ISSN 2664-4150 (Print) & ISSN 2664-794X (Online)
South Asian Research Journal of Engineering and Technology
Abbreviated Key Title: South Asian Res J Eng Tech

| Volume-7 | Issue-5 | Nov-Dec- 2025 | **DOI: https://doi.org/10.36346/sarjet.2025.v07i05.001**

Original Research Article

A Survey and Comparison of Machine Learning-Based Skin Cancer Detection Methods

G. Mahalaxmi[1]*, Dudekula Raziya[2]

[1]Assistant Professor (Ph.D), Department of ECE, Ashoka Women's Engineering College, NH-44, Dupadu, Kurnool-518218
[2]Assistant Professor (Ph.D), G.Pullaiah College of Engineering and Technology, Kurnool

***Corresponding Author:** G. Mahalaxmi
Assistant Professor (Ph.D), Department of ECE, Ashoka Women's Engineering College, NH-44, Dupadu, Kurnool-518218

Article History
Received: 12.10.2025
Accepted: 03.12.2025
Published: 06.12.2025

Abstract: This research looks at different ways computers can automatically spot skin cancer from images of skin problems. They tested two main approaches. The first one pulls out important details from images using basic tools and then feeds those details into various machine learning programs to make a diagnosis. The second approach uses more advanced deep learning methods like convolutional neural networks that can both find the important image details and make the diagnosis all in one go. The researchers tried out several traditional machine learning methods including Logistic Regression, k-Nearest Neighbors, Naive Bayes, Decision Tree, Random Forest, and Support Vector Machine. They also tested four different CNN models: VGG16, ResNet60, InceptionV3, and Inception-ResNetV2. For the first approach, they used four different ways to pull out image features and combined them together to get the best possible feature set, which should lead to better accuracy in spotting cancer. To test how well these methods work, they used a collection of skin lesion images from the ISIC database that includes both harmless and cancerous skin conditions. They measured how effective each algorithm was using standard performance metrics and also kept track of how long each one took to train, giving them insight into both accuracy and computational demands.

Keywords: Convolutional Neural Networks, Image Processing, Machine Learning, Skin Cancer and Automatic Detection.

1. INTRODUCTION

Melanoma is an aggressive type of skin cancer that spreads quickly and can be fatal, affecting many people and causing most deaths from skin cancer [1]. Cancer spots on the skin usually grow slowly over time, but people often miss them or don't pay attention because they don't know what to look for. This happens because skin marks can come from many different causes [2].

Most people have what doctors call pigmented spots, which are typically even on both sides, flat, have smooth edges, and are all one colour. Whether you get these spots depends on your genes. When kids have birthmarks, these marks usually get bigger as they grow up. Around age 35, these skin spots should stop changing - they shouldn't get bigger, smaller, or look different. So if you notice any changes after this age, it's something to worry about since melanoma tends to hit younger and middle-aged people, unlike other cancers that mostly affect older folks [3].

Just because you get a new spot on your skin as an adult doesn't automatically mean its cancer, but you should definitely have a skin doctor check it out to make sure it's not skin cancer. Finding cancerous tumours as early as possible helps doctors start treatment right away and gives patients the best chance of making a complete recovery. The first step in figuring out what's going on is usually a dermoscopic exam, which is pretty straightforward and doesn't hurt at all. Basically, doctors use a special tool to get a really close look at any suspicious spots on your skin, magnifying them about 10 to 20 times so they can see the structure and colours clearly. There's also something called video dermoscopy, where they can actually record what they're seeing, save those images, and then check back later to see if anything has changed over time [4].

Automated skin cancer diagnosis can help people get their results quicker and begin treatment right away, or it can show that a growth is harmless so they can decide against having it removed - which might be a minor procedure, but it's still uncomfortable and comes with some risks [5]. Scientists frequently write about using computers to automatically or semi-automatically diagnose illnesses, including cancer. For example, there's research on automatically spotting whether breast tumours are harmless or dangerous. Doctors also use computer-based diagnostic tools powered by math formulas in other medical areas, like figuring out delays in pulse waves, determining what stage of sleep someone is in by looking at their brain wave patterns, or understanding someone's mental condition by examining specific details pulled from their brain activity recordings [6, 7].

Machine learning and AI systems are becoming increasingly popular for automatically identifying diseases. These smart algorithms have already shown they work really well for sorting and classifying different types of information - like figuring out if someone is male or female from their photo, identifying people through multiple biometric checks, spotting objects as they happen in real-time, or catching spam emails before they reach your inbox. The medical and pharmaceutical fields are also jumping on board with these AI-powered tools [8]. You can now find them being used in medicine dispensing machines that need to recognize different pills and medications, or in healthcare apps that analyze how people walk to help doctors make better diagnoses [9].

