Assessing Deep Learning-Based Prediction Methods for Landslide vulnerability Forecasting

Dr Shashikala Parameshwarappa [#1], Dr B.V.Dhandra [*2],

Computer Sceince and Engineering Department

Government Engineering College ,Raichur -584135

State:Karnataka, Country: India

Dr B V Dhandra

Department of Master of Computer Applications

Gulbarga University Gulbarga 58102

***Abstract*— When it comes to natural catastrophes in Himachal Pradesh, landslides are among the deadliest and most destructive. Utilizing landslide susceptibility maps for areas prone to landslides allows for proactive planning and mitigation of devastating landslide disasters. To represent the potential for landslides, we trained a CNN-DNN, a deep convolutional neural network. We trained and validated the proposed model using the training dataset (80%) and the testing dataset (20%). The training dataset contained relevant information about previous landslides, field notes, and remote sensing photos. The testing dataset included a variety of geomorphological, geological, environmental, and human activity variables. To evaluate the CNN-DNN model's predictive power, we used a number of statistics extracted through confusion matrix and error indices extracted through receiver operating characteristic (ROC) curve. To ensure a comprehensive evaluation of the CNN-DNN model, it was compared to several state-of-the-art benchmark machine learning algorithms, including logistic regression (LR), SVM, BNB, MLP, classifiers. Results showed that compared to benchmark approaches, CNN-DNN model produced more accurate predictions for landslide susceptibility mapping.**

***Keywords*— Landslides, CNN, ROC, SVM, logistic regression (LR), GNB, MLP, BNB and decision tree (DT)**

I. INTRODUCTION

Landslides cause extensive damage and human casualties, earning them a reputation as a top natural disaster threat in Himachal Pradesh. Utilizing landslide susceptibility maps for areas prone to landslides allows for proactive planning and mitigation of devastating landslide disasters.

We created a CNN-DNN, to map likelihood of landslides and tested it on an unprecedented scale in Iran's Isfahan region. Training datasets included 80% of the total data, while testing datasets made up 20%. Both sets of data included pertinent information on past landslides, as well as field recordings and remote sensing photos. Covariates in the model included various geomorphological, geological, environmental, and human activity aspects. There was an extensive battery of tests run on the CNN-DNN model's predictive power utilizing ROC curve error indices and data drawn from the confusion matrix. A number of cutting-edge benchmark machine learning techniques were used to undertake a thorough evaluation of the CNN-DNN model. With an AUC of 90.9%, IRs of 84.8%, MSE of 0.17, RMSE of 0.40, and MAPE of 0.42, the CNN-DNN model outperformed the benchmark algorithms in terms of landslide vulnerability mapping prediction accuracy. According to CNN-DNN map, province's primary Zagros trend is located in the southwest and west, creating a particularly vulnerable region. Isfahan provincial land use planners and landslide risk managers may benefit greatly from these results.

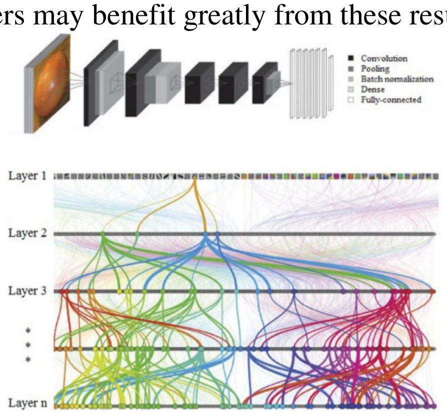

Fig-1: Architecture

Researching the efficacy of a CNN-DNN neural network in landslide risk assessment was the driving force for this investigation. Himachal Pradesh was the site of the assessment. We set the following objectives: I. What are the main factors that increase the likelihood of landslides? II. Can alternative prediction models be more accurate than the CNN-DNN model? is the CNN-DNN model capable of providing the most accurate susceptibility mapping?

