The Depth Estimation and Visualization of Dermatological Lesions: A Novel Methodology

Pranav Parekh, Richard O. Oyeleke

Submitted to: Journal of Medical Internet Research
on: April 23, 2024

Table of Contents

The Depth Estimation and Visualization of Dermatological Lesions: A Novel Methodology

Pranav Parekh[1]; Richard O. Oyeleke[2]

[1]Stevens Institute of Technology Hoboken US

Corresponding Author:
Pranav Parekh

Abstract

Background: Thus far, considerable research has been focused on classifying a lesion as benign or malignant.

Objective: To propose a novel methodology for the depth estimation and visualization of skin lesions.

Methods: We start by doing the same using a CNN, followed by using Explainable AI (XAI) to localize the image features responsible for the CNN output. We apply computer graphics for depth estimation and developing the 3D structure of the lesion. Our novel method, called the red spot analysis, measures the degree of infection based on which a conical hologram is constructed. Physicians can study this hologram via Mixed Reality headsets.

Results: The neural model achieves an accuracy of 85.61%. We successfully obtained 3D representations of lesion depth using the method stated above.

Conclusions: When we map the CNN outputs (benign or malignant) to the corresponding hologram, we observe that a malignant lesion has a higher concentration of red spots (infection) in the upper and deeper portions of the skin. Clinical Trial: We do not perform RCT for this study.

(JMIR Preprints 23/04/2024:59839)
DOI: https://doi.org/10.2196/preprints.59839

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Paper

Enter information for authors (including designations, affiliations, correspondence, contributions) in the online metadata form. Do not use periods after initials, and include degree designations and affiliations for all authors. Trial registration numbers are also filled in on the metadata forms online.

The Depth Estimation and Visualization of Dermatological Lesions: A Novel Methodology

Abstract

Background: Thus far, considerable research has been focused on classifying a lesion as benign or malignant.
Objective: To propose a novel methodology for the depth estimation and visualization of skin lesions.
Methods: We start by doing the same using a CNN, followed by using Explainable AI (XAI) to localize the image features responsible for the CNN output. We apply computer graphics for depth estimation and developing the 3D structure of the lesion. Our novel method, called the red spot analysis, measures the degree of infection based on which a conical hologram is constructed. Physicians can study this hologram via Mixed Reality headsets.
Results: The neural model achieves an accuracy of 85.61%. We successfully obtained 3D representations of lesion depth using the method stated above.
Conclusions: When we map the CNN outputs (benign or malignant) to the corresponding hologram, we observe that a malignant lesion has a higher concentration of red spots (infection) in the upper and deeper portions of the skin.
Trial Registration: We do not perform RCT for this study.

Keywords: Machine Learning (ML), Computer Vision; Neural Networks; Explainable AI (XAI); Computer Graphics; Red Spot Analysis; Mixed Reality (MR)

Introduction

Skin cancer is the abnormal growth of skin cells that most often develops due to exposure to ultraviolet radiation. Based on the affected cells, the skin lesions caused by the cancer are divided into melanocytic and non-melanocytic [1]. Nonmelanoma skin cancers are divided into basal cell carcinoma, squamous cell carcinoma and Merkel cell carcinoma. Basal cell carcinoma is the most common type of skin cancer but is usually treatable. On the other hand, melanocytic skin cancers are divided into melanoma and nevus. Melanoma is a serious skin cancer that can be fatal if not detected early. Melanoma is life-threatening when it grows beyond the skin's dermis, making depth an essential factor while treating melanoma [2].

Based on how deep the cancer has penetrated the skin, melanoma can be classified into 5 stages. Stage 0 is 100% curable and occurs when the lesion is on top of the skin. Stages 1,2 and 3 are curable through surgery (or advanced surgery) and post-medication; however, as the stages increase, so does the difficulty in treating the cancer. Stage 4 is the deadliest of them all, and it occurs when the cancer has spread into lymph nodes and organs. There are low survival rates among patients 1. Therefore, the early detection of melanoma is essential. We want to be able to detect melanoma within the earlier stages [3]. Once melanoma is detected, we want to reconstruct the lesion into a 3D holographic projection to examine the skin lesion's depth. Analyzing the depth tells us the stage of cancer and the treatment required.

Our first step is to detect melanoma from a given input skin lesion using ML and XAI. This study uses ML as an umbrella term for neural models and computer vision. ML is growing rapidly, and its excellent performance has enormous potential in many fields, including healthcare. However, there is a need to explore the interpretability of ML models. They are commonly used as a black box that throws an output based on a specific input data sample. However, for fields like healthcare, where context plays a vital role, recent research has been explored to develop Explainable AI (XAI). XAI methods help explain the decisions and predictions made by the model. This helps us improve our systems and fix our hyperparameters while implementing the models [4]. In the next section, we shall review some XAI methods and use them to detect skin melanoma.

The second step is reconstructing the detected melanoma lesion as a 3D holographic projection. This uses computer graphics concepts like depth map estimation and surface reconstruction. We also developed a novel method called the red spot analysis to quantify the amount of infection with increased depth into the skin. Our final output is a conical structure of the lesion which can be visualized and interacted with as a hologram via a Mixed Reality headset. The reason for proposing the use of MR is to blend the real and virtual worlds so that we have a projection of the lesion within a real-world setting [5].

To summarize, the major contributions of the paper are as follows:

- It gives the physician a tool to estimate how much the lesion proliferated into the skin.
- The hologram is interactive, so certain areas can be zoomed in and studied in detail.
- This enables quick and accurate diagnosis of the patient.

Before moving on to the implementation methodology, we review how ML and MR have impacted the treatment of skin lesions and assisted physicians in making decisions.

Prior Work

The following section discusses the previous studies that are crucial in developing this research. We elaborate on how the two fields have performed individually for skin cancer treatment. This gives us an idea of integrating them to achieve further improvements.

ML models and XAI for melanoma detection

The following study uses ML as an umbrella term that includes computer vision and neural networks. The classification of tumors as benign and malignant has been a familiar problem solved by logistic regression [6]. Numerous studies have extended machine learning algorithms for the classification of skin lesions to detect melanoma [7].

The study by [8] used a computer-aided diagnosis system to classify the two classes of skin lesions - benign and malignant. Classification is performed by 4 ML classifiers: Support Vector Machine (SVM), Hidden Naive Bayes (HNB), Random forest (RF) and logistic regression. The paper by [9] presents a skin lesions classification system based on transfer learning and neural networks. They use the Alex-Net alongside the softmax activation function for the multi-class classification of three types of lesions. They classify the segmented color images into melanoma, nevus and seborrheic keratosis.

Performing segmentation is often used as a pre-classification procedure in several studies. [10]

performs threshold-based segmentation based on Otsu's algorithm. The shape, color and texture features are extracted from the segmentation, which are used in identifying malignant melanoma from Clark nevi. On the other hand, [11] uses supervised mechanisms like the MAP technique for segmentation and G-LoG (Gaussian - Laplacian of Gaussian) for classification. Another standard algorithm for extracting color features before applying logistic regression is the K-Mean Clustering and K-Nearest Neighbours [12].

Apart from detecting melanoma, it is critical to detect the depth of the cancer. The 3 common characteristics of melanoma moles are:

- The outer covering of the moles is ragged, asymmetrical and coarse
- Almost half of the moles present do not resemble the other half of the moles
- The newly formed moles are of different shape, color and texture from the previously existing moles.

Based on the features of the moles, we get an idea about the spreading level and severity of the disease. The Proliferative Index (PI) is the fraction of the total active nuclei present at that instance of time [13]. Its relation to the depth of the lesion is yet to be studied. The study by [14] used the sum rule fusion method and ANN to confirm whether the melanoma stage is critical. However, this method does not clarify the distinction between each stage. To find a clear difference between each stage, we need to estimate the depth of the lesion within the skin and lymph nodes.

