AI-BASED LOCALIZATION AND CLASSIFICATION OF SKIN DISEASE WITH ERYTHEMA

NALAIYA THIRAN IBM PROJECT REPORT

TEAM ID:PNT2022TMID30346

Submitted by

V.JAYAPRIYA	**611419104022**
A.NISHANTHI	**611419104045**
M.SATHYA	**611419104072**
E.SUBASHINI	**611419104085**

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

MAHENDRA ENGINEERING COLLEGE FOR WOMEN,

TIRUCHENGODE,NAMAKKAL-637205.

ANNA UNIVERSITY:CHENNAI 600 025

1. INTRODUCTION

a. ProjectOverview

Now a day's people are suffering from skin diseases, More than 125 million sufferingfromPsoriasisalsoskincancerrateisrapidlyincreasingoverthelastfewdecades especially Melanoma is most diversifying skin cancer. If skin diseases are not treated at an earlierstage,thenitmayleadtocomplicationsinthebodyincludingspreadingofthe infection from one individual to the other. To overcome the above problem we are building amodelwhichisusedforthepreventionandearlydetectionofskincancer,psoriasis. Basically, skin disease diagnosis depends on the different characteristics like colour, shape, texture etc. Here the person can capture the images of skin and then the image will be sent the trained model. The model analyses the image and detect whether the person is having skin disease or not.

b. Purpose

Thediseasesarenotconsideredskindiseases,andskintoneismajorlysufferedfr om theultravioletraysfromthesun.However,dermatologistsperformthemajorityofnon-invasive screening tests simply with the naked eye, even though skin illness is a frequent diseaseforwhichearlydetectionandclassificationareessentialforpatientsuccessand recovery. Thecharacteristic of theskin images is diversified so thatitis achallengingjobto deviseanefficientandrobustalgorithmforautomaticdetectionofskindiseaseandits severity.Automaticprocessingofsuchimagesforskinanalysisrequiresquantitative discriminator to differentiate the diseases.

2. LITERATURE SURVEY

a. Existingproblem

Erythema is a broad category of skin condition that can impact any area of the skinand mucous membranes.Itusuallyoccursinresponsetodiseaseorinfectionin reactiontoadrug.Severityoftherashrangesfrommildtolifethreatening.Itisanabnormal redness of skin or mucous membranes. Capillary congestion causes the condition, and red splotches on the hands or feet are classic examples of it.

b. References

PAPER1: DeepLearninginSkinDiseaseImageRecognition:AReview

PUBLICATIONYEAR: 11November2020

AUTHORNAME:LING-FANGLI1,XUWANG1,WEI-JIANHU1,NEALN.XIONG2, (Senior Member, IEEE), YONG-XING DU 1 , AND BAO-SHAN LI1.

JOUNALNAME:IEEEACCESS

(https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=6287639)

SUMMARY:

Theapplication ofdeep learning methods to diagnosediseases has becomeanew research topic in the medical field. In the field of medicine, skin disease is one of the most

commondiseases,anditsvisualrepresentationismoreprominentcomparedwiththeother typesofdiseases.Accordingly,theuseofdeeplearningmethodsforskindiseaseimage recognitionisofgreatsignificanceandhasattractedtheattentionofresearchers.Inthis study,wereview45researcheffortsontheidentificationofskindiseasebyusingdeep learning technology since 2016. We analyze these studies from the aspects of disease type, data set, data processing technology, data augmentation technology, model for skin disease imagerecognition,deeplearningframework,evaluationindicators,andmodel performance.Moreover,wesummarizethetraditionalandmachinelearning-basedskin diseasediagnosisandtreatmentmethods.Wealsoanalyzethecurrentprogressinthisfield

and predict four directions that may become the research topic in the future. Our results show that the skin disease image recognition method based on deep learning is better than thoseofdermatologistsandothercomputer-aidedtreatmentmethodsinskindisease diagnosis, especially the multi deep learning model fusion method has the best recognition effect.

CONCLUSION:

Forty-fiverelevantpapershavebeenidentifiedtoobtaintheirconcernsaboutthe areas and skin disease types. These papers are utilized as basis to study the used data source, the preprocessing and data expansion techniques, the technical details of the models, and the performanceindicators'overallperformance.DeeplearningmodelsAlexNet,VGG, GoogleNet, and ResNet are widely used in skin disease recognition. Researchers often use the multimodel fusion technology to improve the performance of models. In future work, we plan to apply the concepts and best practices of deep learning described in this survey

toothermedicalfieldsthathavenotfullyutilizedthistechnology.Thissurveyaimsto encourage many researchers to conduct deep learning experiments and apply the model of deep learning in the field of computer vision involving medicine, thereby achieving smart and convenient development for the medical industry.

PAPER2: ProgressiveTransferLearningandAdversarialDomainAdaptationforCross-Domain Skin Disease Classification

PUBLICATIONYEAR:MAY2020

AUTHORNAME: YanyangGu,ZongyuanGe,Member,IEEE,C.PaulBonnington,Senior Member, IEEE, and Jun Zhou, Senior Member, IEEE

JOURNALNAME:IEEE JOURNALOFBIOMEDICALANDHEALTH INFOMATICS

SUMMARY:

Deeplearninghasbeenusedtoanalyzeanddiagnosevariousskindiseasesthrough medical imaging. However, recent researches show thata welltrained deep learning model maynotgeneralizewelltodatafromdifferentcohortsduetodomainshift.Simpledata fusion techniques such as combining disease samples from different data sources are not effectivetosolvethisproblem.Inthispaper,wepresenttwomethodsforanoveltaskof cross-domain skin disease recognition. Starting from a fully supervised deep convolutional neuralnetworkclassifierpre-trainedonImageNet,weexploreatwo-stepprogressive transferlearningtechniquebyfine-tuningthenetworkontwoskindiseasedatasets.We then propose to adopt adversarial learning as a domain adaptation technique to perform invariantattributetranslationfromsourcetotargetdomaininordertoimprovethe recognitionperformance.Inordertoevaluatethesetwomethods,weanalyzegeneralizati on capabilityofthetrainedmodelonmelanomadetection,cancerdetectionandcross-

modality learning tasks ontwoskinimage datasets collected fromdifferent clinical settings
andcohortswithdifferentdiseasedistributions.Theexperimentsprovetheeffectiveness of our method in solving the domain shift problem.

CONCLUSION:

Inthiswork,theyquantitativelyvalidatethemodelgeneralizationfordifferentdatasets fromtwoperspectives.Oneisapplyingparameter-basedprogressivetransferlearningto sharetransferableknowledgefromtask-differentsourcedomainandtask-samebut dataset-differentintermediatedomainwithtargetdomain.Inthesecondmethod,we generalize the model for different datasets by integrating images from other datasets after translating with cycle consistent generative networks (Cycle-GAN). In this way, the model can be generalized for dataset-different domain as well as modality-different domain. Our experimentsshowtheimprovementsforbothoverallmulti-classclassificationaccuracy andbinary classificationaccuracy on both sourcedomaindatasetsand targetdomain datasets. The improvement in binary classification is especially outstanding, which, in real cases, is more expected as the missing rate of melanoma shall be lowered. In the future, to furtherimprovethe classificationperformance,an algorithm can be developedthatmay contain discriminantfeatures fromboth thetraining sets. Thefusion could bedeveloped by constructing a hybrid training parameter set from the twotraining setparameterswhich
wereextractedonindividualdatasets.Althoughthisworkappliesdomainadaptationto skindiseaseimagingdatasetaugmentation,webelievethisschememayinspiremore studiesonapplicationsthatlacktrainingdata,especiallyingeneralmedicalimaging applications.

