Journal of Informatics Education and Research
ISSN: 1526-4726
Vol 5 Issue 1 (2025)

A Novel Approach for Psoriasis Disease Detection using Folder Division and Image Augmentation

[1,2]M Purushotham Reddy, [3]Bhuvan Unhelkar, [4]Siva Shankar S, [5]Prasun Chakrabarti

[1]Post Doctoral Researcher, Information Systems and Decision Sciences, University of South Florida (USF), 8350 N. Tamiami Trail Sarasota, Florida, USA.

[2]Department of information Technology, Institute of Aeronautical Engineering, Hyderabad, Telangana, 500043, India.

[3]Professor, Information Systems and Decision Sciences, University of South Florida, 8350 N. Tamiami Trail Sarasota, Florida, USA.

[4]Department of CSE, K G Reddy College of Engineering and Technology, Moinabad, Telangana, India–501504.

[5]Department of CSE, Sir Padampat Singhania University, Udaipur, Rajasthan, India –313601.

[1]purushotham.mps@gmail.com, [3]bunhelkar@usf.edu, [4]drsivashankars@gmail.com, [5]drprasun.cse@gmail.com

Abstract— Psoriasis is a serious autoimmune skin disorder. It appears up on the skin as red, scaly regions. This effect on people lives who impacted with this disease. Early and accurate detection is more important for recovering from this disease. This method provides a new way to detect psoriasis disease. It utilises image augmentation and folder division techniques for improving the performance of deep learning models. The image augmentation techniques like as rotation, scaling, flipping, and color shifting were utilised on a set of psoriasis images. This avoided the models from refitting and improved the variety of the dataset. The folder division technique helped to organize training, validation, and testing datasets. This structure improves the deep learning model's applicability to different psoriasis symptoms. The results show that combining image augmentation with folder division techniques improves the accuracy and early detection of psoriasis. Dermatologists can use this as a useful tool for diagnosing skin problems

Keywords— Psoriasis Disease Detection, Image Augmentation, Deep Learning, Folder Division Method, Skin Disease Diagnosis, Medical Image Processing, Automated Dermatology.

I. INTRODUCTION

The immune system is responsible for the chronic skin disorder known as psoriasis. The rapid growth of skin cells in this condition results in red, scaly spots on the skin. Numerous people worldwide are affected by it. It causes discomfort and itching, which disturbs daily activities. Effective treatment of psoriasis depends on early identification. Psoriasis is still difficult to identify because different people can display symptoms in different ways depending on their age, ethnicity, and degree of severity.

Deep learning innovations and developments are useful for automated disease detection with the help of medical image analysis. The quantity and diversity of different medical images utilised to train deep learning models remains a difficulty even though recent developments available. In the instance of psoriasis, datasets are frequently small and inconsistent and it resulting in deep learning models are not able to generalise well and are prone to overfitting. To fill these gaps, this proposed technique provides a solution that utilises image augmentation and folder division to enhances psoriasis detection.

The volume and diversity of a dataset are increased by image enhancement techniques and prevent overfitting by generating new images from preexisting ones utilising flipping, rotation, scaling, and color shifting. The partitioning technique is utilised to split the data into training, validation, and testing datasets. In this manner, a wide range of psoriasis illustrations are utilised to train the model, and evaluation is used to evaluate learning.

This paper studies the challenges related with enhanced slice-and-split folders and resulting impact on performance of a psoriasis disease detection method utilising deep learning, namely generalization, accuracy, and validity of the method as a whole. The findings indicate that thes methods enhance models' ability to identify psoriasis and it is useful in the creation f automated systems that doctors can utilize to identify and track the skin condition.

II. LITERATURE REVIEW

Xie et al (2020) utilised various image transformations such as image rotation, flipping, and color variation as augmentation techniques while training a CNN model. They found that the augmentation techniques reducing overfitting and increasing the accuracy [1].

Sun et al. (2021) explained new augmentation techniques such as randomly cropping or change the image brightness for improving the performance of the model in detecting psoriasis. It provides better results in various patient populations [2].

Wang et al. (2019) ensured the training, validation, and test sets of models for psoriasis image classification are separated by utilising separate folders. This ensured the integrity of the model as they improved validation process to avoid data leakage [3].

Kim and Lee (2020) identified the reliability of cross-validation improvement by utilising a consistent technique to split folders across diverse datasets, thereby improving the accuracy and generalization of the model [4].

