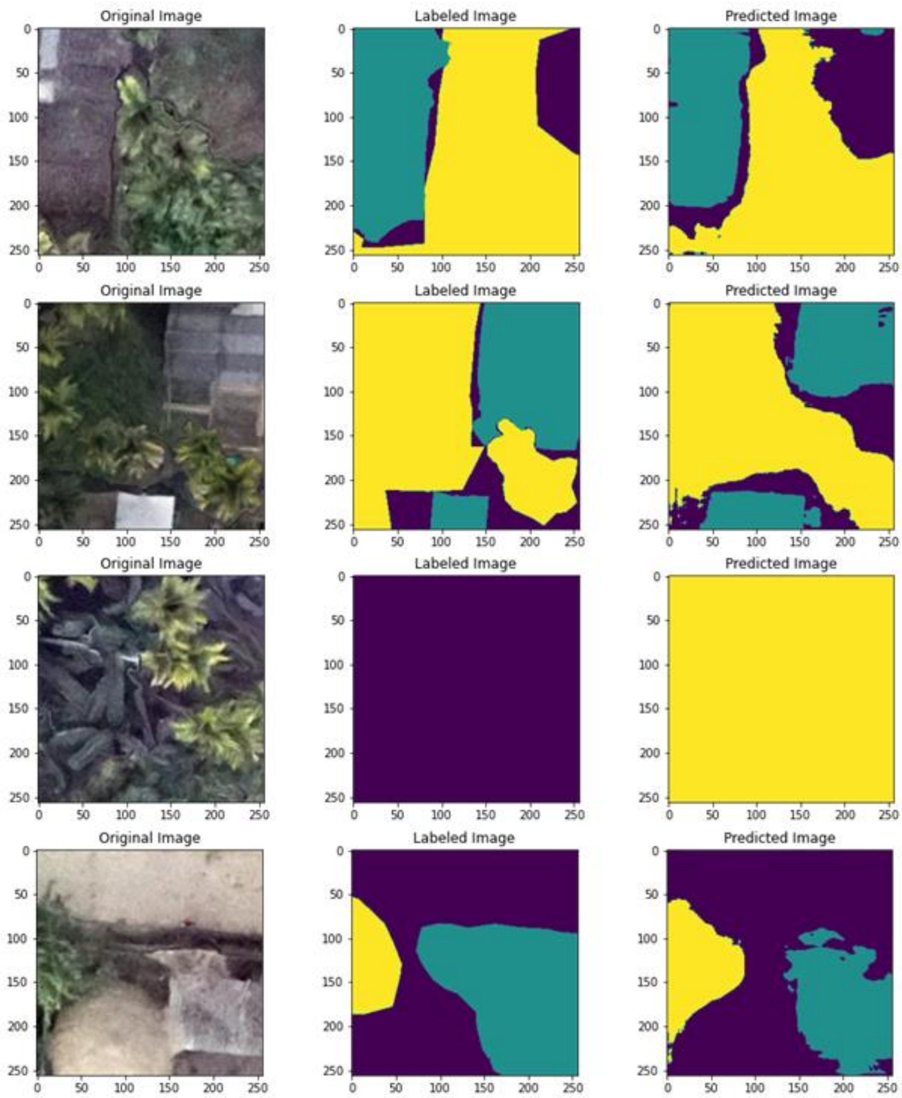

Fig. 6 Prediction of Image Segmentation done by model.

Discussion

The model was also run with the training image subset as input and the results were visualized for understanding the segmentation accuracy attained by the model.

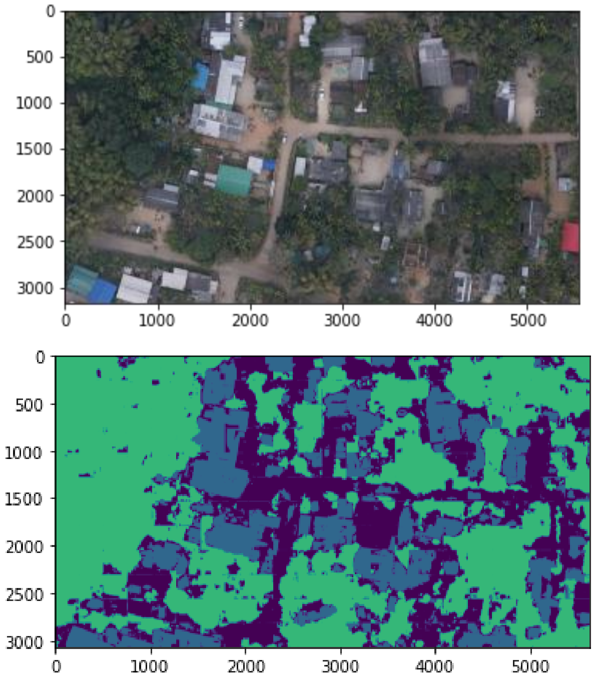

Fig. 7 Prediction of Image Segmentation done by model on Training Image.

Finally, the model was used to extract features from the other dataset selected for this work. The predicted images obtained as output by applying the model was processed using ArcGIS for creating polygons.

Fig. 8 Features extracted by the model.

Conclusions

The motivation for undertaking this work was to gain an understanding for designing, implementing and evaluating a deep learning network for automating feature extraction from UAV image data. The objective was to carry out semantic segmentation, to identify and extract huts in study area selected. The task was accomplished using developing a model based on U-Net architecture to produce semantically segmented classes from the images. The model was then implemented on sample images and the results obtained were analyzed. The evaluation of model and segmented images exhibited the potential of using neural networks in automatic feature extraction and at the same time hint towards taking future work of improvement to achieve the desired level of confidence.

Acknowledgements

My foremost appreciation to Col. K Joshil Raj, HoD Electrical and Mechanical Engineering Department, College of Military Engineering, Pune for his constant supervision, guidance, and encouragement throughout this project. I sincerely thank him for his generous involvement in advising, providing constructive feedback and full support. A very special thanks to Prof. Murali P, Associate Professor, BITS Pilani as well for his invaluable assistance and insights helping me to complete this work. Lastly, I would also like to thank all my friends and family members for encouraging and supporting me whenever I needed them.

References

O Marena, M S N Fitri, O A Hisam, A Kamil, Z M Latif (2021). Accuracy Assessment of Large-Scale Topographic Feature Extraction Using High Resolution Raster Image and Artificial Intelligence Method.

Zhao, Tiebiao & Yang, Yonghuan & Niu, Haoyu & Chen, YangQuan & Wang, Dong. (2018). Comparing U-Net convolutional networks with fully convolutional networks in the performances of pomegranate tree canopy segmentation. 64. 10.1117/12.2325570.

Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28

Shenglong Chen, Yoshiki Ogawa, Chenbo Zhao, Yoshihide Sekimoto, (2023). Large-scale individual building extraction from open-source satellite imagery via super-resolution-based instance segmentation approach, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 195.

Umas P, Liiv Innar (2016) Segmentation of Satellite Imagery using U-Net Models for Land Cover Classification, IEEE Access, Volume 4.

Long J, Shelhamer E, Darrell T (2015) Fully Convolutional Networks for Semantic Segmentation, UC Berkeley

Citation

Negi, R. (2024). Automatic Feature Extraction from UAV Imagery using Deep Learning. In: Dandabathula, G., Bera, A.K., Rao, S.S., Srivastav, S.K. (Eds.), Proceedings of the 43rd INCA International Conference, Jodhpur, 06–08 November 2023, pp. 243–249, ISBN 978-93-341-2277-0.