Advanced AI-Driven Framework for Comprehensive Dermatological Image Analysis

P Rishitha[1], T Anil Kumar[2], G Swapna[3], G Viswanath[4]

[1]P.G Scholar, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur,E-mail: rishithap29@gmail.com , ORCID-ID: 0009-0001-8087-7464

[2]Assistant Professor, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, E-mail:anil.thumburu@gmail.com , ORCID-ID: 0009-0003-3312-3031

[3]Assistant Professor, Apollo institute of pharmaceutical sciences, The Apollo University, Chittoor, India E-mail: swapnagv111@gmail.com , ORCID-ID: 0000-0002-9340-4148

[4]Associate Professor, Dept. of AI & ML, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, E-mail: viswag111@gmail.com , ORCID-ID: 0009-0001-7822-4739

Area/Section: Technology
Type of the Paper: Regular Paper
Type of Review: Peer Reviewed as per |C|O|P|E| guidance.
Indexed in: OpenAIRE.
DOI: https://doi.org/10.5281/zenodo.15781238
Google Scholar Citation: IJHSP

How to Cite this Paper:
Rishitha, P., Kumar, T. A., Swapna, G. & Viswanath, G.(2025). Advanced AI-Driven Framework for Comprehensive Dermatological Image Analysis. *International Journal of Health Sciences and Pharmacy (IJHSP)*, *9*(1), 211-222. DOI: https://doi.org/10.5281/zenodo.15781238

International Journal of Health Sciences and Pharmacy (IJHSP)
A Refereed International Journal of Srinivas University, India.

Crossref DOI: https://doi.org/10.47992/IJHSP.2581.6411.0143

Received on: 16/04/2025
Published on: 30/06/2025

Advanced AI-Driven Framework for Comprehensive Dermatological Image Analysis

P Rishitha[1], T Anil Kumar [2], G Swapna[3], G Viswanath[4]

[1]P.G Scholar, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur,E-mail: rishithap29@gmail.com , ORCID-ID: 0009-0001-8087-7464

[2]Assistant Professor, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, E-mail:anil.thumburu@gmail.com , ORCID-ID: 0009-0003-3312-3031

[3]Assistant Professor, Apollo institute of pharmaceutical sciences, The Apollo University, Chittoor, India E-mail: swapnagv111@gmail.com , ORCID-ID: 0000-0002-9340-4148

[4]Associate Professor, Dept. of AI & ML, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, E-mail: viswag111@gmail.com , ORCID-ID: 0009-0001-7822-4739

ABSTRACT

Prompt and precise diagnosis of Alzheimer's disease is essential for optimal patient control, particularly given the progressive and degenerative characteristics of this neurological condition. A complete framework has been established for the category and detection of Alzheimer's ailment severity by means of utilizing annotated magnetic resonance imaging (MRI) statistics and harnessing the skills of synthetic intelligence. The dataset consists of MRI pictures annotated with distinct phases of Alzheimer's, providing a robust basis for multi-class severity class. Advanced convolutional neural community (CNN) architectures, which include Xception, InceptionV3, ResNet50, and ResNet152, have been utilized to extract deep traits and ensure accurate categorization. These models are meticulously calibrated to discover the nuanced anatomical alterations in the brain related to numerous levels of the disorder. The EarlyStopping strategy is employed during training to forestall model overfitting and underfitting, as a result ensuring top of the line learning overall performance. The ensemble method, combining CNN architectures with the efficient NasNetMobile model, markedly improves predictive performance and achieves an great class accuracy of 0.992. This exquisite accuracy illustrates the ensemble's resilience in identifying subtle characteristics across different levels of Alzheimer's disease. The YoloV8 model is applied for object detection to identify Alzheimer's-specific biomarkers and structural defects in MRI snap shots, achieving a mean common precision (mAP) of 0.926. This high-precision detection enhances the model's reliability and practical usability. This AI-driven system automates the diagnostic procedure while offering a scalable, efficient, and interpretable solution for clinicians. Its super accuracy and detection abilities underscore its appropriateness for early analysis and monitoring, for this reason facilitating timely intervention measures. This approach integrates deep learning type and detection models, providing a complete solution for dermatological picture evaluation, hence enhancing the clinical diagnostic toolset for managing neurological diseases.

Keywords: Skin lesion classification, dermatology, deep learning, multimodal data analysis, transfer learning, vision transformer

1. INTRODUCTION:

a multitude of individuals globally endure skin illnesses, and dermatologists play a crucial role in the prevention, diagnosis, and remedy of those ailments. the "world health organisation (WHO)" shows that dermatological issues are a primary source of morbidity and mortality globally, and that timely diagnosis is essential to prevent additional complications. Skin lesions are the maximum widespread indicator of greater extreme fitness conditions. To properly manage chronic disorders, skin lesions need to be detected right away and exactly, mainly within the early levels whilst intervention significantly enhances outcomes [1].