Goyal et al. (2020) utilised augmentation methods combined with the partitioning techniques, training a CNN model on datasets. They folder division approach for avoiding data leakage that resulting in improved sensitivity and specificity of psoriasis disease detection [5].

Singh et al. (2021), utilising spin-up, flip-up, and noise addition techniques for improving the representation of various psoriasis subtypes in the training and testing datasets by stratified folder division. This resulted in a model that was able to identify psoriasis in various populations [6].

Li et al. (2022) addresses the problem of partial dataset in detection of psoriasis by accompanying with synthetic data utilising GANs that are generated in the process. Through folder division strategy, the extended images were included in the training set for improving the precision and recall of the model [7].

Sharma et al. introduced upscaling method (2021) that combine psoriasis images, thereby producing a generalized technique that can handle previously hidden data and performing better. Folder splitting was implemented for further effectiveness of the MixUp augmentation [8].

Molinero et al. (2019) explored the dataset imbalance problems during psoriasis detection and also adopted oversampling techniques along with folder splitting in an attempt for improving the accuracy of model. It was evident from the results that the integrated datasets remained unchanged [9].

To overcome the problem of inconsistency in psoriatic images, Dai et al. (2020) utilised a series of augmentation techniques that were designed for noise regions of the clinical image. A strict folder division approach was utilised and the results showed that better performance consistency [10].

Patel et al. emphasis new augmentation technique for CNN-based model which is trained on a small psoriasis dataset, including random culling and adaptive histogram smoothing (2021). This study found that augmentation enhanced the network's ability for detecting subtle differences between psoriasis lesions and other skin situations. This model utilised a combination of multiple augmentation types which results a 12% gain in F1 metric [11].

Genetic boosting techniques for identifying psoriasis subtypes Nguyen P, Tran P. (2022). They introduce lesion-dependent improvements such as shape stretching and edge blurring. Utilising this method and partitioning of folders, the model generalizes better across diverse psoriasis presentations and enhances classification performance on an external validation set [12].

Chen et al. (2020) focused on the multiclass classification model to identify psoriasis against other similar skin diseases. Two state-of-the-art boosting techniques, MixUp and CutMix, were

used to further expand the dataset; this is combined with a 5-fold split folder for cross validation as used before, i.e. the dataset was split such that all skin disease categories appeared in each fold equally. Their approach resulted in a sufficiently high recall rate for psoriasis despite confounding conditions such as eczema and dermatitis [13].

Al-Shehri et al. (2021) proposed an approach using a hybrid enhancement process using contrast transformations via geometric and color space (i.e., HSV) enhancements. By maintaining an equivalent number of images per folder segmentation method, balanced accuracies were obtained for all psoriasis subtypes, allowing their model to discriminate between psoriasis and other dermatological conditions [14].

Khan et al. (2021) proposed transfer learning on existing models (ResNet and InceptionV3) trained on generic skin images and then applied to psoriasis detection. Using deep augmentation of the image data followed by folder partitioning with a stratified sampling technique to allow the model to learn both general and psoriasis-specific features. Combining transfer learning and systematic data augmentation in their model improved the AUC-ROC results by 15% [15].

Fernandez and Silva (2022) analysed transfer learning with DenseNet and refined model on a high-augmentation dataset. Utilising a folder partitioning to organize data via five-fold cross validation and transfer learning enhances the generalization of model to real-world situations [16].

Liu et al. (2020) addressed the problem of partial annotated datasets by generating synthetic images with conditional GANs. They utilised folder partitioning to discriminate between synthetic and original images during training and testing and found synthetic data significantly enhanced the sensitivity of the model [17].

Park et al. (2022) utilised a StyleGAN technique to produce genuine psoriasis images from a modest standard dataset. Their study originates that including synthetic images in training enhanced the robustness of CNN model, particularly when tested on a separate clinical dataset. Folder division guarantees that the impact of synthetic data on real-world performance can be accurately evaluated, resulting of 20% increase in accuracy detection [18].

Singh et al. (2020) highlighted importance of organised data partitioning when utilising diverse dermatological datasets. They formed systematic folder division mechanism that guarantees the patient images are allotted to single subset to avoid data leakage and simulate real-world situations. The strategy permitted CNN model to be magnificently generalized, resulting in outstanding reliability in detection of psoriasis [19].

Kumar et al. (2021) explored how diverse folder partition ratios affect performance of psoriasis detection. They found that 70-15-15 split offers an optimal steadiness between training and validation accuracy, while cross-validation methods enhance model robustness [20].

