MSc Artificial Intelligence

Assessment of Eczema in Dark Skin using Generative AI (DCGAN)

Project Documentation – MedVision – AI Research And Development Projects

Module Leader

Dr. Maria Luisa

Student Names and IDs

Chandima Ranasinghe Liyana Arachchige - 24047270

Ashika Srinivasan - 32066221

Bhagya Chandradas Velluva Vedu - 32074965

Assessment of Eczema in Dark Skin using Generative AI (DCGAN)

Introduction

Eczema, also known as atopic dermatitis, is a common inflammatory skin disorder that affects people of all skin colors. However, because of differences in skin appearance, pigmentation, and the poor representation of dark skin in current medical databases, diagnosing eczema in people with darker skin tones offers unique issues. This research tackles these issues by presenting an AI-based approach for reliable eczema evaluation and diagnosis in those with dark skin.

Problem Description

The challenges in accurately diagnosing eczema among individuals with dark skin are multifaceted. These challenges stem from incomplete data, the unique manifestations of the condition on dark skin, and limitations in visual assessment methods. Conventional diagnostic techniques often struggle to differentiate eczema from other skin disorders, potentially causing incorrect diagnoses and treatment delays. To address this, the development of an AI-based evaluation system tailored to recognize the nuances of eczema on dark skin holds the potential to significantly enhance diagnostic precision and ultimately improve patient outcomes.

Moreover, the inadequacy of a diverse dataset encompassing various skin tones, lighting scenarios, and levels of severity poses a hindrance to the comprehensive education and training of nursing students in eczema diagnosis. This deficiency could subsequently lead to instances of overlooked diagnoses and postponed therapeutic interventions.

Implemented Solution

Collecting and creating a high-quality dataset to identify Eczema in dark skin using traditional methods can be challenging due to many reasons including underrepresentation or lack of diversity in existing datasets, variability of eczema symptoms, photographic challenges (Difficulties in capturing the appearance of eczema in dark skin), ethical considerations, social and cultural barriers. Hence, it is crucial to look for state-of-the-art AI technologies to tackle these challenges and generate a synthetic dataset that can be used to diagnose eczema in dark skin accurately. Over recent years, Generative AI has shown promise in generating numerous synthetic medical datasets such as brain MRI scans, chest X-rays, and skin lesions (Chen et al., 2021), Hence, Generative AI technology was used to implement the solution for the project.

Different GAN architectures were experimented on as a part of the research and Deep Convolutional GAN (DCGAN) was identified as the ideal technology to implement the solution due to its stability and the size of the dataset. DCGAN was introduced by Radford et al., (2016) as a more stable GAN architecture that can learn good representations from images to generate synthetic data. Over the years, DCGAN has become a popular GAN architecture for generative modeling.

To tackle the complex challenge of accurately diagnosing eczema in individuals with dark skin, the DCGAN was selected as the appropriate AI technique to address the problem. DCGAN consists of two primary components called the generator and discriminator which employs deep convolutional neural networks designed for generating synthetic datasets. The generator creates synthetic data from random noise to resemble real images, while the discriminator distinguishes between real and synthetic images. They both are trained together in an adversarial manner. Because of this adversarial training, the generator improves its ability to create more realistic images.

The implemented solution for the problem is a DCGAN model optimized for generating eczema images using a smaller dataset. However, the model is not capable of generating eczema images that look like real images due to the limitations such as using a smaller dataset and not having GPU access. The implemented solution is presented as a Google Colab notebook. The parameters for both the generator and the discriminator of the DCGAN were carefully chosen after experimenting with different values to optimize the results even with the limits of the dataset and resources.

