POVERTY PREDICTION FROM SATELLITE IMAGERY WITH IMAGE PROCESSING USING CONVOLUTIONAL NEURAL NETWORK

T.Suryam[1], K.Kavya[2], K.Anusha Pavani[3], K.Devi Sri[4], M.Yuva Bhargavi[5], M.Muskhan[6]

[1]Assistant Professor Department of Computer Science & Engineering, Vignan's Institute of Engineering for Womens, Duvvada, Visakhapatnam

[2,3,4,5,6]Department of Computer Science & Engineering, Vignan's Institute of Engineering for Womens, Duvvada, Visakhapatnam

ABSTRACT

The first of the United Nations' 17 Sustainable Development Goals is the eradication of poverty in all of its manifestations globally by the year 2030. National household surveys are used to gather information about poverty in areas where the majority of the world's poorest people reside. Obtaining accurate and timely information is challenging because running such surveys requires a lot of resources. A workable answer to the data scarcity issue is now possible thanks to advancements in computer vision and public access to abundant data sources like daytime satellite images and nocturnal lights. In order to forecast the distribution of poverty at the village level in nations using more advanced technologies and up-to-date data, this project will build on prior research by using machine learning techniques to process daytime satellite images and nighttime lights. With the availability of aerial satellite data, new and exciting uses, like the per-pixel classification of specific objects, are now possible. This study demonstrates how a convolutional neural network (CNN) can be used to fully, quickly, and accurately classify each pixel of satellite imagery of a small city. The high-level segmentation is then enhanced using the estimated low-levelpixelclasses.TheCNNarchitecture'svariousdesigndecisionsareanalysedandassessed.Theinvestigatedlandareaiscompletelymanuallyclassifiedintothe following five categories: vegetation, ground, roads, buildings, and water. The classification accuracy is then compared to other per-pixel classification studies on other land areas that have a comparable selection of categories. The outcomes of the comprehensive classification and segmentation on a few map segments demonstrate the viability of CNNs for solving the segmentation and object identification problems for remote sensing data. Based on the extracted features and luminosity, image is classified into three categories: low, bright, high and corresponding wealth index will be predicted.

Keywords: aerial satellite data; convolutional neural networks; per-pixel classification; Prediction.

INTRODUCTION

Lack of adequate resources to meet one's basic needs is referred to as poverty. Finding out how inequality and poverty are distributed spatially in a particular region is known as poverty mapping. In order to execute and evaluate programmes to end poverty, measuring poverty is essential. Statistics on poverty are typically gathered from nationwide household surveys. As a result of the high expenses associated with conducting such surveys, there are not enough

statistics on poverty in developing nations. A major obstacle to reducing poverty is this data divide. Additionally, COVID-19 is probably responsible for the first rise in world poverty since 1998 [1]. According to predictions, Sub-Saharan Africa will experience the greatest increase in extreme poverty as a result of COVID-19. Additionally, particularly in the developing world, a consistent lack of data on important economic indicators hinders progress towards that objective. Accurate poverty measurements are beneficial for two primary reasons:

Accurate measurements of poverty level allows philanthropic agencies and governments to identify where resources and interventions are needed, and helps guide the direction of financial aid.

Organizations can better monitor progress towards the Sustainable Development Goals with the help of regular and trustworthy statistics on the distribution and levels of poverty.