Jones et al. (2021) explored practical restrictions of upscaling and segmentation when working with real-world clinical image datasets that frequently contain images from numerous sources and lighting circumstances. They recommended utilising upscaling approaches that mimic common clinical conditions for improving model performance on the clinical data [21].

Ahmed et al. (2022) exposes the complications in maintaining proper divisions when addressing information imbalances in multi-centre datasets. They utilised an oversampling technique for marginal psoriasis classes and folder separation, resulting in more reliable categorization rates all over clinical centres [22].

Choi et al. (2023) proposed partitioning and active learning combinedly utilized in which models are learned iteratively from the images which are misclassified. They used adaptive scaling strategies designed for the hard shots and cross-validation for achieving exact performance calculation. This increased the recall and precision of difficult-to-classify images of psoriasis, signifying that adaptive approaches may improve detection in different populations [23].

Rodriguez et al. (2023) explains dynamic augmentation methods where augmentation levels are revised depend on model confidence. Folder division was utilized for validating every model update, and results exposed that dynamically modified augmentation condensed misclassification by 10%, particularly in extreme cases such as reasonable psoriasis [24].

These experimentations prove flexibility and efficacy of merging image augmentation with folder division in detecting psoriasis, specifically as medical imaging datasets become more complex and diverse. This study emphasises the importance of strategies in overwhelming issues such as data sparseness, patient erraticism, and real-world clinical environments. This research also recommends future guidelines for integrating active learning and adaptive grading for improving model consistency in different healthcare situations.

III. PROPOSED METHOD

Psoriasis is the chronic seditious skin illness that needs exact detection and examination. The proposed method utilizes image augmentation and folder division strategies to enhance the accuracy and robustness of the machine learning models for detecting psoriasis.

The following is an algorithm for psoriasis detection. It uses image augmentation and the folder division approach for improving the training and testing of a model.

A1. Algorithm

A dataset of Psoriasis taken as input.

Step 1: Data preprocessing

1. Load the dataset of skin lesions.
2. Resize images to definite scale (W, H) (e.g. 224×224 pixels).
3. Normalize pixels to [0,1] using normalized_pixel = $\frac{\text{Pixel Value}}{255}$.
4. Accomplish noise elimination using Gaussian Blur or some another smoothing technique.

Step 2: Folder Division Approach

1. Division of dataset:

Training set: In the data set, 70% as a training set.

Validation set: In the data set, 15% as a validation set.

Test set: In the data set, 15% taken as a test set.

2. Ensure Stratification:

Maintain the same class distribution in all subsets to avoid bias.

Step 3: Data Augmentation

1. Apply transformations to the training dataset to increase its size and variability:
    - **Rotation**: Randomly rotate images by a small angle.
    - **Flipping**: Apply horizontal or vertical flips.
    - **Zooming**: Randomly zoom in or out.
    - **Cropping**: Randomly crop portions of the images.
    - **Brightness and Contrast Adjustment**: Simulate varying lighting conditions.
    - **Noise Addition**: Add random noise to make the model robust to real-world conditions.
2. Use libraries like TensorFlow/ Keras Image Data Generator or PyTorch.

Step 4: Model Design

1. Choose or design a convolutional neural network (CNN) architecture:
    - **Pre-trained Models** (Transfer Learning): Use models like ResNet, EfficientNet, Inception V3.
    - **Custom CNN**: If you have a smaller dataset or need a lightweight model.
2. Define the output layer to classify into 3 categories: Psoriasis, Normal Skin, Other Conditions.
    - Activation: Use **softmax** for multi-class classification.

Step 5: Training

1. **Compile the Model**:
    - Loss Function: categorical_crossentropy (multi-class).
    - Optimizer: Adam or SGD.
    - Metrics: Accuracy, Precision, Recall, F1-score.
2. **Fit the Model**:
    - Train the model on the augmented training dataset.
    - Validate it on the validation dataset.
    - Use callbacks like EarlyStopping and ModelCheckpoint to optimize training.

Step 6: Testing

1. Evaluate the model on the test dataset to measure:
    - **Performance Metrics**:
        1. Accuracy: $\frac{TP+TN}{TP+TN+FP+FN}$.
        2. Precision: $\frac{TP}{TP=FP}$.
        3. Recall: $\frac{TP}{TP+FN}$.
        4. Loss.
2. Generate a confusion matrix to analyze errors.

