Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
2326-9865

Auto Skin Tumour Classification Using CNN Framework with Tensorflow and Keras

[1]Shobha Y , [2]Dr. K. V. Prasad, [3]Dr. Anuradha SG, [4]Dr. Hanumesh Vaidya,

[1]Research Scholar, Department of Studies in Computer Science, Vijayanagara SriKrishnadevraya University, Ballari, Karnataka-India.

[2]Department of Studies in Mathematics, Vijayanagara SriKrishnadevaraya University, Ballari, Karnataka-India

[2]Professor, Department of Computer Science and Engineering, Rao Bahadur Y. Mahabaleshwararappa Engineering College, Ballari, Karnataka-India.

[4]Department of studies in Mathematics, Vijayanagara SriKrishnadevaraya University, Ballari, Karnataka-India.

Article Info
Page Number: 172 - 184
Publication Issue:
Vol 72 No. 1 (2023)

Article History
Article Received*: 15 October 2022*
Revised*: 24 November 2022*
Accepted*: 18 December 2022*

Abstract
Skin tumour classification is emergent and most challenging problem in the medical diagnosis because of its similarity in the patterns of tumour cells with other symptoms & diseases on the victim body. An automatic and robust system is required for the early detection of the diseased providing aid in the field of computer assisted medical diagnosis thus decreasing the mortality rate. Machine learning and Deep Learning play a significant role in transforming the health care sector. Convolutional Neural Network (CNN) is used in healthcare sector for assisting to generate massive amount of health records, analyse thousands of health records and further provide insights on clinical decisions to service providers. In this article, we propose robust solution to detect and classify the skin tumour: Benign vs. Malignant using CNN architecture, TensorFlow and Keras on publicly available ISIC (International Skin Imaging Collaboration) data set. The results obtained in the experimental study records 0.34% loss with accuracy of 80.72% in the validation set and 82.75% on the test set.

Keywords: CNN, TensorFlow, Keras, Benign and Malignant

1. INTRODUCTION

Cancer is one of the most significant factors in the world's disease burden in the present decade and is expected to increase of new cases more than 15 million in 2023.It is one of the major root cause for mortality worldwide [1], As reported by WHO, the cancer-related mortality in the Middle East will be increased by 80-100% in coming 15 years. The most common cancer in the world is skin cancer and males are ranked first and females are ranked second [2]. Skin cancer accounts for 25% to 32.7% of all cancers. The source of skin cancer is due to changes in the characteristic feature of

natural skin which become malignant, where cells keep disassociate into irregular appearances and these cells cannot be controlled due to the damage of DNA. On the basis of pathological perception, skin cancer has unnatural formation with cell discrimination in different proportion of chromatin, nucleus, and cytoplasm [3]. Amongst all skin cancers, malignant and benign are the deadliest. Benign tumours are non-cancerous growths in the body. Benign and Malignant tumours can be distinguished as the former are basically harmless, latter are usually harmful.

In contradiction to cancerous tumours, benign tumours do not spread from the genesis to different sections of the body. Skin cancer damages to individual of all skin tones. Skin cancer causes predominantly on areas of sun-exposed skin i.e. exterior body parts that not regularly exposed to the sun, for example palms and soles of the feet, toe nails and other private areas of the human body. Efficient, accurate and early diagnosis of skin cancer plays a critical role in the process of healing, appropriate medical diagnosis, could help in avoiding the worst results of skin cancer and helps in taking preventive measures to prevent deaths. Hence, the system has to detect before time needed that can be facilitated and to create maximum general consciousness in the recognition of different types of skin tumour or other skin disorderliness such as a benign tumour on the skin that appear alike to skin cancer. Timely detection of skin cancer bestow patient with noble chance to get successful skin cancer treatment.

