SkînScreen™

Background &Technical Accomplishments Paper

July 17, 2019

Table of Contents

SkinScreen™

Executive Summary

Artificial Intelligence methods, computational power, and massive datasets continue to expand and evolve. As a result, an opportunity exists within the medical community to leverage these technologies to improve the detection, classification and decision analysis of skin lesions. Currently, humans are unable to leverage the massive amounts of skin image data that are available to assist in classification of skin lesions without the assistance of technology. Traditional decision-making methods for skin lesion classification have not advanced with the advent of these new technologies. Current diagnostic classification methods are subjective and vary by individuals, even though a common training curriculum exists and is studied by well-educated doctors and students of the medical profession.

In addition, humans have a finite ability to detect very slight variations within skin site imagery. This is due to the limitations of human vision. Neural networks have the ability through statistical curve fitting and other mathematical techniques to detect variations that are beyond human detection capabilities. Automation through machine learning algorithms allows for more responsive detection and correct classification of visual data. SkinScreen brings precision and accuracy augmentation to this problem.

SkinScreen™ overcomes many of the limitations of human capabilities and provides a more precise and accurate prediction of classification via Artificial Intelligence (AI). SkinScreen assists the user (Medical Professional) to make a "data supported decision" based on statistical analysis of the most modern methods available. Resistance to this disruptive technology within the profession may occur when introduced, but the benefits derived far outweigh any cost in overcoming change and resistance in the utilization of this tool.

SkinScreen has been designed to automate tedious repetitive tasks, freeing time for complex human decision-making. This is performed by improving the workflow efficiency of dermatologists by prioritizing and accelerating engagements based on the severity of the SkinScreen classification predictions. When an image data is loaded into SkinScreen, the tool is able to apply a multi-faceted approach by automatically detecting and identifying nine of the most common the skin lesion classes or provide a "non sufficient data available" decision. SkinScreen allows for faster outcomes over manual, heuristic methods used today such as the common ABCDE approach taught in medical schools. Confidence measures, reported to the users, back up machine learning results. This tool revolutionizes the processing of visual data into actionable information that surpasses human capability and current AI-powered skin lesion detection solutions in accuracy, precision, speed, and breadth.

I. Current Market Solutions

I.1 Heuristic (Rules-based) Solution

Since 1985, doctors have been using a heuristic, preoperative based approach known as ABCDE (Asymmetry, Border, Color, Diameter, Evolving) in the detection of skin lesions. This manual approach has proven useful in the detection and classification of skin lesions however, it can result in human error even after being trained on the ABCDE approach. This was clinically proven by June K Robinson, MD and Rob Turrisi, PhD and the results are displayed in Table 1. *(Skills Training to Learn Discrimination of ABCDE Criteria by Those at Risk of Developing Melanoma", April 2006)*

Table I. Recognition of ABCD Criteria Before and After Skills Training

Table 1. Recognition of ABCD Criteria on a Transparency of a Pigmented Lesion Before and After Skills Training: Decision to Seek Medical Care by 100 Subjects*

	Before Skills Training		After Skills Training		After 20-min Break	
Feature	Correct	Undecided	Correct	Undecided	Correct	Undecided
Asymmetry	8 (4)	45 (8)	11 (5)	54 (8)	10 (3)	47 (10)
Border irregularity†	63 (12)	9 (4)	93 (3)	4 (2)	93 (4)	6 (2)
Color variation†	55 (15)	8 (2)	91 (7)	6 (3)	89 (2)	6 (2)
Diameter	37 (11)	43 (11)	94 (2)	1 (1)	92 (5)	3 (1)
Decision to see physician	35 (8)	0	88 (10)	0	85 (6)	0

*Data are given as mean (SD) percentages.
†Recognition improved with skills training with scoring; $P<.05$, paired t test.

