Deep Learning based Approach for Land Surface Identification

Hitesh K. Sharma [1], Tanupriya Choudhury [2], Sachi N. Mohanty [3], G R Anil [4], Rohini A [5]

[1,2] *University of Petroleum & Energy Studies (UPES), Dehradun- 248007,India.**
[3] *Adjunct Professor, Singidunum University, Serbia and School of Computer Science & Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India.*
[4] *Vardhaman College of Engineering(Autonomous), Hyderabad, Andhra Pradesh 501218, India.*
[5] *Anil Neerukonda Institute of Technology and Sciences, Vishakapatnam, AndhraPradesh,531162, India.*

Abstract

Impenetrable surface has been perceived as a critical marker in evaluating metropolitan conditions. Nonetheless, precise impenetrable surface extraction is as yet a test. Adequacy of impenetrable surface in metropolitan land-use grouping has not been all around tended to. This paper investigated extraction of impenetrable surface data from Landsat Enhanced Thematic Mapper information dependent on the reconciliation of part pictures from direct ghastly combination examination and land surface temperature. Another methodology for metropolitan land-use characterization, in light of the joined utilization of impenetrable surface furthermore, populace thickness, was created. Five metropolitan land-use classes (i.e., low-, medium-, high-, and extremely focused energy neighborhoods, and business/modern/transportation utilizes) were created in the city of Indianapolis, Indiana, USA. Results showed that the incorporation of division pictures and surface temperature gave generously worked on impenetrable surface picture. Precision evaluation demonstrated that the root mean-square mistake and framework blunder yielded 9.22% and 5.68%, individually, for the impenetrable surface picture. The general arrangement precision of 83.78% for five metropolitan land-use classes was obtained.

Keywords

Land-Use, Classification, , Machine Learning, Convolutional Neural Network (CNN), Deep Learning, LULC.

1. Introduction

Land surface identification is a major use case in metropolitan cities. Exact identification of locality, identification of beach, river, parks. Building etc. provides a major advantage to plan some good initiatives for making the metropolitan cities more organized and less crowded. The ecosystem of various places can also be managed with proper understanding of cities with proper categorization. Machine Learning (ML) techniques and ML based models can help to solve this major issues. The image datasets of various locations of a city can help to analyze the ecosystem of that city and it can be helpful for the government to take some good initiative for betterment of urban lifestyle. Digital Image Processing (DIP) [3] with a basic knowledge of computer vision and IoT can help to solve this issue in metropolitan cities. There are many machine learning algorithms those are very useful in this use case. Land surface identification require the combination of many advanced IT based technologies including Internet of Things (IoT), Image Processing, Computer Vision, Deep Learning etc. Some of the main ML algorithms used for this problem

ACI'22: Workshop on Advances in Computation Intelligence, its Concepts & Applications at ISIC 2022, May 17-19, Savannah, United States
EMAIL: hkshitesh@gmail.com (A.1); tanupriya1986@gmail.com*, (A.2); sachinandan09@gmail.com (A.3); anilgrcse@gmail.com (A.4); rohinaruna@gmail.com (A.5)
ORCID: 0000-0001-6816-0324 (A. 1); 0000-0002-9826-2759 (A. 2); 0000-0002-4939-0797 (A. 3); 0000-0002-4256-5873 (A. 4); 0000-0001-5809-7317 (A. 5)

set are Support Vector Machine (SVM), Random Forest Classifier. Any of the two Supervised and Un-Supervised learning algorithms can be used in this use case.

In this research work, we have used two widely used and popular algorithms named as Support Vector Machine and Random Forest Classifiers. The dataset has been passed to both the algorithmic model and results has been compared.

2. Data Set

Paragraph text Land Use Dataset provided by UC Merced in October 28 [14]. This dataset has 21 different land use classes and it is generated by UC Merced for researchers [14]. Each Class contains 100 images and each image contain a size of 256X256 pixel dimensions. Each image measures 256x256 pixels. These images are captured from many different areas of Urban Area of USGS National Map Imagery from different locations of the country. The pixel resolution of this public domain imagery is 1 foot. The dataset contain diversity in type of images. The dataset contains images of river, beaches, mountains, building, roads etc. The images are captured with a quality device. In figure 1, we have shown some sample images from the dataset.