PAPER 3: Self-PacedBalanceLearningforClinicalSkinDiseaseRecognition

PUBLICATIONYEAR: August 2020

AUTHORNAME: JufengYang,XiaopingWu,JieLiang,XiaoxiaoSun,Ming-MingCheng, Paul L. Rosin, and Liang Wang

JOURNAL NAME: IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNINGSYSTEMS

SUMMARY:

Class imbalance is a challenging problem in many classification tasks. It induces biased classificationresultsforminorityclassesthatcontainlesstrainingsamplesthanothers. Mostexistingapproachesaimtoremedytheimbalancednumberofinstancesamong categoriesbyresamplingthemajorityandminorityclassesaccordingly.However,the imbalancedlevelofdifficultyofrecognizingdifferentcategoriesisalsocrucial,especiallyf or distinguishing samples with many classes. For example, in the task of clinical skin disease recognition,severalrarediseaseshaveasmallnumberoftrainingsamples,buttheyareea sy to diagnose because of their distinct visual properties. On the other hand, some common skin diseases, e.g., eczema, are hard to recognize due to the lack of special symptoms. To addressthisproblem,weproposeaself-pacedbalancelearning(SPBL)algorithminthis paper. Specifically, we introduce a comprehensive metric termed the complexity of image category that is a combination of both sample number and recognition difficulty. First, the complexityisinitializedusingthemodelofthefirstpace,wherethepaceindicatesone iteration in the self-paced learning paradigm. We then assign each class a penalty weight thatislargerformorecomplexcategoriesandsmallerforeasierones,afterwhichthe curriculum is reconstructed by rearranging the training samples. Consequently, the model caniterativelylearndiscriminativerepresentationsviabalancingthecomplexityineach pace.ExperimentalresultsontheSD-198andSD-260benchmarkdatasetsdemonstrate that the proposed SPBL algorithm performs favorably against the state-of-the-art

methods. We also demonstrate the effectiveness of the SPBL algorithm's generalization capacity on various tasks, such as indoor scene image recognition and object classification.

CONCLUSION:

Inthispaper,theyaddresstheclassimbalanceissueandproposeanovelSPBL algorithmthatistrainedusingsamplesfromeasytohard.Theyalsoproposeanovelinsight thatinreal-worldapplications,theclassimbalanceproblemisnotonlyduetothe imbalanceddistributionofclasssizesbutalsotheimbalancedrecognitiondifficulty. Inspired by that, we propose both the PWU and CR strategies that ensure that the model learnsacomprehensivelybalancedrepresentationineachSPLprocedure.Theyconduct experiments on two imbalanced data sets about clinical skin disease recognition tasks and severalotherimbalancedproblems.Theresultsindicatethatbothcomponentsofthe proposed algorithm are effective anddemonstrate the advantageof the SPBL against the state-of-the-art methods.

PAPER 4: A Visually Interpretable Deep Learning Framework for Histopathological Image- Based Skin Cancer Diagnosis

PUBLICATIONYEAR:MAY2021

AUTHORNAME: Shancheng Jiang, Huichuan Li, andZhiJin, Member, IEEE

JOURNALNAME:IEEEJOURNALOFBIOMEDICALANDHEALTHINFORMATICS

SUMMARY:

Owingtothehighincidencerateandthesevereimpactofskincancer,theprecise diagnosis of malignant skin tumors is a significant goal, especially considering treatment is normally effective if the tumor is detected early. Limited

publishedhistopathological image sets and the lack of an intuitive correspondence between the features of lesion areas and a certaintypeofskincancerposeachallengetotheestablishmentofhigh?qualityand interpretablecomputer-aideddiagnostic(CAD)systems.Tosolvethisproblem,alight-weight attention mechanismbased deep learning framework, namely, DRANet, is proposed todifferentiate11typesofskindiseasesbasedonarealhistopathologicalimageset collected by us during the last 10 years. The CAD system can output not only the name of a certain diseasebutalso avisualized diagnostic reportshowing possible areas related to the disease. The experimental results demonstrate that the DRANet obtains significantly better performance than baseline models (i.e., InceptionV3, ResNet50, VGG16, and VGG19) with comparableparametersizeandcompetitiveaccuracywithfewermodelparameters. VisualizedresultsproducedbythehiddenlayersoftheDRANetactuallyhighlightpartoft he

class?specificregionsofdiagnosticpointsandarevaluablefordecisionmakinginthe diagnosis of skin diseases.

CONCLUSION:

Inthisstudy,theyproposeanoveldeeplearning-basedCADsystemforskincancer diagnosis.Theoutputofthissystemisnotonlythelabelofacertaindiseasebutalsoa visualizeddiagnosticreportshowingpossibleareasrelatedtothedisease,whichcanbe consideredasdiagnosticpoints.Thevisualizeddiagnosticreportisgeneratedbythe attentionmechanismembeddedattheupperpartofourdeepstructure,andthewhole frameworkistrainedwithonly classificationlabelsinanend-to-endway.TheSpatial Attention-Orientedmaskbranchhelpstheтrunkbranchtocaptureregionsofinterest relatedtothetargetlabel,andtheChannelAttention-Orientedmaskbranchisabletomodel interdependenciesbetweenthechannelsoffeaturemapsfromthetrunkbranchina

computationallyefficientwaywhileenhancingtherepresentationalpowerofthetrunk branchthroughouteverySEAmodule.Tofurtherenhancethemodeladaptabilityfor differentsettings ofthetraining batchsizeand theconfiguration of hardware,weintroduce FRNlayersforeliminatingthedependencebetweensamplesorchannels ofthesame sample.Inthecasestudy,allcomparativeexperimentsareconductedwithareal histopathologicalimagesetthatwascollectedandmaintainedbyusduringthelast10years. Resultsvalidate the positive effectsofthe combinationoftwomask branchesin the SEA moduleandtherobustnessoftheFRNlayersonasmallscaletrainingset.OurDRANet achievessignificantlybetterperformancethanbaselinewell-knownimageclassification models with comparable parameter size and competitive accuracy with fewer parameters andsmallercomputationalcomplexity.Withlightmodelsizeandrelativelycheap computational costs, our DRANet can be deployed on common PCs or even a tiny mobile chip that might be embedded in medical image equipment. Visualized results generated by theCAMmodule actuallyhighlightpartof theclass-specificregionsofthediagnosticpoints, and they are valuable for decision making in the diagnosis of skin diseases.

PAPER5: Necrolyticmigratoryerythemaisanimportantvisualcutaneousclueof glucagonoma **PUBLICATION YEAR:** August 2022

AUTHORNAME: WeiLi1,6,XueYang1,6,YuanDeng2,YinaJiang2,GuipingXu3,EnxiaoLi4 , YinyingWu4, JuanRen5 , ZhenhuaMa1 , ShunbinDong1, LiangHan1, QingyongMa1 , ZhengWu1 & ZhengWang1

JOURNALNAME:www.nature.com/scientificreports

SUMMARY:

Erythemais anextremelyrareandslow-growingfunctionalpancreatic neuroendocrine tumorarisingfromisletalphacellsinthetailofthepancreas.Itusuallypresentswith

glucagonomasyndromeassociatedwithcharacteristicclinicalsymptoms,including necrolyticmigratoryerythema(NME),diabetesmellitus(DM),stomatitis,anemia,deepvein thrombosis (DVT), weight loss, diarrhea and other symptoms. With the exception of NME,otherclinicalmanifestationsarenonspecific,whichaccountsforthedelayin diagnosisinmostcasesandalsoforthefactthatatleast50%ofcasesalreadyhavemetastatic diseaseatthetimeofdiagnosis.NMEisobservedinapproximately70–90%ofpatients diagnosedwithglucagonoma.Thisrashisusuallywidespread,andthemajorsitesof involvement are the perioral region, trunk, extremities and perineum. The distinguishing featureofNMEisannularerythematousplaqueswithcentralbullous,ulcerativelesions surroundedbybrownpigment,whichareusuallypruriticandpainful.Thehistological featuresofthisskinlesionincludeparakeratosis,hyperkeratosis,spongiosisofthe epidermiswithnecrolysis,lossofthegranularlayer,vacuolizationofkeratinocytes,and perivascularandinterstitialinflammation.Thispapersummarizestheclinical characteristics of seven typical patients with glucagonoma followed at our hospital during thepast10years.Ourcumulativeexperiences(includingdiagnosisandtreatment)mayhelp clinicianstobetterrecognize,diagnoseandtreatglucagonoma.Thisstudywasapproved by the Ethics Committee of the First Affiliated Hospital of Xi'an Jiao tong University and the study was conducted in accordance with the approved guidelines. Informed consent was obtainedfromallsubjectsand/ortheirlegalguardian(s).Wereviewedthedatabaseand collected sevencases of glucagonomainthepast10years. Patients with clinical presentationsofskinmanifestation(theskinrashischaracterizedbyanintense erythematouslesion,whichshowssuperficialepidermalnecrosisandspreadsina centrifugal pattern), glucagonoma syndrome, elevated plasma glucagon,and a

pathological diagnosis of pancreatic islet cell tumor wereincluded inthis cohort. The recordsof theincludedpatientswerereviewed.TumordiameterswereobtainedfromCTscan measurements. Follow?up data, including patients' follow-up status, symptoms (skin rash), recoveryandadministrationofothertherapies,wereacquiredfromhospitalmedical records or by phone interviews with the patients, relatives, or general practitioners.