Although ML models have improved accuracy in melanoma detection, there is a lack of transparency in how these systems obtain their results. XAI systems are used to provide explanations to clinicians, thereby solving the issue of transparency. There are two branches of XAI techniques [15]:

- Intrinsically and inherently understandable algorithms: There could be a trade-off between performance and interoperability, leading to bad results
- Retrospective post hoc algorithms: They are often rejected in the medical field due to the risk of confirmation bias along with the explanations

They developed their own multimodal XAI system that matched the XAI explanations to the clinician's judgments, aligning it well to the medical task [15]. Deep Neural Networks like Alex-Net mentioned above have primarily been seen as black-box predictors. [16] use XAI techniques like the Integrated Gradient Attribution method and SmoothGrad Noise Reduction algorithm to visualize the model's contributing features internally.

Recently, Convolutional Neural Networks (CNNs) have achieved excellent results in detecting and diagnosing melanoma [17]. Deep pre-trained convolutional models have also been used to extract features from skin lesions for necessary classification [18]. Such models consist of convolutional, pooling, and dense, fully connected layers for the required output. A paper by [19] utilizes the Global Average Pooling layer (GAP) to support the localization of objects in an image. They are used to retain the spatial structure of the feature maps, and identify discriminative regions of the image. They performed the global average pooling operation on the feature maps just before the final softmax activation layer, which helped determine the critical regions of the image. The Class Activation Map (CAM) indicates the discriminative region a CNN uses to classify the image into its corresponding class. It does so by projecting back the output layer weights onto the convolutional feature map.

The primary limitation of the CAM method is that architectural constraints bind it: only the architectures with GAP layers before the softmax layer can use CAM visualizations. The modified model must be retrained, which can also slightly trade off the model's performance. Therefore, this falls under the first category of the XAI techniques mentioned above. A more generalized approach proposed by [20] improved the limitation of CAM. The gradient CAM (GradCAM) technique considers the target object's gradients flowing into the final convolutional layer to create a localization mapping that highlights the essential regions of the target image. We use this XAI method to highlight the areas responsible for the classifier's output.

3D depth estimation and MR Visualization of the skin lesion

Depth estimation is the task of measuring each pixel relative to the camera. Concerning skin lesions, the depth of a pixel relative to the skin surface denotes how critical the situation is for the patient. Depth is extracted from either single (monocular) or multiple (stereo) views of an image. Structure from motion [21], stereo vision [22], and depth from focus/defocus [23] are used to estimate depth considering multiple images. In the following study, we have a singular top view of the skin lesion as our input for depth estimation.

The conventional methods for defocus estimation have relied on multiple images [24]. The defocus is measured using a deblurring process over an image set of the same scene captured using multiple focus settings. On the other hand, with constrained image acquisition techniques like active illumination [25] and coded aperture method [26], we can estimate depth using single images that focus on one view. However, their main drawback is that they require additional illumination and camera modification to obtain the defocus map. In [27], a novel technique is used to estimate the defocus occurrence from a single image. Defocus estimation refers to the depth estimated from a defocus blur at the edges of an image. We obtain a full defocus occurrence map by propagating the defocus blur amount into the inner portions of the image. Using the following concept, we estimate the depth of the lesion from a single image.

Consequently, once we have estimated the depth of the lesion, we want to visualize it as a 3D volumetric structure so that it can analyzed correctly. The Gabor filter is a linear filter that combines a sine wave with a Gaussian envelope. The combination of orientation with the Gaussian function makes it well-suited for edge detection [28], local feature extraction [29] and texture analysis [30]. We extend the application to 3D reconstruction by using multiple Gabor filters with different orientation and frequency values to capture the range of structural features. Apart from the frequency and orientation, bandwidth is a crucial parameter that is decided based on the characteristics of the image [31].

Once we obtain the 3D structure, we analyze the diagram and calculate lesion volume units within varying depth ranges under the skin. Our end goal from the testing pipeline is to have an interactive holographic projection visualized on a Mixed Reality (MR) headset. Using MR, the physician can analyze the criticality of the melanoma and under which stage the melanoma could be at that time. MR headsets like the Hololens 2 use the MRTK library onto which the holograms are uploaded for visualization [32]. Inter-operative navigation has been performed in different surgeries using different Extended Reality (XR) techniques. Recent studies involving navigation via holograms include [33] for percutaneous kidney puncture, [34] for laparoscopic cholecystectomy, [35] for craniomaxillofacial surgery, and [36] for laparoscopic nephrectomy. Other studies have tested the feasibility of such navigation systems; a study concerning the skin tested the feasibility of MR-based navigation towards the sentinel node in patients with melanoma [37]. In our study, we want to use MR as an analysis tool for estimating the depth of the lesion within the skin.

Methods

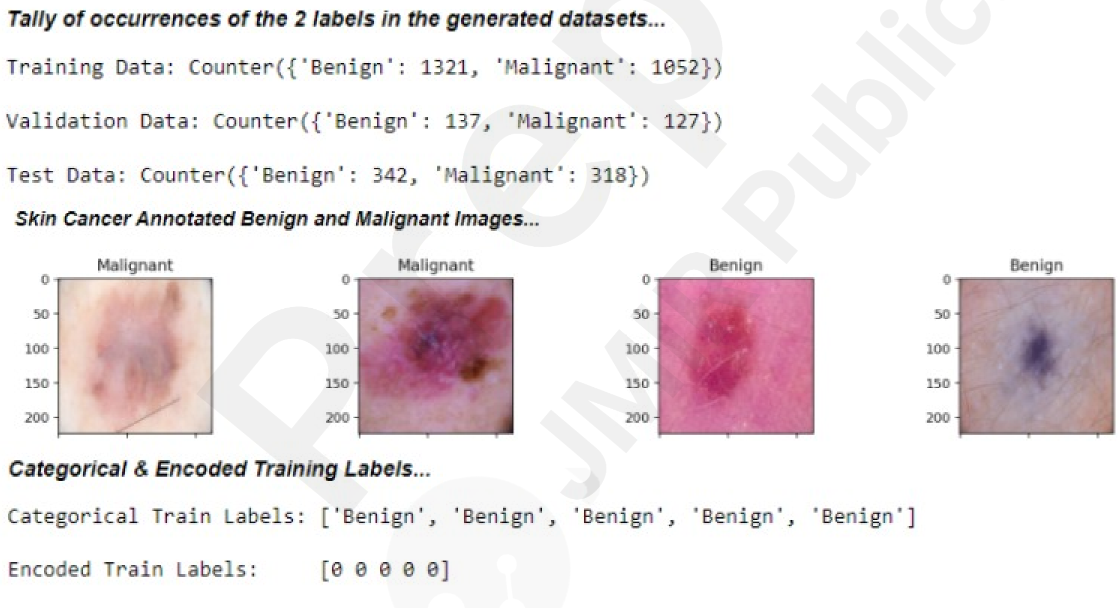
ML + XAI for melanoma detection

As mentioned in the introduction, this is the first step of our study. We detect melanoma from a skin lesion and then use GradCAM as our XAI technique that highlights the essential parameters of the image. We shall elaborate on this further in the following section.

Dataset and Data Processing

For this study, we have used processed skin cancer images from the ISIC archive. The dataset is well-balanced, having 1497 malignant images and 1800 benign images. The malignant images primarily include skin lesions that proliferate under the skin and may even reach the lymph nodes if not treated. In our dataset, nevus and melanoma are malignant by nature. If a particular lesion is labeled as malignant, there is a high chance that it could be melanoma, making it critical to detect malignancy. We also want to give an output expressing the degree of malignancy of the lesion, which tells us how close it could be to melanoma.

We first label each image as "Benign" and "Malignant" since we want a clear idea of what each data point represents during training. Both image categories have the same shape distribution of (224, 224, 3) where the area of the image is 224x224 and three represents the RGB value. Therefore, the model is trained on a uniform distribution, not affecting the output label. We then split the dataset into training, validation, and testing. After the split, we have 2373 samples in training, 264 in validation, and 660 in the test set. We use the LabelEncoder from the sklearn library to convert "Benign" and "Malignant" annotations into categorical labels. (Figure 1) better represents the dataset.