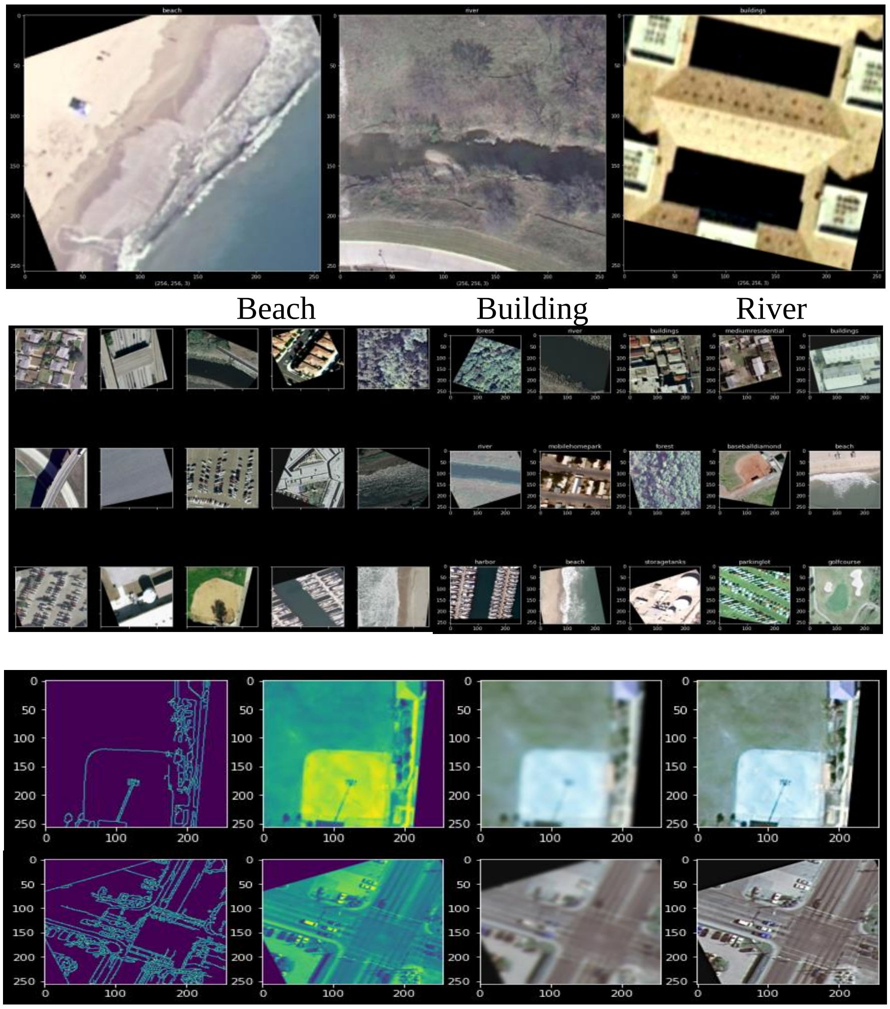


Figure 1. Image with Edges, Colored, Blurred, & Sharp

3. Dataset Standardization

An Image generator is created to standardize the input images. It is used to adjust the images in such way that the mean of pixel intensity will be zero and standard deviation will become 1. The old pixel value of an image will be replaced by new values calculated using following formula. In this formula each value will subtract mean and divide the result with standard deviation.
The formula is given in the equation (Equation 1) as under:-

$$xi = \frac{xi-\mu}{\sigma} \text{ ....(Equation 1)}$$

The pixel intensity is ranging from 0.0 to 1.0. In this case, the mean will never be zero and standard deviation will never be 1. To eliminate this problem above standardization formulation is applied on each pixel value and resultant values after application has used for the work.

4. Proposed CNN Model

4.1. Convolutional Neural Network (CNN)

CNNs are necessary to numerous PC vision applications, for example, object discovery and arrangement. Generally, CNNs [4] are comprised of a few unique sorts of layers consecutively linked together. The two fundamental kinds of layers, the convolutional layer and the completely associated layer, make up an enormous piece of the CNN design (Fig. 2).