This research presents two different ways to tackle automated melanoma detection. Traditional machine learning methods that don't use deep learning aren't really popular for analyzing digital images anymore because they don't perform as well and can't scale up easily. But here's the thing - these older algorithms often work better when you're dealing with smaller amounts of data, plus they're much simpler to build and don't need as much computing power or memory since they have fewer parameters to work with [10]. On the flip side, deep neural networks need really powerful graphics cards to crunch all those numbers in a reasonable timeframe. This creates expensive costs and makes it tough to put these neural network solutions on everyday devices like smart phones [11].

This research makes several key contributions:

- We looked at how machine learning methods can be used to classify cancer lesions from medical images and examined how well different approaches work for this problem [12].
- We studied various ways of describing image features and tested different combinations to see which ones help the classification algorithms perform better, measuring success using standard classification metrics.
- We compared different algorithms to understand how their computational demands affect their classification performance, particularly when these algorithms need to run on mobile devices that don't have much processing power [13].

2. Related Works

Right now, the top approach for recognizing and sorting digital images, especially when it comes to spotting and diagnosing skin cancer, involves deep neural networks called CNNs - that's Convolutional Neural Networks. These networks work with groups of convolutional filters that take input images with pretty big dimensions but shallow depth - typically just 3 channels for red, green, and blue - and turn them into tensors, which are basically deep feature maps that capture what's important in the image.

The researchers in reference [17] suggest using a CNN architecture that has four convolutional layers arranged in pairs, with max pooling layers between each pair. Once they flatten the feature vector, they use two fully connected layers to do the classification. The way they set up their network, along with how they pre processed and segmented the skin lesions, helped them achieve 81% accuracy.

Transfer learning has emerged as another promising approach discussed in research studies. Researchers take convolutional neural networks that have already been trained on large datasets and then tweak the classification component to make the network work for new classification tasks. This method helps achieve better validation results even when working with smaller amounts of data. One study brought up the crucial concern about diagnosing skin cancer across different skin tones, an area that hasn't gotten enough attention in current research, and they explored these using networks like Squeeze Net. Another group of researchers worked with the GoogLeNet network and managed to hit 93.2% accuracy, while yet another study looked at how various neural network designs perform when tackling multi-class classification challenges.

This paper also explores using machine learning algorithms to detect skin cancer. This method requires pulling out image features first using various descriptors. In one study, researchers tested different classifiers like Support Vector

Machine, Multilayer perceptron, k-nearest neighbours, and Adaboost, with Adaboost performing the best at 93% accuracy. Another study looked at several algorithms for diagnosing breast cancer and found that Decision Tree and Random Forest worked best when they tested them.

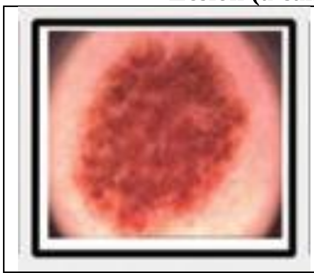
The researchers in reference [24], came up with a hybrid approach. They took advantage of a pre-trained ResNet model [25], to pull out deep features and then used Fisher Vector Encoding to combine all the local deep features into one overall image representation. They also threw in Principal Component Analysis (PCA) to cut down on the size of these deep feature vectors. In the end, they went with SVM for classifying the data, which helped them achieve a classification accuracy that was better than 86%.

3. MATERIALS

The image information we used for our experiments came from the ISIC database, which stands for The International Skin Imaging Collaboration. ISIC is basically a global partnership that brings together researchers and companies working to get better at diagnosing melanoma using advanced digital imaging of the skin. Their main goal is to save lives since melanoma is much easier to treat when it's caught early. They've created a public image collection that keeps growing, and it's designed to help people develop and test new automated systems for diagnosing skin conditions. The project involves several organizations, like the International Dermoscopy Society and the International Society for Digital Imaging of the Skin.

This study worked with training and testing datasets that included two types of skin lesions - those that are harmless and those that are cancerous. For training, they used 2637 images split roughly evenly between the two types, and for testing, they had 660 images. All the images were sized at 224 × 244 pixels in both datasets. You can see examples of both types of skin lesions in Table1.