Tally of occurrences of the 2 labels in the generated datasets...

```
Training Data: Counter({'Benign': 1321, 'Malignant': 1052})

Validation Data: Counter({'Benign': 137, 'Malignant': 127})

Test Data: Counter({'Benign': 342, 'Malignant': 318})
```

Skin Cancer Annotated Benign and Malignant Images...

Categorical & Encoded Training Labels...

```
Categorical Train Labels: ['Benign', 'Benign', 'Benign', 'Benign', 'Benign']

Encoded Train Labels:     [0 0 0 0 0]
```

Figure 1: The Dataset Summary.

.

Model Architecture

The "SkinCancerDetection_VGG19_Model" architecture comprises of two levels:

- The base model: We use a pre-trained VGG19 model with weights used for the ImageNet dataset [38]. Convolution blocks 1 and 2 have two convolution layers, while blocks 3, 4 and 5 consist of four. This is to capture features hierarchically, where the earlier blocks capture low-level features like edges and textures and the latter blocks capture high-level features. Each block consists of one pooling layer. The "include_top" parameter is set to false because we want to exclude the top layers of the model for this specific task.

- The functional model: A model that uses Flatten on the base model's output and then comprises five Dense layers with the ReLU activation function. They perform hierarchical feature reduction with each layer. The final Dense layer is of one unit, with the sigmoid activation function determining the class of the output label. The functional model is depicted in (Figure 2).

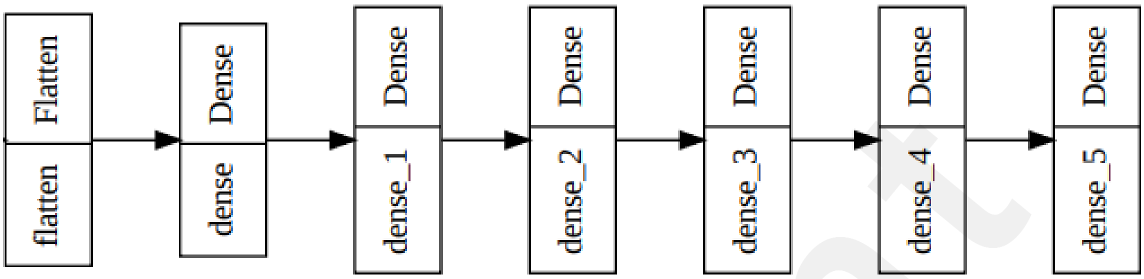


Figure 2: The Functional Model.

Pre-trained models have been trained on large, diverse datasets, imparting valuable knowledge about low-level features. They also help prevent overfitting since they are generalized on a large dataset like the ImageNet. The model loss is calculated as the Binary Cross Entropy, and we use the Adam Optimizer for optimization.

GradCAM for localization

As discussed in the literature review section, GradCAM is an intrinsically implemented XAI technique to localize the image's essential parameters. We use GradCAM instead of standard object detection methods since it gives us a heatmap of the localized area instead of a bounding box. This heatmap is useful for the depth estimation of the region. GradCAM computes gradients of the target label flowing from the final convolution layer, followed by a weighted sum of the feature maps in the final layer to create a localization mapping depicting the important parameters. In our case, the localized portion is the part of the lesion on the skin. The GradCAM is a well-established method that we implement in our GradCamUtility file. We enter the "block5_conv4" attribute into the GradCamUtility class since it is our last convolution layer of the model. The final class activation map depicts the feature maps that contributed most to the corresponding output label. A visual representation of the layers within a convolution block alongside where the GradCAM algorithm is applied is depicted in (Figure 3). An original image and its corresponding class activation map are depicted in (Figure 4).

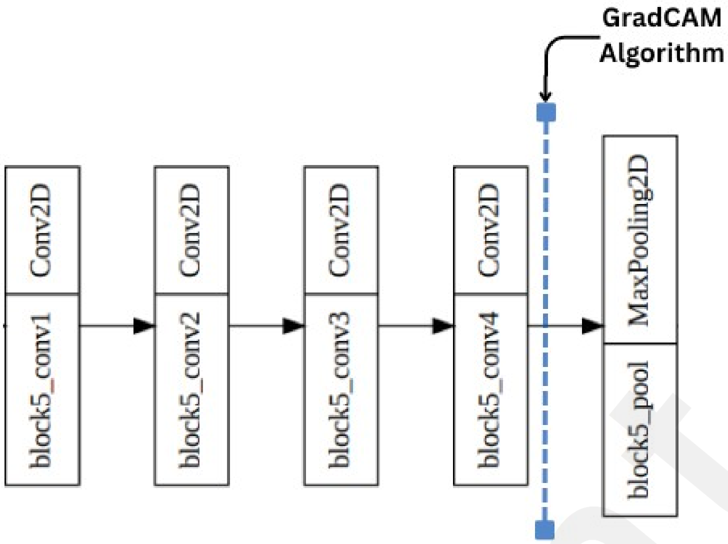


Figure 3: Application of GradCAM.

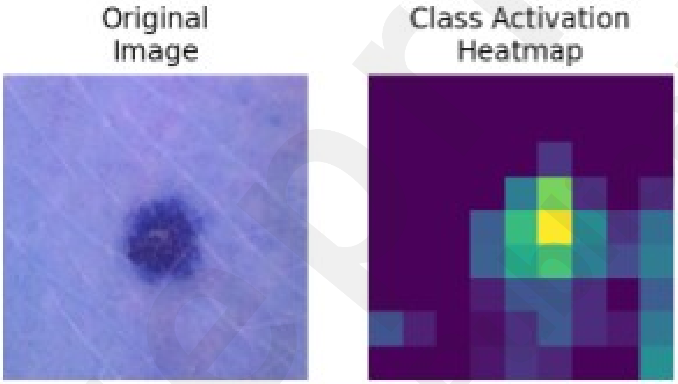


Figure 4: An Image and its Activation Map.

3D Generation of skin lesion depth

As discussed in the introduction, our second step is to generate a 3D hologram that helps physicians analyze the depth of the lesion within the skin. We do this in the three following steps:

- Estimating depth from defocus occurrence
- 3D structure representation of the depth
- Red Spot Analysis

Estimating depth from defocus occurrence

As discussed in the literature review, we use the defocus occurrence method to estimate depth when we have one input image of a particular scene. After GradCAM, we have a blended image with localization of the important parameters as our output. We perform Canny Edge detection on the blended image which extracts the edges and boundaries of the image, thereby giving us an outline of a localized area. We use the threshold values of 50 and 150 to determine strong, weak and no edges. The following values give us the appropriate boundary for the specific image.

We apply the Gaussian Blur over the entire image for smoothing since it can help create coherent depth maps and reduce the impact of noise. We apply the defocus occurrence method, with the edges and blurred values being the two input attributes of the function. Finally, we use Min-Max normalization as our normalization technique on the defocus map. Finally, we have the output defocus map. The algorithm has been performed in the code.

3D Structure representation of the depth

We use Gabor filters for the 3D structure representation of the depth [39]. The getGaborKernel of the cv2 library provides a straightforward way to generate Gabor filters with the required parameters. Tuning hyperparameters is essential for the desirable outcome. Several architectural decisions have to be made while using Gabor filters. The parameters and findings are discussed as follows:

- Number of Gabor filters: We use four at different orientations to construct a 3D structure from the 2D defocus depthmap.
- Orientation: We use Gabor filters from [0, pi/4, pi/2, 3*pi/4] to analyze textures from different orientations. For 0 degrees, the textures are analyzed horizontally, while for 90 degrees, the textures are analyzed vertically.
- Frequency: Frequency is an important parameter since it determines if we want to capture finer details. We have set the frequency values for the 0 and 90-degree filters as 0.1 since we want the finer details horizontally and vertically. We have set the frequency for the angled filters at 0.4 since we want the filters to capture the coarse details.
- Sigma value: The sigma value represents the sigma of the Gaussian distribution. It represents the blur along a particular direction. We select a low value of 0.01 along both axes for a sharper, pixelated output.
- Gamma value: We have selected a low gamma value of 0.5, making the output anti-isotropic for finer texture analysis.
- Size: We have selected a Gabor filter of dimensions (5,5). It is a trade-off size recommended for capturing the coarse and fine details of the image.