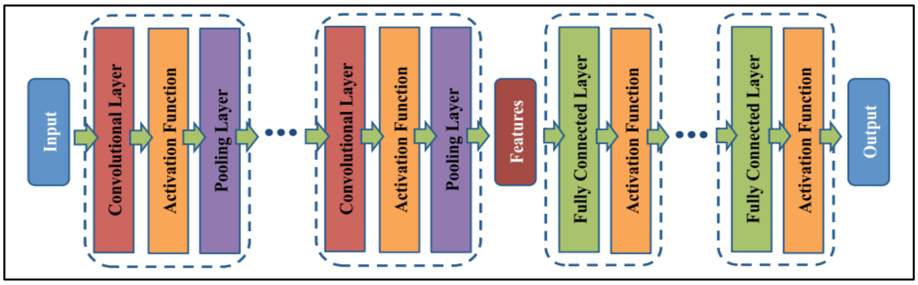


Figure 2: CNN Architecture

In the convolutional layer, the convolution activity is performed between the information and the learned convolutional loads. In the completely associated layer, network augmentation is performed between the element framework and the learned component weight grid. There are likewise different tasks required, for example, applying a non-direct initiation work like corrected straight units (ReLUs) to the yield of a layer, and a pooling layer, for example, greatest or normal pooling, to decrease the component map size [5][8]. CNN is a deep learning algorithm. It is a type of a neural network which is usually applied to 2D images dataset. A convolutional neural network (CNN) is a deep learning algorithm. A CNN contains multiple layers, a basic neural network [12] consists of three layers i.e input layers, hidden layers, and output layers, each layer consists of neurons. A CNN takes images as input adds weights and biases to them. The upper hand for using CNN over other learning algorithms is that it can extract important features from the images without any human assistance, for example, if we want to detect images containing skin cancer it can detect the important features to distinguish from a cancerous skin mole to a non-cancerous skin mole it can detect the important features itself, and the main edge that it has over other learning algorithm is its efficiency and accuracy. [13].

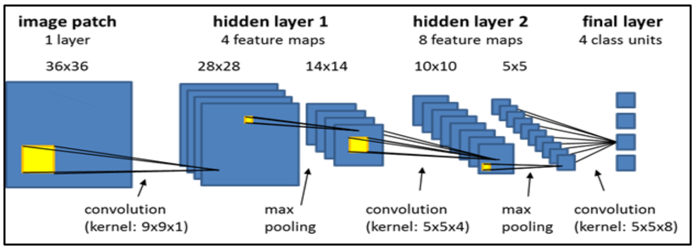


Figure 3. Example of Convolutional Neural Network.

There are five different layers in a CNN (Fig. 3)
1. Input Layer (Layer 1)
2. Convolutional Layer (Layer 2)
3. Max Pooling Layer (Layer 3)
4. Logistic/Softmax Layer (Layer 4)
5. Classification/Prediction/ Output Layer (Layer 5)

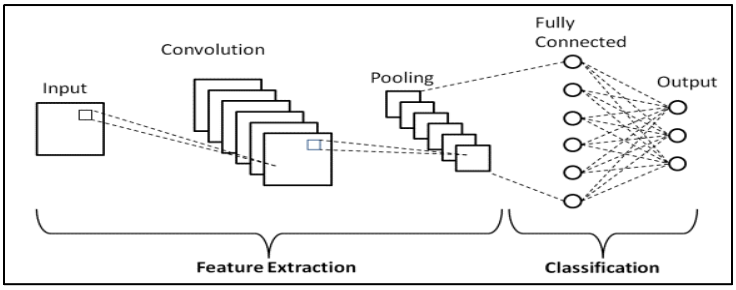


Figure 4: Different layers of Convolutional Neural Network.

The first layer of a CNN is the input layer which has the images and patterns [2] in the form of a matrix and each value of the matrix denotes a pixel and its value depending on the RGB color that pixel has which are further processed to the further layers (figure 4). Convo layer if the feature extraction layer in a CNN it extracts the important features of the images, convo layers contains activation functions such as RELU (changes all the negative value to 0). The pooling layer or the filter layer is very important to make the computation fast, it is used between each convo layer, it reduces the spatial volume in the images thus making the computation faster. The softmax or the logistic layer is used at the end of the neural network depending on the classification [6] we want in our model, if we need binary classification then we use a logistic function and for multiclassification. we have used softmax function (Fig. 5). The output layer is the end of the neural network which provides the value that we need for classification. The second CNN model has been described in Fig. 6.