Table 1: Example images of skin cancer: (a) shows a Benign Lesion (non-cancerous spot), (b) shows Melanoma Lesion (a cancerous spot)

	
Benign Lesion	Melanoma Lesion

4. METHOD

Comparison of Machine Learning Algorithms for Skin Cancer Detection:

Table 2: Comparison between the Random Forest Method and other existing procedures (Convolutional Neural Network (CNN), Naïve Bayes (NB) etc.,). Abbreviations: Logistic Regression (LR), Support Vector Machine (SVM), k-Nearest Neighbors (k-NN), Decision Tree (DT)

Algorithm	Type	Key Features Used	Advantages	Limitations	Accuracy (Typical Range)	Common Use Case / Dataset
Logistic Regression (LR)	Supervised (Classification)	Color, texture, shape features	Simple, interpretable, low computational cost	Limited for nonlinear data	75–85%	Small tabular datasets or handcrafted features
Support Vector Machine (SVM)	Supervised (Classification)	Texture (GLCM), shape, color histograms	Effective for high-dimensional data; robust to overfitting	Requires careful kernel & parameter tuning	80–90%	ISIC, PH2 datasets
k-Nearest Neighbors (k-NN)	Supervised (Instance-based)	Color, texture	Simple, no training phase	Computationally expensive for large data	70–85%	Small sample datasets

Algorithm	**Type**	**Key Features Used**	**Advantages**	**Limitations**	**Accuracy (Typical Range)**	**Common Use Case / Dataset**
Decision Tree (DT)	Supervised (Classification)	Texture and shape	Easy interpretation, fast	Overfits easily	70–80%	Feature-based classification
Random Forest (RF)	Ensemble (Bagging)	Combined texture, color	High accuracy, reduces overfitting	Less interpretable	85–92%	ISIC, HAM10000 datasets
Gradient Boosting / XGBoost	Ensemble (Boosting)	Combined handcrafted features	High predictive power	Sensitive to noise, parameter tuning needed	88–94%	Tabular image features or pre-trained embeddings
Naïve Bayes (NB)	Probabilistic	Color and shape	Fast, works well with small data	Assumes independence among features	70–80%	Early skin lesion datasets
Artificial Neural Networks (ANN)	Deep / Shallow Neural Net	Raw or preprocessed pixels	Learns complex relationships	Needs large datasets	85–92%	ISIC / HAM10000
Convolutional Neural Networks (CNN)	Deep Learning	Raw image pixels	Automatic feature extraction, high accuracy	Requires large labeled datasets, high computation	90–98%	ISIC 2019, HAM10000, Derm7pt
Transfer Learning (e.g., ResNet, VGG16, InceptionV3, EfficientNet)	Deep Learning (Pretrained CNN)	Raw images (fine-tuned)	Excellent accuracy, low training time	Large model size, needs GPU	93–99%	ISIC, HAM10000
Hybrid / Optimized CNN (CNN+SVM / CNN+RF / CNN+Feature Fusion)	Hybrid Deep + Classical ML	CNN extracted features	Combines strengths of both methods	Complex architecture	95–99%	Custom or clinical datasets

The Comparison of Machine Learning Algorithms for Skin Cancer Detection is displayed in the Table 2.

5. CONCLUSION

This research looked at how different machine learning methods perform when automatically figuring out whether skin lesions are harmless or cancerous. The researchers fed the computer algorithms information about the lesions' shape, colour, and texture to help them make these decisions. For each method they tested, they worked out which combination of features worked best to get the most accurate results and catch the most actual cases of cancer in their test samples. It turned out that the Hybrid / Optimized CNN (CNN+SVM / CNN+RF / CNN+Feature Fusion) came out on top, and interestingly, it actually worked better than Convolutional Neural Networks that had been specially adjusted for this particular dataset.

REFERENCES

1. Davis, L.E.; Shalin, S.C.; Tackett, A.J. Current state of melanoma diagnosis and treatment. *Cancer Biol. Ther.* 2019, *20*, 1366–1379.
2. MacGill, M. What to Know about Melanoma. Medical News Today. Available online: https://www.medicalnewstoday.com/articles/154322 (accessed on 15 September 2022).
3. Heistein, J.B.; Archarya, U. Malignant Melanoma. StatPearls. Available online: https://www.statpearls.com/ArticleLibrary/viewarticle/24678 (accessed on 15 September 2022).
4. Blundo, A.; Cignoni, A.; Banfi, T.; Ciuti, G. Comparative Analysis of Diagnostic Techniques for Melanoma Detection: A Systematic Review of Diagnostic Test Accuracy Studies and Meta-Analysis. *Front. Med.* 2021, *8*, 637069.

