ISSN: 2366-1313

# DEEPSKIN: A DEEP LEARNING APPROACH FOR SKIN CANCER CLASSIFICATION

**[1]Mrs. K.Srilatha, [2]A. Bhavana, [3]Ch.Saketh, [4]G.Poojitha**

[1]Assistant Professor, Dept. of CSE, Teegala Krishna Reddy Engineering College, Meerpet, Hyderabad,

kumbamsrilatha@tkrec.ac.in

[2,3,4,]BTech Student, Dept. of CSE, Teegala Krishna Reddy Engineering College, Meerpet, Hyderabad

ayyannagaribhavana@gmail.com, cheekatisakethch@gmail.com, poojithagoshika@gmail.com

## ABSTRACT :

Skin cancer is one of the diseases that spreads around the world most fast because of the scarcity of resources. A precise diagnosis is necessary for the identification of skin cancer in order to implement prevention measures generally. Dermatologists struggle to identify skin cancer in its early stages, and deep learning has been heavily utilized in both supervised and unsupervised learning tasks in recent years. In tests of object detection and classification, Convolutional Neural Networks (CNN), one of these models, has outperformed the rest. The dataset utilized for the experiments is filtered from MNIST: HAM10000, which includes seven distinct types of skin lesions and has a sample size of 10015. Pre-processing methods for the data are used, including segmentation using auto encoders and decoders, sampling and dull razors. The model was trained using transfer learning methods such as DenseNet169 and Resnet50 to provide the desired output. Based on the findings of the experiment, Resnet50's oversampling strategy provided high accuracy and f1-measure, whereas DenseNet169's under sampling technique produced good accuracy and f1-measure. This work will be expanded in the future to incorporate parameter adjustment for improved forecast accuracy.

**Keywords:** Skin cancer, CNN, Data Augmentation, Deep learning.

ISSN: 2366-1313

# I INTRODUCTION

Deep literacy is frequently used, even for unsupervised literacy tasks. Convolutional Neural Networks (CNN) are widely used for object detection and bracketing tasks. As a result, trained end-to-end in a controlled setting.

CNNs eliminate the requirement for people to manually generate feature sets. In recent years, Convolutional Neural Networks (CNNs) have outperformed human experts in categorizing skin cancer lesions. productive. The ceroscopy approach aims to improve melanoma patients' performance. Ceroscopy is a noninvasive skin imaging procedure that magnifies and illuminates the afflicted skin area, eliminating facial reflection. Skin cancer. Early detection is a prized possession. Identifying the difference between benign and malignant skin lesions can be challenging due to their similarities. The sun's UV radiation and UV tanning beds are the leading causes of skin cancer. Dermatologists struggle to identify between melanoma and non-melanoma tumors due to their similar appearance to skin. The fundamental issue with similar opinions is basically It is subjective and difficult to replicate. Robotization and deep literacy can provide early opinion reports, allowing patients to seek therapy from dermatologists. Early detection is critical for skin cancer, as treatment choices are limited. Preventing skin cancer requires precise evaluation and identification of the condition. Deep literacy is frequently used, even for unsupervised literacy tasks. Convolutional Neural Networks (CNN) are widely used for object detection and bracketing tasks. As a result, trained end-to-end in a controlled setting CNNs eliminate the requirement for people to manually generate feature sets. In recent years, Convolutional Neural Networks (CNNs) have outperformed human experts in categorizing skin cancer lesions.

# II. LITERATURE SURVEY

**1. Detection of skin cancer using CNN algorithm:**

https://www.sciencedirect.com/science/article/abs/pii/S0933365719301460

Early detection of skin cancer is crucial for preventing certain types, such as focal cell carcinoma and melanoma. Image processing and machine vision are increasingly used in medical applications. This paper proposes a new image processing-based method for early skin cancer detection using an optimal Convolutional Neural Network (CNN). An

ISSN: 2366-1313

improved wave optimization algorithm is used for optimizing the CNN. The method is compared with other methods on two different datasets, and simulation results show superiority. The proposed method is a promising advancement in skin lesion classification.

2. **Skin lesion classification based on deep convolutional neural networks architectures:**

https://www.researchgate.net/publication/350529813_Skin_Lesion_Classification_Based_on_Deep_Convolutional_Neural_Networks_Architectures

Skin cancer is a primary cancer characterized by various dermatological disorders and can be classified into different types based on morphological features, color, structure, and texture. The mortality rate of skin cancer patients depends on early detection and diagnosis. Current dermoscopic images face limitations like shadow, artifact, and noise, which can hinder detection efforts. Deep learning neural networks have been used to analyze these images for skin cancer detection. This paper reviews the latest deep learning concepts for skin cancer detection and classification.