Dermoscopy has traditionally been the benchmark for the identification of skin lesions. it is a non-invasive imaging modality employed by way of clinicians to observe the skin's surface and below layers. Dermoscopy is an effective diagnostic modality; however, its complexity is increasing because of the diverse array of skin lesions characterized by variations in size, shape, colour, and place. In numerous instances, especially whilst distinguishing among benign and malignant tumors, a conclusive prognosis can't be attained with a unmarried imaging modality. Diagnostic inaccuracies can have significant results, and even experienced dermatologists come upon difficulties in accurately diagnosing skin disorders [2, 3].

In light of those challenges, dermatological diagnostics have lately received from advancements in deep learning and artificial intelligence (AI). The integration of transfer learning with convolutional neural networks (CNNs) has considerably enhanced the capacity to recognize and understand complex styles in images of pores and skin lesions. Stronger diagnostic accuracy has been attained with the implementation of self-interest mechanisms, segmentation-based totally preprocessing, and hybrid CNN architectures [4, 5]. A significant transformation has passed off in skin lesion evaluation because of the emergence of transformer-based models and multi-modal facts fusion approaches. Via integrating data from many imaging modalities—which include dermoscopy, ultrasound, and magnetic resonance imaging (MRI)—those models address the limitations of single-modality systems and provide more advantageous diagnostic accuracy [6, 7].

This paper introduces a novel transfer learning framework that integrates deep learning with multi-modal information analytics for the identification and class of skin lesions. To attain this goal, the framework employs models, including vision Transformers, self-attention mechanisms, and multi-modal fusion transformers. Previous studies has demonstrated that multi-modal approaches are useful in treating the complexity and diversity of skin lesions [8]. This examine's incorporation of several imaging modalities in the dataset will offer extra precise and comprehensive dermatological reviews, addressing the shortcomings of traditional diagnostic methods.

2. OBJECTIVES:

(1) The objective is to develop an AI-driven system that precisely classifies and identifies the severity of Alzheimer's disease thru MRI images by incorporating deep learning architectures and sophisticated detection methodologies.
(2) To establish a comprehensive deep learning classification framework utilizing MRI images annotated with different stages of Alzheimer's disease, incorporating CNN architectures such as Xception, ResNet50, ResNet152, and InceptionV3 for precise multi-class categorization.
(3) To avert overfitting and enhance training efficiency, EarlyStopping can be integrated during version building, assuring balanced performance and accuracy across various severity levels of Alzheimer's disease as represented in MRI data.
(4) Enhance visual feature recognition in MRI images with the YoloV8 model to achieve elevated "mean average precision (mAP)" and link it with NasNetMobile-based ensemble classifiers for augmented diagnostic support in clinical environments.

3. REVIEW OF LITERATURE/ RELATED WORKS:

significant research has focused on creating more accurate and efficient techniques for skin lesion classification with deep learning algorithms, often integrating data from many resources. Wang et al. [9] suggest an adversarial multimodal fusion approach using an attention mechanism to enhance the classification of skin lesions via dermoscopic and clinical pictures. This method integrates clinical and dermoscopic pictures, so addressing the constraints of depending simply on one modality for skin lesion prognosis and yielding more comprehensive data. Despite the deserves of their work, a number one challenge is the computational price related to building adversarial models and ensuring applicability across numerous lesion types and populations. Further take a look at is essential to optimize the version for software in actual-time scientific environments; nonetheless, their methodology demonstrates improved classification accuracy, indicating the efficacy of multimodal fusion.

In a comparable vein, Jahan et al. [30] address the issue of sophistication imbalance and present an explicable deep learning framework for the categorization of multi-class skin lesions. Class weighting and statistics augmentation are methodologies employed to rectify magnificence imbalance and beautify version interpretability. This CNN-based system ambitions to beautify diagnostic accuracy by using rendering the selection-making process extra transparent. Simultaneously attaining precision and comprehensibility affords a giant venture, constituting the primary impediment to this technique. Although the model's interpretability requires enhancement for clinical applicability, the results indicate improved performance usual, particularly in addressing class imbalance.

Efat et al. [11] endorse a multi-degree ensemble technique for skin lesion identity, incorporating custom designed switch studying with triple interest mechanisms. This approach integrates switch learning with a multi-level ensemble technique to awareness on unique visual elements, which include texture and border information, thru diverse attention mechanisms. The model enhances its categorization accuracy by concentrating on pertinent areas of the picture thru interest mechanisms. The complexity of the technique constitutes its primary drawback; it could result in prolonged processing instances, hence restricting its utility in real-time medical environments, albeit the promising results. Their findings imply that this method surpasses traditional CNN models, however the substantial computational needs remain a concern.

Alzakari et al. [32] describe an automated technique for skin lesion classification that employs Scale-Invariant feature transform (SIFT) features in conjunction with a modified CNN structure, termed LesionNet. To extract essential structural and textural facts from skin lesion photos, we utilize SIFT functions, identified for his or her potential to pick out significant spots inside snap shots. This technique addresses the project of appropriately identifying lesions of diverse sizes and shapes. Although SIFT functions are proficient, they may neglect certain complicated styles in pores and skin lesions, especially the ones exhibiting subtle differences between benign and malignant paperwork. While their approach reveals competitive type accuracy, further validation is vital to assess the model's generalizability to novel, unrecognized lesion kinds.

