ISSN 2395-1621

Predict Poverty-Ratio Status of an Area from Satellite Images Using Machine Learning

Mohite Rutuja, Shaikh Rejawana, Bhavsar Simran, Kondhalkar Dipali, Prof. B.B.Bhusare

rutujamohite00@gmail.com,
rejawanashaikh24@gmail.com,
simranbhav0004@gmail.com,
kondhalkardipali000@gmail.com,
bhusare85bapu@gmail.com

Department of Computer Engineering
Navsahyadri Education Society's Group of Institutions Pune.

ABSTRACT

Estimating Socio economic and developmental parameters like poverty levels of rural and urban areas from satellite images is challenging problem. Also obtaining efficient data about local livelihoods in developing countries is expensive and such data are rare. Traditionally that data come from household surveys. It is possi-ble to predict poverty in both regions with advance technology. Here system which we are designing using machine learning supervised method has the ability to identify some major factors they are Agriculture, Water resources, Building, Roads and Roof. Roof top is a very essential factor for our system for prediction of socio-economic status in rural areas. Our system takes satellite image as a input which is pre-trained on large dataset of satellite images. and then this satellite images is compared with our trained model which contains all these major factors present within it and after comparing these factors we get prediction of the status of that satellite image in the form of percentage and by considering this percentages we are predicting poverty in rural and ur-ban areas.

Keywords: Convolutional Neural network, High Resolution Satellite Images, Economic Status prediction, Computational Sustainability.

ARTICLE INFO

Article History

Received: 14th May 2020

Received in revised form : 14th May 2020

Accepted: 16th May 2020

Published online : 16th May 2020

I. INTRODUCTION

There are so many regions in the world where humans are exist but they have no facilities for their livelihood. They don't even have basic necessity of life like water, food and so on. Some region has lack of only one factor and some regions have lack of all the factors. Some region has water but not electricity while another region has home but not any other necessities. For such type of regions, some organizations are ready to help them with the support of government of that country but due to lack of communication from that region, the organization knows only the location of that region. They don't even know what the basic ne-cessities of that region are?

In that case, the organization can only have the sat-ellite image of the region and they try to determine necessities by observing satellite image. But by only observing that region through satellite image we cannot estimate the presence of the factors on that region. So to solve this kind of problem we are introducing an application to predict socioeconomic status of a region.

To overcome this problem, we present the system which has the high resolution satellite imagery. In this paper we train Convolutional Neural Networks (CNNs) to estimate poverty directly from high and medium resolution satellite images. These images compared with the train model and identify some major factors like Agriculture, Water resources, Buildings, Roads and Roof top. After that we get the predictions of those images in the form of percentages and based on the final results we identify the rural and urban areas.

We conclude that CNNs can be trained end-to-end on satellite imagery to estimate poverty and gives the better prediction results. The proposed framework utilizes segmentation for extracting the feature and CNN for analyzing poverty areas.

In this paper, we describe our methods to investigate how the supervised machine learning algorithms dis-tinguish the spatial characteristics of different satellite images with the help of CNN. The organization of this document is as follows. In Section 2 (Literature Sur-vey), we enlisted details of all research existed. In section 3 present the data require to the system and perform pre-processing. In Section 4 (System Architecture), present architecture of proposed

system. In section 5 (Result and Discussion) present experimental results compare with the previous prediction models .In section 6 (Conclusion) discussed conclusion and future work required to improve our system.

II. LITERATURE SURVEY

In 2017, Anthony Perez and Christopher Yeh have done the project on Poverty Prediction with Public Landsat 7 Satellite Imagery and Machine Learning[1]. Based on this paper they presented results show that the current state-of-the-art in satellite-based poverty prediction lends itself to predicting relative wealth within a single country where some ground truth data is available, but may struggle with extrapolating across country borders. Using some combination of nightlights and predictions from the proposed models may yield further improvements.

