Pilot Study on Dermatological Disease Classification using Transfer Learning

Muhammed Sarim[a], Hunaid Hameed[b]

[a]*Federal Urdu University of Arts, Science and Technology, Karachi, Pakistan*
[b]*Institute of Business Administration Karachi*

Abstract

Objective
Pakistan faces a shortfall of doctors because of a low literacy rate and an overwhelming inclination of doctors to move to foreign countries. Skin diseases, especially those prevalent in the under developed world run rampant. This causes a major lack of health care, especially in rural areas. To address this problem an expert system based on Deep Learning is proposed to identify skin diseases.

Materials and Methods
We collected a dataset of images of skin diseases at Civil Hospital, Karachi, Pakistan by taking pictures of incoming patients with their consent. The diseases were identified by senior doctors. We used the 18 layer Residual Network, designed by Microsoft Research, to prepare a decision support system for skin disease identification.

Results
An accuracy of 86% was achieved on 10 skin diseases. A comparison of the model with human doctors on a subset of the validation set found the best performing doctor and the model to have a 74% accuracy.

Discussion
The model uses a deep neural network originally created for large scale image recognition and reused for disease identification. The accuracy is also comparable, however this work has used significantly less data.

Conclusion
This work has provided a proof of concept that shortage of skilled professionals in medicine in third world countries can be addressed with artificial intelligence.

Keywords: Deep Learning, ResNet, Computer Vision, Dermatology, Skin Diseases Classification

---

Email addresses: `msarim@fuuast.edu.pk` (Muhammed Sarim), `hhameed@khi.iba.edu.pk` (Hunaid Hameed)

1. Introduction

Pakistan is a third world country with low literacy rates and rapidly increasing population. Medical professionals are in small amounts and the population is rising rapidly. Doctors in public hospitals are over worked because majority of the people are unable to pay for premium healthcare and obtain healthcare from public hospitals. This leads to misdiagnosis which further escalates the problem.

In such circumstances, there was a need for computer aided diagnosis system. To address this need, we collected images of skin diseases at the Civil Hospital Karachi, which is a public hospital in the metropolis city of Pakistan, Karachi. The images were labelled by a senior dermatologist of 30 years of experience. The diagnosis was cross checked from another doctor and only images for which there was consensus on the diagnosis were included.

We performed feature extraction on a pre-trained 18 layer ResNet deep neural network and were able to obtain 86% accuracy on 10 unique skin diseases. The performance of the network was determined to be on par with the ability of human doctors. This was determined by testing the ability of the model and human doctors on a subset of the validation set.

2. Background

Skin problems are a common occurance among children and women in Pakistan. A study conducted by Khalida Naz Memon et al.[1] at an Out Patient Department of Dermatology in Pakistan observed that 82.5% of the children of age 10 years or less visiting the facility suffered from skin diseases, 45.5% of which is Scabies. 77.2% of the patients belong to rural or slum areas, 68.9% are of low socio-economic strata, and 82% are living in overcrowded families. A strong association between skin infections and water inadequacy and scabies and overcrowding was found.

Another study conducted by Ijaz Ahmed et al.[2] at the Hamdard Hospital, in the metropolis of Pakistan, found a similar pattern with variation. 1733 patients were enrolled. Infections were most common, next were Scabies (18%), Eczema (18%) and Acne (13%).

Mallon et al.[3] found that severely affected patients of Acne have social, psychological and emotional problems that were as severe as those reported by patients with chronic disabling asthma, epilepsy, diabetes, back pain or arthritis.

A survey[4] conducted among 275 recent medical graduates from Lahore found that only 14.2% intented to return to Pakistan after training from abroad. 10% never intended to return to Pakistan and 37% intended to stay temporarily. Two of the reasons were financial conditions of doctors and job opportunities.

Harnessing recent advances in Artificial Intelligence could alleviate the problems caused by rampant occurances of skin diseases and remove impact caused by departure of doctors from Pakistan. This will also alleviate the problem of over burdened doctors, especially in the public sector hospitals in Pakistan.

2.1. Description of Diseases

Scabies is a parasatic disease with 300 million occurances worldwide[5]. It is a major public health issue in developing countries and common in resource deprived communities.

In a Bangladeshi urban slum, 952 out of 1000 children had scabies. Appearance of scabies includes papules, vesicles, pustules, itching (especially at night), and a positive family history[6]. In children, scabies can be mistaken for pyoderma[7]. Secondary infections in Scabies are a common occurance[8].

Lichen Planus commonly occurs among middle aged and elderly with (30 to 70 years old). Occurance among children are rare. It occurs in dense band like lesions[9]. Clinically, Lichen Planus must be differentiated from Psoriasis[9].

Herpes Zoster causes a painful blistering rash and can lead to complications. There is 30% life time risk of contracting the disease. A population based cohort study over a 60 year period in Minnesota found a 4 fold increase in Herpes Zoster[10]. Another study reported a 50% chance of occurance of Herpes Zoster for people who live upto 85 years. Pain caused by Herpes Zoster may persist months or years after the resolution of lesions and cause anorexia, weight loss, fatigue, depression, withdrawal from social activities and employment, and loss of independent living[11].

Leishmaniasis is the ninth largest disease burden among individual infectious diseases[12]. Leishmaniasis affects 88 countries, 72 of which are developing countries. Majority of the cases are from South Asia and Africa. It is serious threat to 350 million people. There are 1.5 to 2 million cases and 70,000 deaths each year[13].

Psoriasis is chronic inflammatory skin disease[14]. The occurance of psoriasis is few units per 10,000 people-years[15].

Pemphigus Vulgaris is an autoimmune, blistering disease[16] and comprises of 70% of all occurances of Pemphigus[17]. A study found a highly significant increase in prevalance of autoimmune diseases among first degree family members of patients of Pemphigus Vulgaris compared with a control group[18].

Alopecia Areata is characterized by non-scarring patches on scalp and body hair. There is no cure. There is a 2% lifetime incidence worldwide[19]. It can cause feelings of vulnerability, nakedness and low self-esteem. it is often suggested to patients that their disease is caused by stress which causes feelings of guilt and responsibility[20]. A study found that the overall prevalence of Alopecia Areata was 0.1% to 0.2% in the 1970s in the US population[21].

Acne is considered an ‘adolescent’ disorder[22] but it is prevalent in adults as well. A study found 54% women and 40% men in a sample to have acne, and its prevalence did not decrease till after 44 years[23]. The global prevalence of acne is 9.4%, making it the eighth most prevalent disease[24].

Epidermolysis Bullosa Simplex is a congenital blistering disease, in which blistering occurs due to a minor trauma[25]. The prevalence is 6 - 30 per million live births. However, many cases remain undiagnosed, suggesting that the actual number of cases would be higher[26]. It is aggravated by high temperatures and sweating.

