Hairfall and Scalp Disease Detection Using Deep Learning and AI Thesis Description

Hira Farman[1*], Usman Amjad[2], Alisha Farman[1], Jaish Kadiwal[1], Usman Saleem[1], and Hamood[1]

[1]Department of Computer Science, Iqra University, Karachi, Pakistan
[2]Computer Science and Information technology Department, Ned University of Engineering and Technology, Karachi, Pakistan

ABSTRACT Millions of people worldwide suffer from common health issues like hair loss and scalp disorders, which can cause psychological discomfort and, in extreme situations, necessitate medical attention. Numerous factors, such as genetics, lifestyle, environmental effects, and underlying health disorders, contribute to the complexity in diagnosing these illnesses. Because of their diverse and erratic presentations, dermatologists and trichologists encounter significant difficulties in correctly diagnosing and treating these conditions. Recent developments in artificial intelligence (AI) have opened up new avenues for improving dermatology diagnostic accuracy. This study makes use of cutting edge deep learning and machine learning methods to better accurately identify scalp conditions and hair loss trends. Random Forest, K Nearest Neighbors (KNN), Logistic Regression, Convolutional Neural Network (CNN), Simple Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), and Artificial Neural Network (ANN) are among the models whose performance we assess. Standard classification metrics including Accuracy, Precision, Recall, F1-Score, and Area under the Curve (AUC) are used to gauge each model's efficacy. With an AUC of 0.975, the results show that Logistic Regression has the best accuracy for class separation, demonstrating its potent ability to distinguish between circumstances. Furthermore, with an accuracy of 0.15, precision of 0.067, recall of 0.15, and F1-Score of 0.08, the RNN model with a TANH activation function was the best performer on a number of criteria. By addressing the vanishing gradient problem, a prevalent difficulty in recurrent models, the TANH function which maps inputs between -1 and 1 proves beneficial and improves predictive stability. These results highlight how AI-driven models have the potential to greatly enhance therapy planning and diagnostic precision in the management of hair loss and scalp disorders. AI has the potential to be a crucial tool in dermatology with future development, enabling more precise, effective, and early detection that will ultimately improve patient outcomes and enable more focused therapies.

INDEX TERMS hairfall prediction, scalp disease detection using deep learning, convolutional neural networks, recurrent neural networks, long short-term memory.

I. INTRODUCTION

Hair fall prediction can contribute significantly to helping individuals make sophisticated decisions regarding hair care, treatment options, and overall health. Factors such as genetics, hormonal imbalances, nutritional deficiencies, stress, and environmental conditions significantly impact hair health. In particular, hair care

*Corresponding Author: hira.farman@iqra.edu.pk

and treatment industries benefit from accurate hair fall predictions, as they enable better product development and customized treatment plans. Hair loss affects over 80 million individuals in the U.S, essentially due to aging, stress, medication, or genetics. Early diagnosis is often delayed due to dependence on professional dermatologists and costly tests. Deep learning techniques like convolutional neural networks (CNNs) enable early detection of scalp conditions, making the process more available and effective [1].

A System for Scalp Health Diagnosis and Inspection Based on Machine Learning which provides poor daily habits often result in common scalp problems such as greasy hair, folliculitis, dandruff, and hair loss. AI-based techniques, particularly machine learning (ML), allows for exact categorization of scalp damage and healthy hair, providing a productive approach to early detection of hair loss. Similarly, early prediction of hair fall is crucial. It helps individuals take protective measures, make informed decisions about hair care and treatment, and even avoid advanced stages of hair damage [2].

Alopecia Areata (AA) prediction with machine learning methods provides a dataset of 1,000 images of healthy hair that are collected through web scraping, along with the Figaro 1k dataset. The research utilizes SVM, KNN, and CNN algorithms, highlighting the importance of computer-aided diagnosis in improving the accuracy of AA prediction and classification. Today, artificial intelligence techniques are increasingly used for data analysis and prediction across different fields, including hair and scalp health. In various studies for hair fall, various Data Mining and Machine Learning (ML) approaches have been used for prediction that are applied, resulting in improved accuracy. In this study, a deep learning methodology was employed to analyze hair loss data. Deep learning beat at handling large datasets and complex patterns [3]. Hair Fall, sometimes referred to as baldness or alopecia, directs to a hair loss resulting from part of the head. Typical varieties consist of male-or alopecia areata, and hair loss with a feminine pattern, while hair thinning is referred to as telogeneffluvium. An Ordinal Logistic Regression study attributes behavioral factors to promote FPHL (Female Pattern Hair Loss) seriousness, and suggests avoiding alcohol, ponytails, and improving sleep to prevent worsening [4]. Evaluation of Patients with Alopecia, the research mentions the precise tools for diagnosing and determining alopecia, including structured interviews, questionnaires, and clinical examinations [5].

A. CAUSES OF HAIR FALL

Numerous factors, such as genetic predisposition, hormonal fluctuations, nutritional deficiencies, stress, illnesses, and certain medical condition as shown in fig 1, can contribute to hair loss. Particularly in diseases like androgenetic alopecia, the most prevalent kind of hair loss, hereditary factors are important. Increased shedding may also result from hormonal changes brought on by pregnancy, menopause, or thyroid problems. Hair follicles might also become weaker due to a lack of vital nutrients like protein and iron. Physical or emotional stress can cause telogen effluvium, a temporary loss of hair. Hair loss may also be exacerbated by some medical problems, such as infections of the scalp or autoimmune diseases like alopecia areata. Lastly, the fact that hair loss is a side effect of some medications used to treat different health concerns highlights the complex nature of this condition [6].

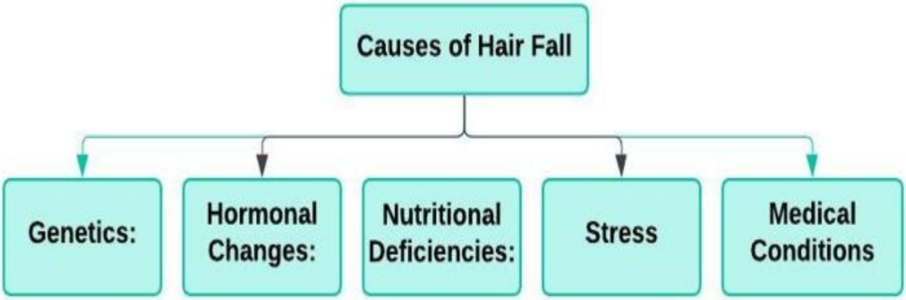

FIGURE 1. Causes of hair fall

B. PROBLEM STATEMENT

The primary objective of the problem-solving approach is to develop an advanced Artificial Intelligence learning-based system for detecting hair fall and scalp diseases to improve the precision of diagnosing hair fall and various scalp diseases by leveraging advanced deep learning models, to streamline the diagnostic process, reducing the need for manual interference and increasing the workflow and offer detailed and actionable insights based on the analysis, enabling timely and appropriate treatment decisions. Hair fall can be caused by a variety of factors, including genetic, environmental, and medical conditions, leading to diverse patterns and symptoms. Scalp diseases include a range of conditions such as dandruff, psoriasis, eczema, and fungal infections, each with distinct symptoms. The aim of this study is to provide advanced AI machine learning and deep learning techniques to increase the accuracy of detecting specific conditions and their severity, and automate the analysis of images and data to reduce manual effort and improve the efficiency of the diagnostic process. The detection of hairfall is possible by creating a deep learning model that can analyze scalp images.

C. RESEARCH QUESTIONS

RQ1: How productive are deep learning models in detecting different types of hair fall and scalp diseases compared to standard diagnostic methods

RQ2: Which machine learning or deep learning model provides the highest accuracy and reliability for predicting hair fall patterns and scalp diseases, and what factors contribute to its success?

RQ3: How do different activation functions (SIGMOID, RELU, TANH) impact the performance of various deep learning models (such as CNN, RNN, LSTM, and ANN) in accurately classifying hair and scalp diseases?

II. RELATED WORK

One aspect of the human body that significantly influences a person's beauty and look is their hair. In order to identify and limit hair follicles and assess the degree of hair loss in microscopic pictures, the Hair Loss Severity Estimation using Mask R-CNN suggests an intelligent system that combines multitask learning with Mask R-CNN. Three classes healthy, normal, and severe—are used to group the photos. An artificial neural network (ANN) is used to forecast hair loss over time. These methods were used to identify hair loss in various scalp diseases [7]. CNN's prognosis of

Alopecia Areata, it introduces a brand-new CNN architecture built for effective detection using a dataset of images. CNN achieves 98% accuracy on the dataset when used to analyse the dreadful autoimmune condition known as Alopecia Areata, which causes bald areas, brittle nails, and hair destruction. This gives way to a model that uses information on scalp health to forecast the degree of hair loss. Factors including follicle health, environmental circumstances, hormone levels, hair density, and scalp texture were all included in the dataset, which was gathered from dermatological clinics [8].

Deep Learning-Based Hair Loss Stage Prediction, using practical grouping, a convolutional neural network (CNN), was used to forecast the stages of hair loss. In order to improve early diagnosis and individualized therapy, this work demonstrates that deep learning can readily detect the phases of hair loss from frontal scalp photos. This hybrid approach was used to forecast hair loss up to seven stages in advance. 94% accuracy was attained by the technique [9]. This study, Automated Measurement of Hair Density utilizing Deep Neural Networks, is on utilizing deep neural networks to automate the identification of Hair Density Measurement (HDM) objects. For precise analysis, the dataset included 4,492 RGB photos of male patients with hair loss, together with information on the kinds and locations of each hair follicle [10].

In order to enable objective tracking and early detection of scalp and hair problems using cutting-edge deep learning techniques, the Intelligent Healthcare Platform for detection of Scalp and Hair problems investigates three deep learning models: ResNet-152, Net B-6, and ViT-B/16. It demonstrates how well data-driven methods work to comprehend hair loss trends and get a 70% accuracy rate [11]. Identification of Hair and Scalp Disorders enhances patient diagnosis and treatment through the use of deep learning, which focusses on incorporating machine learning into online apps. It seeks to highlight how technology is changing healthcare and patient management in the future. For this prediction, a deep convolutional CNN, or neural network, was used [12].

Using webcam and microscope sensors, this intelligent hair and scalp analysis system uses the Norwood-Hamilton Model and camera sensors to evaluate the health of hair and scalp by creating precise feature photos. The technology evaluates hair loss and scalp problems with 90% accuracy by using a deep learning model [13]. Deep Learning to Determine the Degree of Hair Loss in Face Photos uses a suggested matching approach to automatically classify facial photos using a training dataset with different degrees of baldness. In order to accurately forecast hair loss, the study examined four factors: age, follicle size, scalp condition, and hair density. Metrics like Error Squared Root Mean (RMSE) and a confusion matrix were used to gauge the prediction accuracy [14].