Various imaging techniques like demographic image processing techniques, optical coherence tomography, magnetic resonance imaging, and confocal scanning laser microscopy are currently being used to diagnose skin tumour. These processed images are conspicuously scanned by the skin specialists [4], that leads time-consuming and tedious procedure. This laborious process can be reduced with the improvement in identifying efficiently Computer-aided Diagnosis (CAD) systems are being used [5]. Nowadays, CAD is considered essential factor of the regular health examination in the clinic, which consists of natural unfinished image acquisition, pre-processing, Region of Interest (ROI) and feature extrication and lastly recognition [8,9]. The eventual part can be considered as classification which is crucial and challenging part due to varied dimension of the lesion region, inter-texture, skin tone, and the existence of various artifacts like reflections, hair, rolling lines, air bubbles, non-uniform vignetting, shadows, and markers [5, 6]. Currently several methods have been adopted for skin cancer detection. CNN is amongst most prominent algorithm in feature learning and object classification. CNN is equipped with the capability of training a network on a large-scale dataset to perform in a better way. In our research we have proposed a model AutoSTC (Auto Skin Tumour Classification). The Dataset to train AutoSTC is adopted from Kaggle, Skin Cancer Malignant Vs Benign archived from ISIC(International Skin Imaging Collaboration).

The remaining part of the paper is organized as follows: Section2 presents about the related work in the field of skin tumour classification. Section 3 provides the detailed explanation of pipelined CNN framework. Section4 provides insights on Materials and Methods. Section 5 provides the explanation on the proposed frame work of AutoSTC using CNN, Keras and TensorFlow. Results and Discussions are presented in Section 6 followed by conclusion.

2. RELATED WORK

Numerous researchers in the past have contributed their work in the field of skin tumour classification. Table 1 provides the brief overview on research gap identified in the area of computer assisted medical diagnosis for skin tumour classification. The most challenging issues in the medical field are to deal with the imbalance in the datasets associated with various healthcare records and the availability of the ground truth records for accurate diagnosis.

Table 1: Research Gap identified in the area of skin tumour classification

AUTHOR	METHODOLGY	RESEARCH GAP
Houssam Benbrahim et al., [7]	CNN with TensorFlow and Keras.	● To enhance feasibility and reliability of the model. ● To maximize the system performance with the use of TensorFlow on Spark. To port the experimental study on Big data environment
Aishwarya Dutta et al.,[8]	DCNN with Transfer Learning.	● The proposed framework will be tested on different datasets of dermoscopic images and to employ this framework on different domains for recognition to verify adaptability and generality.
AdekanmiAdegun et al., [9]		● To provide solutions to the challenges in analyzing skin lesions images fine-grained appearance of the skin lesion images.
Karl Thurnhofer-Hemsi et al., [10]	Deep Network with DenseNet201,GoogleNet, InceptionV3,MobileNetV2	● To achieve solution for imbalance of dataset used in training and testing dataset.
Ardan Adi Nugroho et al., [11]	Convolution Neural Network.	● System found difficult to identify bcc disease ● System found suitable to identify akiee disease
Taher M.Ghazal et al.,[12]	DCNN with AlexNet	● To improve the accuracy of model.
Muhammad AttiqueKhan et al.,[13]	DCNN, RESNET-50, RESNET-101 and SVM.	● Feature selection based on auto encoder.

Yuexiang Li et al.,[14]	CNN	● To create a benchmark for subsequent related research.

3.CONVOLUTION NEURAL NETWORK(CNN)

Convolution layer consists of stack of filters known as kernels. These filters when tested with an input image the resultant is generation of feature map of an input image. The task of convolution layer is to extract the feature representation of their input image. CNNs are utilized in the recognition of an abject with in image, classifying of images and to point the region of interest of an image. The generic CNN framework with input layer, convolutional layer, pooling layer, fully connected layer and output layer is represented in the figure1 as shown below:

