iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.61648

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Pioneering Dermatological Disease Detection with CNNs

Dr. Krishna Samalla[1], Kamma Vignesh[2], Nadipally Dheeraj[3], Dusharla Shivagowtham[4]

Abstract: ***Skin cancer poses a significant public health concern globally, with its incidence steadily rising. Early and accurate detection plays a pivotal role in improving patient outcomes and survival rates. With leveraging advancements in deep learning, this project explores the efficacy of Convolutional Neural Networks (CNNs) in automating the classification and detection of skin cancer from dermatoscopic images. The methodology employed in this project involves the development and optimization of a CNN architecture tailored for skin cancer analysis. The proposed model harnesses the hierarchical feature learning capabilities of CNN to automatically extract intricate patterns and discriminative features from dermatoscopic images, enabling accurate differentiation among various skin lesion types, including melanoma, basal cell carcinoma, and benign nevi. In addition to classification, the developed CNN architecture integrates a lesion detection mechanism. This auxiliary functionality enables the identification and delineation of potentially cancerous regions within skin lesions. The proposed model's detection capabilities are evaluated through bounding box prediction and segmentation masks, exhibit notable performance in localizing suspicious areas, aiding dermatologists in targeted examinations. With the help of this project, we can prevent the cause of damage to skin at an early stage.***

Keywords: Adam optimization, Data augmentation, Class imbalance, Under fitting, Over fitting, Sequential model.

I. INTRODUCTION

Skin cancer, characterized by abnormal skin cell growth due to DNA mutations, encompasses several common types such as basal cell carcinoma (BCC), squamous cell carcinoma (SCC), and melanoma. Additionally, non-cancerous growths like seborrheic keratosis and dermatofibroma are prevalent. The significance of early detection cannot be overstated, yet this is complicated by class imbalance within datasets, where benign cases often outnumber malignant ones, posing challenges for accurate diagnosis.

To address these challenges, a pioneering project is underway to advance skin cancer diagnosis using cutting-edge techniques such as convolutional neural networks (CNNs), Adam optimization, data augmentation, and class imbalance detection. Adam optimization dynamically adjusts learning rates during model training, leading to enhanced performance. Data augmentation techniques, including rotation and brightness adjustments, diversify the training dataset, aiding in model generalization. Moreover, class imbalance detection ensures a balanced representation of benign and malignant cases, thereby mitigating biased model training and improving diagnostic accuracy. The project's objectives are multifaceted, aiming to design optimized CNN architectures tailored for skin cancer diagnosis. Experimentation with different configurations and hyperparameters will be conducted to maximize model performance. Evaluation metrics such as accuracy, sensitivity, specificity, and AUC-ROC will be utilized to assess model performance on independent test datasets. Furthermore, integration into clinical practice will involve developing user-friendly interfaces for dermatologists, streamlining the diagnostic process and facilitating timely and accurate diagnosis.

However, numerous challenges must be addressed along the way. These include:

1) *Overcoming Class Imbalance:* Strategies to balance representation of benign and malignant cases in training data.
2) *Accessing Diverse Datasets:* Obtaining a wide range of skin lesion images to enhance model generalization.
3) *Mitigating Overfitting Risk:* Ensuring models do not become overly specialized to the training data.
4) *Quality Assurance in Data Augmentation:* Ensuring augmented data remains relevant and accurate for training.
5) *Addressing Algorithmic Bias:* Preventing models from favoring dominant classes over minority ones, which could lead to less effective diagnoses.
6) *Ensuring Interpretability:* Making CNN decisions understandable to clinicians to foster trust in the diagnostic process.

By confronting these challenges head-on, the project endeavors to revolutionize skin cancer diagnosis, ultimately leading to improved patient outcomes and reduced healthcare burdens. This innovative approach holds the promise of transforming the landscape of skin cancer diagnosis, leveraging state-of-the-art technology to benefit patients, healthcare providers, and the broader medical community.

©IJRASET: All Rights are Reserved



The motivation behind pioneering skin cancer diagnosis through convolutional networks using class imbalance detection lies in addressing the critical need for accurate and efficient early detection of skin cancer. Skin cancer is one of the most prevalent types of cancer lobally, with rising incidence rates posing significant challenges to healthcare systems worldwide. Early detection is paramount for successful treatment outcomes, yet diagnosing skin cancer accurately can be challenging, particularly due to the presence of class imbalances within datasets used for training machine learning models.