3. **Skin cancer classification using image processing and machine learning:**

https://www.researchgate.net/publication/350935822_Skin_Cancer_Classification_Using_Image_Processing_and_Machine_Learning

This paper presents a method for classifying skin lesions as benign or malignant using image processing and machine learning. It proposes a new contrast stretching method for dermoscopic images and uses the OTSU thresholding algorithm for segmentation. The paper extracts features like Gray level Co-occurrence Matrix (GLCM), histogram of oriented gradients (HOG) object, and color identification from the segmented images. Principal component analysis (PCA) reduction and Synthetic minority oversampling technique (SMOTE) sampling are used to reduce dimensionality and address class imbalance problems. The feature vector is then standardized and scaled. A wrapper method is used for feature selection, and classifiers like Quadratic Discriminant, SVM, and Random Forest are used. The system achieved a classification accuracy of 93.89% on ISIC-ISBI 2016.

4. **Skin cancer detection using machine learning techniques:**

ISSN: 2366-1313

https://www.researchgate.net/publication/343673685_Skin_cancer_detection_and_classification_using_machine_learning

Skin cancer is a dangerous disease with a high mortality rate due to lack of knowledge about symptoms and prevention. Early detection is crucial to prevent cancer spread. Melanoma, Basal cell carcinoma, and Squamous cell carcinoma are the most hazardous types. This project aims to detect and classify skin cancer using machine learning and image processing tools. Dermoscopic images are used as input, and the Dull razor method removes unwanted hair particles. Gaussian filters smooth the image, while Median filters preserve lesion edges. Color-based k-means clustering is used for segmentation, and statistical and texture features are extracted using Asymmetry, Border, Color, Diameter, (ABCD) and Gray Level Co-occurrence Matrix (GLCM). The experiment was conducted on the ISIC 2019 Challenge dataset, and the Multi-class Support Vector Machine (MSVM) was used for classification, achieving an accuracy of about 96.25.

# III SYSTEM ANALYSIS

## EXISTING SYSTEM

The literature suggests using Convolutional Neural Networks to predict and classify four types of skin lesions. A website is created for real-time use of the model, which predicts the three most common categories of skin lesions for a certain images. The experiment was conducted using the MNIST: HAM10000 dataset, which includes 100,000 annotated images. Another researcher developed an algorithm to detect skin lesions early by extracting features using the ABCD rule, GLCM, and HOG. Preprocessing is used to improve skin lesions quality and clarity are used to reduce artifacts such as skin color, hair, and so on. Geodesic Active Contour (GAC) was used to segment the lesion, allowing for distinct feature extraction. The ABCD scoring system was utilized to extract attributes like symmetry, border, color, and diameter. HOG and GLCM were used to extract textural information. The retrieved features are used to categorize skin lesions as benign or melanoma using machine learning algorithms including SVM, KNN, and Naïve Bayes. The project utilized skin lesion images from the International Skin imaging Collaboration, including 328 benign images and 672 melanoma images.

**Limitations of Existing system**

- The existing model, which classifies only four skin lesions, may not offer a comprehensive diagnosis, potentially leading to misdiagnoses detections of certain skin conditions.
- The existing work's data pre-processing techniques are not as advanced, which could impact the quality and robustness of the model's predictions.
- The existing work might not take full advantage of state-of-the-art transfer learning approaches, potentially affecting its model's accuracy and generalization.
- The existing work might not have explored segmentation, potentially limiting the quality of its features.
- The second existing work focuses only on classifying between benign and melanoma lesions. This limited scope in classification might lead to missed diagnoses or inadequate identification of other significant skin conditions.
- The use of Geodesic Active Contour (GAC) for segmentation in the second existing work might be less accurate or detailed compared to the segmentation approach. This could affect the quality of the extracted features.

## PROPOSED SYSTEM

The study proposes a Convolutional Neural Networks (CNN) model that outperforms others in object detection and classification tests. The dataset screened from MNIST: HAM10000 comprises seven types of skin lesions with a sample size of 10015. Data pre-processing techniques like sampling and segmentation are used, and transfer learning techniques like DenseNet169 and Resnet 50 are used to train the model.