For many application areas, such as land inventory and vegetation monitoring, aerial and satellite imagery collection is crucial. The methods for collecting aerial images have been extensively studied. Due to decreased data collection expenses and improved technology, data collection has advanced quickly in recent years.But unless it is processed and useful data is recovered, imagery is meaningless. Recently, a number of methods for automating information retrieval from satellite imagery have been developed, with various application areas being the focus. Segmentation and feature extraction are frequently used in existing methods for completing the object recognition job The inapplicability of many developed algorithms to other domains is one of their primary problems. For one particular problem, and frequently for one specific area with well-known seasonal imagery changes, many methods for image segmentation and classification are ideal. It should be noted that image segmentation is frequently not of interest and that image data can be cropped into regions with fixed dimensions due to the advancement in acquisition technology stated above and excessive resolution. Using deep learning algorithms, such as a convolutional neural network (CNN), for picture classification is one potential solution to the issue of algorithm generalisation. The main advantage of CNN-based algorithms is that they don't need to collect features prior , which leads to better classification abilities. CNNs have recently been demonstrated to be effective in a variety of tasks, including object identification object detection , scene parsing , and scene classification In this study, we look into the per-pixel classification of very high resolution (VHR) satellite images using a CNN. One of the many deep learning designs is a convolutional neural network (CNN). A series of feed-forward layers in the widely used CNN are used in recent years to handle a variety of complicated problems, including image recognition and classification. Similar to a conventional neural network, a CNN is composed of neurons with biases and weights that can be learned. It is feasible to think of the neurons in a feed-forward artificial neural network as receiving a set of inputs and conducting some non-linear processing . Images are used as inputs in convolutional network designs, allowing the encoding of specific properties into the architecture. Convolutional, pooling, and full connection layers are among the layers that make up the standard CNN structure. It is possible to say that this particular neural network is a special case of the neural network because it consists of fully connected layers that mathematically sum up a weight, pooling/ subsampling layers that make the features robust against distortion and noise, and one or more convolutional layers that extract low-level features like lines, edges, and corners. In this paper, convolutional neural networks (CNNs) with Keras Sequential API were used as a pre-trained model for features extraction and trained on the image Net dataset.

LITERATURE SURVEY

A literature review is a report that evaluates the data from works of literature that are relevant to the field of study you have chosen. The evaluation is structured in such a way that it ought to provide a theoretical framework for the research and assist you in identifying the focus of your own investigation. A literature review is a crucial step in the research process because it summarises all prior studies on the relevant issue statements and establishes the framework for the current study. Without first analysing the prior research that has been conducted on the subject, no new research can be accepted seriously. Based on the provided problem definition, the current study work is a review of the literature on satellite image processing methodologies and approaches. It gives the analyst information on different satellite image categorization techniques and describes them in detail. The automatic classification of satellite images and the processing methods and approaches that use algorithms are highlighted in the current literature study. Our research focuses on categorising satellite images into low, bright, and high luminosity and determining corresponding wealth indices, regardless of the colour of the image used as input. In our search, we discovered several papers that provided details on various approaches that are already being used for this purpose. In our study, we discovered that dividing an image into several bands and using them to categorise things into different classes is the best method for doing so.

Numerous applications benefit from the interpretation of these satellite images, including environmental conservation and management, water resource research, soil quality studies, environmental study after natural disasters, meteorology simulations, obtaining information on land use and land cover, preventing natural disasters, and researching the evolution of climatic change. The methods used to retrieve data from remote sensing photos vary. The most effective method for understanding images and extracting information is the classification technique [2]. Satellite image classification organises the image's pixels into a number of predefined classes [3]. Based on the digital values derived from the satellite photos, the pixels are grouped. Satellite image classification organises the image's pixels into a number of predefined classes [3]. Based on the digital values derived from the satellite photos, the pixels are grouped. In the case of grayscale photographs, the extracted pixel values may be a single value, but for multispectral, multimodal, or multi temporal images, they may be a multivariate value. The classification helps in the extraction of the data from the various bands [4] of the satellite sensor.

