EXTRACTION OF MAN-MADE OBJECT FROM REMOTE SENSING IMAGES USING GABOR ENERGY FEATURES AND NEURAL NETWORKS

Abstract

This chapter presents a novel approach for man-made object extraction in Remote Sensing (RS) images. This paper focuses on the design and implementation of a system that allows a user to extract multiple objects such as buildings or roads from an input image without much user intervention. The framework includes five main stages: 1) Pre-processing Stage. 2) Extraction of Local energy features using edge information and Gabor filter followed by down sampling to reduce the redundant information. 3) Further reduction of the size of feature vectors using Wavelet decomposition. 4) Classification and recognition of man-made structures using Probabilistic Neural Network (PNN) 5) NDVI based post-classification refinement. Experiments are carried out on a dataset of 200 RS images. The proposed framework yields Overall Accuracy (OA) of 93%. Experimental results validate the effective performance of the suggested method for man-made objects extraction from RS images. Compared with other methods; the proposed framework exhibits significantly improved accuracy results and computationally much more efficient. Most notably, it has a much smaller input size, which makes it more feasible in practical applications.

Keywords: Remote Sensing Image, Man-made Object Extraction, Gabor Wavelets, Probabilistic Neural network

Authors

Md. Abdul Alim Sheikh
Department of Electronics and Communication Engineering
Aliah University
Kolkata, India
alim.sheikh16@gmail.com

I. INTRODUCTION

Remote Sensing (RS) is a key tool for extracting usable data from earth surface features. The rapid advancement of RS technology and high-speed imaging sensors has made it possible to obtain a large number of high-resolution RS images every day, and the availability of high-resolution RS imagery has expanded dramatically [1]. These valuable data have opened up a new window for understanding the earth surface and has opened the fields of exploration and application. These valuable data hold a lot of potential for meaningful and precise terrestrial object interpretation. Among various terrestrial object, man-made objects like building, road are the most important types of terrestrial objects. Object information extraction from RS images is a computer vision; photogrammetry and remote sensing task that plays a vital role in a wide range of RS applications that need the creation of maps or the study of urban changes.

The precise and up-to-date man-made objects e.g., buildings, roads have a strong positive correlation with economic development and the growth of nation. Roads and buildings are the significant man-made structure to be extracted automatically. Extraction of man-made objects like buildings, roads from RS Imagery plays a vital role in several RS applications, such as urban planning and reconstruction, change detection, environment monitoring and disaster management, estimation of human population, 3D city modelling, traffic management and land use analysis, real-estate management, illegal building survey, geographic information systems, and military reconnaissance and many other geospatial related applications like map generation and update etc. [2]-[7].

Although man-made objects like buildings and road detection can be achieved manually by human experts, but it is very time-consuming, labour-intensive, error prone and expensive to extract buildings from RS images. Manual analyses of satellite and aerial imagery were previously possible due to the low volume of images available; however, this is no longer the case. With the large amount of data, we deal with today; object extraction from RS images becomes an issue. As a result of the dramatic growth in the amount of RS imagery available in recent years, interpretating and identifying of terrestrial objects e.g., buildings, road in RS images has become a challenging problem at scale. The traditional image processing-based strategy, which is overly reliant on human feature extraction, is incapable of addressing the challenges of large-scale dataset interpretation and does not fulfil the requirements of nowadays practical applications [8]. As a result, numerous attempts have been made to develop automatic, accurate, and computationally fast methodologies to extract objects e.g., buildings and roads, from RS images [9].

However, extraction of buildings and other man-made structures accurately and efficiently from RS imagery is still a challenging task with several difficulties. The fundamental issue with these approaches is that the building is confused with other objects having similar spectral reflectance. A number of strategies for extracting man-made objects from RS imagery have been presented in recent years. Some comprehensive reviews on man-made objects extraction from satellite can be found in [10-12]. Buildings and roads are one of the most important groups of man-made objects and automatic/semi-automatic extraction of building and roads can minimize the human labour in the application of map generation and updates. The fundamental issue with these approaches is that the building is confused with

other objects having similar spectral reflectance. The diverse characteristics and nature of objects appearing in urban environment such as colour, sizes and shapes, material, and interference of building shadows and trees make precise and reliable building extraction more complex and challenging [12]. Many objects in high-resolution RS images, such as highways and parking lots, appear to be quite similar to buildings [13]. In practice, this has been the most significant stumbling block for RS applications.

In recent years, machine learning (ML) technologies have shown to be strong contenders in this area. DL has illuminated this issue with its revolutionary growth. Rather than manual extraction of task specific features for learning-based image analysis tasks, DL has shown how hierarchies of features can be extracted automatically during the training process and learned from the data itself [13]. The Convolutional Neural Network (CNN), and the Fully Convolutional Neural (FCN) network are widely used for these purposes [13].

This chapter focuses on the design and implementation of a system that allows a user to extract multiple objects such as a building or a road from an input image without much user intervention. A novel technique for detection and recognition of man-made and natural objects (e.g., vegetation, natural water body, and seashore) from RS images is presented. Most of the man-made object, such as buildings and roads, have regular shapes with largely straight lines and consistent texture, whereas natural objects, such as vegetation and lakes, have irregular shapes with disorderly boundaries and textures [12]. Compared to scenes with natural structures, photographs of man-made structures are far more likely to have more edges segments, vertical and horizontal straight-line segments, and the presence of these straight lines is not coincidental. We attempt to take use of some non-accidental features of photos from both groups as a result.