Their clinical study yielded when an individual performs a skin self exam (SSE) the level of accuracy after being trained on the ABCDE approach varies based on the sex and age of the individual. ***As a result, this technique is subjective in its approach, inaccurate at times, and results vary by individual.***

I.2 Current AI-Powered Solutions

With the release of public datasets on skin lesions, such as the HAM10000, an opportunity exists to develop AI-powered applications to assist in the detection of skin lesions. However, the current solutions lack privacy, accuracy, precision, and breadth. Current systems also require a greater level of human interaction then is necessary. These systems are in their infancy and can be easily fooled which provides a false faith in their overall technological capabilities. For example, you can submit a picture of a 'giraffe' and some solutions will classify the giraffe as having a skin lesion when this is not true. Other solutions require users to upload personal pictures of the skin lesion to their servers in order to perform the detection of the skin lesion. With this type of black-box approach the user is unsure of the safeguards in place and whether privacy and HIPAA regulations are being adhered too in the storage and processing of that image. These shortcomings are not present within SkinScreen.

2. Key Design Elements of SkinScreen

In order to provide value to the end user as well as society in general, several key elements were adopted into the SkinScreen application during the design phase:

1. **User privacy –** All images are to be processed on the user's device and no images are uploaded to SkinScreen servers. This is performed with the MobileNetV2 architecture and Tensorflow.js libraries.
2. **Skin lesion presence –** Prior to the classification, the application will determine whether a skin lesion can be detected in the image.
3. **Detection of 9+ classes of skin lesions –** Detect more classes of benign and malignant skin lesions and provide better feedback for each user that what is currently available. Continue to update the models based on new, verified raw images.
4. **Higher accuracy and precision rates than current solutions –** Through the use of 180,000 images, the SkinScreen application is able to account for permutations in different skin lesion classes.
5. **Real-time feedback** – Provide back results in under two seconds since all processing is performed locally on the user's device and involves no network latency.
6. **User-friendly tools** – Provide useful support tools regardless of the user's background and skillsets. Provide for multi-linguistic capabilities within the application.
7. **Multi-platform use** – Include support for a web application, mobile devices, and API REST service calls by third-party consumers.
8. **Expand upon academic and industry research to maximize performance** – This was performed through the utilization of the latest versions of Tensorflow and Keras libraries and with supervised machine learning techniques in the development of the SkinScreen models.