CONCLUSION:

Surgicalremovalisconsideredtobetheonlydefinitiveandcurativetreatmentfor pancreatic glucagonoma and NME7 . Optional operations included simple enucleation (< 2 cm) with peripancreatic lymph dissection, pancreaticoduodenectomy with lymphdissection,distalpancreatectomywithperipancreaticlymphdissectionand splenectomy.However,morethanhalfofallglucagonomaspresentwithmetastaticdiseas e, mostcommonlylivermetastasis.Ithasbeenreportedthatsynchronousresectionof pancreatic neuroendocrine tumors and liver metastasis (more than 30% of the liver tissue retained) providesamorefavorableoutcome.Livertransplantationmaybeconsidered asa potentialtherapeuticapproachforunresectablehepaticmetastasesarisingfrompancrea tic glucagonoma20.TACEmightalsobeasafetherapeuticapproachforlivermetastasisarisi ng from NETs because of the highly vascular and blood supply that primarily derives from the hepatic artery21. In addition, RFAis usually performed in combination with surgery,which hascertainadvantagesinremovingisolatedmetastases22.Medicaltherapyfor glucagonoma, including chemotherapeutics, somatostatin analogous, PRRT and molecular targeted drugs, are also effective in controlling clinical symptoms and

tumour growth7,16. In conclusion, erythema is a rare type of functional NET. Since NME might be the only clue for the early detection of this tumour, it is very important to correctly diagnose NME in a timely mannerCurrently, surgicalintervention is theonly definitive for disease.Medicaltherapyiseffectiveforsymptomcontrolandmetastaticdisease management.

c. **ProblemStatementDefinition**

➤Now a day's people are suffering from skin diseases, More than 125 million people suffering from Psoriasis also skin cancer rate is rapidly increasing over the last few decades especially Melanoma is most diversifying skin cancer.

➤Ifskindiseasesarenottreatedatanearlierstage,thenitmayleadtocomplications inthebodyincludingspreadingoftheinfectionfromoneindividualtotheother.

➤The skin diseases can be prevented by investigating the infected region at an early stage.

➤The characteristic of the skin imagesis diversifiedso that it is a challenging job to devise an efficient and robust algorithm for automatic detection of skin disease and its severity.

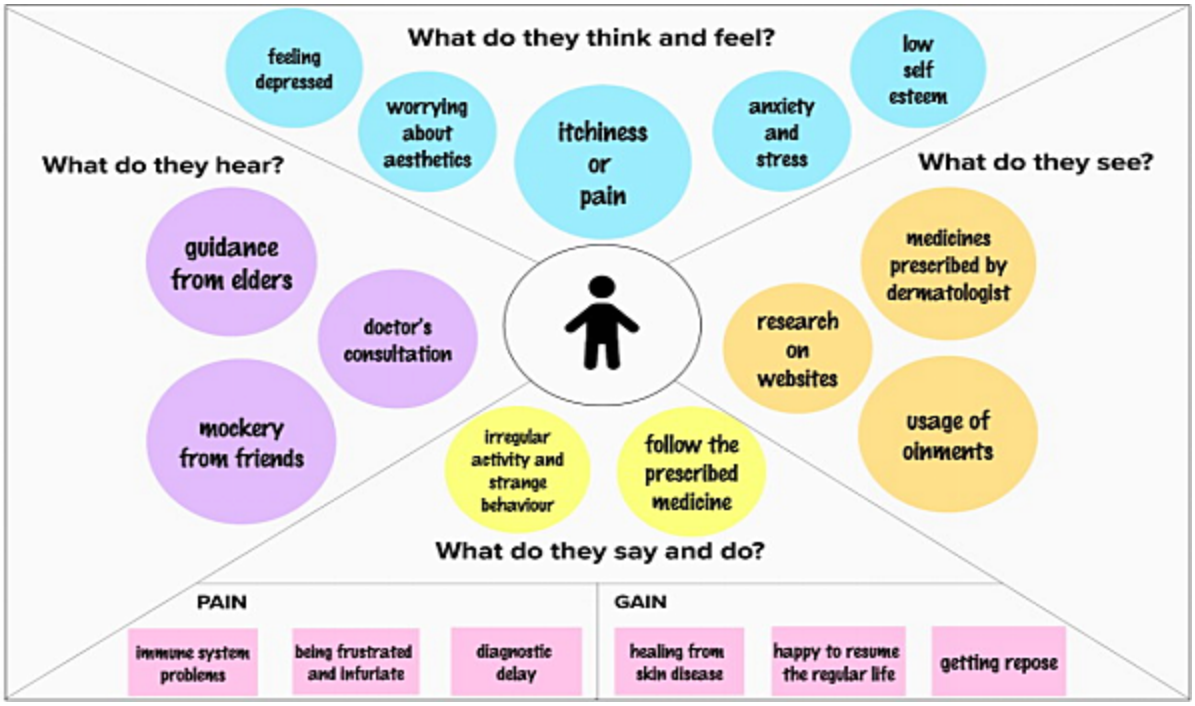
➤Skintoneandskincolourplayanimportantroleinskindiseasedetection.Colourand coarsenessofskinarevisuallydifferent.

➤Automaticprocessingofsuchimagesforskinanalysisrequiresquantitative discriminator to differentiate the diseases.

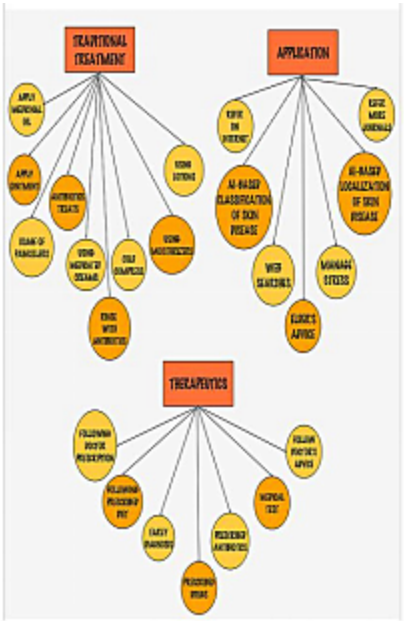
IDEATION&PROPOSEDSOL UTION

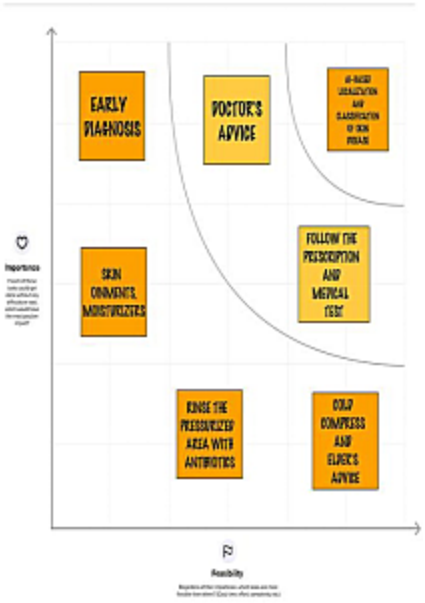
a.