Justification

1. Data Preprocessing

Data augmentation is a crucial step in enhancing the diversity and quality of the dataset to train the DC GAN model effectively. Given the challenge of acquiring an adequately diverse collection of images featuring eczema on dark skin, we have implemented augmentation techniques to amplify our dataset's richness. Acquiring high-quality images featuring eczema on dark skin tones proved to be a formidable task. Data collection efforts centered around sourcing images from multiple reliable resources, primarily provided by the client. These resources included the Mind the Gap Website, Mind the Gap Clinical Handbook, Brown Skin Matters Instagram, and Do not Forget the Bubbles - Skin Deep (Mukwende, et al., 2020) (*Eczema - Skin Deep*. 2020). While these sources promised diversity regarding skin tones, they also presented their challenges.

One of the most pressing challenges was the limited availability of images showcasing eczema on dark skin. This scarcity reflects a broader issue in medical databases, where dark skin is often underrepresented. This scarcity hindered our attempts to create a sufficiently diverse dataset that accurately encapsulated the variations in eczema presentation on darker skin tones. Furthermore, the images obtained from these resources often required substantial preprocessing. Many images suffered from poor focus, lighting irregularities, and varying backgrounds. These factors posed significant obstacles to accurately identifying and isolating eczema features on dark skin. These artifacts could introduce noise into the dataset and affect the AI model's performance.

A meticulous preprocessing strategy was devised to address these challenges and ensure the dataset's quality. An essential step was the removal of backgrounds to isolate the skin regions of interest. This process demanded manual intervention to ensure accuracy. Additionally, automatic cropping techniques were implemented, but their limitations led to a combination of manual and automated cropping to achieve optimal results. Through these efforts, we managed to curate a dataset that featured diverse dark skin tones and adhered to high-quality standards necessary for DC GAN training. This dataset, although relatively small due to the scarcity of suitable images, formed the foundation on which data augmentation techniques were applied.

To facilitate data augmentation, we employed the Augmentor library, which offers a range of image transformation options. Augmentor allowed us to apply various modifications to the original images, thereby generating multiple augmented versions. Techniques such as rotation, zooming, horizontal and vertical flipping, and random cropping were selectively employed to simulate different eczema conditions on dark skin. Moreover, we implemented advanced augmentation techniques that leverage deep learning concepts to enrich the dataset further. Using PyTorch's torchvision.transforms library, we designed a pipeline that applies a combination of transformations to each image, including random horizontal flips, vertical flips, rotations, color jittering, and random affine transformations. This process resulted in multiple variations of each image, effectively increasing the dataset size.

Despite the challenges faced during data collection and preprocessing, our commitment to building a robust dataset remained unwavering. Finally, the data collection phase of our project was marked by a proactive approach to overcoming challenges. The scarcity of images and the need for extensive preprocessing did not deter us from our goal of creating a reliable and representative dataset. Through collaboration with reputable resources and persistent efforts, we developed a dataset that forms the cornerstone of our AI-based synthetic data generation of eczema that helps nursing students to diagnose eczema in dark skin.

2. DCGAN Model

DCGAN employs convolutional layers to capture hierarchical features from images (Behara et al., 2023). Hence, they are capable of extracting features from images at numerous scales, which is crucial for generating high-quality and diverse image datasets.

DCGAN has also shown potential in generating high-quality images even with a smaller dataset as it employs the minibatch discrimination technique that divides the training dataset into smaller batches and uses each batch to train the GAN individually (Fang et al., 2018). This leads the DCGAN to learn from a smaller dataset without overfitting. In addition, DCGAN has convolutional layers that can capture spatial patterns and features such as skin textures effectively. Furthermore, DCGAN employs progressive training that allows it to start the training process with lower-resolution images and increase the resolution of the images gradually. Hence, the learning process becomes more stabilized leading to better convergence even with a smaller training dataset.

The weights of the DCGAN should be randomly initialized from a normal distribution with standard deviation = 0.02 and mean = 0 (Radford et al., 2016). Hence, the weights of both transpose convolutional layer and the convolutional layer were initialized based on what the authors suggested in the DCGAN paper (Radford et al., 2016).