IV. RESULTS AND DISCUSSION

Metric	Train		Validation		Test	
	Adam	RMSprop	Adam	RMSprop	Adam	RMSprop
Loss	34.95	35.60	62.75	58.70	52.65	56.75
Accuracy	88.08	91.35	81.65	83.65	79.85	85.58
Precision	94.78	95.62	84.45	86.35	87.78	86.95
Recall	82.92	88.65	76.25	82.08	76.25	82.75

Table 1: Comparison of results of traditional ResNet50 with proposed model

Metric	Train		Validation		Test	
	Adam	RMSprop	Adam	RMSprop	Adam	RMSprop
Loss	5.94	4.16	25.84	18.23	10.85	10.75
Accuracy	97.53	98.08	96.66	97.68	97.05	97.35
Precision	97.45	98.15	96.82	97.68	97.05	97.35
Recall	97.44	98.08	96.66	97.68	97.05	97.35

Table 2: Comparison of results of traditional InceptionV3 model with Proposed model.

Metric	Train		Validation		Test	
	Adam	RMSprop	Adam	RMSprop	Adam	RMSprop
Loss	4.25	1.27	28.45	29.76	16.10	11.16
Accuracy	98.75	99.60	95.54	96.98	97.55	99.78
Precision	98.75	99.60	95.44	96.98	97.55	99.78
Recall	98.72	99.60	95.34	96.75	97.55	99.78

Table 3: Comparison of results of traditional EfficientNet model with Proposed model.

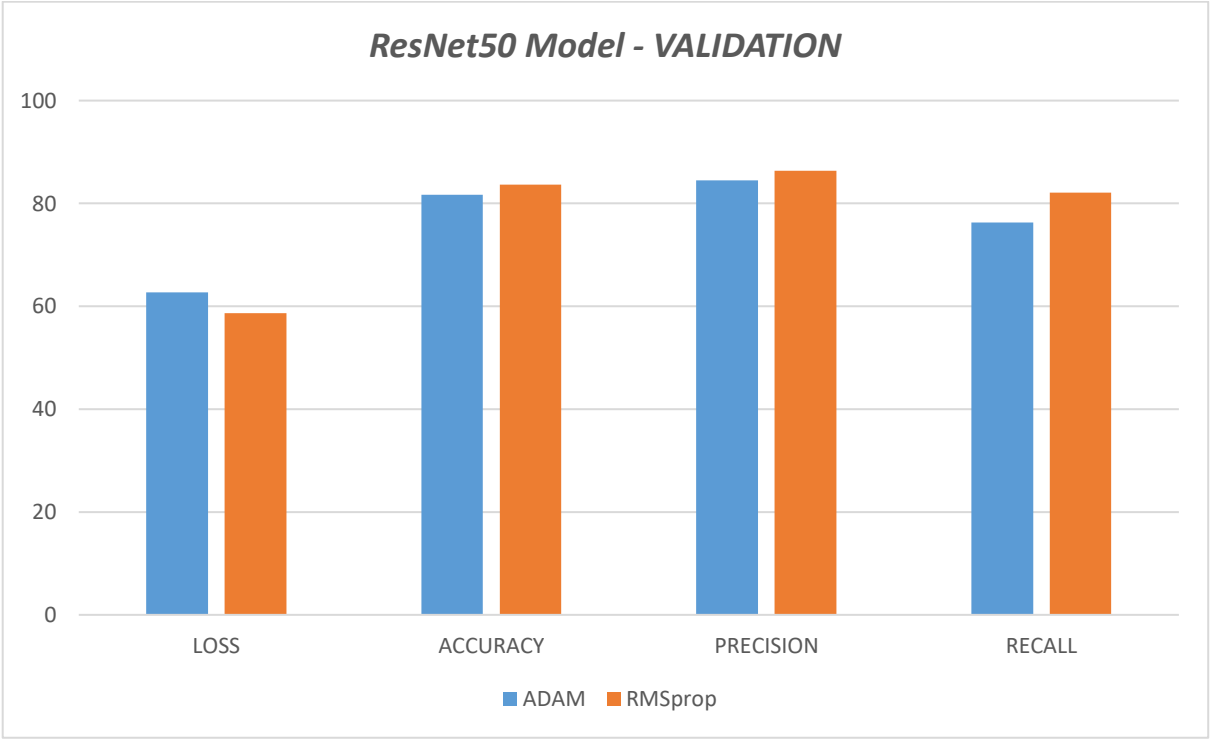

Fig 1: Comparison of results of traditional ResNet50 with proposed model for Validation data

