4. Image Classification using Deep Learning Features (Deep CNNs)

We separated the data for each dataset into a training and a test set at a ratio of 80/20 to train and evaluate deep CNNs on the proposed innovative and introduced existing datasets. We made sure the split was applied to each class separately. Except for the typical vertical flipping, we used no data augmentation in our studies. We fine-tuned ResNet-50 and ResNet-101 [He et al.(2016), He et al.(2016)] CNN models for each dataset, which were pre-trained on the ILSVRC-2012 image 205 classification dataset. We have trained the last layer using a learning rate of 0.001 in all evaluations. Then, using a modest learning rate of 10^{-6} to 10^{-5}, we fine-tuned the entire network. To evaluate the performance of the different CNN models on the analyzed datasets, we computed the total classification accuracy.

5. Results and Discussion

The EuroSAT dataset was utilized as the basis for this investigation. Here, all 10 classes were taken into account. Each class has between 2000 and 3000 labeled images. As a result, the experiment used a total of 27,000 images. The features are extracted from the images using HLAC. Various kernel functions, such as the linear kernel, polynomial kernel, and RBF kernels, were used. The grid search algorithm was used to select the various kernel parameters to improve the results. For multi-class classification, SVM employs a one-vs-all classification technique. The accuracy and F1-score criteria are used to assess the performance. We have used pre-trained models ResNet-50 and ResNet-101 were trained on the EuroSAT dataset. To test the model's performance on the EuroSAT dataset, we evaluated the total classification accuracy.

Classifier	Method used	Performance Evaluation Criteria	Classification Accuracy (%) with Training/Testing split 80/20		
SVM (benchmark results)	SIFT (BoVW)	Accuracy (%)	58.55 (k=10)	67.22 (k=100	70.05 (k=500)
SVM (results of proposed approach)	HLAC	Accuracy (%)	87.84 (RBF)	88.35 (Lin)	89.35 (Poly)
		F1-score	85.37 (RBF)	86.14 (Lin)	86.26 (Poly)

Table 1. Comparison of classification accuracy of handcrafted features with benchmark

Table 1 shows the classification accuracy and F1-score for 10 classes in the EuroSAT dataset. The outcomes were compared to the benchmarks established in [Helber et al. (2019)]. With a train and test splitting ratio of 80/20 and the SIFT (BoVW) technique, the optimum classification accuracy with SVM was 70.05 % in the benchmark. Using HLAC feature extraction, the SVM classifier was able to efficiently distinguish between separate classes, improving accuracy to 89.35 % with a train and test splitting ratio of 80/20. The accuracy has improved by 19.3% with the benchmark.

To improve accuracy, many kernel methods were utilized, including the linear kernel, polynomial kernel, and RBF kernel. The RBF kernel had the lowest accuracy (87.84%), the linear kernel was in the middle (88.35%), and the polynomial kernel had the best accuracy (89.35%). The minimum F1-score for the RBF kernel was 85.37, and the maximum F1-score for the polynomial kernel was 86.26. Figure 5 additionally depicts the accuracy of the reported result vs. the benchmark as a graph.

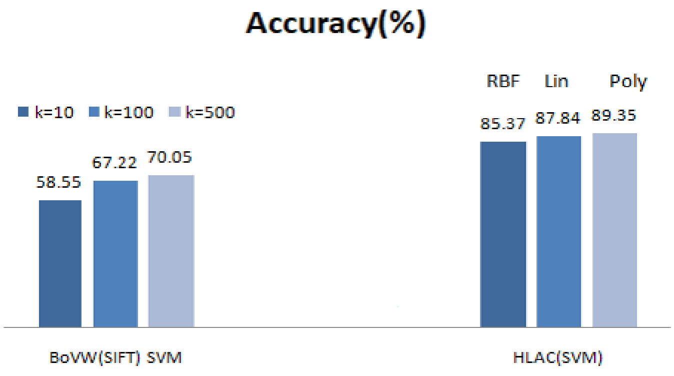

Fig.5. Comparison of classification accuracy

The classification results for the various deep CNN models are listed in Table 2. The authors in [Helber et al. (2019)] employed a pre-trained ResNet-50 model [He et al. (2016), He et al. (2016)] that was trained on the EuroSAT dataset and attained a classification accuracy of 98.57 %. We utilized a pre-trained ResNet-50 and ResNet-101 model and trained on the EuroSAT dataset and attained an accuracy of 98.69%. On the EuroSAT dataset [Helber et al. (2018), Helber et al. (2019)], the deep CNNs produce state-of-the-art results, outperforming earlier results by 0.12% [15, 20]. Figure 6 depicts the confusion matrix of this best-performing network, which allows for a class-level evaluation. The classifier occasionally confuses the River and Highway classes, the AnnualCrop and PermanentCrop classes, as well as the PermanentCrop and HerbaceousVegetation classes even if this occurs infrequently.