In 2017, Boris Babenko, Jonathan Hersh, David New-house, Anusha Ramakrishnan, Tom Swartz have done the project on Poverty Mapping Using Convolutional Neural Networks Trained on High and Medium Reso-lution Satellite Images, With an Application in Mex-ico[3]. In this paper Presents the CNN predictions for urban areas using imagery for either Digital Globe or Planet, using the 10% withheld validation sample. R2 estimates that show the correlation between predicted poverty and benchmark poverty as measured in the 2015 Intercensus. R2 is estimated at 0.6using the Digi-tal Globe imagery, and 0.54 using Planet imagery. The drop in performance is modest but not server,

Especially considering that Planet imagery offers daily revisit rates of the earth's landmass.In 2018, Shailesh M. Pandey, Tushar Agarwal, Na-rayanan C Krishnan have done the project on "Multi-Task Deep Learning for Predicting Poverty from Satellite Images". Based on this paper it gives two step for predicting poverty in rural regions of India from satellite imagery. First, train a multi-task fully convolutional model to predict three developmental parameters roof top, source of lighting and source of drinking water from satellite imagery. Using only satellite imagery as input, they are able to estimate income and poverty close to the true values collected on the ground level.

In 2018, Barak Oshri, Annie Hu, Peter Adelson, Xiao Chen, Pascaline Dupas, Jeremy Weinstein, Marshall Burke, David Lobell, Stefano Ermon have done the project on Infrastructure Quality Assessment in Africa using Satellite Imagery and Deep Learning[4]. Based on this paper Data on infrastructure quality outcomes in developing countries is lacking, and this work explored the use of globally available remote sensing data for predicting such outcomes. Using Afro barometer survey data, it introduced a deep learning approach that demonstrates good predictive ability. The quality of a deep model heavily relies on adequate data available, and a large focus should be towards making better use of existing image and survey data, through strong cataloging and collating efforts. However, our results demonstrate the proof of concept that satellite imagery can be used to predict infrastructure quality.

III. DATA AND PRE-PROCESSING

The analysis for this project was performed on imagery supplied by satellite imagery. We use satellite imagery provided by the map puzzle software map which contain high resolution satellite images and it will save the images in jpg, gif, png, bmp, tiff format which is shown in figure 1..map puzzle software is free and It works on Windows 7, Windows XP, Windows Vista.

This imagery include Three band of Red, Green, Blue(RGB) values we experiment this with training model to include additional information[3]. This data-set consist only RGB images. Most existing CNN models are designed to work with 3-channel RGB images and thus are not directly compatible with mul-ti-band satellite images. Thus, we adapted several ex-isting architectures to work on multi-band satellite images: 18- and 34-layer ResNets [5] and VGG-F/ [6]. We trained each model using all bands and using only the RGB bands. When using only the RGB bands, we initialized the CNNs with weights pre-trained on the ImageNet dataset [7].

Fig A: Original image

Fig B: Grayscale Image Fig C: Bilateral Image Fig D: Canny Edge Detection

IV. PROPOSED SYSTEM

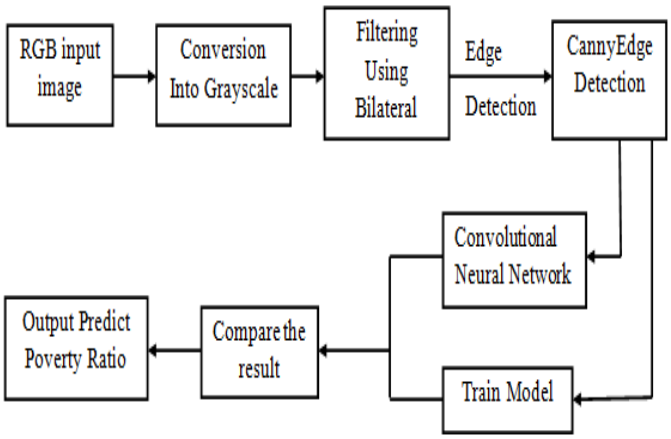

Fig1: System Architecture

Step 1: Basic Image processing:

1.1 Gray scale Conversion:

Our system takes RGB satellite images as a input then convert these satellite images into grayscale form.