The Prediction of Hair Fall Patterns in a Person Using Artificial Intelligence for Better Care and Treatment describes how deep learning algorithms may be used to predict hair loss patterns linked to telogen efflux and alopecia areata, with an accuracy of 85% [15]. In the article Survey Based Machine Learning Approaches to Diagnose Hair Fall Disorder in the Bangladeshi Community, a range of machine learning techniques, such as Random Forest, K Nearest Neighbors (KNN), and Support Vector Machine (SVM), are used to evaluate accuracy when analyzing datasets related to hair fall disorders. Metrics such as F measure, accuracy, precision, and recall were used to assess the performance, which showed that it was effective in detecting hair loss diseases with a prediction accuracy of 89.54% [16].

The article, An Analysis of Alopecia Areata Classification Framework for Human Hair Loss Based on VGGSVM Approach presents a framework for evaluating neural networks to distinguish between alopecia and non-alopecia situations. This framework distinguishes between alopecia and healthy hair using a dual model method. Seventy percent of the image dataset is utilized to construct a machine learning model using Support Vector Machines (SVM) 23 after features are extracted using the VGG19 model. Machine Learning on Classification of Healthy and Unhealthy Hair provides answers to concerns regarding the reliability and suitability of deep learning models in critical applications, such as medical diagnostics, with an impressive accuracy of 96.63% [12].

A Healthy Scalp Inspection and Diagnosis System Ung Multiple of Deep Learning-Based Modules presents a novel multimodal deep learning-oriented system for scalp inspection and diagnosis. Diagnostic accuracy is increased with the application of AI-driven item recognition approaches. In the article, Trichoscopic [17] Characteristics of Hair Loss in Women, according to a one-year hospital-based cross-sectional study, trichoscopy may effectively identify early female pattern hair loss (FPHL) and differentiate it from other conditions even in the absence of hormone testing. Combining AI and deep learning models significantly reduces the need for comprehensive scalp inspections [18].

The trichoscopic features of female hair loss include: Even without hormone testing, trichoscopy may successfully identify early FPHL (Female Pattern Hair Loss) and differentiate it from other conditions, according to a one-year hospital-based cross-sectional research [19]. According to the study "Effect of Behavioral Factors on Severity of Female Pattern Hair Loss: An Ordinal Logistic Regression Analysis," avoiding alcohol, wearing ponytails, and getting more sleep can all help prevent the condition from getting worse. A paper presenting the quantitative classification [20] for androgenetic alopecia and its application to hair transplantation provides a new numerical PRECISE scale for classifying (Androgenetic Alopecia) AGA with subjective methods, recommending 1500 follicular units per score for hair transplantation to improve results [21].

The paper, Hair Transplantation in the United States: A Population-based Survey of Female and Male Pattern Baldness, gives the details that Americans see hair loss as a major issue and value hair transplantation, but more affordable, gender-specific options and public opinions on HT (Hair Transplantation) need more analysis. A research on types of hairline [22] recession in androgenetic alopecia and perceptions of aging in Asian males, mentions that cranial hair loss is connected to aging and affects age perception in Western males, but the impact of hairline slowdown in individuals with PFA (Perceived Facial Age) is not well known [23]. Evaluation of Patients with Alopecia, highlights the need for refined, precise tools for diagnosing and determining alopecia, including structured interviews, questionnaires, and clinical examinations [8].

Association of Hair Loss With Health Utility Measurements Before and After Hair Transplant Surgery in Men and Women, shows that hair transplant surgery improves health utility scores for androgenetic alopecia in both men and women compared to unfinished cases [3]. Power of Molecule on Hair Growth A Clinical Study, mentions that Hair follicles produce keratin, the main protein in hair. Regular use of hair serum with NX35 growth molecule improves hair density, volume, and thickness. This paper maintains that hair loss affects 50% of men and 15–30% of women. A research [24] about hair loss among transgender and gender-non binary patients: a cross-sectional study, examines how gender-specific hormones impact scalp hair loss, underlining the need for personalized diagnosis and treatment. MHT (Masculinizing Hormone Therapy) uses testosterone for male features, while FHT (Feminizing Hormone Therapy) uses estrogen and antiandrogens for female features [25].

A research titled, reliability of horizontally sectioned scalp biopsies in the diagnosis of chronic diffuse telogen hair loss in men, provides the understanding of male pattern baldness that is critical for successful hair transplant surgery, which depends on proper patient selection and good results of Hair Transplantation. Another research titled, Classification [26] of the types of androgenetic alopecia (common baldness) occurring in females, categorizes female androgenetic alopecia phases which presented to enable early diagnosis and treatment with antiandrogens. A paper titled, A New Classification of Male Pattern [27] Baldness and a Clinical Study of the Anterior Hairline provides the baselines for hair restoration by identifying six types of male pattern baldness and highlights the need for systematic measurements of hairline structures.

The paper on the prevalence [28] and types of androgenetic alopecia in Korean men and women, mentions that 48.5% of men and 45.2% of women have a family history of baldness, Type III vertex connection is most common between ages 30-70, with type VI existing after 70. Korean men often have more frontal hairline, and female patterns are seen in 11.1% of cases. The commonness of AGA Androgenetic Alopecia (Norwood III or above) in Korean men is 14.1%. Another paper titled, Trichoscopic Patterns of No scarring Alopecia's, analyzes a study group with a mean age of 26. Trichoscopy discloses common features such as broken hair and black dots (48% each), and honeycomb pigmentation (26%). Alopecia areata (AA) is more frequent in males (41.8%), while females are uniformly affected by AA and female pattern hair loss (29.8%). Early diagnosis and treatment of hair loss, a major cause of psychological stress, are critical. 80 million Americans suffer from hair loss due to aging, stress, medication, or genetics. The hair-related diseases when diagnosed are often faced with delays due to the need for professional dermatologists and medical tests. Convolutional neural networks (CNNs) validate early-stage detection. AI gives the details of Machine Learning-Based Scalp Hair Inspection and Diagnosis System for Scalp Health. Poor daily habits lead to common scalp and hair issues like dandruff, folliculitis, hair loss, and oily hair. ML-based techniques enable accurate classification of healthy hair and damaged scalp, offering an effective solution for detecting hair loss. These techniques present the Prediction of Alopecia Areata using Machine Learning Techniques. A dataset of 1,000 images of healthy hairs are collected through web scraping and the Figaro 1k dataset. In order to effectively detect Alopecia Areata, a persistent autoimmune illness that causes bald spots, brittle nails, and hair destruction, this study suggests a revolutionary CNN architecture. On the used picture dataset, the CNN model achieves an amazing 98% accuracy rate. Predicting the stages of hair loss using CNN to automatically recognize various stages from frontal photographs is another component of the research, which advances methods for diagnosis and treatment. Additionally, a dataset of 4,492 RGB photographs of male patients with annotated hair follicle locations and kinds was used to construct an automated hair density measuring (HDM) system employing deep neural networks.

III. COMPARATIVE ANALYSIS OF PAST STUDIES FOR HAIR FALL DETECTION USING AI AND MACHINE LEARNING

A comparative descripton of earlier research on the classification of hair diseases is shown in Table 1, which summarizes important elements such the goals of each study, the models used, the main conclusions, any gaps or limits found, and their applicability to the current investigation. It summarizes the different machine learning and deep learning models that have been utilized, describes the objectives of previous researches, and emphasizes noteworthy results in terms of model accuracy and performance. It also highlights the shortcomings of earlier research, such as inadequate data or problems with generalizability, and explains how these conclusions apply to the current study. This well-organized summary facilitates the identification of research contributions and areas for improvement in the use of deep learning methods for the classification of hair diseases.

IV. GAP ANALYSIS

A. SOURCE TABLE

Table-II presents a carefully selected group of literature that shows the development of a deep learning model for recognizing scalp diseases and hair loss. These sources provide valuable understanding of many machine learning and image processing approaches related to scalp health, featuring both advanced deep learning frameworks and traditional techniques. The overview highlights the connection of these studies for the hair loss detection using AI, machine learning and deep learning.

TABLE I

COMPARATIVE ANALYSIS

Ref	Objective of study	Models	Key Finding	Gaps/Limitation	Relevance to our work
[1]	To classify hair follicles and estimate hair loss severity.	Utilized Res Net-50 and ResNet-101 with Mask R-CNN.	ResNet-50 had higher misclassification ion rates than ResNet-101.	Focuses primarily on accuracy; does not explore user interface or practical application.	Supports the application of CNN and Mask R-CNN in dermatology.
[2].	to use machine learning to identify disorders of the ha ir and scalp	Deep learning with CNN for scalp disease detection.	Achieved 96.2% training accuracy, and 91.1% validation accuracy Identifying diseases accurately.	Limited datasets hinder the generalizability of findings.	Directly applicable to the developm ent of systems for detecting scalp di sorders.
[3].	To suggest a methodology for alopecia early detection.	Two-layer feed-forward network with back propagation.	Using 100 samples, a 91% training accuracy was attained.	Limited sample size; focus on only one type of hair disease.	Provides a foundation for improving early detection techniques.
[4].	Investigating deep learning developments for scalp disease di agnosis	Reviewed the integration of CNN and FCN for diagnosis	Enhanced diagnostic Potential as a result of devel opments in computer vision and processing.	highlights the need for more re search and developments in deep learning Applications.	Supports the ongoing development of AI-driven diagnostic tools for scalp health.
[5].	Using machine learning to disting uish between healt hy hair and alopecia areata	Implemented a machine learning classification framework.	Created a model that was highly accurate and focused on particular hair conditions.	Other scalp conditions might n ot be fully covered by the Frame work.	Discusses classification methods th at are pertinent to hair Disorders.
[6].	The goal of the research study is to improve the diagnosis and categorization of scalp illnesses by creating a deep learning model with Res Net.	Features retrieved by Res Net, localized follicles using ROI alignment, density-based classification, and scores standardized across circumstances Making Use of Mask R-CNN	Overall, ResNet-50 outperformed ResNet-101 in misclassifying labels as severe, normal, and healthy.	The short sample size, lack of diversity, robustness problems, high computing demands, and interpretability issues are some of the study's weaknesses.	This study paper addresses issues of accuracy and improves clinical decision-making for scalp problems by utilizing deep learning to improve hair loss diagnosis.