Empathy Map Canvas



Ideation&Brainstorming





.ProposedSolution

➤To overcome the above problem we are building a model which isused for the prevention and early detection of skin cancer, psoriasis.

➤Basically, skin disease diagnosis depends on the different characteristics like colour, shape, texture etc. Here the person can capture the images of skin and then the image will be sent the trained model.

➤Themodelanalysestheimageanddetectwhetherthepersonishavingskindiseaseor not.

Novelty:

➤Thenovelty proposedin this approachiswehavecollected thedataseton ourown.

➤Wehavealso annotatedtheimages by ourselves.

SocialImpact:

➤Themodelwhichwillbebuiltbyusisveryusefulfortheuserstofindthetypeofdisease quickly and get the correct the medicine as soon as possible.

➤Weensuretotheusersthatourmodeldiagnosesthediseaseswell.

BusinessModel:

➤HealthCareSector(Hospitals).

➤Can generaterevenuethroughdirectcustomers.

➤Cancollaboratewithhealthcaresectorandgeneraterevenuefromtheircustomers.

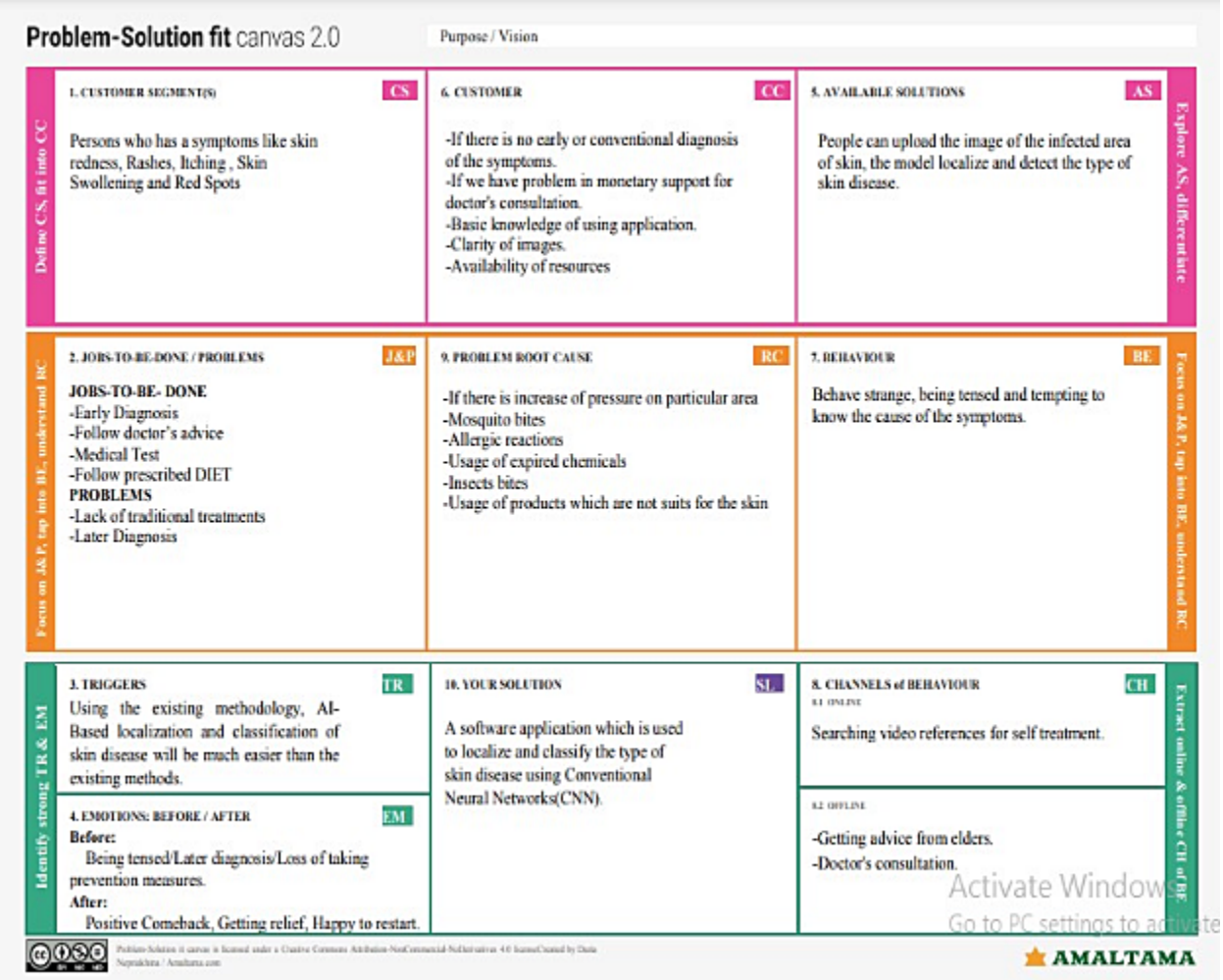
➤Contributing thecorporatesocial responsibility byprovidingbetter solutionsto the healthcare and to patients.

Scalabilityofthesolution:

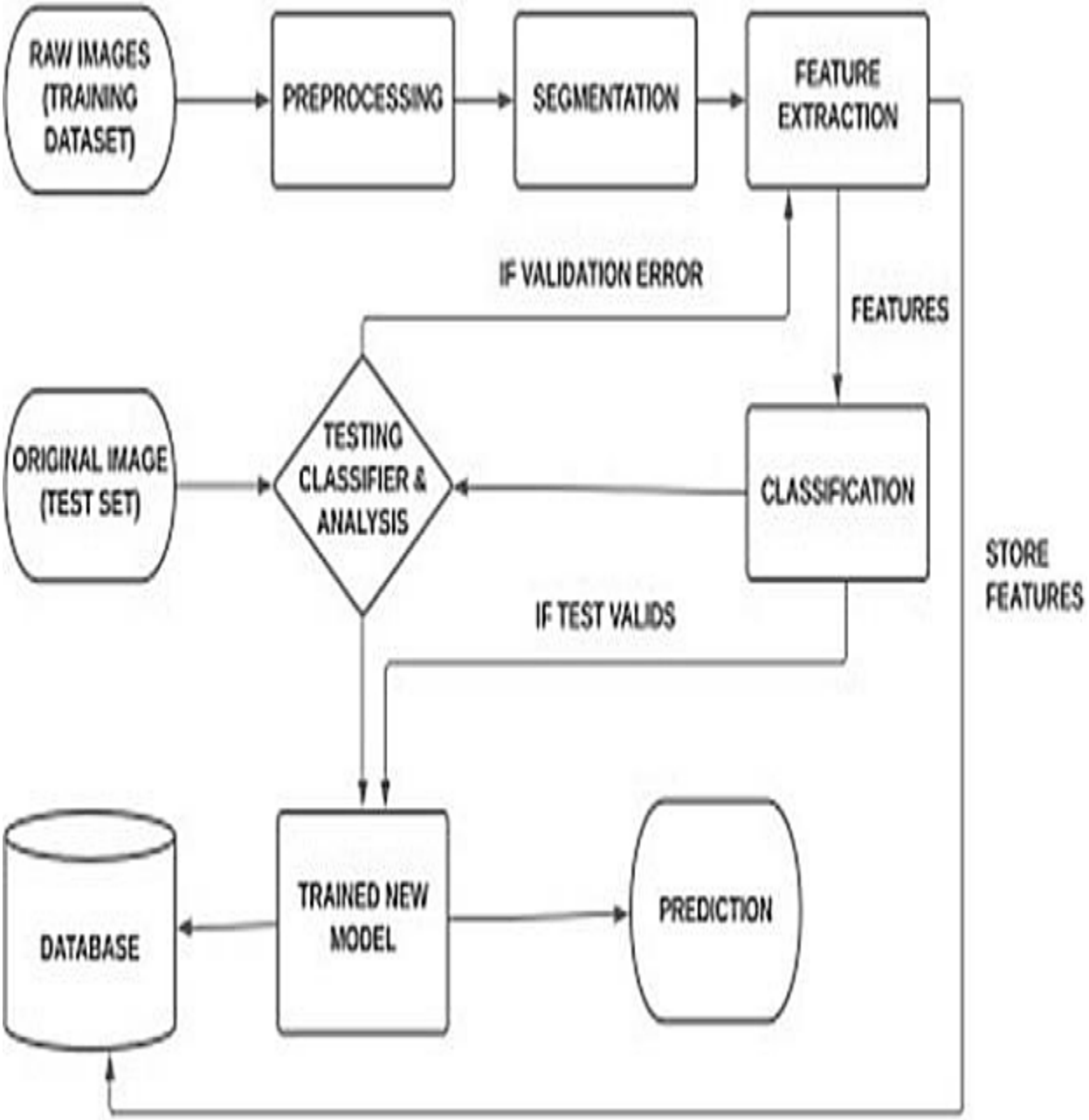
➤Todetecthowmuchareaisaffectedbytheskindiseasewithaccuracy.