We used several well-known machine learning methods to evaluate the CNN-DNN model. These methods included decision tree (DT) classifiers, multilayer perceptron's (MLP), Gaussian naïve Bayes (GNB), logistic regression (LR), support vector machines (SVM), and Bernoulli naïve Bayes (BNB).

After the many landslide "covariates" (or variables) significant to landslide occurrence in the research region were established, the geographic susceptibility to landslides was predicted using the algorithms. Subsequently, we dug more into the areas that were most vulnerable. To evaluate the prediction results, we used receiver operating characteristic curves (ROC) and confusion matrices, which include overall accuracy, precision, recall, and F1-score.

II. Related Work

This study examines the potential for landslides in the Ebantu area in the Oromia regional state in western Ethiopia (Firomsa, M., & Abay, [19] 2019). Field observations and interpretations using Google Earth revealed 92 landslides. In order to create the landslide susceptibility zonation map, data sets were first produced as layer into spatial GIS database. The geology was changed using geological map of Nekemte. Land use map was generated using digital image processing methods from the 2015 Landsat +ETM satellite. We used the statistical index approach to pinpoint potential landslide hotspots, and then we used statistical analysis to determine what elements had a role in previous landslides. We further divide the causal factor map into many categories according to their impact on mass movement, and we give each category a grade based on how much of an impact it has on slope instability. After the weight values are assigned, the map overlay procedure is carried out using Arc GIS 10.3. Lastly, the overlay approach produces a landslide danger map exhibiting different zones.

In 2019, Milevski and Dragićević published a study. This study method determines the relative importance of these aspects according to the judgment of experts. Summarizing the factor layers into harmonized raster grids yields the LS model. The last step is to use the quantiles and natural breaks method to categorize the grid model's values. About 40% of the country's landmass is at high or extremely high risk of landslides, according to the generated maps, which validate the findings using validation techniques and ROC analysis. The use of statistical approaches in a hybrid model may further enhance this strategy.

Bharati, P., [21]. To undertake scientific investigations in mountainous terrains to reduce the socio-economic impacts of landslides, landslide susceptibility zonation (LSZ) has usually been considered the proper step to take in 2020. There has been much use of combining geographic information systems (GIS) with machine learning (ML) for the purpose of making very accurate spatial predictions of landslide vulnerability. Nevertheless, there is a lack of investigation into the suitability of ML models for township level LSZ endeavours in the ML and GIS-based LSZ literatures, which leads to a lack of clarity about the appropriate selection of ML method from among several state-of-the-art approaches. This paper uses a case study to address this issue and provide a reliable method that may serve as a standard for such investigations in the future and for comparisons across other ML genres. In order to achieve this, the current endeavour has relied on four distinct supervised ML algorithms: an ANN, ELM from neural network (NN) genre, a classical ML algorithm from SVM genre, and a neuro-fuzzy system called an ELANFIS. As a case study, we decided to look at the famous hill station of Uttarakhand. Produced were thirteen landslide susceptibility maps (LSM). By using the research area's landslide inventory, we were able to assess and statistically verify the spatial performance of these maps. Out of all the LSMs tested, the one with the most agreement with the validation measures was LSM-ELANFIS-VII of the ELANFIS model, which has eleven membership functions (MF). The ELANFIS-generated LSMs not only perform satisfactorily during validation, but they also show a distinct geomorphological fingerprint and practical dispersion of landslide vulnerability classes, which bodes well for the eventual transition from GIS-based LSZ to ensemble neuro-fuzzy ML models.

In order to forecast the likelihood of landslides in Yanshan County, China, this work presents 4 heterogeneous ensemble-learning methods. (Fang, Z[22], 2021). The goal of these methods is to provide trustworthy results while avoiding model selection issues by strategically combining several state-of-the-art classifiers from CNN, recurrent neural networks, support vector machines, and logistic regression. Three primary procedures make up the research. The first stage involves creating a geographical database with 380 past landslide sites and 16 landslide conditioning variables. Step two involves selecting grid cells that correspond to landslide and non-slide sites in the research region at random to form training & testing databases, respectively. Building suggested heterogeneous ensemble-learning approaches for landslide vulnerability mapping is the last stage. While comparing with aforementioned individual classifiers, suggested ensemble-learning approaches provide statistically superior prediction accuracy. With an precision of 80.70 percent, the blending ensemble-learning technique outperforms all of the other ensemble-learning approaches.