Ref	Objective of study	Models	Key Finding	Gaps/Limitation	Relevance to our work
[7].	The aim of deep learning is to diagnose hair-related diseases more accurately and automatically.	CNN Convolutional neural network (CNN) model was applied to the processing and analysis of a dataset including 150 scalp pictures.	The model achieved high accuracy (96.2% training accuracy and 91.1% validation accuracy) with precision and recall scores for alopecia, psoriasis, and folliculitis, suggesting efficacy.	Limitations include a small sample size, limitations in dataset variety, and minimal current research in this topic.	This work, which presents comparable approaches and difficulties, is in line with your focus on deep learning for the diagnosis of scalp diseases.
[8].	Provide a precise way to distinguish between healthy hair and injured scalps.	Using pictures of both healthy and ill scalps, the VGG-19 model, a convolutional neural network (CNN), is used in this study to train and test the categorization of various hair conditions.	Using a machine learning algorithm linked to hair fall illnesses, it provides a 90% accurate categorization of healthy hair and damaged scalp.	Difficulties in acquiring a variety of datasets could impact the accuracy and generalizability of the model.	Corresponds with your emphasis on the identification and categorization of scalp diseases using machine learning.
[10]	To use deep learning to create a non-distressing, effective solution for the early identification and detection of dermatological disorders that impact the hair and scalp.	Convolutional Neural Networks (CNNs) are used in this study to analyze dermatological diseases for images.	98% accuracy in identifying scalp disorders is achieved, for showing the usefulness of CNNs in the processing of medical images.	Short explanation of the quantity and distinction of the dataset, possible over fitting of the model, and absence of clinical validation across a range of demographics.	It focuses on the technical features and thorough process of the system, by showcasing this innovation's transformative potential for patient care and delivery, it hopes to improve global healthcare outcomes.
[21]	It provides the differences of two Optimized CNN's with existing models for affected hair	CNN Model with the datasets of healthy hair images are applied for the analysis and classifying hair	While Applying CNN Algorithm, it provides 95% accuracy for identifying hair disease AA (Alopecia Areata).	It may not show the full range of Alopecia Areata changes and other hair conditions and while relying on one CNN Model should need to improve classification accuracy.	This indicates that the suggested algorithms like CNN provides an actual framework for Alopecia Areata Classification.

Ref	Objective of study	Models	Key Finding	Gaps/Limitation	Relevance to our work
	images.	disease Alopecia Areata			
[27]	To examine the relationship between the trichoscopic characteristics of female pattern hair loss (FPHL) and the degree of hair loss	Clinical and microscopic examination of 110 patients; results were compared with controls; 89% accuracy was achieved with deep learning and machine learning models.	It was discovered that Trichoscopic characteristics such as hair diameter diversity, white spots, tiny scaling, and honeycomb coloring were important markers.	Only Trichoscopic observations; lacks hormonal and histological analysis	Similar to your project on scalp condition diagnosis, this one is pertinent to the use of trichoscopy and AI models for early hair loss detection.
[34]	Will conduct clinical trials to examine the NX35growthTM molecule's efficacy in hair growth	Hair density and thickness were analyzed using deep learning models, and data was gathered from 51 subjects	With an accuracy of 85%, NX35growthTM serum improved hair density, volume, and thickness following 28 and 56 days of use.	Small participant numbers, no control group, and hazy long-term results	Important to your investigation of the effects of deep learning models and scalp treatments on hair development, which is connected to your concentration on the identification of hair diseases
[14]	It gives the accuracy while reviewing datasets.	Machine Learning model with SVM, KNN and Random Forest for the diagnosis of hair fall disorder in the community of Bangladesh people.	By Applying SVM, KNN and Random Forest it provides 92%, 90% and 84% accuracy respectively.	The Models while estimating accuracy doesn't tells that which specific metrics (e.g Precision, Recall, F1 Score) to determine the model performance.	It illustrates the arrangement of techniques based on model performance in identifying hair fall, improving diagnostic precision and effectiveness.

TABLE II

SOURCE TABLE

S. No	Source Name
S1	"Androgenetic alopecia" [1].
S2	"Preprocessing with image denoising and histogram equalization for endoscopy image analysis using texture analysis" [2].
S3	"Genetic prediction of male pattern baldness based on large independent datasets" [3].
S4	"Alopecia areata: review of epidemiology, clinical features, pathogenesis, and new treatment options" [4].
S5	"Leveraging deep neural networks to uncover unprecedented levels of precision in the diagnosis of hair and scalp disorders" [5].
S6	"Classification framework for healthy hairs and alopecia areata: a machine learning (ml) approach" [6].
S7	"Intelligent Healthcare Platform for Diagnosis of Scalp and Hair Disorders" [7].
S8	"Hair follicle classification and hair loss severity estimation using mask R-CNN" [8].
S9	"Hair and scalp disease detection using deep learning" [9].
S10	"A mobile device-based hairy scalp diagnosis system using deep learning techniques" [10].
S11	"A Machine Learning Algorithm Applied to Trichoscopy for Androgenic Alopecia Staging and Severity Assessment" [11].
S12	"The prediction of hair fall pattern in a person using artificial intelligence for better care and treatment" [12].
S13	"Effects of Exposure Time to Sun on Hair Fall During Lockdown in Covid Pandemic" [13]

The work has so far been done on hair fall detection. Some studies use datasets from dermatology clinics, hair loss dataset and other relevant sources, employing machine learning and deep learning tactics.

B. FEATURE TABLE

The deep learning model of identifying scalp diseases and hair fall contains essential features to enhance performance and user interaction in applications. A comprehensive review of studies highlights key factors for effective hair fall prediction, with each feature contributing to better saving plans. Table-3 summarizes these obtained features, offering valuable awareness into the methodologies used in deep learning and machine learning for sending this issue.

TABLE III

FEATURE TABLE

F#	Name	Description
FT1	Machine Learning	The paper employs the concept of machine learning in its research meth-odology [2]-[4].
FT2	Deep Learning	The research methodology in the paper integrates the use

F#	Name	Description
		of Deep Learning [6], [7], [10].
FT3	Dataset Images	Image Datasets are used for training the model like CNN, ML Models etc. [7], [9], [21].
FT4	Video Detecting	Hair disease through video detecting and wireless cameras for product directions [11], [17].
FT5	Sensors	Microscope sensors to evaluate hair and scalp status through feature images [11].
FT6	Google Map	Users can easily access Google Maps to search for hair fall information on their smart phones [16], [17].
FT7	Desktop / Web App	This provides Web App so any person can check its Hair fall by simply uploading an image [17], [19]
FT8	Mobile Application	A React Native mobile app can detect and predict hair loss for users [12], [17].
FT9	Camera	The webcam for feature images and then Using Deep Learning Models [11], [17].
FT10	KNN Algorithm	KNN is used to evaluate accuracy while reviewing datasets [4], [14].
FT11	Augmentation	While Working on the images with Deep Learning Models the hairfall/scalp image dataset increases [4], [7], [9].
FT12	Color Segmentation	The Model check the features including color, texture of user from the uploaded image [2], [5], [6].
FT13	CNN Algorithm	It is used in the hair fall detection as it provides 98% accuracy while working on image datasets [1], [6], [20]

TABLE IV
MAPPING TABLE (SOURCE AND FEATURES)

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	Proposed Work
FT1	×	✓	✓	✓	×	×	×	×	×	×	×	×	×	✓
FT2	×	×	✓	×	×	✓	✓		×	✓	×	×		✓
FT3	×	✓	×	✓	✓	✓	✓	×	✓	✓	×	✓	×	✓
FT4	✓	✓	×	×	×	×	×	×	×	×	✓	×	×	×
FT5	✓	×	×	×	×	×	×	×	×	×	✓	×	×	×
FT6	✓	×	×	×	×	×	✓	×	×	✓	×	×	×	×
FT7	×	✓	✓	✓	✓	×	×	×	×	×	×	×	✓	×
FT8	×	✓	✓	×	✓	×	✓	×	✓	×	×	✓	×	✓
FT9	✓	✓	×	×	×	×	×	×	✓	×	✓	×	✓	×
FT10	×	✓	×	✓	×	×	×	✓	×	×	×	×	×	✓
FT11	✓	✓	×	✓	✓	×	×	×	✓	×	×	✓	×	✓
FT12	×	✓	×	×	×	×	×	×	✓	×	×	✓	×	✓
FT13	✓	×	×	✓	✓	✓	×	×	✓	✓	×	×	×	✓

C. MAPPING TABLE

Table-4 displays the connection between numerous research roots and the features in the proposed deep learning model for diagnosing hair loss and scalp conditions. It highlights the advantages and disadvantages of previous studies by mapping specific features and showing how they increase the model's functionality. This comparison also helps to identify the unique features of our research, displaying patterns and connections among different machine learning and deep learning projects in hair loss detection.

V. METHODOLOGY

The Hair Fall and Scalp Disease Detection project engages a structured methodology consisting of several steps. Initially, the hair fall dataset is sourced from Kaggle and arranged for training, validation, and test sets. Data preprocessing is critical for effective model training, involving the handling of missing values both by replacement or removal—and the application of min-max normalization to enhance accuracy and accelerate linking between them. Following preprocessing, the dataset is dividing into training and testing blocks, and various models, including 'Artificial Neural Networks (ANN)', 'Convolutional Neural Networks (CNN)', and 'Recurrent Neural Networks (RNN)', are trained on this data. The performance of these models is then calculated using the testing blocks. The ANN is managed through a Multi-Layer Perceptron (MLP), which processes the training data, computes errors, initializes, and updates weights by using back propagation until best possible results are achieved. RNNs are especially utilized for time-series predictions due to their ability to maintain information from previous outputs. To conquer challenges such as disappearing gradients, Long Short- Term Memory (LSTM) networks are implemented, manipulating gating methods to functionally maintain memory and improve prediction accuracy.

The study aims to address the difficulty and cost associated with diagnosing hair-related diseases like alopecia, psoriasis, dandruff, and other scalp conditions. Traditional methods for diagnosing hair diseases often require multiple consultations, and expensive testing, and are time-consuming. Therefore, there is a need for an automated, accessible, and efficient system that can diagnose these diseases early, reducing the time for treatment and improving patient outcomes.

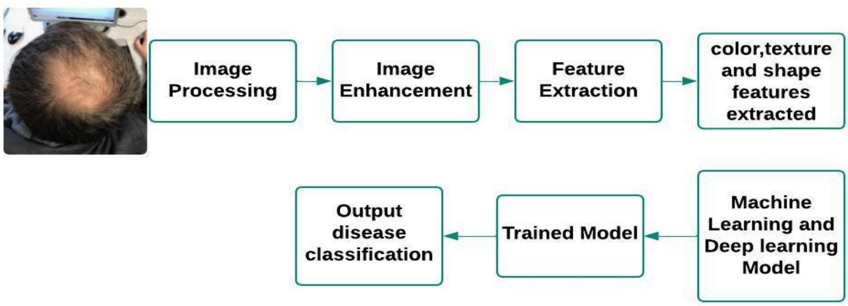

FIGURE 2. Hair fall forecasting framework

Figure 2 illustrates a proposed methodology for classifying diseases using image analysis techniques. It involves preprocessing, enhancement, feature extraction, and deep learning algorithms to classify images into different disease categories.

