Real-time Skin Cancer Detection Using Neural Networks on an Embedded Device

Rowin Veneman
University of Twente
PO Box 217, 7500 AE Enschede
The Netherlands
r.veneman@student.utwente.nl

Abstract—Skin cancer is one of the most common types of cancer there is. Skin cancer develops on the skin and can spread to other areas of the body possibly causing extensive damage. The earlier the cancer is treated the better the changes are of survival. Since skin cancer grows on the outer layer of the skin, it can be diagnosed by a trained eye. Each type of skin cancer has certain visual characteristics which makes it possible for Artificial Intelligence (AI) to determine whether a tumour is a specific type of cancer or whether it is a benign tumour. However, many people do not know how to recognize skin cancer and people might not be inclined to visit a doctor for a simple discolouring on the skin. This research paper will try and develop a smartphone application that can detect skin tumour in real-time using object detection. Additionally, it will also be able to classify skin cancer from a single image using a trained Convolutional Neural Network (CNN).

*Index Terms*—Skin cancer, Convolutional Neural Network, Embedded devices, smartphone

I. INTRODUCTION

Skin cancer is a common type of cancer with approximately 22.000 new cases in the Netherlands in 2020 and approximately 127.200 cases in total[9]. Skin cancer develops on the skin and depending on the type of skin cancer, it can look like a mole or a dry spot of skin[4]. A spot on the skin is called a skin tumour and these can be malignant which means cancerous, or benign which is not cancerous. Moles and warts are two common types of benign tumours[10]. Among the malignant tumours there are various types, however this project will not differentiate between them. On the surface, the tumour can look harmless, but it can grow into the skin and even spread to other parts of the body. The only way to prevent this is to start treatment as soon as possible, and this requires getting the diagnosis as soon as possible. Making it easier for people to diagnose skin cancer themselves is helpful in getting an early diagnosis. People might not be inclined to have every tumour on their skin checked out by a doctor, but if all they have to do is take a picture or film themselves with their smartphone then people can detect cancer sooner. As with any cancer, the earlier it is diagnosed the better the chances of survival. There are several types of skin cancer and they have different characteristics, these types vary in appearance and effect on the body. Logically, some types are more dangerous than others.

An application will be developed that uses an object detection model for real-time skin tumour detection and a trained convolutional neural network(CNN) that can classify skin cancer based on a single image. The object detection model allows the user to scan large parts of their body quickly for skin tumours that can be either benign or malignant. The classification model allows the users to take or select a picture and determine whether that image contains a cancerous tumour. Both models have different use cases and can help the user in detecting potential skin cancer.

Figure 1. Several examples of malignant and benign tumours.
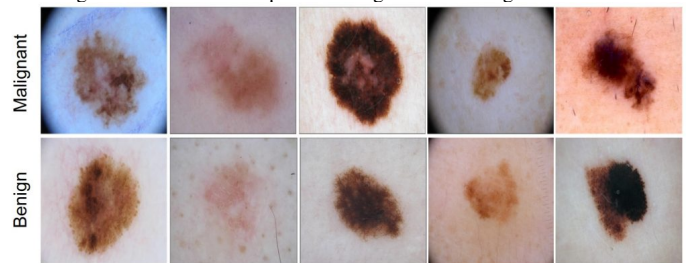
II. PROBLEM STATEMENT

Skin cancer needs to be detected as soon as possible. The sooner the cancer is diagnosed, the more effective treatment is. When skin cancer grows it can spread to other parts of the body making it hard to treat. Therefore diagnosing skin cancer should be made as simple as possible, that is where this application comes in. Whilst doctors will still have to do the actual diagnosis, the application can give the user an indication of a tumour being benign or malignant. This application is possible, because skin cancer can be diagnosed by an dermatologist by examining the skin. The diagnosis can then be confirmed by performing a skin biopsy. The most important visual characteristics of a cancerous tumour is change over time, in addition, asymmetrical shape and uneven colours are important indicators too. CNNs can be trained to recognize the patterns of both malignant and benign tumours allowing them to perform an initial diagnosis. There are existing application that run on smartphones that can detect skin cancer with high accuracy, however these process a single image. This research will try to make the process real-time on a smartphone, this allows users to scan multiple tumours at once. In addition, the regular possibility of processing a single image will also be available. If this research is successful, then

it can be added to other (embedded) devices too. The target group of the application is elderly people, since they are more likely to get skin cancer.