If not treated properly, diseases can lead to severe depression, low self-esteem and quality of life and loss of independant living. The magnitude of their effects is comparable with asthma, arthritis and epilepsy.

2.2. Machine Learning in Skin Disease Detection

Machine Learning is an application of artificial intelligence through which systems get the ability to learn patterns in labelled data and apply them to unseen data. Computers learn patterns automatically, however the features provided have to be hand picked and pre-processed by a human engineer. This poses itself as a limitation as the work is limited by the knowledge and discretion of the human user.

Machine learning has been used in the context of skin diseases classification to identify diseases, lesions or both. All pieces of work initially use signal processing techniques to extract the lesion from an image and the some go a step further to classify it to a disease using machine learning on visual and non-visual features of the extracted lesion. In all bodies of work, there has been a limited number and type of diseases that were catered to.

P. Kharazmi and others[27] segmented skin lesions using the density of the hemoglobin and melanin quanities. Classification using Mahalanobis distance on these components, combined with non-visual shape and colour information gave an accuracy of 96.5%. Another use of image processing and machine learning on dermatological diseases was the application of the sobel operator to extract edges and thresholding using Otsu's method. Feature extraction was tailored around different diseases, for example, bumps caused by hair folicles are absent around lesions of pityriasis rubra pilaris. Visual features and site information was fed to a shallow neural network and an accuracy of upto 95% was achieved[28].

Maximum a posteriori probability estimates, based on the Bayes theorem, have been used to classify lesions. Lesions were segmented using Linear Discriminant Analysis and K-means++ algorithm. Area Under the Curve (AUC) accuracy of 94.2% was achieved[29].

Another body of work[30] also used Bayesial Analysis (MAP (Maximum A Posteriori) estimation). The model segments the skin lesion, detects the nearby hair and identifies the lesion as a network of pigmentation ('pigment network'). Pigment network is labelled as either 'background', 'absent' or 'present'. S. Joseph and J. R. Panicker[31] devised two algorithms which used technique also mentioned in the above articles to remove hair from melanoma lesions and classify them. 200 images were used, of which 40 were melanoma moles, 80 atypical moles and 80 normal moles. The highest overall accuracy was 93.5%.

Zortea et al.[32] found a simple statistical classifier to have slightly better specificity and sensitivity scores than dermatologists on detection of melanoma using dermoscopic images. The classifier used features of asymmetry, color, border, geometry, and texture.

In machine learning the body of work focuses on both dermatological diseases and skin cancer. However, due to the limitation of providing features of the data to image processing and machine learning at human discretion, a diverse range and quanitity of skin diseases are not catered to.

2.3. Deep Learning in Skin Disease Detection

Deep Learning is application of Artificial Neural Networks which has at least 5 and upto more than 100 layers of a neural network. Deep learning is also called heirarchical learning because the layers of the deep neural network learn features heirarchically, starting from higher level features and eventually learning the low level details. Unlike Machine Learning, a deep neural network does not need to be provided hand picked features to learn to identfy a pattern because the algorithm has the ability to learn discriminant features on its own. This removes the need for a human engineer to handpick discriminant attributes for pattern recognition.

Deep learning has been used widely in the domain of dermatology for a broad range of diseases and cancers. Most studies have used a dataset which is available at a nearby research center or hospital. This adds a geographical limitation, the occurance and distribution of skin diseases varies across regions and as a consequence a model prepared in one part of the world is not as useful in another.

I.-M. Syu et al.[33] trained a Convolutional Neural Network (CNN) with 5700 normal and psoriasis lesion images to identify psoriasis lesions. A validation accuracy of 95% was achieved after 40 epochs of training.

Liao H.[34] used 23,000 skin disease images from the Dermnet dataset to train a Neural Network. OLE dataset from the New York Department of Health with 23,000 images. VGG16, VGG19 and GoogleNet architectures were used. 73.1% Top-1 accuracy was achieved on the Dermnet and 31.1% accuracy was achieved on the OLE Dataset. Zhang et al.[35] used a pretrianed GoogleNet v3 to classify 4 diseases - melanocytic nevus, seborrheic keratosis, basal cell carcinoma (BCC) and psoriasis. Data was obtained from Peking Union Medical College and split in training and validation sets. The accuracy was 86.54% $\pm$ 3.63% on the training set and 85.86% $\pm$ 4.649% on the validation set.

Scientists at Stanford[36] fine tuned the GoogleNet, already trained on 1.28 million ImageNet[37] images, with a dataset of 757 disease classes. The taxonomy of these diseases broke down into 2,032 diseases. The data was from 18 different clinicial-curated, online repositories and the Stanford University Medical Center. The accuracy was 72.1% $\pm$ 0.9%.

A dataset was obtained by Codella et al.[38] from the International Skin Imaging Collaboration containing 2624 images of melanoma, atypical nevi, and benign lesions. Ensemble models of SVM, sparse coders and CNN achieved an accuracy of 93.1% on classification of melanoma vs. non-melanoma lesions and 73.9% accuracy on classification of melanoma vs. atypical legions. Han et al.[39] classified 12 skin diseases more than 19,000 images on the ResNet-152 model. The performance of the model was comparable with 16 dermatologists.

An Internation pool of doctors or various experience levels were outperformed by a CNN in the task of melanoma detection[40]. Google's Inception v4 CNN was tested in a 300 image test set extracted from University of Heidelberg's Department of Dermatology. Dermatologists of less than 2 (Beginner), 2 - 5 (Skilled) and more than 5 (Expert) years of experience were given the same images. The Receiver Operating Characteristic (ROC) areas were 0.79 and 0.82 for levels 1 and 2, respectively.

Most of the body of work in Deep Learning is focused on Skin Cancer detection. Benign lesions are rarely touched upon, only work conducted by Haofu Liao[34] catered to a broad range of benign diseases. Since the appearance, frequency and distribution of diseases varies among regions, our dynamic required original work.

2.4. Deep Learning in other Domains of Medicine

Deep Learning has widespread use in medicine. Pulmonary Tuberculosis has been detected in chest radiographs using ensembles of the AlexNet and GoogleNet with an AUC of 0.99[41]. Another study used a Deep CNN to classify radiographs between frontal and lateral with high throughput. 100% of the test set was classified correctly[42]. A domain specific and a pre-trained neural network was used to classify lesions in breast MRIs with an accuracy of 83% and 71% respectively[43]. Multi-modal preoperative brain images were used to predict the survival time of patients of glioma, the most aggressive form of brain tumor. The highest accuracy achieved was 89.9%[44].

3. Materials and Methods

3.1. Data Collection

Images of diseases were collected at Dermatology Department of Civil Hospital Karachi. Incoming patients were asked for their consent to have a picture of their lesion taken. Images of 10 diseases were collected in substantial quantity to be used in the study. Table 1 has the frequency of images for each disease.