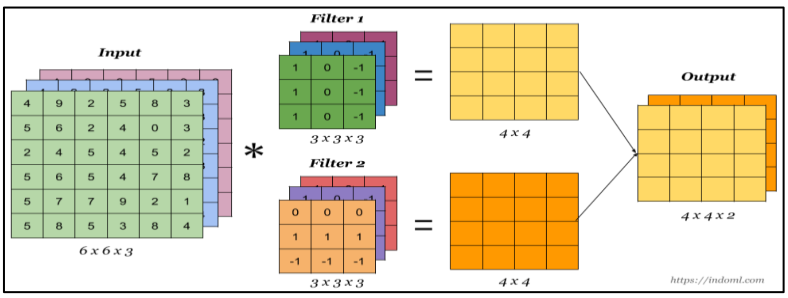
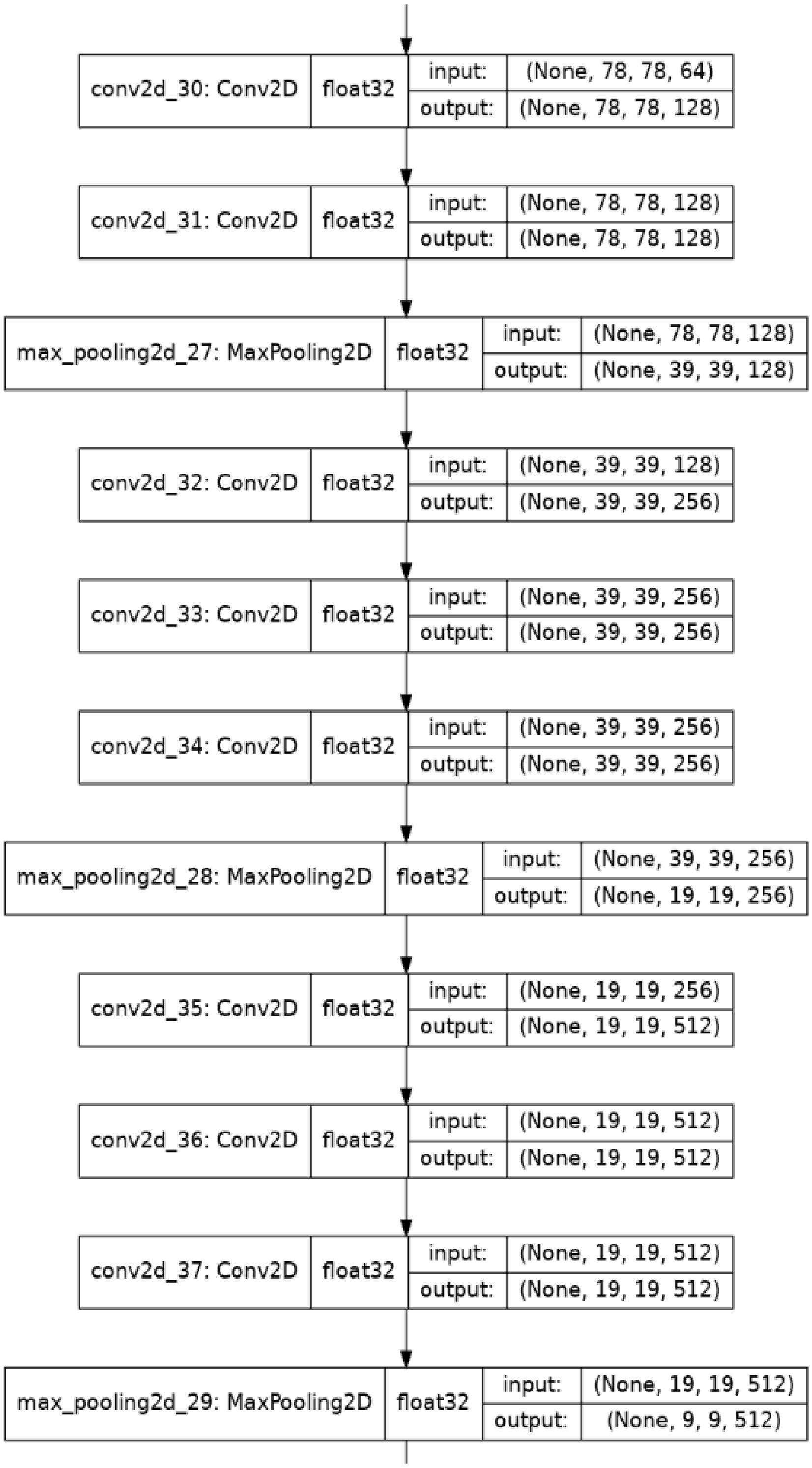
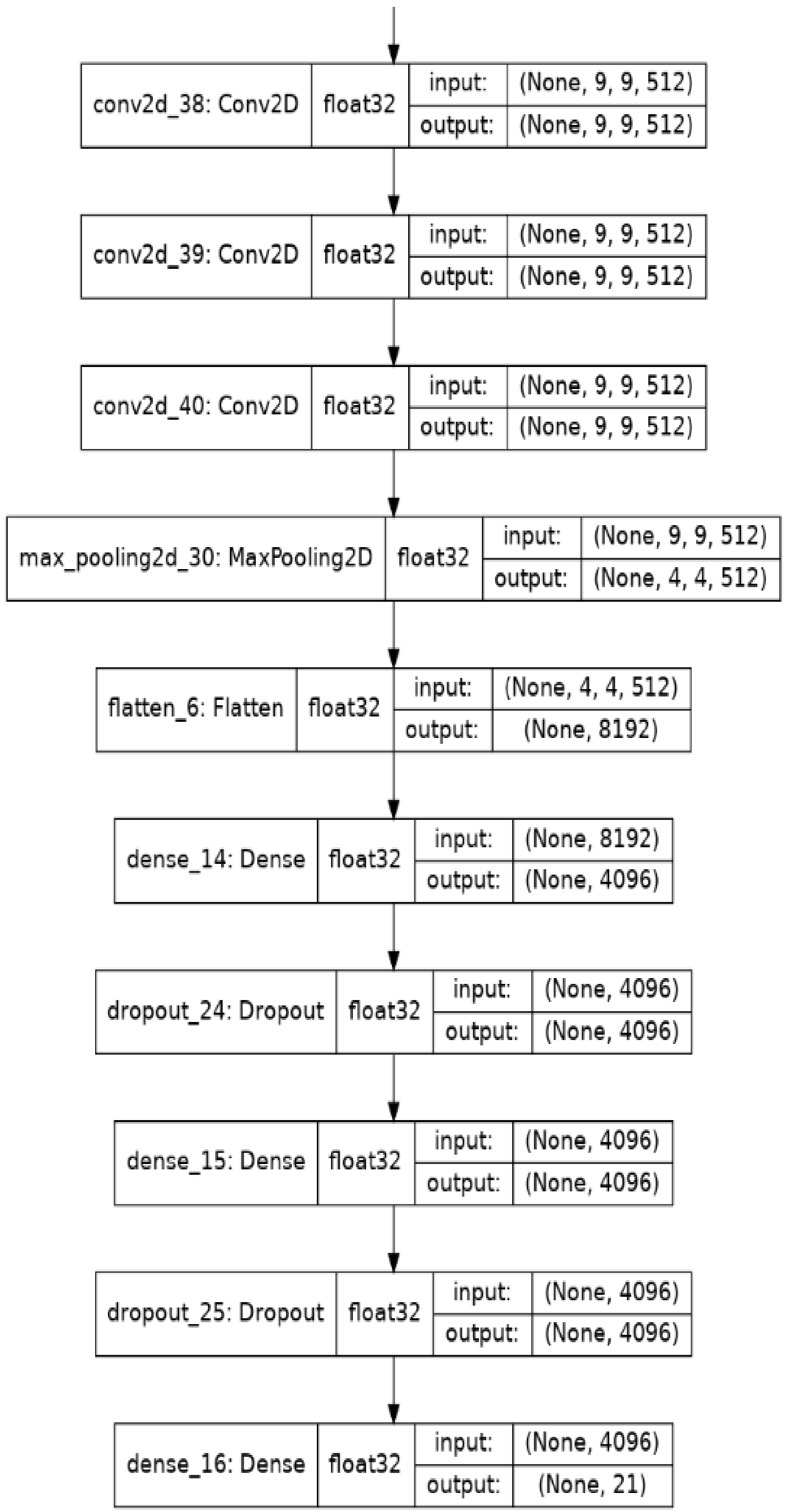


