

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Volume 2, Issue 2, April 2022

Machine Learning Algorithm for Detection of Deadliest Forms of Skin Cancer

Mohammad Awais¹, Adiya Parveen², Afrin Sheikh³, Shifa Sheikh⁴, Iqra Qureshi⁵, Prof. Mohsina Anjum⁶, Prof. Syed Irfan Ali⁷

Students, Department of Electronics and Telecommunication Engineering^{1,2,3,4,5} Assistant Professor, Department of Electronics and Telecommunication Engineering^{6,7} Anjuman College of Engineering and Technology, Nagpur, Maharashtra, India

Abstract: Skin cancer is one of the most growing types and dangerous cancer in the world. The early diagnosis of melanoma and other ski cancer is a critical issue for dermatologists. In this paper, we using Machine Learning Algorithm for Detection of Deadliest Forms of Skin Cancer. This project aims to develop a skin cancer detection ML Model which can classify the skin cancer types and help in early detection. The ML Model is developed in Dot Net (. Net). The model is developed and tested with different network architectures by varying the type of layers used to train the machine. Basically our model use DNN (Deep Neural Network) and ResNet50 for detection of skin cancer. The model will be tested and trained on the dataset collected from the International Skin Imaging Collaboration (ISIC).

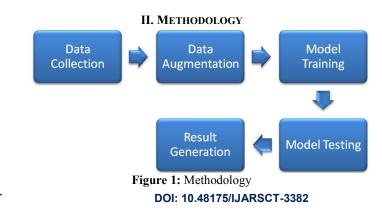
Keywords: Skin cancer, Melanoma, Machine learning, DNN (Deep Neural Network), ResNet50

I. INTRODUCTION

Dermatology is the branch of medicine that concerned with the diseases' diagnosis of skin, hair and nails, the skin is the most important part in the human body which protects the internal parts from the outside world. The skin cancer is the most important of these diseases; it can be growing at any part in the body and occurring from non-pigmented cells [1].Skin cancer is one of the most growing types and the most dangerous in the world of cancer; the important of these tumors is malignant melanoma, the rates of melanoma have been rising for at least 30 years.

The main risk of melanoma is could be spread entire the body by lymphatic vessels and blood vessels [2], thus the main strategy is the early diagnosis of melanoma and removal of thin melanoma; it is the most common cancer if diagnosed at an early stage can be cured without complications. Therefore, the early diagnosis of skin cancer and melanoma is a critical issue and the main challenges for dermatologists to reduce mortality and morbidity [3]. It is very difficult the diagnosis of skin cancer by naked eye using the features that recognize a benign from skin cancer when using the clinical practice.

Nowadays the computer technology in medical decision support is used widespread and pervasive across a wide range of medical area, such as cancer research, dermatology[4], thus the computerized methods analysis of dermoscopy images allows overcome various issues which help the dermatologist to take less time and high efficiency in diagnosis of skin cancer. It is important to develop of various computerized method for clinical diagnosis using machine learning system, which these system gives to dermatologist different analysis texture, pigmentation, colour by using classification methods.



Copyright to IJARSCT www.ijarsct.co.in



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Volume 2, Issue 2, April 2022

III. PROPOSED SYSTEM

The proposed system should be able to:

- **Data Collection:** For facilitating the computer-aided diagnosis related work on melanoma many datasets can be found online. The International Skin Imaging Collaboration (ISIC) is an academia and industry partnership designed to facilitate the application of digital skin imaging to help reduce melanoma mortality.
- Data Augmentation: Data augmentation is a strategy that enables practitioners to significantly increase the diversity of data available for training models, without actually collecting new data. Data augmentation techniques used to train large neural networks. As the modern neural networks contain millions of parameters, it takes a proportional amount of data to result considerable performance [5]. So, it is understandable that deep learning models require a large dataset of images to provide desired performance. But it is a wild goose chase to look for novel images in a huge quantity. Instead, minor alterations to the existing data will pull the trick. A poorly trained neural network considers these slightly modified images as distinct ones.
- Mode Training: As the modern deep neural networks require millions of images to train the entire model, it is
 inconvenient to collect that amount of data for a particular task. Therefore, a method called transfer learning has
 been getting attention for training a deep neural network with a limited amount of data. Transfer learning is the
 method of taking a pretrained model and fine-tune the model with a smaller dataset. ResNet-50 is a deep neural
 network that is 50 layers deep. You can load a pretrained version of the network trained on more than a million
 images from the ImageNet database [6]. The pretrained network can classify images into 1000 object categories