A. DATASET DESCRIPTION

The dataset of hair fall has been obtained from the Kaggle separated has been used in this study. Hairfall data has been used for prediction purposes. This dataset, which includes pictures of different hair and scalp states, is a useful tool for deep learning and medical image analysis applications. Ten main illnesses are included in the dataset. A predetermined amount of photos from each of these categories can be utilized to train deep learning models and perform classification tasks. The photos are suitable for jobs involving automatic diagnosis of scalp ailments since they have undergone pre-processing utilizing techniques including denoising and augmentation. The 12000 photos in the collection are distributed among several disease categories but have to train 9600 images. To make training models on particular conditions easier, the dataset is arranged so that each disease has a folder with the corresponding photos. Deep learning models focusing on the early diagnosis and categorization of scalp illnesses can be developed with this dataset. It can help dermatologists identify patients more quickly and accurately. All diseases contain 960 images data

- Alopecia Areata [44].
- Contact Dermatitis [44].
- Folliculitis, it has also
- Head Lice
- Lichen Planus
- Male Pattern Baldness
- Psoriasis
- Seborrheic dermatitis
- Telogen Effluvium
- Tinea Capitis

A great number of hair and scalp diseases [29] are included in the model, they are characterized by different manifestations and causes. In alopecia areata, the immune system is directed towards attacking the hair follicles, and hair loss is therefore patchy. Changed skin or hair treatment triggers contact dermatitis, which produces redness and itchiness on the scalp. Folliculitis is an inflammation of the hair follicles in which small red pimples are commonly observed on the skin of the scalp. The signs of head lice infestation are itching and the presence of pest lesions such as head lice or eggs. Lichen plans is an inflammatory disease, which can manifest as tiny, inflammatory papules on the scalp that can lead to hair loss at times. Pattern for Men Baldness is a genetic disease that only occurs in men and leads to normal balding pattern by gradual hair loss. Maps of red lesion on the scalp that is scaly, itchy or uncomfortable is an indicator of psoriasis. Seborrhoeic dermatitis therefore qualifies as a fairly ordinary condition for which one can develop dandruff, as well as oily, scaly skin. Telogen Stress or hormonal changes lead to many hair follicles to shift to what is known as resting phase, causing effluvium, a temporary form of hair loss. Finally, tinea capitis is a fungal infection of the scalp that causes circular bald patches with scaling in most cases. This diversity of the diseases demonstrates how effectively the model may help in diagnosing several types of scalps and hair health problems [30]-[32].

B. DATA PREPROCESSING AND LOADING

Data preprocessing is essential for enhancing the accuracy of hair fall detection models. In this study, missing values in the dataset were either dropped or replaced with the mean of the respective features. The dataset is organized and fed into the models for training, validation, and testing, ensuring that the images are resized to the required input shape of 512x512. Data scaling and normalization are applied to standardize the data, with min-max normalization used to scale attributes between 0 and 1. This normalization helps to ensure faster combination during model training, enhancing overall performance. The Equation 1, to normalize input data effectively, is mentioned below.

$$y_{normalized} = \frac{y - y_{min}}{y_{max} - y_{min}} \quad (1)$$

Here, y stands for normalized data, y is the actual hair fall data value that needs to be normalized, y_{min} for the minimum hair fall data value, and y_{max} for the maximum hair fall value, respectively.

C. TRAINING AND TESTING DATA

The hair fall dataset is split into test and training sets, with the later consisting of more recent data and the former of older data. Prior to training the model, ANN, CNN, and Simple RNN is, followed by further training with LSTM networks. Its performance is then evaluated on the test set to analyse prediction accuracy.

D. MODELS UTILIZED IN THE STUDY

This study inspects many algorithms for hair fall detection as shown in fig 3, deploying both machine learning and deep learning techniques. Each algorithm is tested with different features to achieve ideal accuracy. A detailed description of each model is used more and more for hair fall detection and analysis, because machine learning (ML) and deep learning (DL) algorithms can analyse and interpret complicated data patterns.

Structured data, such as patient demographics, hair loss trends, and clinical parameters, are analyzed using machine learning techniques like logistic regression, decision trees, random forests, and support vector machines. By training on labelled datasets, these algorithms are able to detect correlations and forecast the likelihood of hair loss, allowing dermatologists to evaluate each patient's unique risk factors and suggest the best course of action.

Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), two types of deep learning techniques, are very good at processing unstructured input, such as sequences and images. While RNNs can be used for time-series data to analyse patterns in hair fall over time, CNNs are especially good at analyzing scalp photos to find indications of hair thinning or loss by learning spatial hierarchies in the data. By identifying complex patterns in huge datasets, these models increase classification precision and offer insightful information about disorders of the hair and scalp.

In this study, the researchers implemented ML and DL procedures to improve dermatology's diagnostic capabilities, allowing for more precise and individualized evaluations of scalp disorders and hair loss [33], [34].

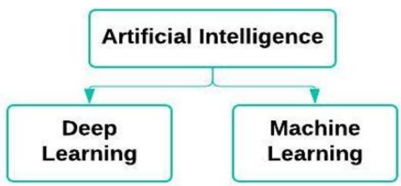

FIGURE 3. Hairfall disease detection AI

Techniques

E. MACHINE LEARNING MODELS

Machine learning (ML) is a part of artificial intelligence (AI) that influences data and algorithms to increase accuracy by copying human learning. In this study on hair fall detection, three predicting models are applied, which are K-Nearest Neighbors (KNN), Random Forest, and Logistic Regression. Each model is designed to analyze data successfully and improve detection results over time. The study underlines that computer-aided diagnosis raises the accuracy of alopecia areata prediction and classification through machine learning models [35], [36].

1) K –NEAREST NEIGHBORS (KNN) MODEL

K-Nearest Neighbors (KNN) is a simple, non-parametric, and lazy learning algorithm used for classification and regression tasks [36]. It identifies a new data point based on the majority class of its k nearest neighbors in the feature space. The most commonly used distance metric in KNN is the Euclidean distance, given by:

$$\sqrt{\sum_{l=1}^{n}(x_{il} - x_{jl})2} \quad (2)$$

Where,

- x_i and x_j are two data points with n features.
- x_{il} *and* x_{jl} are the l-th features of the points xi and xj.
- The Distance Metric can also be used as $d(x_i, x_j) = \sum_{l=1}^{n} |x_{il} - x_{jl}|$

Working: When a user uploads a photo, the backend extracts features from it using the same method as during training, shown in fig 4. The KNN algorithm then calculates distances between these features and those in the training dataset to identify the 'k' nearest neighbors. It allocates the most common label among these neighbors to the image. Finally, the classification result is sent back to the mobile app for user display.

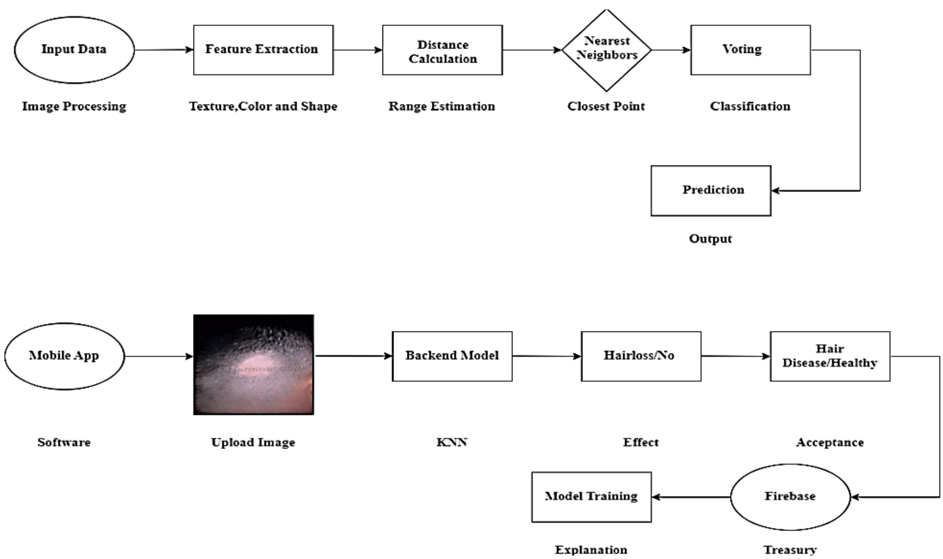

FIGURE 4. KNN model working & project execution

2) LOGISTIC REGRESSION:

Logistic regression is a numerical method that is used for binary classification tasks shown in fig 5. It is a supervised learning algorithm that predicts the chances that a given input belongs to a definite class. Inspite of its name, logistic regression is used for classification rather than reversal. In logistic regression, the model is represented as a linear combination of input features, followed by the application of the sigmoid function to map the output to a probability. The logistic regression model outputs a probability p that the input belongs to class 1. This is given by the sigmoid of the linear combination shown in equation (3 and 4).

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \qquad (3)$$

Where,

β_0I is the intercept (bias term),

$\beta_1, \beta_2, \dots, \beta_n$ Are the coefficients (weights) associated with the input features$x_1, x_2, \cdots x_n$.

Z is the linear predictor.

$$p = \sigma(z) = \frac{1}{1+e^{-z}} \qquad (4)$$

σ (z) represents the sigmoid function and z is the input feature of the linear combination.

Working: The user uploads an image through the React Native app, which is then prepared, and features like texture, color and patterns are extracted, potentially using a CNN Model. These features are augmented into a logistic regression model, trained on labeled scalp images, to predict the probability of hair loss or specific hair diseases based on the output possibilities. The uploaded images, along with the model's predictions are stored in a database.

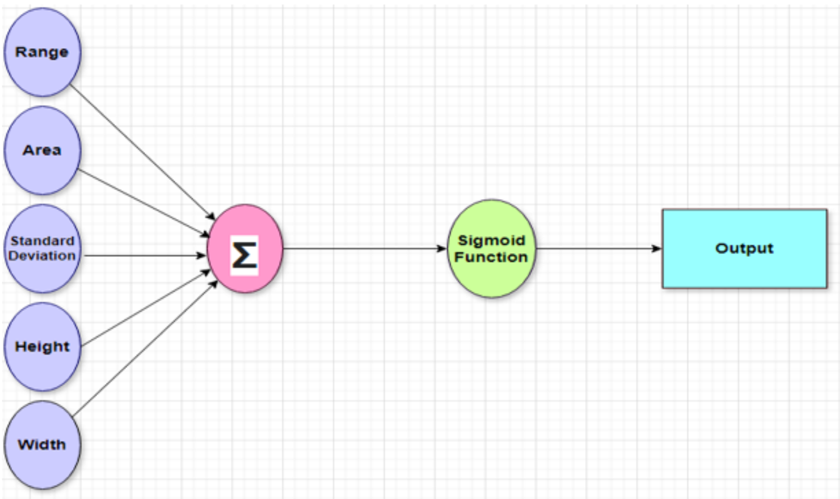

FIGURE 5. Logistic Regression Working & Project Execution

3) RANDOM FOREST

Random Forest

Random Forest is an object learning method used for classification and reversal tasks. It works by constructing multiple decision trees during training and returning the mode (classification) or mean (regression) of the individual trees' predictions. The variance reduction criterion is used to measure how well a split reduces the variance of the target values, shown in equation (5).

$$Variance\ Reduction = Var(D) - \frac{N_L}{N}Var(D_L) - \frac{N_R}{N}Var(D_R) \quad (5)$$

Where,

- Var (D) is the variance of the target values in the dataset D,
- D_L and D_R are the variances in the left and right child nodes
- N_L and N_R are the number of samples in the left and right child nodes,
- N is the total number of samples.