II. LITERATURE SURVEY

The veracity of the data pertaining to computer-aided techniques' accuracy rate has been scrutinized and evaluated by specialists. Databases from IEEE, SpringerLink, and ScienceDirect were accessed. Techniques for classifying and segmenting skin lesions were examined, with important drawbacks highlighted. In, a refined method for diagnosing melanoma skin cancer was introduced. To generate synthetic images of melanoma, an implantation manifold with nonlinear embeddings was utilized. A new set of skin melanoma datasets was constructed using the data augmentation approach using dermatoscopic scans from the publicly available PH2 dataset

For the purpose of detecting skin cancer, Catarina Barata and Jorge S. Marques [Barata2019] used hierarchical structures with deep learning in [1]. They discovered that dermatologists classify skin lesions according to a hierarchical structure. However, automated systems do not take this hierarchical approach into account. They consider every variety of skin ailment. In order to execute the hierarchical diagnosis of skin cancer, a deep learning method was considered in the research. They mentioned how easy it was to analyze and arrange the skin lesion photos according to a structural hierarchy. They used a deep learning technique to provide a significant evaluation criterion for the diagnosis .

Deep learning-based skin cancer diagnosis was also studied by Le Thu Thao and Nguyen Hong Quang [Thao2017] [2]. The focus of the investigation was skin lesions rather than the identification of melanoma. Two issues with the deep learning approach for detecting skin cancer from photos of skin tumors were resolved as a result of the research. The datasets were trained utilizing conventional information from the 2017 International Skin Imaging Collaboration (ISIC) databases. In comparison to the previously recommended deep learning methodologies, the conclusion demonstrated superior accuracy and performance.

For the purpose of identifying and diagnosing skin cancer, Yoonsik Kim, Insung Hwang, and Nam Ik Cho [Kim2017] [3] recommended using two CNN and training techniques.

The first type of network was a 20-layer, 3x3 filter VGG network. The second involved utilizing 20 NiN layers in the original conception model. Both patch-based and image- based training are taken into consideration for cancer diagnosis .

The improved photos were used to train the Squeeze Net deep learning network. The experiments showed a considerable improvement in melanoma identification accuracy of 90.18%. It was proposed in to use the VGG-SegNet algorithm to extract a skin melanoma (SM) region from a digital dermatoscopy image. Important performance metrics were then determined by contrasting the recovered segmented SM with the ground truth (GT). Using the common ISIC2016 database, the suggested plan was assessed and confirmed.

Researchers have classified skin cancer using a combination of artificial and human intelligence. A CNN and a total of 112 German dermatologists divided 300 skin lesions with biopsy verification into five groups. The two independently acquired sets of diagnoses were combined using gradient boosting to produce a single classifier. 82.95% multiclass accuracy was achieved by both human and machine. Benign and malignant tumors can be identified using the deep learning-based InSiNet method.

Recent Work	Data Size	Data Set	Techniques Used	Number of Classes
[4]	300	HAM10000	CNN with XGBoost	Five
[5]	1323	HAM10000	InSiNet	Two
[6]	48373	Dermnet, ISIC Archive	MobileNetV2	Two
[7]	7470	HAM10000	ResNet50	Seven
[8]	1000	ISIC	SVM + RF	Eight
[9]	279	ImageNet	DCNN VGG-16	Two
[10]	10,015	HAM10000	Alexnet	Seven

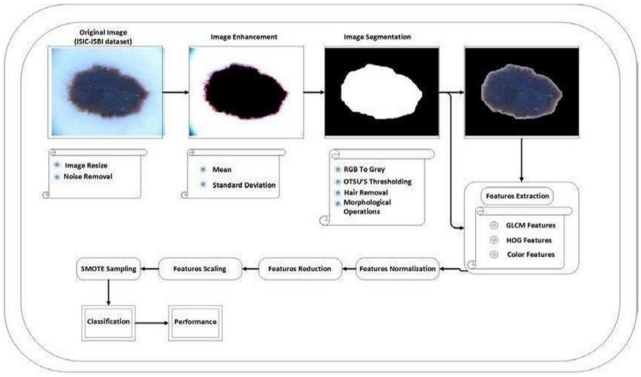
III. METHODOLOGY

A. Class Imbalance Detection method

When there are unequal numbers of instances in each class inside a dataset, it is referred to as class imbalance. i.e. there are noticeably more instances of some classes than others. An unbalanced problem is one where there are relatively few occurrences of the positive class in comparison to the negative class.

Managing class imbalance can help machine learning models perform better to be more robust and ensure that they adapt well to new data. Additionally, it may result in assessment metrics that are more meaningful and precise and accurately represent the model's actual performance in each class. There are several methods like under sampling, oversampling (SMOTE), ensemble method, class weights and focal loss for resolving class disparities.