A Sony DSC-W320 camera, with a lens of focal length of 26 to 105mm, was used to take pictures at VGA (640 x 480 pixels) resolution. The camera was hand held and a picture of the lesion was taken without regard to the body part. The environment and lighting conditions for taking pictures were not controlled.

Each picture was assigned a label by a senior dermatologist of 30 years of experience and cross checked by another dermatologist. An image was only used if there was consensus among the doctors on its diagnosis.

Table 1: Count of Images for each Disease

Disease	**Number of Images**
Psoriasis	108
Lichen Planus	32
Scabies	29
Pemphigus Vulgaris	16
Herpes Zoster	16
Acne	20
Leishmaniasis	13
Infected Scabies	11
Alopecia Areata	8
Epidermolysis Bullose Simplex	13

Crops of just the infected area were obtained from these images and train and validation splits were obtained for training. Table 2 has the number of train and validation images for each disease. Some captured images had more than 1 lesions which resulted in a greater number of cropped images. Arbitrary selection was used to balance the classes.

The size of the images used for training was 224 x 224 pixels in RGB channels.

Table 2: Count of Images for each Disease

Disease	**Training Images**	**Validation Images**
Psoriasis	40	10
Lichen Planus	26	15
Scabies	40	13
Pemphigus Vulgaris	19	6
Herpes Zoster	19	7
Acne	16	5
Leishmaniasis	7	3
Infected Scabies	7	3
Alopecia Areata	4	1
Epidermolysis Bullose Simplex	18	6

3.2. Model

The deep residual neural network[45], pretrained on the imagenet dataset[37], was obtained. This model was optimzed to do well in the large scale image recognition task of ILSVRC. The final sigmoid layer of the model was removed and the weights of the rest of the layers were kept frozen. Our data was passed through this neural network to obtain discriminant features which were used to train only one sigmoid layer from scratch. Within the scope of transfer learning, this technique is called fixed feature extraction and gives better performance than a untrained neural network with randomly initialized weights[46].

Architecture of the ResNet 18 is shown in Figure 2. The ResNet trains residual connections between layers which allow for efficient learning. Residual connections are hosted in residual blocks, as shown in Figure 1. The connections allow for gradient to flow without passing through activation functions. It was first place in the ILSVRC 2015 task[45].

A weighted Cross Entropy cost function, as shown in Equation 1, was used. The weights acted as a penalty and allowed the classifier to learn quicker on diseases with fewer images. This allowed us to balance the varying number of images among the classes.

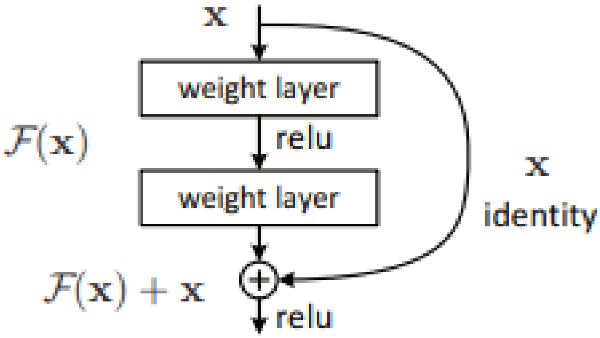


Figure 1: Residual Block[45]

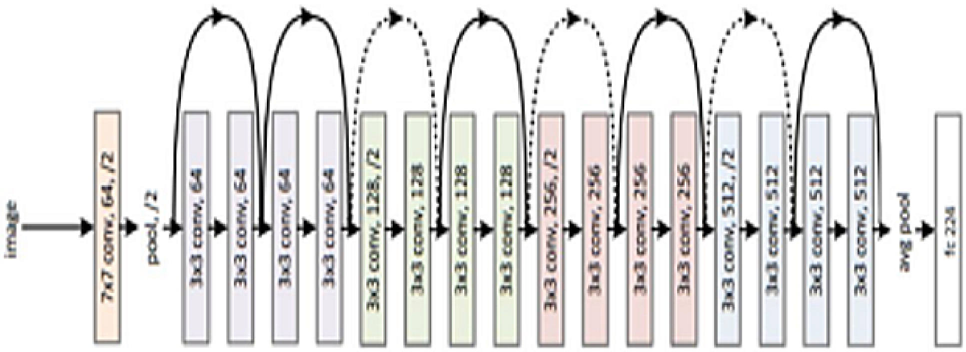


Figure 2: 18 Layer ResNet Architecture[47]

$$E_{\text{entropy}} = -\sum_{n}^{N}\sum_{k}^{c} t_k^n \ln y_k^n \tag{1}$$

3.3. Training

A Lenovo ThinkPad x240s with Intel Core i5-4250U CPU with clock rate 1.30GHz and 8GB RAM was used for training. Average training time was 14 minutes on 196 training images.

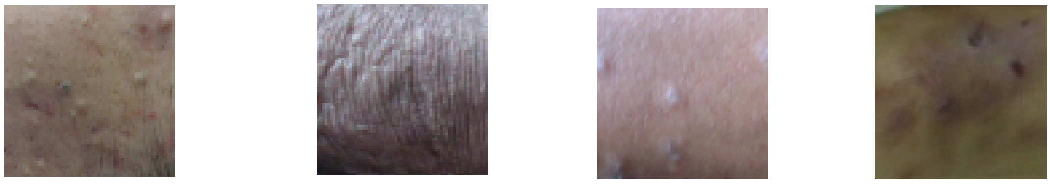
The model was trained for 25 epochs, with a batch size of 6. The data set was normalized prior to training because the environment was uncontrolled when taking pictures of the diseases, normalization allowed for the varying brightness to be handled. The learning rate was 0.01 and momentum was kept at 0.9. The learning rate decay was 0.1 with a step size of 7. Many combinations of these parameters were tried and these were chosen empirically.

4. Results

4.1. Performance

The model was designed to select the top two choices for each image. The average accuracy achieved with the first choice on the validation set was 86%. We trained multiple models and the accuracy ranged from 82% to 91%. Figure 5 is a correlation matrix of the classification. Figure 4 shows some successful classifications.

Figure 3 shows some incorrect classifications. The difference between Scabies and Infected Scabies is of colour. Similarly, if the lesions of Lichen Planus are not dark, they looks similar to Scabies. Epidermolysis Bullose Simplex and Pemphigus Vulgaris are similar looking diseases, with large red lesions and the classifier also confuses the two. In many cases the correct class is chosen as a second choice. Confusion generally exists between two diseases that look similar.