The 3D output of the Gabor filter is constructed as a scatter plot using the Plotly3D library in Python. We get a 3D heatmap on the 'RdBu' color scale, where the red spots denote the deeper-lying area of the lesion. We now introduce the red spot analysis to estimate the depth of the corresponding lesion and construct a 3D conical structure for the same.

Red Spot Analysis

The output from the Gabor filter has been constructed using the "RdBu" color scale. The "RdBu" scenario contrasts two extremes where the red values indicate areas with higher distances from the viewer, and the blue regions are closer to the imaging device. For a given depth d, the spots above it are blue, while those below it are the red spots. The red spots quantify the amount of infection since a malignant lesion could have red spots over greater depths as compared to benign cases. We check the red spots for consecutive depths and how much they decrease as we go deeper into the skin. This shows us how the number of red spots decreases with each depth range, and we can note where there are 0 red spots, which signifies the end of the lesion. A sample study of the Red Spot analysis done for four cases is depicted in (Table 1). We have two cases that are benign with lower probabilities of malignancy and two cases that are malignant. It shows a change in the number of red spots with depth for all four cases. "Below 100" represents the depth on the colormap, and the corresponding values are the red spots that proliferate deeper than 100 units. In contrast, 0 represents the skin's surface.

Depth Threshold	10 % Benign	98 % Malignant	1.6 % Benign	99.14 % Malignant
Below 0	176770	180211	178262	176263
Below 100	1809	35059	2605	21288
Below 200	470	12925	1222	11870
Below 300	235	4208	588	6573
Below 400	91	1132	283	3414
Below 500	27	211	126	1745
Below 600	2	19	6	728
Below 700	0	4	0	254
Below 800	0	0	0	132
Below 900	0	0	0	51
Below 1000	0	0	0	0

Table 1: The Red Spot Analysis.

We can infer the following observations from the depicted red spot analysis:

- We observe that the number of red spots decreases as we go deeper within the skin. This is because the lesion volume is more significant in the upper layers of the skin.
- Once we get 0 red spots, we can say that the lesion has not surpassed that particular depth threshold. For example, the first test case does not have spots above 700, signifying that the lesion depth is less than 700 units of the heatmap.
- For the malignant cases (higher probability), we notice a significantly higher volume of spots beneath the skin depth of 100 units and onwards.
- We also see that malignant tumors are deeper and can extend to "Below 900" depth units as seen in the fourth test case.

The final step in the pipeline is to represent the red spot values for each depth range as conical slices. These conical slices, when connected, represent a 3D conical structure of the lesion beneath the skin. We have developed a code that takes the red spot values for a test case as our input and gives the 3D conical structure as output. The 3D conical structure intends to depict the depth of the lesion and how much it has proliferated within the skin. A physician can view the structure as an interactive hologram on the MR headset.

Results

In this section, we display the evaluation metrics of the model on the testing data and the results from the testing pipeline output. We consolidate and summarize all the required outputs from the experiment.

Model Evaluation

As a part of the evaluation, we display the confusion matrix for each label which gives us the True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) values. We calculate the precision, recall and F1 score based on the matrix's values. We calculate the Matthews Correlation Coefficient (MCC), which is more reliable since it gets a high score only if the prediction obtains satisfying results in all four categories of the confusion matrix. The MCC is a more informative score in evaluating binary classifications than the accuracy or F1 score. We also study the AUC-ROC curve, which represents the TP-FP trade-off, and the AUC-PR score, which represents the precision-recall trade-off.

We have the following conclusions about the VGG19-GradCAM model based on the evaluation scores:

- We have an MCC score of 0.7123. Since it ranges from 0 to 1 where 0 represents random guessing and 1 represents perfect prediction. A score of 0.7123 is good since it indicates a strong positive correlation between the model predictions and class labels.
- The accuracy depends on the correctly labeled data (true data), and a score of 0.8561 is satisfactory.
- The precision, recall and F1 score depend on the false and true data points, and all three are satisfactory values. We have a precision value of 0.8410 and a recall of 0.8648. The F1-score represents the harmonic mean (balance) between the two values, taking the value of 0.8527.
- The ROC and PR scores are calculated from their respective curves, taking the values of 0.8564 and 0.8855. Since our dataset is mildly imbalanced towards the negative class (0 or benign), they are essential scores.

It is important to state that we have yet to record the trade-off in the metrics due to the addition of the Grad-CAM calculations after the final convolution layer. There may have been a mild decrease in the recorded metrics due to the presence of these manipulations. Our goal for this experiment is to estimate depth through the 3D reconstructions of the lesion.