The imbalanced-learn package can be used to handle different approaches. The original test data should be used to evaluate the generated model, and these techniques should only be used on the training set.



B. SMOTE flow Model to Detect Skin Cancer

However, SMOTE creates artificial samples, which may include anomalies and noise in the data as shown in figure 4. SMOTE application can be expensive if the dataset is big. When there is a serious class imbalance and a notable under representation of the minority class, SMOTE is especially helpful. It can also generate more synthetic samples to enhance model performance in situations where there is a dearth of data for the minority class.

C. Design Approach for Skin Cancer Detection

1) Problem Understanding and Requirements Gathering

- Conduct thorough research on skin cancer detection methodologies, including image processing and machine learning techniques.
- Collaborate with medical professionals to understand the specific requirements and challenges in skin cancer diagnosis.
- Define the scope of the project, including the types of skin conditions to be detected and the target accuracy levels.

2) Data Collection and Preprocessing

- Collect a diverse dataset of skin images representing various types of skin conditions.
- Ensure proper labeling and annotation of the dataset to facilitate supervised learning.
- Preprocess the images to standardize size, format, and resolution, and remove noise or artifacts.

3) Model Selection and Architecture Design

- Choose an appropriate deep learning architecture for image classification tasks, such as Convolutional Neural Networks (CNNs).

4) Training Strategy

- Split the dataset into training, validation, and test sets to evaluate the model's performance.

5) Evaluation and Validation

- Evaluate the trained model using performance metrics like accuracy, precision, recall, and F1 score.

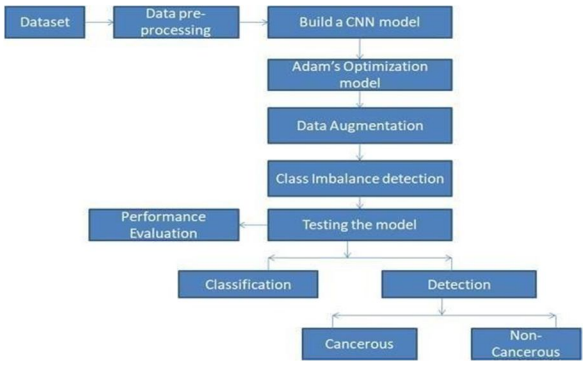
6) Integration and Deployment

- Integrate the trained model into a user-friendly interface for easy access by medical professionals.

7) Continuous Improvement

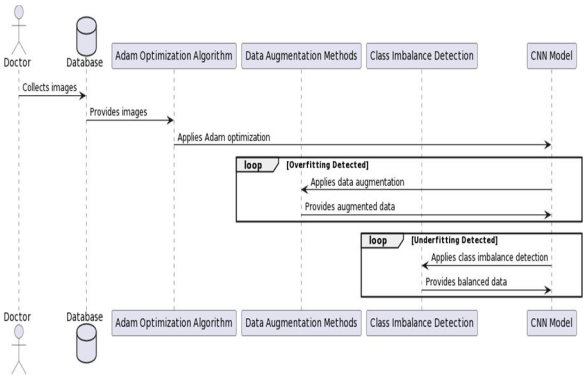
- Collect feedback from users and stakeholders to identify areas for improvement.
- Regularly update the model with new data and retrain it to adapt to evolving trends and challenges in skin cancer detection.

D. Architectural Design

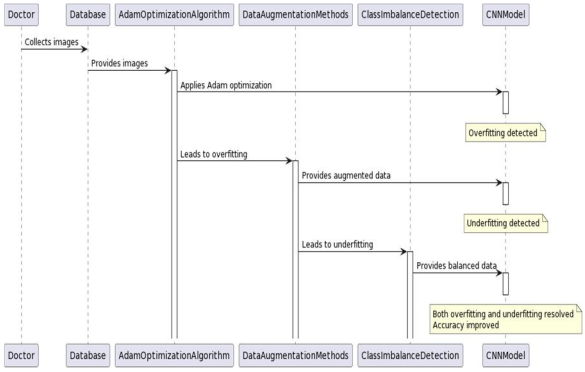


By following this design approach, the skin cancer detection project can effectively leverage cutting-edge technologies to develop an accurate and reliable system for early diagnosis and treatment of skin cancer.