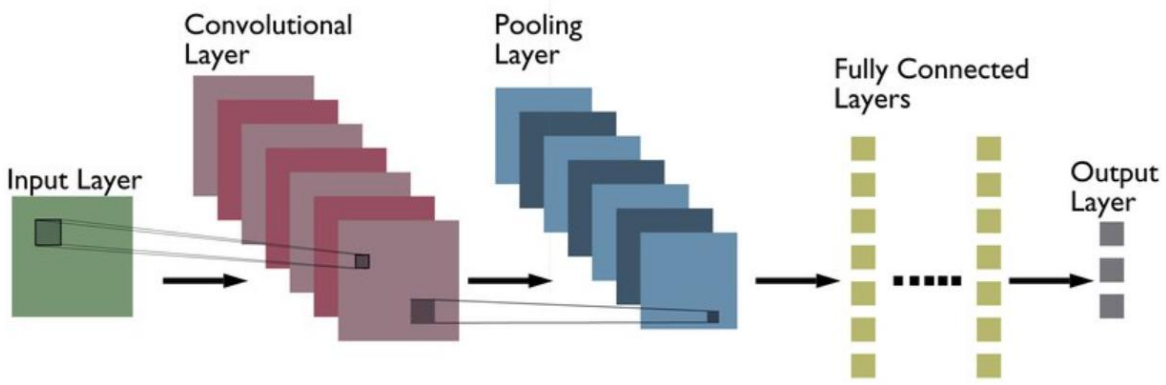


Figure 1: Generic CNN Framework

- **Pooling layer:** The motive behind the pooling layer is to lessen spatial resolution of an input which makes process effortless with limited usage of memory.
- **Fully Connected Layer:** Each one neuron in CNN is connected to every single neuron from their precedent layer. Fully connected layers customarily headed to the rearmost of a CNN. Fully connected layer interprets the feature representation and yield top level reasoning.

4.MATERIALS AND METHODS

Dataset

The dataset consists 3600 images of benign skin moles and malignant skin moles. The dataset is balanced. The dataset contains 2 major folders namely test and train. Both the train and test folder contain 2 folders namely benign and malignant. Inside the train folder, the benign folder has 1440 images while the malignant folder has 1197 images. Inside the test folder, the benign folder has 360 images while the malignant folder has 300 images. All the images have the same dimension – 224×244.

5.PROPOSED FRAMEWORK

AutoSTC consists of input layer Gaussian noise layer followed by 9 Convolutional layers in 3 blocks, each block encompasses with a max pooling layer which reduce the computational complexity by reducing the number of connections between convolution layer to minimize the over fitting. Each block of convolution undergoes batch normalisation. During the process of training AutoSTC, a deep neural network with collective hidden layers that are dissimilar in pattern and their connectivity are being used.

The first layer undergoes convolution operation to the input data, to identify various patterns or characteristic attributes from an input image, next, it disseminates the results to the succeeding layer. The subsequent, second layer is Gaussian noise layer. Keras aids in the inclusion of Gaussian noise via a distinct layer called the Gaussian Noise layer. This layer can be used to add noise to an existing model.

The third layer is convolution layer. Convolution operation is applied to the input data, to discover various patterns or features from an input image, then the resultant are passed to the next layer. So, this layer defines a filter (kernel). Pooling layers are used in CNN for integration of features learned by the convolutional layer feature map. It primarily used in the minimization of overfitting during the training period of the model by condensing the features in the feature map. Different pooling layers make use of different pooling techniques such as maximum, average, gated and tree pooling to name a few. The most popular approach is Max Pooling. Each convolution block ends with batch normalisation in which permits each layer of the network to learn more independently. It is used to normalize the output of the previous layers. With the batchnormalisation learning becomes efficient and also it can be utilized for regularization to avoid overfitting of the model. This layer gets added to the sequential model to standardize the input or the outputs. It can be used at different points in between the layers of the model. It is repeatedly located after defining the sequential model and after the convolution and pooling layers.

The third block of Convolution succeeded with Fully-Connected layers which take the high-level filtered images from the previous layer and convert them into a vector. Dropouts are the regularization technique that is used to prevent overfitting in the model. Dropouts are to enhance the learning of the model. It is advised not to use Dropouts after the convolution layers, and is oftenly used after the dense layers of the network.

Table.2 Description related to various layers of AutoSTC

Type of layers	Output Shape	Parameters
input_3 (Input Layer)	(None, 224, 224, 3)	0
To detect different features from an input image, and transmit the result to the next layer.		
(GaussianNoise)	(None, 224, 224, 3)	0