III. RESEARCH QUESTION

The problem statement can be described using the following research question:

Can a smartphone do real-time skin cancer detection with accuracy on par with human experts and other applications? This can be answered with the following sub-questions:

- How to classify skin cancer using a neural network?
- How to detect potential skin cancer in real-time?
- How to deploy to the skin cancer classification and detection to a smartphone?

IV. RELATED WORK

In order to obtain related work, research domain Google Scholar and Scopus were used. The used search terms were: "skin cancer", "skin cancer detection", "neural networks" and "object detection". There has already been some research on this subject.

Related works have shown that convolutional neural networks are able to classify skin cancer at the levels of human experts. Study [14] has shown that their CNN is better than beginner and intermediate raters, and on par with expert level raters. Study [2] does a systemic review of different skin cancer classifiers using CNNs, they analysed other studies and discuss the techniques used by these studies. Furthermore, [2] also discusses the validity of the studies, one of the most important aspects is the amount of data available for each class classified. In addition, [2] concludes that the development of large public archives are required, since existing archives mostly consist of skin lesions of light-skinned people. Study [6] developed a neural network, that detected skin cancer on a pre-processed image with an overall accuracy of 0.76. Study [6] discusses their approach clearly, and their approach can be studied and used to improve the results of this paper. Similarly to study [6], study [11] developed a CNN that can classify skin cancer using an image. Study [11] classifies two cases, benign and malignant and reached an accuracy of 0.81. Furthermore, study [5] also developed a CNN that reached an average accuracy of 0.94 identifying four types of skin cancer. Finally, study [8] developed a CNN that can detect 134 disorders, that study also accurately predicted malignancy and could suggest primary treatment options. Their CNN is used as an addition to the diagnosis of medical professionals to improve the reliability of the overall diagnosis. Study [1] discusses acYou Only Look Once which is an object detection algorithm based on a convolutional neural network. In particular, study [1] discusses YOLOv4 which can run in real time whilst still producing accurate results.

The research done in this field has already achieved great results. Many of the studies achieved results on par with human experts.

Additionally, study [15] managed to increase the validation accuracy by 12.4% of a CNN detecting melanoma's by using data augmentation techniques. Data augmentation techniques can be used to increase the amount of available data by creating variations of the already available data. This gives the model more data to be trained on which results in a better model.

V. METHODS

A. Dataset

Data was obtained from the international skin imaging collaboration(ISIC). The ISIC database contains in total 5714 images of malignant tumours and 47684 images of benign tumours. Dimensions range from 640x480 to 6000 x 4000. All diagnosis were confirmed using one of several methods, and in all cases of malignancy diagnosis was confirmed histopathologically(tissue analysis using a microscope). Furthermore, all data was anonymised by ISIC and included some metadata such as age, sex and location of the tumour.

Additionally, a COCO data set was downloaded[3]. This was used as the third class for the image classification. This data set contains all sorts of images, from bread to humans to food. This data set was added in case the user took pictures of something else entirely.

Data augmentation was used to increase the amount images of malignant tumours.

B. Model Architecture

The image classification model was built using a pre-trained MobileNetV3. In addition to transfer learning, fine-tuning was also used to make the model more relevant for this specific task. MobileNetV3 was chosen because it is an accurate, yet lightweight network. For resource-constrained embedded devices, such as smartphones, it is important that the network is not too big. This architecture requires images to be scaled down to a resolution of 224 x 224, images were scaled down bi-linearly. The object detection model was built on a pre-trained YOLO model, specifically YOLOv4-tiny. YOLOv4-tiny is a pre-trained object detection model that reaches near real-time speeds. YOLOv4-tiny was used, because it is a small model capable of high-speeds and high accuracy. YOLO can be trained to recognize any object. The YOLO model outputs a bounding box that indicates where the object is found in the image, together with the class of the object and a confidence level.

C. Training

The image classification model was trained on all 6000 malignant images, 6000 benign images and 6000 images from the COCO dataset..

YOLOv4-tiny was trained on 11740 self-labelled images. 5740 Malignant and 6000 benign. These were split randomly, with a chance of 0.2, this resulted in 9801 images for training and 1939 images for validation.