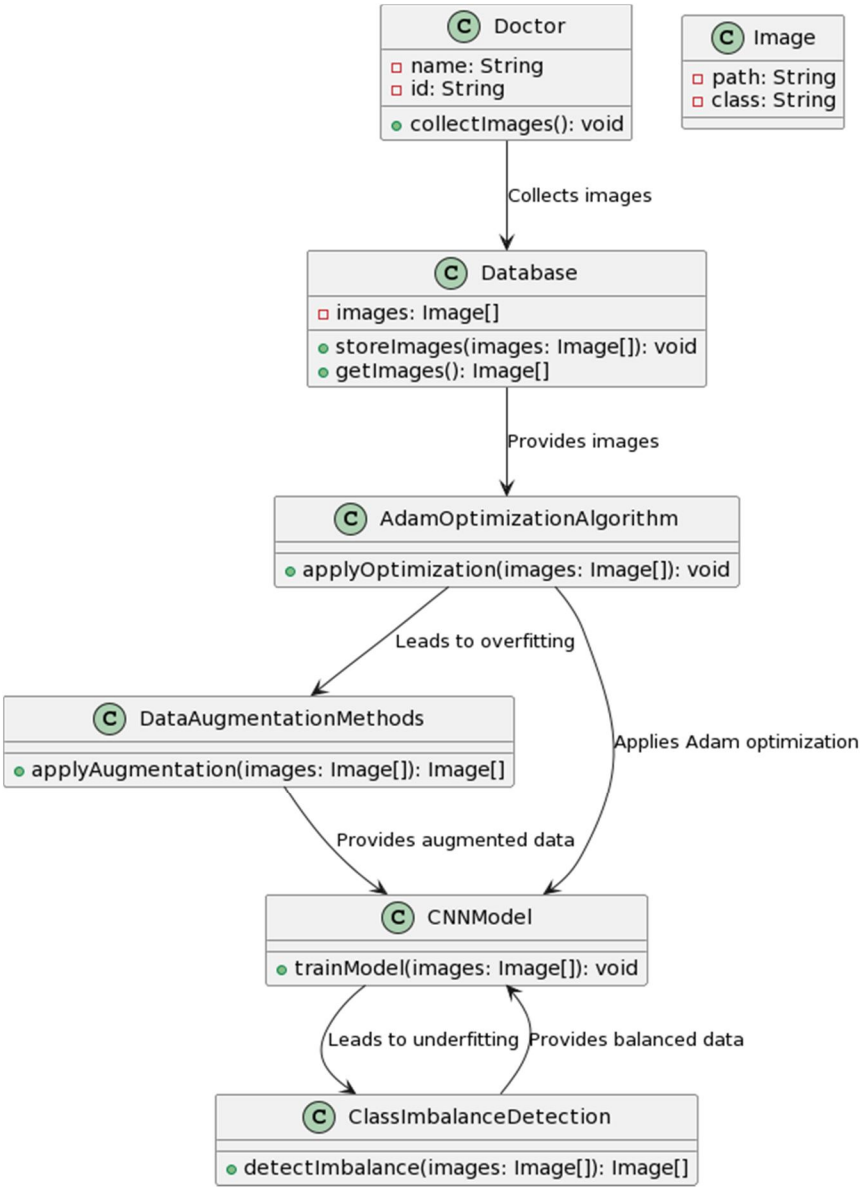
E. Implementation Methodologies



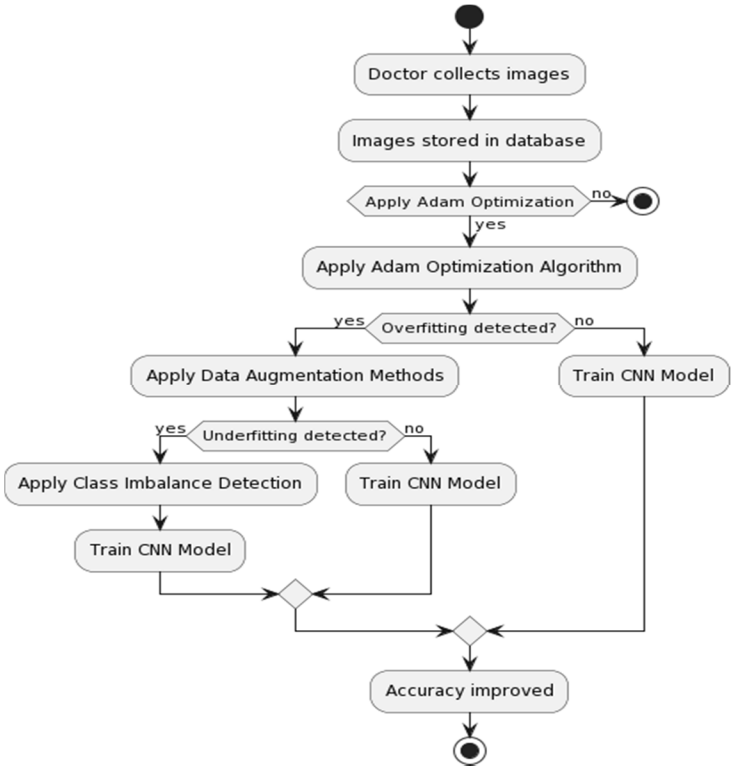
F. Use case flow of Skin Cancer Detection