Working: Users upload photos of their scalp or hair through the React Native app, which preprocesses these images by changing size, normalization, and extracting features like texture and color [37]-[39]. These features are used to train a Random Forest model with labeled data showing hair loss or disease. When a new image is inspected, the model predicts the existence and type of hair condition based on the extracted features, with results displayed to the user, showing both hair loss, a specific disease, and a normal condition, as shown in fig 6.

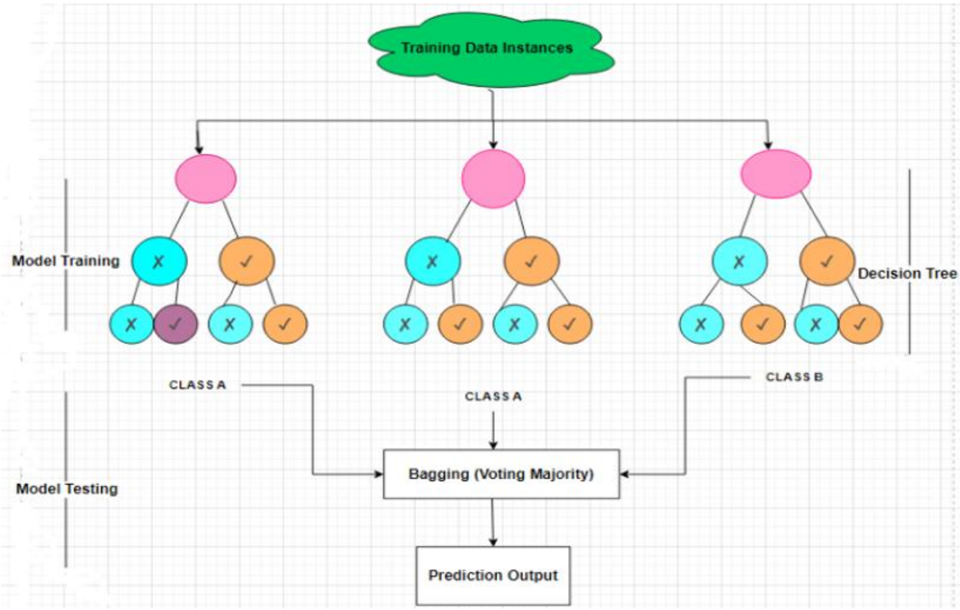

FIGURE 6. Random forest working & project execution

F. DEEP LEARNING MODELS

Deep learning, a part of machine learning, depends on neural networks and representation learning, designing inspiration from organic neurology. The process involves collecting artificial neurons in layers and training them to process data helpfully. In this study on hair fall detection, several deep learning techniques are utilized to increase prediction accuracy. Deep Learning can automatically identify hair loss stages from frontal photos, training to advance diagnosis and treatment [40]-[43].

1) CONVOLUTIONAL NEURAL NETWORK MODEL

A Convolutional Neural Network (CNN) is also used for a deep learning method that processes network data, such as images, by learning geometric features through convolutional layers. It is skillful for image recognition and object detection tasks. The mathematical formula, for the convolution operation applied to an output pixel at position yi, j can be expressed as shown in equation (6) and (7):

$$yi, j = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x_{i+mj+n \cdot w_{m,n}} + \text{b} \qquad (6)$$

Where,

-x is the input image (referred to as the output, from the preceding layer.),

-w is the filter (also known as the kernel),

-b is the subjective term,

-y is the resulting feature chart.

In a fully connected layer, each neuron is connected to every neuron in the previous layer. This layer classically emerges around the end of the network.

$$y = Wx + b \qquad (7)$$

Where,

The weight matrix is represented by the letter W, x represents the (compressed feature chart) input vector, b is the subjective term, y is the output.

Working: A CNN algorithm is trained on scalp and hair fall images that examines user-uploaded photos through a mobile application (Based on React Native) to identify hair conditions and provide personalized treatment suggestions.

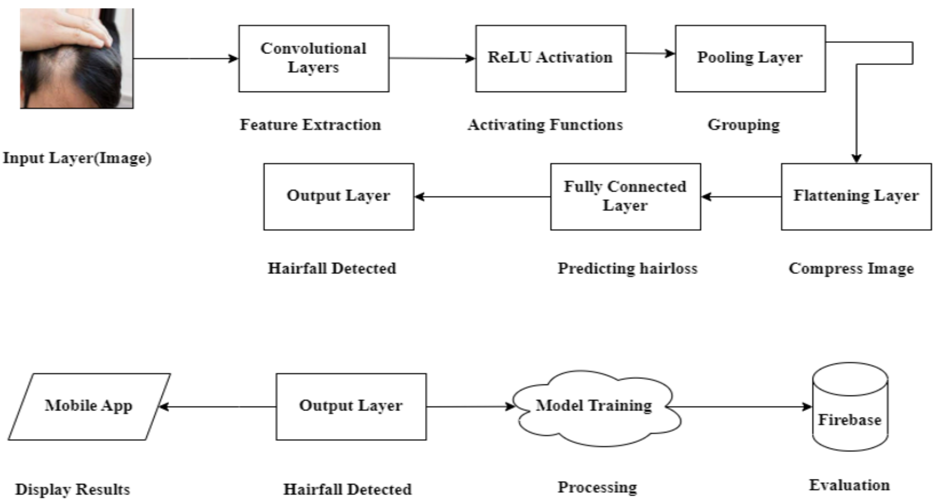

FIGURE 7. CNN model working & project execution

2) LONG SHORT-TERM MEMORY (LSTM) MODEL

RNNs face issues of blowing and ending gradients, which can delay learning. To address this, Gated Recurrent Units (GRUs) and Long Short-Term Memory (LSTM) networks are used, with LSTM being particularly effective for time series prediction, including hair fall prediction in this study. LSTM reduces the fading gradient problem and offers improved accuracy over traditional RNNs. Its architecture includes three gates, that is: an input gate, an output gate, a forget gate; and a cell state, as illustrated in Figure 19.

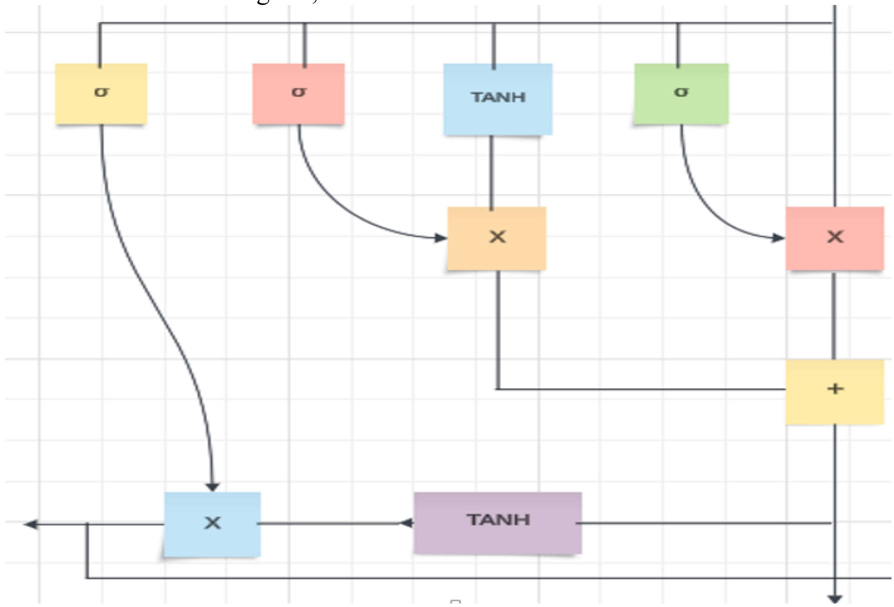

FIGURE 8. Structure of LSTM unit

The Forget Gate decides which data should be removed from the block as it is no longer needed. The output is transmitted through the sigmoid activation to the cell state, after the forget gate f(t) multiplies the input and the output of the previous state by the corresponding weights.

The input gate determines which input values should be written to the memory state. Using sigmoid activation, the input gate (it) processes the input from the earlier time stamps as well as the present input.

The working for an LSTM unit are as shown in equation (8) to (12)

Forget Gate:

$$f(t) = \text{sigmoid}(W_f * [h(t-1), x(t)] + bf) \quad (8)$$

Input Gate:

$$i(t) = \text{sigmoid}(Wi * [h(t-1), x(t)] + bi) \ \& \ \hat{C}(t) = \tanh(W_c * [h(t-1), x(t)] + b_c) \quad (9)$$

Update Cell State:

$$c(t) = f(t) * c(t-1) + i(t) * \hat{C}(t) \quad (10)$$

Output Gate:

$$o(t) = \text{sigmoid}(W_o * [h(t-1), x(t)] + b_o)) \quad (11)$$

Hidden State:

$$h(t) = o(t) * \tanh(c(t)) \quad (12)$$

3) ARTIFICIAL NEURAL NETWORK (ANN) MODEL:

An Artificial Neural Network (ANN) is a numerical model influenced by the way the human brain processes information. It consists of interconnected neurons that work in parallel. There are three major types of neural network architectures: Single-layer feed-forward (with only input and output layers), Multi-layer feed-forward (with input, hidden, and output layers), and Recurrent Neural Networks. Multi-layer Perceptron (MLP), a type of multi-layer feed-forward network, was used in this study. MLPs are typically trained using the algorithm for back propagation, as shown in fig 8.