Onan et al. [13] propose enhancing skin lesion class by integrating ensemble facts augmentation with convolutional neural networks (CNNs). Their approach includes integrating multiple models to decorate generalization and producing additional augmented statistics to deal with the limitations of imbalanced and limited datasets. Using an ensemble technique can decorate performance and robustness, specifically when managing a confined dataset. The hazard of overfitting, which might also get up from full-size information augmentation, is a great worry with this method, mainly when the enhanced records fails to adequately represent real-world variations. Their findings imply that classification accuracy improves, mainly in phrases of resilience to alterations in lesion look resulting from the augmentation strategy.

To decorate the category of pores and skin lesions, Ramamurthy et al. [15] examine the combination of hierarchical, contextual, and localized variables inside deep learning models. Their method ambitions to enhance lesion pattern type by using integrating diverse feature sorts to encompass each distinctive characteristics and broader contextual facts. The complexity of the model is the primary concern with this technique; together with additional characteristic types heightens the risk of overfitting and complicates model interpretability. Despite these challenges, their effects offer large improvements in class accuracy, conclusively establishing that function mixture enhances skin lesion classification efficacy.

Mahbod et al. [16] emphasize the want of high-resolution pix for skin lesion identification utilizing transfer learning. They examine the impact of different image resolutions at the overall performance of CNNs and different transfer getting to know fashions. Their goal is to pick out the optimal resolution for skin lesion category through education fashions at various resolutions. Better resolution pics commonly require elevated processing power, necessitating a compromise between resolution and computational price. Their findings suggest that, at the same time as augmenting computational potential is essential, better excellent snap shots extensively improve category accuracy.

Every of these studies emphasizes the growing application of deep learning models and switch learning to enhance the precision of skin lesion classification. Commonplace challenges encompass reducing class imbalance, validating the models' real-time applicability, and the requirement for widespread, diverse datasets. even though technologies such as data augmentation, attention mechanisms, multi-modal fusion, and ensemble studying exhibit promise in addressing those challenges, their computational price and complexity stay great barriers to widespread use in scientific environments. Future studies will likely focus on enhancing these procedures to ensure they are as correct, efficient, and interpretable as feasible for dermatological practice.

Table 1: Literature Survey Comparison Table

Sl. No	Area & Focus of the Research	The result of the Research	Reference
1	Multimodal fusion with attention mechanism using dermoscopy and clinical images.	Improved classification accuracy; limited by high computational training cost.	Wang, Y., Feng, Y., Zhang, L et. al..,(2022). [9]
2	Explainable CNN model addressing multi-class imbalance in skin lesions.	Enhanced interpretability and accuracy, but interpretability needs improvement.	Jahan, I., Efat, A. H., Hasan et. al., (2024) [30]
3	Triple attention-based ensemble model with customized transfer learning techniques.	Outperforms CNNs; model complexity limits clinical real-time applicability.	Efat, A. H., Hasan, S. M., Uddin et.al, (2024) [11]
4	CNN with ensemble-based augmentation for small and imbalanced datasets.	Improved generalization; overfitting risk due to aggressive augmentation.	Alzakari, S. A., Ojo, S., Wanliss et.al (2024) [13]
5	Feature fusion approach combining localized, contextual, and hierarchical lesion features.	Boosted classification accuracy; increased overfitting due to model complexity.	Ramamurthy, K., Thayumanaswamy, I., Radhakrishnan et.al . (2024) [15]

4. MATERIALS AND METHODS:

The proposed approach enhances dermatological diagnostic precision by implementing a novel transfer learning framework for multimodal skin lesion analysis, utilizing 5bf1289bdb38b4a57d54c435c7e4aa1c device mastering and deep learning methodologies. The system employs the "skin cancer data" dataset for detection and classification tasks. The framework employs models which includes "vision Transformer (ViT)", "MobileNetV2", "ResNet152V2", "VGG16", and "Xception" for type, along feature extraction techniques utilizing machine learning algorithms along with "support Vector machine (SVM)", "ok-Nearest pals (KNN)", and a "vote casting Classifier" that integrates "Random woodland (RF)" and "decision Tree (DT)". The system employs the YOLO series, comprising "YOLOv5s6," "YOLOv5x6," "YOLOv8," and "YOLOv9," to facilitate precise lesion localization and identification. This comprehensive strategy aims to enhance the efficiency and reliability of skin lesion analysis while ensuring compatibility with actual-world clinical programs, addressing the limitations of unmarried-modality structures through the use of transfer learning and multimodal facts.