G. Sequence Flow of Skin Cancer Detection



H. Class Diagram Flow for Skin Cancer Detection



IV. DESCRIPTION

This use case describes the process of detecting skin cancer using a combination of image processing techniques and machine learning algorithms. The process begins with a doctor collecting images of the patient's skin, which are then stored in a database system. The stored data consists of images categorized into 9 different types of classes representing various types of skin conditions.
The collected images are used as input for the skin cancer detection system. Initially, an Adam optimization algorithm is applied to a convolutional neural network (CNN) model to optimize the learning process. However, due to the complex nature of the data, overfitting is observed in the results.
To address the overfitting issue, data augmentation methods are employed. These methods involve rescaling and rotating the images to increase the diversity of the training dataset. Despite this augmentation, the CNN model exhibits underfitting, indicating a lack of generalization.
To mitigate underfitting, class imbalance detection techniques are applied. Extra 500 images are added to each class in the dataset to ensure balanced representation. The balanced dataset is then used as input for the CNN model.
The CNN model trained on the balanced dataset successfully resolves both underfitting and overfitting issues. As a result, the accuracy of the skin cancer detection system is significantly improved compared to the existing VGG16 model.

A. Actors

- Doctor: Collects images of the patient's skin.
- Skin Cancer Detection System: Processes and analyzes the collected images.

B. Preconditions

- The doctor has access to a device capable of capturing high-quality images of the patient's skin.
- The skin cancer detection system is properly configured and integrated with the database system.

C. Postconditions

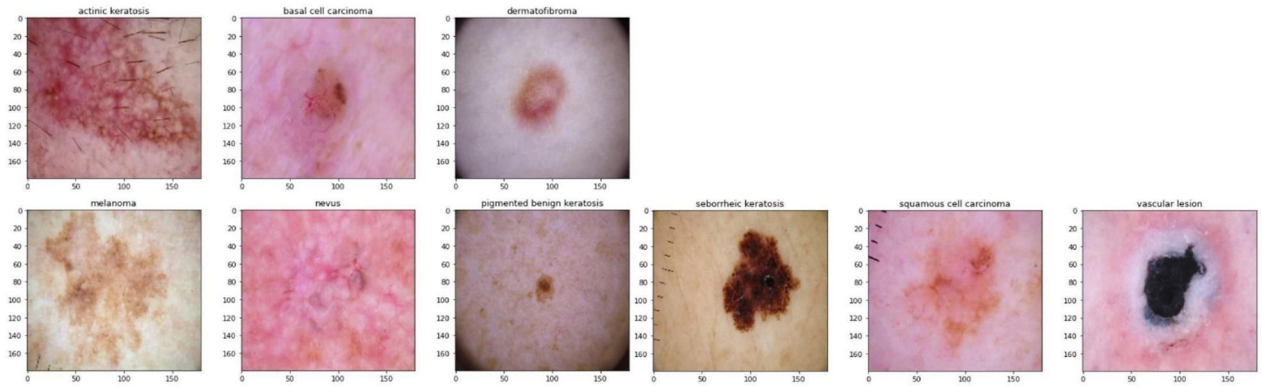
- The skin cancer detection system produces accurate diagnoses based on the input images.
- The CNN model is trained and optimized to detect various types of skin cancer with improved accuracy.

D. Main Flow

- The doctor collects images of the patient's skin using a suitable imaging device.
- The collected images are stored in a database system for further analysis.
- The skin cancer detection system retrieves the stored images from the database.
- An Adam optimization algorithm is applied to a CNN model to optimize the learning process.
- Overfitting is observed in the results obtained from the CNN model.
- Data augmentation methods, including rescaling and rotating the images, are applied to address the overfitting issue.
- Despite data augmentation, the CNN model exhibits underfitting, indicating a lack of generalization.
- Class imbalance detection techniques are applied to ensure a balanced representation of different classes in the dataset.
- Extra 500 images are added to each class in the dataset to mitigate underfitting.
- The balanced dataset is used as input for the CNN model.
- The CNN modeltrained on the balanced dataset successfully resolves both underfitting and overfitting issues.
- The skin cancer detection system produces accurate diagnoses based on the input images, with improved accuracy compared to the existing VGG16 model.