5. Silva, T.A.E.d.; Silva, L.F.d.; Muchaluat-Saade, D.C.; Conci, A.A. Computational Method to Assist the Diagnosis of Breast Disease Using Dynamic Thermography. *Sensors* 2020, *20*, 3866.
6. Fanizzi, A.; Losurdo, L.; Basile, T.M.A.; Bellotti, R.; Bottigli, U.; Delogu, P.; Diacono, D.; Didonna, V.; Fausto, A.; Lombardi, A.; et al. Fully Automated Support System for Diagnosis of Breast Cancer in Contrast-Enhanced Spectral Mammography Images. *J. Clin. Med.* 2019, *8*, 891. [
7. Sieczkowski, K.; Sondej, T.; Dobrowolski, A.; Olszewski, R. Autocorrelation algorithm for determining a pulse wave delay. In Proceedings of the 2016 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), Poznan, Poland, 21–23 September 2016; pp. 321–326. [Google Scholar] [CrossRef]
8. Hasan, M.J.; Shon, D.; Im, K.; Choi, H.-K.; Yoo, D.-S.; Kim, J.-M. Sleep State Classification Using Power Spectral Density and Residual Neural Network with Multichannel EEG Signals. *Appl. Sci.* 2020, *10*, 7639.
9. Hasan, M.J.; Kim, J.-M. A Hybrid Feature Pool-Based Emotional Stress State Detection Algorithm Using EEG Signals. *Brain Sci.* 2019, *9*, 376.
10. Lin, Y.; Xie, H. Face Gender Recognition based on Face Recognition Feature Vectors. In Proceedings of the 2020 IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE), Dalian, China, 27–29 September 2020; pp. 162–166.
11. Alay, N.; Al-Baity, H.H. Deep Learning Approach for Multimodal Biometric Recognition System Based on Fusion of Iris, Face, and Finger Vein Traits. *Sensors* 2020, *20*, 5523
12. Kanimozhi, S.; Gayathri, G.; Mala, T. Multiple Real-time object identification using Single shot Multi-Box detection. In Proceedings of the 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Gurugram, India, 6–7 September 2019; pp. 1–5.
13. Rapacz, S.; Chołda, P.; Natkaniec, M. A Method for Fast Selection of Machine-Learning Classifiers for Spam Filtering. *Electronics* 2021, *10*, 2083.
14. Jianu, S.R.S.; Ichim, L.; Popescu, D. Automatic Diagnosis of Skin Cancer Using Neural Networks. In Proceedings of the 2019 11th International Symposium on Advanced Topics in Electrical Engineering (ATEE), Bucharest, Romania, 28–30 March 2019; pp. 1–4.
15. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. *J. Big Data* 2016, *3*, 1–40.
16. Pillay, V.; Hirasen, D.; Viriri, S.; Gwetu, M. Melanoma Skin Cancer Classification Using Transfer Learning. In *Advances in Computational Collective Intelligence. ICCCI 2020. Communications in Computer and Information Science*; Hernes, M., Wojtkiewicz, K., Szczerbicki, E., Eds.; Springer: Cham, Switzerland, 2020.
17. Vasconcelos, C.N.; Vasconcelos, B.N. Convolutional neural network committees for melanoma classification with classical and expert knowledge based image transforms data augmentation. *arXiv* 2017, arXiv:1702.07025.
18. Huang, H.W.; Hsu, B.W.Y.; Lee, C.H.; Tseng, V.S. Development of a light-weight deep learning model for cloud applications and remote diagnosis of skin cancers. *J. Dermatol.* 2020, *48*, 310–316.
19. Sabri, M.A.; Filali, Y.; el Khoukhi, H.; Aarab, A. Skin Cancer Diagnosis Using an Improved Ensemble Machine Learning model. In Proceedings of the 2020 International Conference on Intelligent Systems and Computer Vision (ISCV), Fez, Morocco, 9–11 June 2020; pp. 1–5.
20. Tumuluru, P.; Lakshmi, C.P.; Sahaja, T.; Prazna, R. A Review of Machine Learning Techniques for Breast Cancer Diagnosis in Medical Applications. In Proceedings of the 2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 12–14 December 2019; pp. 618–623.
21. Yu, Z.; Jiang, X.; Zhou, F.; Qing, J.; Ni, D.; Chen, S.; Lei, B.; Wang, T. Melanoma Recognition in Dermoscopy Images via Aggregated Deep Convolutional Features. *IEEE Trans. Biomed. Eng.* 2019, *66*, 1006–1016.
22. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
23. The International Skin Imaging Collaboration. Available online: https://www.isic-archive.com/#!/topWithHeader/wideContentTop/main (accessed on 5 January 2021).
24. Marta Bistron and Zbigniew Piotrowski., Comparision of Machine Learning Algorithms used for Skin Cancer Diagnosis., vol.12, issue 19., 10.3390/app12199960.