

Real Estate House and Land Property rate Prediction using CNN

Manoj Gandhi, Agnihotri

Department of IT, Jeybee Institute of Technology, Sriperumpudur, Kanchipuram, India.

Article Information

Received : 06 Oct 2023
Revised : 14 Oct 2023
Accepted : 10 Oct 2023
Published : 16 Nov 2023

Corresponding Author:

Abstract— Identifying the physical aspect of the earth's surface (Land cover) and also how we exploit the land (Land use) is a challenging problem in environment monitoring and much of other subdomains. One of the most efficient ways to do this is through Remote Sensing (analyzing satellite images). For such classification using satellite images, there exist many algorithms and methods, but they have several problems associated with them, such as improper feature extraction, poor efficiency, etc. Problems associated with established land-use classification methods can be solved by using various optimization techniques with the Convolutional neural networks(CNN). The structure of the Convolutional neural network model is modified to improve the classification performance, and the overfitting phenomenon that may occur during training is avoided by optimizing the training algorithm. This work mainly focuses on classifying land types such as forest lands, bare lands, residential buildings, Rivers, Highways, cultivated lands, etc. The outcome of this work can be further processed for monitoring in various domains.

Keywords: *Convolution Neural Networks(CNN), Deep Learning, Land Classification.*

Copyright © 2023: Manoj Gandhi, Agnihotri, This is an open access distribution, and reproduction in any medium, provided Access article distributed under the Creative Commons Attribution License the original work is properly cited License, which permits unrestricted use.

Citation: Manoj Gandhi, Agnihotri, "Real Estate House and Land Property rate Prediction using CNN", Journal of Science, Computing and Engineering Research, 6(11), 19-20, November 2023.

I. INTRODUCTION

Land-cover and land-use change (LUCC), that is associated with imbalances in ecosystems, biodiversity, and global climate changes, show the consequences of human activities and climatic changes on the ecological atmosphere of the Earth's surface. The international community gives more significance to LUCC because of the key content of worldwide environmental change research. Since the 1990s, various research institutions such as Food and Agriculture Organization(FAO), International Geosphere-Biosphere Project(IGBP), International Institute for Applied Systems Analysis(IIASA), etc. have launched several LUCC related projects [15]. Remote sensing is a potential tool for keeping track of the Earth's surface and also a fundamental element for works that use classification and identification technologies to look over the land-use status [16].

While performing field surveys is more comprehensive and authoritative, it's an upscale task and mostly takes an extended time to update. It is made easy with remote sensing. Land use or land cover change detection is very important for a greater understanding of topography dynamics. Land use and land cover changes could also be an omnipresent and accelerating process, primarily driven by natural circumstances and anthropogenic undertakings, that successively reflects changes that may affect natural ecosystems. Remote sensing images assist us in discovering solutions to the increasing number of environmental challenges that we face today. It helps not only in getting a bird's eye view of what's around but also in discovering

parts of the world that are hardly seen. The potential of categorizing land use and land cover allows humans to more competently utilize natural resources and thereby decreasing waste and deprivation. In spite of its potential to be incredibly useful, satellite data is very huge and confusing, and making sense out of it requires complex analysis. Understanding land use/land cover has become more and more vital because the country plans to overcome the issues of hazard, unrestrained development, loss of superior agricultural lands, worsening environmental standards, demolition of important wetlands, and so forth. If standards and living conditions are to be upgraded or maintained at modern-day ranges, land use records are needed in the analysis of environmental activities.

II. RELATED WORK

Land cover and Land use are two separate terms that we often use interchangeably. Land cover refers back to the bodily traits of the earth's floor, captured within the distribution of flora, soil, water, and different physical characteristics of the land, together with the ones created completely via human activities.