V. RESULTS

A. Activity Flow Diagram of Skin Cancer Detection



Visualized Images of Each type of Skin Cancer

These are the images obtained in visualization of the different types of skin cancer. Dataset consists of 9 classes and above are the images of each class.

Model: "sequential"

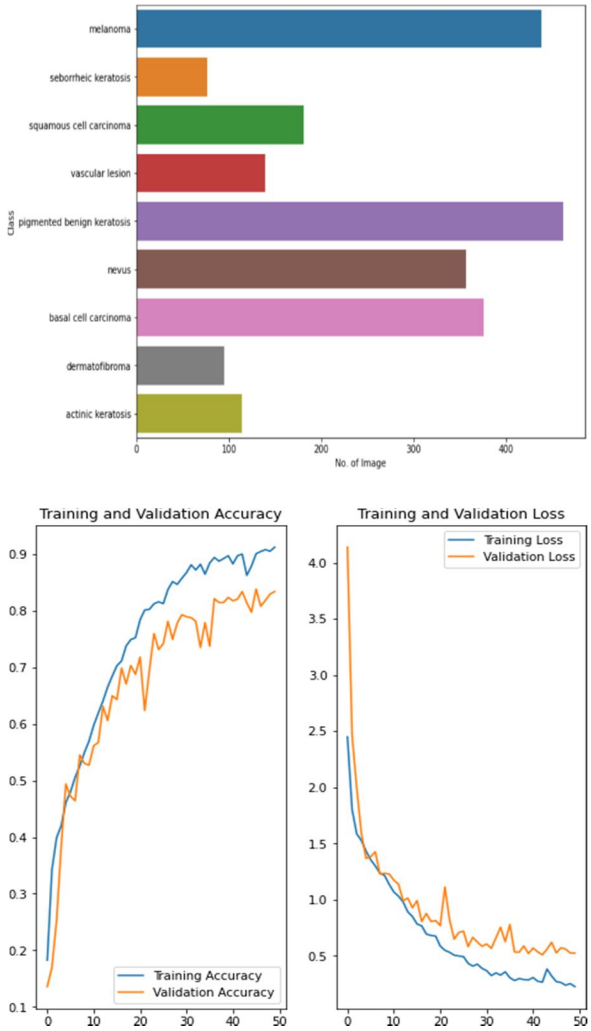
Layer (type)	Output Shape	Param #
rescaling (Rescaling)	(None, 180, 180, 3)	0
conv2d (Conv2D)	(None, 178, 178, 32)	896
max_pooling2d (MaxPooling2D)	(None, 89, 89, 32)	0
conv2d_1 (Conv2D)	(None, 87, 87, 64)	18496
max_pooling2d_1 (MaxPooling2	(None, 43, 43, 64)	0
conv2d_2 (Conv2D)	(None, 41, 41, 128)	73856
max_pooling2d_2 (MaxPooling2	(None, 20, 20, 128)	0
flatten (Flatten)	(None, 51200)	0
dense (Dense)	(None, 512)	26214912
dense_1 (Dense)	(None, 128)	65664
dense_2 (Dense)	(None, 9)	1161

...
Total params: 26,374,985
Trainable params: 26,374,985
Non-trainable params: 0

Layers Present in CNN Model

	Class	No. of Image
0	melanoma	438
1	seborrheic keratosis	77
2	squamous cell carcinoma	181
3	vascular lesion	139
4	pigmented benign keratosis	462
5	nevus	357
6	basal cell carcinoma	376
7	dermatofibroma	95
8	actinic keratosis	114

Number of images in each class



Graph between class vs its number of image Result obtained by class imbalance detection

- As per the final model (model3) Training accuracy and validation accuracy increases
- Modeloverfitting issue is solved
- Class rebalance helps in augmentation and achieving the best Training and validation accuracy.

Type of Algorithm used	Accuracy obtained	Problem Araised
Adam Optimization	88.84%	Over Fitting
Data Augmentation	53.63%	Underfitting
Class Imbalance detection	91.21%	Above two problems are resolved.