D. Mobile Application

The application was developed using Flutter[7]. Flutter is an open-source toolkit for building cross-platform applications. Flutter can be deployed to both Android and IOS, however it can also deploy to web, desktop and embedded devices. The application should be easy to use, since the target group consists of elderly people.

To deploy the models the tflite_flutter plugin[12] will be used. This allows any tflite model to be interpreted directly on the smartphone. Additionally, the tflite_flutter_helper[13] plugin will be used for processing of input and output. The tflite_flutter plugin comes with examples, which will be used to implement the model in Flutter.

VI. RESULTS

A. Image Classification Model

Figure 2. The image classification model reached high accuracy for all three classes. The left-hand side is the classifier result, the top is the actual class.

Confusion Matrix for the Image Classification Model				
	Malig.	Benign	Other	Precision
Malig.	832	67	1	92.44%
Benign	88	810	2	90.00%
Other	2	1	897	99.67 %
Recall	90.24%	92.26%	99.67%	-

The overall accuracy of the model is 94.04%.

The image classification model includes a third class, called 'others', because the model would classify random images as either benign or malignant with high confidence. This happened because one of the final layers is a softmax layer, which causes the sum of all probabilities to equal one. Therefore, the 'other' class was added, so that random objects would not be classified as malignant or benign with a confidence of 1.0.

B. Object Detection Model

The object has a mean accurate precision(mAP) of 89.34% at a confidence threshold of 0.50. Furthermore, at a confidence level of 0.25, the precision was 0.81, the recall was 0.89 and the F1-score was 0.85.

C. Mobile Application

Both models were successfully implemented in the mobile application, see the appendix for screenshots of the application. The image classification model takes about a second to process.

The object detection model has to reach high speeds, since it is supposed to run in real-time. The object detection model has two important steps, image pre-processing, where the image is loaded and resized to the correct format and the inference step, where the image is run through the model. The pre-processing takes 36.07ms(N=15), inference takes 232.14ms(N=15). The total elapsed time for a single frame to be processed is 353.18ms(N=15). This means the YOLO model can run at 2.83 frames per second.

VII. DISCUSSION

The accuracy's of both models are promising. Both models show the potential of CNNs to recognize skin cancer on similar levels to human experts. These models can be used in unison, the object detection model is good for a quick scan whereas the image classification model is more accurate and better suited for a single scan. Combining both models gives users the best of both worlds, fast scanning and accurate results.

The most important metric for both models is their recall, images that are predicted to be benign but in reality are malignant. False benign cases pose a tremendous risk for the user's health, because malignant tumours should be treated as soon as possible and if they go undetected then the cancer has more chance to spread. The precision should be as low as possible too, but it is not as harmful as having a bad recall.

The object detection model has good results. The mAP and recall are the most important values, and these are both good. The model does not run at real-time speeds, but still fast enough to be useful in real-life application.

These models can be used by people at home, however they can also be used in AI-assisted diagnosis. The doctor would independently provide a diagnosis, and use these models as a second opinion. If either diagnosis is confident that the tumour is malignant, then a skin biopsy could be performed to confirm the diagnosis.

There are several ways this project can be improved. These are discussed in the next chapter.

VIII. RECOMMENDATIONS

There are several ways to improve the performance of the models. Firstly, metadata should be taken into account. Data such as age, location of the tumour on the body and genetic disposition to cancer could all be added to the model and could help improve accuracy. This would require more work on the part of the user, but it could be add to the accuracy of the model. In addition, the models could be improved if they take into account the change of tumours over time. For this to work, the model needs to remember what a specific tumour looks like and then compare it to the same tumour when scanned next time. The user should scan their body at a certain interval too. This helps increase performance, since change over time is an important visual characteristic of malignant tumours. Moreover, the model is trained on dermoscopic images taken by doctors and dermatologists, which are high quality and clear images, but not reflective of what the images taken by users. The model would therefore be more consistent if it was trained on images taken by users, instead of the high quality professional ones.

Furthermore, the data used to train both models is limited to images taken from light skinned people. As a consequence, the performance has not been tested on tumours on darker skins. To solve this problem, more data should be collected from different tumours on different colors of skin.

The use case of this application is to check from home whether a skin tumour is malignant or benign, in this use case it is not a necessity to know which type of skin cancer it is.