To aim this, an automatic manmade object detection and recognition method is proposed, based on Gabor wavelet and Neural Network (NN). The features are extracted using the Gabor wavelets followed by down sampling by a factor to reduce the redundant information. In addition, dimension reduction method is used to further reduce the size of the feature vectors. Finally, the features are applied to a classifier for recognition. Spectral information is also used as a source for man-made object extraction by eliminating trees. For further refining of the building extraction, NDVI is utilized to eliminate the tree-generated lines. The main novelties of the proposed approaches consist of

1. Capable to extract multiple objects e.g., roads and buildings.
2. Techniques based on the efficient differentiation of sharp edges present or absent serve as a discriminative characteristic in separating man-made items from natural objects.
3. Exploiting domain knowledge and local interaction to achieve correct classification percentages with high accuracy
4. Large size feature difficulty is avoided. The system is sped up by it. Real RS image numerical results provide to validate the suggested technique.
5. The proposed framework's high computational efficiency and ease of implementation demonstrate its promise for any object extraction from RS images.

The rest of the chapter is as follows: in section 2, the suggested technique for man-made object extraction is described in detail. Then experimental results are provided in

section 4.5 with a brief overview of dataset and experimental setup. Finally, section 4.6 summarizes the chapter with future scope.

II. PROPOSED APPROACH

The suggested method's data flow diagram is given in Fig. 1, which clearly illustrates the many steps. Pre-processing, feature extraction, categorization, and refinement of building structures make up the steps. The following sections include detailed descriptions of each stage.

1. **Pre-processing** A pre-processing pipeline is adopted to cope with input images of varying quality, resolution, and channels to remove of noises and undesirable objects. In view of both noise diminution and edge preservation, bilateral filtering and Histogram Equalization (HE) are performed to preprocess the input image [12]. Bilateral filter performs noise reduction and nonlinear smoothing on images by keeping the edge information by means of a nonlinear combination of nearby image values. Fig. 2(b) shows the pre-processing result after bilateral filtering on the image shown in Fig. 2(a).

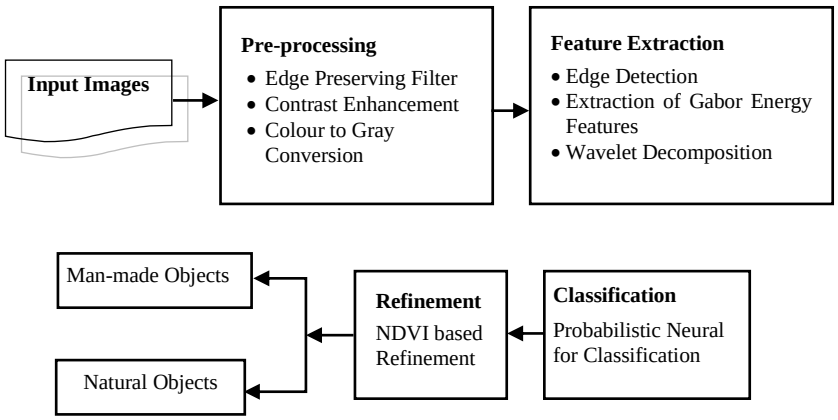


Figure 1: The Methodology of the Proposed Technique

Histogram Equalization (HE) technique is used for image enhancement which adjusts the intensity histogram to approximate a uniform distribution. Fig.2(c) shows a HE processed image. After enhancement, the colour images are converted to gray scale images for feature extraction.

(a) (b)

(C)

Figure 2: (a) Input Image of Emerging Suburban Area (b) Result of Bilateral Filtering (c) Histogram Equalization Processed Image

2. **Feature extraction:** Local energy features are extracted using edge information and Gabor wavelets followed by down sampling to reduce redundant information and wavelet decomposition technique.

    - **Edge detection:** The Sobel operator is employed to take advantage of the non-accidental occurrence of edges and straight-line segments. When compared to the other operators, this operator provides significantly larger output values for similar edges [12]. The presence of edge information of man-made and natural images is shown in Fig. 3.

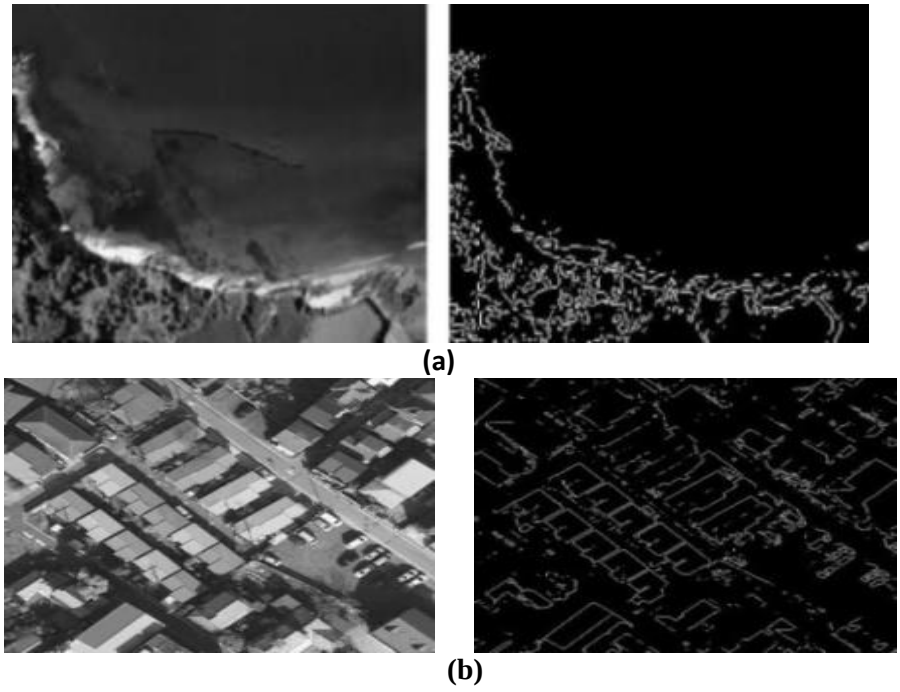


Figure 3: (A) Images and its Sobel Edge Output (A) Natural Scene (B) Manmade Structures

- **Feature extraction using gabor wavelets:** Gabor filters may be thought of as orientation and scale tunable edge and line detectors, making them an excellent tool for detecting geometrically limited linear features from RS imagery, such as buildings and roads. The invariance of Gabor filters to rotation, scale, and translation is its most significant advantage. They are also resistant to photometric disturbances such as illumination changes and image noise [14].