	it Project Build Debug T	est Analyze Tools Ext	ensions Window			P test1		M – Ø
- 🎯 🍄 • 🗃 🔛 🖀	タ・ペー Debug ・ Any	CPU • • test1 •	> - (-) 🛱 🔤					년 Live Share
MLModel1.mbconfig* +	Program.cs							
Data	Training setup summan	$^{\prime}$						
Train 중 Feedback	Cancel training	C Training						
Dutput								+ û
	nine Learning - MLModel1	• 1.2.1 • •	×≡ ;;					
	C:\Users\MY_PC\AppData\Local\	T> MI VCT] ->]> MI M-d						
I Trainer		curacy MacroAccuracy Du						
[Source-AutoML, Kind				t., Kind-Infol D	wnloading resnet v2 50	299.meta from https://aka.ms/mlnet	resources/meta/resnet v2 50	299, meta to C:\Users\MY PC\An
						wnloaded 4096 bytes out of 1026169		
						wnloaded 10264576 bytes out of 102		
						wnloaded 20525056 bytes out of 102		
						wnloaded 30785536 bytes out of 102		
						wnloaded 41050112 bytes out of 102		
						wnloaded 51310592 bytes out of 1020 wnloaded 61571072 bytes out of 1020		
						whileaded 01571072 bytes out of 1020 whileaded 71835648 bytes out of 1020		
						wnloaded 82096128 bytes out of 102		
						wnloaded 92356608 bytes out of 1020		
						wnloaded 102616931 bytes out of 10		
[Source=AutoML, Kind	-Info] [Source-ImageClassifica	tionTrainer; Ensuring met	files are preser	t., Kind=Info] r	esnet v2 50 299.meta: Do	wnload complete		
	=Trace] [Source=ImageClassific							
						: Validation, Batch Processed Count		
	-Trace] [Source-ImageClassific							
	<pre>-Trace] [Source-ImageClassific =Trace] [Source=ImageClassific</pre>					I: Validation, Batch Processed Count I: Train, Batch Processed Count		
						: Validation, Batch Processed Count		
	-Trace] [Source-ImageClassific							
						: Validation, Batch Processed Count		
	-Trace] [Source-ImageClassific							
[Source=AutoML, Kind	=Trace] [Source=ImageClassific	ationTrainer; ImageClassi	icationTrainer, k	ind=Trace] Phase	Training, Dataset used	: Validation, Batch Processed Count	: 1, Epoch: 4, Accuracy	: 1, Cross-Entropy:
	-Trace] [Source-ImageClassific							: 0.9666667, Cross-Entropy:
						: Validation, Batch Processed Coun		
	-Trace] [Source-ImageClassific							
						: Validation, Batch Processed Coun		
	=Irace] [Source=ImageElassific					I: Irain, Batch Processed Count I: Validation, Batch Processed Count		
	=Trace] [Source=ImageClassific =Trace] [Source=ImageClassific							
						: Validation, Batch Processed Count		
	-Trace] [Source-ImageClassific							
Course Alacian 192-3	T			·····	TITLE BALLINE	. V.113.43 B.4.6 B		
ady							↑ Add to Source Cor	atrol 🔺 🚸 Select Repository 🔺
P Type here to	an a walk	0 11		-	^		- 1345 11	03:01 PM
	searcn	O 🗄 🔳					🥥 43°C Haze ∕	 Image: A constraint of the second seco

IJARSCT



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Volume 2, Issue 2, April 2022

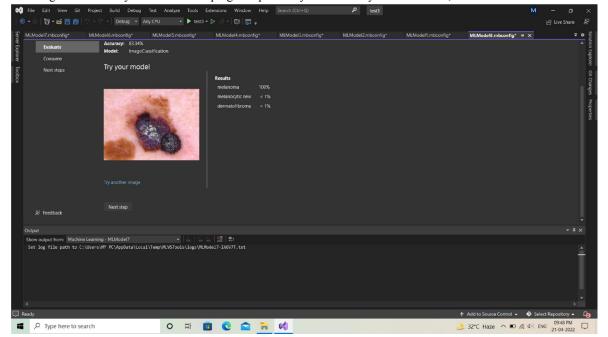
0	File	Edit View Git	Project Build Debug	Test Analyze Tools Ex	tensions Window	Help Search (C	(trl + Q)	۶ _{test3}		N	- 0	×	
3		1 🔁 - 🖬 🖻 🛛		Any CPU 🔹 🕨 test3 🔹							년 Live Share		
Serv	MLMod	del7.mbconfig*	MLModel6.mbccnfig*	MLModel5.mbconfig*	MLModel4.mbconfig	MLMode	el3.mbconfig* M	ILModel2.mbconfig*	MLModel1.mbccnfig*	MLModel8.mbconfig* 🕫 🗙	Ŧ	alos 1 1 2	
Server Explorer Toolbox												tion E	
lorer		Scenaric	Train									on Explorer	
8		Environment	Model Builder aut	omatically sets the training time ba	ised on the size of your i	dataset.							
Š		Data	Training setup sun	Training setup summary 🗸									
		Train	Train again	Training complete								Git Changes Properties	
		Evaluate										Prop	
		Consume	Training res									erties	
		Next steps	Best accuracy: Best model:	83.34% DNN + ResNet50									
			Training time:	1580.20 seconds									
			Models explored Generated code-b	(total): 1 whind: MLModel8.consumption.cs	s, MLModel8.training.cs								
			Next step										
		Feedback											
		recuback											
		output from: Machine		cal\Temp\//L/STocls\logs\//LMoc									
	Sec 1	log file path to C:	(Users (HT PC (AppData (Lo	car/remp/nLvsrocis/rogs/nLnoc	IEI/-IADV/I.txt							<u> </u>	
	Ready									Add to Source Control 🔺 🚸		G	
		Type here to sea	arch	O H I	0	H 📢				32°C Haze ∧ ■ @ Φ)	0947 PM		
	- /	.,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,								SE C TIME THE AN AN	ENG 21 04 2022	~	

Figure 2: Model Training

• **Model Testing:** This step utilizes the weights achieved from training step to predict the classes for the images in the test-set. As the training has been done using a five-fold cross validation approach, ML Model needs to execute the testing in the similar manner. ML Model also provides an option for the researcher to visualize the detection on test data. In this work, this monitoring approach has been proved to be useful for finding the difficulties the model is facing while going through this test-phase. For example augmentation had been found to generate some artifacts in the augmented data which was hurting the performance of the model.