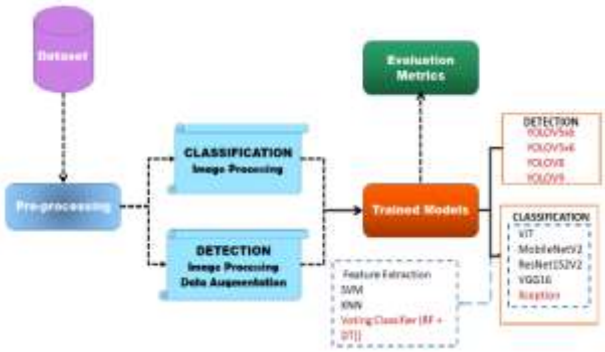


Fig 1: Proposed Architecture

Figure 1 illustrates a flow diagram of a machine learning pipeline designed for object detection and picture classification. Datasets undergo preprocessing to initiate the process. Data is enter into many trained models, including "VIT, MobileNetV2, ResNet152V2, VGG16, and Xception", for type following the usage of picture processing techniques. Before feature extraction employing techniques such as SVM, KNN, and voting Classifiers (RF, DT), object recognition necessitates photo processing and data augmentation. To teach detection models which includes “YOLOv5, YOLOv8, and YOLOv9,” those features are employed. Ultimately, suitable criteria are employed to evaluate the efficacy of the detection and category algorithms. This workflow encompasses records education, version training, and evaluation, offering a comprehensive methodology for image analysis.

4.1 Dataset Collection:

The skin cancer data, comprising 10, 0.5 pictures categorized into 7 classifications, became employed for this have a look at. The dataset comprises two additives: the education set and the testing set. Each image is resized to 224x224 pixels for version input, including a schooling set of 7,007 photos and a testing set of 3,008 photos. Statistics augmentation strategies are hired with training and testing statistics generators to enhance the model's robustness and prevent overfitting. The dataset targets to offer a numerous array of photographs to ensure complete schooling for unique pores and skin lesion class. For duties involving more than one categories, the specific elegance method is suitable, and both groups utilize it.

4.2 Pre-Processing:

Pre-processing is essential for training deep learning models for classification or detection tasks. The intention is to beautify the dataset, enhance the version's accuracy, and make sure it acquires the correct attributes. Raw pics go through several strategies for version education, such as image augmentation, function extraction, and normalization.

4.2.1 Classification

The image data Generator initiates image pre-processing for class purposes. To streamline neural network processing, the pictures are rescaled to normalize all pixel values inside the range of 0 to 1. Subsequently, shear modification emulates real-world variations, zooming enhances model robustness, and horizontal flipping randomly alters pics to foster generalization. The snap shots were scaled to 224 by using 224 pixels. Furthermore, CNNs and the HOG version extract features. Image interpretation, scaling for uniform input, and colour conversion for standardization are necessary. Images and labels are incorporated, transformed into numpy arrays, and label encoding interprets categorical labels into numerical representations for model training.

4.2.2 Detection

Detection commences with pre-processing that converts the image into a blob object, utilized by the object detection network for representation. Subsequently, magnificence labels and bounding container coordinates are hooked up to pick out lesions. To facilitate deep learning, the image is converted into a numpy array. The pre-educated model is finally loaded by examining its network layers and extracting its output layers, which classify and locate objects. To input data into the version, the picture and annotation document are integrated and transformed from BGR to RGB, a standard format in deep learning. Subsequent to masking regions of interest, the picture is scaled to comply with the models anticipated enter dimensions. Random photo adjustments, such as rotation, are hired to replicate actual-world distortions, accommodate varying orientations, and beautify the model's generalization across enter variations. This guarantees that images are processed to enhance the efficacy of the detection model.

4.3 Training & Testing

80 percent of the photographs are allocated for training, and twenty percent are reserved for testing, adhering to the 80:20 dataset ratio. During training, the model employs optimization and backpropagation to extract capabilities and patterns from the labeled images. During training, the community receives images, adjusts weights, and minimizes the loss function. To evaluate the model's generalizability, its performance is measured at the 20% of photographs that remain unseen. Constant performance on new statistics is assured via assessing the model's accuracy, precision, and recall the usage of the testing set.

4.4 Algorithms:

4.4.1 Classification Models

In the realm of image type, the "vision Transformer" (ViT) is the preferred deep learning architecture. The primary concept involves segmenting images into smaller components and eventually analyzing the interrelations among them through a transformer mechanism. This technique is typically applied for sequential data, including text. Especially for tough skin lesion classification tasks, ViT's self-attention mechanism permits it to efficiently recognize worldwide context. In this context, ViT is employed to research pores and skin lesions, and its capacity to model spatial dependency enhances type accuracy [17]. MobileNetV2 is a deep learning version characterized by its lightweight architecture and efficient construction, making it very effective for mobile and embedded tool programs. To reduce computational needs while preserving excessive accuracy, it utilizes depthwise separable convolutions and linear bottleneck layers. The yr 19 this version's rapid, actual-time processing and minimal aid consumption decorate skin lesion classification duties. Cellular health applications can enhance accessibility by accurately classifying skin lesions using MobileNetV2, even on gadgets with limited processing electricity. To mitigate the problem of vanishing gradients in deep neural networks, ResNet152V2, part of the ResNet structure, is designed utilizing residual learning and functions as a deep convolutional community. [18] Inside It employs bypass connections to beautify gradient float at some point of the network. The capacity to differentiate between benign and malignant skin conditions is based on ResNet152V2's skillability in studying complicated styles, a capability afforded by its depth. Utilizing it improves performance for unique lesion categorization and ensures the model's resilience in managing extensive datasets. VGG16 is a 16-layer convolutional neural network (CNN) architecture recognized for its depth and simplicity. In [20], the performance of characteristic extraction and the simplicity of implementation are each improved via its uniform architecture of layered 3x3 convolutional layers. VGG16 is hired to extract intricate details from lesion pictures for the class of pores and skin lesions. The model's depth enhances type results by facilitating the extraction of advanced representations of skin lesions. This approach is reliable for dermoscopy picture evaluation due to its simplicity and effectiveness.