Gray-level images are directly used to extract Gabor filter-based features. Gabor functions can be obtained with a Gaussian window multiplied by a complex sinusoidal wave [14]. A 2-D Gabor function is defined as:

$$g_{f,\theta,\varphi,\sigma,\gamma}(x,y) = exp\left(-\frac{x^{'2}+\gamma^2 y^{'2}}{2\sigma^2}\right) exp\{2\pi f x^{'} + \varphi\} \quad (1)$$

Where $x^{'} = xcos\theta + ysin\theta$ and $y^{'} = -xsin\theta + ycos\theta$

Where *f* is the frequency of the sinusoidal factor, θ (between 0 and π) is the orientation, φ is the phase offset, σ denotes standard deviation and γ is the spatial aspect ratio which clarifies the ellipticity of the cooperation of the Gabor function.

Different filters can be generated with varying values for orientation and scale. As a result, a Gabor filters bank is created, which is made up of a set of Gaussian filters with various radial frequencies and orientations that cover the frequency domain. The whole frequency spectrum, both amplitude and phase are captured by the

Gabor filter family. In this step, forty-eight Gabor filters are employed in six scales and eight orientations as shown in Fig. 4.

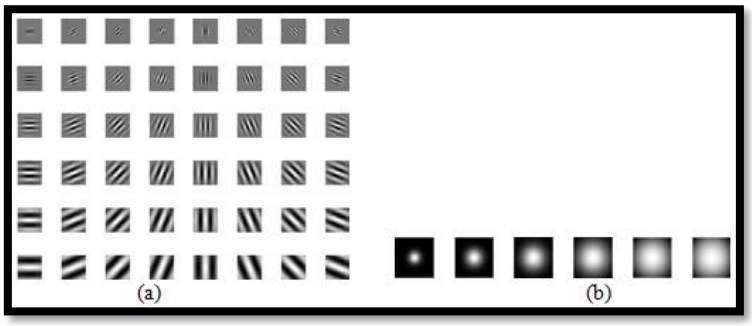


Figure 4: Gabor wavelets. (a) The real part of the Gabor kernel at scales = 6 and orientations = 8 (b) the magnitude of Gabor kernels at six different scales

- **Gabor feature illustration:** The Gabor wavelet representation of the image is obtained by convolving the image *I* of dimension M × N with every Gabor filter of the Gabor filter family as defined in Eqn. (2) at every pixel (*x; y*). Each pixel is distinguished from other pixels using the power spectrum of the filtered picture at that pixel's location.

$$r^2_{f,\theta,\varphi,\sigma,\gamma} = i(x,y) * g_{f,\theta,\varphi,\sigma,\gamma} =$$

$$\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} i(m,n)\, g(x-m, y-n) \qquad (2)$$

Where * is the convolution operation. To create the Gabor energy quantity, the filtered responses of the symmetric and anti-symmetric filters are combined. After that, the vector sum of the corresponding filtered outputs from each filter bank is used to create the Gabor energy matrix E for each orientation.

$$E_{f,\theta,\varphi,\sigma,\gamma}(x,y) = \sqrt{r^2_{f,\theta,\sigma,\gamma,0}(x,y) + r^2_{f,\theta,\sigma,\gamma,\frac{-\pi}{2}}(x,y)} \qquad (3)$$

Where $r^2_{f,\theta,\sigma,\gamma,0}(x,y)$ and $r^2_{f,\theta,\sigma,\gamma,\frac{-\pi}{2}}(x,y)$ are outputs of symmetric and anti-symmetric filters.

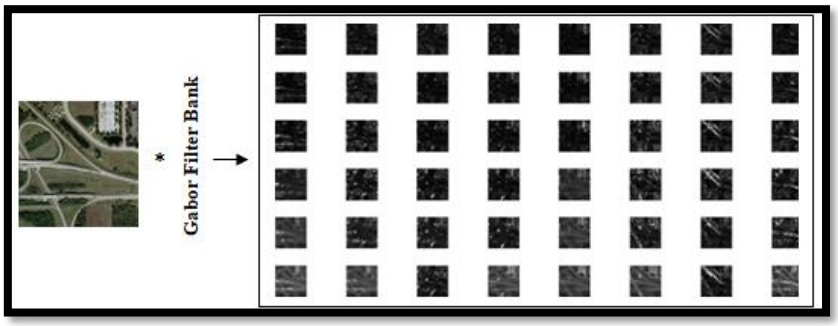


Figure 5: Results of Feature Extraction. To Create a Gabor Filter Image, the Input Image is convolved With the Gabor Filter Banks. The Gabor Filtered Images' Amplitudes at 6 Scales and 8 Orientations are Shown

The input images used in our experiments are 256×256 pixels in size. The Gabor wavelets depicted in Fig. 4 are used to extract the features. The Gabor wavelet representation of a sample image from our database is shown in Fig. 5. The location, size, and orientation characteristics of these representations match those of the Gabor wavelets in Fig. 4.