IV. RESULT GENERATION

In this MLModel shows the result by analysing and evaluating the image. For every sets of observation, the observation with the highest sensitivity value while keeping the specificity and accuracy above 80.00%, has been chosen as the best one.



Copyright to IJARSCT www.ijarsct.co.in

IJARSCT



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

90	File Edit View Git Proje	ct Build Debug Test Analyze Tools Ex	tensions Window Help	Search (Ctrl+Q)	P test3		M – Ø	×
)•0 御•醫醫 ♡•	🖓 - Debug 🔹 Any CPU 🔹 🕨 test3 •	▷ ő - 🛱 🖥 =				Live Share	
Sen	MLModel5.mbconfig* + × MU			MLModel1.mbconfig*	MLModel8.mbconfig*			• • S
ver Exp	Evaluate	Accuracy: 81.67% Model: ImageClassification						4
Explorer	Consume	model. indyclassification						xploi
	Next steps	Try your model						e e
Toolbox			Results					Sit Ch
			melanocytic nevi 92%					ppne
			melanoma 6%					a v
	R feedback	Try enother image Next step	dematofibroma 3%					Solution Explorer Git Changes Properties
							- q	×
	Show output from: Machine Learn	ng - MLModel5 •	. 🛎 🗱					
	Set log file path to C:\User	<pre>\/WY PC\AppOata\Local\Temp\MLVSTools\logs\MLMon</pre>	del5-EBJ0T1.txt					
R	4 					🛧 Add to Source Control 🔺 🚸	Celect Perperitory	- Go
							00:45 044	
	P Type here to search	O 🖽 🖪				232°C Haze ∧ ■ @ Φ	ENG	

Volume 2, Issue 2, April 2022

Figure 3: Result Generation

V. CONCLUSION

The proposed work was aimed to develop a model for detecting melanoma, Melanocytic and Dermatofibroma in images by using a deep learning model. By the end of this work, a model, named "ML Model", has been developed. ML Model uses the DNN and ResNet50 network as the architecture for the deep learning model. An available dataset of dermoscopic images, provided by International Skin Imaging Collaboration (ISIC), has been utilized to train and evaluate the performance of MelNet on the dermoscopic data. The digital dataset, provided by ISIC, has been used to train and analyse MelNet's performance on the digital data. Every image from the datasets was annotated according to the clinical classification of the mole. For making the system robust, the data was augmented by operating several augmentation methods.

During the training phase,

ACKNOWLEDGMENT

We would like to thank **Prof. Ruhina Quazi** Project Co-ordinator and **Prof. Mohsina Anjum** and for her invaluable supervision during the course of this project. Our gratitude extends to the Faculty of **A.C.E.T** for the opportunities provided at the Department of **Electronics And Telecommunication**. Additionally, we would like to express gratitude to **Prof. Syed Irfan Ali** for his treasured supervision, support, and tutelage which was really influential in shaping our experiment methods and critiquing our results. We would like to thank our friends for a cherished time spent together in the lab, and in social settings. Our appreciation also goes out to our family and friends for their encouragement and support all through our studies.

REFERENCES

- [1]. K. Korotkovand R. Garcia, "Computerized analysis of pigmented skin lesions: A review", Artificial Intelligence in Medicine, 56 (2),2012,pp. 69–90.
- [2]. J. Scharcanski, M. E. Celebi and S. Service, "Online. Computer Vision Techniques for the Diagnosis of Skin Cancer", 2014, pp. 193–219.
- [3]. L. Li, Q. Zhang, Y. Ding, H. Jiang, B. H. Thiers and J. Z. Wang, "Automatic diagnosis of melanoma using machine learning methods on a spectroscopic system", BMC Medical Imaging, 14 (1),2014, 36
- [4]. E. Barati, M. Saraee, A. Mohammadi, N. AdibiandM.R. Ahamadzadeh, "A Survey on Utilization of Data Mining Approaches for Dermatological (Skin) Diseases Prediction", Journal of Selected Areas in Health Informatics, 2 (3),2011,pp. 1–11.

Copyright to IJARSCT www.ijarsct.co.in DOI: 10.48175/IJARSCT-3382



International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Volume 2, Issue 2, April 2022

- [5]. L. Perez and J. Wang, "The Effectiveness of Data Augmentation in Image Classification using Deep Learning," Dec. 2017.
- [6]. ImageNet. http://www.image-net.org