a sophisticated convolutional neural network known as Xception (extreme Inception) was developed to leverage depthwise separable convolutions. This version plays successfully on photo popularity obligations due to its utilization of separable convolutions, which are more efficient than standard convolutions. Xception is fine for skin lesion categorization since it attains superior effects with fewer parameters and decreased computational assets. The model's capacity to extract detailed facts from pix enhances both diagnostic velocity and accuracy, facilitating the separation of diverse types of skin lesions. For photo class and other supervised machine learning responsibilities, aid Vector device (SVM) is the optimal choice. Support vector machines (SVMs) facilitate feature extraction and decorate pattern popularity through diminishing the dimensionality of a dataset. Pores and skin lesion categorization can be achieved by support vector machines (SVMs) by classifying extracted features into malignant and benign categories. To provide advanced accuracy regardless of records imbalance or noise, it provides a sturdy answer by way of maximizing the margin between lessons.

k-Nearest neighbors (KNN) is a good supervised learning technique that categorizes data factors based on their proximity to neighboring ones. KNN identifies common styles in the feature area to extract traits from pores and skin lesion pictures. KNN utilizes functions extracted from CNN models to categorise pores and skin lesions primarily based on their proximity to labeled times within the training dataset. This approach for skin lesion class is both comprehensible and notably effective.
The vote casting Classifier enhances class accuracy with the aid of utilizing an ensemble gaining knowledge of method that integrates multiple models. The vote casting Classifier enhances precision and resilience by means of amalgamating predictions from "decision Tree (DT) and Random forest (RF)". This composite methodology effectively integrates Random forest and decision trees for the classification of skin lesions, enhancing decision-making and performance on complex lesion datasets. This hybrid model optimizes classification even as simultaneously mitigating overfitting and enhancing generalization [18].

4.4.2 *Detection Models*

YOLOv5s6 is a revised iteration of the "You only look once (YOLO)" series, designed for efficient and specific actual-time item detection. It identifies and locates items in pix with exceptional precision in a single forward pass. One application of YOLOv5s6 is the precise identification and classification of skin lesions in dermoscopic photos. It is right for real-time analysis as a result of its pace and precision, facilitating the swift identification and diagnosis of skin lesions.
The YOLOv5x6 series, an enhanced iteration of the YOLOv5 series, is characterised via advanced accuracy and object detection performance. It enhances recognition capabilities, particularly for intricate and diminutive items, by the utilization of an increased model structure. Leveraging its remarkable detection capabilities, YOLOv5x6 is nice in pores and skin lesion identity for as it should be figuring out and localizing a various array of lesions, even those with subtle variations. Complex skin lesion photographs, wherein precision is important for diagnosis, are preferably matched to its enhanced architecture.
YOLOv8 is an enhanced model of the YOLO series designed for speedy and specific item detection. Its primary benefit is the efficient processing of substantial volumes of image facts. YOLOv8 allows accurate and rapid localization of pores and skin lesions in dermoscopic pics. The performance is important in actual-time packages, permitting doctors to obtain immediate remarks on lesion detection and utilize it for advanced diagnostics.
The latest addition to the YOLO own family, YOLOv9, has been created to enhance the recognition of complicated and nuanced tasks. Even minuscule or challenging-to-diagnose cutaneous lesions can be exactly identified with the assistance of YOLOv9. It enhances lesion analysis through expediting and refining image analysis with superior algorithms. YOLOv9's efficacy is assured by means of its fast processing of huge datasets, even if addressing a diverse array of skin lesion types.

5. RESULTS AND DISCUSSION:

Accuracy: If the test can consistently differentiate between healthy individuals and patients, it is deemed correct. figuring out the test's accuracy necessitates calculating the ratio of cases with valid results to those without. It appears theoretically as follows:

$$"Accuracy = \frac{\mathrm{TP + TN}}{\mathrm{TP + FP + TN + FN}} (1)"$$

Precision: Accuracy is defined as the proportion of correctly identified positive occurrences or samples expressed as a percent. the subsequent formula was employed to check this value:

$$"Precision = \frac{\text{True Positive}}{\text{True Positive } + \text{False Positive}} (2)"$$

Recall: The metric in metric learning assesses the model's capacity to identify all pertinent times of the class. by comparing the amount of accurately predicted positive instances to the total number of real positives, we may assess the version's efficacy in identifying the class instances.

$$"Recall = \frac{\mathrm{TP}}{\mathrm{TP + FN}} (3)"$$

Sensitivity: test sensitivity evaluates the efficacy of a test or instrument in diagnosing an issue. it is determined by contrasting the quantity of positive exams with the actual impacted population.