The feature images obtained from Gabor filters are further down sampled by a factor 8 to reduce the redundant information as the adjacent pixels in an image are usually highly correlated. The feature vector will have a size of (256×256×6×8)/(8×8) = 49,152 in total. After that, the vectors are normalized to have a zero mean and unit variance. In addition to down sampling, dimensionality reduction method is used to further minimize the size of the feature vectors.

- **Feature reduction by wavelet decomposition:** The size of extracted features in remote sensing image analysis is often large, which increases computing complexity and reduces system performance. To address these issues, Wavelet decomposition is used to reduce feature dimension and their redundancies to a level that is easy for applying to classifier. Here, at first the available feature data is transformed to a wavelet form with a substantial proportion of its total energy packed into a small number of transform coefficients. So, taking these few coefficients, while neglecting the rest, we can retain the image feature. This is how the transformed data is reduced to a lower dimensional space.

  Due to the sparseness of the wavelet transform, the important coefficients of the transformed data have a larger magnitude than the unimportant coefficients. Coif3 wavelet is used due to its easy implementation, fast speed, shorter filter and easy to describe small texture structure, good resolution and smooth traits. The feature size is reduced to 16×1 using sixth level decomposition.

3. **Classification:** A PNN network [13] is adopted to classify the input feature vectors into a specific class. PNN network is adopted for its many advantages e.g. a) training is easy and instantaneous b) training speed is many times quicker than back propagation c) an inherently parallel structure d) No local minima issues e) The ability to add or remove training data without requiring substantial retraining f) It is also resistant to instances with noise.

   The network structure in our proposed algorithm is illustrated in Fig. 6. The PNN used here has four layers: the input layer, Radial Basis Layer, the Competitive Layer, and the output layer.

   - **Input layer:** This layer's sole purpose is to distribute input to all neurons in the pattern layer. The black vertical bar in Fig. 6 represents the input vector, designated as P. It has an R×1 dimension. R = 16 in this study.

   - **Radial basis layer:** It calculates the vector distance between an input feature vector p and the weight vector resulting from each row in weight matrix W.

- **Competitive layer:** *C*ompetitive layer has no bias. The binary output of competitive function is denoted by *a2*. The vector '*a1*' is multiplied with layer weight matrix $LW_{2,1}$ in this layer, yielding an output vector *a2*.

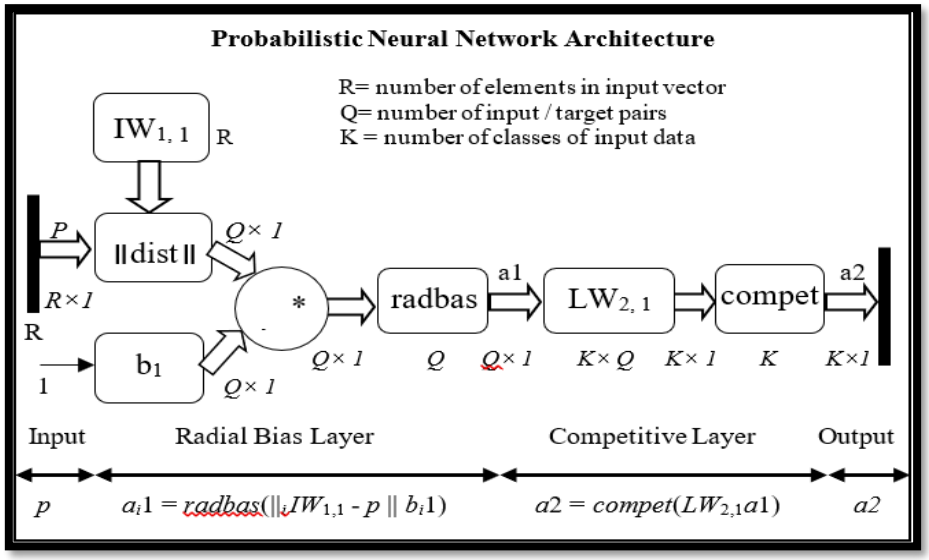


Figure 6: Architecture of the PNN, R=16, Q=108, K=2

The selected neural network has two hidden layers with 108 and 16 nodes between the input and output layers, respectively. It contains a single-node output layer that generates the classification decision and a 16-node input layer that accepts the 16 characteristics of each picture region (8 Gabor energy matrices).

4. **Refinement of results:** The initial man-made object extraction results produced in the previous step may contain some non-building objects, e.g. trees. The NDVI is one key parameter, which is used here to differentiate between vegetated and non-vegetated objects [12]. The NDVI values show how much green vegetation is present in each pixel. The classification is based on the simple assumption that objects with an NDVI of more than a specific value must be trees; and NDVI of man-made class is low. Trees were excluded from the categorization result generated in the previous stage because their NDVI values exceeded the threshold values. The NDVI is calculated as:

$$NDVI = \frac{NIR+RED}{NIR+RED} \qquad (6)$$

Where NIR: Near-infrared reflectance value; RED: Visible red reflectance value

Fig.7 indicates the refinement process employed for extraction of buildings and roads from images. The general rule for this refinement process is that man-made class will be found when NDVI is low, and vegetation class e.g. trees will be found NDVI is very high.