However there are different types of malignant and benign skin tumours, and if these models were to assist doctors in their diagnosis having more classes could be useful.

The image classification model requires resolutions of 224x224, this can be a problem. Results could be better if the input image has more detail, one way to solve this would be to increase the allowed resolution of the image, however then the model would need to be re-designed. Another solution would be to use object detection to find where the tumour is on the image, and crop the image before re-scaling to surround the bounding box. This would reduce the amount of useless information in the image.

The softmax layer of the MobileNetV3 model should be removed. This required the model to be re-trained and re-designed somewhat, but then the probability of the results are more nuanced. Nuanced probabilities make sense for cancer detection, since even with a 30% chance of malignancy, the user should still go to the doctor.

Before the application can be released, the application should be tested in real-life on skin tumours, both benign and malignant.

IX. CONCLUSION

The potential for artificial intelligence(AI)-assisted diagnosis, or even fully autonomous AI diagnosis is clear. In case of the latter many improvements can still be made, such as the ones mentioned in the recommendations section. Image classification and object detection models will continue to improve, which would mean better performing models allowing for higher accuracy and lower recall. Even in the current state, these models can aid doctors in raising the confidence in their diagnosis, and this is not limited to skin cancer, this technology could be applied to many other diseases. This will be something we will be seeing more of in the future.

Moreover, this project also shows the potential for artificial doctors on embedded devices. These AI doctors can be deployed to smartphones, but could also be deployed to other embedded devices. There are many possible applications for this technology, allowing for many diseases to be diagnosed by embedded devices around the house. One such application could be devices that automatically diagnose stool and urine samples, as many diseases could be diagnosed this way. This is one specific example, but there are a lot of possibilities. This project shows that with relative ease, a model can be created that is on par with human experts.

REFERENCES

[1] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. "Yolov4: Optimal speed and accuracy of object detection". In: *arXiv preprint arXiv:2004.10934* (2020).

[2] Titus Josef Brinker et al. "Skin cancer classification using convolutional neural networks: systematic review". In: *Journal of medical Internet research* 20.10 (2018), e11936.

[3] *COCO common objects in Context*. URL: https://cocodataset.org/#home.

[4] American Academy of Dermatology Association. *Types of skin cancer*. URL: https://www.aad.org/public/diseases/skin-cancer/types/common.

[5] Ulzii-Orshikh Dorj et al. "The skin cancer classification using deep convolutional neural network". In: *Multimedia Tools and Applications* 77.8 (2018), pp. 9909–9924.

[6] Pratik Dubal et al. "Skin cancer detection and classification". In: *2017 6th International Conference on Electrical Engineering and Informatics (ICEEI)*. 2017, pp. 1–6. DOI: 10.1109/ICEEI.2017.8312419.

[7] *Flutter. Design Beautiful Apps*. URL: https://flutter.dev/.

[8] S Han et al. "991 Deep neural networks empower medical professionals in diagnosing skin cancer and predicting treatment options for general skin disorders". In: *Journal of Investigative Dermatology* 139.5 (2019), S171.

[9] *Huidkanker. Cijfers & Context*. URL: https://www.volksgezondheidenzorg.info/onderwerp/huidkanker/cijfers-context/huidige-situatie.

[10] *Skin Cancer (Non-Melanoma): Introduction*. URL: https://www.cancer.net/cancer-types/skin-cancer-non-melanoma/introduction.

[11] Serban Radu Stefan Jianu, Loretta Ichim, and Dan Popescu. "Automatic Diagnosis of Skin Cancer Using Neural Networks". In: *2019 11th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*. Vol. 11. 2019, pp. 1–4. DOI: 10.1109/ATEE.2019.8724938.

[12] *tflite_flutter 0.8.0*. URL: https://pub.dev/packages/tflite_flutter.

[13] *tflite_flutter_helper*. URL: https://github.com/am15h/tflite_flutter_helper.

[14] Philipp Tschandl et al. "Expert-level diagnosis of non-pigmented skin cancer by combined convolutional neural networks". In: *JAMA dermatology* 155.1 (2019), pp. 58–65.

[15] Yifan Yang. "Data Augmentation to Improve the diagnosis of Melanoma using Convolutional Neural Networks". In: *Proceedings of the 2021 International Conference on Bioinformatics and Intelligent Computing*. 2021, pp. 151–158.

