Academic Literature Review

- What academic work is relevant to your project topic? Pick 3 papers, ask us for help if you need it.
    - Methodology for diagnosing of skin cancer on images of dermatologic spots by spectral analysis. [https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4605047]
    - The melanoma skin cancer detection and classification using support vector machine. [https://ieeexplore.ieee.org/abstract/document/8257738]
    - Computer-assisted diagnosis techniques (dermoscopy and spectroscopy-based) for diagnosing skin cancer in adults (Review) [https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6517147/pdf/CD013186.pdf]

- What makes these papers important/relevant?
    - The first paper states the types of skin cancers that someone may come across. It also includes the severity of obtaining skin cancer whether it's the most common and least dangerous to the least common and most dangerous. This paper also states that in the early stages of each type of skin cancer is almost always curable. This goes along with our plan to catch it in the early stages to refrain from skin cancer being deadly.
    - The second research paper involves explanation of how image processing has been proposed for identification of melanoma skin cancer and its increasing use as a diagnostic tool for classifying the medical images with accurate interpretation. The contribution and work describe the efficiency of methods used for classifying the lesions using various segmentation, feature extraction and classification techniques. This helps in identifying how the performance is sensitive to selecting particular features for classification.
    - The work and contributions of Ruffano et al. (of paper 3) are key elements for understanding the place of digital image processing in the medical sciences, specifically in classifying and diagnosing surface distortions of the foremost layers of skin. It is a systematic review that covers a large amount of work in the field. The study introduces two methods of skin cancer detection, dermoscopy and spectroscopy using the CAD system and assesses their effectiveness in classifying and determining the characteristics of skin cancers of various forms. It contrasts the two

methods and provides insights on the shortcoming and benefits of using either methodology.

- What are their results and how did they achieve these results?
    - Paper 1 - The results were with the methodology applied to the dermatologic images which were 332 provided by a medical specialist, 260 images were calculated as skin cancer with 72 being benign lesions. The classic, inverse and k-law nonlinear filter using different color transformations was applied to each image to get the results.
    - Paper 2 - An Image database with two categories (Normal, Melanoma) was generated by gathering images from multiple sites. The proposed algorithm will enhance the contrast of the image and reduce the noise which helps for more accurate segmentation and classification. Later, segmentation is done to retrieve the lesion region(ROI) and extracts the color, border irregularity, circulation, Energy, correlation, homogeneity and contrast, entropy, skewness, kurtosis features. Principal Component Analysis (PCA) is used to identify the most efficient features from the extracted features as some of them may be ineffective on accuracy. These features are fed in the Support Vector Machine (SVM) algorithm to classify the image into binary classes as benign and malignant. The proposed method works efficiently on the dermoscopy images with 92.1% accuracy. Same accuracy is achieved when the set of the features selected by PCA or the entire set of 11 features are used, but with lower computational complexity.
    - Paper 3 - The study reviewed various datasets and studies utilizing Computer Assisted Diagnostics (CAD) systems and their efficacies in several experiments. New discoveries surrounding the field of discourse for CAD research were made, however the initial research scope was not met. Some of the reviewed studies were low quality, and involved poor reporting related to highly selected participant groups. Nevertheless, several things were determined during the review process. CAD systems can identify melanoma within selected groups. The specificity of the groups selected means that though this is a sure way to find melanoma in selected cohorts, it may not be a viable standalone option I the field. The results were not indicative of primary-care settings as they were selected for controlled environments. In the environments CAD was utilized, it was just as sensitive as the manual determinations made by industry professionals. In some instances, CAD was up to as much as 18% more

effective. There is still too little evidence in the base of investigations to recommend for clinical practice.

- What's different/unique about these approaches?

  These papers provide analysis on different algorithms implemented for images captured using various technologies. Although these papers use similar skin cancer detection pipeline for classification, they undergo different approaches in each stage for segmentation, feature extraction or classification. These approaches help gather information about detection of skin cancer using different algorithms using image processing to achieve results. There are many algorithms and image processing techniques to achieve this but the idea itself is different and like no other.

Open Source

- What open-source code is available that are relevant to your topic?

  The following link can be used to find a wide range of repositories which are available for Classification of Skin cancer detection: https://github.com/topics/melanoma-detection

- What data is available for testing and/or training algorithms?

  A large collection of multi-sources dermatoscopic images is available which can be used for training or testing the algorithm: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DBW86T

- Is labeled data available? How much? How is the data licensed? Is it under copyright protection?

  Dataset contains images with malignant multiple skin cancer labels namely Actinic keratoses and intraepithelial carcinoma / Bowen's disease (akiec), basal cell carcinoma (bcc), benign keratosis-like lesions (solar lentigines / seborrheic

keratoses and lichen-planus like keratoses(bkl), dermatofibroma (df), melanoma (mel), melanocytic nevi (nv).

The dataset consists of 10015 dermatoscopic images which can serve as a training set for classification of skin cancer using machine learning.

When using the ISIC 2018 dataset in research, following 2 works are to be mentioned in it.

[1] Noel Codella, Veronica Rotemberg, Philipp Tschandl, M. Emre Celebi, Stephen Dusza, David Gutman, Brian Helba, Aadi Kalloo, Konstantinos Liopyris, Michael Marchetti, Harald Kittler, Allan Halpern: “Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC)”, 2018; https://arxiv.org/abs/1902.03368

[2] Tschandl, P., Rosendahl, C. & Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci. Data* **5**, 180161 doi:10.1038/sdata.2018.161 (2018).

Other Questions:

- What problem will your Computer Vision solution solve, and for whom?

  Our computer vision model will solve problems that are related to detecting skin cancer. Our model would be able to classify between different skin cancers. By making such classification, we can give people the information required to make

decisions about the urgency of them seeing a dermatologist. It can be used by people to screen if an unusual mole or a birthmark is prone to any type of cancer.

- Has anyone reverse engineered these products?

  As far as we know it has never been reverse engineered as doing that would require an immense amount of research and funding. Scientists and doctors would have to work together to make that happen in real life.

- If you have access to the product, what can you learn from using it?

  If we have access to the product we can learn the difference between different types of skin cancers and stop them from spreading by eradicating the problem by detecting them at a very early stage. This would save a lot of lives every year.

- It can be tough to tell exactly how proprietary solutions work, but what can you find on the internet?

  I think the chance of having this solution work is very high as there are already multiple applications developed to detect skin cancer types which allow the patient to self exam and store the data in a photographic library which can be then analyzed to detect moles/lesions and can suggest a potential solution.

- How much would a customer be willing to pay for your product?

  A normal skin cancer screening costs anywhere around $300 to $500 so by using this app people can save that money. This type of application can go anywhere from $5 to $15 on play store or app store. When it comes to saving money people won't mind saving more money by spending some.

- What companies are solving similar problems to yours?

Apps like UMSkinCheck and SkinVision in the UK are solving similar problems to ours. These two apps kinda do a similar job to our project. They use computer vision to detect skin cancer by using images from users.

Reference

1. Ferrante di Ruffano L, Takwoingi Y, Dinnes J, Chuchu N, Bayliss SE, Davenport C, Matin RN, Godfrey K, O'sullivan C, Gulati A, Chan SA, Durack A, O'connell S, Gardiner MD, Bamber J, Deeks JJ, Williams HC. Computer-assisted diagnosis techniques (dermoscopy and spectroscopy-based) for diagnosing skin cancer in adults. Cochrane Database Syst Rev 2018;2018(12). PMID:30521691