III. Proposed System

Researchers evaluated danger of landslides in research area by using the proposed CNN-DNN approach. OA and ROC were deciding factors in projected model's output. In a graphical form, the ROC curve shows how a binary classifier system's diagnostic capability changes in response to changes in its discriminating threshold. So, area under curve (OA) and AUC from the ROC plot show how accurate the classifiers are.

IV. Methodology

1) Multiple phases made up the research. Initial steps included conducting a ground survey to document and estimate the frequency of landslides in the region under investigation. Additionally, by combining CNN’s feature extraction with the DNN's classification skills, we were able to pinpoint potentially very accurate regions that were quite vulnerable. The next step was to put the model through its paces using ROC curve, error models, and performance criteria.

2) Here, we assess whether the suggested CNN-DNN approach is appropriate for creating comprehensive landslide vulnerability maps. Using a battery of relevant statistical metrics, we compared the CNN-DNN's performance to that of many high-quality benchmark methods. To train the CNN-DNN, we used 15 landslide covariates, which we categorized into 4 sets. The data was prepared for landslide susceptibility analysis by normalizing all covariate layers before entering them into the model. While the CNN extracted features, the DNN classified pixels as either highly susceptible or lowly susceptible. You may find the study's hyperparameters in Table 2. Improving precision of machine learning model predictions is typical goal of optimizing the fitting process using hyperparameters. Finding the sweet spot for the evaluation values is the goal of hyperparameter selection. Because some optimizers produce more precise results than others61, we utilized a variety of them to tweak the hyperparameters. The evaluations in the offered research were based on the grid search method. When it came time to training and testing machine learning models, we used the hyperparameters producing best outcomes as of accuracy.

By combining the confusion matrix with the algorithm performance matrix, we were able to assess how well the suggested technique worked.

You can see how well a prediction algorithm is doing by looking at its performance matrix, which is table which displays projected values and includes sensitivity, 1-specificity characteristics. In order to evaluate the performance of a classifier with that of reliable outside opinions, classification tasks make use of metrics like true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The proportion of pertinent instances (TP) out of all the recovered instances is known as precision, which is also known as the positive predictive value.

Precision = TPTP +FP (1)

A measure of sensitivity, recall is the sum of all relevant instance fractions.

Recall/Sensitivity = TPTP+FN (2)

Precision and recall, thus, are relevant-ness-based metrics. One way to get the false-positive rate is to subtract the specificity from the total. Specificity is defined as:

Specificity = TNTN+FP (3)

On unbalanced datasets, accuracy could be a deceptive measure. Making all the values in a prediction set negative yields an accuracy score of 0.95 in a scenario where there are 95 positive values and 5 negative ones. But when the two values are near to one another, F1-score—harmonic mean of recall and precision—provides an approximation of 2 averages and, more generally, harmonic mean.

F1-score=2□Precision□ Recall Precision + Recall (4)

Total precision (TP) + total noise (TN) divided by total sample size (n) is the overall accuracy (OA), which is a measure of how likely it is that a test would properly categorize a person.

Accuracy = TP+TNTP+TN+FP+FN (5)

With that in mind, OA is also the mean of the two metrics 'sensitivity' and 'specificity'. You can learn a lot about the reliability of a classifier by using the performance matrix.