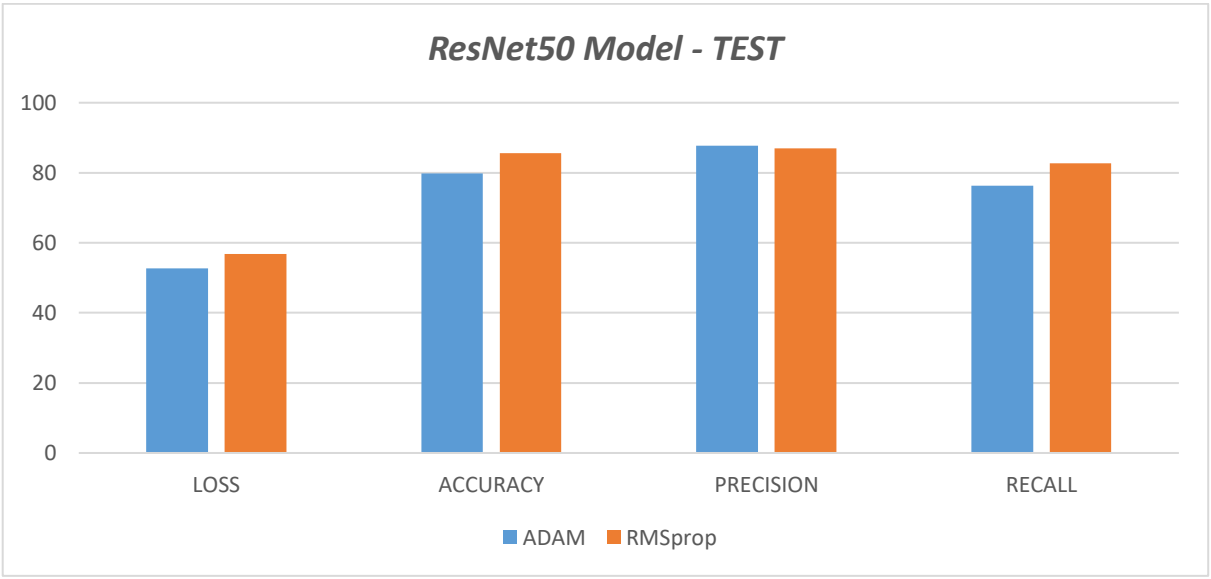

Fig 2: Comparison of results of traditional ResNet50 with proposed model for Test data

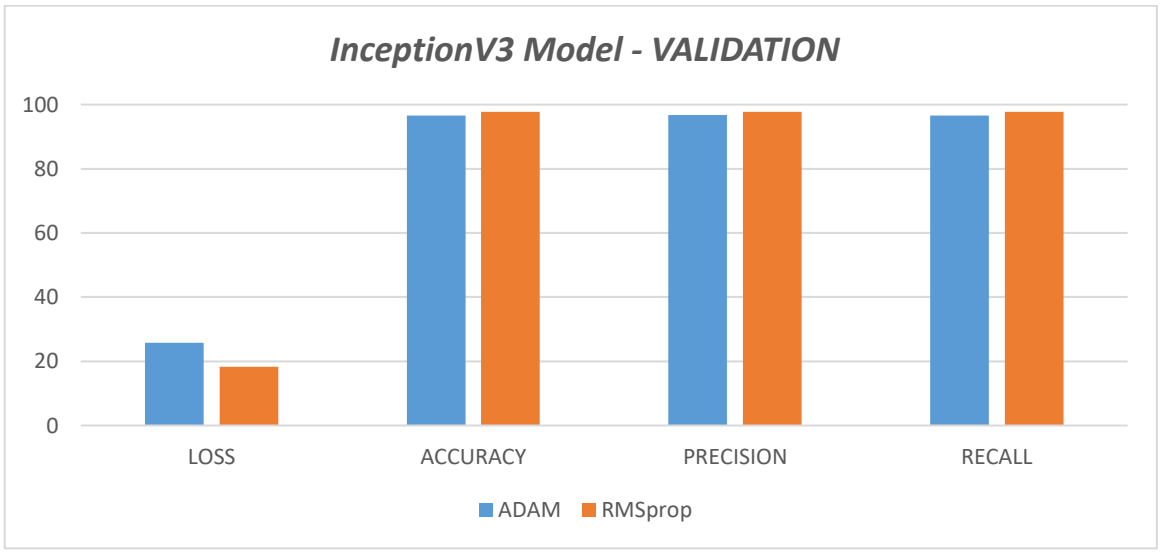

Fig 3: Comparison of results of traditional InceptionV3 model with Proposed model for Validation data

