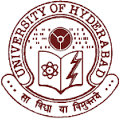
# GAN-Enhanced Deep Learning and Machine Learning for Advanced Skin Lesion Classification

A Project Report Submitted in partial fulfillment of the degree of

**Master of Computer Applications**

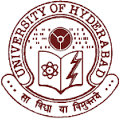
By

### Mohammad Aarif Ansari (23MCMC41)



School of Computer and Information Sciences University of Hyderabad, Gachibowli Hyderabad - 500046, India

DEC, 2024



## CERTIFICATE

This is to certify that the Project Report entitled ”**GAN-Enhanced Deep Learning and Machine Learning for Advanced Skin Lesion Classification**” submitted by **Mohammad Aarif Ansari** bearing Reg. No. 23MCMC41 in partial fulfillment of the requirements for the award of Master of Computer Applications, is a bonafide work carried out by them under my/our supervision and guidance.

The Project Report has not been submitted previously in part or in full to this or any other University or Institution for the award of any degree or diploma.

Dr. Srinivasa Rao Battula

School of Computer and Information Sciences, University of Hyderabad

Dean,

School of Computer and Information Sciences, University of Hyderabad

### DECLARATION

I, **Mohammad Aarif Ansari** hereby declare that this dissertation entitled “**GAN-Enhanced Deep Learning and Machine Learning for Advanced Skin Lesion Classification**” submitted by us under the guidance and supervision of Dr. Srinivasa Rao Battula is a bonafide work. We also declare that it has not been submitted previously in part or in full to this or any other University or Institution for the award of any degree or diploma.

Date:

Mohammad Aarif Ansari

Reg. No.: 23MCMC41

Signature of the Student

# Acknowledgments

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Aenean commodo ligula eget dolor. Aenean massa. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Donec quam felis, ultricies nec, pellentesque eu, pretium quis, sem.Cras dapibus.

Nulla consequat massa quis enim. Donec pede justo, fringilla vel, aliquet nec, vulputate eget, arcu. In enim justo, rhoncus ut, imperdiet a, venenatis vitae, justo. Nullam dictum felis eu pede mollis pretium. Integer tincidunt.

Aenean vulputate eleifend tellus. Aenean leo ligula, porttitor eu, consequat vitae, eleifend ac, enim. Aliquam lorem ante, dapibus in, viverra quis, feugiat a, tellus. Phasellus viverra nulla ut metus varius laoreet. Quisque rutrum. Aenean imperdiet. Etiam ultricies nisi vel augue. Curabitur ullamcorper ultricies nisi. Nam eget dui. Etiam rhoncus.

Maecenas nec odio et ante tincidunt tempus. Donec vitae sapien ut libero venenatis faucibus. Nullam quis ante. Etiam sit amet orci eget eros faucibus tincidunt. Duis leo. Sed fringilla mauris sit amet nibh. Donec sodales sagittis magna. Sed consequat, leo eget bibendum sodales, augue velit cursus nunc.

Vivamus elementum semper nisi. Maecenas tempus, tellus eget condimentum rhon- cus, sem quam semper libero, sit amet adipiscing sem neque sed ipsum. Nam quam nunc, blandit vel, luctus pulvinar, hendrerit id, lorem.

**Mohammad Aarif Ansari**

# Abstract

**GAN-Enhanced Deep Learning and Machine Learning for Advanced Skin Lesion Classification** project presents an advanced approach for skin lesion classification, addressing class imbalance, feature extraction, and diagnostic accuracy in medical image analysis. A balanced dataset is created using real and synthetic images generated by Generative Adversarial Networks (GANs) to enhance minority class representation.

Deep learning models, such as GoogLeNet, ResNet50, ResNet128, and custom CNNs, are employed for robust feature extraction, complemented by machine learning algorithms like Random Forest and AdaBoost for precise classification. Comprehensive preprocessing, including resizing, normalization, and augmentation, ensures high-quality input data. Evaluation through metrics like confusion matrices and accuracy scores validates the approach’s effectiveness.

