

Intelligent Mirror: Detecting Skin Cancer (Melanoma) using Convolutional Neural Network with Augmented Reality Feedback

Pratap Salunkhe
Cybercity, Tower IV,
Magarpatta City, Hadapsar,
Pune, Maharashtra 411013, India

Vishal Mehta
Cybercity, Tower IV,
Magarpatta City, Hadapsar,
Pune, Maharashtra 411013, India

ABSTRACT

Human: Is that mole developing on the face leads to some risk towards the skin?

Mirror: Yes.

Human: I don't know if I was not aware of the fact until now.

Mirror: Shall I go ahead and fixup an appointment with the dermatologist on this weekend: Saturday 11am?

Mirror: Sounds Good!

Please check your email the appointment confirmation.

In this paper we propose a system (a health mirror) an application of internet of health (IOH) which monitors the sun exposed parts of the human body, pulls the attention of the user if it detects something and recommends to see the doctor and also has the capability to classify the type of skin cancer (if detected) using the current state-of-art convolutional neural network in deep learning.

General Terms

CNN

Keywords

Convolutional Neural Network - CNN

1. INTRODUCTION

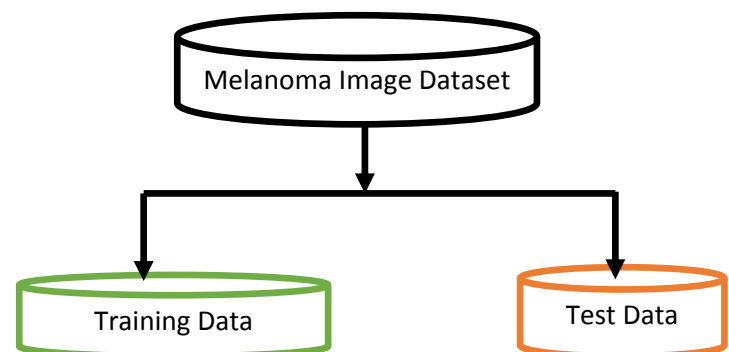
In this system we have a regular mirror embedded with usb camera connected to the embedded board. The system is deployed with the learned neural network models using which it's capable of classify the given static image taken of the user by the camera as one of the class of the skin cancer. The user stands in front of the mirror. The camera takes the picture of the subject (user) and after analyzing the image by comparing through its learned model it classifies the image as one of the cancer classes or most probably the skin infection (if abnormality detected). Convolutional Neural Networks today has become the state of art in computer vision and natural language processing applications because of its accuracy. Be it Apple Siri personal assistant or Microsoft Cortana or recently released Google Allo a chat bot which is an Artificial Intelligent based chat bot which has the capability to answer any of the users question. The real power comes into the play when Convolutional Neural Networks (CNN) are interfaced to unstructured data like videos and images.

A user stands in front of the mirror. The web cam set up on the top center of mirror gets ON and by detecting the face and starts scanning the skin if it's able to identify anything abnormal on the skin (sometimes not visible from the naked eye) Once the diagnose gets completed it will augmented the

summary report on the mirror using the augmented reality user interface and will ask for setting up an appointment with the nearby dermatologist for the further investigation. The model is capable of identifying the lowest critical skin infection to any of the melanoma class cancer.

Convolutional Neural Network (CNN) is a form of neural network which has the capability to automatically extract the features from the raw data (image, speech signals) reduce the dimensionality of the image data and then finally vote for the confidence score to make a prediction on the content to be classified in the image data. There are various CNN architectures like AlexNet[1], GoogleNet[2] which have proven very powerful for the classification of the image content.

Here in the given problem in hand we have collected 1000's of raw images of different skin cancer categories (one per category per folder). And made our model learn the various features of all the different categories on the training dataset on the Amazon cloud AMI with multi GPU support. The dataset is divided into training and test set. Once the model is trained on the training set we need to test the accuracy of the model on the test set of the image data.