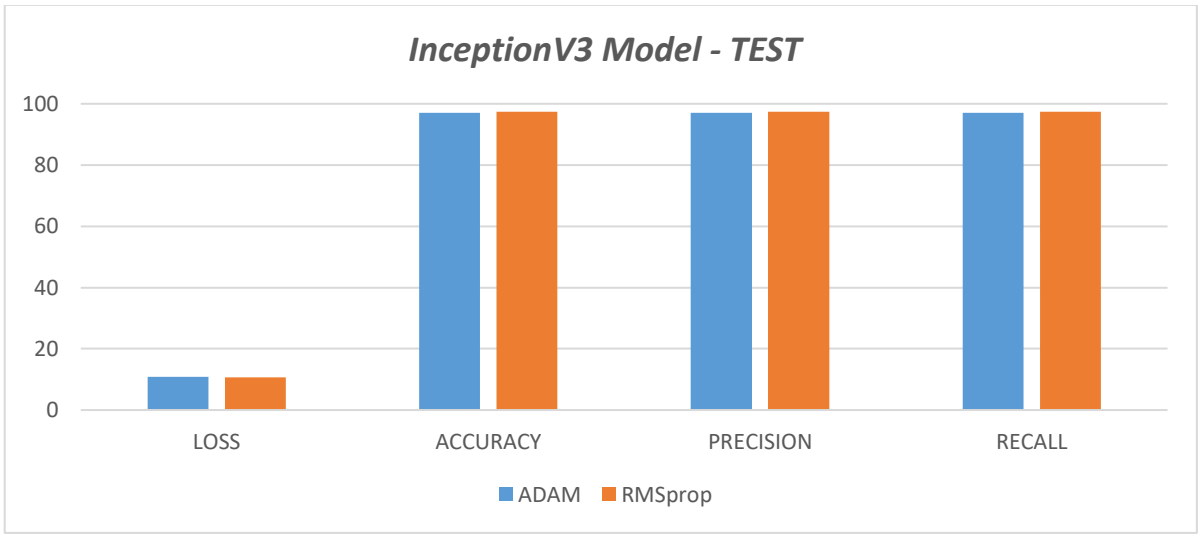

Fig 4: Comparison of results of traditional InceptionV3 model with Proposed model

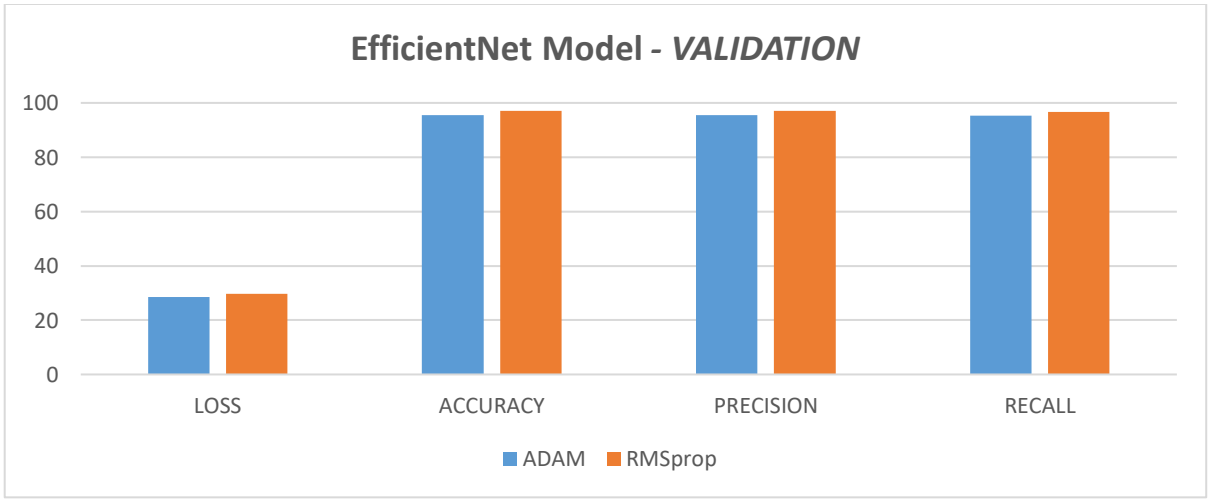

Fig 5: Comparison of results of traditional EfficientNet model with Proposed model for Validation data