Method	Accuracy (%)
ResNet-50 (benchmark)	98.57
ResNet-50 (our result)	98.50
ResNet-101 (our result)	98.69

Table 2. Comparison of classification accuracy of deep learning models with benchmark

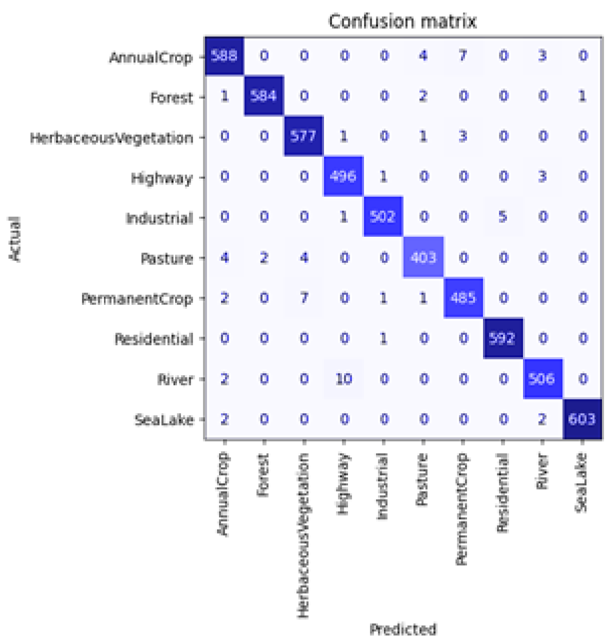

Fig.6. Confusion matrix of ResNet-101 model

6. Conclusion and Future Work

In this study, we compared various models based on deep learning, transfer learning, and models trained on hand-crafted features. HLAC features were used with the SVM classifier to classify satellite images. SVM provided good classification accuracy using the HLAC features extraction methods. The increased accuracy of the SVM classifier showed that it could effectively distinguish between distinct classes when features were extracted using HLAC. Using other kernels and altering their settings improves the outcomes much more. Using a polynomial kernel, the accuracy was improved by 19.3 % for ten classes.

The ResNet-50 and ResNet-101 models were used to evaluate the performance of deep learning features, and the performance was enhanced by 0.12% using the ResNet-101 model compared with benchmark results.

In future work, the order of the HLAC features might be increased to improve the accuracy. Furthermore, by evaluating all of the spectral bands of the satellite images, the results can be enhanced. To improve classification, handcrafted features can be merged with deep learning models.

References

[1] Battineni, G., Chintalapudi, N., Amenta, F. (2019). Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM). Informatics in Medicine Unlocked, 16, 100200.
[2] Bulugu,I.; Ye, Z.; Banzi,J. (2017): Higher-order local autocorrelation feature extraction methodology for hand gestures recognition, In 2017 2nd International Conference on Multimedia and Image Processing (ICMIP), pp. 83-87.
[3] Cao, J., Cholakkal, H., Anwer, R. M., Khan, F. S., Pang, Y., & Shao, L. (2020). D2det: Towards high quality object detection and instance segmentation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 11485-11494).
[4] Cortes, C.; Vapnik, V. (1995): Support-vector networks, Machine learning 20, no. 3, pp. 273-297.
[5] Csurka, G.; Dance C.R.; Humenberger M. (20.18): From handcrafted to deep local features, arXiv preprint arXiv: 1807.10254, 2018.
[6] Dalal, N. ; Triggs, B. (2005): Histograms of oriented gradients for human detection, In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), vol. 1, pp. 886-893.
[7] Deepak, R. S., Uday, S. I., & Malathi, D. (2017). Machine learning approach in stock market prediction. International Journal of Pure and Applied Mathematics, 115(8), 71-77.
[8] Dewangkoro, H. I.; Arymurthy, A. M. (2021): Land Use and Land Cover Classification Using CNN, SVM, and Channel Squeeze & Spatial Excitation Block, In IOP Conference Series: Earth and Environmental Science, IOP Publishing, vol. 704, no. 1, pp. 012048.
[9] Haralick, R. M.; Shanmugam, K.; and Dinstein, H. (1973): Textural features for image classification, IEEE Transactions on systems, man, and cybernetics, vol.3, no.6, pp. 610-621.
[10] Hayashi, H. ; Uchida, S. (2018): A trainable multiplication layer for auto-correlation and co-occurrence extraction," In: Asian Conference on Computer Vision, Springer, pp. 414–430.
[11] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
[12] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Identity mappings in deep residual networks. In European conference on computer vision (pp. 630-645). Springer, Cham.
[13] Helber, P., Bischke, B., Dengel, A., & Borth, D. (2018, July). Introducing eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, pp. 204-207.
[14] Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12(7), 2217-2226.
[15] Humeau-Heurtier, A. (2019). Texture feature extraction methods: A survey. IEEE Access, 7, 8975-9000.
[16] Ishihara,T.; Otsu, N.(2004): Gesture recognition using auto-regressive coefficients of higher-order local auto-correlation features," In: Sixth IEEE International Conference on Automatic Face and Gesture Recognition Proceedings, pp. 583–588.
[17] Jivane, N.J. ; Soundrapandiyan, R. (2017): Enhancement of an algorithm for oil tank detection in satellite images, International Journal of Intelligent Engineering System, vol. 10, no.3, pp. 298–305.
[18] Kobayashi, T. ; Otsu, N. (2009): "Color image feature extraction using color index local auto-correlations," In: 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1057–1060.
[19] Kupidura, P. (2019): The comparison of different methods of texture analysis for their efficacy for land use classification in satellite imagery, Remote Sensing, vol. 11, issue 10, pp. 1233.
[20] Kurita T.; Hayamizu, S. (1998): Gesture recognition using hlac features of parcor images and hmm based recognizer, In: Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, pp. 422–427.
[21] Kussul, N., Shelestov, A., Lavreniuk, M., Butko, I., & Skakun, S. (2016). Deep learning approach for large scale land cover mapping based on remote sensing data fusion. In 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 198-201). IEEE.
[22] Li, K., Wan, G., Cheng, G., Meng, L., & Han, J. (2020). Object detection in optical remote sensing images: A survey and a new benchmark. ISPRS Journal of Photogrammetry and Remote Sensing, 159, 296-307.
[23] Liu, Q.; Basu S.; Ganguly, S.; Mukhopadhyay, S.; DiBiano, R.; Karki, M.; Nemani, R. (2020): Deepsat v2: feature augmented convolutional neural nets for satellite image classification, Remote Sensing Letters, vol.11, no. 2, pp. 156–165.
[24] Lowe, D.G. (2004): Distinctive image features from scale-invariant keypoints, International journal of computer vision vol. 60 no.), pp. 91–110.
[25] Noi, P. T.; Kappas, M. (2018): Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery, Sensors, vol .18(1).
[26] Oh, S. I., & Kang, H. B. (2017). Object detection and classification by decision-level fusion for intelligent vehicle systems. Sensors, 17(1), 207.
[27] Otsu, N.; Kurita, T. (1988): A new scheme for practical flexible and intelligent vision systems, In: MVA, pp. 431–435.
[28] Quinlan , J.R. (1986): Induction of decision trees, Machine learning, vol. 1, no.1, pp. 81–106.