$$\text{"}Sensitivity = \frac{TP}{(TP + FN)}(4)\text{"}$$

Specificity: The calculation involves dividing the quantity of individuals who test poor for a condition through the total range of individuals without the condition, which includes fake positives.

$$\text{"}Specificity = \frac{TN}{(TN + FP)}(5)\text{"}$$

mAP: quality rating metric "mean average Precision (MAP)". the quantity of pertinent recommendations and the ranking role are taken into account. MAP at ok represents the arithmetic mean of average Precision (AP) at k for all users or queries.

$$\text{"}mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k(6)\text{"}$$

Desk 1 contrasts the algorithms according to their performance metrics: accuracy, precision, recall, sensitivity, and specificity. The overall performance metrics of every set of rules—"mAP, consider, and precision"—are offered in table 2. The optimal results for detection and classification are achieved the use of "Xception, FS-SVM, and YoloV8." Furthermore, measurements from alternative algorithms are also included for comparative analysis.

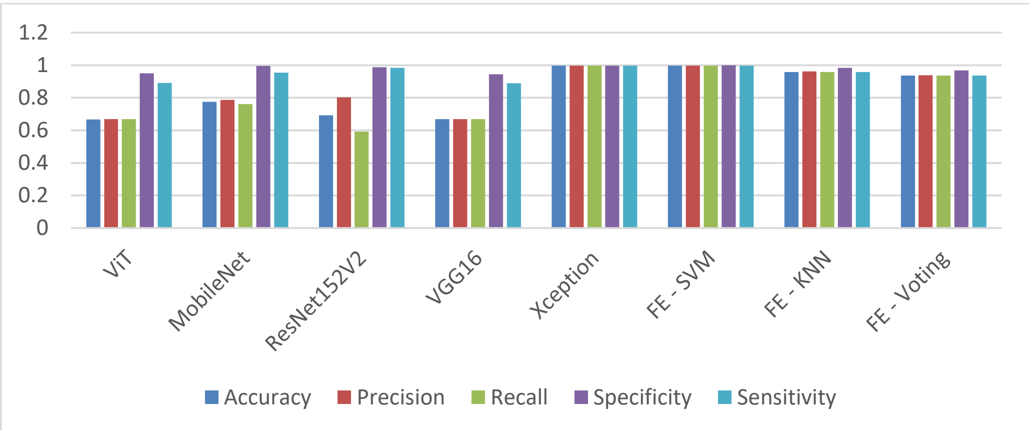
Table 2: Performance Evaluation Metrics of Classification

Model	Accuracy	Precision	Recall	Specificity	Sensitivity
ViT	0.668	0.669	0.669	0.950	0.891
MobileNet	0.776	0.788	0.761	0.996	0.954
ResNet152V2	0.693	0.803	0.592	0.988	0.985
VGG16	0.669	0.669	0.669	0.945	0.890
Xception	**0.997**	**0.997**	**0.997**	**0.997**	**0.997**
FE - SVM	**0.997**	**0.997**	**0.997**	**0.999**	**0.997**
FE - KNN	0.959	0.962	0.959	0.984	0.959
FE - Voting	0.936	0.939	0.936	0.969	0.936

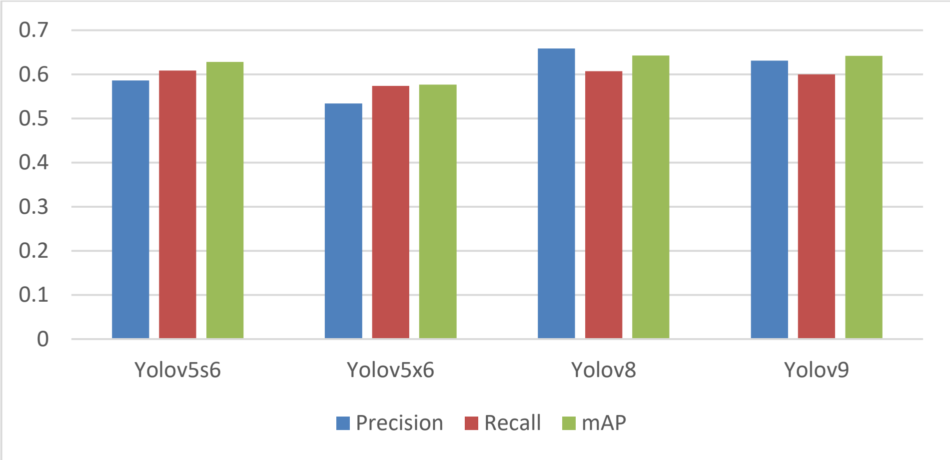
Table 3: Performance Evaluation Metrics of Detection

Model	Precision	Recall	mAP
Yolov5s6	0.586	0.609	0.628
Yolov5x6	0.534	0.574	0.577
Yolov8	**0.659**	**0.607**	**0.643**
Yolov9	0.631	0.600	0.642

Graph 1: Comparison Graphs of Classification



Graph 2: Comparison Graphs of Detection



In Graph 1, blue represents accuracy, maroon signifies precision, green indicates recall, violet implies specificity, and mild blue reflects sensitivity. In Graph 2, the colours blue, maroon, and green represent precision, recall, and violet, respectively. Xception, FS-SVM, and YoloV8 demonstrate superior performance relative to the other algorithms across all metrics. The aforementioned graph visually represents these details.