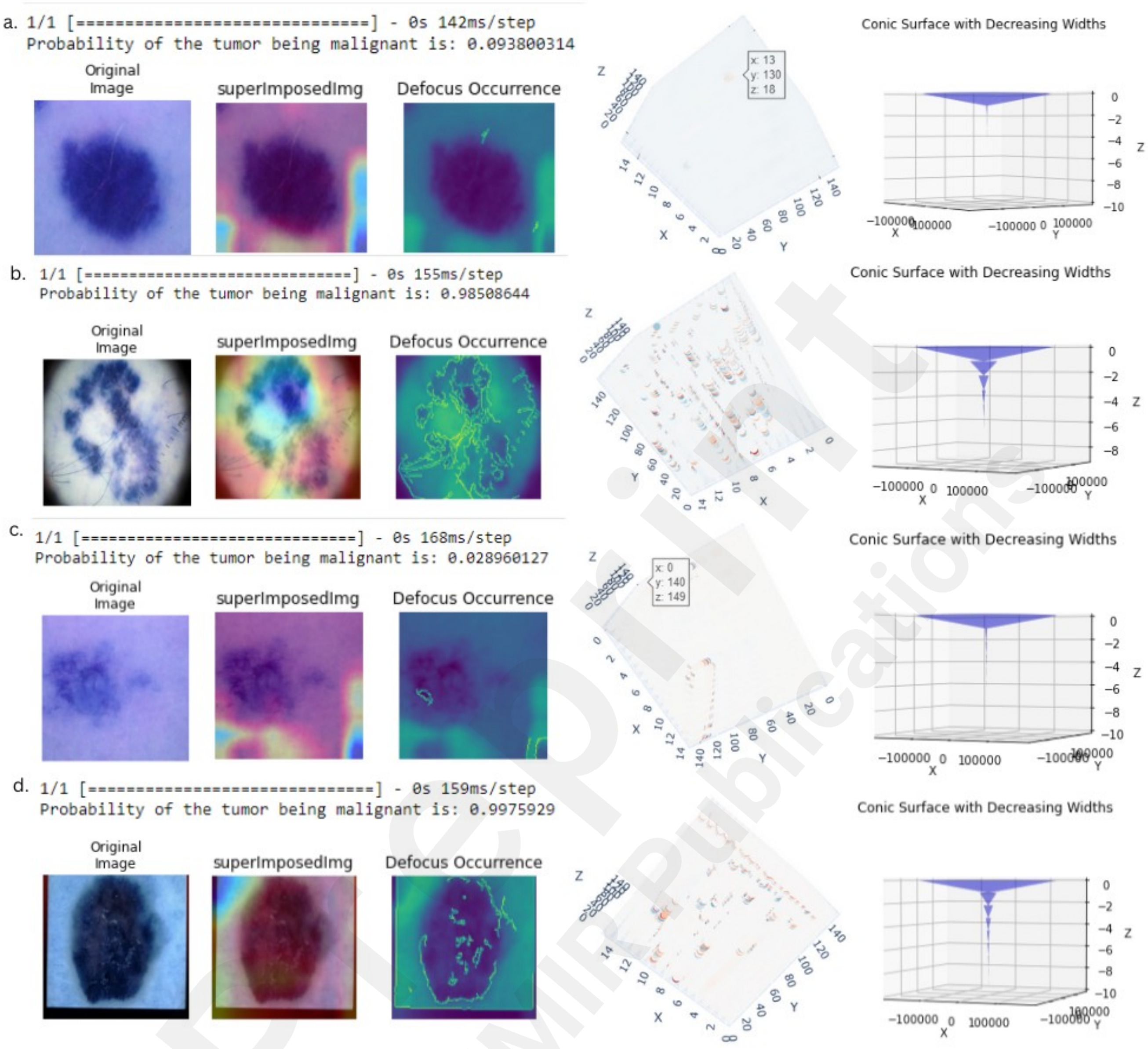
Results from the pipeline

We display the end-to-end output of every step mentioned in the study in Figure 5. The four cases correspond to those in (Table 1) since we have performed the red spot analysis for them. We train the model over the single test data point and get the probability of malignancy as the output. We apply Grad-CAM on the original image, which gives us the "superImposedImg" as the output and performs depth estimation, giving us the defocus occurrence map as shown on the "viridis" colormap. The 3D output of the depth is obtained after using the Gabor filter. After the red spot analysis, we get the final 3D conical representation of the lesion "Conic Surface with Decreasing Widths." The decreasing widths represent the number of red spots greater than a certain depth unit.

We can infer the following conclusions based on the outputs of the testing pipeline in (Figure 5):

- Non-malignant cases: A lighter-colored lesion is not heavily distinguished from the rest of the skin. We do not have a significant defocus occurrence boundary as seen in the cases of (a.) and (c.). Due to the lack of a boundary, our 3D representation is mildly covered by red or blue spots. Most spots are neutral, and our conic surface is heavily represented between the 0 and -2 Z values.
- Malignant cases: They are better distinguished by the GradCAM algorithm, and their defocus map takes a specific boundary-like structure, as seen in cases (b.) and (d.). The 3D representation of the depth shows a significantly greater number of colored spots. The conic structure has a significant concentration after the -2 Z value, extending to a greater depth than the other two cases.
- The (d.) case has an additional black outline, so we discard the line of red spots between the range of $135 < y < 145$. Such an anomalous line can give us faulty output and more red spots than there are.

The 3D conic surface and depth map are to be visualized on a MR headset by a physician since they can interact with the generated hologram and better determine the diagnosis in case of melanoma.



a.

```
1/1 [==============================] - 0s 142ms/step
Probability of the tumor being malignant is: 0.093800314
```

b.

```
1/1 [==============================] - 0s 155ms/step
Probability of the tumor being malignant is: 0.98508644
```

c.

```
1/1 [==============================] - 0s 168ms/step
Probability of the tumor being malignant is: 0.028960127
```

d.

```
1/1 [==============================] - 0s 159ms/step
Probability of the tumor being malignant is: 0.9975929
```

Figure 5: Outputs of the Testing Pipeline.

Discussion

Limitations and Challenges

The study comprises computer vision, graphics and an MR headset, resulting in a few limitations along the pipeline. Concerning the vision model, we have yet to quantify how much the GradCAM algorithm has affected the model performance. If the difference is significant, we must consider using GradCAM++ or other alternate segmentation methods. Although computer models have achieved a higher accuracy during melanoma detection, it would be the next step to match such accuracy while using XAI methods alongside the model.

The estimation of depth using the defocus method heavily relies on the color and width of the lesion as seen on the skin surface. Despite being malignant, we may not get an appropriate depth map output if our lesion is not sufficiently dark or large. This method also needs to extend to cases where the image is zoomed out, and we can see the entire body part on which the lesion is present. For example, one test case had the presence of a thumb on which the lesion was present, resulting in a slightly inaccurate depth output. We get additional red spots due to the presence of the thumb, making the lesion deeper than it is.

As computer science researchers, it would not be our place to state the importance of this study in a practical setting. To extend this research to practical applications, it would be crucial to know a physician's opinion of the hologram when visualized by them through the MR headset. Apart from that, it is also essential to evaluate the generated holograms on the MR headset. For its evaluation, further research must be performed to analyze the validity and fidelity of the hologram rendered onto the

Future Work

The mentioned limitations could facilitate further research for depth estimation of melanoma skin lesions. The proposed approach is qualitative. and it needs to be extended to more cases. We must know if this is valid for all other benign/malignant cases, and this metric must be quantified somehow. We also want to extend this experiment to evaluate the generated holograms for an MR device using the concepts of validity and fidelity. This will be explored in further research until we develop a practical tool for improving melanoma treatment.

Researchers who pursue computer vision problems could study the trade-off between the model accuracy and XAI methods. We could extend computer graphics research to estimate the depth of images that comprise a body part with a lesion. An interesting concept that was introduced in the review section was that of the PI value. Calculating the PI value requires different datasets and methodologies, and it would be insightful to study its relationship with the lesion's depth.

Conflicts of Interest

None Declared.

Abbreviations

All are explained in the main body of the manuscript.

References

1. T Y Satheesha, D Satyanarayana, M G Prasad, and K D Dhruve. Melanoma is skin deep: a 3d reconstruction technique for computerized dermoscopic skin lesion classification. IEEE journal of translational engineering in health and medicine; 5:1–17; 2017.
2. K Shimizu, H Iyatomi, M E Celebi, K A Norton, and M Tanaka. Four-class classification of skin lesions with task decomposition strategy. In IEEE transactions on biomedical engineering; pages: 274–283; 62(1); 2014.
3. P Bhati and M Singhal. Early-stage detection and classification of melanoma; pages: 181–185; November 2015.
4. M Kłosok and M Chlebus. Towards better understanding of complex machine learning models using explainable artificial intelligence (XAI): Case of credit scoring modelling. University of Warsaw; Faculty of Economic Sciences, Warsaw; 2020.
5. D Sahija. Critical review of mixed reality integration with medical devices for patientcare;

2022; 10, 2022.

6. L Khairunnahar, M A Hasib, R H B Rezanur, M R Islam, and M K Hosain. Classification of malignant and benign tissue with logistic regression. Informatics in Medicine Unlocked; 16(10018):9; 2019.
7. I A Ozkan and M Koklu. Skin lesion classification using machine learning algorithms; pages: 285–289; International Journal of Intelligent Systems and Applications in Engineering; 5(4); 2017.
8. R Garnavi, R Aldeen, and M J Bailey. Computer-aided diagnosis of melanoma using border- and wavelet-based texture analysis. In IEEE Trans. Inf. Technol. Biomed. vol. 16; no. 6; pp: 1239–1252; 2012.
9. K M. Hosny, M A Kassem, and M M Foaud. Classification of skin lesions using transfer learning and augmentation with alex-net. PloS one; 14:5; 2019.
10. JF Alcón, C Ciuhu, Ten Kate, Heinrich W, Uzunbajakava A, & de Haan G N, Krekels. Automatic imaging system with decision support for inspection of pigmented skin lesions and melanoma diagnosis; pages: 14–25; IEEE journal of selected topics in signal processing 3(1); 2009.
11. P Wighton, T K Lee, H Lui, D I McLean, and M S Atkins. Generalizing common tasks in automated skin lesion diagnosis. IEEE Transactions on Information Technology in Biomedicine; 15(4):622–629; 2011.
12. M E Celebi and A Zornberg. Automated quantification of clinically significant colors in dermoscopy images and its application to skin lesion classification; In IEEE systems journal; pages 980–984 8(3); 2014.
13. S Alheejawi, H Xu, R Berendt, N Jha, and M Mandal. Novel lymph node segmentation and proliferation index measurement for skin melanoma biopsy images; Computerized Medical Imaging and Graphics; 73:19–29; 2019.
14. N S Kumar, K Hariprasath, S Tamilselvi, A Kavinya, and N Kaviyavarshini. Detection of stages of melanoma using deep learning. Multimedia Tools and Applications; 80:18677–18692; 2021.
15. T Chanda, K Hauser, S Hobelsberger, T C Bucher, C N Garcia, and & Brinker T J Wies; C Dermatologist-like explainable AI enhances trust and confidence in diagnosing melanoma.; 2023.
16. Z Papanastasopoulos, R K Samala, H P Chan, L Hadjiiski, C Paramagul, M A Helvie, and C H Neal. Explainable ai for medical imaging: deep-learning cnn ensemble for classification of estrogen receptor status from breast mri. In Medical imaging Computer-aided diagnosi; Vol. 11314, pp: 228-235; SPIE; 1131:228–235; March 2020.
17. E Nasr-Esfahani, S Samavi, N Karimi, S M R Soroushmehr, M H Jafari, K Ward, and K Najarian. Melanoma detection by analysis of clinical images using convolutional neural network. In 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); pages 1373–1376; IEEE; August 2016.
18. H K Gajera, D R Nayak, and M A Zaveri. A comprehensive analysis of dermoscopy images for melanoma detection via deep cnn features. Biomedical Signal Processing and Control; 79(10418):6; 2023.
19. B Zhou, A Khosla, A Lapedriza, A Oliva, and A Torralba. Learning deep features for discriminative localization. In Proceedings of the IEEE conference on computer vision and pattern recognition; pages: 2921–2929; 2016.
20. R R Selvaraju, A Das, R Vedantam, M Cogswell, D Parikh, and D Batra. Grad-cam: Why did you