The data is recovered in terms of digital numbers, which are then translated to a category. Using the decision tree technique, S. Muhammad et al. [5] suggested a supervised satellite image categorization system. This technique uses pixel colour and intensity to extract features from satellite photos. The objects in the satellite photos can be identified with the use of extracted features. The technique uses a decision tree and identified items to classify satellite photos. Extremely high-resolution satellite image classification using fuzzy rule-based systems was introduced by J. Shabnam et al. in their paper "Very High- Resolution Satellite Image Classification Using Fuzzy Rule-Based Systems" [6] using supervised satellite image classification techniques. With this technique, satellite photos are divided into five main categories: bare ground, road, building, plant, and shadow. At the edges of objects, fuzzy approaches are utilised to increase categorization accuracy. Bjorn Frohlich's "Land Cover Classification of Satellite Imagery Using Contextual Knowledge" [7] explains a technique for categorising satellite photos into several types of land cover. The segment-level categorization used in this automated method is supported by a training set. The contextual aspects of many preset classes are used into the classification methods to increase classification accuracy. Selim put forth a classification strategy utilising the Bayesian technique in the work "Spatial technique for Picture Classification" [8]. The technique classifies high-resolution satellite photos using spatial information. Two phases make up the method's classification process. In phase 1, discrete nonparametric density models are used to train Bayesian classifiers. Each pixel's spectral and textural properties are retrieved. Phase 2: The pixel-level categorization maps are transformed into contiguous regions using the iterative split-and-merge procedure. The most popular unsupervised satellite classification methodology is described in "Comparison of Four Classification Techniques to Extract Land Use and Land Cover from Raw Satellite Imagery for Certain Far Arid Regions, Kingdom of Saudi Arabia" [9]on ISODATA. On a satellite image, it generates a predetermined number of unlabeled clusters or classes. Subsequently, the clusters are given meaningful labels. The number of clusters and iterations that must be executed can be controlled by a number of factors in ISODATA. Clusters typically include pixels from various classes. A non-parametric, unsupervised statistical classification method is the Support Vector Machine (SVM) [10]. Maps of land use can be extracted using this technique. SVM operates under the presumption that there is no knowledge of how the whole amount of data should be distributed. SVMenhances accuracy, speed, and cost of satellite categorization. A statistically supervised way for finding patterns is the maximum method. Based on the probability values of the pixels, it assigns pixels to the proper classes [11]. The satellite image's pixels can be classified effectively using the maximum likelihood method. Yet it takes time, and using little ground truth data leads to subpar outcomes.

PROPOSED WORK

In this study, we use a CNN sequential model, which is suitable for a simple stack of layers with precisely one input tensor and one output tensor for each layer. We calculate the loss function, which generates a category index of the most likely matching category, using sparse categorical cross entropy. Indeed, it works with numbers, but these integers must be class indices rather than actual values. Just output indexes that the ground truth indicates are used to construct the logarithm in this loss. Since other values are instantly multiplied by zero in the categorical cross entropy's standard version, this has no effect on the result

Figure1: **a)** Picture with high luminosity **b)** Picture with medium luminosity **c)** Picture with low luminosity.

This results in higher efficiency because it only computes the logarithm once per instance and skips the summing.

We update network weights iteratively based on training data using the adam optimizer, an optimization algorithm that can be employed in place of the conventional stochastic gradient descent method.

1. Data Acquistion and Cleaning:

Economic Varible:

Wealth Index: For the socioeconomic indicators, we used the 2017 Demographic and Health Survey (DHS) in Burundi as a gauge of reality. The DHS dataset's "Wealth Index" or "Wealth Index Factor Score" was the subject of our attention. The "Wealth Index", which is calculated as the first principal component of qualities relating to common asset ownership on a per-household level, is the main indicator of socioeconomic well-being. With the help of the sklearn Min Max Scaler module, we converted the "Wealth Index" to the range [0,1]. The median wealth index value for each cluster, as provided in the DHS data set, was then determined.

Access to Electricity: The number of yes responses to the survey question on access to electricity is disclosed in the DHS dataset. This figure was calculated as the sum of the values for each cluster's households.

Access to Water: The total travel time, measured in minutes, to reach the water source is included in the DHS dataset. Time is set to zero if there is an on-site water supply. We calculate the average amount of time taken by each cluster's households to reach a water source.

Education: The DHS collects data on the total number of school years that household members older than six have completed. This figure was calculated as the sum of the values for each cluster's households.