The pixel values of the training dataset were normalized to be in the range of -1 to 1. This prevents the DCGAN from becoming too sensitive to certain features of eczema images enhancing the convergence and the stability of the training process.

The transpose convolution is employed in the generator of the model to upsample and expand incoming feature maps. This increases the dimensions of the input feature maps leading to improvement in the quality of the output. Similarly, the discriminator model was built using the strided convolution to distinguish between real and fake eczema images.

The generator initiates with a dense (fully connected layer) and is followed by transpose convolution. Batch normalization is applied to the output and leaky ReLu is used as the activation function. Batch-normalization can stabilize the training of the model by normalizing the input of each layer. Similarly, leaky ReLu was chosen as it can address the vanishing gradient problem (Jacob, 2022).

At the end of the generator, there is a convolution layer and three filters are used on it. Because the generator is used to synthesize realistic eczema RGB images, three filters are used to align with the three channels of an RGB (Red, Green, and Blue) image, allowing the generator to create images with the necessary color information The activation function used to generate the RGB images is tanh, that maps the output between -1 and 1 facilitating accurate color representation of RGB images (Pixel value of RGB is 0 – 255). Hence, these fine-tuning techniques and parameter values are carefully chosen to generate RGB eczema images with accurate colors, visual coherence, and representation of real eczema images.

The discriminator is a simple binary classification network that takes both the real and the fake image and outputs a probability of whether the given image is real or fake.

The discriminator of the model is a binary classification network. It can take real and fake eczema images as the input and provide the probability of being real or fake as the output for the given input image. In the discriminator, a series of strided convolutions is employed and the activation function is leaky ReLu. To prevent overfitting a dropout layer of 0.3 is used as the regularization technique. In the final step, the feature maps are flattened and converted to a fully connected layer of 1 neuron to produce the output which classifies the input as real or fake. The activation function chosen for the fully connected layer is sigmoid which outputs the probability range between 0 and 1, signifying the probability of the input being real.

100 epochs are used to train the model, and it is visible that the generator starts synthesizing eczema images that look like real images when it completes the 100 epochs compared to the first 10 epochs. Even though the image quality can be improved by increasing the epochs value, the model starts overfitting. The objective of the project is to build a model that can be generalized to new data. Training the model too long by increasing the value of epochs leads to overfitting and preventing the model’s ability to generalize well to new data.

The proposed model can be further enhanced with a high-quality larger training dataset that is diverse and representative of various dark skin tones and GPU access to generate synthetic eczema images that look like real images.

3. Validation Techniques

Validation and performance measurement were integral in shaping AI outcomes, using Frechet Inception Distance (FID) and Inception Score (IS) as validation techniques for evaluating image quality and diversity. The IS metric accommodates the diverse manifestations of eczema across various dark skin tones. Furthermore, FID enables cross-domain comparisons, aiding in discerning eczema-specific attributes within dark skin patterns. This approach aligns with established GAN practices, lending credibility to the assessment of eczema using DCGAN. Chong and Forsyth's paper discusses evaluation metrics in generative models, specifically GANs, and proposes unbiased FID and IS metrics. They show enhanced accuracy and reduced bias compared to traditional methods. Tracking DCGAN progress over training epochs through FID and IS scores helps identify optimal image quality and diversity for eczema evaluation.

Based on the implementation, the Inception Score (IS) of 72.9 at epoch 1 is notably high, suggesting that the generated images are effectively recognised by the Inception model as belonging to their respective classes. However, the abrupt surge in IS to an extremely high value at Epoch 10 raises concerns and prompts a closer investigation. This unexpected spike was caused while evaluating generative models at earlier epochs and will not reflect their stabilised state, leading to unreliable scores. Additionally, the Fréchet Inception Distance (FID) score of 21.3 implies that the generated images are reasonably similar to real images within the chosen feature space, signifying promising progress. During the validation process with epochs set to 60, a conflict arises, indicating that the model's performance in terms of Fréchet Inception Distance (FID) and Inception Score (IS) is conflicting with the system's limitations. The improvement observed in these scores as epoch count increases is indicative of the model's tendency to overfit the training data. This overfitting is particularly evident in the enhancement of image quality with extended epochs. Unfortunately, the absence of sufficient resources hinders the optimisation of output quality.