APPENDIX

Figure 3. The home page, from here users can either scan or classify an image.
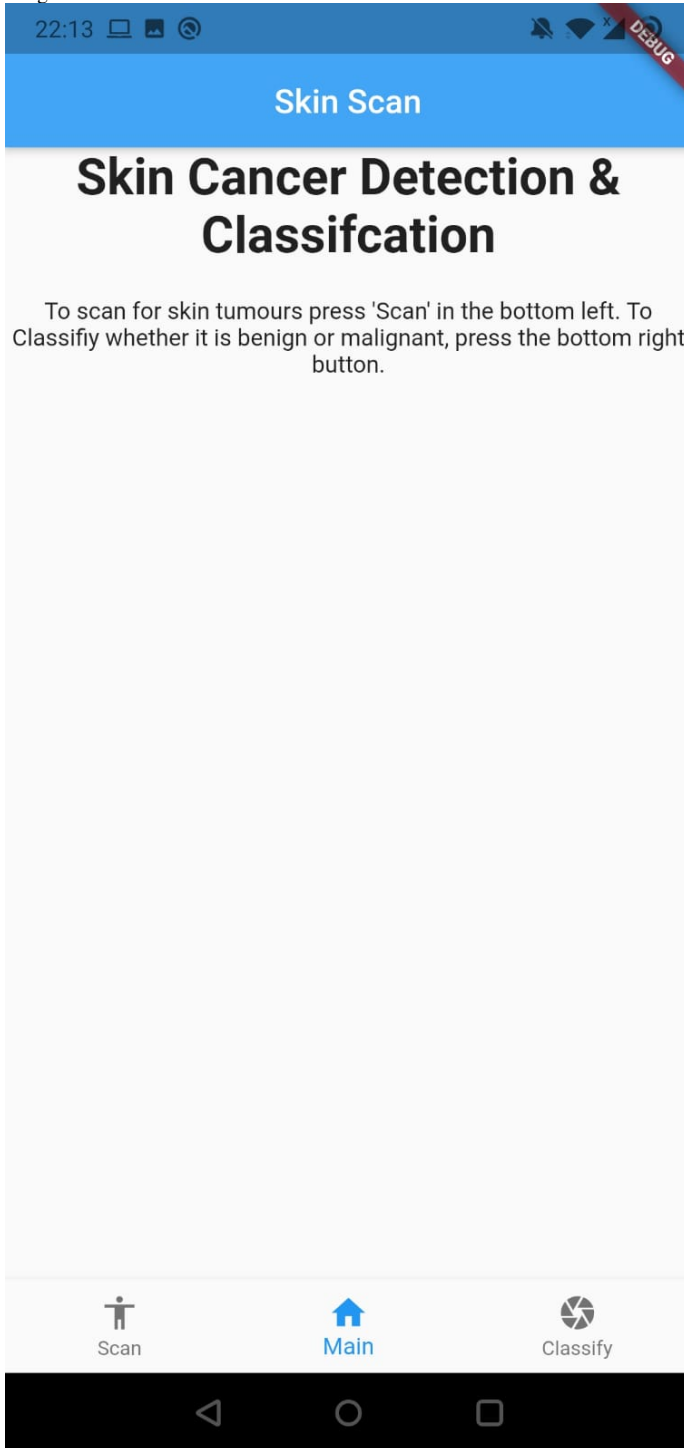