Nighttime Luminosity :

The National Centre for Environmental Information's 2020 satellite imagery is used to provide the nighttime luminosity (or light intensity) statistics. The information comprises a continuous brightness level for Burundi that ranges from 0 to 63, with 0 representing the darkest pixel.We calculate the median wealth index for each cluster using DHS survey data. The nighttime satellite image has a resolution of around 1 km per pixel, and we used 10 pixels by 10 pixels to average the luminance of each cluster. The mean nighttime illumination and wealth index were then combined at the cluster level. For the regression model to wealth index from luminosity, we employed power transform on the wealth index prior to the modeling.

Distribution of Wealth Index:

The distribution of wealth index is the next thing we examine. We compare its value with zero luminosity and with luminosity greater than zero (Figure 1.6.4). We can see that the wealth indices are almost 0 when the places are fully black at night (luminosity is zero). The value of the wealth index rises when brightness rises over zero.

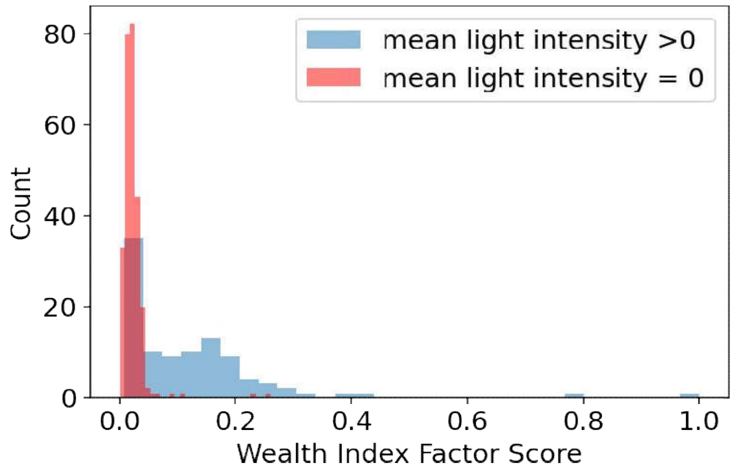

Figure2: Distribution of wealth Index of Burundi

3. Regression Model to predict Wealth Index from Night time Imagery

Previous research has used machine learning techniques to forecast wealth using night time satellite data in both non-African and sub-Saharan African nations. According to the countries and models, the predictive models published in the literature often produced r-squared values between 0.51 and 0.75.In this study, we investigated various regression models to estimate Burundi's wealth index using the illumination of the night sky. Despite the fact that we performed a power conversion on the wealth index before modelling, the regression models do not have a good fit in terms of r-squared, with the best result coming from the Random Forest Regressor (0.54). The poverty map at the cluster level was then recreated using the best anticipated wealth index.