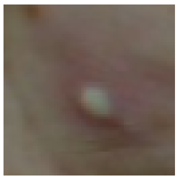


A: Acne	A: Psoriasis	A: Lichen Planus	A: Scabies
1: Herpes Zoster	1: Alopecia Areota	1: Scabies	1: Infected Scabies
2: Pemphigus Vulgaris	2: Psoriasis	2: Lichen Planus	2: Scabies

Figure 3: Incorrect Predictions by the Model

A: Epidermolysis Bullosa Simplex
1: Epidermolysis Bullosa Simplex
2: Pemphigus Vulgaris

A: Leishmaniasis
1: Leishmaniasis
2: Pemphigus Vulgaris



A: Infected Scabies
1: Infected Scabies
2: Herpes Zoster

A: Acne
1: Acne
2: Herpes Zoster

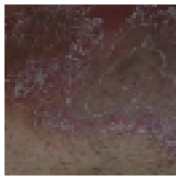
A: Alopecia Areota
1: Alopecia Areota
2: Lichen Planus

A: Infected Scabies
1: Infected Scabies
2: Scabies

A: Herpes Zoster
1: Herpes Zoster
2: Pemphigus Vulgaris

A: Lichen Planus
1: Lichen Planus
2: Psoriasis

A: Herpes Zoster
1: Herpes Zoster
2: Psoriasis



A: Psoriasis
1: Psoriasis
2: Acne

Figure 4: Correct Predictions by the Model. Actual (A), Top Prediction (1), 2nd Prediction (2)

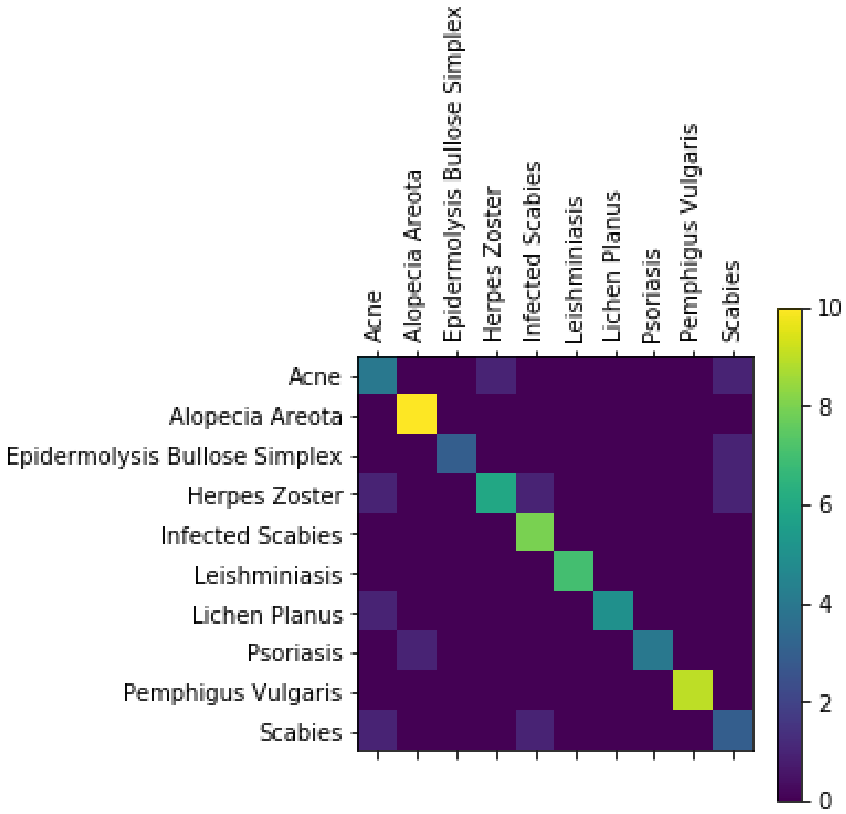


Figure 5: Correlation Matrix

4.2. Avoidable Bias

Dermatologists of diverse degrees of experience were given a subset of 19 images of the validation set. This subset had images of all classes. The doctors were asked to select up to two diseases, in the order of decreasing likelihood for images of the skin diseases with no knowledge of the part of the body where the disease occured. Their decision was purely based on the appearance of the lesions.

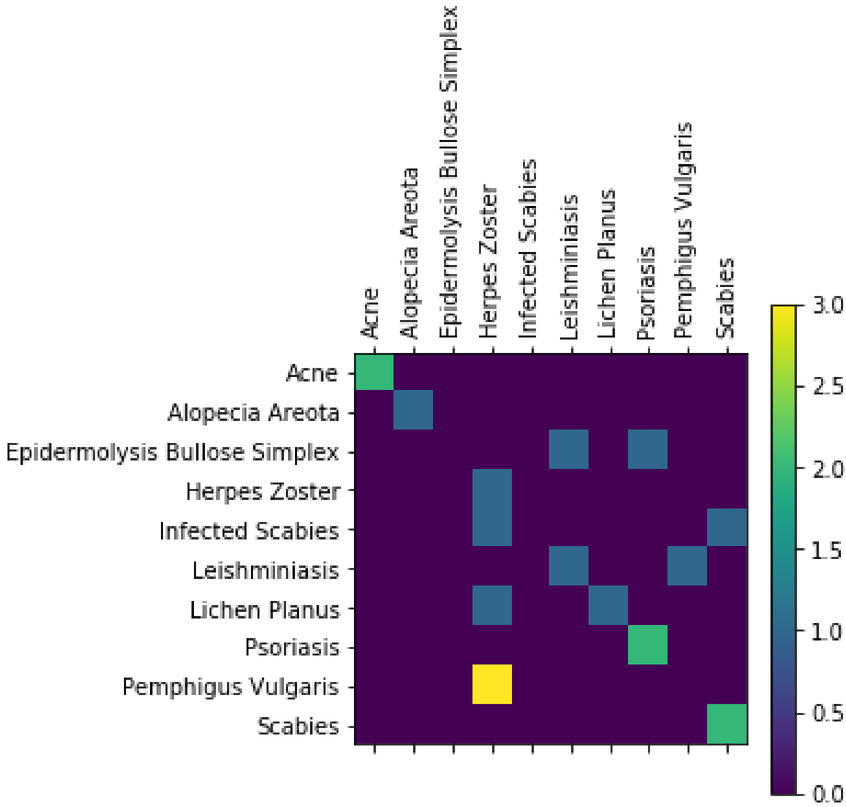
A visualization of the performance of the model and the doctors in the form of a confusion matrix helped to signify the similarities in their approach and confusion between diseases. It signified how quickly artificial intelligence can reach the level of educated human beings.

On this subset, our model has an accuracy of 74% on the first choice and 85% on the second choice.