VI. CONCLUSION

The conclusion of this project marks a significant achievement in the field of skin cancer detection, where meticulous attention to algorithmic optimization has led to substantial improvements in accuracy and model robustness. Initially grappling with challenges of overfitting and subsequently encountering underfitting due to data augmentation, the project team demonstrated resilience and adaptability in their approach. Through the implementation of class imbalance detection techniques, notably the addition of extra images to each class in the dataset, the project achieved a commendable accuracy of 91.2%. This outcome not only surpassed previous benchmarks but also addressed the dual issues of underfitting and overfitting, indicating a harmonious balance in model performance. By resolving these fundamental challenges, the project has laid a solid foundation for the development of more accurate and reliable skin cancer detection systems, poised to make significant strides in linical practice and public health initiatives.

VII. FUTURE SCOPE

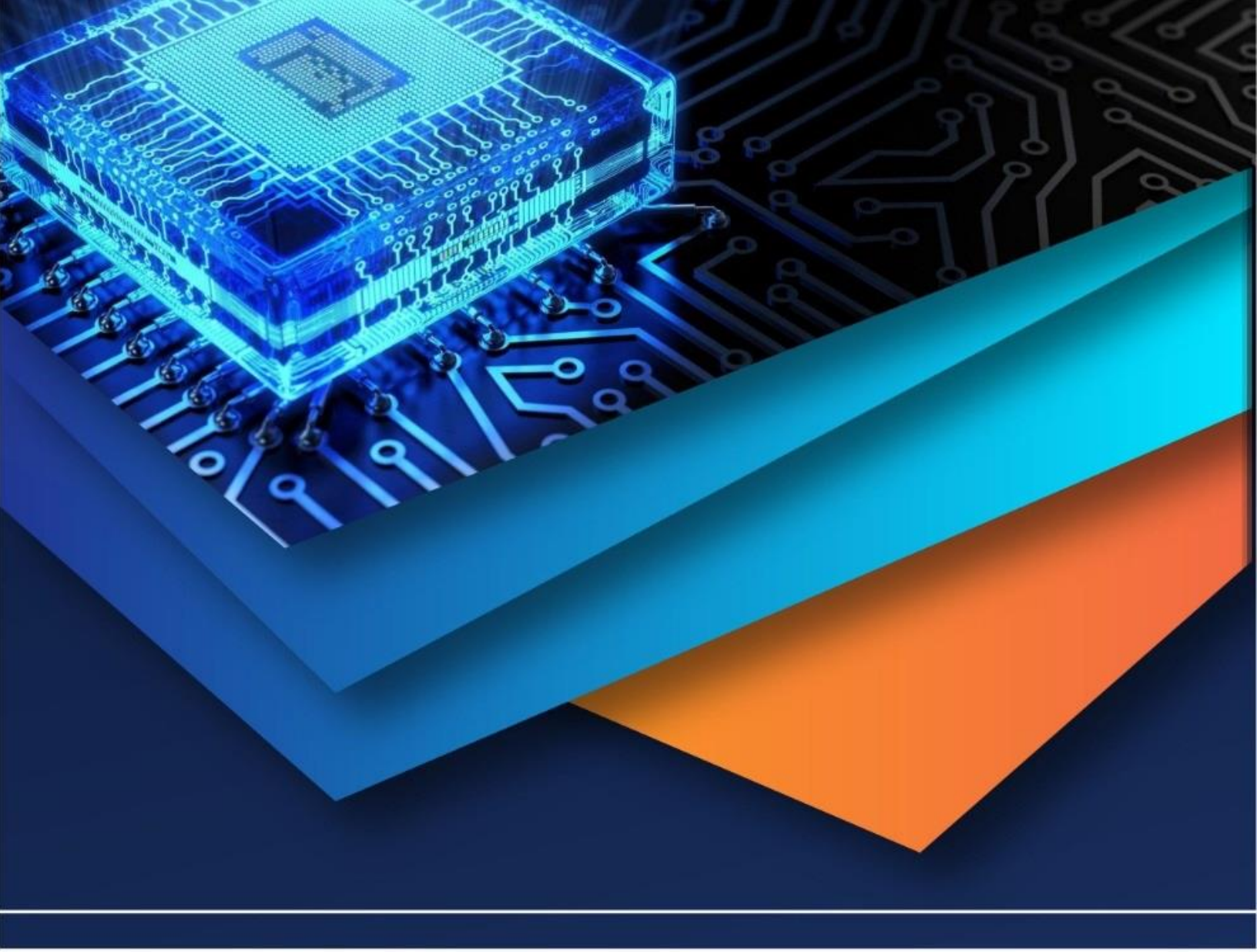
The success of this project in achieving a high accuracy rate and resolving issues of underfitting and overfitting opens up promising avenues for future research and development in the field of skin cancer detection. Moving forward, one potential area of focus could be the integration of additional advanced techniques, such as transfer learning or ensemble methods, to further enhance the performance and robustness of the model. Moreover, continued refinement of data augmentation strategies and exploration of novel augmentation techniques could lead to even greater improvements in model generalization and accuracy. Additionally, the project could benefit from expanding the scope of the dataset to include more diverse and representative samples, including images from different demographics and skin types, to ensure broader applicability and inclusivity of the detection system. Furthermore, efforts to optimize computational efficiency and scalability will be crucial for deploying the system in real-world clinical settings and population-wide screening programs. Collaboration with healthcare professionals and stakeholders will also be essential to validate the system's effectiveness in clinical practice and ensure its seamless integration into existing healthcare workflows. Overall, the project lays a solid foundation for continued innovation and advancement in skin cancer detection, with the potential to make a significant impact on patient care and public health initiatives in the future.