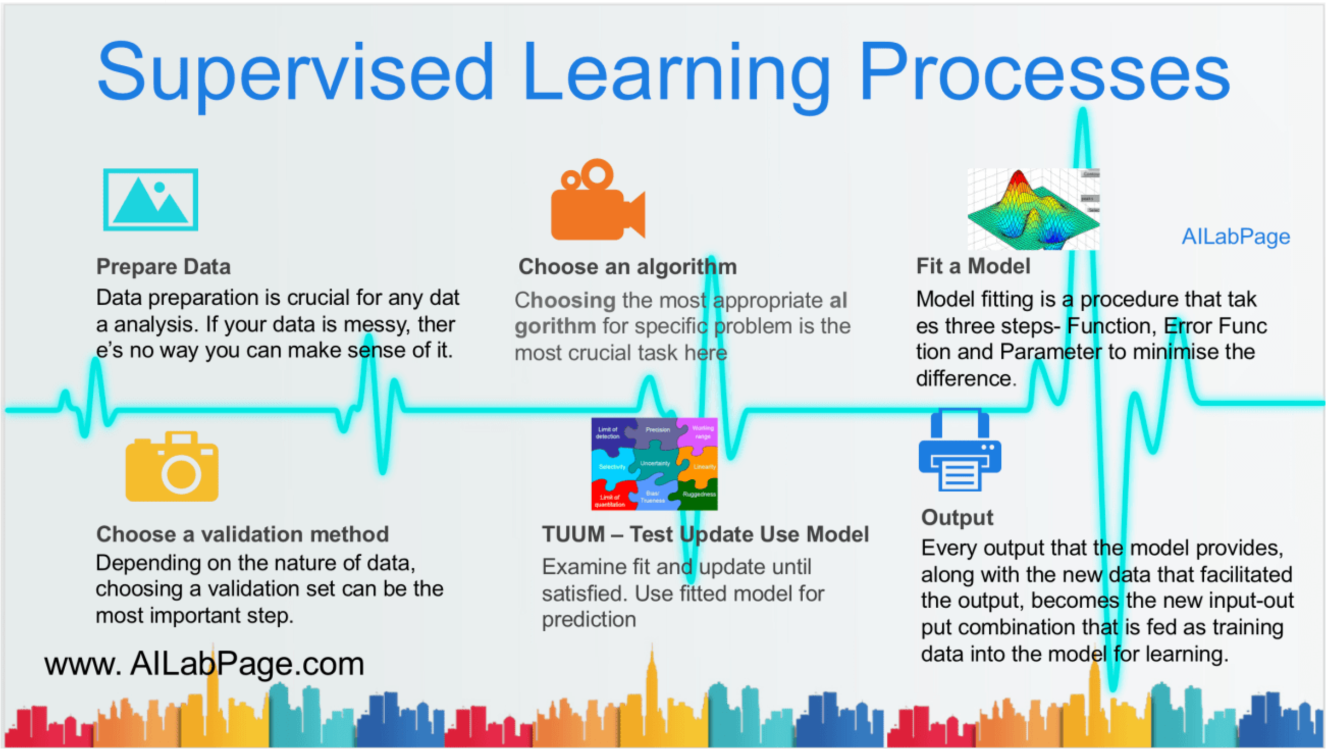
3. SkinScreen Feature Extraction

Our approach to "feature extraction" focuses on efficient computation of the right features for the specific neural network-determined class in order to provide high precision and accuracy. Image Identification of unique data often requires additional processing and transformations to produce distinguishing features. Many significant features exist to separate one image type from another. This computation can be costly and time consuming; a brute force comparison across all possible features is not feasible in most cases.

4. Leverage Supervised Machine Learning

During the development of SkinScreen it was decided that a supervised learning training approach will be prudent in the development of the models.

Figure 1. Supervised Learning Processes



Supervised learning, involves defining both input and output for the system in our kernels, and training the machine learning algorithm to learn the relationship. Given a new input, the trained system is able to perform this learned transform to predict the output. The training process requires many data examples with both signal observations (input) and the ground truth classification (output). These training examples are iteratively processed, and the system parameters are incrementally adjusted to better fit this input/output relationship. During this iterative process the machine learning model is monitored to ensure performance is consistent. Once this relationship is trained, a new input is tested, and the correct output calculated. These training, validation, and inference steps are typical for machine learning systems.

Modern machine learning approaches commonly use neural networks. Neural nets are so called due to the interconnected neuron elements used for calculations. This is where the similarities with the human brain end. Layers of neurons are computed as realization of matrix math. Matrix math is done very efficiently on modern processors and hardware accelerators. SkinScreen™ uses proven Convolutional Neural Network (CNN) technology. Our machine learning model builds on academic and industry research to maximize performance. We instantiate this model into software using the TensorFlow™, Keras, and Tensorflow.js open source libraries.

5.0 Testing

Once a neural network model has been trained with inline validation, it is important to continue to test performance against new examples. When a new input is presented to the system, a prediction is made. This computed inference should be reviewed by an expert to ensure correctness.

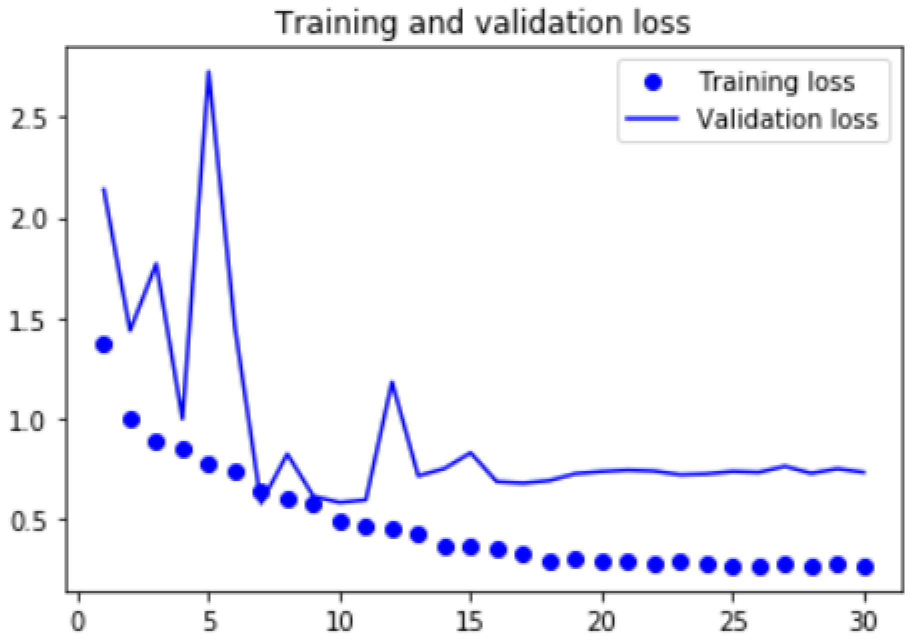
We continually test SkinScreen™ to ensure detection performance and classification accuracy. We ensure all images are being detected and isolated correctly, and the predicted machine learning classifications make sense.