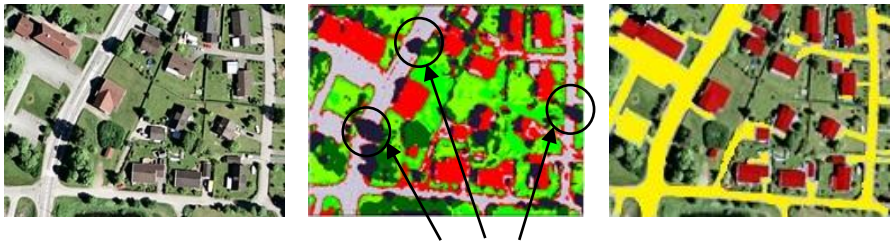


Figure 7: Input Image, Classified Image and Extracted Buildings and Roads after Refinement

Knowledge engineer, which is included in ERDAS Imagine Software, was used to create an expert system for post-classification refinement. Building class is discovered when the NDVI is less than -0.038, road class is found when the NDVI is less than 0.02 and tree class is found when the NDVI is larger than 0.1, as in our example.

III. DATASET AND EXPERIMENTAL RESULTS

To assess the suggested method, the experiment is carried out on a balanced set of two categories of 200 RS imagery of size 256×256 each. Among them, total of 120 images, 60 from each class, are used for training phase and 40 images, 20 man-made and 20 natural scene images, are used for testing as shown in Table 1. Remaining 40 images, 20 man-made and 20 naturals, are used for validation purpose. For creating the data set, we used selected parts (256×256 pixels) of sceneries from Massachusetts buildings dataset [31], Massachusetts roads dataset [31]. A ground truth (buildings/roads) map was manually labeled for each image in the data collection.

For training and testing the suggested approach, a PC with an Intel (R) Core(TM) i5-4590 CPU running at 3.30 GHz and 4GB of RAM is utilized. The method is realized in MATLAB. Few samples feature vector of man-made and natural class images of training and testing sets are tabulated in Table 4 and Table 5. We set the parameters of the Gabor filter bank as follows.

1. The No. of scales: 6
2. The No. of orientations: 8
3. The No. of rows and columns: 39
4. The factors of down-sampling along the rows and along the columns: 4

Through Gabor filtering, it has been shown that Gabor energy maps for man-made buildings typically exhibit strong orientation traits.

Table 1: Number of Samples per Class for the Training, Validation and Test Set

Land Cover	Data set (200 images)		
	Training	Testing	Validation
Man-made structures	60	20	20
Non-Manmade	60	20	20

1. **Qualitative evaluation:** For qualitative evaluation, proposed method is applied for extracting two types of manmade objects, i.e., building and roads from the images in databases. To demonstrate the effectiveness of the suggested method, object extraction results are shown here.

    - **Road network detection:** For assessing the effectiveness of the proposed technique for road object extraction, the qualitative segmentation results are presented on the different images from Massachusetts road dataset in Fig.8. Roads were mostly homogeneous and not disturbed by shadows or occlusions. The region surrounds man-made structures (roads) with near-green colors. The final result overlaid on the initial image and the ground truth data as shown in Fig.8. The algorithm is successfully detecting the whole road network with high accuracy.