1.2 Bilateral Filtering:

After grayscale conversion filtering is done using bilateral for the purpose of noise removal. e.g. If we download the image through map puzzle software during downloading process of image if some point of images meets more speed then we get blur image so we cant get accurate clear image so to remove those noise here bilateral filtering is done.

1.3 Smoothing:

Smoothing is done for better accuracy.

Step 2: Canny edge Detection:

In this step for detecting edges of images like agriculture, water resources, roads based of that edges we identify that these image is of these type so for these purpose we use canny edge detection algorithm.

Step 3: CNN:

After detecting edges CNN algorithm is perform for feature extraction. CNN extract the features of satellite image automatically.

The elements enter into CNN operation:

Step 1: Input layer:

Input layer in CNN contain the image data. It passes the data to the first hidden layer.

Step 2: Convolutional layer:

Convolutional layer is layer of extract the feature from input image. Convolution preserves the relationship between pixel by learning image feature using the input data.

Step 3: RELU layer:

RELU stand for rectifier linear unit. It is the most generally utilized activation function and it is defined as max(0,s). It gives the positive output s if S>0.

Step 4: Pooling Layer:

The pooling operation involve sliding a two dimensional filter over each Channel of feature. feature map having dimensions nh x nw x nc, the dimensions of output obtained after a pooling layer is

$$(nh - f + 1) / s \times (nw - f + 1)/s \times nc$$

Step 5: Fully Connected layer:

In this part, everything that we trained throughout the section will be merged together. It performs the classification based on feature extracted by the previous layer.

Step 6: Soft-max layer:

It is optimization Functions for CNN model. to calculate final accuracy and losses.

Step 4: Compare the result:

In this step compare the results with train model then predict the poverty ratio in the form of percentage and determine the socio economic status.

V. EXPERIMENTAL RESULT AND COMPARISON

A. System Setup:

The whole architecture is made by PyQT4 library used in python language. PyQT library gives all necessary stuff related to GUI design. Another task is to pre-process the input image which can be done by OpenCV library in python. By using this library, image is converted into grayscale image, contour image and smoothen image and run in environment with System consist Processor min I3, Ubuntu 32 bit machine with 4GB of RAM.

B. Results:

In the Proposed system, we using supervised CNN Which improves the accuracy of the prediction.CNN is proved for better accuracies with supporting to the complemented with light weight library in python for image processing as OpenCV which help us to classify the image and improves the speed of execution .System has use various parameter likes Agriculture, Water Road and Building to calculate economical area with better accuracy.

Comparative results of existing and proposed system is as follows.

Parameter	Existing System	Proposed System
AI based approach	No	Yes
Use of satellite image	No	Yes
Use of open CV	No	Yes
CNN	No	Yes
Improved speed	No	Yes
Light Weight	No	Yes

Table 1:Comparative Result

Figure A: show the accuracy of that predictions based on the training and testing samples. Figure B: show the loss occur in the results based on the increasing the number of epochs and Figure C: show the color heat map that represents how many percentage of RGB colors contains in it.