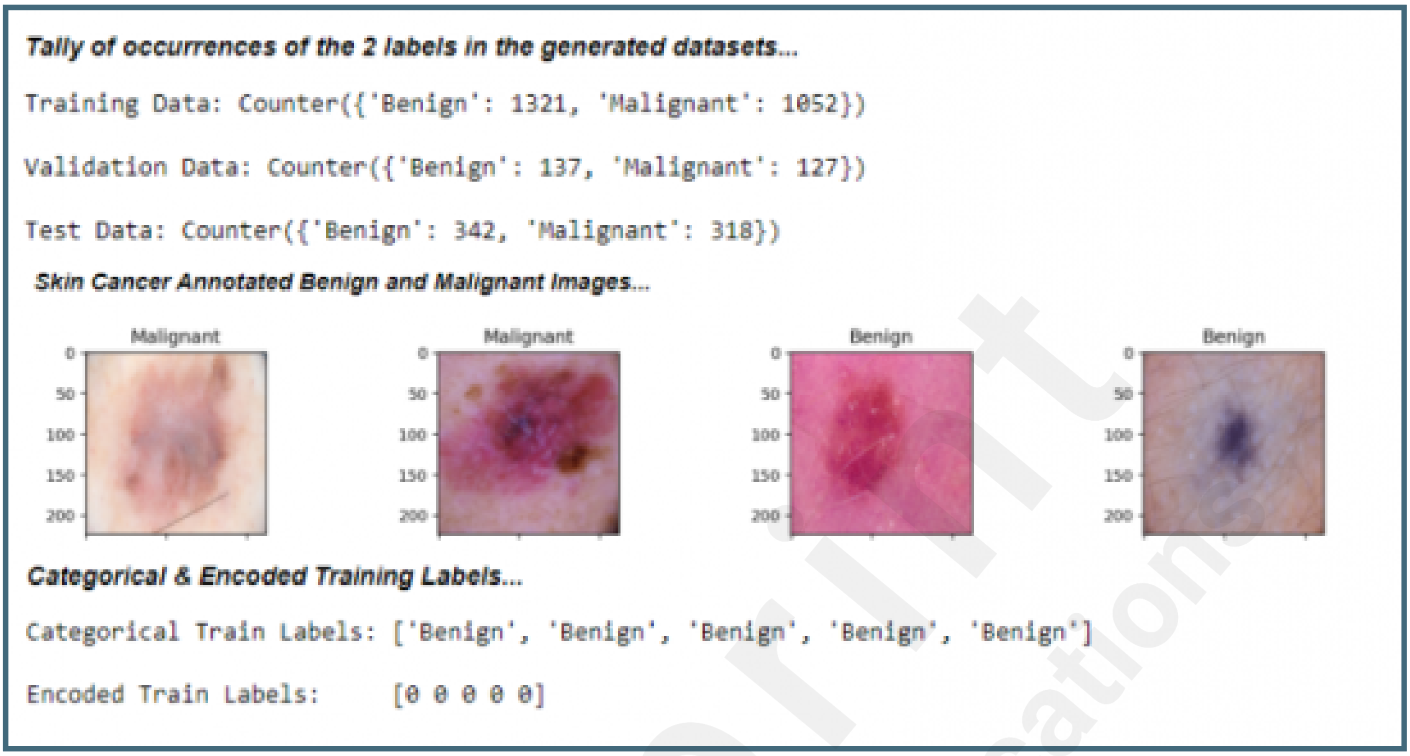
say that?; arxiv; In preprint; 2016.
21. J L Schonberger and J M Frahm. Structure-from-motion revisited. In Proceedings of the IEEE conference on computer vision and pattern recognition; pages: 4104–4113; 2016.
22. R A Hamzah and H Ibrahim. Literature survey on stereo vision disparity map algorithms. Journal of Sensors; 2016.
23. M Subbarao and G Surya. Depth from defocus: A spatial domain approach. International Journal of Computer Vision; 13(3):271–294; 1994.
24. P Favaro and S Soatto. A geometric approach to shape from defocus. IEEE Transactions on Pattern Analysis and Machine Intelligence; 27(3):406–417; 2015.
25. F. Moreno-Noguer, P. N. Belhumeur, and S. K. Nayar. Active refocusing of images and videos. ACM Transactions On Graphics (TOG); 26(3):67–es; 2007.
26. A Levin, R Fergus, F Durand, and W T Freeman. Image and depth from a conventional camera with a coded aperture. ACM transactions on graphics (TOG); 26(3): 70–es; 2007.
27. S Zhuo and T Sim. Defocus map estimation from a single image. Pattern Recognition; 44(9):1852–1858; 2011.
28. R Mehrotra, K R Namuduri, and N Ranganathan. Gabor filter-based edge detection. Pattern recognition; 25(12):1479–1494; 1992.
29. X Wang and & Liu C Ding, X Gabor filters-based feature extraction for character recognition. Pattern recognition; 38(3):369–379; 2005.
30. V S Vyas and P Rege. Automated texture analysis with gabor filter. GVIP journal; 6(1):35–41; 2006.
31. J R Movellan. Tutorial on gabor filters. Open source document; 40:1–23; 2002.
32. S Ong, V K Siddaraju, S Ong, and V K Siddaraju. Introduction to the mixed reality toolkit. Beginning Windows Mixed Reality Programming: For HoloLens and Mixed Reality Headsets; pages: 85–110; 2021.
33. F Porpiglia, E Checcucci, D Amparore, D Peretti, F Piramide, and C De Cillis, S & Fiori. Percutaneous kidney puncture with three-dimensional mixed-reality hologram guidance: from preoperative planning to intraoperative navigation. European Urology; 81(6):588–597; 2022.
34. M Kitagawa, M Sugimoto, H Haruta, A Umezawa, and Y Kurokawa. Intraoperative holography navigation using a mixed-reality wearable computer during laparoscopic cholecystectomy. Surgery; 171(4):1006–1013; 2022.
35. EZ Cai, Y Gao, KY Ngiam, and TC Lim. Mixed reality intraoperative navigation in craniomaxillofacial surgery; Plastic and Reconstructive Surgery; 148(4):686e–688e; 2021.
36. G Li, J Dong, J Wang, D Cao, X Zhang, Z Cao, and G Lu. The clinical application value of mixed-reality-assisted surgical navigation for laparoscopic nephrectomy; Cancer Medicine; 9(15):5480–5489; 2020.
37. Engelen T, Brouwer O R, Mathéron H M, Valdés-Olmos R A, Nieweg O E & van Leeuwen F W, van den Berg N S. A pilot study of spect/ct-based mixed-reality navigation towards the sentinel node in patients with melanoma or merkel cell carcinoma of a lower extremity. Nuclear Medicine Communications; 37(8):812–817.
38. M Simon, E Rodner, and J Denzler. Imagenet pre-trained models with batch normalization. Arxiv; preprint; 2016.
39. L A Camuñas-Mesa, T Serrano-Gotarredona, S H Ieng, R B Benosman, and B Linares-Barranco; On the use of orientation filters for 3d reconstruction in event-driven stereo vision; Frontiers in neuroscience; 8:48; 2014.