V. System Architecture

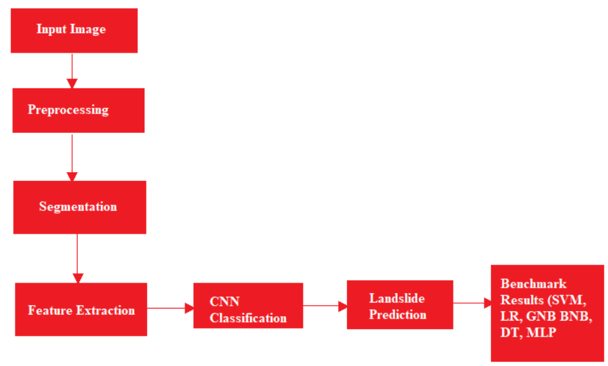

Figure 2: System Architecture

VI. Implementation

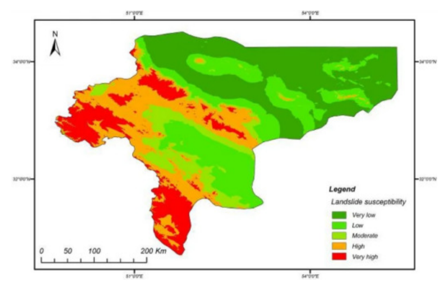

Figure 3: Landslide susceptibility map for the proposed model

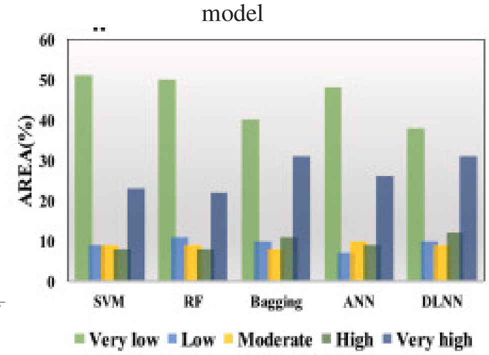

Fig-4: Area wise percentage of susceptibility groups using various models.

VII. Conclusion

One of the trickiest parts of geo-hazard assessment is making maps of landslide vulnerabilities. Analysis using state-of-the-art deep learning methods may be useful in this setting. To determine the likelihood of landslides in Iran's Isfahan province, we used a new CNN-DNN prediction model. Data on landslides in the past, which included a variety of land sliding kinds, and a number of potential causes were all input into the model. Outperforming several benchmark techniques, the suggested CNN-DNN model achieved very high accuracy.