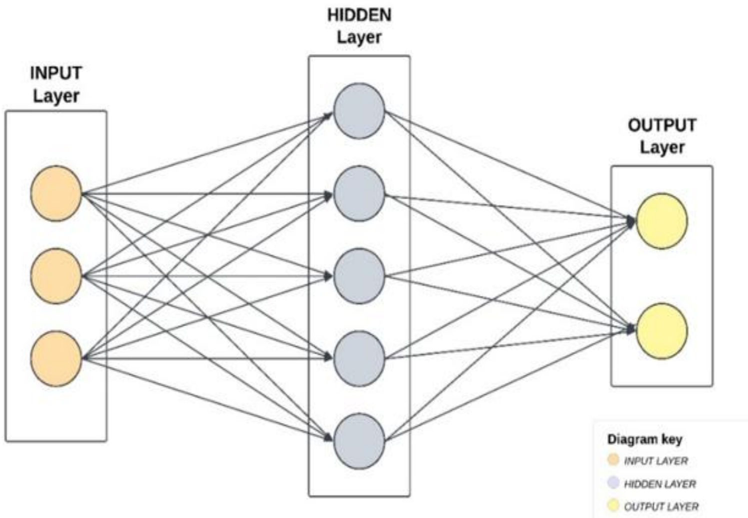

FIGURE 9. Representation of artificial neural network

4) RECURRENT NEURAL NETWORK (RNN) MODEL

In feed-forward neural networks (FFNNs), each output is independent and doesn't keep previous information. In contrast, Recurrent Neural Networks (RNNs) can remember past outputs, making them ideal for time series predictions. In RNNs, the input for the current state (Ct) is both the new input and the output from the previous time step (Ct-1). For the next state (Ct+1), it combines the new input and the previous output. RNNs learn using back propagation through time, as shown in Equation (13) .

$$h_t = g\ (h_t\text{-}1,\ x_t\ ) \qquad (13)$$

When xt is a new input at time step t, g is a recursive function, and ht-1 represents the output from the previous state. The current state is indicated by h_t. Equation 14 provides the formula for using the activation function.

$$h_t = tanh\ (w_h\ \ h_{t-1} + w_x\ \ x_t) \qquad (14)$$

Here, weight at the input neuron is represented by w_x, and weight at the recurrent neuron by h_t. Equation 15 contains the output calculation formula

$$= w_y h_t \qquad (15)$$

Where w_y stands for weight at the output neuron and y_t stands for output.

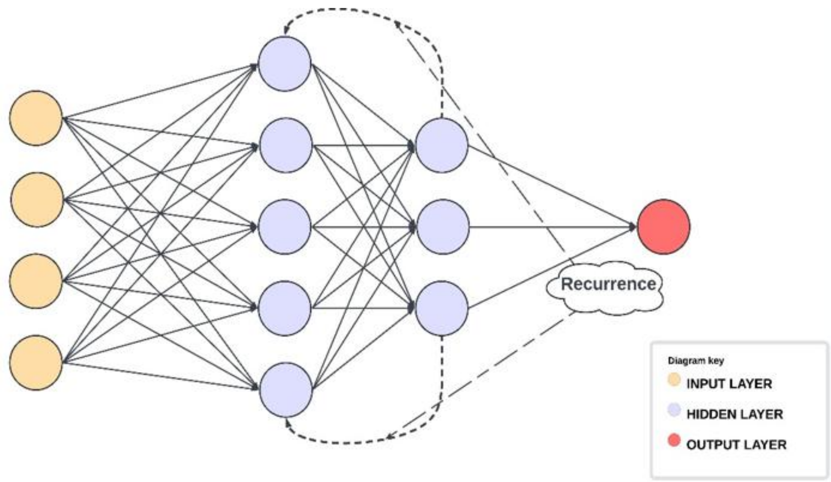

FIGURE 10. Recurrent neural network

VI. RESULTS AND DISCUSSION

In this study, machine learning and deep learning models are utilized for hair fall disease classification prediction. Metrics like AUC (Area under Curve), F1 score, Precision, and Recall are used in Table 5 to summarize the model's findings. The Logistic Regression model performs well in class distinction, as evidenced by its greatest AUC value of 0.975. K-Nearest Neighbor (KNN), on the other hand, performs best in terms of F1 score (0.852), precision (0.859), and recall (0.853), indicating that it strikes a good compromise between precision and true positive rates. While KNN performs exceptionally well in F1, Precision, and Recall, Logistic Regression offers the best AUC overall, making it maybe the better option for balanced performance across measures.

The hair fall disease classification dataset was used to evaluate the performance of many deep learning models and activation functions, shown in table 6. The findings from each model CNN, RNN, LSTM, and ANN were different in terms of metrics including precision, recall, and F1-score. The Tanh activation function in Model 2 (RNN) produced the best results in terms of F1-score, recall, accuracy, and precision. In particular, it had an F1-score of 0.08, a precision of 0.0667, a recall of 0.15, and an accuracy of 15%. This suggests that compared to other model-activation combinations, RNN with Tanh performed better at classifying cases of hair fall illness for this dataset. The output range of the Tanh function (-1 to 1) probably offered a superior gradient for RNN's sequential data processing skills, assisting in more accurately capturing the temporal dependencies in the data. The performance of other models was subpar. For instance, Model 1 (CNN) continuously obtained poor accuracy and minimal precision, recall, and F1-scores (0.1 accuracy across all activations), irrespective of the activation function (ReLU, Tanh, or Sigmoid). With low metrics across activations, Models 3 (LSTM) and 4 (ANN) likewise demonstrated poor efficacy. Particularly in Model 3 (LSTM), the sigmoid activation function performed poorly, yielding zero scores for every metric. This could be because of problems such as disappearing gradients that interfered with the model's ability to learn. The figures 12, 13 and 14 describe the model's performance in graphical representation by using evaluation metrics f score and precision.

The CNN model achieves an overall accuracy of 87.97% with a test loss of 0.82, indicating strong performance in classifying a variety of disorders connected to the hair and scalp, as shown in table 7, describing various hair diesases. With F1-scores ranging from 0.80 to 0.92, it performs especially well for hairfall-specific conditions like Telogen Effluvium, Male Pattern Baldness, and Alopecia Areata, suggesting that it may accurately differentiate these conditions. Furthermore, it exhibits extremely high recall and precision for classes such as head lice and contact dermatitis, indicating that it can detect these cases with few false positives or negatives. The model may, however, occasionally overlook cases of Alopecia Areata (0.71) and Seborrhoeic Dermatitis (0.96), as evidenced by its somewhat poorer recall for these illnesses. Increasing the dataset or improving the model may be necessary to address this. Overall, the findings demonstrate the model's great potential as a diagnostic tool for hair loss-related conditions, supporting accurate and efficient classification.

TABLE V
MACHINE LEARNING MODEL RESULT TABLE

no	Model name	AUC	F1	Precision	Recall
1.	K nearest neighbors	0.953	0.852	0.859	0.853
2.	Random forest	0.873	0.563	0.589	0.562
3.	Logistic regression	0.975	0.812	0.820	0.853

TABLE VI
DEEP LEARNING MODEL COMPARISON TABLE ON HAIRFALL DATASET DISEASE CLASSIFICATION

Mode no	Model name	Activation	Accuracy	Precision	Recall	F1-Score
Model 1	CNN	RELU	0.10	0.020000	0.10	0.03
		TANH	0.10	0.010000	0.10	0.01
		SIGMOID	0.10	0.010000	0.10	0.01
Model 2	RNN	RELU	0.10	0.025000	0.10	0.03
		TANH	0.15	0.066667	0.15	0.08
		SIGMOID	0.15	0.045098	0.15	0.00
Model 3	LSTM	RELU	0.10	0.010000	0.10	0.01
		TANH	0.10	0.023611	0.10	0.03
		SIGMOID	0.00	0.000000	0.000	0.000
Model 4	ANN	RELU	0.10	0.010000	0.10	0.018
		TANH	0.10	0.010526	0.10	0.019
		SIGMOID	0.10	0.041026	0.10	0.053

TABLE VII
CLASSIFICATION REPORT MODEL TESTING (CNN)

Condition	Precision	Recall	F1 score	Support
Alopecia Areata	0.93	0.71	0.80	227
Contact Dermatitis	1.00	1.00	1.00	125
Folliculitis	0.93	0.73	0.82	201
Head Lice	1.00	1.00	1.00	125
Lichen Planus	0.97	0.90	0.93	124
Male Pattern Baldness	0.95	0.90	0.92	135
Psoriasis	0.85	0.98	0.91	126
Seborrheic Dermatitis	0.66	0.96	0.78	125
Telogen Effluvium	0.79	0.95	0.86	134
Tinea Capitis	0.84	0.90	0.87	124

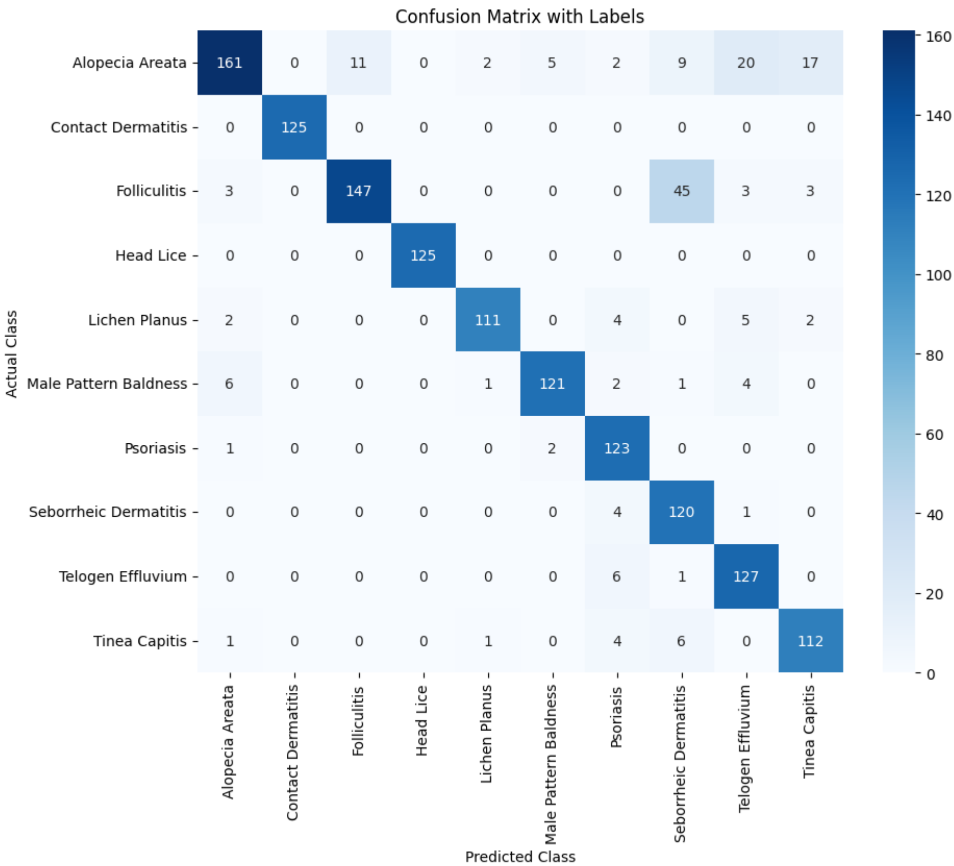

FIGURE 11. Confusion matric of various diseases

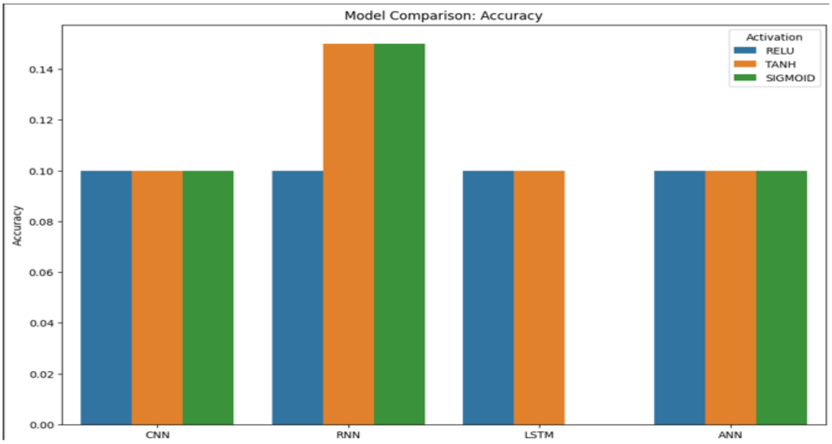

FIGURE 12. Deep learning model comparison by applying different activation function