6. CONCLUSION:

This study proposed an innovative approach to dermatology by redefining skin lesion type via the integration of deep learning and multimodal statistics analytics. The device markedly more desirable its categorization and detection abilities by way of integrating superior algorithms with the "skin cancer facts" dataset. The categorization framework finished a remarkable performance of 99.7 percent with its "Xception" and "feature Extraction – assist Vector system (FE-SVM)" modules. Reaching a "mean average Precision (mAP)" of 64.3%, "YOLOv8" surpassed the other YOLO models within the detection phase of the proposed system. The findings imply that multimodal analytics, in conjunction with switch learning, can surpass the limitations of unmarried-modality processes and provide a dependable and efficient option for dermatological diagnosis. The design demonstrates that deep learning complements dermatological patient effects by increasing diagnostic accuracy and allowing scalable, dependable, and interpretable clinical solutions.

Destiny research will focus on enhancing the model's generalizability by augmenting the dataset's variety via the inclusion of extra skin types and uncommon lesion categories. The amalgamation of photo and non-image facts, encompassing genetic data and patient histories, might be stronger through the research of sophisticated multimodal facts fusion methodologies. Efforts can be undertaken to enhance the framework's computational efficiency for real-time clinical implementation. Medical applicability and reliability can be guaranteed through further evaluation in collaboration with dermatologists in real-world settings.

REFERENCES:

[1] Esposito, C., Ficco, M., & Gupta, B. B. (2021). Block chain-based authentication and authorization for smart city applications. *Inf. Process. Manage*, 58(2), 145631-145652.

[2] Viswanath, G. (2022). A Smart Recommendation System for Medicine using Intelligent NLP Techniques. *International Conference on Automation, Computing and Renewable Systems (ICACRS)*, 5(2), 1081-1084

[3] Khouloud, S., Ahlem, M., Fadel, T., & Amel, S. (2022). W-Net and inception residual network for skin lesion segmentation and classification. *Appl. Intell.,* 52(4), 3976–3994.

[4] Swapna, G., & Bhaskar, K. (2024). Early-Stage Autism Spectrum Disorder Detection Using Machine Learning. *International Journal of HRM and Organizational Behavior*, 12(3), 269-283,

[5] Heroza, R. I., Gan, J. Q., & Raza, H. (2024). Enhancing skin lesion classification: A self-attention fusion approach with vision transformer. *In Annual Conference on Medical Image Understanding and Analysis,* 14860(1), 309-322.

[6] Viswanath, G. (2024). Multiple Cancer Types Classified Using CTMRI Images Based On Learning Without Forgetting Powered Deep Learning Models. *International Journal of HRM and Organizational Behavior,* 12(3), 243-253.

[7] Patwary, A. N., Fu, A., Battula, S. K., Naha, R. K., Garg, S., & Mahanti, A. (2020). Fogauthchain: A secure location-based authentication scheme in fog computing environments using block chain. *Comput. Commun,* 162(2), 212-224.

[8] Viswanath, G., & Swapna, G. (2024). Health Prediction Using Machine Learning with Drive HQ Cloud Security. *Frontiersin Health Informatics*. 13(8), 2755-2761.

[9] Duong, H., Roccuzzo, A., Stähli, A., Salvi, G.E., Lang, N.P., & Sculean, A. (2022). Oral health-related quality of life of patients rehabilitated with fixed and removable implant-supported dental prostheses. *Periodontology,* 88(4), 201–237.

[10] Viswanath, G., & Swapna, G. (2025). Data Mining-Driven Multi-Feature Selection for Chronic Disease Forecasting. *Journal of Neonatal Surgery*, 14(5s), 108-124.

[11] Muramatsu, C., Morishita, T., Takahashi, R., Hayashi, T., Nishiyama, W., Ariji, Y., Zhou, X., Hara, T., Katsumata, A., Ariji, E., et al. (2021). Tooth detection and classification on panoramic radiographs for automatic dental chart filing: Improved classification by multi-sized input data. *Oral Radiol*, 37(5), 13–19.

[12] Viswanath, G. (2024). Personalized Breast Cancer Prognosis through Data Mining Innovations. *Cuestiones de Fisioterapia*, 53(2), 538-548.

[13] Liu, Z., Xiong, R., & Jiang, T. (2023). CI-Net: Clinical-Inspired Network for Automated Skin Lesion Recognition. *IEEE Transactions on Medical Imaging,* 42(3), 619-632.

[14] Viswanath, G. (2021). Adaptive Light Weight Encryption Algorithm for Securing Multi-Cloud Storage. *Turkish Journal of Computer and Mathematics Education*, 12(9), 545-554.

[15] Wang, G., Liu, L., Lauri, F., & El-Hassani, A. H. (2024). Multi-Scale DCNN with Dynamic Weight and Part Cross-Entropy Loss for Skin Lesion Diagnosis. *Big Data Mining and Analytics*, 7(4), 1347-136.