Utilization of land by people and their habitat refers to land-use, usually with the emphasis at the utilitarian position of land for economic utilizations eg. construction of buildings, agricultural lands, etc. Land use or land cover mapping has been a beneficial and specific way to improve the selection of regions designed to agricultural, urban, and business regions of a community with the discovery of

remote sensing and Geographical Information System(GIS) approaches. Applications of satellite data made it possible to study the adjustments in land cover at low cost, in less time, and with better accuracy in integration with GIS that acts as an effective platform for records evaluation or analysis, replacement, and retrieval. Several types of land-use classification algorithms or standards exist, which include several classes and are reasonable for the complex features of land-use/and-cover types, and these traits pose complexities for precise classification. Determining the classification method, followed by selecting the precise classifier is essential in the classification of remote sensing images. Classification methods consist of supervised/unsupervised algorithms, the authentic use of feature extraction from spectral information and, hard or soft classification. Due to differences in the basic unit of classification, classification strategies can be divided into pixel-based and object-oriented classifications[6]. Regarding classifier selection, the conventional technique used is the statistical procedure for low-level feature extraction, along with distance[7], K-nearest neighbor[8], maximum likelihood[9], and logistic regression[10] classifiers. High-resolution remote sensing (HRRS) images are applied in land-use classification with the fast evolution of aerospace, computer, and sensor technologies[13]. The range of objects inside a given class expands as does the similarity of objects in various classes because of spectral confusion in these high-resolution remote sensing images. These properties decrease the effectiveness of conventional classification techniques based on low-level features/attributes. Therefore, the method built on modeling the mid-level features has evolved based on the low-level feature technique[14]. The bag-of-visual-words (BoVW) is one of the three types of mid-level feature extraction methods that describe image features. The performance of BoVW-based procedures relies on the extraction of features in practical applications. The other two mid-level feature extraction techniques are latent Dirichlet allocation(LDA) and machine learning model. The machine learning models unaided perform data analysis and feature extraction and remove the pattern of the extracted features in line with pre-established rules. Thus, these models are able to obtain enhanced classification results when put in with the complex images. Sparse coding, support vector machines, neural networks, and deep learning are commonly used machine learning methods[11]. The deep learning networks are made of various non-linear layers, which constitute a state-of-the-art method of intelligent pattern recognition and are critical new conduct in the province of remote sensing image processing. Research on applying Convolution neural networks to remote sensing images has surfaced over recent years. The Hinton team successfully topped the ImageNet image classification contest and minimized the error rate(top-5) of 1000 images from 26.2 to 15.3%[12]. Hu et al was the first to use a Convolution neural network model to classify HRRS images[5]. Chen et al integrated a CNN classification technique that utilizes spatial information, pixel spectral information, and explored the significance of spatial

information in the classification of High-resolution remote sensing(HRRS) images[17].

III. PROPOSED TECHNIQUES

In this work, the implementation is carried out as steps mentioned below which are also depicted in Fig. 1. A. Data Collection and Data Preparation B. Simple CNN model C. Improvised CNN D. Using Transfer Learning These steps are explained in the following subsections.

A. Data Collection and Data Preparation For this work, we have used open-source EuroSAT Sentinel-2 satellite images from the German Research Center for Artificial Intelligence, which can be downloaded locally [link]. The dataset which we have considered has about 27,000 labeled images. And these images are of 10 different land use classes, which are Highway, Forest, Herbaceous Vegetation, Annual Crop, Industrial, etc. Each multispectral image consists of 13 different color bands that represent different wavelengths of light or color and different resolutions. These different light bands help distinguish parts of the landscape that reflect certain sorts of light, especially ways. Preprocessing of satellite images involves radiometric calibration and geometric correction. The motive of preprocessing is to convey the information contained in the activity of image synthesis further to furnish it in the neighborhood to the actual image. The elimination of image contortion caused by radiation errors is voided by radiation correction. Geometric correction ensures that the total error of the corrected position is no more than one pixel. And also images are translated from 'tif' form to 'jpeg' form as required using rasterio. And it is very important to split the dataset into train, tests, and valid folders.

B. Simple CNN Model The first convolutional neural network explored in this work was a simple CNN with a three-layer architecture: one convolutional input layer, one flatten layer in the middle and one dense output layer as shown in Fig. 2. The simplicity of the first model is a relatively fast test that the data preparation was sufficient and that the neural network was set up with the right parameters. However, this model has a very low accuracy of around 30% as shown in Table-III. Since satellite data is pretty complex this model will require more layers for it to learn the distinct features of each category of the dataset. So we have improvised the CNN model by adding a few more layers in the next step.