Figure 5: Filter layer breakdown in a CNN Model.


conv2d_30: Conv2D
float32
input:
(None, 78, 78, 64)
output:
(None, 78, 78, 128)
conv2d_31: Conv2D
float32
input:
(None, 78, 78, 128)
output:
(None, 78, 78, 128)
max_pooling2d_27: MaxPooling2D
float32
input:
(None, 78, 78, 128)
output:
(None, 39, 39, 128)
conv2d_32: Conv2D
float32
input:
(None, 39, 39, 128)
output:
(None, 39, 39, 256)
conv2d_33: Conv2D
float32
input:
(None, 39, 39, 256)
output:
(None, 39, 39, 256)
conv2d_34: Conv2D
float32
input:
(None, 39, 39, 256)
output:
(None, 39, 39, 256)
max_pooling2d_28: MaxPooling2D
float32
input:
(None, 39, 39, 256)
output:
(None, 19, 19, 256)
conv2d_35: Conv2D
float32
input:
(None, 19, 19, 256)
output:
(None, 19, 19, 512)
conv2d_36: Conv2D
float32
input:
(None, 19, 19, 512)
output:
(None, 19, 19, 512)
conv2d_37: Conv2D
float32
input:
(None, 19, 19, 512)
output:
(None, 19, 19, 512)
max_pooling2d_29: MaxPooling2D
float32
input:
(None, 19, 19, 512)
output:
(None, 9, 9, 512)