REFERENCES

[1] C.Barata, C. and J. S. Marques, Deep Learning for Skin Cancer Diagnosis with Hierarchical Architectures. In IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), pages 841-845, 2019.
[2] L.T. Thao and N.H. Quang, Automatic skin lesion analysis towards melanoma detection, In 21st Asia Pacific Symposium on Intelligent and Evolutionary Systems (IES), pages 106-111, 2017.
[3] Y. Kim, I. Hwang and N. I. Cho, Convolutional neural networks and training strategies for skin detection. In IEEE International Conference on Image Processing (ICIP), pages 3919- 3923, 2017.
[4] Rajput G., Agrawal S., Raut G., Vishvakarma S.K. An accurate and noninvasive skin cancer screening based on imaging technique. Int. J. Imaging Syst. Technol. 2022;32:354–368. doi: 10.1002/ima.22616. [CrossRef] [Google Scholar].
[5] Guan Q., Wang Y., Ping B., Li D., Du J., Qin Y., Lu H., Wan X., Xiang J. Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: A pilot study. J. Cancer. 2019;10:4876. doi: 10.7150/jca.28769. [PMC free article] [PubMed] [CrossRef] [Google Scholar].
[6] Hekler A., Utikal J.S., Enk A.H., Hauschild A., Weichenthal M., Maron R.C., Berking C., Haferkamp S., Klode J., Schadendorf D., et al. Superior skin cancer classification by the combination of human and artificial intelligence. Eur. J. Cancer. 2019;120:114–121. doi: 10.1016/j.ejca.2019.07.019. [PubMed] [CrossRef] [Google Scholar]
[7] Reis H.C., Turk V., Khoshelham K., Kaya S. InSiNet: A deep convolutional approach to skin cancer detection and segmentation. Med. Biol. Eng. Comput. 2022;7:1–20. doi: 10.1007/s11517-021-02473-0. [PubMed] [CrossRef] [Google Scholar].
[8] Ech-Cherif A., Misbhauddin M., Ech-Cherif M. Deep neural network based mobile dermoscopy application for triaging skin cancer detection; Proceedings of the 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS); Riyadh, Saudi Arabia. 1 May 2019. [Google Scholar]

[9] Le D.N., Le H.X., Ngo L.T., Ngo H.T. Transfer learning with class-weighted and focal loss function for automatic skin cancer classification. arXiv. 20202009.05977 [Google Scholar]
[10] Murugan A., Nair S.A.H., Preethi A.A.P., Kumar K.S. Diagnosis of skin cancer using machine learning techniques. Microprocess. Microsyst. 2021;81:103727. doi: 10.1016/j.micpro.2020.103727. [CrossRef] [Google Scholar]
[11] Craythorne, E. and Al-Niami, F., 2017. Skin cancer. Medicine, 45(7), pp.431-434.
[12] Moussa, R., Gerges, F., Salem, C., Akiki, R., Falou, O. and Azar, D., 2016, October. Computer aided detection of Melanoma using geometric features. In 2016 3rd Middle East Conference on Biomedical Engineering (MECBME) (pp. 125- 128). IEEE.
[13] W. Sterry and R. Paus. Thieme clinical companions' dermatology. In Thieme Verlag. Stuttgart, New York., 2006.
[14] Matthew Hoffman,"Picture of the skin and skin cases,". [online]. http://www.webmd.com/skin-problems-and-treatments/picture-of-the-skin#1 [Accessed:March 14, 2019].

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)