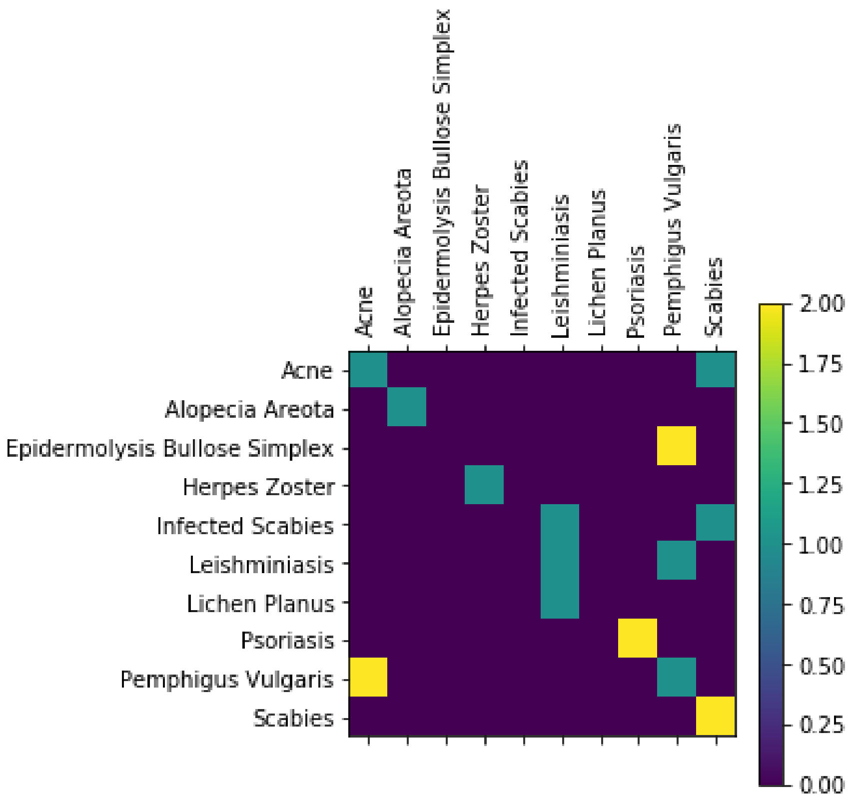
Table 3 mentions the accuracy and years of experience of the doctors and the model on the mentioned 19 images. Figures 6 are correlation matrices of the doctors on their first choice. Senior doctors of 30 and 20 years of experience only selected one choice.

Table 3: Performance of Doctors on subset of Validation set.

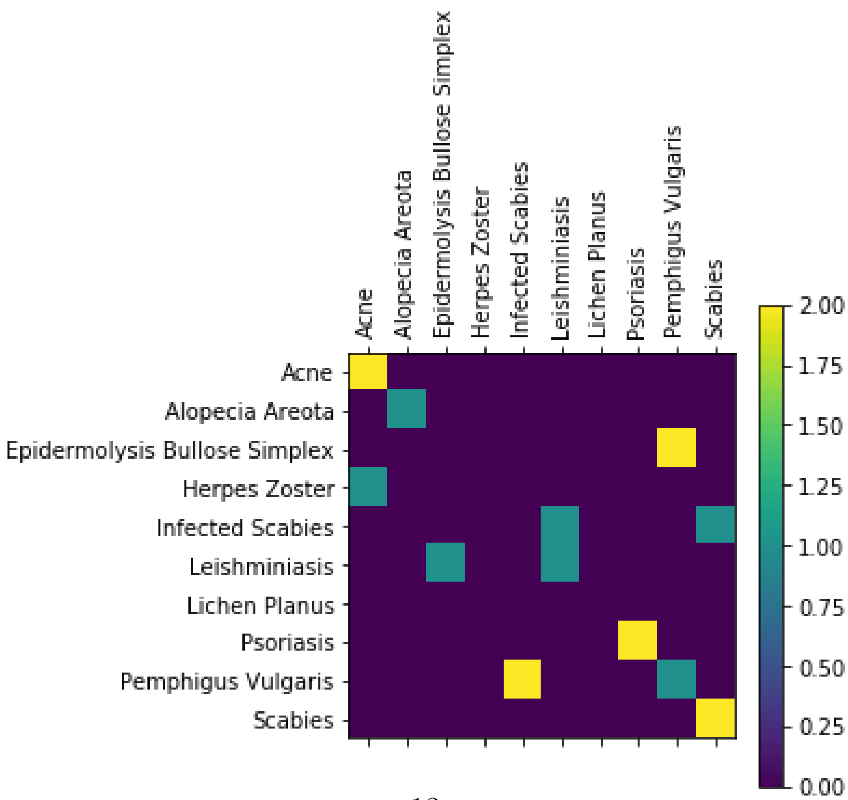
Candidate	Accuracy on First Choice (%)	Accuracy on Second Choice
Doctor (30 years Experience)	54	54
Doctor (20 years Experience)	54	54
Doctor (3 years Experience)	68	73
Doctor (3 years Experience)	47	57
Doctor (2 years Experience)	74	79
Doctor (2 months Experience)	37	37
Model	74	85



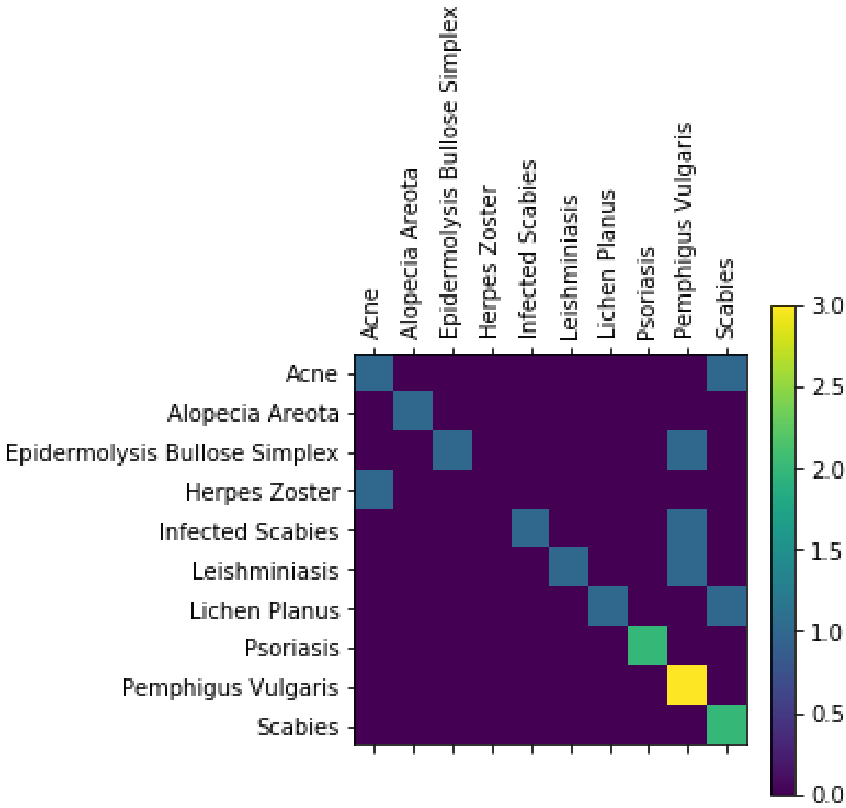
(a) Correlation Matrix of Doctor with 30 Years of Experience