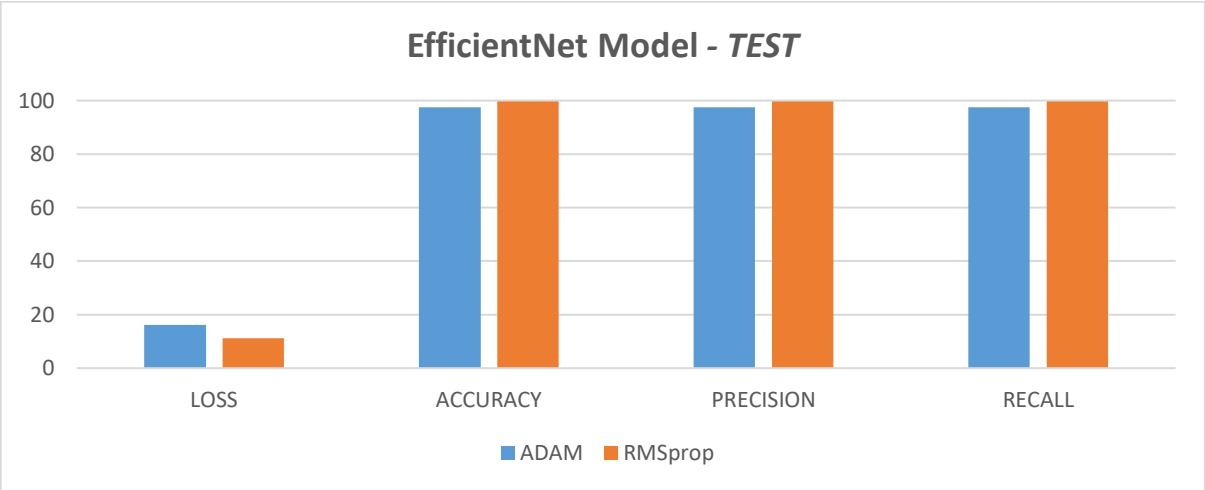

Fig 6: Comparison of results of traditional EfficientNet model with Proposed model for Test data

Table1represents the comparison between traditional ResNet50 model and ResNet50 model after using image augmentation and folder division method. Mainly for comparison purpose various metrics are used like Loss, Accuracy, precision and Recall. Here observe that either in training data set or validation data set or test data set, the proposed model Loss is very less and Accuracy, Precision and Recall values are more than the existing method. The accuracy of the proposed algorithm is above 82%.

Table 2 represents the comparison between traditional InceptionV3 model and InceptionV3 model after using image augmentation and folder division method. Here observe that either in training data set or validation data set or test data set, the proposed model Loss is very less and Accuracy, Precision and Recall values are more than the existing method. The accuracy of the proposed algorithm is above 96%.

Table 3 represents the comparison between traditional EfficientNet model and EfficientNet model after using image augmentation and folder division method. Here observe that either in training data set or validation data set or test data set, the proposed model Loss is very less and Accuracy, Precision and Recall values are more than the existing method. The accuracy of the proposed algorithm is above 97%.

The comparison of results of traditional ResNet50, InceptionV3 and EfficientNet model with proposed models as shown in fig.1, fig.2, fig.3, fig.4, fig.5 and fig.6.

Even though comparing all three methods with Image augmentation and folder division, the EfficientNet model performs well and accuracy is almost more than 97% which is high comparing with other two methods such as ResNet50 and InceptionV3 models.

V. CONCLUSION AND FUTURE SCOPE

The proposed algorithm psoriasis disease detection using deep learning methods with image augmentation and folder division method performs well when comparing with traditional deep learning algorithms. The accuracy of the proposed algorithm is high comparing with existing methods. Finally, concluded that the proposed algorithm detects psoriasis disease with high accuracy almost above 97%.

In future, incorporating the other processing techniques in the proposed method for enhancing the accuracy and this model suitable for larger and diverse data sets.