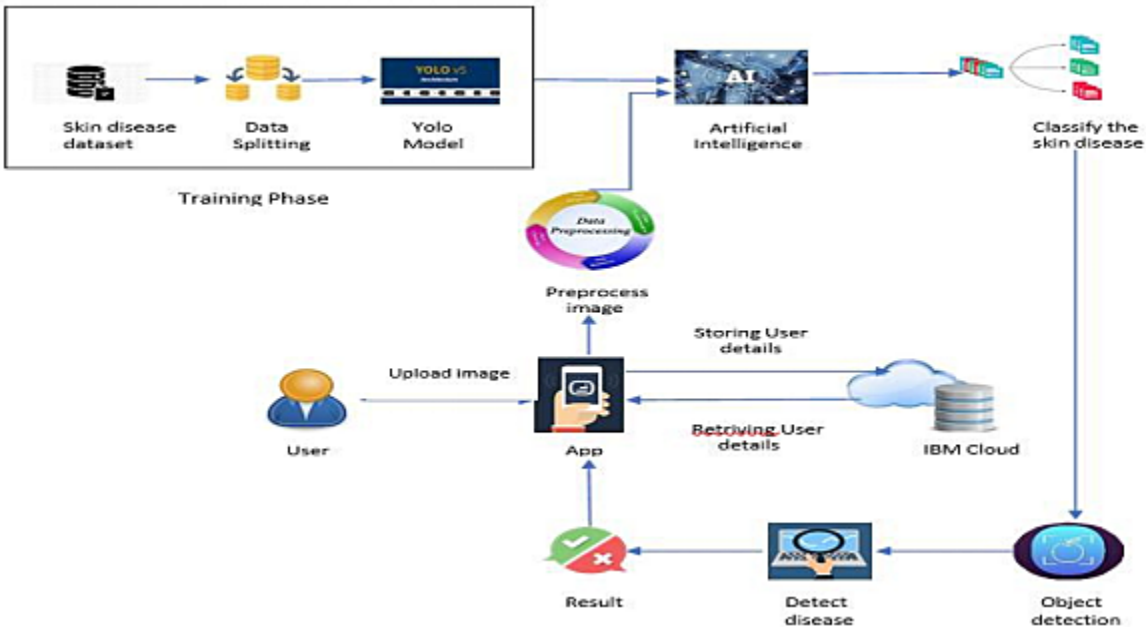
➤Classification ofSkindiseasewill helpdoctorstodiagnosethediseaseeffectively.

3.4.ProblemSolutionFit



a. **Dataflowdiagrams**


RAW IMAGES (TRAINING DATASET)
PREPROCESSING
SEGMENTATION
FEATURE EXTRACTION
IF VALIDATION ERROR
FEATURES
ORIGINAL IMAGE (TEST SET)
TESTING CLASSIFIER & ANALYSIS
CLASSIFICATION
STORE FEATURES
IF TEST VALIDS
DATABASE
TRAINED NEW MODEL
PREDICTION



TechnologyArchitecture:

In the past, the skills required to make an accurate dermatological diagnosis have required exposure tothousands of patients over many years. However, in recent years, artificial intelligence (AI) has made enormous advances, particularly in theareaofimage classification.

This has led computer scientists to apply these techniques to develop algorithms thatareabletorecognizeskinlesions,particularlymelanoma.Since2017,therehavebeen numerousstudiesassessingtheaccuracyofalgorithms,withsomereportingthatthe accuracy matches or surpasses that of a dermatologist.While the principles underlying thesemethodsarerelativelystraightforward,itcanbechallengingforthepractising dermatologist to make sense of a plethora of unfamiliar terms in this domain.

Here we explain the concepts of AI, machine learning,neural networks and deep learning,andexplore theprinciplesofhowthesetasksareaccomplished.Wecritically evaluatethestudiesthathaveassessedtheefficacyofthesemethodsanddiscuss limitations and potential ethical issues. The burden of skin cancer is growing within the Westernworld,withmajorimplicationsforbothpopulationskinhealthandtheprovision of dermatology services.AI has the potential to assist in the diagnosis of skin lesions and mayhaveparticularvalueattheinterfacebetweenprimaryandsecondarycare.The emerging technology represents an exciting opportunityfordermatologists, who are the individuals bestinformed to explore the utility of this powerfulnovel diagnostic tool, and facilitate itssafe and ethical implementation within several healthcare system.

FIGURE1:

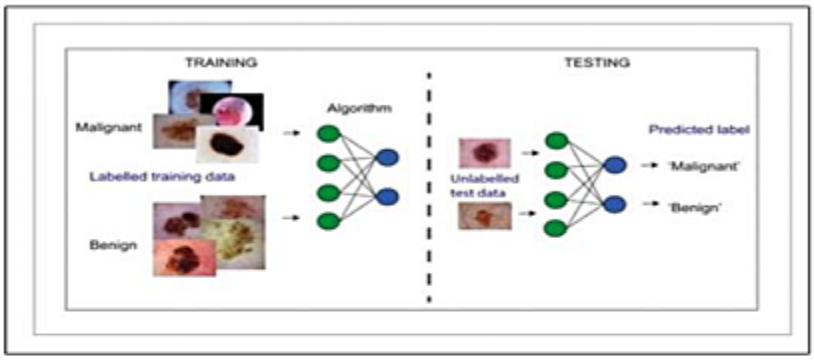


FIGURE2:

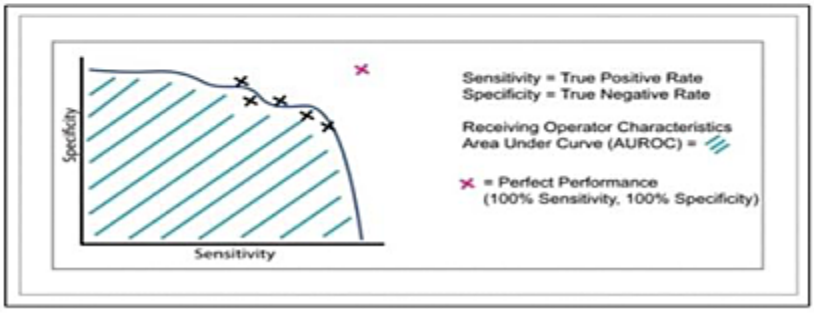


FIGURE3:

