ISSN 2395-1621

Predicting poverty status of area from satellite images using CNN

[#1]Jyotsana Pote, [#2]Sneha Ugile, [#3]Pooja Narale, [#4]Prof. Vajid Khan

[1]jyotsna.pote10061998@gmail.com,
[2]snehaugile01@gmail.com,
[3]naralepooja11@gmail.com

[#1234]Department of Computer Engineering,
JSCOE, Hadapsar, Pune, Maharashtra, India.

ABSTRACT

The government is unable to estimate socio-economic status of a remote area and also they are unable to help them. Because government only has their satellite image as a record and they can only see that area through map but through this image they cannot get status about that area. So, considering this satellite image of an area, there is a profound need to detect status of the remote area. In this project, we propose an advanced framework to identify socio-economic status of area through satellite image. We are considering some major factors or attributes like water supply, roof tops, electricity and agriculture field and we are going to train some data sets through CNN technique then input satellite image is compare with train data sets and if there is presence of this factors in input image then we classify status of the area as poor, rich or medium.

Keywords- PyQT, NumPy, OpenCV, Satellite Image, LandSat 7, Google Earth, MySQL

ARTICLE INFO

Article History

Received: 14th November 2019

Received in revised form :

14th November 2019

Accepted: 18th November 2019

Published online :

18th November 2019

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 necessities of that region are?

In that case, the organization can only have the satellite 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.

The system which we are designing has the ability to identify some major factors which are very basic necessities of a region and they are electricity, water supply, agricultural field. One more factors we are using for estimate status of region is roof top of the house. Roof top is a very essential factor for our system. For prediction of socio-economic status, our system takes satellite image and then this satellite image 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 of presence of factors in the image and by considering this percentages of factors we are predicting socio-economic status.

To achieve required result, application is design through python language and using its libraries. So, to design user friendly desktop application, PyQT library method is used in python language. To preprocess data sets of satellite images, we are using OpenCV library method and through preprocessing of image, we converts our input satellite image into grayscale image,

contour image and smoothen image. To authenticate the user, we are using MySQL database connectivity.

The data sets of satellite image is collected from LandSat 7 and Google Earth website and along with this platform, there is another application (Magic Puzzle) on which by providing longitude and latitude we get satellite image of that area.Some times, satellite image is downloaded in the form multiple titles and by combining all the tiles, we are getting a single whole image. In this way, we are collecting data sets of satellite image.

The overview of survey paper is that it contain literature survey and then it contain System Architecture and then conclusion and references along with URL of data sets of satellite images.

II. LITERATURE SURVEY

They propose a two-step approach for predicting poverty in rural regions of India from satellite imagery. First, they train a multi-task fully convolutional model to predict three developmental parameters – the main material of the roof, 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 by significant manual effort and monetary expense. Their main contribution is a two step approach for poverty prediction. First, they engineer a multi-task fully convolutional model to predict the material of roof, source of lighting and source of drinking water from the satellite imagery of a village. The results presented by them clearly support the effectiveness of their approach. In this way, their experiments suggest that predicting poverty levels from multiple developmental parameters is more reliable than using a single parameter. Their main focus on only three factors which are the material of roof, source of lighting and source of drinking water present in input satellite image.[1]

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 Afrobarometer survey data, introduced a deep learning approach that demonstrates good predictive ability. Their results demonstrate the proof of concept that satellite imagery can be used to predict infrastructure quality.[2]

Their 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.[3]

Presents the CNN predictions for urban areas using imagery for either Digital Globe or Planet, using the validation sample. They present R2 estimates that show the correlation between predicted poverty and benchmark poverty as measured in the 2015 Intercensus. The drop in performance is modest but not severe. Poverty estimates for urban areas in Mexico are mapped. Their main focus is on comparing predicted poverty level with actual poverty level.[4]

The similarity between all research papers is that they all are using CNN methods for predicting socio-economic status. The reason behind using CNN is its accuracy of predicting result.

III. SYSTEM ARCHITECTURE

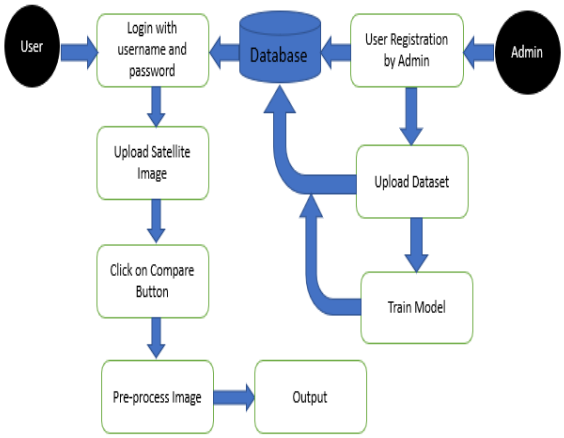

Figure No. 3.1 system Architecture

In the diagram, there is flow of our project.

1. The whole architecture is made by PyQT library used in python language. PyQT library gives all the necessary stuff related to GUI design. PyQT provides us display screen, buttons and so on. So, In this way PyQT helps us in design GUI.

2. After designing of GUI, another task is to authenticate valid user for operating application. To deal with this task, we are using MySQL database to store data of username and password and through this, user can authenticate easily.

3. Another task is to preprocess the input image which can be done by OpenCV library of python. By using this library, image is converted into grayscale image, contour image and smoothen image.

4. The major task of this survey paper is to collect data sets of satellite image and to achieve this result, we are working on Google Earth images, LandSat 7 images and also take help of magic puzzle application on which, we are providing latitude and longitude of a particular area and as a result, we are getting satellite image of that area. In this way, we achieve our all the tasks to achieve our project goal.

IV. CONCLUSION

In our application, we have first collect data sets of satellite image and after that make a desktop application so that user can able to predict socio-economic status. To predict status of a satellite image, we have use preprocessing of a input image so that features can be easily detected from input mage and to achieve this we are using opencv library. In this way, we are successfully implementing all the tasks of the survey paper.

V. FUTURE WORK

In future work, we will apply CNN methods for training and testing model through collected data sets of satellite images and try to predict socio-economic status of input satellite image. Along with CNN methods, we can use decision tree algorithm or naïve baiyes to predict that the area is poor, rich or medium in terms of occurrence of factors in that input image.

VI. REFERENCES

[1] Pandey, S. M., Agarwal, T., & Krishnan, N. C. (2014). Multi-Task Deep Learning for Predicting Poverty from Satellite Images. Ropar: The Thirtieth AAAI Conference on Innovative Applications of Artificial Intelligence (IAAI-18).

[2] OSHRI, B., HU, A., ADELSON, P., & LOBELL, D. (2018). Infrastructure Quality Assessment in Africa using Satellite Imagery and Deep Learning. Stanford: Association for Computing Machinery.

[3] Perez, A., Yeh, C., Azzari, G., Burke, M., Lobell, D., & Ermon, S. (2017). Poverty Prediction with Public Landsat 7 Satellite Imagery and Machine Learning.

California: 31st Conference on Neural Information Processing Systems.

[4] Babenko,B.,Hersh,J.,Newhouse,D.,Ramakrishnan, A., & Swartz, T. (2017). Poverty Mapping Using Convolutional Neural Networks Trained on High and Medium Resolution Satellite Images, With an Application in Mexico.

California: 31st Conference on Neural Information Processing Systems.