**Proposed system Advantages:**

- Our work utilizes advanced data pre-processing techniques, including sampling, dull razor, and segmentation using auto encoder and decoder.
- This approach is bolstered by a larger dataset, potentially leading to improved model generalization and better performance when compared to smaller datasets used in existing works.
- Furthermore, our dataset may encompass a more diverse set of lesion types, thereby enhancing the model's ability to generalize across a wider range of skin conditions.

- our methodology incorporates techniques such as segmentation using auto encoder and decoder, which can provide more detailed and accurate feature extraction.

## IV IMPLEMENTATION

**Architecture:**

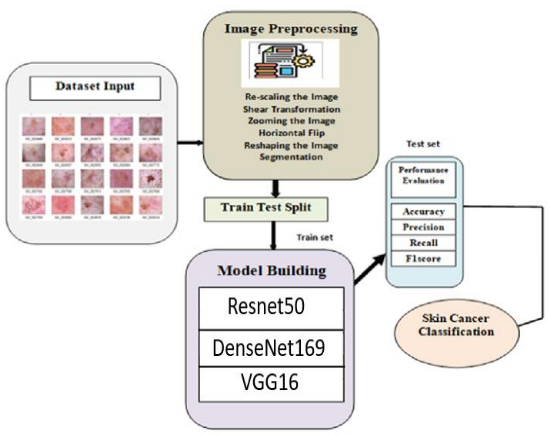



Fig-1. Architectures of the system model

There are various components for system architecture:

**1. Dataset:**

The HAM10000 (Human Against Machine with 10000 training images) dataset is a widely used resource in the field of skin cancer classification. It comprises of 10,015 dermatoscopic images which are categorized into seven different diagnostic categories, including various types of benign and malignant skin lesions.

These are the types of skin lesions present in HAM10000:

1. Actinic keratosis and intraepithelial carcinoma / Bowen's disease (akiec)
2. Basal cell carcinoma (bcc)
3. Benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus like keratoses) (bkl)
4. Dermatofibroma (df)
5. Melanoma (mel)
6. Melanocytic nevi (nv)
7. Vascular lesions (vasc)

This dataset is valuable for training and evaluating machine learning models for skin cancer classification tasks. It has significantly contributed to the advancement of automated skin cancer diagnosis, aiding in the development of deep learning models and computer-aided diagnostic systems. Researchers and practitioners in the field leverage this dataset to develop and validate algorithms that can assist dermatologists in accurately identifying and classifying skin

ISSN: 2366-1313

lesions, thereby improving early detection and patient outcomes in dermatology.

## 2. Image preprocessing:

Image processing is essential for skin cancer classification as it enables the extraction of vital features from images, aiding in the diagnosis and classification of various skin lesions. By enhancing the quality of input data and highlighting relevant details, image processing significantly contributes to the accurate identification of different types of skin lesion and improving the diagnostic process for skin cancer.

Feature extraction are employed to improve the quality of input data. Preprocessing steps, including resizing, normalization, and color space conversion, are used to standardize the images and highlight important visual cues. Through sampling and digital representation, image processing allows for efficient digital analysis and classification of skin cancer, ultimately contributing to the development of reliable machine learning models for automated diagnosis and classification of skin lesions.

- **Segmentation:** This process is critical for accurate classification as it enables the extraction of essential features such as color, texture, and shape characteristics specific to the lesions. Various techniques, including thresholding, edge detection, region growing, and machine learning-based approaches, are employed for effective segmentation. By separating the lesions from the surrounding healthy tissue, segmentation facilitates targeted analysis, improving the accuracy of classification and aiding in the differentiation between benign and malignant lesions. Overall, segmentation plays a vital role in extracting pertinent information from dermatoscopic images, contributing to the development of reliable machine learning models for precise skin cancer classification.

## 3.Model Building:

The dataset is divided into training and testing sets, typically using an 80-20 split, where the training set is used to train the model and the testing set is reserved for evaluating its performance. By using the ResNet50,DenseNet169 and VGG16 it train the model.

ResNet-50, a powerful deep convolutional neural network (CNN) architecture, is instrumental in skin cancer classification due to its exceptional ability to extract intricate features from dermatoscopic images. With 50 layers, including residual blocks, ResNet-50 can effectively capture hierarchical features crucial for discriminating between benign and malignant skin lesions. Its deep architecture enables the automatic learning of discriminative features, such as textural patterns and irregularities, essential for accurate classification. Additionally, the implementation of transfer learning using pre-trained ResNet-50 models, fine-tuned with dermatoscopic images, significantly boosts model convergence and classification accuracy, particularly in scenarios with limited training data. ResNet-50's capability to mitigate the vanishing gradient problem through residual connections allows for effective training of deeper networks, which is vital for learning intricate features from complex dermatoscopic images.