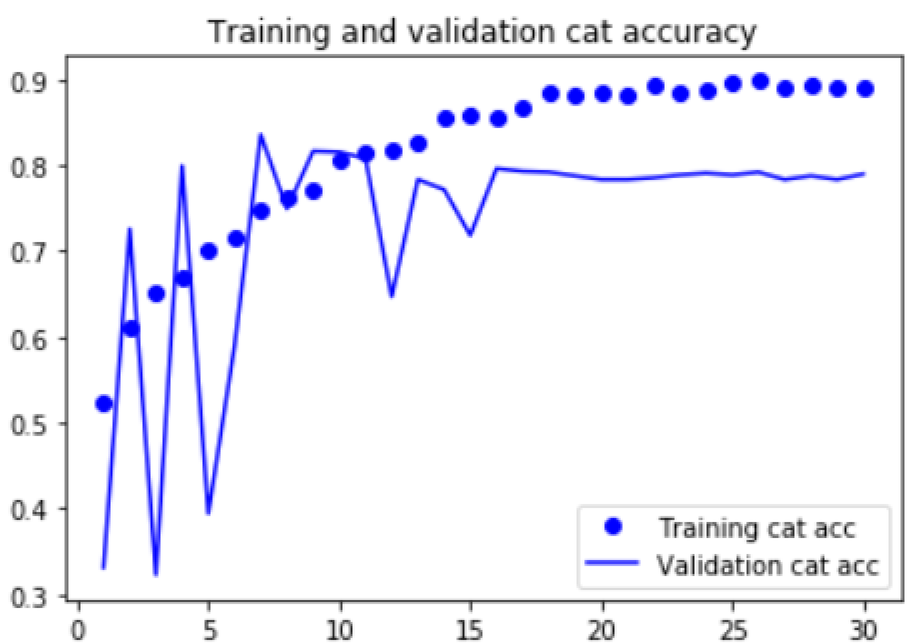
6.0 Validation

During training of a machine learning model, it is important to monitor incremental performance for a separate set of data. The validation set should be representative of real-world testable examples, but unique from the training data. This can be real-world data with known truth information for the image type. When a machine learning model trains on a set, it reduced the error when classifying only that training set. A similar level of performance on the withheld validation data must be verified. This is a key challenge of machine learning: Generalization from operation on a small set of data to new input examples. In many cases a machine learning model over specializes to the training data, not able to generalize to new data. Overfitting a model is a condition where the statistical model begins to describe the random error in the data rather than the relationships between the variables. This overfitting problem is overcome through various machine learning architecture and training parameter methods.

Once the models are trained, we then determine the accuracy of the model and whether it is an appropriate fit and not over-fitted. We then plotted the models against our validation data set to visualize an appropriate fit (see Figure 2) based on the loss and category accuracy metrics, generating a confusion matrix to determine the level of precision and recall, and then reviewed the F1 score for accuracy.

Figure 2. Loss and Category Accuracy Metrics for Non-Sufficient Data Classifier





7. Post-processing

If each of the model's results meet our threshold for success we then convert our model's topology and weights from the Keras format over to the Tensorflow.js

Layers format. This step is performed in order to allow the browser to consume the format and process the sharded weights through an optimized loading function. Once this process is complete, we then deploy our models to our operational environments. See Figure 3 and Figure 4 for the results of different images in an operational state via the SkinScreen web application.