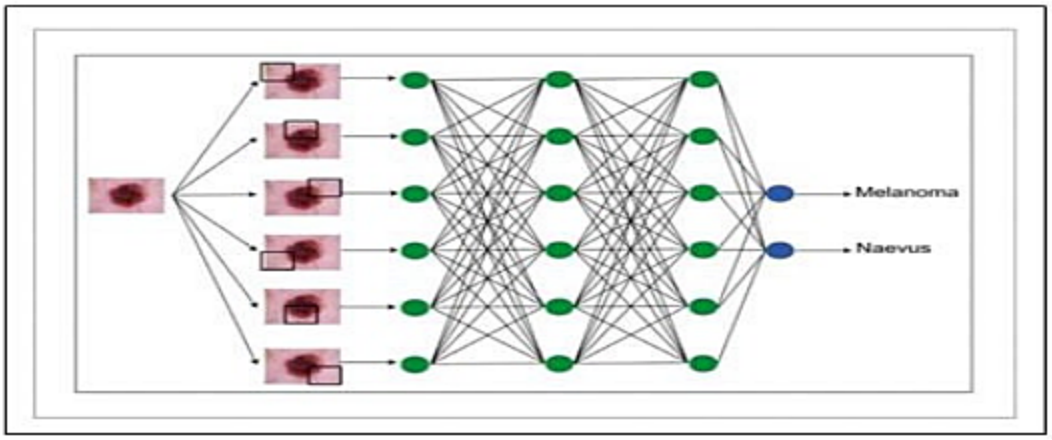


FIGURE4:

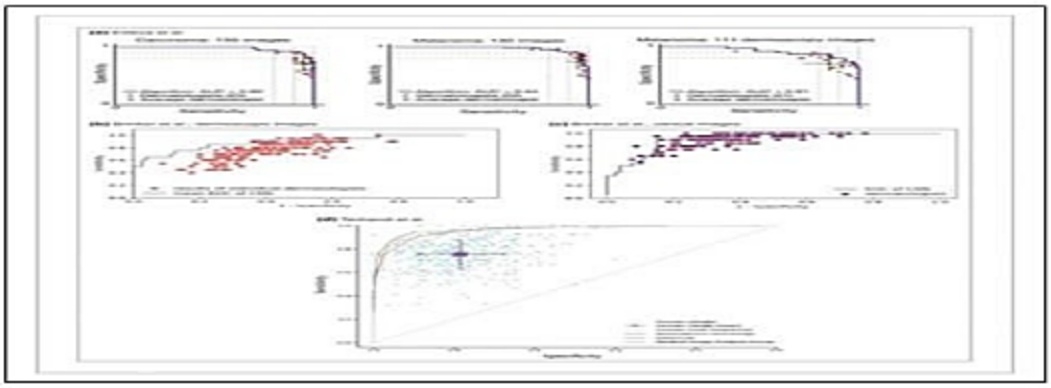
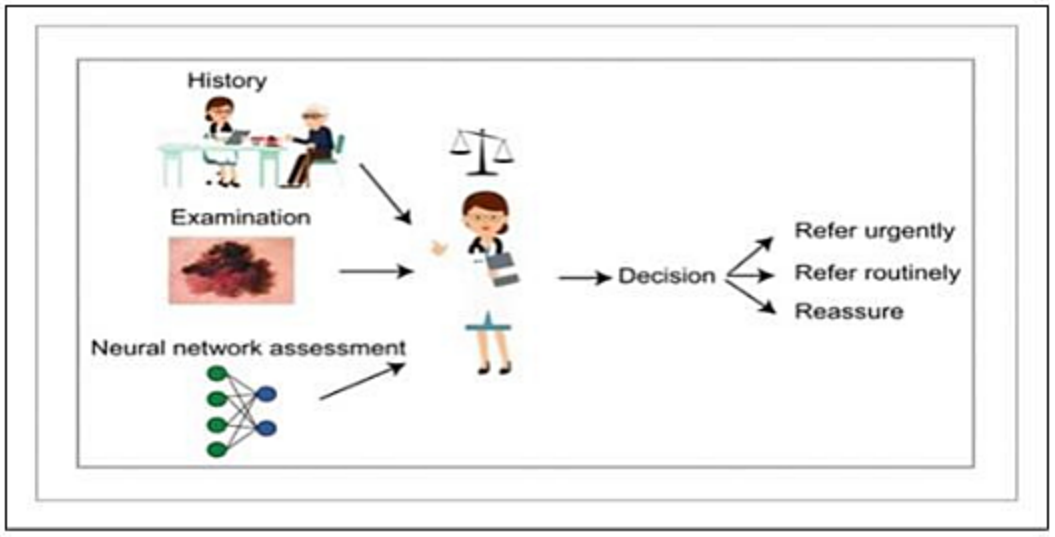


FIGURE5:



PROJECTPLANNING&SCHEDULING

SprintPlanning&Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story /Task	Story Points	Priority	Team Members
Sprint-1	Create Dataset	USN-1	Create the datasetwith50 imagesperskin disease	3	High	Jayapriya.V, Nishanthi.A, Sathya.M Subashini. E,
Sprint-2	Annotate images	USN-2	Annotateimages using MicrosoftVOTT intofourphases	3	High	Jayapriya.V, Nishanthi.A, Sathya.M Subashini. E,
Sprint-3	Training Yolo and Build python code	USN-3	DownloadandConvert pre-trained weights.Train Yolov3 detectorand buildthe sourcecode	3	High	Jayapriya.V, Nishanthi.A,
Sprint-4	CloudantDB	USN-4	Create cloud account, create serviceinstance, launch cloudant DB and create the database	3	High	Sathya.M Subashini. E,

Sprint-4	Registration	USN-5	Asauser,Ican registerforthe application by entering my email, password, and confirmingmy password.	3	High	Nishanthi .A
Sprint-4		USN-6	Asauser,Iwill receive confirmation emailonceI haveregisteredfor theapplication	2	Medium	Sathya.M
Sprint-4		USN-7	Asauser,Ican registerforthe application throughmobilenumber	3	High	Subashini.E
Sprint-4		USN-8	Asauser.Iwill Receive confirmation SMS	3	High	Jayapriya.V
Sprint-4	Login	USN-9	Asauser,Ican logintothe applicationby entering login credentials	3	High	Nishanthi .A Sathya.M

Sprint-4	Dashboard	USN-10	Asauser,Ican uploadmy imagesandget mydetailsof skin diseases	3	High	Sathya.M Subashini .E
Sprint-4	Logout	USN-11	Asauser,Ican logout successfully	2	Medium	Jayapriya.V Nishanthi.A

a. **SprintDeliverySchedule**

TITLE	DESCRIPTION	DATE
Literature Survey & Information Gathering	Literature survey on the selected project & gathering informationbyreferringthe, technical papers,research publications etc.	27SEPTEMBER2022
PrepareEmpathyMap	Prepare Empathy Map Canvas to capture the user Pains&Gains,Preparelistof problemstatements	10SEPTEMBER2022
Ideation	List the by organizing the brainstormingsession andprioritizethetop3 ideasbasedonthe feasibility & importance.	10OCTOBER2022
ProposedSolution	Prepare the proposed solution document, which includes the novelty, feasibilityofidea,business model, social impact, scalability of solution, etc.	11OCTOBER2022
ProblemSolutionFit	Prepareproblem-solution fit document.	10OCTOBER2022

a. **Feature1**

CODING&SOLUTIONING

In this project we are implementing a feature which is creating a website for AI-based localization and classification of skin diseasewith erythema. By using this feature wecan login into the website and upload photos.

Codeforloginpage:

```
<!DOCTYPEhtml>
<html>
<head>
<metacharset="UTF-8">
<metaname="viewport"content="width=device-width,initial-scale=1">
<title>SKINCLASSIFICATION</title>
<linkhref='https://fonts.googleapis.com/css?family=Pacifico'rel='stylesheet'type='text/css'>
<linkhref='https://fonts.googleapis.com/css?family=Arimo'rel='stylesheet'type='text/css'>
<linkhref='https://fonts.googleapis.com/css?family=Hind:300'rel='stylesheet'type='text/css'>
<linkhref='https://fonts.googleapis.com/css?family=Open+Sans+Condensed:300'rel='styl
esheet' type='text/css'>
<linkrel="stylesheet"href="{{url_for('static',filename='css/style.css')}}">
<linkhref='https://fonts.googleapis.com/css?family=Merriweather'rel='stylesheet'>
<linkhref='https://fonts.googleapis.com/css?family=JosefinSans'rel='stylesheet'>
<linkhref='https://fonts.googleapis.com/css?family=Montserrat'rel='stylesheet'>
<style>
.header{
                        top:0;
                        margin:0px;
                        left:0px;
                        right:0px;
                        position:fixed;
                        background-color:#566573; color: #7FE228;
                        box-shadow:0px8px4pxgrey;
                        overflow: hidden;
                        padding-left:20px;
```

```
                font-family:'JosefinSans';
                font-size:2vw;
                width:100%;
                height:8%;
                text-align:center;
            }
    </div>
        </form>
</div>
</body>
</html>
```

a. **Feature2**

In this project we are using CNN.CNN is usedfor effective classification.CNN is computionallyefficient and we can classify the images without much human intervention.