C. Improvised CNN In this network, we have added a few more layers as shown in Table-I to learn the distinct features. Each image will be fed into the CNN model. The first layer, with a 5x5 kernel size, is a Convolutional 2D layer. The results of convolution are feature maps with size 28x28. After that, the Max Pooling layer is performed as subsampling with kernel 2x2 (non-overlapping). The results of the max-pooling layer are feature maps with size 14x14 (because of non-overlapping so stride or movement of the kernel is 2). The convolution layer and Max Pooling layer are then performed. Each output in the convolutional layer is often activated by using ReLU. The dropout technique is

incorporated here as a regularization technique. Using this CNN model the accuracy is successfully improved to 90%. But to improve the accuracy furthermore, transfer learning with CNN is used in the next step.

IV. RESULTS AND DISCUSSIONS

After the data preparation, the first convolutional neural network(CNN) explored was a simple CNN with a three-layer architecture. The model was compiled using Adam optimizer, which is a stochastic optimization function that's computationally efficient. The accuracy of the model during training stayed constant at 30% for almost all training epochs. The confusion matrix of the model is evaluated and it turned out that this model is way too simple. This model must not have learned much, because it guessed every test image as the same category. No wonder the accuracy was around 30%, which is what one would expect if they were randomly guessing for a dataset with 10 categories.

Since the simple model was not very good, the next step was to use an improvised CNN in which we personalized the architecture to improve the performance of the CNN model by adding a few more layers. Using this CNN model, accuracy of 90% is achieved as shown in Table-III. To improve the accuracy of the classification, the next step was to use transfer learning to improve the performance of this model on the dataset. In this case, the ImageNet VGG16 model was used. Data Augmentation and Dropout regularization are incorporated to boost the effectiveness of the model. Data augmentation increases the size of the dataset by creating multiple versions of the same image that are slightly transformed. Even though this model took very long to run for ten epochs, it out-performed the simple CNN by an impressive amount. This model achieved 95% accuracy, guessing most categories in the test set correctly. We then experimented with the Batch Normalization technique in this model. Batch normalization improves the speed and performance of a neural network by normalizing the data to a standard scale in the model's learning process. With ten epochs of training, the tuned model achieved a 56% accuracy. Batch normalization is known to speed up model training, allowing for larger learning rates and fewer epochs. With Batch Normalization, the model took less time to compile than before. Despite the difference in running time, the tuned model's accuracy is continuing to improve up until the last epoch, and would likely continue improving if given more epochs to train. For the purpose of this work, all model parameters are kept the same, but it might have tuned the model much faster, and reap the benefits of batch normalization if the learning rate is increased to something larger than 0.0001. A. Experimental Setup For experimental setup, we have using NSL-KDD intrusion dataset and computerized data analyzing tool WEKA is used to perform the classification testing. Because the Weka tool is data mining process. It contains the clustering, pre-processing, classification, regression and feature

V. CONCLUSION

This paper, we proposed a naive Bayes, SVM and decision tree algorithms using NSL-KDD dataset. In this dataset 41

attributes are other. In this work, the first simple model is explored just as a relatively fast test and as it has very few layers, it has failed to learn the features. To better the efficiency of the CNN model or to learn the features more effectively, few more layers are added to simple CNN which is improvised CNN. It successfully improved the classification accuracy, but to further improve the accuracy, CNN with Transfer learning is explored. Some optimization techniques such as data augmentation, dropout regularization, etc. have been used.