DenseNet169, with its densely connected layers and efficient feature extraction capabilities, indeed plays a crucial role in accurately differentiating between benign and malignant skin lesions. Additionally, the use of transfer learning with pre-trained DenseNet169 models can significantly enhance classification accuracy, particularly in scenarios with limited training data. The dense connectivity pattern in DenseNet169 promotes feature reuse and propagation, addressing the vanishing gradient problem and contributing to the development of robust and interpretable skin cancer classification models.

It effectively highlights the significance of VGG16 in extracting discriminative features from dermatoscopic images, aiding in the differentiation between benign and malignant skin lesions. The mention of VGG16's straightforward architecture and uniform convolutional layer design adds clarity to its implementation. Additionally, the emphasis on transfer learning with pre-trained VGG16 models for improved classification accuracy, particularly in scenarios with limited training data, underscores its practical relevance.

**4.Test Set:**

The test set comprises a collection of dermatoscopic images that are exclusively used to evaluate the performance and generalization ability of a classification model. It serves to provide an unbiased assessment of the model's accuracy and effectiveness in classifying benign and malignant skin lesions on previously unseen data, ensuring its real-world applicability and preventing over fitting.

ISSN: 2366-1313

- **Accuracy:** The proportion of correctly classified dermatoscopic images out of the total images in the test set. It serves as a key performance metric, indicating the model's overall effectiveness in accurately distinguishing between benign and malignant skin lesions. The formula for calculating accuracy in a skin cancer classification test set is :Accuracy = (true positives + true negatives)/total.

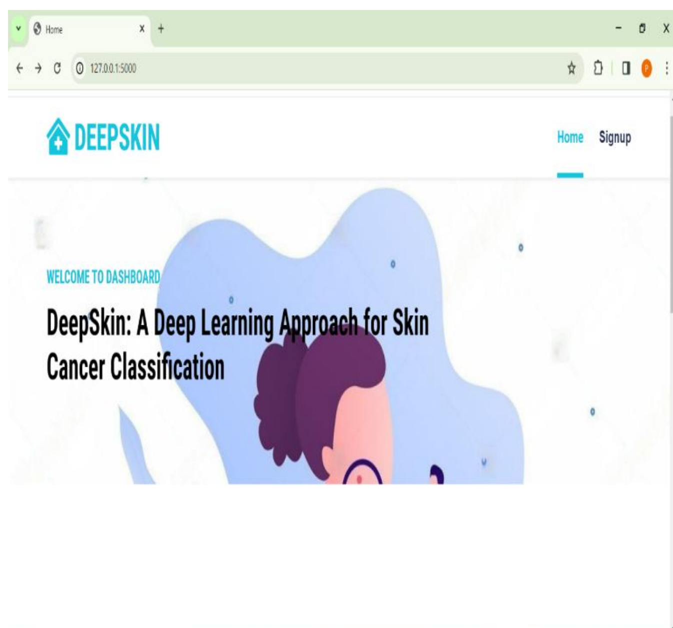
- **Precision:** The proportion of accurately predicted malignant cases out of all the cases that were predicted as malignant by the model. It serves as a crucial metric for assessing the model's ability to avoid misclassifying benign cases as malignant, its specificity in identifying true positives. The formula for calculating precision is Precision=True positives/(True positives +False positives).

- **Recall:** Recall refers to the proportion of actual positive cases that are correctly identified as such by the classification model. In other words, it measures the model's ability to correctly identify all the actual cases of skin cancer, minimizing the number of false negatives. The formula for calculating recall is Recall= True Positive(TP)/True Positive(TP)+False Negative(FN).

- **F1 score:** F1 score is a measure of a model's accuracy that takes both precision and recall into account. It is the harmonic mean of precision and recall, providing a balance between the two metrics and offering a single value that represents the model's performance in identifying both positive and negative cases of skin cancer.