This layer can be used to add noise to an existing model		
conv2d_12 (Conv2D)	(None, 111, 111, 16)	448
conv2d_13 (Conv2D)	(None, 55, 55, 16)	2320
To detect various patterns of an input image. Whenever the program glides the filter over the image , the dot product of the filter to the image is applied which results into new image with verges.		
max_pooling2d_6 (MaxPooling2D)	None, 27, 27, 16)	0
*minimizes the number of parameters within the model. It turns the low-level data into higher-level information		
batch_normalization_8	None, 27, 27, 16)	64
**Batchnormalisation during this process each layer of the network is allowed to learn more independently.		
conv2d_14 (Conv2D)	(None, 25, 25, 32)	4640
conv2d_15 (Conv2D)	(None, 23, 23, 32)	9248
*max_pooling2d_7(MaxPooling2D)	(None, 23, 23, 32)	0
**batch_normalization_9	(None, 22, 22, 32)	128
conv2d_16 (Conv2D)	None, 20, 20, 64	18496
conv2d_17 (Conv2D)	(None, 18, 18, 64	36928
*max_pooling2d_8(MaxPooling 2D)	(None, 17, 17, 64)	0
**batch_normalization_10	(None, 17, 17, 64)	256
flatten_2	(None, 18496)	0
Flattening converts the data into oneDimensional array to create a single long feature vector connected to fully connected layer		
dense_4 (Dense)	(None,16)	295952
***It is used to classify image based on output from convolution layer		
**batch_normalization_11	(None, 16)	64
dropout_2 (Dropout)	(None, 16)	0
***Dense_5 (Dense)	(None,1)	17

A. KERAS

Keras is a kind of most prominent high-level neural networks APIs. It is written in Python and supports multiple back-end neural network computation engines. “Keras is an open-source software library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow library. Up until version 2.3, Keras supports multiple back ends, including

TensorFlow, Microsoft Cognitive Toolkit, Theano, and PlaidML". It emerged as part of the research hardship of project "ONEIROS (Open-ended Neuro-Electronic Intelligent Robot Operating System)", and its original author and developer is François Chollet, a Google engineer. The contemporary version of Keras is 2.9.0(May 13, 2022)

Features of Keras

Keras supports various optimization techniques to make high level neural network API easier and more efficient. It supports the following features −

- Consistent, simple and extensible API.
- Minimal structure - easy to achieve the result with no ruffles.
- It supports multiple platforms and backends. It runs on both CPU and GPU.

Applications of Keras

- Keras is utilized for creation of deep models which can be productized on smartphones.
- Keras also aids in distributed training of deep learning models.
- Keras is used by companies such as Netflix, Yelp, Uber, etc.

B. TENSORFLOW:

TensorFlow tutorial is developed for both beginners and professionals. It is a kind of popular deep learning framework, developed by Google Team. It is a free and open-source software library and designed using python programming language. It is a software framework, planned by the Google team for the implementation of machine learning and deep learning concepts in the easiest manner. It is associated with computational algebra of optimization techniques for ease of calculation for several mathematical expressions. TensorFlow aids in training and rundeep neural networks for handwritten digit classification, image recognition, word embedding and creation of various sequence models. "TensorFlow is a free and open-source software library for machine learning and artificial intelligence"

- TensorFlow was created and developed by the Google Brain team which was basically created for internal Google utilisation in research production. The primary version was released under the Apache License 2.0 in 2015 and updated version is Tensorflow2.0 was released by Google in September 2019.TensorFlow is cross platform which has got flexible architecture that allows to run on CPUs, GPUs and TPUs. It provides APIs in multiple languages (Python, C++, JAVA).

C. SYSTEM EVALUATION

The proposed AutoSTC was evaluated using confusion matrix of "True Positive (TP)", False positive (FP)", "True Negative (TN)" and "False Negative (FN)". The accuracy, recall, precision andF1-Score performance measures as represented in the equation 1,2,3,4 of the system are used to classify Benign and Malignant type of tumour. The precision value of the system indicates percentage of correctly classified positive patients from all positive recognition. The harmonic

mean between the precision and recall is computed by F1-score as represented in equation.4. Further the ranking of the overall system is measured through the accuracy factor as represented in equation.1

True Positive (TP) indicates the condition while the data is positive and correctly predicted as positive. True Negative (TN) indicates the condition where the data is negative and correctly predicted as negative. False Positive (FP) indicates the conditions where the data is negative but incorrectly detected as positive, whereas False Negative (FN) indicates the conditions where the data is positive, but incorrectly detected as negative.

$$Accuracy = \frac{(Tn + Tp)}{(Tn + Fp + Tp + Fn)} - eqn.1$$

$$Recall = \frac{Tp}{Tp + Fn} - - - - - - - - eqn.2$$

$$Precision = \frac{Tp}{Tp + Fp} - - - - - - eqn.3$$

$$F1 - Score = 2\,X\frac{Precision\ *\ recall}{Precision\ +\ recall} - eqn.4$$

6. RESULTS AND DISCUSSION

In this article, a method based on CNNs with TensorFlow and Keras model was proposed for the extraction of the skin cancer in JPG images. The input layer of the architecture consists of 224x224 images with the output layer of two neurons specifying two classes of skin cancer respectively. MaxPooling, batch normalization and dropout operations are performed in the pipelined architecture in various layers of CNN. The architecture succeeded in the classification of dermoscopic images related to skin lesions in two classes Benign and Malignant. Confusion matrix as represented in the figure.2is implemented to study the adaptability and generality of the machine learning model.