[16] Tian, Q., Yan, L., Zhang, X., Zhang, X., Hu, Y., Han, Y., Liu, Z., Nan, H., Sun, Q., Sun, Y., Yang, Y., Yu, Y., Zhang, J., Hu, B., Xiao, G., Chen, P., Tian, S., Xu, J., Wang, W., & Cui, G. (2018). Radiomics strategy for glioma grading using texture features from multiparametric MRI. *J. Magn. Reson. Imag*., 48(6), 1518–1528.

[17] Emblem, K. E., Zoellner, F. G., Tennoe, B., Nedregaard, B., Nome, T., Due-Tonnessen, P., Hald, J. K., Scheie, D., & Bjornerud, A. (2008). Predictive modeling in glioma grading from MR perfusion images using support vec tor machines. *Magn. Reson. Med*., 60(4), 945–952.

[18] Jabeen, T., Ashraf, H., & Ullah, A. (2021). A survey on healthcare data security in wireless body area networks. *J. Ambient Intell. Humanized Comput*., 12(2), 9841–9854.

[19] Zou, L., Luo, X., Zhang, Y., Yang, X., & Wang, X. (2023). HC-DTTSVM: A network intrusion detection method based on decision tree twin support vector machine and hierarchical clustering. *IEEE Access*, 11(1), 21404–21416.

[20] Yue, G., Wei, P., Zhou, T., Jiang, Q., Yan, W., & Wang, T. (2022). Toward multicenter skin lesion classification using deep neural network with adaptively weighted balance loss. *IEEE Transactions on Medical Imaging,* 42(1), 119-131.

[21] Swapna, G. (2023). A Drug-Target Interaction Prediction Based on Supervised Probabilistic Classification. *Journal of Computer Science*, 19(3), 1203-1211.

[22] Emre Celebi, M., Iyatomi, H., Schaefer, G., & Stoecker, W. V. (2009). Lesion border detection in dermoscopy images. *Computerized Medical Imaging and Graphics,* 33(2), 148-153.

[23] Viswanath, G. (2024). Machine-Learning-Based Cloud Intrusion Detection. *International Journal of Mechanical Engineering Research and Technology,* 16(5), 38-52.

[24] Samanta, P. K., & Rout, N. K. (2021). Skin lesion classification using deep convolutional neural network and transfer learning approach. *Advances in Smart Communication Technology and Information Processin. Springer,* 165(1), 327–333.

[25] Swapna, G., & Bhaskar, K. (2024). Malaria Diagnosis Using Double Hidden Layer Extreme Learning Machine Algorithm With Cnn Feature Extraction And Parasite Inflator. *International Journal of Information Technology and Computer Engineering,* 12(4), 536-547.

[26] Adebanji, O. O., Gelbukh, I., Calvo, H., & Ojo, O.E. (2022). Sequential models for sentiment analysis: A comparative study. *Mexican Int. Conf. Artif. Intell. Springer*, 13613(2), 227–235.

[27] Viswanath, G., & Swapna, G. (2025). Diabetes Diagnosis Using Machine Learning with Cloud Security. *Cuestiones de Fisioterapia,* 54(2), 417-431.

[28] Vachmanus, S., Noraset, T., Piyanonpong, W., Rattananukrom, T., & Tuarob, S. (2023). DeepMetaForge: A Deep Vision-Transformer Metadata-Fusion Network for Automatic Skin Lesion Classification. *IEEE Access.* 11(1), 145467-145484,

[29] Viswanath, G. (2024). Improved Light GBM Model Performance Analysis and Comparison for Coronary Heart Disease Prediction. *International Journal of Information Technology and Computer Engineering,* 12(3), 658-672.

[30] Tan, C., Talreja, K., Shivam, A., Nag, A., Miranda, M., Raghav, S., Rajaram, A., Lutz, M., Kumar, S., & Kumar, A. (2022). Machine learning in clinical text classification: Specialty identification and COVID-19 risk. *13th Int. Conf. Inf. Commun. Technol. Converg. (ICTC),* 2022(1), 1366–1370.

[31] Viswanath, G. (2024). Enhancing Cloud Security: A Block chain-Based Verification Framework for Multi-Cloud Virtual Machine Images. *Frontiers in Health Informatics*, 13(3), 9535-9549.

[32] Ojo, O. E., Ta, H. T., Gelbukh, A., Calvo, H., Adebanji, O. O., & Sidorov, G. (2023). Transformer-based approaches to sentiment detection. *in Proc. 8th World Conference Soft Computing, Baku, Azerbaijan. Springer,* 423(2), 101–110.

[33] Viswanath, G. (2021). Hybrid encryption framework for securing big data storage in multi-cloud environment. *Evolutionary intelligence*, 14(2), 691-698.

[34] Onan, A., Bulut, V., & Ezgi, A. (2023). Enhancing Skin Lesion Classification with Ensemble Data Augmentation and Convolutional Neural Networks. *International Congress of Electrical and Computer Engineering,* 2023(1), 131-145.

[35] Viswanath, G. (2023). A Real-Time Case Scenario Based On URL Phishing Detection Through Login URLS. *Material science and technology,* 22(9), 103-108.