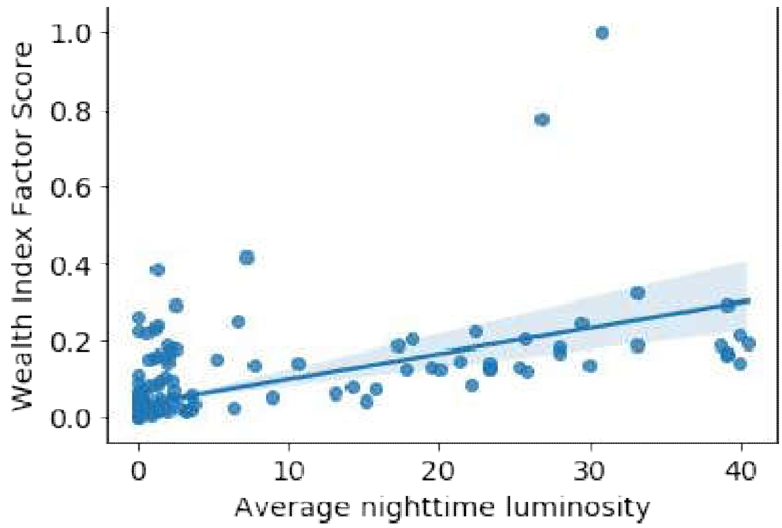

Figure3: Wealth Index vs Night time luminosity

ARCHITECTURE OVERVIEW

CNN is made up of several processing layers. Convolution filter families that recognise picture features make up each layer. The CNN eventually creates a set of predicted probabilities, one for each class, by combining the detector outputs in fully linked "dense" layers near the conclusion of the series. When it trains, the network itself learns which features to detect and how to do so. Convolutional neural networks (CNNs or ConvNets) [13] are built to handle natural signals that appear as numerous arrays. They have had success with both 2D and 3D structured multiple array tasks, including video and challenges involving object detection in pictures [14] and speech recognition from audio spectrograms . By utilising local connections and tied weights to mimic the characteristics of real signals, CNNs are easier to train since they have fewer parameters than a fully-connected network. CNNs also have the advantage of learning slightly translational- and rotationally-invariant features due to the use of pooling, which is a desired quality for natural signals. By stacking many CNNs, a deep CNN can be produced in which low-level features (such as edges and corners) are integrated to form mid-level features with a greater level of abstraction (objects). The final CNN layer is often coupled to one or more fully-connected layers (FC-layers).

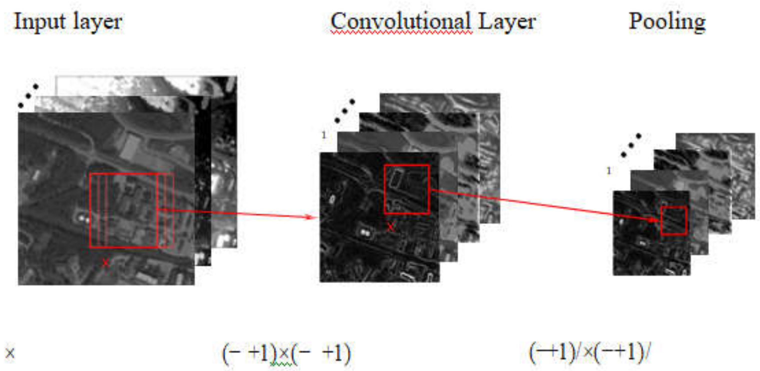

Figure.4. The three layers that make upon eCNN layer: input layer, convolutional layer and pooling layer. The input layer has *c* color channels, and the convolutional and pooling layer has *k*
Feature maps, where *k* is the number of filters. The size of the input image is $m \times m$ and is further decreased in the convolutional and pooling layer with the filter size *n* and pooling dimension *p*.

Per-Pixel Classification Using a Single CNN

Figure 5 illustrates the CNN classification of a single pixel in a satellite image. The first CNN layer's input consists of c spectral bands with a contextual size of n*m, the centre of which is the pixel that needs to be categorised. The fully-connected (FC) layer is the second layer after the softmax classifier in the full architecture, which is made up of L numbers of stacked standard CNN layers .One feature and pooling map for each filter makes up the convolutional and pooling layer of the first CNN. After the convolutional step, (x) = max, rectified linear units (ReLU) are utilised as the non-linear activation function (0, x).
After the non-linear activation function step, the non-saturated output brought on by the ReLU activation function is normalised using local contrast normalisation (LCN). choosing the contextual region surrounding the pixel to be classified, the quantity of CNN layers, the number of filters, the size of the filters, and the pooling dimension for each CNN layer. The pooling layer of the previous stacked CNN is used as input for a fully connected layer that is attached.