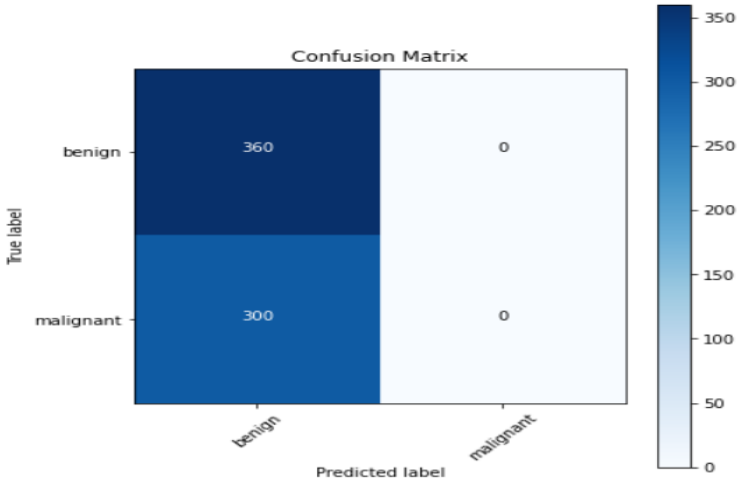


Figure 2: Confusion Matrix representation for Test Case

In the experimental study, out of the 660 data samples in ISIC datasets, 360 samples are of Benign (Non Tumour cells) and 300 samples are of Malignant (Tumour cells). Out of 360 Benign samples, 360 samples are classified correctly and thus FP samples are 0. Similarly, out of 300 Malignant samples all are classified correctly and thus TN is 0. The output of the confusion matrix in AutoSTC mapped to four different categories –True Positive, True Negative, False Positive and False Negative.

- **True Positive (TP)**: A patient who is actually affected with any of the two tumours.
- **True Negative (TN**): A Patient who is not actually affected with any of the two tumours.
- **False Positive (FP):**A patient who is not actually affected with tumour but classified as tumour
- **False Negative (FN)**: A patient who is actually affected with any of the two tumours but not classified as tumour. The results obtained on the ISIC test dataset (see in the Table 3) for the classification of skin tumour are presented.

Table 3: Classification report for AutoSTC with the performance metrics

	PRECISION	RECALL	F1 SCORE	IMAGES
BENIGN	0.55	1.00	0.71	360
MALIGNANT	0.00	0.00	0.55	300

Training loss and validation loss in the experimental work are as represented in the figure.3