[29] Rana V.K.; Suryanarayana, T.M.V.(2020): Performance evaluation of mle, rf and svm classification algorithms for watershed scale land use/land cover mapping using sentinel 2 bands, Remote Sensing Applications: Society and Environment, vol 19, pp. 100351.
[30] Russakovsky, O; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; Berg, A. C.; Fei-Fei, L. (2015): ImageNet Large Scale Visual Recognition Challenge, International Journal of Computer Vision (IJCV), vol. 115, no. 3, pp. 211–252.
[31] Tianyu, Z.; Zhenjiang, M.; Jianhu, Z. (2018): Combining cnn with hand-crafted features for image classification, In: 2018 14th IEEE International Conference on Signal Processing (ICSP 2018), pp. 554–557.
[32] Uehara, K. ; Sakanashi, H.; Nosato, H.; Murakawa, M.; Miyamoto, H.; Nakamura, R. (2017): Object detection of satellite images using multi-channel higher-order local autocorrelation, In: 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, pp. 1339–1344.
[33] Vapnik, V.(1998): Statistical learning theory New York, NY: Wiley.
[34] Varma, M.K.S. ; Rao, N.K. (2017): Satellite image classification using genetic algorithm based on svm classifier, International Journal of Control Theory and Applications, vol. 10, no.26, pg. 12–20.
[35] Zhang, X.; Cui, J. ; Wang, W.; Lin, C. (2017): A study for texture feature extraction of high-resolution satellite images based on a direction measure and gray level co-occurrence matrix fusion algorithm, Sensors vol. 17(7), pg. 1474.
[36] Zhong, Y., Fei, F., Liu, Y., Zhao, B., Jiao, H., & Zhang, L. (2017). SatCNN: satellite image dataset classification using agile convolutional neural networks. Remote sensing letters, 8(2), 136-145.

Authors Profile

Khushbu R. Joshi is a research scholar in the department of Sankalchand Patel University Visnagar, India. She is also working as a lecturer in the Electronics and communication department of LDRP Institute of Technology and Research, Gandhinagar, India. She obtained her BE degree in Electronics and Communication Engineering in 2007 from Saurashtra University, Rajkot and ME degree in Electronics and Communication (CSE) in 2009 from Gujarat University, Ahmedabad. Her research interests include Machine Learning, Deep Learning and Computer Vision System. She is a member of IETE, India.

Manish I. Patel is working as an Assistant Professor in Electronics and Communication Engineering Department since November 2019. He has more than 15 years of teaching experience. He obtained his BE degree in Electronics and Communication Engineering in 2003 from North Gujarat University, Patan and MTech degree in Electronics and Communication (VLSI Design) in 2010 from Nirma University, Ahmedabad. Dr Manish obtained his PhD in 2017 from Gujarat Technological University, Ahmedabad. His areas of interest include signal and image processing, machine learning and VLSI design. He has published more than fifteen papers in journals and conference proceedings. He has guided PG students. He is a life member of ISTE and IETE, India.