Detection of Skin Cancer using Inception V3 And Inception V4 Convolutional Neural Network (CNN) For Accuracy Improvement

Bhimanadhula Nandini¹; Dr. R. Puviarasi²*

¹Research Scholar, Department of Electronics and Communication Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India.

²Project Guide, Department of Electronics and Communication Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India.

¹bhimanadhulanandini17@saveetha.com
²puviarasi@saveetha.com

Abstract

Aim: The main aim of this work is to measure the accuracy for automated detection of dermal cell images using the Convolutional Neural Network (CNN) algorithm.

Materials and Methods: The skin images the dataset collected from International skin images collaboration (ISIC). In this framework 1200 images are used out of which (80%) are trained and (20%) are used for testing for the detection of skin cancer. 1200 images are used for group I (Inception V4) in comparison with Inception V3 and statistical analysis done using SPSS analysis. The sample size of two groups is calculated using G power with pretest power of 80 and alpha value 0.05 (error rate) with inputs 2400 (1200*2). Results: The inception V4 using CNN shows better results in mean Accuracy of 92.34±0.87 followed by inception V3 produces an accuracy of 90.34±0.13 with the significance value of <0.001. Conclusion: It is concluded that based on the execution analysis, the Inception V4 appears to be better accuracy compared with the Inception V3 algorithm.

Key-words: Innovative Detection, Convolutional Neural Network (CNN), Inception V3, Inception V4 CNN, Accuracy, Deep Learning.

1. Introduction

Accuracy measurement in skin cancer cell detection is one of the primary and challenging factors in the medical field. (Kadambur and Al Riyadi 2020). It is important for automated classification of skin lesions and prediction based on images can increase the accuracy of early
prediction to diagnose and treat the cancer as soon as possible. (Esteva et al. 2017). Much of the medical research is concentrating on the accuracy improvement of skin cancer detection. The advanced methods of image processing are used for analysing the performance of the medical research study. Detection of skin cancer using machine & deep learning algorithms, results in accurate prediction with less time and manual functions in the medical research. (Silver et al. 2016). (Haenssle et al. 2018).

In recent years, many researchers concentrating in skin cancer detection using machine learning techniques. Nearly 38,892 papers in science direct and 3,950 papers in pubmed were published in the last five years. In this (Jain, Jagtap, and Pise 2015) paper, for the detection of Melanoma Skin Cancer a computer aided method is analysed using various Image Processing tools. The Lesion Image analysis tools check for the various Melanoma parameters Like Asymmetry, Border, Colour, Diameter, (ABCD) and it produces an accuracy about 75.84%. In this study, using Neural Network and ABCD rule methodology detects and identifies skin lesions as benign or malignant. It is trained to classify the lesions for a high degree of accuracy. It shows the overall classification accuracy of 76.9% on a dataset of 463 images (Dubal et al. 2017). This research develop a simple method capable of detecting and classifying skin lesions using dermoscopy images on ABCD rules and it is implemented in the MATLAB environment. The experiment is based on PH2 database containing suspicious melanoma skin cancer and the overall accuracy of the developed approach is 90% (Zghal and Derbel 2020). In this paper, a model-driven architecture in the cloud, that uses deep learning algorithms in its core implementations, is used to construct models that assist in predicting skin cancer with improved accuracy and to build deep learning models to classify dermal cell images and detect skin cancer. CNN classifier is used and the deep learning models built here are tested on standard datasets, and the metric area under the curve of 99.77% was observed (Kadampur and Al Riyaeed 2020). Based on various analyses of existing methodology this article shows Area Under the Curve (AUC) of 99.77% in the detection of skin cancer by applying Deep learning Studio (DLS) in CNN classifiers (Kadambur and Al Riyaeed 2020).

Previously our team has a rich experience in working on various research projects across multiple disciplines (Sathish and Karthick 2020; Varghese, Ramesh, and Veeraiyan 2019; S. R. Samuel, Acharya, and Rao 2020; Venu, Raju, and Subramani 2019; M. S. Samuel et al. 2019; Venu, Subramani, and Raju 2019; Mehta et al. 2019; Sharma et al. 2019; Malli Suresh Babu et al. 2019; Krishnaswamy et al. 2020; Muthukrishnan et al. 2020; Gheena and Ezhilarasan 2019; Vignesh et al. 2019; Ke et al. 2019; Vijayakumar Jain et al. 2019; Jose, Ajitha, and Subbaiyan 2020). Now the growing trend in this area motivated us to pursue this project.
Based on the existing research analysis it is important to know which deep learning algorithm produces high performance with less significant value (<0.05) in order to achieve a better outcome. Hence the main aim of this work is to compare the mean accuracy of Inception V4 CNN and Inception V3 CNN algorithms in detection of skin cancer.