Figure 4. The image classification page without an image processed. The user can take a picture or select one from the gallery.
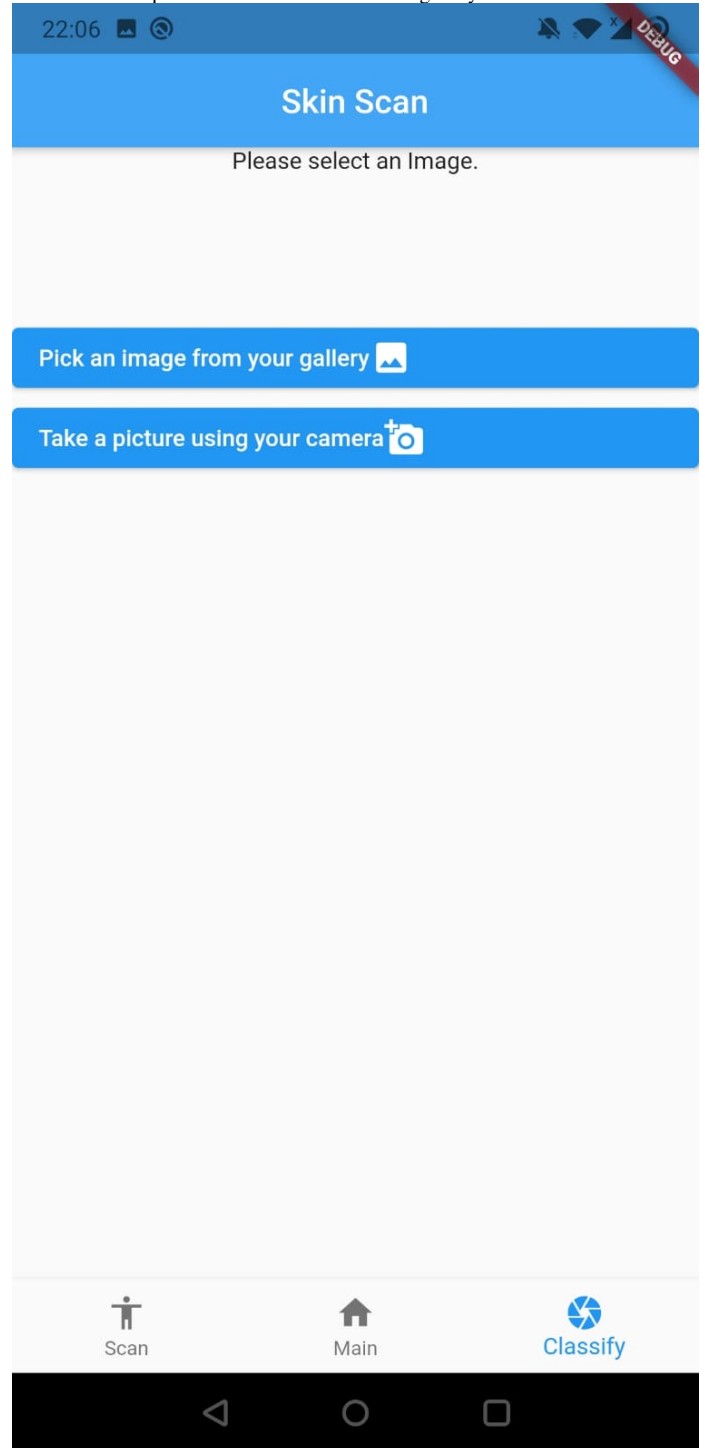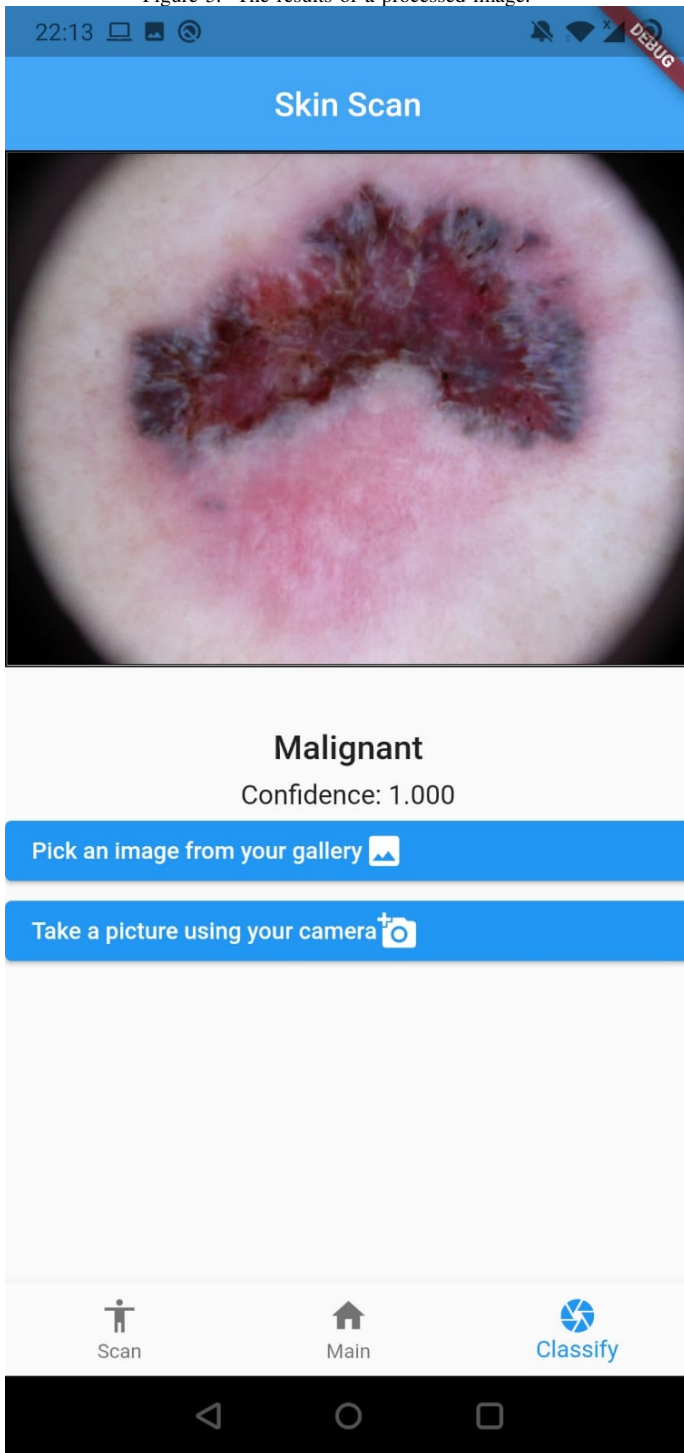
Figure 5. The results of a processed image.