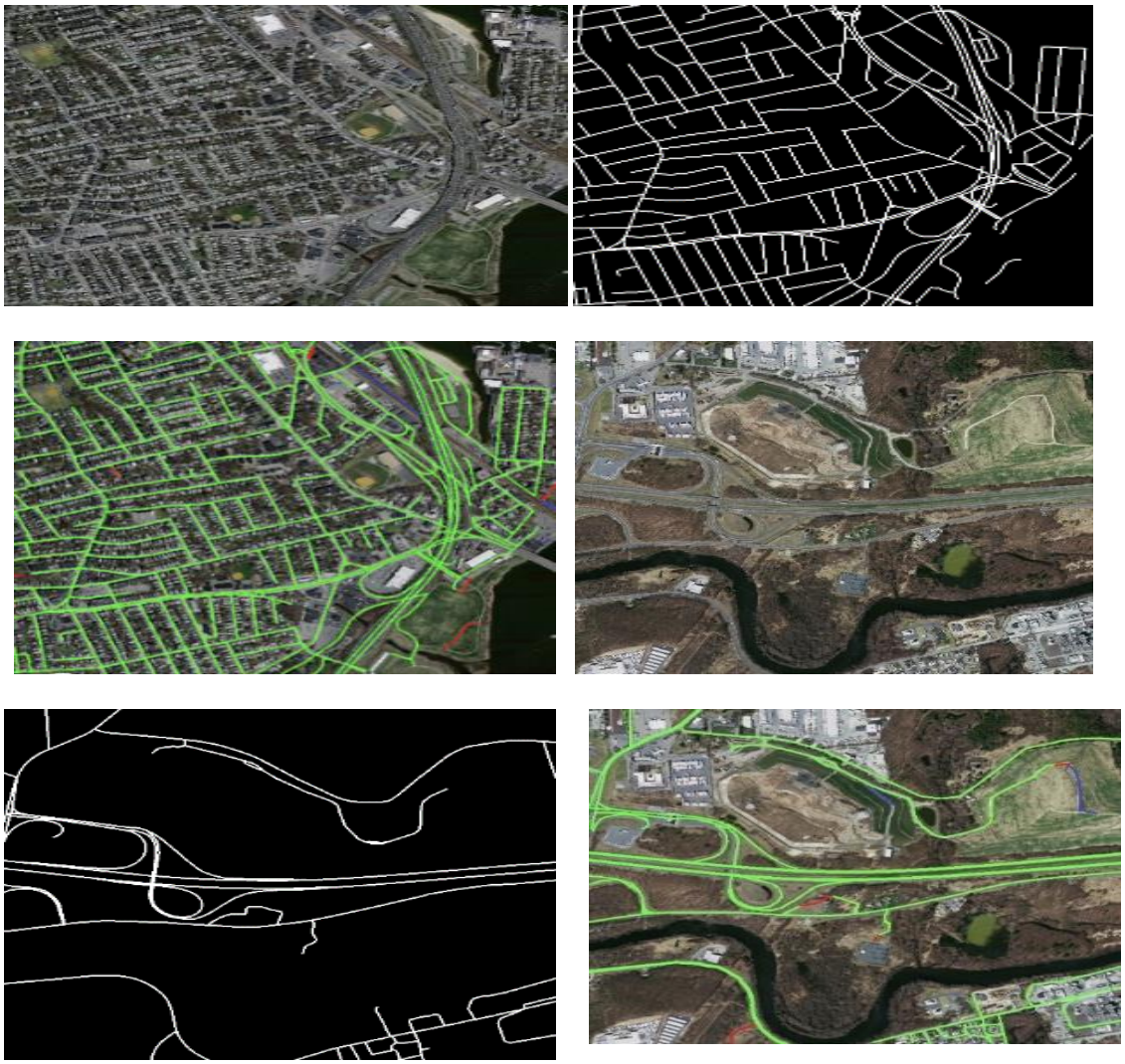


Figure 8: Road extraction results from the suggested technique obtained from the Massachusetts Road Dataset. The input images and the related ground truth images are shown in the first and second columns, respectively. The road extraction results are displayed in the third column. The colours green, blue, and red, respectively, stand for TPs, FPs, and FNs.

- **Building extraction:** For assessing the usefulness of the proposed technique for building object extraction, the qualitative results are depicted on the different images from datasets in Fig.9. The dataset contains all three bands (i.e., R, G, B) and rich information including roads, various buildings, vegetation etc. Fig.9 depicts the building objects extraction results from Massachusetts building dataset attained by suggested method.

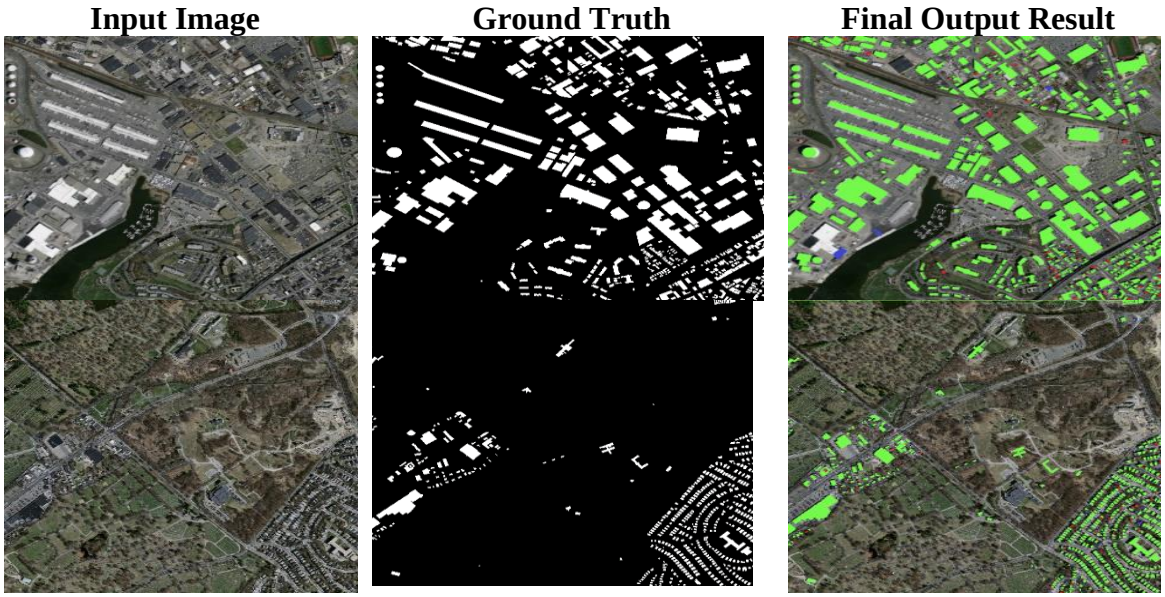


Figure 9: Building extraction results from the suggested technique obtained from the Massachusetts building dataset. The input images and the related ground truth images are shown in the first and second columns, respectively. The road extraction results are displayed in the third column. The colors green, blue, and red, respectively, stand for TPs, FPs, and FNs

2. **Quantitative evaluation:** For assessing the performance of the suggested object extraction technique, the following five evaluation metrics are used [32].

$$TPR = \frac{TP}{TP+FN} \quad (7)$$

$$TNR = \frac{TN}{TN+FP} \quad (8)$$

$$FNR = \frac{FN}{TP+FN} \quad (9)$$

$$FPR = \frac{FP}{FP+TN} \quad (10)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (11)$$

Where TP: True Positive; TN: True Negative; FP: False Positive, and FN: False Negative. The Classification Overall Accuracy (OA) is used to measure the rate of images that were correctly classified. The higher the value of an evaluation metric, the better the method's performance.

The confusion matrix and the Receiver Operator Characteristic (ROC) curve are used to assess the proposed method's performance. The confusion matrix and the ROCs is shown

in Fig.10. 186 images were correctly identified and 14 images were misclassified among total 200 images by this proposed method as shown in 11(a). The proposed framework resulted in OA of 93%. ROC graph depicted in Fig. 10(b) is a plot of TP rate vs. FP rate. The ROC graph of the suggested model for all data, as shown in Fig.10(b), indicates an exceptional classification between the two classes.