conv2d_38: Conv2D
float32
input:
(None, 9, 9, 512)
output:
(None, 9, 9, 512)
conv2d_39: Conv2D
float32
input:
(None, 9, 9, 512)
output:
(None, 9, 9, 512)
conv2d_40: Conv2D
float32
input:
(None, 9, 9, 512)
output:
(None, 9, 9, 512)
max_pooling2d_30: MaxPooling2D
float32
input:
(None, 9, 9, 512)
output:
(None, 4, 4, 512)
flatten_6: Flatten
float32
input:
(None, 4, 4, 512)
output:
(None, 8192)
dense_14: Dense
float32
input:
(None, 8192)
output:
(None, 4096)
dropout_24: Dropout
float32
input:
(None, 4096)
output:
(None, 4096)
dense_15: Dense
float32
input:
(None, 4096)
output:
(None, 4096)
dropout_25: Dropout
float32
input:
(None, 4096)
output:
(None, 4096)
dense_16: Dense
float32
input:
(None, 4096)
output:
(None, 21)

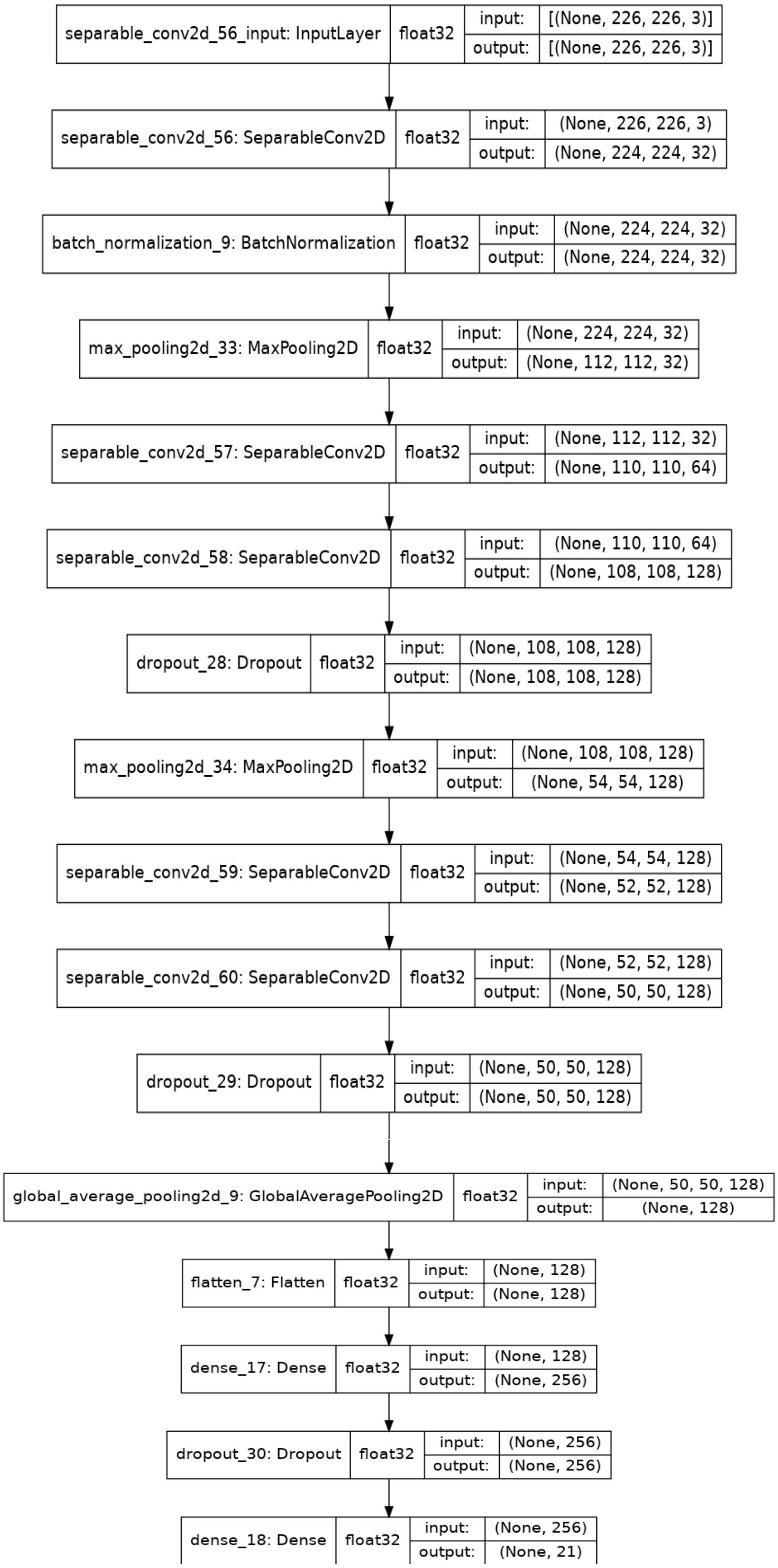


Figure 6: CNN Model 2

4.2. Dataset Specifications

This dataset has 21 different land use classes and it is generated by UC Merced for researchers [14]. Each Class contains 100 images and each image contain a size of 256X256 pixel dimensions. Each image measures 256x256 pixels (Table 1).