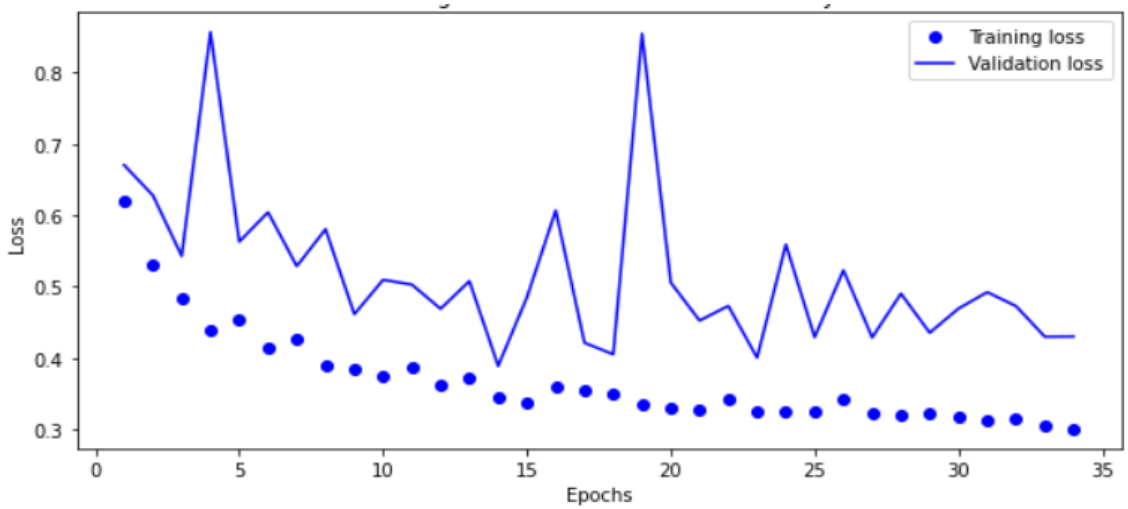


Figure 3: Training and Validation Loss Representation per Epoch

Similarly training accuracy and validation accuracy of the classification problem are represented in the figure.4

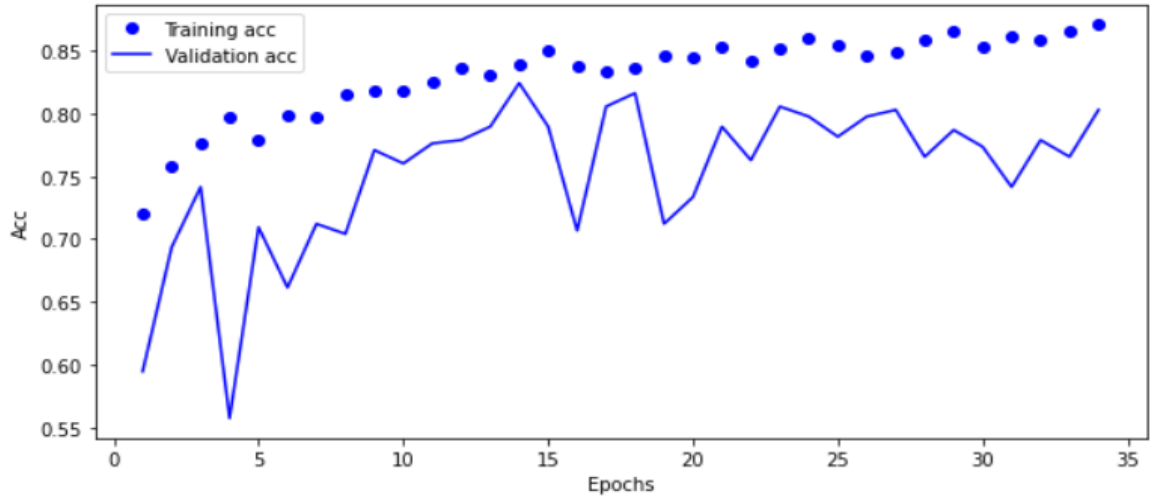
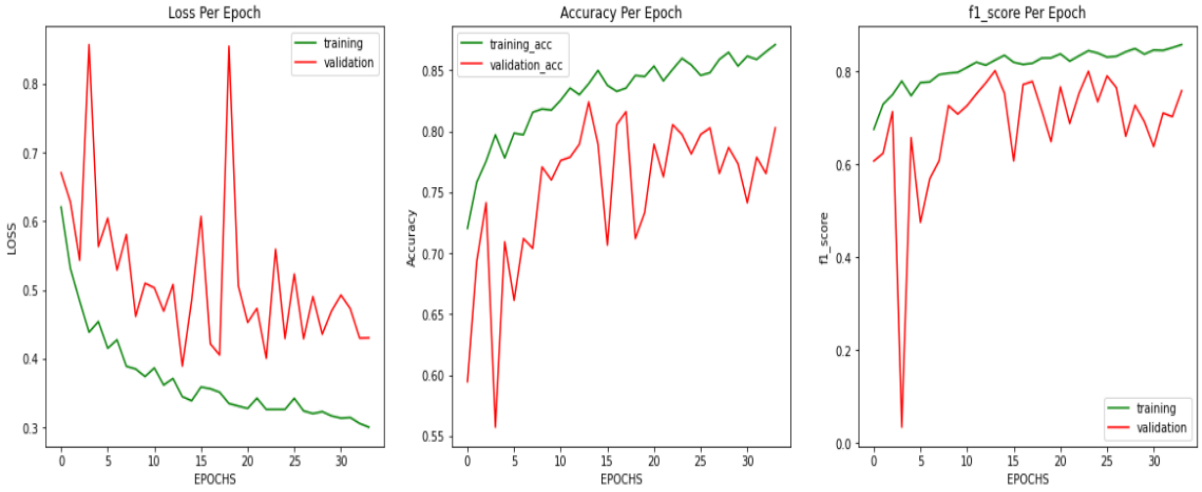


Figure 4: Training and Validation Accuracy Representation per Epoch

The experimental study proved demonstrates that AutoSTC can reach a very high accuracy of 80.72% in the validation set and 82.75% in the test set with epochs equal to 100 as shown in fig.3.The system subjected to testing yields loss of 0.34 in the classification of tumour.