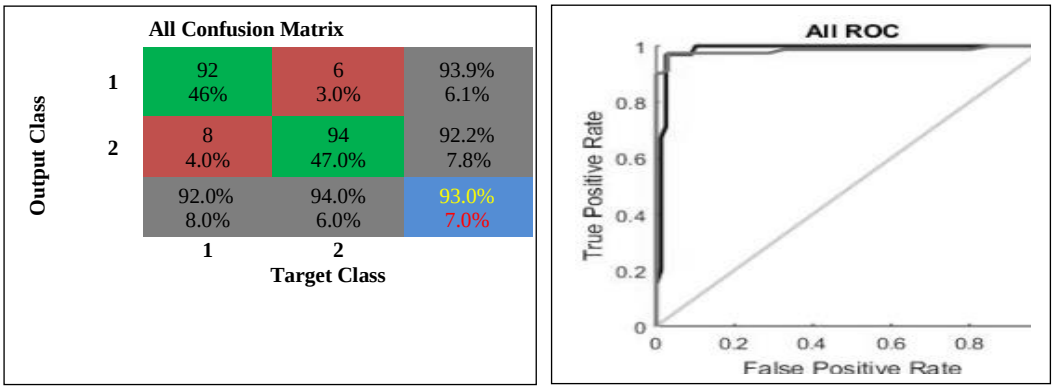


Figure 10: (a) Confusion Matrix and (B) ROC Plot of all Data

Table 2 lists the OA, TPR, TNR, FNR and FPR of the proposed method of training, testing, validation and all data, respectively. As seen from the tables, the proposed system incurs the acceptable level of performance with the mean values of no less than 93 and 92.55 and 94.1 for overall accuracy, TPR and TNR, respectively. The overall performance analysis is depicted graphically in Fig.11.

Table 2: Performance Analysis of Gabor Feature of Training, Testing, Validation and all Data

Proposed Gabor Energy Feature and PNN					
	Accuracy	**TPR**	**TNR**	**FNR**	**FPR**
Training (%)	94.3	92.2	95.4	7.8	4.6
Testing (%)	93.4	93.5	94.9	6.5	5.1
Validation (%)	92.5	92.3	93.8	7.7	6.2
All data (%)	93.0	92.0	94.0	8.0	6.0

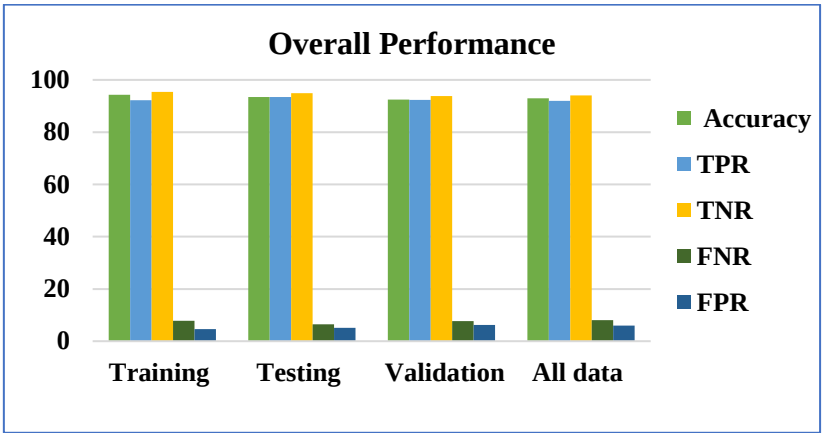
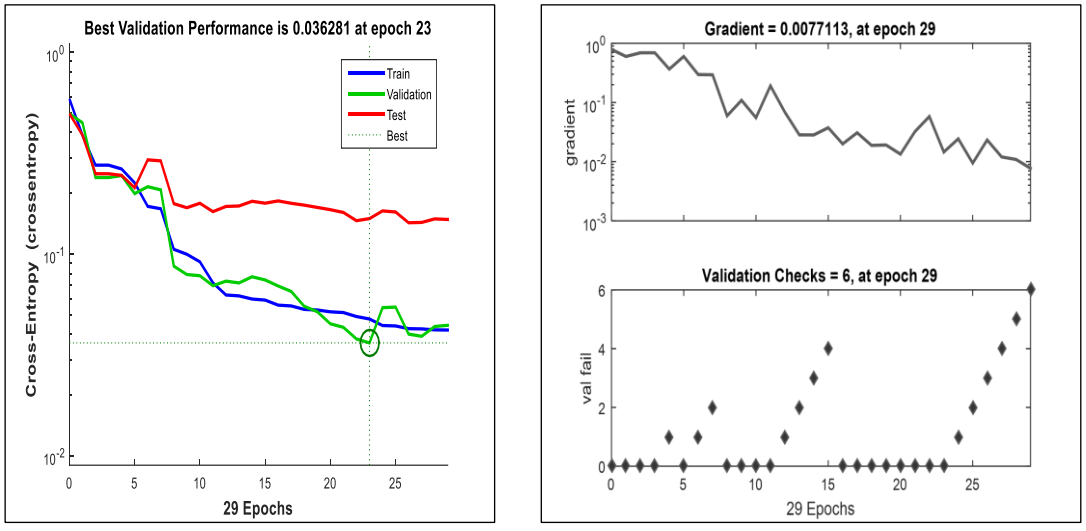


Figure 11: Overall Performance Analysis of Training, Testing, Validation and all Data Sets

Cross-entropy (CE) is a measure of the network's quality. Fig. 12 illustrates validation performance based on the cross-entropy error (a). The training was terminated at performance 0.0420 on iteration 29. As seen in the performance graph in Fig. 12, cross-entropy is reduced for effective categorization. The greatest validation performance before epoch 29 was 0.036 at epoch 23, which is inferior to the final 0.0420. The neural network training state's dynamics are shown in Fig.12(b) as a logarithmic cross-entropy gradient. For this collection of data, it is fair to terminate at the gradient at the terminus, which was 7.7113×10−3.