40. C Lee, G A Rincon, G Meyer, T Höllerer, and D A Bowman. The effects of visual realism on search tasks in mixed reality simulation. IEEE transactions on visualization and computer graphics; 19(4):547–556; 2013.

Supplementary Files

Figures

The Dataset Summary.



Tally of occurrences of the 2 labels in the generated datasets...

```
Training Data: Counter({'Benign': 1321, 'Malignant': 1052})

Validation Data: Counter({'Benign': 137, 'Malignant': 127})

Test Data: Counter({'Benign': 342, 'Malignant': 318})
```

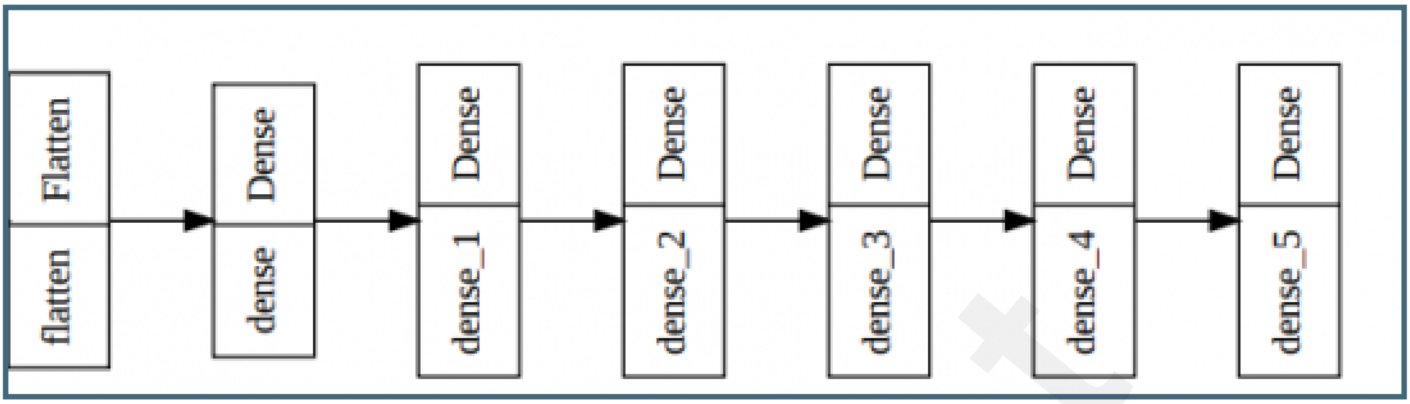
Skin Cancer Annotated Benign and Malignant Images...

Categorical & Encoded Training Labels...

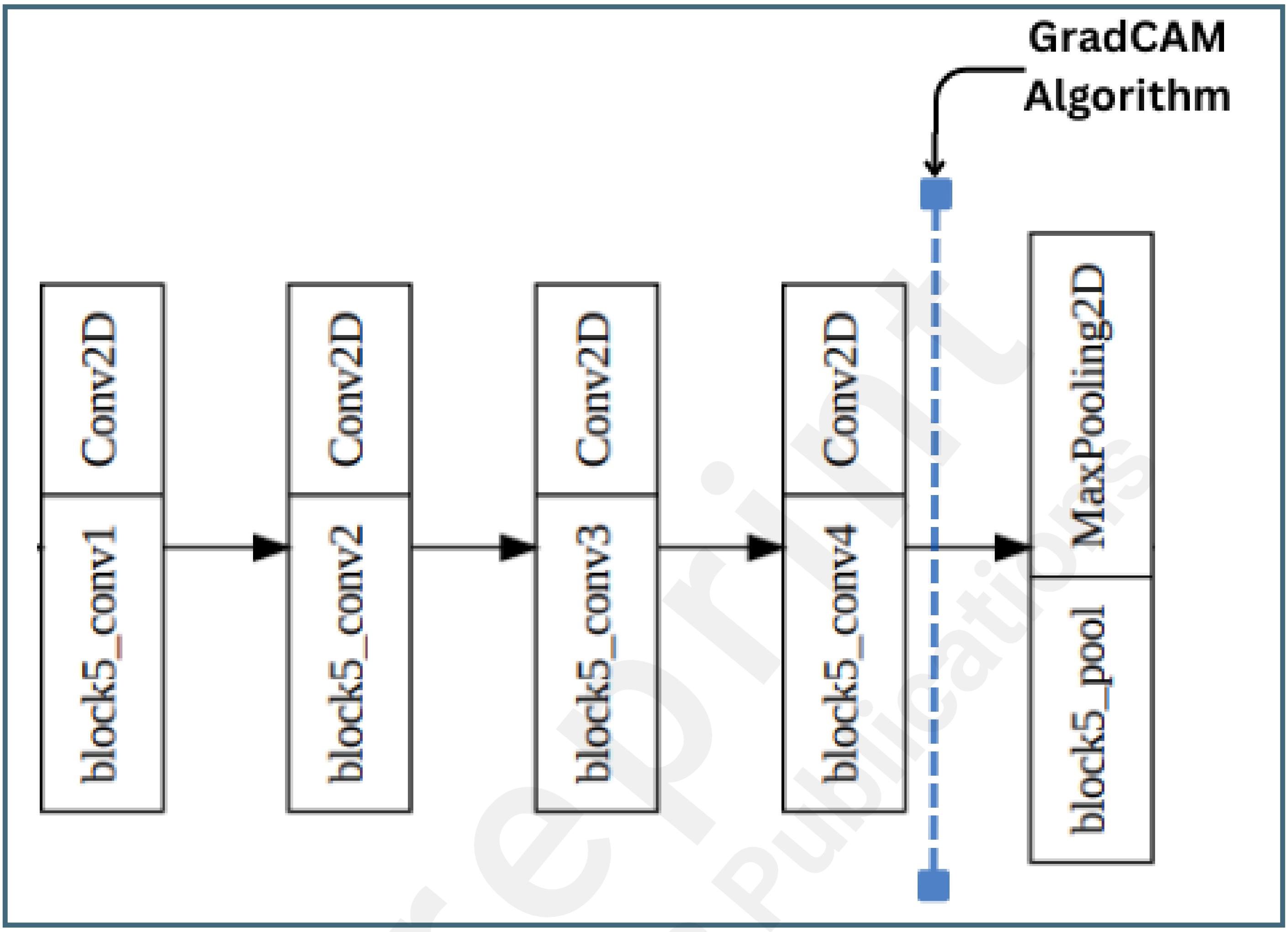
```
Categorical Train Labels: ['Benign', 'Benign', 'Benign', 'Benign', 'Benign']

Encoded Train Labels:     [0 0 0 0 0]
```