REFERENCES

- [1]. P. Nirmala, T. Manimegalai, J. R. Arunkumar, S. Vimala, G. Vinoth Rajkumar, Raja Raju, "A Mechanism for Detecting the Intruder in the Network through a Stacking Dilated CNN Model", *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 1955009, 13 pages, 2022. <https://doi.org/10.1155/2022/1955009>.
- [2]. J.R.Arunkumar, Dr.E.Muthukumar,"A Novel Method to Improve AODV Protocol for WSN" *Journal of Engineering Sciences*" Volume 3, Issue 1, Jul 2012. ISSN NO: 0377-9254
- [3]. J. R. Arunkumar, S. Velmurugan, B. Chinnaiyah, G. Charulatha, M. Ramkumar Prabhu et al., "Logistic regression with elliptical curve cryptography to establish secure iot," *Computer Systems Science and Engineering*, vol. 45, no.3, pp. 2635–2645, 2023.
- [4]. P. K. Devi, D. Arulanantham, C. Kalaivanan, N. Gomathi, J. R. Arunkumar and G. Ramkumar, "An Secure and Low Energy Consumption based Intelligent Street Light Managing System using LoRa Network," 2022 6th International Conference on Electronics, Communication and Aerospace Technology, Coimbatore, India, 2022, pp. 638-645, doi: 10.1109/ICECA55336.2022.10009408.
- [5]. Prathima Chilukuri , J.R. Arun Kumar , R. Anusuya , M. Ramkumar Prabhu. "Auto Encoders and Decoders Techniques of Convolutional Neural Network Approach for Image Denoising In Deep Learning" *Journal of Pharmaceutical Negative Results*, 13(4), 1036–1040. <https://doi.org/10.47750/pnr.2022.13.04.142> ,November 4, 2022.
- [6]. R. Yugha, V. Vinodhini, J. R. Arunkumar, K. Varalakshmi, G. Karthikeyan and G. Ramkumar, "An Automated Glaucoma Detection from Fundus Images based on Deep Learning Network," 2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Dharan, Nepal, 2022, pp. 757-763, doi: 10.1109/I-SMAC55078.2022.9987254.
- [7]. E. Thenmozhi, A. Karunakaran, J. R. Arunkumar, V. Chinnammal, C. Kalaivanan and G. Anitha, "An Efficient Object Detection and Classification from Restored Thermal Images based on Mask RCNN," 2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Dharan, Nepal, 2022, pp. 639-645, doi: 10.1109/I-SMAC55078.2022.9987422.
- [8]. L. Saravanan, W. Nancy, K. P. Chandran, D. Vijayanandh, J. R. Arunkumar and R. T. Prabhu, "A Novel Approach for a Smart Early Flood Detection and Awareness System using IoT," 2022 8th International Conference on Smart Structures and Systems (ICSSS), Chennai, India, 2022, pp. 1-4, doi: 10.1109/ICSSS54381.2022.9782286.
- [9]. S. Bharathi, A. Balaji, D. Irene. J. C. Kalaivanan and R. Anusuya, "An Efficient Liver Disease Prediction based on Deep Convolutional Neural Network using Biopsy Images," 2022 3rd International Conference on Smart Electronics and