Table 1 Classification Used

Class	Class Indices	Dataset Distribution
Agricultural	0	350
Airplane	1	350
baseball diamond	2	350
Beach	3	350
Buildings	4	350
Chaparral	5	350
dense residential	6	350
Forest	7	350
Freeway	8	350
golf course	9	350
Harbor	10	350
Intersection	11	350
medium residential	12	350
mobile home park	13	350
Overpass	14	350
parking lot	15	350
River	16	350
Runway	17	350
sparse residential	18	350
storage tanks	19	350

5. Experimental Results

The following are the results of the CNN model. The results are shown in figure in Fig. 7 (Model accuracy vs. loss).

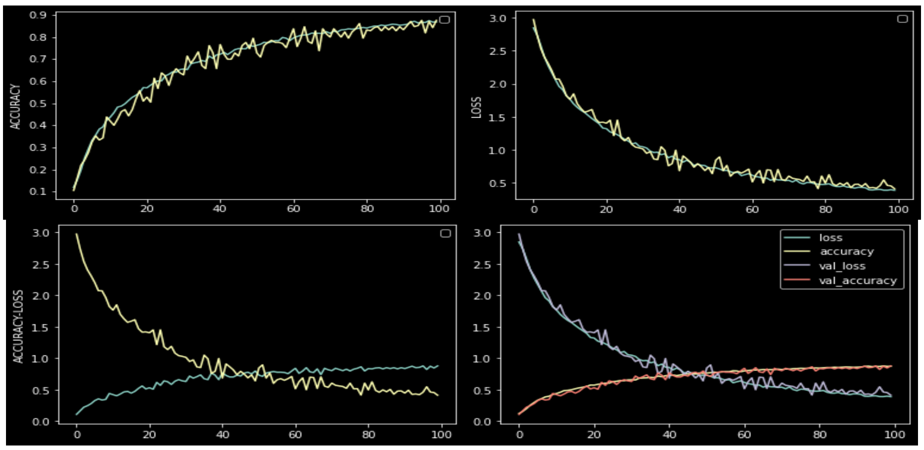


Figure 7: Test Result Proposed model

6. Application Areas

Probably the main changes to the climate, like urbanization, deforestation, and farming development, happen at the size of scenes and straightforwardly sway biological system measures (O'Neill et al. 1997 [9]; Belmaker et al. 2015 [11]). Then again, biotic communications can genuinely change scenes and produce spatial examples in that, a wonder named environment designing (Hastings et al. 2007 [10]). In this manner, map the Land Cover Land Utilization (LCLU) at the scene scale to screen and deal with these changes. Order utilizing satellite information gives an urgent beginning stage to this undertaking.

7. Conclusion

21 land-use classes can be effectively grouped utilizing medium spatial goal distantly detected information with a general arrangement exactness of 83.78%. The basic issue is to foster an excellent impenetrable surface picture. The combination of land surface temperature and LSMA derived division pictures has been shown to be compelling for refining the impenetrable surface picture, which has a by and large RMSE of 9.22% and a framework blunder of 5.68%.

8. References

The coordination of impenetrable surface and populace thickness gives another understanding to metropolitan land use grouping and the methodology created in this paper can be applied to other metropolitan conditions [1]. The authors [14] have explored bag-of-visual-words (BOVW), an innovative way to deal with land use characterization in high-goal overhead symbolism. We have considered a standard non-spatial portrayal where the frequencies yet not the areas of quantized picture highlights are utilized to separate between classes undifferentiated from how words are utilized for text record arrangement regardless of their request for event. We then , additionally think about two spatial augmentations, the set up spatial pyramid match portion which considers irrefutably the spatial course of action of the picture highlights, just as a ingenious approach which we have termed the spatial co-event part that thinks about the relative game plan. These expansions are persuaded by the significance of spatial construction in geographic information. The strategies are evaluated by utilizing a large ground truth picture dataset of twenty-one land-use classes. Not-withstanding correlations with standard methodologies, we have performed broad assessment of various designs, for example, the size of the visual word references used to determine the BOVW portrayals and the scale at which the spatial connections are thought of. We have shown that despite the fact that BOVW approaches don't really perform better compared to the best standard methodologies in general, they are addressed a hearty elective that is more successful for specific land-use classes [6][7]. We likewise show that broadening the BOVW approach with our proposed spatial co-occurrence part reliably further develops execution. This algorithm deals with different statistical analyses of geospatial databases for improving the accurateness in the recognition of the cancer of the breast which is based on the Convolutional Neural Network algorithm. From the GIS data center, the dataset is gained. It is composed of 9,109 GIS images of different locations of various cities. It is analyzed that the data can be validated, tested and then trained. Lastly, the error histogram has been figured from the dataset and thus we have come to the confusion matrix, so that the accuracy level can be projected and achieve a validated accuracy which is up to 96% to 98%.