2. Materials and Methods

This study was carried out in the Department of Electronics and Communication Engineering at Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, India. The skin images the dataset collected from International skin images collaboration (ISIC). In this framework 1200 images (Kadampur and Al Riyaaee 2020) are used out of which (80%) are trained and (20%) are used for testing for the innovative detection of skin cancer. The proposed work is based on a computerized program in python using image samples of 2 groups. The sample size of two groups is calculated using G power with pretest power of 80 and alpha value 0.05 (error rate) with inputs 2400 (1200*2) fed from ISIC the research by (Kadampur and Al Riyadi 2020). Group 1 performance analysis is done using Inception V4 CNN algorithm and group 2 is by using Inception V3 CNN algorithm in python platform.

Fig. 1 - Process of the Convolutional Neural Network (CNN). The process flow includes image preprocessing, image segmentation and feature extraction process to produce an output. The softmax layer will decide the desired outcome (yes(1)/no(0)).

The model was built using google colab, an online python software. The work progress starts with image preprocessing, Image segmentation, Classification and Detection. Fig. 1 explains the test
procedure. According to this, the input image is given to the system and the Pre-processing includes the image resizing and an improvement of the image data. Image segmentation is performed by using automatic thresholding segmentation and masking operation in red, green and blue (RGB) planes. Threshold segmentation converts an image into a binary image (White or Black). Entropy, shape, energy features are extracted. The 80% of the images from the dataset is used for training and 20% of images is used for testing. The model is trained using Inception V4 CNN and Inception V3 CNN classifiers and accuracy is calculated. All the statistical analysis is done in SPSS software.

The input images were taken from the International skin images collaboration (ISIC) dataset for the accuracy analysis. The dataset consists of 1200 skin cancer images with a resolution 224*224 pixel values. The accuracy measurement is performed and average values are taken and it is compared with existing algorithms. For each sample the performance metric is calculated based on accuracy. Five samples for each group are considered and mean accuracy is calculated as mentioned in Table 1. These samples are used for the statistical analysis to calculate the Mean, standard deviation and significance values.

Table 1 - Sample Dataset of SPSS Software. Group 1 samples are obtained from Inception V4 and Group 2 samples are obtained from Inception V3. The Comparison shows the more accuracy value for Inception V4 than Inception V3. These five samples are used for statistical analysis in the SPSS tool

<table>
<thead>
<tr>
<th>S.no</th>
<th>Group</th>
<th>Sample</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Sample 1</td>
<td>92.23</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>Sample 2</td>
<td>92.17</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>Sample 3</td>
<td>92.14</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>Sample 4</td>
<td>92.08</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>Sample 5</td>
<td>92</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>Sample 1</td>
<td>90.34</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>Sample 2</td>
<td>90.27</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>Sample 3</td>
<td>90.15</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>Sample 4</td>
<td>90.06</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>Sample 5</td>
<td>90</td>
</tr>
</tbody>
</table>

3. Statistical Analysis

Statistical analysis is done using an independent sample test in SPSS Software. Descriptive statistics (Mean, Standard Deviation and Standard Error Mean) were carried out for each model. The analysis of this research work is done using an independent sample T-test which is used to compare the accuracy of two groups. The input dataset and the epoch value are independent variables and the dependent variables are the accuracy.
4. Results

The comparison of accuracy values on trained and test dataset using python is represented in Fig. 2. This gives the trained dataset accuracy value of 92.12. As per the result the test dataset accuracy is 90.32%. The accuracy of training and test dataset can be increased by increasing the number of iterations. From this graph it is showing least discrepancy with significance value (<0.05). The comparison of loss analysis results on trained and test dataset using python is illustrated in Fig 3. The loss of test and train dataset can be reduced by increasing the layers and input images. Initially training loss and testing losses are very high which are reduced as epochs are progressed.

Fig. 2 - Accuracy analysis of trained and tested dataset. Blue line indicates the Trained data Accuracy and Yellow line represents the test data accuracy. X axis: iterations (no. of times); Y axis: Accuracy. The graph shows higher accuracy for trained data than tested data.

Fig. 3 - Loss analysis of trained and tested dataset. X axis: iterations (no. of times); Y axis: Loss. The graph shows low loss values for trained data than tested data. For less than 40-50 iterations it shows less than 0.4.
Table 2 - SPSS Statistical Analysis of Inception V3 and Inception V4 models. Mean Accuracy, Sensitivity, Specificity, Standard Deviation and Standard Errors Mean are obtained for 10 iterations. The performance represents better outcomes for Inception V4 and Inception V3

<table>
<thead>
<tr>
<th>Group Statistics</th>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std.Deviation</th>
<th>Std.Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>Inception V4</td>
<td>5</td>
<td>92.1260</td>
<td>0.08678</td>
<td>0.03881</td>
</tr>
<tr>
<td></td>
<td>Inception V3</td>
<td>5</td>
<td>90.1780</td>
<td>0.13046</td>
<td>0.05834</td>
</tr>
</tbody>
</table>

Table 3 - Significance, Mean difference and Std. Error difference of accuracy on 2 groups. Levene’s test for Equality of variances showed a significance of 0.418 for accuracy and T-test for Equality of Means showed a significance of <0.001 for accuracy for the two groups in Skin Cancer detection. It includes 95% CI of the difference.