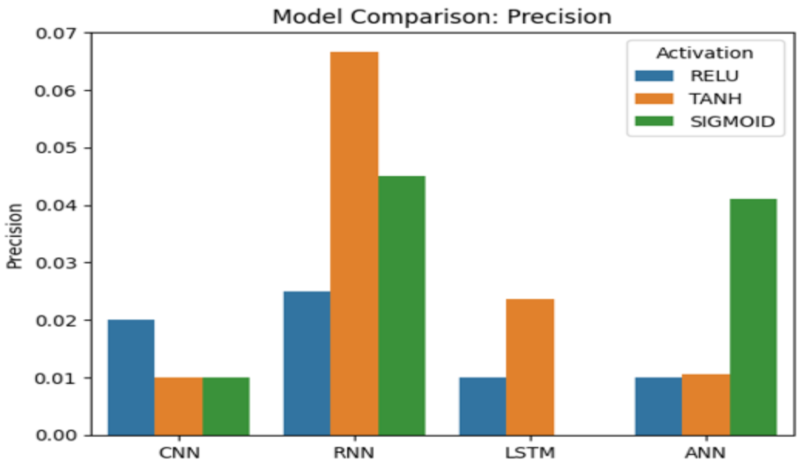

FIGURE 13. Model Evaluation precision

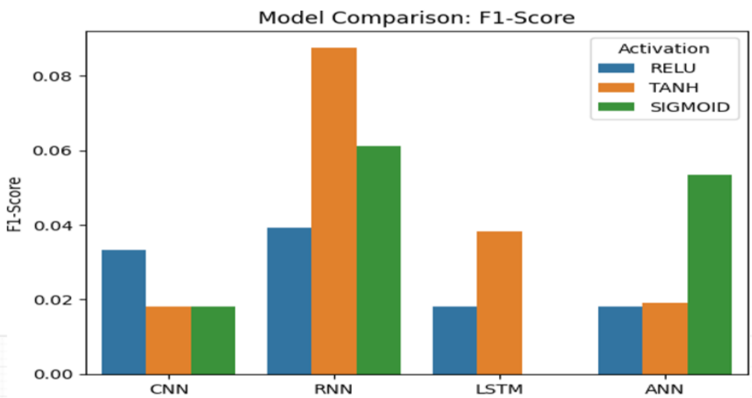

FIGURE 14. Model Comparison F1-Score

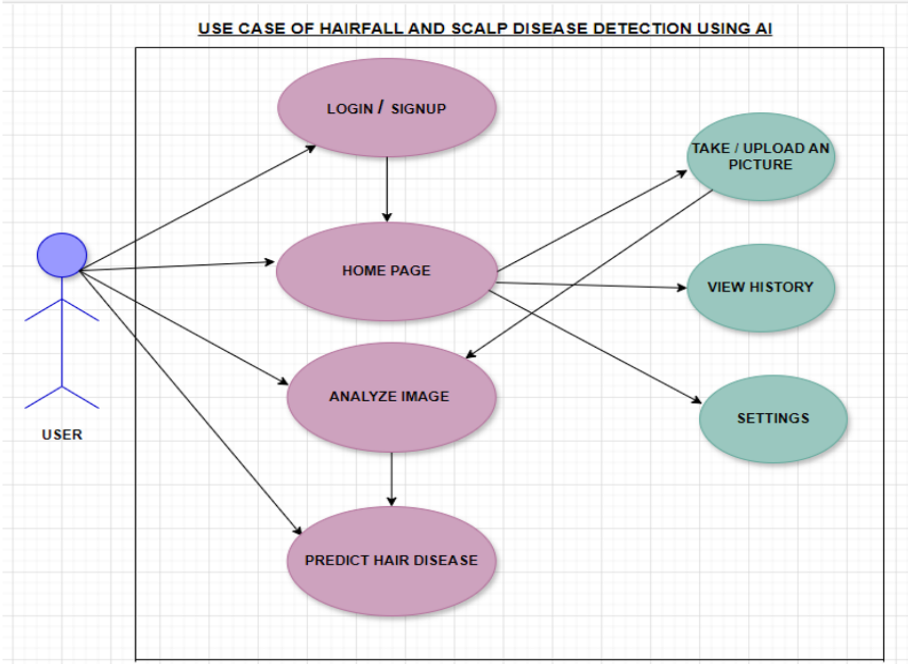

FIGURE 15. Use case diagram of the study

The primary interactions in an AI-powered hair loss and scalp disease diagnosis system are depicted in this use case diagram figure 15. The user accesses the home page by first joining up or logging in. The user can access a number of functions from the home page. They can use the "Analyze Image" tool to process an image that they have taken or uploaded for examination. Based on the study, the system can forecast possible conditions affecting the hair and scalp, and the results are shown in the "Predict Hair Disease" section. Users can also change settings for personalization and check their history of analysis. In addition to allowing image-based disease prediction and history monitoring, this structure offers an intuitive user experience.

VII. EVALUATION METRICS

Performance metrics are measurements that indicate how successfully a model is operating on a particular task. In this study, we have utilized the following performance metrics:

A. ACCURACY.

All it measures is the frequency with which the classifier makes accurate predictions. The ratio of the number of accurate forecasts to the total number of predictions (see equation 16) can be used to determine accuracy.

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \quad (16)$$

The "Number of Correctly Classified Instances" indicates the number of data records that the algorithm correctly classified. The "Total Number of Instances" parameter indicates the total number of data records in the dataset.

B. PRECISION (POSITIVE PREDICTIVE VALUE)

It explains why a large number of cases

that were correctly predicted turned out to be positive. Precision is useful in the scenarios illustrated in equation 17, where False Positives are more problematic than False Negatives.

$$Precision = \frac{TP}{TP+FN} \quad (17)$$

C. RECALL (SENSITIVITY, TRUE POSITIVE RATE)

This prescribes how many actual positive cases our model was able to predict with absolute certainty. Recall comes is handy whenever False Positive is of higher concern than False Negative (as shown in equation 18).

$$Recall = \frac{TP}{TP+FP} \quad (18)$$

7.4. F1 Score

The F1 score is the harmonic mean of precision and recall, providing a balance between the two metrics shown in equation 19.

$$F1\,Score = \frac{2 * (Precision * recall)}{Precision + recall} \quad (19)$$

D. CONCLUSION AND FUTURE RECOMMENDATION

The study shows how deep learning and Artificial Intelligence (AI) can greatly improve the precision and effectiveness of identifying scalp conditions and hair loss. We discovered that AI-driven methods provide significant diagnostic precision by evaluating a number of models, including Random Forest, K-Nearest Neighbors (KNN), Logistic Regression, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Artificial Neural Network (ANN). With a remarkable Area under the Curve (AUC) of 0.975, which indicates strong performance in class distinction, logistic regression was found to be the most effective model for class separation. Furthermore, the RNN model addressed the usual vanishing gradient issue in recurrent networks and produced reliable classification predictions, especially when combined with a TANH activation function. These results highlight the potential of AI models to provide trichologists and dermatologists with trustworthy diagnostic instruments, with the ultimate goal of enhancing patient care and facilitating individualized treatment strategies. In order to increase model generalization, future studies can expand on these findings by expanding the dataset and adding a wider variety of scalp diseases and hair loss patterns. To further increase diagnostic accuracy and lessen potential biases, transfer learning and ensemble techniques could be investigated. By incorporating real-time picture analysis from dermatoscopes or mobile devices, these AI models may be able to reach underserved areas and enable remote diagnoses. Last but not least, imprfoving the interpretability of models—for example, by implementing explainable AI—would increase the predictability and transparency of the results for medical professionals, encouraging a wider use of AI-driven diagnostic tools in clinical settings. The goal of this future path is to create a thorough, useful, and extremely accurate AI-based framework that facilitates early, accurate diagnosis, and enables prompt, focused therapies for health conditions relating to hair and scalp.

CONFLICT OF INTEREST

The author of the manuscript has no financial or non-financial conflict of interest in the subject matter or materials discussed in this manuscript.

DATA AVAILABILITY STATEMENT

Data will be provided by corresponding author upon reasonable request.

FUNDING DETAILS

This research received no external funding.

REFERENCES

[1] T. Hiroyasu, K. Hayashinuma, H. Ichikawa, and N. Yagi, "Preprocessing with image denoising and histogram equalization for endoscopy image analysis using texture analysis," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 789–792, doi: https://doi.org/10.1109/EMBC.2015.7318480.

[2] Y. Chen, P. Hysi, C. Maj, S. Heilmann-Heimbach, T. D. Spector, F. Liu, and M. Kayser, "Genetic prediction of male pattern baldness based on large independent datasets," *Eur. J. Hum. Genet.*, vol. 31, no. 3, pp. 321–328, Mar. 2023, doi: https://doi.org/10.1038/s41431-022-01201-y.

[3] E. Darwin, P. A. Hirt, R. Fertig, B. Doliner, G. Delcanto, and J. J. Jimenez, "Alopecia areata: Review of epidemiology, clinical features, pathogenesis, and new treatment options," *Int. J. Trichol.*, vol. 10, no. 2, pp. 51–60, Apr.–Jun. 2018, doi: https://doi.org/10.4103/ijt.ijt_99_17.

[4] R. Behal, P. Priya, A. Kapoor, and A. Mishra, "Hair loss stage prediction using deep learning," in *Proc. IEEE Int. Conf. Computing, Power and Commun. Technol. (IC2PCT)*, Feb. 2024, pp. 1263–1267, doi: https://doi.org/10.1109/IC2PCT60090.2024.10486442.

[5] N. Jain, B. Doshi, and U. Khopkar, "Trichoscopy in alopecias: Diagnosis simplified," *Int. J. Trichology*, vol. 5, no. 4, pp. 170–178, Oct.–Dec. 2013, doi: https://doi.org/10.4103/0974-7753.130385.

[6] I. Al-Aradi and M. Al-Ghareeb, "Hair fall: Common causes and treatment modalities," *Treatment*, vol. 16, no. 1, pp. 9–15, 2005.