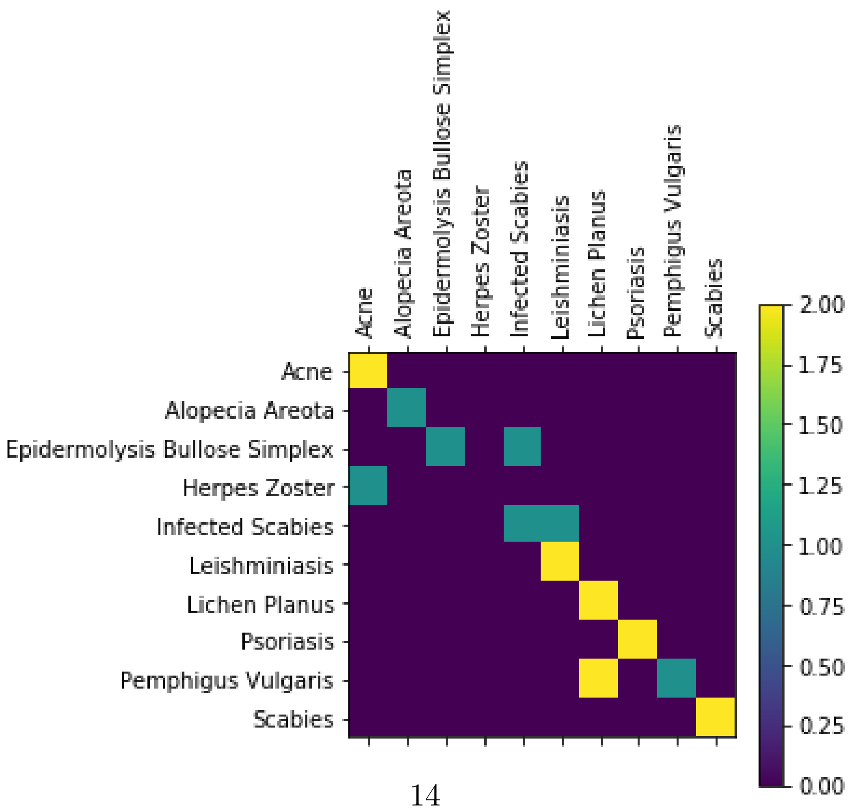
(b) Correlation Matrix of Doctor with 20 Years of Experience



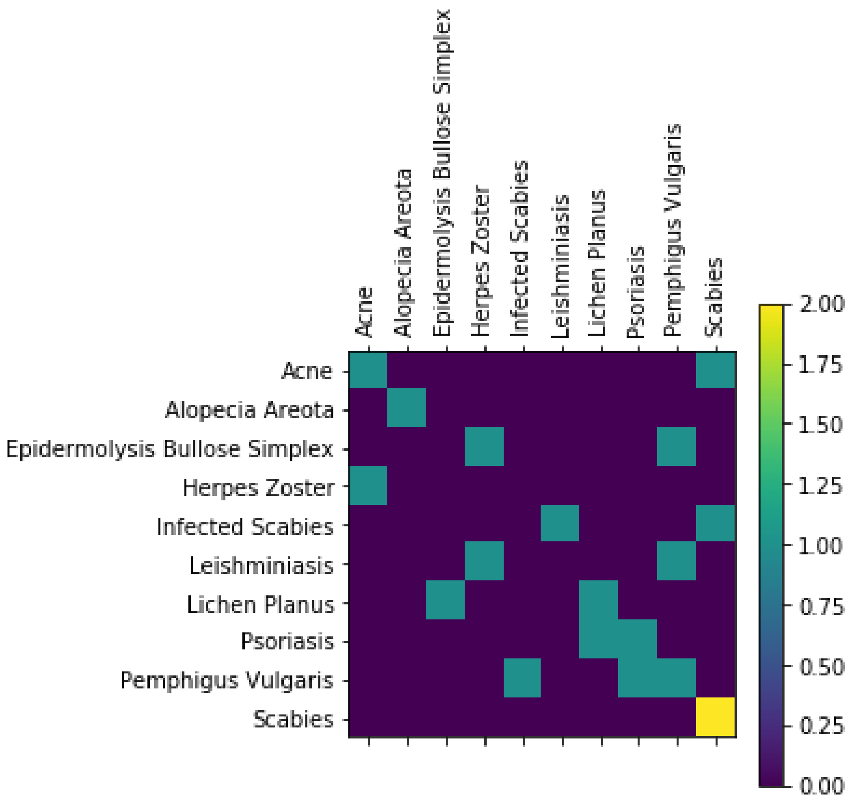
(c) Correlation Matrix of Doctor with 3 Years of Experience



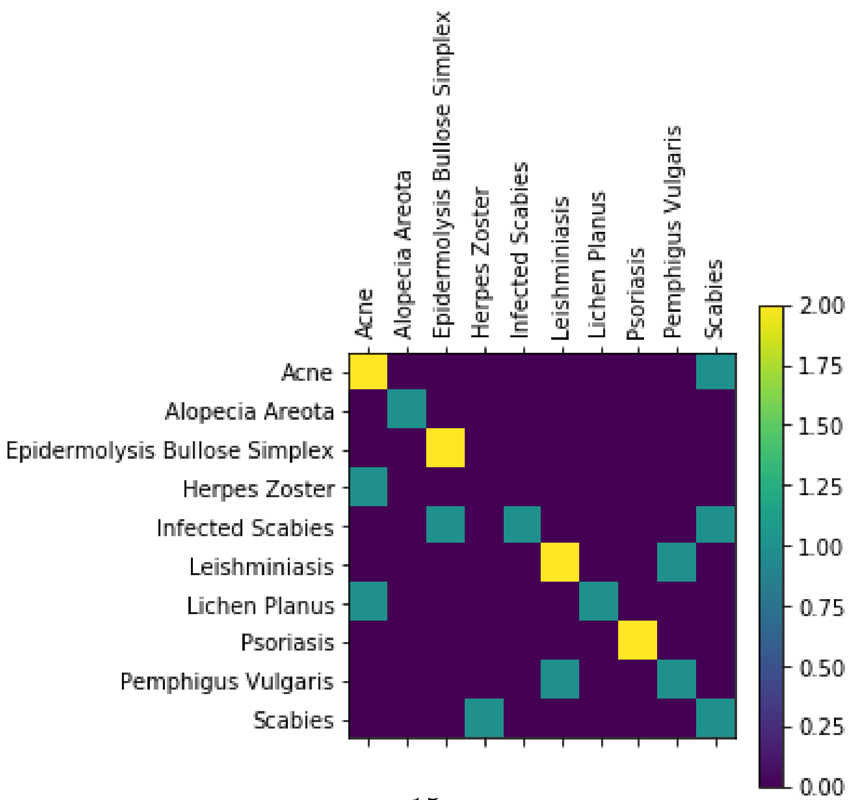
(d) Correlation Matrix of Doctor with 3 Years of Experience