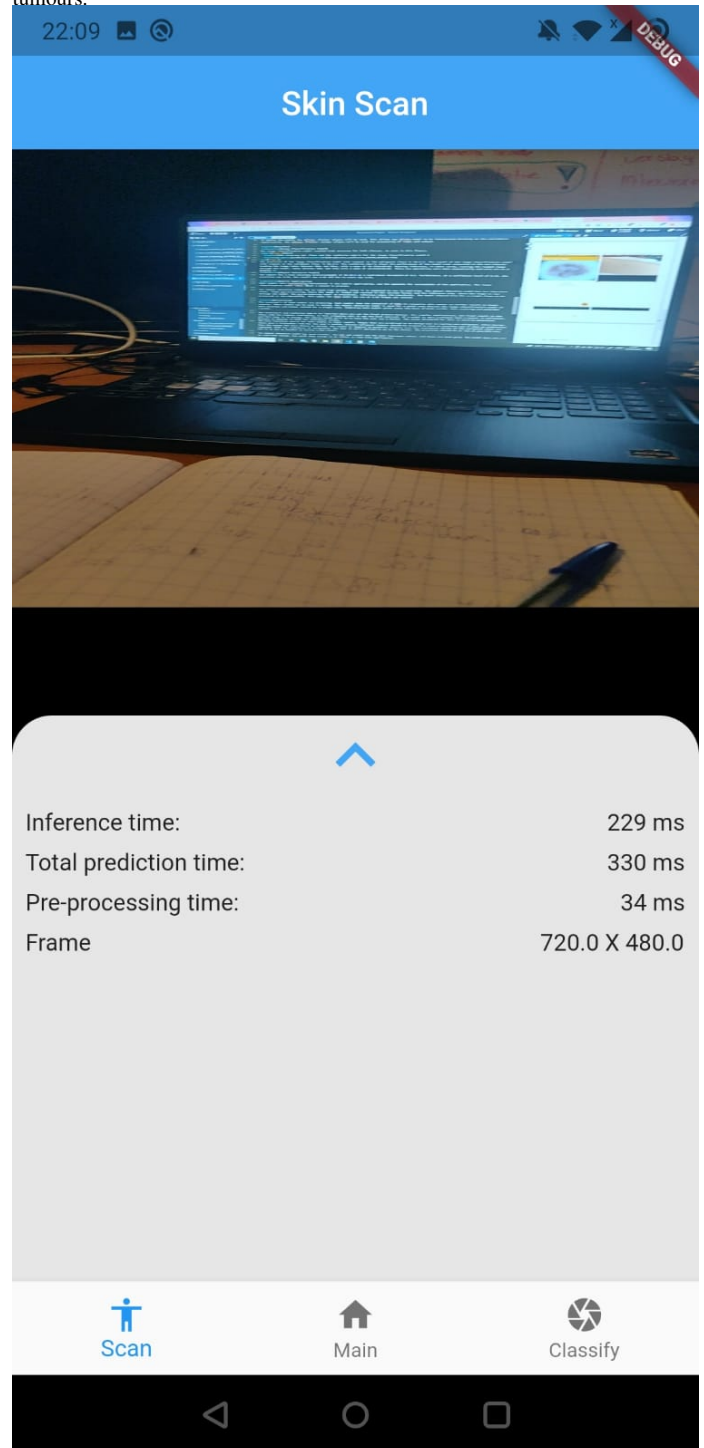Figure 6. The object detection page. The user uses this page to scan for skin tumours.