Figure 3. Non Sufficient Data Classifier

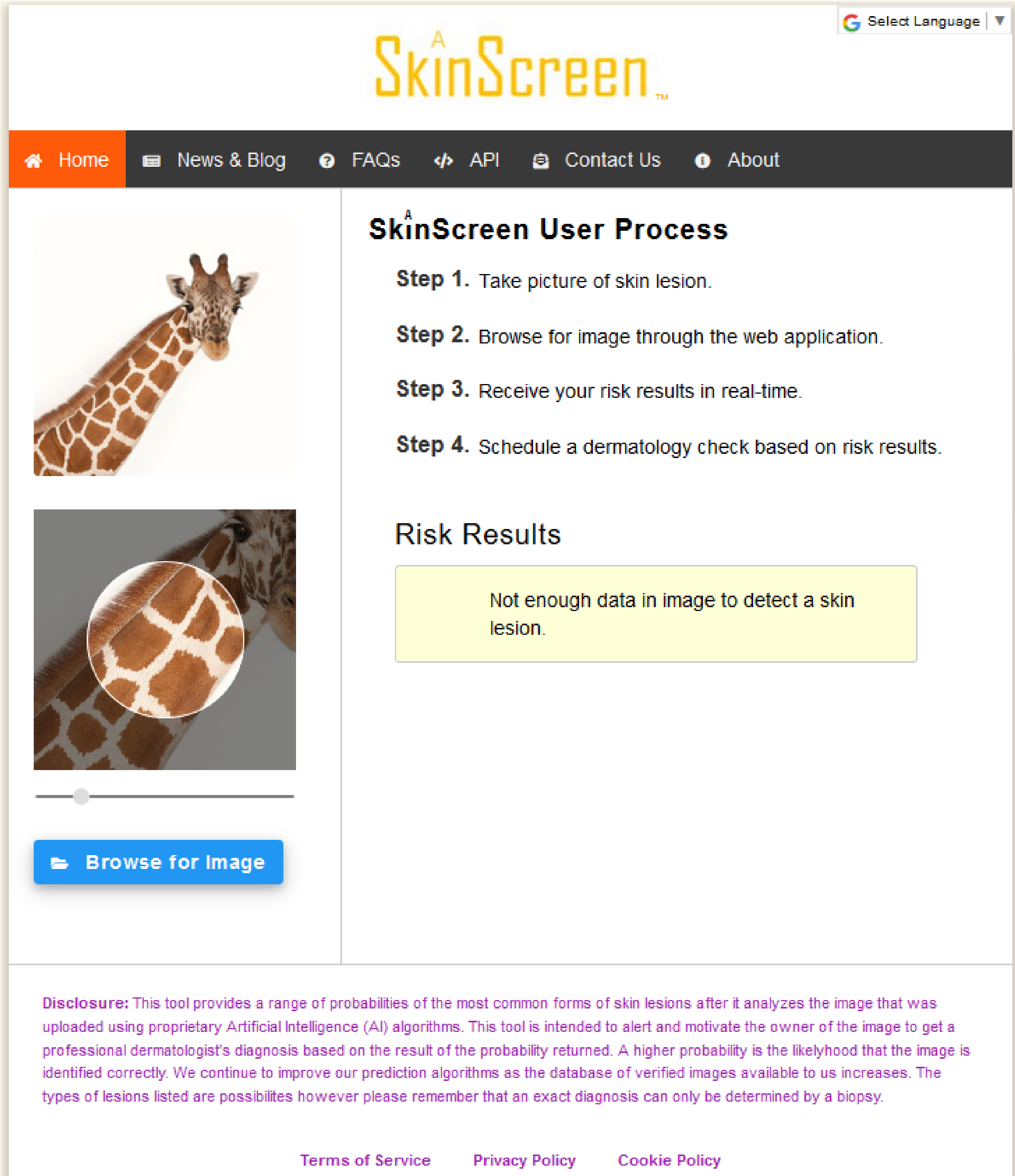
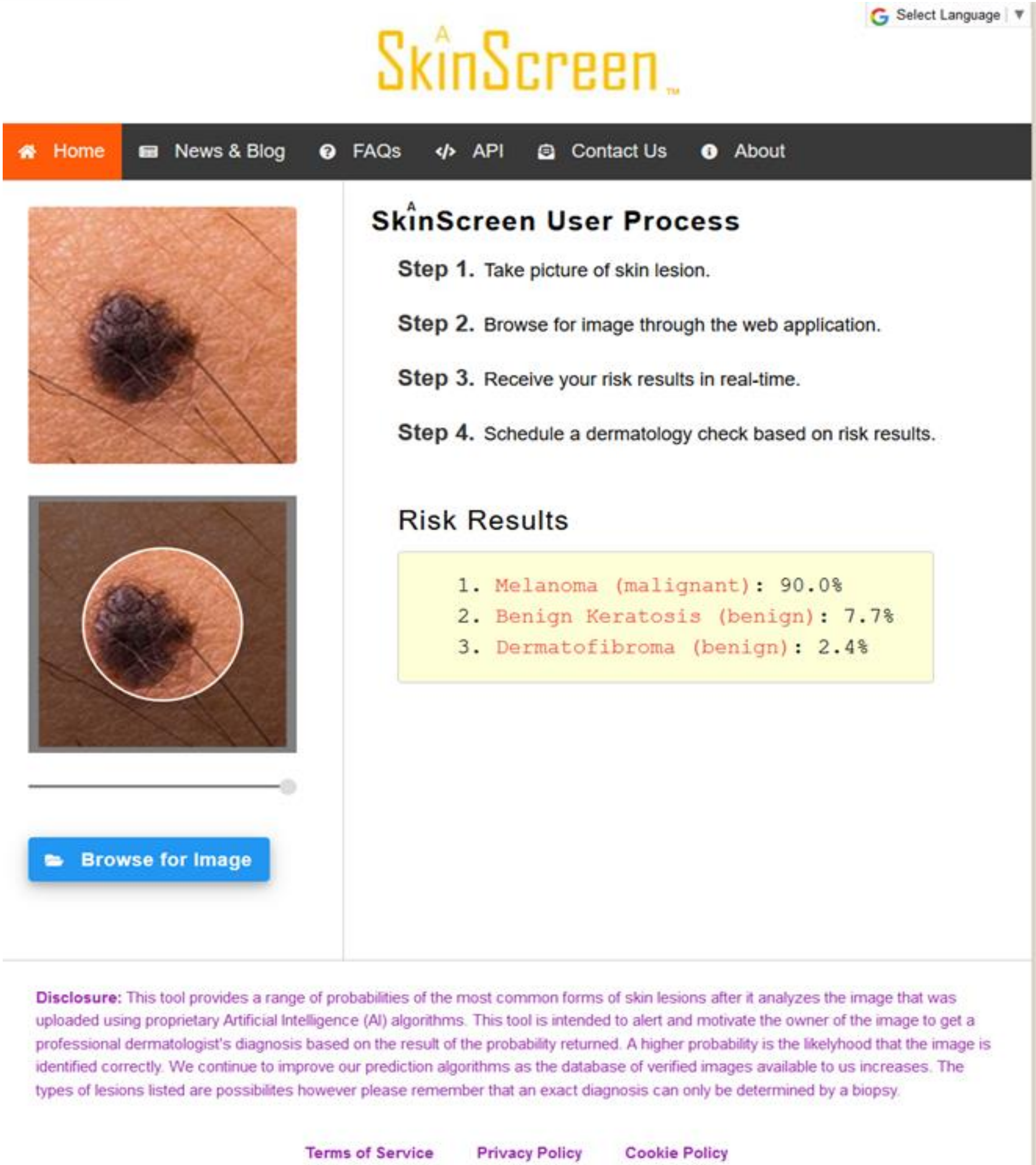


Figure 4. Skin Lesion Classifier



8. Summary

SkinScreen™ automates the detection, isolation, and classification of skin lesions using advanced machine learning techniques and libraries. It offers the capability to detect malignant and benign skin lesions in real-time through a highly accurate and precise solution that was generated through supervised machine learning techniques. The solution leverages the power of deep learning, a method under artificial intelligence, to allow for quicker and more accurate predictions than previously were available through manual, heuristic approaches or other AI-powered solutions.

References

AILabPage.com, “Supervised Learning Processes”, July 2019
https://ailabpage.com

Robinson, June K and Rob Turrisi, “Skills Training to Learn Discrimination of ABCDE Criteria by Those at Risk of Developing Melanoma”, April 2006.
https://jamanetwork.com/journals/jamadermatology/fullarticle/404521