References

1. Colesanti, C. & Wasowski, J. Investigating landslides with space-borne Synthetic Aperture Radar (SAR) interferometry. Eng. Geol. 88, 173–199. https://doi.org/10.1016/j.enggeo.2006.09.013 (2006).
2. Highland, L. & Bobrowsky, P. T. The Landslide Handbook: A Guide to Understanding Landslides (US Geological Survey Reston, 2008).
3. Chen, Z. et al. Landslide research in China. Q. J. Eng. Geol. Hydrogeol. 49, 279–285. https://doi.org/10.1144/qjegh2016-100 (2016).
4. Tang, H., Wasowski, J. & Juang, C. H. Geohazards in the three Gorges Reservoir Area, China-Lessons learned from decades of research. Eng. Geol. 261, 105267. https://doi.org/10.1016/j.enggeo.2019.105267 (2019).
5. Wasowski, J. et al. Recurrent rock avalanches progressively dismantle a mountain ridge in Beichuan County, Sichuan, most recently in the 2008 Wenchuan earthquake. Geomorphology 374, 107492. https://doi.org/10.1016/j.geomorph.2020.107492 (2021).
6. Azarafza, M., Ghazifard, A., Akgün, H. & Asghari-Kaljahi, E. Landslide susceptibility assessment of South Pars Special Zone, southwest Iran. Environ. Earth Sci. 77, 805. https://doi.org/10.1007/s12665-018-7978-1 (2018).
7. Cascini, L. Applicability of landslide susceptibility and hazard zoning at different scales. Eng. Geol. 102, 164–177. https://doi.org/10.1016/j.enggeo.2008.03.016 (2008).
8. Pham, V. D., Nguyen, Q.-H., Nguyen, H.-D., Pham, V.-M. & Bui, Q.-T. Convolutional neural network: Optimised moth flame algorithm for shallow landslide susceptible analysis. IEEE Access 8, 32727–32736. https://doi.org/10.1109/ACCESS.2020.2973415 (2020).
9. Abella, E. A. C. & Van Westen, C. J. Qualitative landslide susceptibility assessment by multicriteria analysis: a case study from San Antonio del Sur, Guantánamo, Cuba. Geomorphology 94, 453–466. https://doi.org/10.1016/j.geomorph.2006.10.038 (2008).
10. Lee, S. & Choi, J. Landslide susceptibility mapping using GIS and the weight-of-evidence model. Int. J. Georg. Inf. Sci. 18, 789–814. https://doi.org/10.1080/13658810410001702003 (2004).
11. Manzo, G., Tofani, V., Segoni, S., Battistini, A. & Catani, F. GIS techniques for regional-scale landslide susceptibility assessment: The Sicily (Italy) case study. Int. J. Geogr. Inf. Sci. 27, 1433–1452. https://doi.org/10.1080/13658816.2012.693614 (2013).
12. Feizizadeh, B. & Blaschke, T. An uncertainty and sensitivity analysis approach for GIS-based multicriteria landslide susceptibility mapping. Int. J. Geogr. Inf. Sci. 28, 610–638. https://doi.org/10.1080/13658816.2013.869821 (2014).
13. Firomsa, M. & Abay, A. Landslide assessment and susceptibility zonation in Ebantu district of Oromia region, western Ethiopia. Bull. Eng. Geol. Environ. 78, 4229–4239. https://doi.org/10.1007/s10064-018-1398-z (2019).
14. Milevski, I. & Dragićević, S. Landslides susceptibility zonation of the territory of north macedonia using analytical hierarchy process approach. Contrib. Sect. Nat. Math. Biotechn. Sci. 40, 115–126. https://doi.org/10.20903/csnmbs.masa.2019.40.1.136 (2019).
15. Peethambaran, B., Anbalagan, R., Kanungo, D., Goswami, A. & Shihabudheen, K. A comparative evaluation of supervised machine learning algorithms for township level landslide susceptibility zonation in parts of Indian Himalayas. CATENA 195, 104751. https://doi.org/10.1016/j.catena.2020.104751 (2020).
16. Fang, Z., Wang, Y., Peng, L. & Hong, H. A comparative study of heterogeneous ensemble-learning techniques for landslide susceptibility mapping. Int. J. Geogr. Inf. Sci. 35, 321–347. https://doi.org/10.1080/13658816.2020.1808897 (2021).
17. Yan, Y. et al. Volunteered geographic information research in the first decade: A narrative review of selected journal articles in GI Science. Int. J. Geogr. Inf. Sci. 34, 1765–1791. https://doi.org/10.1080/13658816.2020.1730848 (2020).

18. Rahman, M. et al. Development of flood hazard map and emergency relief operation system using hydrodynamic modeling and machine learning algorithm. J. Clean. Prod. 133, 127594. https://doi.org/10.1016/j.jclepro.2021.127594(2021) (2021).

19. Rahman, M. et al. Flood susceptibility assessment in Bangladesh using machine learning and multi-criteria decision analysis. Earth Syst. Environ. 3, 585–601. https://doi.org/10.1007/s41748-019-00123-y (2019).
20. Dewan A.M., Hazards, risk, and vulnerability. In: Floods in a Megacity, 35–74. https://doi.org/10.1007/978-94-007-5875-9_2 (2013).

21. Adnan, M. S. G. et al. Improving spatial agreement in machine learning-based landslide susceptibility mapping. Remote Sens. 12, 3347. https://doi.org/10.3390/rs12203347 (2020).

22. Zêzere, J., Pereira, S., Melo, R., Oliveira, S. & Garcia, R. A. Mapping landslide susceptibility using data-driven methods. Sci. Total Environ. 589, 250–267. https://doi.org/10.1016/j.scitotenv.2017.02.188 (2017).