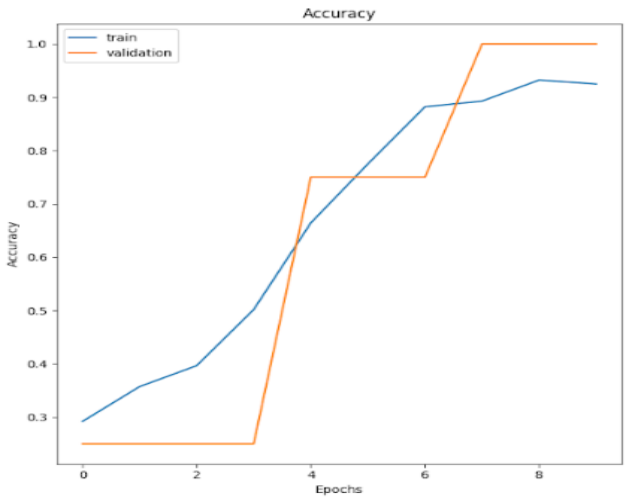

Fig A: Accuracy Result

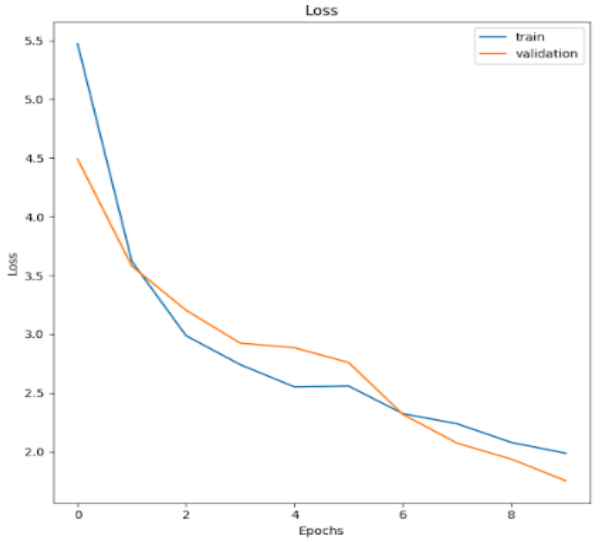

Fig B: Loss Result

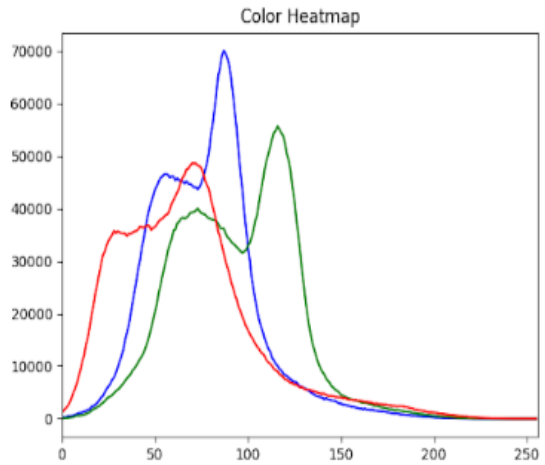

Fig C : Color heat map

VI. CONCLUSION

Predicting socio economic and developmental parameters like poverty levels of rural and urban areas from satellite images is challenging problem. Here we develop a desktop based classifier system that utilizing deep neural network architecture with supervised learning, feature extraction, and segmentation.

Using the satellite images as a input and some combination of the images features extracted through the CNN we are able to estimate the poverty close to the truth value. It utilized different pre-processing technique and utilized CNN classifier for detecting poverty of urban and rural areas.

VII.REFERENCES

1. Anthony Perez and Christopher Yeh "Poverty Prediction with Public Landsat 7 Satellite Imagery and Machine Learning", 2017. Landsat 7:Description of spectral Bands. https://landstat.usgs.gov.[Online;accessed8- June-2017]

2. Panchromatic Image Sharpening of Landsat 7 EMT https://landsat.usgs.gov./panchromatic-image-sharpening-landsat-7-etm.[Online;accesed 9-June-2017].

3. Boris Babenko, Jonathan Hersh, David Newhouse, Anusha Ramakrishnan, Tom Swartz "Poverty Mapping Using Convolutional Neural Networks Trained on High and Medium Resolution Satellite Images, With an Application in Mexico", 2017.

4. Barak Oshri, Annie Hu, Peter Adelson, Xiao Chen, Pascaline Dupas, Jeremy Weinstein, Marshall Burke, David Lobell, Stefano Ermon "Infrastructure Quality Assessment in Africa using Satellite Imagery and Deep Learning",2018.

5. K. He, X. Zhang, S. Ren, and J. Sun. Identity mappings in deep residual networks. CoRR, abs/1603.05027, 2016.

6. K. Simonyan and A. Zisserman. Very deep con-volutional networks for large-scale image recog-nition. CoRR, abs/1409.1556, 2014.

7. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In CVPR09, 2009.