2. SYSTEM DESIGN

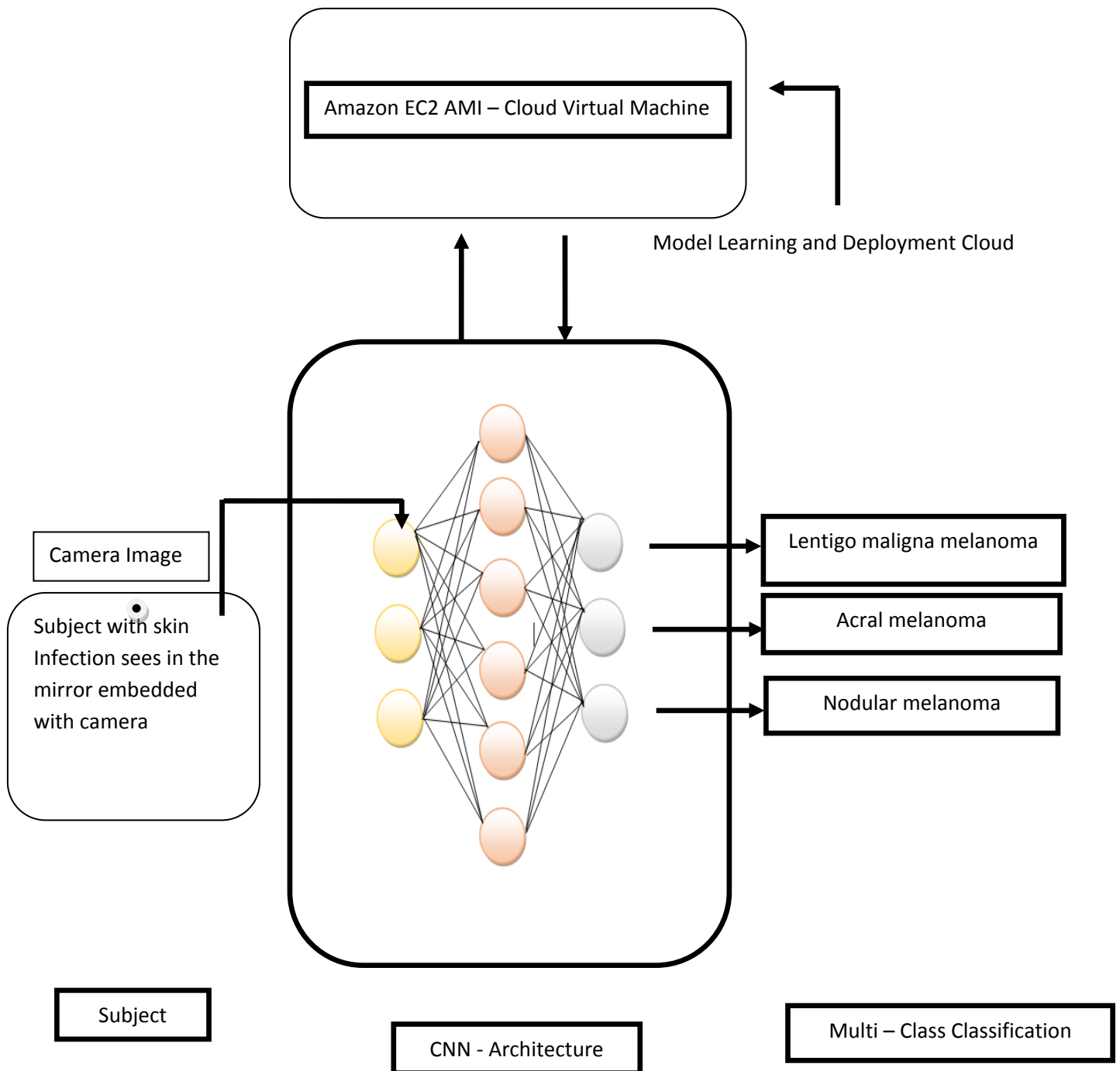


Fig 1

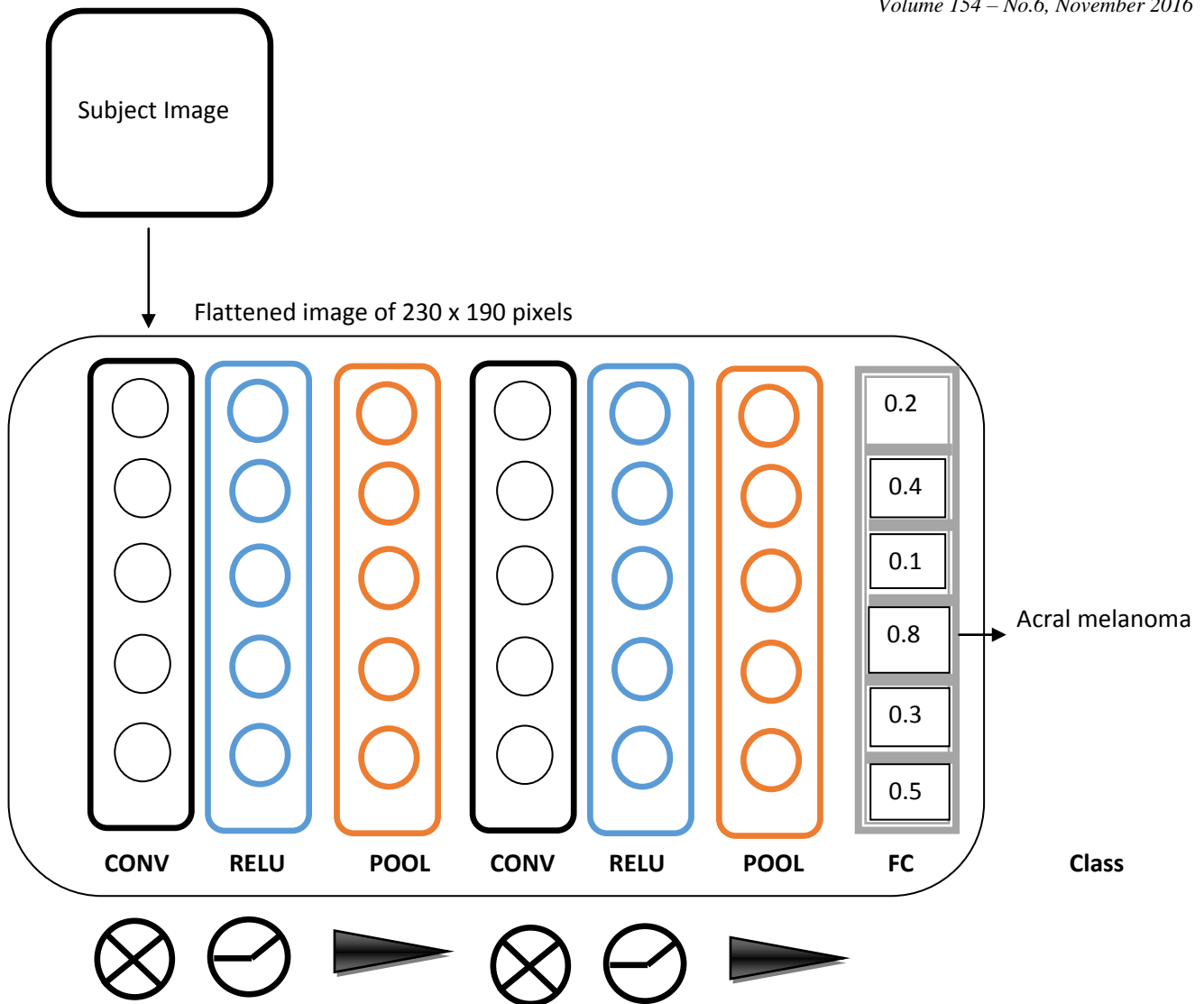


Fig 2

The system takes in the batch of images to train the CNN from the training set. The images are the RGB images with height, width and depth dimensions (number of channels, for RGB image is 3). Here the system is input by image of $230 \times 190 \times 3 = 131100$ pixels. The image is flattened to 131100 features image which gets inputted for the first convolutional layer. In this layer the majority of the heavy lifting work takes place of image feature extraction. A small patch of the image which is also called filter or kernel is placed on the top left most corner of the image and then the filter is multiplied by the corresponding intensity values of the image and a new feature map is created after the multiplication while shifting the filter from left to right by a constant stride (the number of pixels to be skipped and shifted to the left is controlled by the deep learning expert) of one pixel Fig 3. The number of feature maps outputted from the convolutional layer are equal to the number of filters in that layer and this process is called convolution.