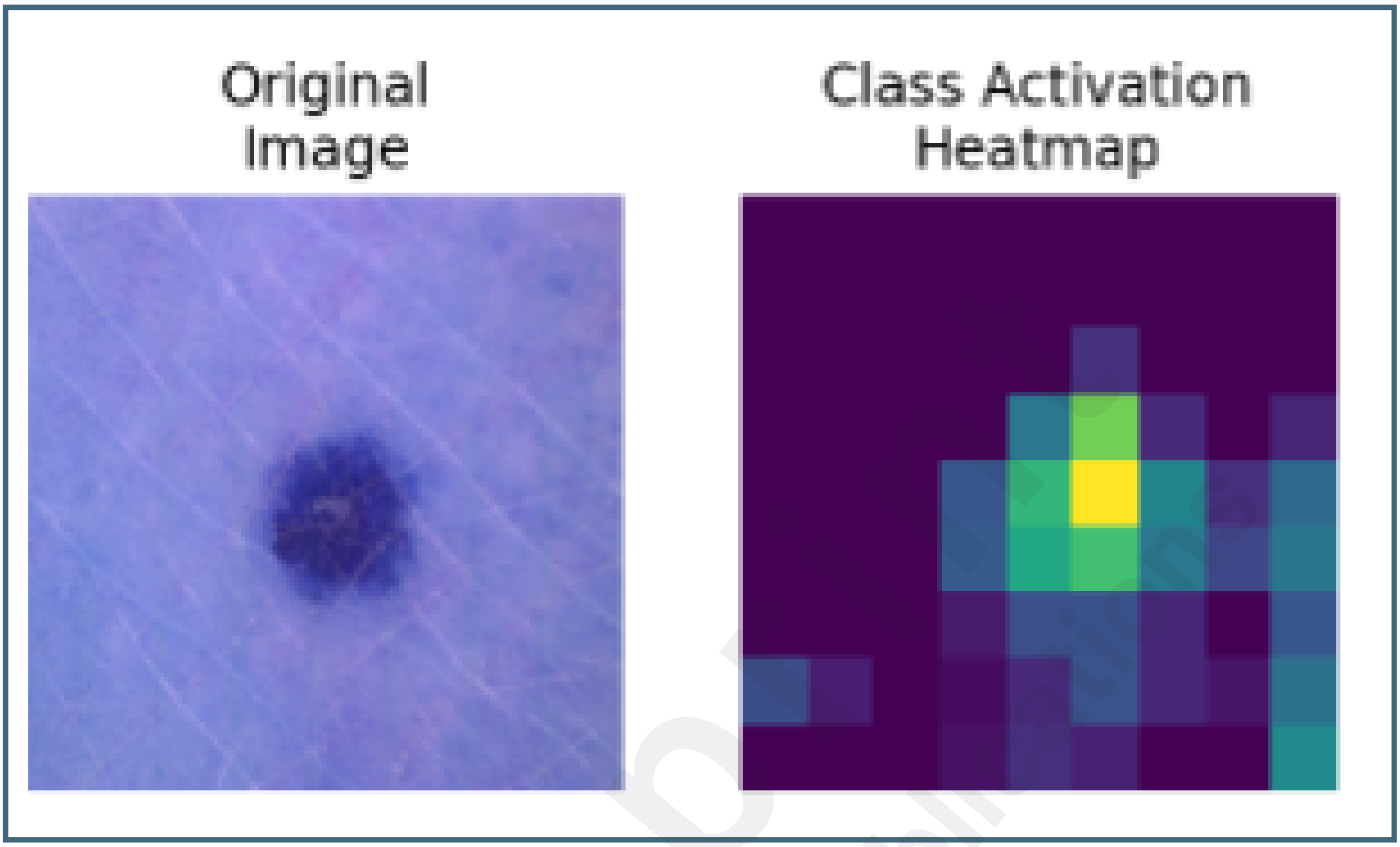
The Functional Model.



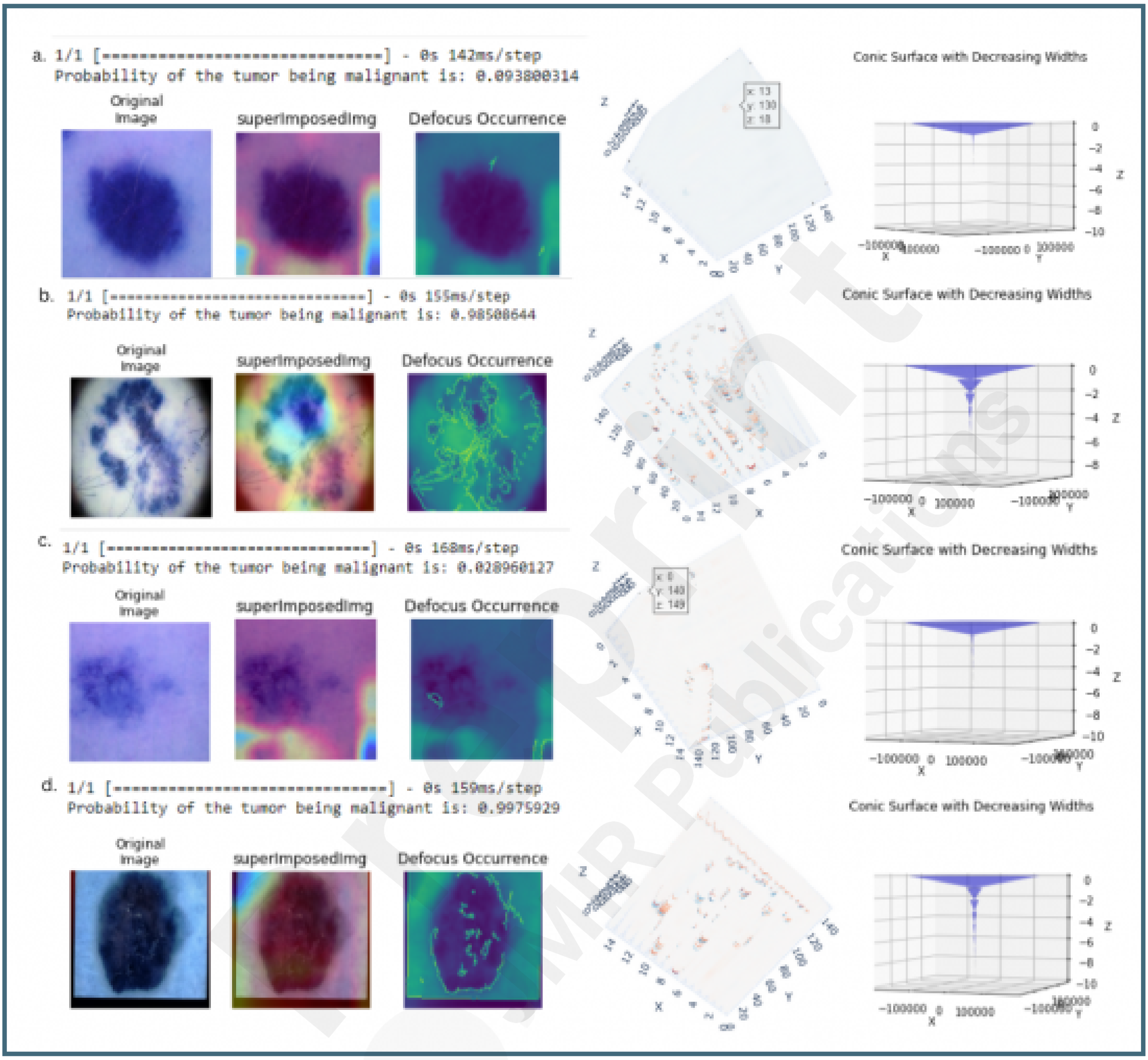
Application of GradCAM.



An Image and its Activation Map.



Outputs of the Testing Pipeline.



a.
```
1/1 [==============================] - 0s 142ms/step
Probability of the tumor being malignant is: 0.093800314
```

b.
```
1/1 [==============================] - 0s 155ms/step
Probability of the tumor being malignant is: 0.98508644
```

c.
```
1/1 [==============================] - 0s 168ms/step
Probability of the tumor being malignant is: 0.028960127
```

d.
```
1/1 [==============================] - 0s 159ms/step
Probability of the tumor being malignant is: 0.9975929
```