In the future, a potential avenue for improvement in the Inception Score is using a higher-quality dataset and GPU resources. This approach enhances model performance, generates higher-quality outputs, and leads to more accurate evaluation scores for metrics like FID and IS.

A visual inspection will assess the effectiveness of a generative model and the diversity encapsulated within its generated images. Nonetheless, when confronted with a restricted sample size of generated images, these plots yield inconclusive or imprecise results, potentially falling short of accurately depicting the genuine distribution of Inception Scores across the entirety of the generated dataset.

The Gaussian Mixture Model (GMM) is a highly suitable technique for evaluating eczema in dark skin using DCGAN. GMM's capacity to model complex data distributions via multiple Gaussian components effectively captures diverse eczema patterns and colour variations across skin tones. By identifying distinct modes in generated eczema images, GMM portrays nuanced manifestations and accommodates variability. Parameters like means and covariances offer quantitative insights into DCGAN's performance. GMM's mode detection aids in recognising unique eczema traits, and its components indicate image diversity.

But here the GMM output's quality is questionable, evident from the analysis of component means and covariances, prompting concerns. Specifically, the component mean values of 0.46 and -0.06, coupled with the associated covariance values of 0.04 and 0.37, respectively, underscore potential difficulties in accurately capturing the underlying data distribution. Additionally, the performance of the GMM is deeply linked with the quality of both the real and generated datasets. Unfortunately, due to limited GPU resources, conducting essential iterative experiments to enhance the model's performance becomes challenging. This limitation hinders our ability to thoroughly refine the GMM's parameters and configurations, potentially restricting its capacity to generate meaningful insights from the data.

References

Behara, K., Bhero, E., & Agee, J. T. (2023). Skin Lesion Synthesis and Classification Using an Improved DCGAN Classifier. *Diagnostics*, *13*(16), 2635. https://doi.org/10.3390/diagnostics13162635

Chen, R. J., Lu, M. Y., Chen, T. Y., Williamson, D. F. K., & Mahmood, F. (2021). Synthetic data in machine learning for medicine and healthcare. *Nature Biomedical Engineering*, *5*(6), 493–497. https://doi.org/10.1038/s41551-021-00751-8

Chong, M. J., & Forsyth, D. (2020, June 15). *Effectively Unbiased FID and Inception Score and where to find them*. ArXiv.org. https://doi.org/10.48550/arXiv.1911.07023

Eczema - Skin Deep. (2020, August 18). https://dftbskindeep.com/all-diagnoses/eczema/

Fang, W., Zhang, F., S. Sheng, V., & Ding, Y. (2018). A Method for Improving CNN-Based Image Recognition Using DCGAN. *Computers, Materials & Continua*, *57*(1), 167–178. https://doi.org/10.32604/cmc.2018.02356

Jacob, T. (2022, February 25). *Vanishing Gradient Problem, Explained*. KDnuggets. https://www.kdnuggets.com/2022/02/vanishing-gradient-problem.html

Mind the gap : a handbook of clinical signs in black and brown skin / Malone Mukwende, Peter Tamony, Margot Turner. (n.d.). Wellcome Collection. https://wellcomecollection.org/works/ndx5vuhy/items

Mukwende, M., Tamony, P., & Turner, M. (2020). *Mind the Gap*. Black & Brown Skin. https://www.blackandbrownskin.co.uk/mindthegap

Radford, A., Metz, L., & Chintala, S. (2016). *UNSUPERVISED REPRESENTATION LEARNING WITH DEEP CONVOLUTIONAL GENERATIVE ADVERSARIAL NETWORKS*. https://arxiv.org/pdf/1511.06434.pdf