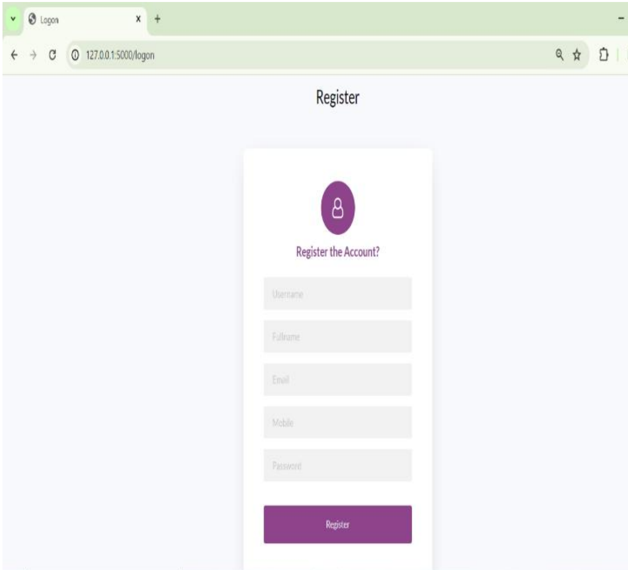
## V RESULT AND DISCUSSION

Home page:

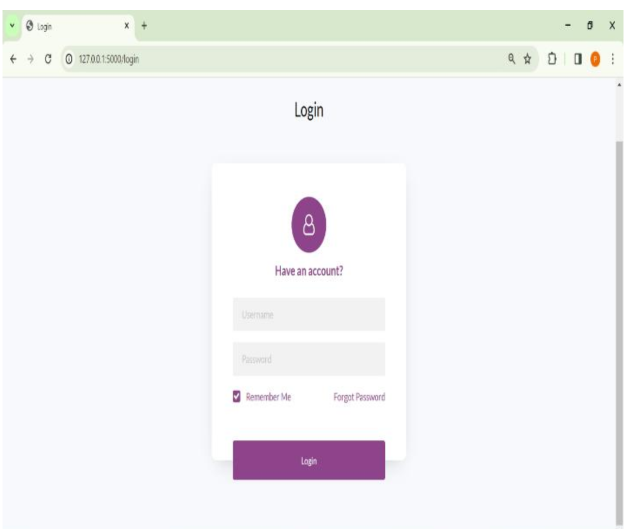



Registration page:

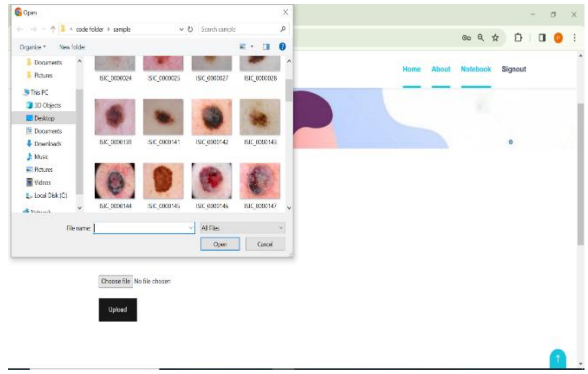
ISSN: 2366-1313



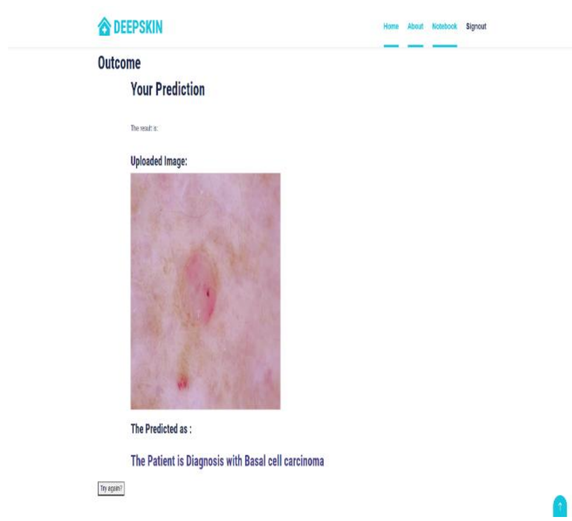

Login page:




Input data:




Output results:




## VI CONCLUSION

Skin cancer is one of the world's fastest growing diseases. Skin cancer is mostly caused by exposure to UV radiation from the sun. Given the limited resources available, early detection of skin cancer essential. Effective skin cancer prevention measures require accurate diagnosis and identification. Dermatologists struggle to detect early-stage skin cancer. Deep learning is now widely used for both supervised and unsupervised applications. CNNs are highly effective for object recognition and categorization (CNN). The dataset was filtered from MNIST: HAM10000, with a sample size of 10015 and seven categories of skin lesions. Data preprocessing techniques involve sampling, segmentation using auto encoder and decoder, and the use of a dull razor. The model was trained utilizing transfer learning approaches, including DenseNet169 and Resnet 50. The training and assessment ratios were 80:20, 70:30, and 40:60, respectively. DenseNet169's under sampling strategy achieved 91.2% accuracy with a f1-measure of 91.7%, whereas Resnet50's oversampling technique delivered 83% accuracy with a f1-measure of 84%. The study's future extension will entail boosting forecast accuracy through parameter adjustment.