[1] P. Anderson, B. Fernando, M. Johnson, and S. Gould. Spice: Semantic propositional image caption evaluation. In ECCV, 2016.

[2] Xinlei Chen, C. Lawrence Zitnick; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 2422-2431.

[3] Simao Herdade, Armin Kappeler, Kofi Boakye, Joao Soares: Image Captioning: Transforming Objects into Words.

[4] J. Mao, W. Xu, Y. Yang, J. Wang, and A. L. Yuille. Deep captioning with multimodal recurrent neural networks, 2015.Link (https://arxiv.org/abs/1412.6632)

[5] Sharma, H.K., Choudhury, T. & Kandwal, A. Machine learning based analytical approach for geographical analysis and prediction of Boston City crime using geospatial dataset. GeoJournal (2021). https://doi.org/10.1007/s10708-021-10485-4.
[6] Choudhury, T., Kumar, V., Nigam, D., & Mandal, B. (2016, September). Intelligent classification of lung & oral cancer through diverse data mining algorithms. In 2016 International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE) (pp. 133-138). IEEE.
[7] Khanchi, Ishu and Ahmed, Ezaz and Sharma, Hitesh Kumar, Automated Framework for Real-Time Sentiment Analysis (March 1, 2020). 5th International Conference on Next Generation Computing Technologies (NGCT-2019), Available at SSRN: https://ssrn.com/abstract=3702238 or http://dx.doi.org/10.2139/ssrn.3702238.
[8] Sharma, Hitesh Kumar, et al. "Detecting hate speech and insults on social commentary using nlp and machine learning." Int J Eng Technol Sci Res 4.12 (2017): 279-285.
[9] O'Neill, R. V., C. T. Hunsaker, K. B. Jones, K. H. Riitters, J. D. Wickham, P. M. Schwartz, I. A. Goodman, B. L. Jackson, and W. S. Baillargeon. 1997. "Monitoring Environmental Quality at the Landscape Scale: Using Landscape Indicators to Assess Biotic Diversity, Watershed Integrity, and Landscape Stability." BioScience 47 (8): 513–519. doi:10.2307/1313119.
[10] Hastings, A., J. E. Byers, J. A. Crooks, K. Cuddington, C. G. Jones, J. G. Lambrinos, T. S. Talley, and W. G. Wilson. 2007. "Ecosystem Engineering in Space and Time." Ecology Letters 10 (2): 153–164. doi:10.1111/j.1461-0248.2006.00997.x.
[11] Belmaker, J., P. Zarnetske, M.-N. Tuanmu, S. Zonneveld, S. Record, A. Strecker, and L. Beaudrot. 2015. "Empirical Evidence for the Scale Dependence of Biotic Interactions." Global Ecology and Biogeography 24 (7): 750–761. doi:10.1111/geb.12311.
[12] Mittal, V., Gupta, S., Choudhury, T. (2018). Comparative Analysis of Authentication and Access Control Protocols Against Malicious Attacks in Wireless Sensor Networks. In: Satapathy, S., Bhateja, V., Das, S. (eds) Smart Computing and Informatics . Smart Innovation, Systems and Technologies, vol 78. Springer, Singapore. https://doi.org/10.1007/978-981-10-5547-8_27.
[13] A. Agarwal, S. Gupta and T. Choudhury, "Continuous and Integrated Software Development using DevOps," 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE), 2018, pp. 290-293, doi: 10.1109/ICACCE.2018.8458052.
[14] Yi Yang and Shawn Newsam, "Bag-Of-Visual-Words and Spatial Extensions for Land-Use Classification," ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS), 2010. Shawn D. Newsam Assistant Professor and Founding Faculty Electrical Engineering & Computer Science School of Engineering University of California, Merced Voice: (209) 228-4167 Email: snewsam@ucmerced.edu Web: http://faculty.ucmerced.edu/snewsam.