- Communication (ICOSEC), Trichy, India, 2022, pp. 1141-1147, doi: 10.1109/ICOSEC54921.2022.9951870.
- [10].Prathima, C. H., Anusuya, R., & Prabhu, M. R. K. (2022). Comprehensive Design Analysis of Digital Marketing in Agriculture Sector. *International Journal of Early Childhood Special Education*, 14(5), 2022.
- [11].Atul Kumar Dwivedi, Deepali Virmani, Anusuya Ramasamy,Purnendu Bikash Acharjee, Mohit Tiwari” Modelling And Analysis Of Artificial Intelligence Approaches In Enhancing The Speech Recognition For Effective Multi-Functional Machine Learning Platform – A Multi Regression Modelling Approach ” *Journal of Engineering Research - ICMET Special Issue*, 2022-04-06.
- [12].M.Ramkumar Prabhu, A.Rajalingam, J.R.Arunkumar, Dr.R.Anusuya” Microstrip Patch Antenna Using Combined Slots for Bandwidth Enhancement and Size”, *Journal of Engineering Sciences*, Vol 11, Issue 1, Jan, 2020, ISSN NO: 0377-9254.
- [13].Anusuya Ramasamy, Abel Adane Changare” Hybrid Fuzzy Knowledge Based Prediction Model for the Software Development and Maintenance Quality in Software Engineering Approach” *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, ISSN: 2278-3075, Volume-9 Issue-10, August 2020.
- [14].J.R.Arunkumar,” Chaotic African Buffalo Optimization Based Efficient Key Mechanism in Categorized Sensor Networks:, *International Journal of Engineering and Advanced Technology (IJEAT)*, ISSN: 2249 – 8958, Volume-9 Issue-3, February, 2020.
- [15]. R. Anusuya, M. Ramkumar Prabhu, Ch. Prathima, J. R. Arun Kumar” Detection of TCP, UDP and ICMP DDOS attacks in SDN Using Machine Learning approach” *Journal of Survey in Fisheries Sciences*, Vol. 10 No. 4S (2023): Special Issue 4.
- [16].M.Ramkumar Prabhu, A.Rajalingam, J.R.Arunkumar, R.Anusuya, “Microstrip Patch Antenna Using Combined Slots For Bandwidth Enhancement And Size” *Journal of Engineering Sciences (JES)*, Vol 11, Issue 1, Jan / 2020, ISSN NO: 0377-9254.
- [17].M.Ramkumar Prabhu, J.R.Arunkumar, A.Rajalingam, R.Anusuya “A Modified Square Patch Antenna with Rhombus slot for High bandwidth” *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* ISSN: 2278-3075, Volume-8 Issue-9, July 2019.
- [18].Revanes, M., Gundal, S.S., Arunkumar, J.R.Arunkumar et al. Artificial neural networks-based improved Levenberg–Marquardt neural network for energy efficiency and anomaly detection in WSN. *Wireless Netw* (2023). <https://doi.org/10.1007/s11276-023-03297-6>.
- [19].I. Chandra, K. V. Karthikeyan, R. V, S. K, M. Tamilselvi and J. R. Arunkumar, "A Robust and Efficient Computational Offloading and Task Scheduling Model in Mobile Cloud Computing," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-8, doi: 10.1109/ICECONF57129.2023.10084293.
- [20].Jangam Raghunath, S Kiran, G Siva Nageswara Rao, JR Arun Kumar, R Anasuya, C Siva Kumar,” A MACHINE LEARNING TECHNIQUE TO DETECT BEHAVIOR BASED MALWARE”, *Semiconductor Optoelectronics*, Vol. 42 No. 1 (2023), 1268-1278
- [21].Dr.J.R.Arunkumar. “Enhanced Dynamic Authorized Secured Protocol for Wireless Sensor Networks, “*Journal of Science, Computing and Engineering Research*, 1(1), 07-11, Mar-Apr 2020.
- [22].Anitha Gopalan,O. Vignesh,R. Anusuya,K. P. Senthilkumar,V. S. Nishok,T. Helan Vidhya, Florin Wilfred,” Reconstructing the Photoacoustic Image with High Quality using the Deep Neural Network Model”, *Contrast Media & Molecular Imaging*, Volume 2023 | Article ID 1172473 | <https://doi.org/10.1155/2023/1172473>.
- [23].R. Anusuya, N. Anusha, V. Sujatha, R. Radhika and S. Inian, "Machine Learning based Landslide Detection System," 2023 7th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2023, pp. 319-323, doi: 10.1109/ICCMC56507.2023.10084226.
- [24].S. Sivakumar, R. Anusuya, V. Nagaraju, L. P. Narendruni and R. Thamizhamuthu, "QoS Based Efficient Link and Consistent Routing in Wireless Sensor Network," 2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), Bengaluru, India, 2023, pp. 1241-1246, doi: 10.1109/IITCEE57236.2023.10091080.
- [25].I. Chandra, G. Sowmiya, G. Charulatha, S. D, S. Gomathi and R. Anusuya, "An efficient Intelligent Systems for Low-Power Consumption Zigbee-Based Wearable Device for Voice Data Transmission," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ICECONF57129.2023.10083856.
- [26].R. Meena, T. Kavitha, A. K. S, D. M. Mathew, R. Anusuya and G. Karthik, "Extracting Behavioral Characteristics of College Students Using Data Mining on Big Data," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ICECONF57129.2023.10084276.
- [27].G. Karthikeyan, D. T. G, R. Anusuya, K. K. G, J. T and R. T. Prabu, "Real-Time Sidewalk Crack Identification and Classification based on Convolutional Neural Network using Thermal Images," 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), Pudukkottai, India, 2022, pp. 1266-1274, doi: 10.1109/ICACRS55517.2022.10029202.
- [28].S. Bharathi, A. Balaji, D. Irene, J. C. Kalaivanan and R. Anusuya, "An Efficient Liver Disease Prediction based on Deep Convolutional Neural Network using Biopsy Images," 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2022, pp. 1141-1147, doi: 10.1109/ICOSEC54921.2022.9951870.
- [29].Dr. R.Anusuya. —Stacking Dilated CNN Authorized Secured Protocol for IoT Security, —*Journal of Science, Computing and Engineering Research*, 1(1), 01-07, May- June 2022.