<table>
<thead>
<tr>
<th>Levene’s Test for Equality of Variances</th>
<th>T-test for Equality of Means</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>Sig.</td>
<td>t</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Equal variances assumed</td>
<td>0.728</td>
</tr>
<tr>
<td>Equal variance not assumed</td>
<td>27.800</td>
<td>6.960</td>
</tr>
</tbody>
</table>

Table 2 shows that the comparison of two groups is made with parameters mean, Standard deviation, and standard error mean. It declares that the Inception V4 performed appears to be better when compared with Inception V3. Table 3 shows independent sample T-tests have a significance of 0.001 which is less than the level of standard significance range (0.005). Hence it is proved that Group 1 and Group 2 are significantly different from each other.

Fig. 4 - Comparison of Mean accuracy and the standard error rate of Inception V4 and Inception V3. X axis represents Group 1 & 2 Algorithms: Y axis represents Mean Accuracy values in %. The Error Bars represent 95% CI and +/- 1 SD
In Fig. 4 the bar graph is plotted which represents the relationship between the 2 groups’ mean accuracy with error bar. In this graph the value of Error Bars is 95% CI and +/-1 SD. Based on this graph it concludes that Inception V4 has more accuracy and a less error rate than the Inception V3.

5. Discussions

In this framework, it is observed that the values of the mean accuracy using the Inception V4 Convolutional Neural Network (CNN) algorithm appears to be compared to Inception V3 CNN algorithm with significance p value is 0.418 in innovative detection of skin cancer.

The statistical analysis produces the mean accuracy, standard deviation and standard error mean of Inception V4 (group 1) are 92.1260, ±0.08678 and 0.03881 followed by Inception V3 (group 2) 90.1780, ±0.13046 and 0.05834 respectively and its shows that the Inception V3 produced least discrepancy of mean accuracy with significance value p<0.05.

Some of the previous studies by (Kadampur and Al Riyae 2020) demonstrated skin cancer disease based model architecture to classify dermal cell images. In this study it is proved that by comparing Inception V3 and the other CNN based algorithms with the results of accuracy, precision and F1 score and ROC AUC values Inception V3 showing better performance. The performance is showing least difference between the groups with the significance of<0.05. (Ech-Cherif, Misbahuddin, and Ech-Cherif 2019) in this research work CNN and DNN models are analysed deeply for skin cancer detection and the accuracy achieved by CNN model showed better performance compared to DNN model and analysis is done for RGB images of skin cancer using CNN model. (Jayalakshmi and Sathiesh Kumar 2019) In this study it is proved that by comparing CNN and BN (Batch Normalization) CNN models with the results of accuracy, loss, precision, recall and F1 score values the BN CNN based system produces acceptable results and analysis is done for CT images of skin cancer detection. (Sedigh, Sadeghian, and Masouleh 2019) in this study it is demonstrated by comparing CNN and GAN (Generative Adversarial Network) algorithms with the results of accuracy, sensitivity, specificity and F1 score values the CNN algorithm showing better performance compared to DNN algorithm. The analysis is done for RGB images in classifying skin cancer. (Xia, Xu, and Nan 2017) in this study flower classification was done on Oxford - 17 flowers and Oxford - 102 flowers dataset using Inception V3 and it is reported that this technique is performing better compared with other methods. (Emara et al. 2019) in this study used CNN model and Inception V4 and it is described that CNN model shows better performance compared to Inception V4.
Our institution is passionate about high quality evidence based research and has excelled in various fields (Vijayashree Priyadarshini 2019; Ezhilarasan, Apoorva, and Ashok Vardhan 2019; Ramesh et al. 2018; Mathew et al. 2020; Sridharan et al. 2019; Pc, Marimuthu, and Devadoss 2018; Ramadurai et al. 2019). We hope this study adds to this rich legacy.

Although the proposed algorithm appears to be better than the conventional methods it has some limitations. The proposed method deals with pathologically proven data and it overcomes the class imbalance and overtraining issues, the performance level can be improved to a greater extent if the input data has more attributes. The performance metrics can be improved by modifying the existing algorithm by increasing or reducing the convolution stages. The proposed algorithm can be improved by considering input attributes like age, gender, stage of observation etc.

6. Conclusion

The innovative detection of skin cancer dermal cell analysis the inception V4 has appeared to be better compared with inception V3 CNN algorithm. The deviation between those 2 groups are acceptable ranges hence, it is concluded that the inception V4 CNN algorithm and inception V3 CNN algorithm can be applied in the medical field for the early prediction of skin cancer. It is concluded that an innovative detection of skin cancer using deep learning algorithms shows good performance in the medical field.

Declarations

Conflict of Interest: No Conflict of Interest in this manuscript.

Author Contribution

BN carried out detection of skin cancer studies, collected data, performance and analysis, statistical analysis and drafted the manuscript. RP was involved in the study, coordinated, guided and helped to draft the manuscript.

Acknowledgement

The authors would like to express their gratitude towards Saveetha School of engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.
**Funding:** We thank the following organizations for providing financial support that enabled us to complete the study.

1. SR Lingam clinic, Chennai.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

**References**


ISSN: 2237-0722
Vol. 11 No. 4 (2021)
Received: 16.05.2021 – Accepted: 08.06.2021