(e) Correlation Matrix of Doctor with 2 Years of Experience



(f) Correlation Matrix of Doctor with 2 Months of Experience



(g) Correlation Matrix of Model

Figure 6: Correlation Matrix of Doctors and Model on Subset of Validation Set

4.3. Significance

Our work has reused a model created for the ILSVRC like the works of Esteva et al.[36], which used the GoogleNet, and the work of Han et al.[39], which used the ResNet-152, while our work used the ResNet-18. The lowest number of images used so far is 2624 images used by Codella et al.[38] and the largest is 129,450 used by Esteva et al.[36] while our work only used only 196 images for training. The datasets of all the other studies were already curated, either online or at a research center or hospital, while our work required the data to be collected.

The model prepared by Esteva et al.[36] was tested against the ability of 21 dermatologists, of which most are not as accurate as the model. The study conducted at University of Heidelberg[40] compared the ability of 58 dermatologists against a model, most of whom were out performed. Our work was compared against 6 dermatologists of various training programs and experience levels. Most of the doctors, like the rest of the studies, were out performed by the model. Only 1 doctor was on par with the model.

Performance of the software was shared with doctors at the Civil Hospital Karachi as a demo on live patient images. They were of the opinion that the it can help undergraduate trainees, house officers and general practitioners in their diagnosis and differential diagnosis. It will specially be helpful to above mentioned doctors in confirming the right choice. The model can assist in detecting secondary bacterial infection in commonly occuring scabies.

They thought it could especially be used in telemedicine, to cater to the large number of their patients which come from remote areas without easy access to a dermatological facility. It will reduce the load on hospitals in metropolitan areas and patients from rural and remote areas will not be required to travel for healthcare.

5. Conclusion

This study has provided a proof of concept that recent advances in Deep Learning can be harnessed to solve the problem of a lack of skilled professionals. We shall invest more time with multiple professionals to collect data all across Pakistan to build a larger dataset which will be used to improve the predictive model further.

Such an application will address the shortage of doctors and also provide an assistance to doctors, especially juniors. It can also be used by students and institutes for traning purposes.

Hospitals with complete facilities are only available in metropolitan areas and most people need to travel from rural areas to find access to healthcare. Such an application can effectively provide an initial diagnosis to professionals working in a healthcare center in rural areas to provide primary medication to avoid a disease from getting worse.

6. Summary Table

What was already known on the topic:

- Skin diseases can be identified using computers with high accuracy and performance comparable with doctors.
- Tuning and reuse of general image recognition models can be harnessed for specific image recognition tasks.

What this study added to our knowledge:

- Healthcare crisis in the 3rd world can be addressed using AI powered telehealth.
- A broad and thorough Skin Disease Classification Decision Support System is possible if image data on a wide range of diseases is collected in diverse geographical regions.
- The quality of training programs and individuals in healthcare can be assessed by comparing trainees with artificially intelligent software.

7. Acknowledgements

The Dermatology Department of Civil Hospital, Karachi, Pakistan played a major role in providing us access to patients to collect data with their consent and Dr. Abdul Aziz for providing diagnosis for the images.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

The authors do not have competing interests.

Data will be made available on request.

References

[1] K. N. Memon, R. A. Soomro, M. S. Ansari, Pattern of skin diseases in patients visiting a tertiary care health facility at hyderabad, pakistan, Journal of Ayub Medical College Abbottabad 23 (4) (2011) 37–39.

[2] I. A. Humaira Maryum, M Zafar Alam, Pattern of skin diseases in a tertiary care private hospital, karachi, Journal of Pakistan Association of Dermatologists 24 (4) (2014) :292–297.

[3] E. Mallon, J. Newton, A. Klassen, S. Stewart-Brown, T. Ryan, A. Finlay, The quality of life in acne: a comparison with general medical conditions using generic questionnaires., The British journal of dermatology 140 (4) (1999) 672–676.

[4] N. Imran, Z. Azeem, I. I. Haider, N. Amjad, M. R. Bhatti, Brain drain: post graduation migration intentions and the influencing factors among medical graduates from lahore, pakistan, BMC research notes 4 (1) (2011) 417.

[5] O. Chosidow, Scabies, New England Journal of Medicine 354 (16) (2006) 1718–1727.

[6] J. Heukelbach, H. Feldmeier, Scabies, The Lancet 367 (9524) (2006) 1767–1774.

[7] H. Feldmeier, G. Singh Chhatwal, H. Guerra, Pyoderma, group a streptococci and parasitic skin diseases–a dangerous relationship, Tropical Medicine & International Health 10 (8) (2005) 713–716.

[8] I. Brook, Microbiology of secondary bacterial infection in scabies lesions., Journal of clinical microbiology 33 (8) (1995) 2139–2140.

[9] A. S. Boyd, K. H. Neldner, Lichen planus, Journal of the American Academy of Dermatology 25 (4) (1991) 593–619.