CodeforbuildingCNN:

```
#Part 1-BuildingtheCNN
#importingtheKeraslibrariesan
dpackages from keras.models
import Sequential
fromkeras.layersimportCon
volution2D
fromkeras.layersimportMa
xPooling2D from
keras.layers import Flatten
fromkeras.layersimportDen
se,Dropout from keras
import optimizers

# Initialing
the CNN
classifier=S
equential()
```

```
#Step 1-Convolution Layer
classifier.add(Convolution2D(32,3,3,input_shape=(64,64,3),activation='relu'))

#step 2 - Pooling
classifier.add(MaxPooling2D(pool
_size=(2,2)))

# Adding second convolution layer
classifier.add(Convolution2D(32,3,3,activation='re
lu'))
classifier.add(MaxPooling2D(pool_size=(2,2)))

#Adding 3rd Concolution Layer
classifier.add(Convolution2D(64,3,3,activation='re
lu'))
classifier.add(MaxPooling2D(pool_size=(2,2)))



#Step 3 - Flattening
classifier.add(Flatten())

#Step 4 - Full Connection
classifier.add(Dense(256, activation
= 'relu'))
classifier.add(Dropout(0.5))
classifier.add(Dense(10,activation='
softmax'))

#CompilingTheCNN
classifier.compile(
        optimizer ='adam',
        loss='categorical_crossent
        ropy', metrics =
        ['accuracy'])

#Part2FitttingtheCNNtotheimage
fromkeras.preprocessing.imageimportImageDataGenerator
```

```
train_datagen = ImageDataGenerator(
        rescale=1./255,
        shear_range=0.2,
        zoom_range=0.2,
        horizontal_flip=Tr
        ue)
test_datagen=ImageDataGenerator(rescal

e=1./255) training_set =

train_datagen.flow_from_directory(

        'Data/train',
        target_size=
        (64, 64),
        batch_size=
        32,
        class_mode=
        'categorical')

test_set=test_datagen.flow_from_director
        y( 'Data/test',
        target_size=(64,
        64),
        batch_size=32,
        class_mode='cate
        gorical')

model=classifier.fit_ge
        nerator(
        training_set,
        steps_per_epoch=
        100, epochs=100,

        validation_data=test_set, validation_steps = 6500
      )
```

```
#Savingt
hemodel
import
h5py
classifier.save('Trained_Model.h5')

print(model.histor
y.keys())
importmatplotlib.
pyplotasplt

#summarizehistoryfor
accuracy
plt.plot(model.history['
acc'])
plt.plot(model.history['
val_acc'])
plt.title('model
accuracy')
plt.ylabel('accuracy')
plt.xlabel('epoch')
plt.legend(['train','test'],loc='upperle
ft') plt.show()

# summarize
history for loss
plt.plot(model.histor
y['loss'])
plt.plot(model.histor
y['val_loss'])
plt.title('model
loss')
plt.ylabel('loss')plt.x
label('epoch')
plt.legend(['train','test'],loc='upperle
ft') plt.show()
```

SourceCode

```
pip3 installtensorflowtensorflow_hubmatplotlibseabornnumpypandassklearnimblearn

  import tensorflow as tf
  importtensorflow_hubas hub
  importmatplotlib.pyplot as pltimport numpy as np import pandas as pd import
  seaborn as
  sns
  fromtensorflow.keras.utilsimportget_file
fromsklearn.metricsimportroc_curve,auc,confusion
_matrix from imblearn.metrics import
sensitivity_score, specificity_score

  import os import glob importzipfile import


  random


  #togetconsistentresultsaftermultiple runstf.random.set_seed(7) np.random.seed(7)
  random.seed(7)


  #0forbenign,1formalignant class_names = ["benign", "malignant"]


  PreparingtheDataset

  defdownload_and_extract_dataset():
   #datasetfromhttps://github.com/udacity/der
   matologist- ai# 5.3GB
   train_url="https://s3-us-west-1.amazonaws.com/udacity-dlnfd/datasets/skin-
   cancer/train.zip"# 824.5MB
   valid_url="https://s3-us-west-1.amazonaws.com/udacity-dlnfd/datasets/skin-
   cancer/valid.zip"# 5.1GB
   test_url = "https://s3-us-west-1.amazonaws.com/udacity-
   dlnfd/datasets/skin- cancer/test.zip"for i, download_link in
   enumerate([valid_url, train_url, test_url]): temp_file =
   f"temp{i}.zip"
    data_dir=get_file(origin=download_link,fname=os.path.join(os.getcwd(),temp_file))
    print("Extracting", download_link)
    withzipfile.ZipFile(data_dir,"r")asz: z.extractall("data")
```

```
    # remove the temp file os.remove(temp_file
    )

  #commentthebelowlineifyoualreadydownloadedthe
  datasetdownload_and_extract_dataset()
  #preparingdata
  #generateCSVmetadatafiletoreadimgpathsandlab
  elsfrom itdef generate_csv(folder, label2int):
    folder_name = os.path.basename(folder)labels = list(label2int)
    #generateCSVfile
    df=pd.DataFrame(columns=["fil
    epath", "label"])i = 0
    forlabelinlabels:
      print("Reading",
      os.path.join(folder, label, "*"))
      forfilepathinglob.glob(os.path.join
      (folder,label,
        "*")):df.loc[i]=[filepath,labe
        l2int[label]] i += 1
    output_file=f"{folder_name}.csv"

    print("Saving",output_file) df.to_csv(output_file)
```

#generateCSVfilesforalldataportions,labelingnevusandseborrheickeratosis

```
  #as0(benign),andmelanomaas1(malignant)
  #youshouldreplace"data"pathtoyourextracteddatasetpath
  # don't replace if you used download_and_extract_dataset() function
  generate_csv("data/train",{"nevus":0,"seborrheic_keratosis":0,"melanoma":1})
  generate_csv("data/valid",{"nevus":0,"seborrheic_keratosis":0,"melanoma":1})
  generate_csv("data/test",{"nevus":0,"seborrheic_keratosis
  ":0, "melanoma": 1})# loading data
  train_metadata_filena
  me=
  "train.csv"valid_meta
  data_filename
  ="valid.csv"#loadCSVfile
  sas DataFrames
```

```
df_train = pd.read_csv(train_metadata_filename) df_valid = pd.read_csv(valid_metadata_filename)
n_training_samples = len(df_train) n_validation_samples = len(df_valid)
print("Numberoftraining
samples:",n_training_samples)
print("Number of validation samples:",
n_validation_samples)
train_ds =
tf.data.Dataset.from_tensor_slices((df_train["filepath"],
df_train["label"])) valid_ds =
tf.data.Dataset.from_tensor_slices((df_valid["filepath"],df
_valid["label"])
```

Output:

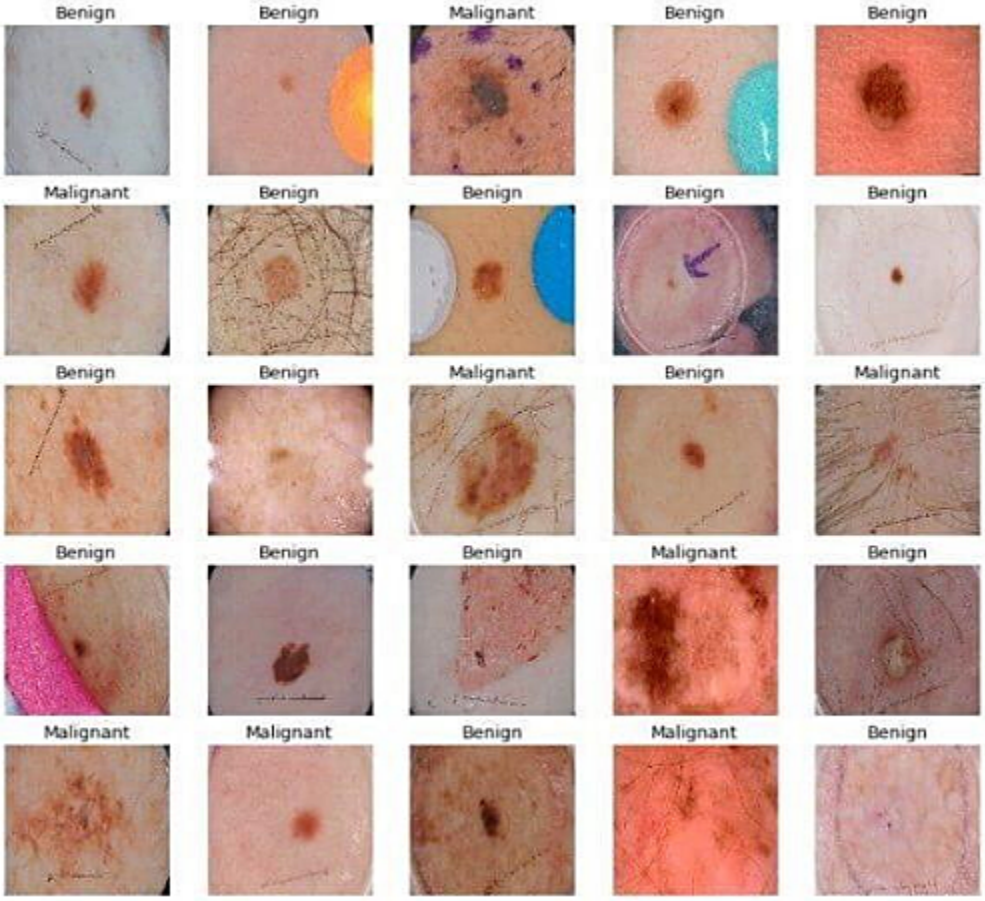
```
Number of training samples: 2000
Numberofvalidationsamples:150
```

#preprocessdata

```
def decode_img(img):

#convertthecompressedstringtoa3
 Duint8 tensorimg =
 tf.image.decode_jpeg(img,
 channels=3)
 #Use`convert_image_dtype`toconverttofloatsinthe[0,1]
 range.img=tf.image.convert_image_dtype(img,tf.float32) # resize the image to the desired
 size.return tf.image.resize(img, [299, 299])

defprocess_path(filepath,label):
 #loadtherawdatafromthefileasa stringimg = tf.io.read_file(filepath) img =
 decode_img(img) return img, label
valid_ds=
valid_ds.map(process_path)train_ds=
train_ds.map(process_path)# test_ds = test_ds
forimage,labelintrain_ds.take(1):
   print("Imageshape:",image.shap
   e) print("Label:",label.numpy())
Imageshape:
parameters batch_size=64 optimizer = "rmsprop"
```

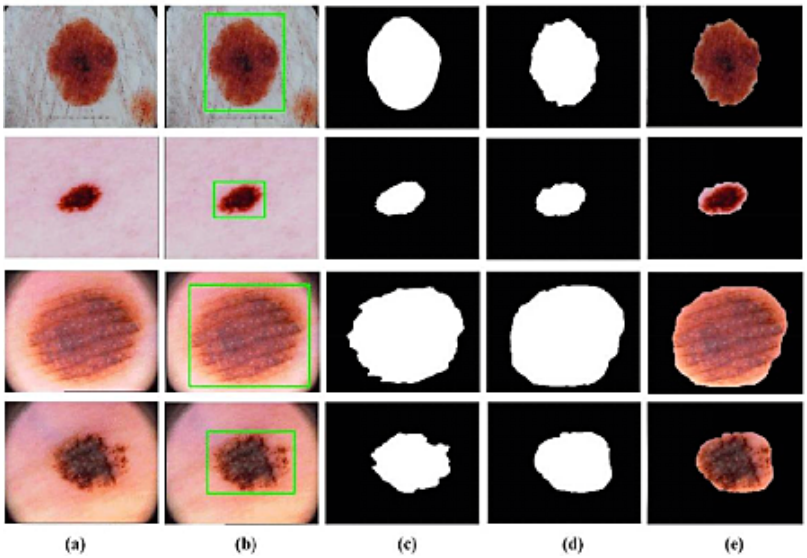
```
defprepare_for_training(ds,cache=True,b
atch_size=64,
shuffle_buffer_size=1000):if cache:
     ifisinstance(cache,str):
       ds = ds.cache(cache)else:
       ds=ds.cache()
   #shufflethedataset ds =
   ds.shuffle(buffer_size=shuffle_buffer_size)#
   Repeat forever ds=ds.repeat() # split to batches
   ds=ds.batch(batch_size)
   #`prefetch`letsthedatasetfetchbatchesinthebackgroundwhi
   lethe model# is training.
   ds =
   ds.prefetch(buffer_size=tf.data.experimental.
   AUTOTUNE) return ds

 valid_ds = prepare_for_training(valid_ds, batch_size=batch_size,
 cache="valid-cached- data") train_ds=prepare_for_training(train_ds,
 batch_size=batch_size,cache="train- cached-data") batch =
 next(iter(valid_ds))

 def show_batch(batch):
  plt.figure(figsize=(12,12
  ))forninrange(25):
     ax = plt.subplot(5,5,n+1) plt.imshow(batch[0][ n])
     plt.title(class_names[batch[1][n].num
     py()].title()) plt.axis('off')

 show_batch(batch)
```

Output:



RESULT

Thefinalresultsarebasedontheaccuracyresultsintheformofthemelanomaand the non-melanoma skin diseases classifications.



1. ADVANTAGESANDDISADVANTAGES

a. Advantages

Instant Response, improves prediction of Skin Disease, no referral needed, SavesMoneyandTime,andConfidentialAdvice.

b. Disadvantages

NetworkConnectivityandAccuracy.

2. CONCLUSION

Wehaveshownthatevenwithoutalargedatasetandhigh-qualityimages,itis possibleto achieve sufficientaccuracyrates.Inaddition,wehaveshownthatcurrent state-of-the-artCNNmodelscanoutperformmodelscreatedbypreviousresearch, throughproperdatapre-processing,self-supervisedlearning,transfer learning, and special CNNarchitecturetechniques.Furthermore,withaccuratesegmentation,wegain knowledge of the location of the disease, which is useful in the pre-processing of data used inclassification,asitallowstheCNNmodeltofocusontheareaofinterest.Lastly,unli ke previous studies, our method provides a solution to classify multiple diseases within a single image. With higher quality and a larger quantity of data, it will be viable to use state-of-the-art models to enable the use of CAD in the field of dermatology.

FUTURESCOPE

ThisimplementationoftheStructuralCo-Occurrencematricesforfeature extractionintheskindiseases classification and the pre-processing techniques are handled by using the Median filter, this filter helps toremove the salt and pepper noise in the image processing; thus, it enhances the quality of the images, and normally, the skin diseases are considered as the risk factor in all over the world. Our proposed approach provides 97%oftheclassificationoftheaccuracyresultswhileanotherexistingmodelsuchasF FT
+ SCM gives 80%, SVM + SCM gives 83%, KNN + SCM gives 85%, and SCM + CNN gives 82%. Future work is dependent on the increased support vector machine's accuracy in classifying skin illnesses, and SCM is used to manage the feature extraction technique.