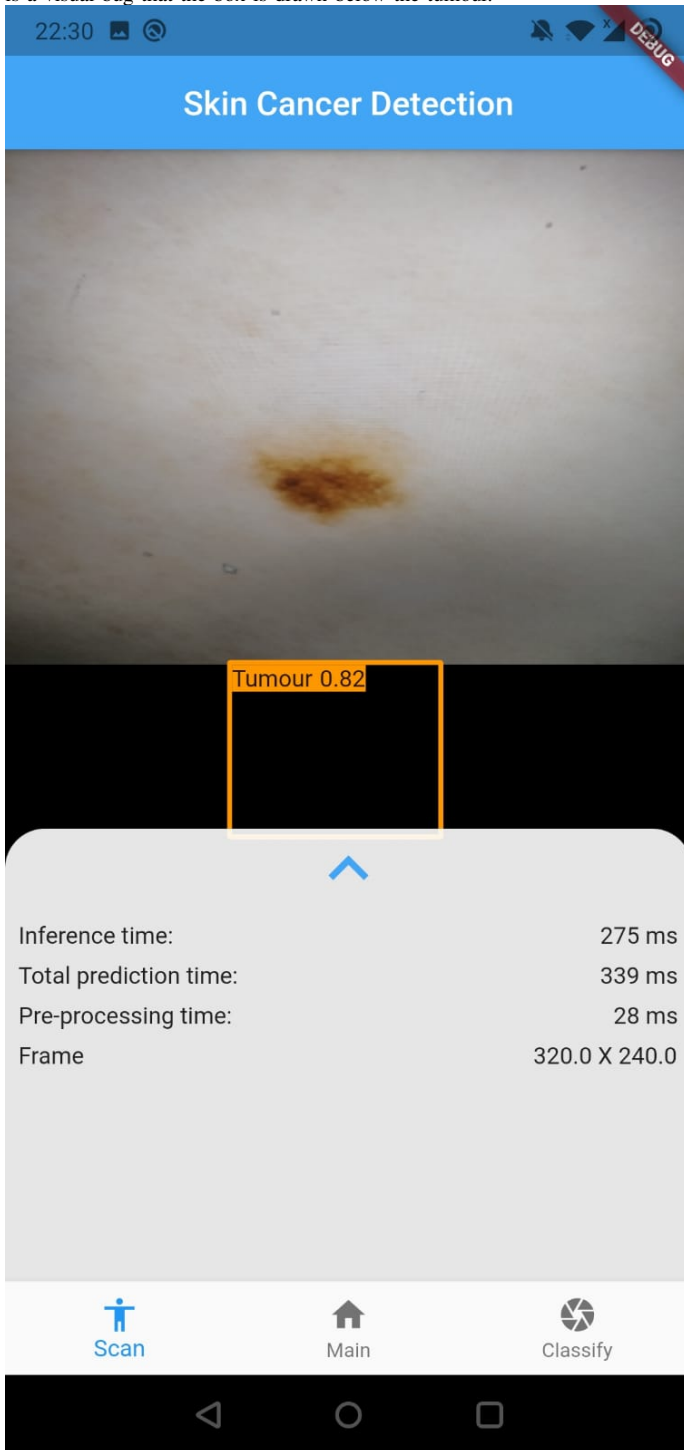Figure 7. The bounding box drawn by the object detection model. Note: there is a visual bug that the box is drawn below the tumour.

Figure 8. Training loss and accuracy