(a) **(b)**

Figure 12: a) Best Validation Performance b) Neural Network Training State

IV. CONCLUSION

This chapter suggests a method that, with minimal user involvement, enables the user to extract man-made items from input RS images, such as buildings and roads. We also use the numerical outcomes of actual RS photos to verify the proposed technique. By putting the

suggested method to the test on various photos, its universality is shown. The need for the current type of work in the fields of photogrammetry and computer vision is always rising, and its potential is limitless. The experimental findings demonstrate the usefulness and dependability of the suggested algorithm. The proposed approach is computationally significantly more efficient than state-of-the-art algorithms while attaining superior performance. Most notably, they have a far smaller number of input parameters, which makes them more practical in practice for not only object extraction, but for other applications such as urban monitoring, socioeconomic parameter analysis, etc.

Even if our findings are promising, the suggested strategy may yet be strengthened by combining deep characteristics with structural ones in next research. Deep learning techniques are increasingly being used to extract objects from RS data. The feature learning process will be sped up in next work using graphics processing units. Furthermore, due of its sensitivity to contrast, the suggested technique might not constantly retain the same categorization percentage. It would be promising to integrate texture features with other discriminative features e.g. spectral and radiometric characteristics, geometrical and contextual information to classify man-mad object and natural objects. Future improvements will also focus on improved object selection methods to lessen classification errors brought on by the similar spectral properties of buildings and other locations like roads and grounds.

REFERENCES

[1] J. Lin, W. Jing, H. Song, and G. Chen, “ESFNet: Efficient network for building extraction from high-resolution aerial images,” *IEEE Access*, vol. 7, pp. 54285-54294, 2019, doi3: 10.1109/access.2019.2912822.

[2] Y. Liu, Z. Li, B. Wei, X. Li, and B. Fu, “Seismic vulnerability assessment at urban scale using data mining and GIScience technology: Application to urumqi (China),” *Geomatics, Natural Hazards Risk*, vol. 10, no. 1, pp. 958-985, Jan. 2019.

[3] T. Panboonyuen, K. Jitkajornwanich, S. Lawawirojwong, P. Srestasathiern, and P. Vateekul, “Semantic segmentation on remotely sensed images using an enhanced global convolutional network with channel attention and domain specific transfer learning," *Remote Sens.*, vol. 11, no. 1, p. 83, 2019.

[4] Abdollahi, A., Pradhan, B., Gite, S., & Alamri, A., “Building footprint extraction from high resolution aerial images using generative adversarial network (GAN) architecture,” *IEEE Access*, Article 209517-209527,2020. http://dx.doi.org/10.1109/ACCESS.2020.3038225.

[5] S. E. Park, Y. Yamaguchi, and D. J. Kim, “Polarimetric SAR remote sensing of the 2011 Tohoku earthquake using ALOS/PALSAR,” *Remote. Sens. Environ.*, vol. 132**,** pp. 212-220, 2013.

[6] X. H. Tong, Z. H. Hong, S. J. Liu, X. Zhang, H. Xie, Z. Y. Li, S. L. Yang, W. A. Wang, and F. Bao, Building-damage detection using pre- and post-seismic high-resolution satellite stereo imagery: A case study of the May 2008 Wenchuan earthquake, *ISPRS J. Photogramm. Remote Sens.*, vol. 68, pp.13-27, 2012.

[7] Wenzao Shi, Z. Mao & Jinqing Liu, “Building area extraction from the high spatial resolution remote sensing imagery,” Earth Science informatics, (2018) https://doi.org/10.1007/s12145-018-0355-5.

[8] K. Bittner, F. Adam, S. Cui, M. K¨orner, and P. Reinartz, “Building footprint extraction from VHR remote sensing images combined with normalized DSMs using fused fully convolutional networks,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 11, no. 8, pp. 2615–2629, 2018.

[9] Q. Zhu, C. Liao, H. Hu, X. Mei, and H. Li, "MAP-Net: Multiple Attending Path Neural Network for Building Footprint Extraction from Remote Sensed Imagery," IEEE Trans. Geosci. Remote Sens., vol. 59, no. 7, pp. 6169-6181, July 2020.

[10] Wenjin Wu, Huadong Guo, and Xinwu Li, "Urban Area SAR Image Man-Made Target Extraction Based on the Product Model and the Time–Frequency Analysis," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 8, no. 3, pp. 943-952, 2015.

[11] Z. Li, W. Shi, Q. Wang, and Z. Miao, "Extracting man-made objects from high spatial resolution remote sensing images via fast level set evolutions," IEEE Transactions on Geoscience and Remote Sensing, vol. 53, no. 2, pp. 883–899, 2014.

[12] Abdul Alim Sheikh, S. Mukhopadhyay, Noise Tolerant Classification of Aerial Images into Manmade Structures and Natural- Scene Images based on Statistical Dispersion Measures, 2012 Annual IEEE Conference (INDICON), 2012, pp. 653-658 DOI: 10.1109/INDCON.2012.6420699

[13] Md. Abdul Alim Sheikh, Tanmoy Maity, Alok Kole "IRU-Net: An Efficient End-to-End Network for Automatic Building Extraction from Remote Sensing Images," IEEE Access, vol. 10, pp. 37811-37828,2022. DOI:10.1109/ACCESS.2022.3164401

[14] M. A. A. Sheikh, A novel self-assessed approach for classification of manmade objects and natural scene images from aerial images, 2011 Annual IEEE India Conference, 2011, pp. 1-7, doi: 10.1109/INDCON.2011.6139328.