(a) **(b)** **(c)**

Figure 5: (a), (b), (c) Loss, Accuracy and F1-Score Graphical representation per Epoch

Table 4: State-of-art comparison with AutoSTC with reference to training, testing and validation on ISIC 2017 data set

Authors	Year	Dataset	Recall	Precision	F1-Score

AlexNet[15]	2018	ISIC-2017(3 tumour classes)	0.34	0.65	-
ResNet-101[16]	2018	ISIC-2017(3 tumour classes)	0.34	0.71	-
Method-1 [18]	2018	ISIC-2017(3 tumour classes)	0.61	-	-
Method-2 [19]	2018	ISIC-2017(3 tumour classes)	0.57	0.68	-
GR [20]	2019	ISIC-2017(3 tumour classes)	0.0	0.0	0.0
SLC [21]	2020	ISIC-20173 tumour classes	0.73	0.73	0.76
AutoSTC	2022	ISIC(2 tumour classes)	1.00	0.55	0.71

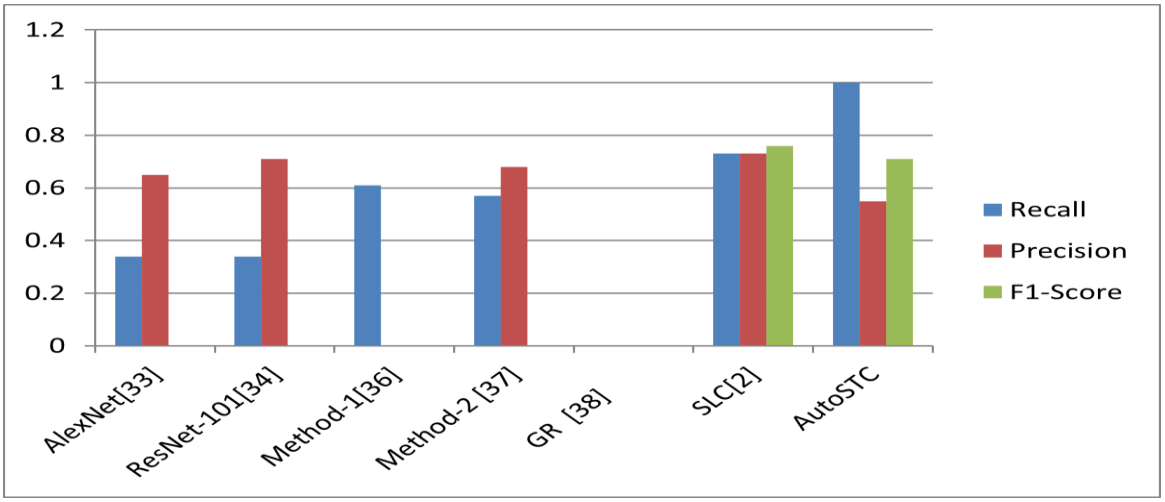


Figure 6: Precision, Recall, F1-score (Auto –STC) comparison with state-of-the-art methods

CONCLUSION

In this study, CNN architecture with Keras and TensorFlow framework is developed to classify two classes of skin tumours: Malignant and Benign. The goal is to build an automatic and robust solution that can assist the doctors to diagnose, detect and identify the diseased patient and further classify the tumours. The proposed framework was trained and tested in the open-source dataset ISIC (International Skin Imaging Collaboration) and has undergone evaluation via broad range of experiments. The results obtained in the experimental study show that the accuracy of the

adopted model with Gaussian noise layer and dropout layer provides 80.75% in validation set and 82.57 % in testing set with epochs equal to 100.Thus serves as a better configuration in comparison with state-of –art methods. To evaluate the overall efficiency of the model, confusion matrix is implemented resulting in 1.0 (recall), 0.71 (F1score) and 0.55 (F1 score) for Benign tumour class and malignant class respectively. Our future work heads in two directions – first to conduct the experimental study on huge dataset with increased number of tumour classes. Later is to build the framework for Big Data environments with the prime concern to prove the reliability and feasibility of the study.