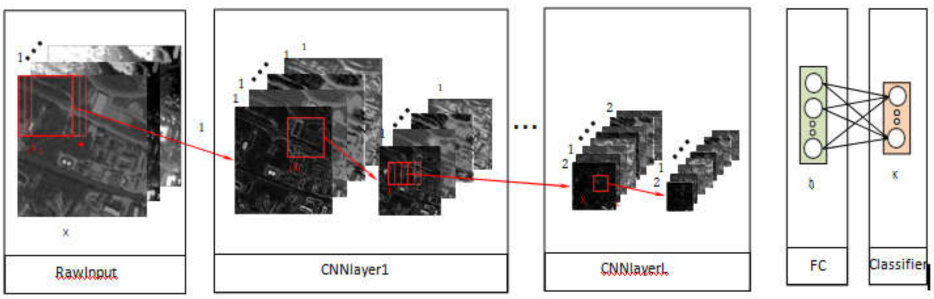

Figure5. Overview of the method used for per-pixel classification. The $k1$ filters of size $n1 \times n1$ from the first CNN layer are convolved over the pixel to be classified and its contextual area of size $m \times m$ with *c* color channels to create $k1$ feature maps. The feature maps are pooled over an area of $p1 \times p1$ to create the pooling layer. The process is repeated for *L* number of CNN layers. The pooling layer of the last CNN is the input to

a fully-connected auto-encoder. The hidden layer of the auto-encoder is the input to the soft max classifier.

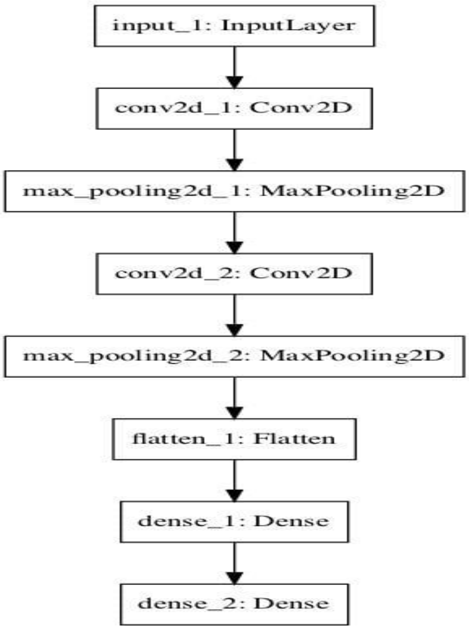

Figure.6.. Keras Sequential Model

RESULTS AND DISCUSSIONS

Analysis of Burundi Country:

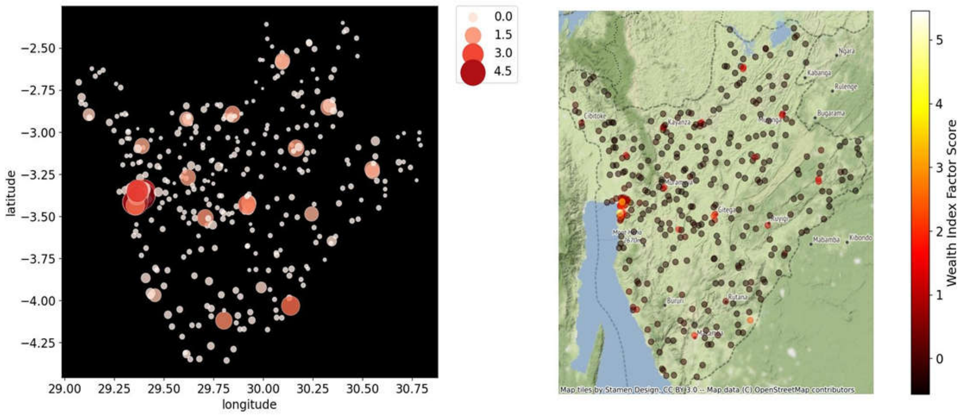

Figure 7: Burundi Wealth Distribution graphs

WEALTH PREDICTION

Figure.8. Wealth Prediction

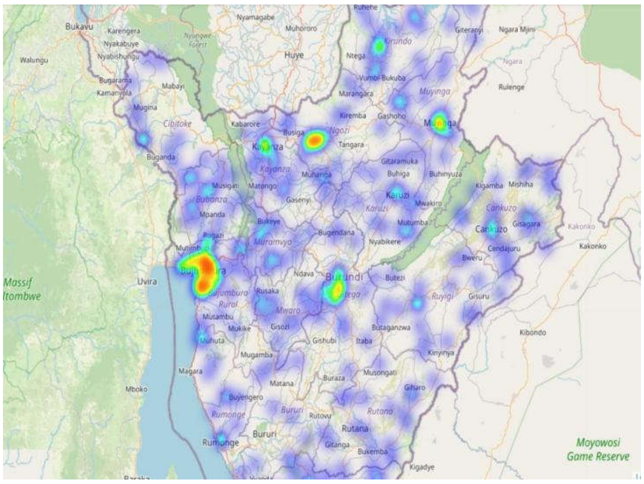

Figure 9: Heat Map of Burundi based on Wealth Index

LIMITATIONS AND FUTURE WORK

Although this research represents a significant improvement in the field of computer vision and satellite imagery-based poverty prediction, there are a few drawbacks that require discussion. The data is first and foremost inadequate by its very nature. This paper can still be improved by adding GPS access and a map that, when clicked on, retrieves information about a specific area's wealth index and luminosity category. The closest daytime and nighttime satellite image to the wealth index's locations was selected, though this location wasn't always the most advantageous. In practise, though, this might not be a huge deal because nearby cities typically have similar wealth indices.

CONCLUSION

This study looked at a new method of predicting poverty that used satellite photos taken during the day and at night. Even if there is more work to be done, we believe our work can provide the basis and demonstrate the ability of poverty prediction. The first step in reducing poverty is to identify its regions, and we think this work has helped with that. This paper can still be improved by adding GPS access and a map that, when clicked on, retrieves information about a specific area's wealth index and luminosity category.

REFERENCES

1. Bhabatosh Chanda, Dwijesh Dutta Majumder, Digital Image processing and analysis.
2. W.G. Rees, Physical Principles of Remote Sensing, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press.
3. Muhammad, S., Aziz, G., Aneela, N. and Muhammad, S. 2012. "Classification by Object Recognition in Satellite Images by using Data Mining". In Proc. Proceedings of the Congress on Engineering (WCE 2012), Vol I, July 4 - 6, London, U.K.
4. Shabnam Jabari and Yun Zhang. "Very High-Resolution Satellite Image Classification Using Fuzzy Rule-Based Systems", Algorithms, vol.6, no.4, pp. 762- 781.
5. Bjorn Frohlich., Eric Bach., Irene Walde., Soren Hese.,Christiane Schmullius, and Joachim Denzler. 2013. "Land Cover Classification of Satellite Images using Contextual Information", ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume II-3/W1, pp. 1-6.
6. Selim Aksoy. "Spatial Techniques for Image Classification," in C. H. Chen, ed., Signal and Image Processing for Remote Sensing, CRC Press, pp.491- 513.
7. Al-Ahmadi, F, S. and Hames, A, S. 2009. "Comparison of Four Classification Methods to Extract Land Use and Land Cover from Raw Satellite Images for Some Remote Arid Areas, Kingdom of Saudi

Arabia", Journal of King Abdulaziz University-Earth Sciences, Vol. 20, No.1, pp: 167-191
8. Jensen, J, R. 2005. "Introductory Digital Image Processing: A Remote Sensing Perspective", 3rd Edition, Upper Saddle River: Prentice-Hall, 526 p.
9. N. L. Tun, A. Gavrilov, and N. M. Tun, "Multi-classification of satellite imagery using fully convolutional neural network," Proc.- 2020 Int.Conf. Ind.
10. Eng. Appl. Manuf. ICIEAM 2020, pp.7–11, 2020, doi:10.1109/ICIEAM48468.2020.9111928
11. sVan Etten, "Satellite imagery multiscalerapid detection with windowed networks, "Proc. - 2019 IEEE Winter Conf. Appl. Comput.Vision, WACV 2019, pp. 735–743, 2019.
12. Van Etten, "You Only Look Twice: Rapid Multi-Scale Object Detection in Satellite Imagery," 2018.
13. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. IEEE Proc. 2278–2324.
14. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. In Advances in Neural InformationProcessing Systems; Curran Associates: North Miami Beach, FL, USA pp. 1097–1105.
15. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. IEEE