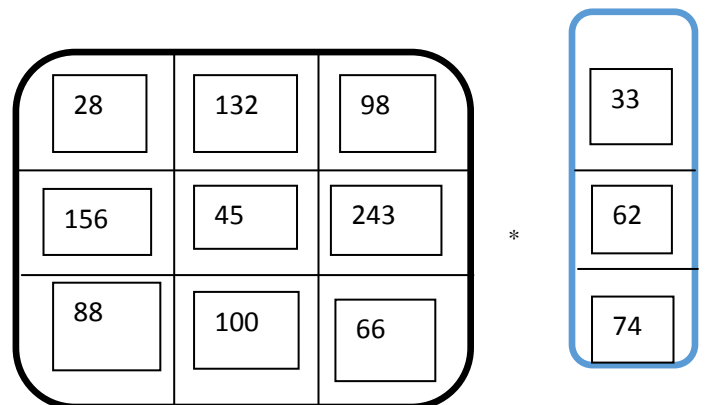


Fig 3

The second layer is the rectified linear unit nonlinear function also called as RELU which converts all the negative values to positive values retaining all the features in the feature map. Once all the feature map values are converted to positive. Then the output of RELU is fed into the next layer called POOL layer defines a 2×2 window and place this window on the left top corner onto the feature map. The pooling layer does the max pool i.e. it pulls out the max value inside the 2×2

matrix and then slides to the next 2x2 matrix onto the right. This helps in two way, first it does the dimensionality reduction of the images second it retains all the features as in the feature map by a single representing number that's the max in the matrix.

And finally the Fully Connected layer FC where everything gets converted into set of votes. The class which has highest number of votes with the highest probability score is classified for that particular class as one of the skin cancer.

3. CONCLUSION

In this paper we have proposed an application of Internet of things and Artificial Intelligence focused on health care domain. In the future work of this is to start collecting the images and annotations and create a training and test dataset with the help of dermatologist on the Amazon cloud. Following which start experimenting with the existing CNN architectures for their accuracy and train the CNN from the scratch for better results. Also once the model is ready will also work upon android application.

4. REFERENCES

- [1] ImageNet Classification with Deep Convolutional Neural Networks Geoffrey E. Hinton University of Toronto Hinton@cs.utoronto.ca, Alex Krizhevsky University of Toronto Kriz@cs.utoronto.ca Ilya Sutskever University of Toronto Ilya@cs.utoronto.ca.
- [2] Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Title: Going Deeper with Convolutions." September 17, 2014. Accessed November 11, 2016. <https://arxiv.org/abs/1409.4842>.
- [3] Karpathy, Andrej, George Toderici, Sanketh Shetty, Thomas Leung, Rahul Sukthankar, and Li Fei-Fei. "Large-Scale Video Classification with Convolutional Neural Networks." 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014. doi:10.1109/cvpr.2014.223..
- [4] Ranft, Benjamin, and Christoph Stiller. "The Role of Machine Vision for Intelligent Vehicles." IEEE Transactions on Intelligent Vehicles 1, no. 1 (2016): 8-19. doi:10.1109/tiv.2016.2551553.
- [5] DenseCap: Fully Convolutional Localization Networks for Dense Justin Johnson*, Andrej Karpathy*, Li Fei-Fei Stanford University
- [6] Deep Fragment Embeddings for Bidirectional Image-Sentence Mapping, Andrej Karpathy, Armand Joulin, Li Fei-Fei. Stanford University
- [7] ImageNet Large Scale Visual Recognition Challenge, Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, Li Fei-Fei. Stanfprd University.