**FUTURE ENHANCEMENT**

The future enhancement of skin cancer classification involves refining the model's forecast accuracy through parameter tuning. This process entails fine-tuning the various parameters of the classification model to optimize its performance in accurately identifying different types of skin cancer. By adjusting parameters such as learning rate, regularization strength, or the number of layers and nodes in the neural network, the model can be optimized to make more precise predictions. Additionally, exploring advanced hyper parameter optimization techniques like grid search or Bayesian optimization can further improve the model's ability to discern between benign and malignant skin lesions. Through parameter tuning, the study aims to elevate the classification model's accuracy, thus enhancing its potential for early and accurate skin cancer diagnosis, which is crucial for timely medical intervention and improved patient outcomes.

## VII REFERENCES

[1] Y. C. Lee, S.-H. Jung, and H.-H. Won, ‘‘WonDerM: Skin lesion classification with fine-tuned neural networks,’’ 2018, arXiv:1808.03426.

[2] U. Jamil, M. U. Akram, S. Khalid, S. Abbas, and K. Saleem, ‘‘Computer based melanocytic and nevus image enhancement and segmentation,’’ BioMed Res. Int., vol. 2016, pp. 1–13, Jan. 2016.

[3] A. Mahbod, G. Schaefer, C. Wang, R. Ecker, and I. Ellinge, ‘‘Skin lesion classification using hybrid deep neural networks,’’ in Proc. IEEE Int. Conf. Acoust.,

Speech Signal Process. (ICASSP), May 2019, pp. 1229–1233.

[4] K. Pai and A. Giridharan, ‘‘Convolutional neural networks for classifying skin lesions,’’ in Proc. TENCON IEEE Region 10 Conf. (TENCON), Oct. 2019, pp. 1794–1796.

[5] A. S. Shete, A. S. Rane, P. S. Gaikwad, and M. H. Patil, ‘‘Detection of skin cancer using CNN algorithm,’’ Int. J., vol. 6, no. 5, pp. 1–4, 2021.

[6] M. Vidya and M. V. Karki, ‘‘Skin cancer detection using machine learning techniques,’’ in Proc. IEEE Int. Conf. Electron., Comput. Commun. Technol. (CONECCT), Jul. 2020, pp. 1–5.

[7] H. Nahata and S. P. Singh, ‘‘Deep learning solutions for skin cancer detection and diagnosis,’’ in Machine Learning with Health Care Perspective. Cham, Switzerland: Springer, 2020, pp. 159–182.

[8] P. Wighton, T. K. Lee, H. Lui, D. I. McLean, and M. S. Atkins, ‘‘Generalizing common tasks in automated skin lesion diagnosis,’’ IEEE Trans. Inf. Technol. Biomed., vol. 15, no. 4, pp. 622–629, Jul. 2011.

[9] J. Saeed and S. Zeebaree, ‘‘Skin lesion classification based on deep convolutional neural networks architectures,’’ J. Appl. Sci. Technol. Trends, vol. 2, no. 1, pp. 41–51, Mar. 2021.

[10] Y. Li, A. Esteva, B. Kuprel, R. Novoa, J. Ko, and S. Thrun, ‘‘Skin cancer detection and tracking using data synthesis and deep learning,’’ 2016, arXiv:1612.01074.

[11] V. Badrinarayanan, A. Kendall, and R. Cipolla, ‘‘SegNet: A deep convolutional encoder–decoder architecture for image segmentation,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

## AUTHORS

**Mrs. K. Srilatha Assistant Professor** Dept. of CSE, Teegala Krishna Reddy Engineering College, Meerpet, Hyderabad.

Email: kumbamsrilatha@tkrec.ac.in

**Miss. A.Bhavana**, Dept. of CSE, Teegala Krishna Reddy Engineering College, Meerpet, Hyderabad.

Email: ayyannagaribhavana@gmail.com

**Mr. Ch.Saketh** ,Dept. of CSE, Teegala Krishna Reddy Engineering College, Meerpet, Hyderabad.

Email: cheekatisakethch@gmail.com

**Miss. G.Poojitha**, Dept. of CSE, Teegala Krishna Reddy Engineering College, Meerpet, Hyderabad.

Email: poojithagoshika@gmail.com