[7] C. H. Ho, T. Sood, and P. M. Zito, "Androgenetic alopecia," in *StatPearls*. Treasure Island, FL, USA: StatPearls Publishing, 2017.

[8] M. S. Chowdhury *et al.*, "Leveraging deep neural networks to uncover unprecedented levels of precision in the diagnosis of hair and scalp disorders," *Skin Res. Technol.*, vol. 30, no. 4, 2024, Art. no. 13660, doi: https://doi.org/10.1111/srt.13660.

[9] C. S. Shakeel, S. J. Khan, B. Chaudhry, S. F. Aijaz, and U. Hassan, "Classification framework for healthy hairs and alopecia areata: A machine learning approach," *Comput. Math. Methods Med.*, vol. 2021, 2021, doi: https://doi.org/10.1155/2021/1102083.

[10] K. Sultanpure, B. Shirsath, B. Bhande, H. Sawai, S. Gawade, and S. Samgir, "Hair and scalp disease detection using deep learning," *arXiv preprint*, arXiv:2403.07940, 2024.

[11] C. Ha, T. Go, and W. Choi, "Intelligent healthcare platform for diagnosis of scalp and hair disorders," *Appl. Sci.*, vol. 14, no. 5, 2024, Art. no.1734 doi: https://doi.org/10.3390/app14051734.

[12] M. Di Fraia *et al.*, "A machine learning algorithm applied to trichoscopy for androgenic alopecia staging and severity assessment," *Dermatol. Pract. Concept.*, vol. 13, no. 3, 2023, Art. no. 2023136, doi: https://doi.org/10.5826/dpc.1303a136.

[13] W. J. Chang *et al.*, "A mobile device-

based hairy scalp diagnosis system using deep learning techniques," in *Proc. IEEE 2nd Global Conf. Life Sciences and Technologies (LifeTech)*, Mar. 2020, pp. 145–146, doi: https://doi.org/10.1109/LifeTech48969.2020.1570617332.

[14] S. A. Farooq, A. Ali, and A. Bashir, "The prediction of hairfall pattern in a person using artificial intelligence for better care and treatment," in *Proc. 4th Int. Conf. Inno. Pract. Technol. Manag.*, Feb. 2024, pp. 1–6.

[15] S. K. Khokhar, A. Qamar, and Y. Mahar, "Effects of exposure time to sun on hair fall during lockdown in COVID pandemic," *J. Gandhara Med. Dent. Sci.*, vol. 10, no. 4, pp. 64–67, 2023.

[16] Y. Yacoob and L. S. Davis, "Detection and analysis of hair," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 7, pp. 1164–1169, Jul. 2006, doi: https://doi.org/10.1109/TPAMI.2006.139.

[17] D. S. Park, J. E. Koo, S. Y. Oh, and H. J. Choi, "Physiological profiles of hair and scalp conditions in Korea population: Survey, hairological expert-diagnosis and new diagnostic device," *Adv. Bioeng. Biomed. Sci. Res.*, vol. 4, no. 2, pp. 63–71, 2021, doi: https://doi.org/10.1080/09546634.2024.2337908.

[18] S. Y. Jhong, P. Y. Yang, and C. H. Hsia, "An attention-based expert inspection system for smart scalp," in *Proc. Asia-Pacific Signal and Information Processing Association Annual Summit and Conf. (APSIPA ASC)*, Dec. 2021, pp. 1678–1681.

[19] X. Liu, C. H. Chen, M. Karvela, and C. Toumazou, "A DNA-based intelligent expert system for personalised skin-health recommendations," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 11, pp. 3276–3284, Nov. 2020, doi: https://doi.org/10.1109/JBHI.2020.2978667.

[20] C. R. Kothari and S. Patil, "Trichoscopic features in female pattern hair loss: A 1-year hospital-based cross-sectional study," *Clin. Dermatol. Rev.*, vol. 8, no. 2, pp. 95–101, 2024, doi: https://doi.org/10.4103/cdr.cdr_123_21.

[21] C. S. Shakeel and S. J. Khan, "Machine learning techniques as effective methods for evaluating hair and skin assessments: A systematic review," *Proc. Inst. Mech. Eng. H*, vol. 238, no. 2, pp. 132–148, 2024, doi: https://doi.org/10.1177/09544119231216290.

[22] C. Saraswathi and B. Pushpa, "Ab-mtedeep classifier trained with AAGAN for the identification and classification of alopecia areata," *Eng., Technol. Appl. Sci. Res.*, vol. 13, no. 3, pp. 10895–10900, 2023, doi: https://doi.org/10.48084/etasr.5852.

[23] A. Egger, M. Tomic-Canic, and A. Tosti, "Advances in stem cell-based therapy for hair loss," *CellR4 Repair Replace. Regen. Reprog.*, vol. 8, 2020, Art. no. 2894.

[24] W. C. Wang, L. B. Chen, and W. J. Chang, "Development and experimental evaluation of machine-learning techniques for an intelligent hairy scalp detection system," *Appl. Sci.*, vol. 8, no. 6, 2018, Art. no. 853, doi: https://doi.org/10.3390/app8060853.

[25] E. K. Ross, C. Vincenzi, and A. Tosti, "Videodermoscopy in the evaluation of hair and scalp disorders," *J. Amer. Acad. Dermatol.*, vol. 55, no. 5, pp. 799–806, Nov. 2006, doi: https://doi.org/10.1016/j.jaad.2006.04.058.

[26] R. Nagar and R. Dhudshia, "Utility of trichoscopy to diagnose early female pattern hair loss in resource-poor setting: A cross-sectional study," *Indian J. Dermatol. Venereol. Leprol.*, vol. 85, 2019, Art. no. 681.

[27] J. H. Kim, S. Kwon, J. Fu, and J. H. Park, "Hair follicle classification and hair loss severity estimation using Mask R-CNN," *J. Imaging*, vol. 8, no. 10, 2022, Art. no. 283, doi: https://doi.org/10.3390/jimaging8100283.

[28] D. Banerjee, V. Kukreja, D. Bordoloi, and A. Choudhary, "Enhanced hair disease classification using deep learning," in *Proc. 11th Int. Conf. Signal Proc. Integr. Networks*, Mar. 2024, pp. 58–63.

[29] L. P. Liu, M. A. Wariboko, X. Hu, Z. H. Wang, Q. Wu, and Y. M. Li, "Factors associated with early-onset androgenetic alopecia: A scoping review," *PLoS One*, vol. 19, no. 3, Mar. 2024, Art. no. 0299212, doi: https://doi.org/10.1371/journal.pone.0299212.

[30] R. Xiao and L. N. Lee, "Updated review of treatment of androgenetic alopecia," *Facial Plast. Surg. Clin.*, vol. 32, no. 3, 417–423, 2024, doi: https://doi.org/10.1016/j.fsc.2024.02.006.

[31] N. Natarelli, N. Gahoonia, and R. K. Sivamani, "Integrative and mechanistic approach to the hair growth cycle and hair loss," *J. Clin. Med.*, vol. 12, no. 3, Jan. 2023, Article no. 893, doi: https://doi.org/10.3390/jcm12030893.

[32] B. M. Piraccini *et al.*, "Efficacy and safety of topical finasteride spray solution for male androgenetic alopecia: A phase III randomized controlled clinical trial," *J. Eur. Acad. Dermatol. Venereol.*, vol. 36, no. 2, pp. 286–294, Feb. 2022, doi: https://doi.org/10.1111/jdv.17738.

[33] C. C. Chen, M. V. Plikus, P. C. Tang, R. B. Widelitz, and C. M. Chuong, "The modulatable stem cell niche: Tissue interactions during hair and feather follicle regeneration," *J. Mol. Biol.*, vol. 428, no. 7, pp. 1423–1440, Apr. 2016, doi: https://doi.org/10.1016/j.jmb.2015.07.009.

[34] P. Bouhanna, "Multifactorial classification of male and female androgenetic alopecia," *Dermatol. Surg.*, vol. 26, no. 6, pp. 555–561, Jun. 2000.

[35] L. Srinivasan, A. Jeevika, R. K. Navina, and S. Priyadharshini, "An enhanced stress-based hairfall detection and prevention using KNN and machine learning techniques," in *Proc. 7th Int. Conf. Trends Electr. Inf.*, Apr. 2023, pp. 1110–1115.

[36] H. Pant, G. Joshi, J. Pant, D. Singh, S. Maurya, and H. R. Goyal, "Advancing dermatology: Deep convolutional neural networks for timely hair disease diagnosis," in *Proc. IEEE 3rd World Conf. Appl. Intell. Comput.*, Jul. 2024, pp. 761–765.

[37] O. O. Ayanlowo, "Scalp and hair disorders at the dermatology outpatient clinic of a tertiary hospital," *Port*

Harcourt Med. J., vol. 11, no. 3, pp. 127–133, 2017, doi: https://doi.org/10.4103/phmj.phmj_32_16.

[38] F. Canpolat, "Infections, infestations and neoplasms of the scalp," in *Hair and Scalp Disorders*, Z. Kutlubay and S. Serdaroglu, Eds., 2017, pp. 199–219.

[39] H. Wolff, T. W. Fischer, and U. Blume-Peytavi, "The diagnosis and treatment of hair and scalp diseases," *Dtsch. Ärztebl. Int.*, vol. 113, no. 21, pp. 377–386, May 2016, doi: https://doi.org/10.3238/arztebl.2016.0377.

[40] A. Rajab *et al.*, "Flood forecasting by using machine learning: A study leveraging historic climatic records of Bangladesh," *Water*, vol. 15, no. 22, Nov. 2023, Art. no. 3970, doi: https://doi.org/10.3390/w15223970.

[41] H. Farman, A. W. Khan, S. Ahmed, D. Khan, M. Imran, and P. Bajaj, "An analysis of supervised machine learning techniques for churn forecasting and component identification in the telecom sector," *J. Comput. Biomed. Inform.*, vol. 7, no. 1, pp. 264–280, 2024.

[42] H. Farman, D. Khan, S. Hassan, M. Hussain, and S. A. A. Usmani, "Analyzing machine learning models for forecasting precipitation in Australia," *J. Comput. Biomed. Inform.*, vol. 7, no. 1, pp. 439–458, 2024.

[43] R. Happle, K. J. Kalveram, U. Büchner, K. Echternacht-Happle, W. Göggelmann, and K. H. Summer, "Contact allergy as a therapeutic tool for alopecia areata: Application of squaric acid dibutylester," *Dermatology*, vol. 161, no. 5, pp. 289–297, 1980, doi: https://doi.org/10.1159/000250380.

[44] H. Farman, S. Ahmed, M. Imran, Z. Noureen, and M. Ahmed, "Deep learning-based bird species identification and classification using images," *J. Comput. Biomed. Inform.*, vol. 6, no. 1, pp. 79–96, 2023.