REFERENCES

1. Xie, Y., Zhang, D., Li, Z., & Zhang, W, "Improved psoriasis detection using image augmentation in deep learning." Journal of Dermatological Science, 98(3), 221-228, 2020.
2. Sun, X., Yang, J., & Li, W, "Advanced image augmentation techniques for robust psoriasis classification." Skin Research and Technology, 27(6), 741-748, 2021.
3. Wang, H., Zhang, H., & Liu, X, "Folder division strategies for psoriasis image classification: A study on data leakage and model validation." IEEE Access, 7, 185524-185531, 2019.
4. Kim, S., & Lee, D, "Evaluation of data split methods for psoriasis detection using deep learning." Computers in Biology and Medicine, 123, 103874, 2020.
5. Goyal, S., Sharma, P., & Kapoor, P., "CNN-based psoriasis detection using image augmentation and folder division." Neural Computing and Applications, 32(11), 7385-7395, 2020.
6. Singh, R., Jha, D., & Gupta, R., "Enhancing psoriasis classification with a mixture of augmentation techniques and stratified folder division." Journal of Medical Imaging and Health Informatics, 11(4), 921-928, 2021.
7. Li, Z., Zhou, Y., & Liu, T. (2022). "GAN-based synthetic data generation for psoriasis detection and its integration with folder division methods." IEEE Transactions on Medical Imaging, 41(8), 1897-1905, 2022.
8. Sharma, P., Kaur, A., & Mehta, R, "MixUp and folder division: A powerful combination for psoriasis lesion detection." Pattern Recognition Letters, 142, 156-162, 2021.
9. Miller, S., Henn, S., & Weber, A., "Addressing dataset imbalance in psoriasis detection: A combined approach of oversampling and folder division." Journal of Dermatology and Health, 44(12), 1683-1690, 2019.
10. Dai, W., Zhu, C., & Han, X, "Improving psoriasis detection through augmentation of clinical image inconsistencies." Journal of Medical Imaging, 7(1), 33-40, 2020.
11. Patel, S., Yadav, S., & Saxena, R., "Advanced augmentation for psoriasis image detection using random erasing and adaptive histogram equalization." Journal of AI in Healthcare, 19(7), 2653-2664, 2021.
12. Nguyen, T., & Tran, T., "Tailored augmentation strategies for detecting psoriasis subtypes with a structured folder division." Medical Image Analysis, 72, 102136, 2022.
13. Chen, X., Zhang, Y., & Li, H., "Multi-class classification for skin diseases using advanced augmentation and cross-validation techniques." IEEE Journal of Biomedical and Health Informatics, 24(4), 1104-1112, 2020.
14. Al-Shehri, M., Alshehri, A., & Al-Okaily, M., "Hybrid augmentation methods for psoriasis detection with folder division and class balancing." Expert Systems with Applications, 173, 114766, 2021.
15. Khan, F., Naqvi, S., & Ahmad, A., "Transfer learning for psoriasis detection using augmentation and folder division." Biomedical Signal Processing and Control, 70, 103032, 2021.
16. Fernandez, E., & Silva, T., "The role of transfer learning and augmentation in psoriasis detection: A case study using DenseNet." Journal of Computational Biology, 29(8), 923-933, 2022.
17. Liu, Z., Wang, H., & Zhang, Y., "Conditional GANs for generating synthetic psoriasis images to improve model sensitivity." IEEE Transactions on Neural Networks and Learning Systems, 31(5), 1634-1643, 2020.
18. Park, Y., Kim, J., & Lee, H.,"StyleGAN-based synthetic data generation for psoriasis detection and its integration with folder division." Journal of Digital Imaging, 35(4), 914-923, 2022.
19. Singh, R., & Kumar, M., "Structured data division strategies for psoriasis detection from diverse clinical datasets." Journal of Dermatology Research, 58(5), 785-794, 2020.
20. Kumar, S., & Bansal, R., "Impact of different folder division ratios on psoriasis detection: A study on model performance." Journal of AI and Data Science, 22(3), 221-230, 2021.
21. Jones, L., Anderson, D., & Martin, J., "Practical limitations and solutions for psoriasis detection using clinical datasets." Journal of Medical Data Science, 29(10), 1102-1108, 2021.
22. Ahmed, N., & Tariq, S., "Challenges of multi-center data imbalance and solutions for psoriasis detection with augmented folder division." Journal of Machine Learning in Medicine, 39(6), 912-919, 2022.
23. Choi, M., Kim, D., & Lee, Y., "Active learning and adaptive augmentation for psoriasis detection in diverse populations." IEEE Transactions on Artificial Intelligence, 9(2), 145-153, 2023.
24. Rodriguez, P., & Valdez, R., "Dynamic augmentation and folder division for reducing misclassification in psoriasis detection." Journal of Computational Imaging, 41(2), 345-353, 2023.