[10] K. Kawai, B. P. Yawn, P. Wollan, R. Harpaz, Increasing incidence of herpes zoster over a 60-year period from a population-based study, Clinical infectious diseases 63 (2) (2016) 221–226.
[11] J. I. Cohen, Herpes zoster, New England Journal of Medicine 369 (3) (2013) 255–263.
[12] J. Alvar, I. D. Velez, C. Bern, M. Herrero, P. Desjeux, J. Cano, J. Jannin, M. den Boer, W. L. C. Team, et al., Leishmaniasis worldwide and global estimates of its incidence, PloS one 7 (5) (2012) e35671.
[13] P. Desjeux, Leishmaniasis: current situation and new perspectives, Comparative immunology, microbiology and infectious diseases 27 (5) (2004) 305–318.
[14] R. Parisi, D. P. Symmons, C. E. Griffiths, D. M. Ashcroft, et al., Global epidemiology of psoriasis: a systematic review of incidence and prevalence, Journal of Investigative Dermatology 133 (2) (2013) 377–385.
[15] L. Naldi, Psoriasis: Epidemiology, World Clinics: Dermatology: Psoriasis: Volume 2, Number 1 2 (1) (2018) 1–10.
[16] A. Y. Chang, Pemphigus vulgaris, in: Inpatient Dermatology, Springer, 2018, pp. 355–358.
[17] K. Kridin, Pemphigus group: overview, epidemiology, mortality, and comorbidities, Immunologic research 66 (2) (2018) 255–270.
[18] A. Firooz, A. Mazhar, A. R. Ahmed, Prevalence of autoimmune diseases in the family members of patients with pemphigus vulgaris, Journal of the American Academy of Dermatology 31 (3) (1994) 434–437.
[19] A. C. V. Fricke, M. Miteva, Epidemiology and burden of alopecia areata: a systematic review, Clinical, cosmetic and investigational dermatology 8 (2015) 397.
[20] V. C. Fiedler, Alopecia Areata: A Review of Therapy, Efficacy, Safety, and Mechanism, Archives of Dermatology 128 (11) (1992) 1519–1529. arXiv:https://jamanetwork.com/journals/jamadermatology/articlepdf/554546/archderm_128_11_016.pdf, doi:10.1001/archderm.1992.01680210097016.
URL https://doi.org/10.1001/archderm.1992.01680210097016
[21] K. Safavi, Prevalence of alopecia areata in the first national health and nutrition examination survey, Archives of dermatology 128 (5) (1992) 702–702.
[22] B. Dreno, F. Poli, Epidemiology of acne, Dermatology 206 (1) (2003) 7.
[23] V. Goulden, G. Stables, W. Cunliffe, Prevalence of facial acne in adults, Journal of the American Academy of Dermatology 41 (4) (1999) 577–580.
[24] J. Tan, K. Bhate, A global perspective on the epidemiology of acne, British Journal of Dermatology 172 (S1) (2015) 3–12. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/bjd.13462, doi:10.1111/bjd.13462.
URL https://onlinelibrary.wiley.com/doi/abs/10.1111/bjd.13462
[25] T. W. Cooper, E. A. Bauer, Epidermolysis bullosa: a review, Pediatric dermatology 1 (3) (1984) 181–188.
[26] E. Sprecher, Epidermolysis bullosa simplex, Dermatologic clinics 28 (1) (2010) 23–32.
[27] P. Kharazmi, M. I. AlJasser, H. Lui, Z. J. Wang, T. K. Lee, Automated detection and segmentation of vascular structures of skin lesions seen in dermoscopy, with an application to basal cell carcinoma classification, IEEE Journal of Biomedical and Health Informatics 21 (6) (2017) 1675–1684. doi:10.1109/JBHI.2016.2637342.
[28] V. B. Kumar, S. S. Kumar, V. Saboo, Dermatological disease detection using image processing and machine learning, in: 2016 Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR), 2016, pp. 1–6. doi:10.1109/ICAIPR.2016.7585217.
[29] P. Wighton, T. K. Lee, H. Lui, D. I. McLean, M. S. Atkins, Generalizing common tasks in automated skin lesion diagnosis, IEEE Transactions on Information Technology in Biomedicine 15 (4) (2011) 622–629. doi:10.1109/TITB.2011.2150758.
[30] P. Wighton, T. K. Lee, H. Lui, D. I. McLean, M. S. Atkins, Generalizing common tasks in automated skin lesion diagnosis, IEEE Transactions on Information Technology in Biomedicine 15 (4) (2011) 622–629. doi:10.1109/TITB.2011.2150758.
[31] S. Joseph, J. R. Panicker, Skin lesion analysis system for melanoma detection with an effective hair

segmentation method, in: 2016 International Conference on Information Science (ICIS), 2016, pp. 91–96. doi:10.1109/INFOSCI.2016.7845307.

[32] M. Zortea, T. R. Schopf, K. Thon, M. Geilhufe, K. Hindberg, H. Kirchesch, K. Møllersen, J. Schulz, S. O. Skrøvseth, F. Godtliebsen, Performance of a dermoscopy-based computer vision system for the diagnosis of pigmented skin lesions compared with visual evaluation by experienced dermatologists, Artificial intelligence in medicine 60 (1) (2014) 13–26.

[33] I.-M. Syu, C.-H. Lai, G.-S. Lin, S.-K. Chai, Psoriasis detection based on deep neural network (2018) 1–2.

[34] H. Liao, A deep learning approach to universal skin disease classification.

[35] X. Zhang, S. Wang, J. Liu, C. Tao, Computer-aided diagnosis of four common cutaneous diseases using deep learning algorithm (2017) 1304–1306.

[36] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (7639) (2017) 115.

[37] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, Imagenet: A large-scale hierarchical image database, in: 2009 IEEE conference on computer vision and pattern recognition, Ieee, 2009, pp. 248–255.

[38] N. Codella, J. Cai, M. Abedini, R. Garnavi, A. Halpern, J. R. Smith, Deep learning, sparse coding, and svm for melanoma recognition in dermoscopy images, in: L. Zhou, L. Wang, Q. Wang, Y. Shi (Eds.), Machine Learning in Medical Imaging, Springer International Publishing, Cham, 2015, pp. 118–126.

[39] S. S. Han, M. S. Kim, W. Lim, G. H. Park, I. Park, S. E. Chang, Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm, Journal of Investigative Dermatology 138 (7) (2018) 1529–1538.

[40] R. study level I, level II Groups, A. Enk, C. Fink, F. Toberer, R. Schneiderbauer, H. A. Haenssle, T. Buhl, A. Blum, A. Kalloo, A. B. H. Hassen, L. Thomas, L. Uhlmann, Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists, Annals of Oncology 29 (8) (2018) 1836–1842. arXiv:http://oup.prod.sis.lan/annonc/article-pdf/29/8/1836/25682281/mdy166.pdf, doi:10.1093/annonc/mdy166.
URL https://doi.org/10.1093/annonc/mdy166

[41] P. Lakhani, B. Sundaram, Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks, Radiology 284 (2) (2017) 574–582.

[42] A. Rajkomar, S. Lingam, A. G. Taylor, M. Blum, J. Mongan, High-throughput classification of radiographs using deep convolutional neural networks, Journal of digital imaging 30 (1) (2017) 95–101.

[43] G. Amit, R. Ben-Ari, O. Hadad, E. Monovich, N. Granot, S. Hashoul, Classification of breast mri lesions using small-size training sets: comparison of deep learning approaches, in: Medical Imaging 2017: Computer-Aided Diagnosis, Vol. 10134, International Society for Optics and Photonics, 2017, p. 101341H.

[44] D. Nie, H. Zhang, E. Adeli, L. Liu, D. Shen, 3d deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 212–220.

[45] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition (2016) 770–778.

[46] J. Yosinski, J. Clune, Y. Bengio, H. Lipson, How transferable are features in deep neural networks?, in: Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 27, Curran Associates, Inc., 2014, pp. 3320–3328.
URL http://papers.nips.cc/paper/5347-how-transferable-are-features-in-deep-neural-networks.pdf

[47] PyTorch, Resnet-18 (Dec 2017).
URL https://www.kaggle.com/pytorch/resnet18