REFERENCES

[1].R. L. Siegel, K. D. Miller, and A. Jemal, Cancer statistics, 2015, CA: a cancer journal for clinicians, 2015, vol. 65, no 1, pp. 5-29

[2].Siegel, R. L., Miller, K. D., Jemal, A.: Cancer statistics, 2020. CA: A Cancer Journal for Clinicians 70(1), 7–30 (2020).

[3].BuljanM, Bulana V and Sandra S 2008Variation in Clinical Presentation of Basal Cell Carcinoma (Kroasia: University Department of Dermatology and Venereology Zagreb Croatia) p 25-30.

[4] .Smith, L., MacNeil, S.: State of the art in non-invasive imaging of cutaneous melanoma. Skin Res. Technol 17(3), 257–269 (2011)

[5] .Hasan, M. K., Dahal, L., Samarakoon, P. N., Tushar, F. I., Mart´ı, R.: DSNet: Automatic dermoscopic skin lesion segmentation. Computers in Biology and Medicine 120, 103738 (2020)

[6] .Mishraa, N. K., Celebi, M. E.: An overview of melanoma detection in dermoscopy images using image processing and machine learning. arXiv:1601.07843 (2016)

[7].Benbrahim, H., Hachimi, H., & Amine, A. (2020). Deep convolutional neural network with tensorflow and keras to classify skin cancer images. *Scalable Computing: Practice and Experience*, *21*(3), 379-390.

[8]. Dutta, A., Hasan, K., & Ahmad, M. (2021). Skin lesion classification using convolutional neural network for melanoma recognition. In *Proceedings of International Joint Conference on Advances in Computational Intelligence* (pp. 55-66). Springer, Singapore.

[9]. Adegun, A., &Viriri, S. (2021). Deep learning techniques for skin lesion analysis and melanoma cancer detection: a survey of state-of-the-art. *Artificial Intelligence Review*, *54*(2), 811-841.

[10].Thurnhofer-Hemsi, K., & Domínguez, E. (2021). A convolutional neural network framework for accurate skin cancer detection. *Neural Processing Letters*, *53*(5), 3073-3093.

[11]. Nugroho, A. A., Slamet, I., &Sugiyanto. (2019, December). Skins cancer identification system of HAMl0000 skin cancer dataset using convolutional neural network. In *AIP Conference Proceedings* (Vol. 2202, No. 1, p. 020039). AIP Publishing LLC.

[12]. Ghazal, T. M., Hussain, S., Khan, M. F., Khan, M. A., Said, R. A., & Ahmad, M. (2022). Detection of Benign and Malignant Tumors in Skin Empowered with Transfer Learning. *Computational Intelligence and Neuroscience*, *2022*.

[13]. Khan, M. A., Javed, M. Y., Sharif, M., Saba, T., & Rehman, A. (2019, April). Multi-model deep neural network based features extraction and optimal selection approach for skin lesion

classification. In *2019 international conference on computer and information sciences (ICCIS)* (pp. 1-7). IEEE.

[14]. Li, Y., & Shen, L. (2018). Skin lesion analysis towards melanoma detection using deep learning network. *Sensors*, *18*(2), 556.

[15].Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, *25*.

[16]. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

[17]. Li, Y., Shen, L.: Skin lesion analysis towards melanoma detection using deep learning network. Sensors 18(2), 556 (2018).

[18]. Yang, J., Xie, F., Fan, H., Jiang, Z., & Liu, J. (2018). Classification for dermoscopy images using convolutional neural networks based on region average pooling. *IEEE Access*, *6*, 65130-65138.

[19] Sultana, N. N., Mandal, B., &Puhan, N. B. (2018). Deep residual network with regularised fisher framework for detection of melanoma. *IET Computer Vision*, *12*(8), 1096-1104.

[20].Serte, S., &Demirel, H. (2019). Gabor wavelet-based deep learning for skin lesion classification. *Computers in biology and medicine*, *113*, 103423.

[21].Ries, L. A., Harkins, D., Krapcho, M., Mariotto, A., Miller, B. A., Feuer, E. J., ... & Edwards, B. K. (2006). SEER cancer statistics review, 1975-2003.