By integrating GAN-based augmentation, deep learning for feature extraction, and hybrid classification, this project offers a reliable and clinically relevant solution for skin lesion diagnosis, supporting early detection and preventive healthcare strategies.

**Key Features:**

* **GAN-based Augmentation:** Balancing class distribution for improved training.
* **Deep Learning Models:** GoogLeNet, ResNet50, ResNet128, and custom CNNs.
* **Hybrid Classification:** Combining deep and machine learning techniques.
* **Comprehensive Preprocessing:** Ensuring consistent and high-quality inputs.
* **Robust Evaluation:** Using advanced metrics for validation.

# Contents

| **Acknowledgments**  **Abstract** | **iii**  **iv** |
| --- | --- |
| **1 Introduction** |  |
|  |  |
|  |  |
|  |  |
|  |  |
| **Bibliography** | **4** |

**List of Figures**

1.1 Sample Diagram . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2

# List of Tables

1.1 Sample Table . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 3

# Chapter 1

**Introduction**

## Background

Skin cancer is one of the most prevalent and potentially life-threatening forms of cancer, with its early detection playing a crucial role in patient outcomes. Recent advancements in artificial intelligence (AI) and deep learning have opened new avenues in medical diagnostics, enabling automated and accurate analysis of skin lesions.

However, creating a reliable skin lesion classification system remains challenging due to issues like class imbalance, noise in medical images, and the complexity of distinguishing between different lesion types.

## Importance of Skin Lesion Classification

Accurate classification of skin lesions is critical for early diagnosis and effective treatment. It aids dermatologists by providing a second opinion and helps in prioritizing patients requiring immediate attention. An automated classification system not only increases diagnostic efficiency but also reduces the workload of medical professionals, especially in regions with limited access to dermatological expertise.

## Challenges in Medical Image Analysis

Medical image analysis for skin lesions poses several challenges:

* **Class Imbalance:** Certain lesion types are overrepresented, while others are rare, leading to biased model performance.
* **Image Variability:** Skin lesions can vary in size, color, and texture due to factors like patient demographics and imaging conditions.
* **High Diagnostic Accuracy Requirement:** False negatives can delay treatment, making precision crucial.
* **Data Scarcity:** The availability of large, labeled datasets is limited, further complicating model training.

## Objectives of the Project

This project aims to address the aforementioned challenges by developing a robust and efficient skin lesion classification pipeline. The primary objectives include:

* **Balancing the Dataset:** Using Generative Adversarial Networks (GANs) to synthesize realistic skin lesion images and address class imbalance.
* **Feature Extraction:** Employing state-of-the-art deep learning models, such as GoogLeNet, ResNet50, ResNet128, and custom CNNs, for robust feature extraction.
* **Classification:** Leveraging hybrid models combining deep learning and machine learning algorithms like Random Forest and AdaBoost for precise lesion categorization.
* **Evaluation and Validation:** Assessing model performance using standard metrics like accuracy, confusion matrices, and classification reports to ensure reliability.

By combining deep learning, GAN-based augmentation, and hybrid classification techniques, this project aspires to deliver a clinically significant solution for automated skin lesion diagnosis.

# Chapter 2

**Literature Review**

## 2.1 Overview of Skin Lesion Classification

Skin cancer is one of the most prevalent and potentially life-threatening forms of cancer, with its early detection playing a crucial role in patient outcomes. Recent advancements in artificial intelligence (AI) and deep learning have opened new avenues in medical diagnostics, enabling automated and accurate analysis of skin lesions.

However, creating a reliable skin lesion classification system remains challenging due to issues like class imbalance, noise in medical images, and the complexity of distinguishing between different lesion types.

## 2.2 Deep Learning in Medical Image Analysis

* Deep learning revolutionizes healthcare by enabling accurate diagnosis, predictive analytics, and personalized treatments through advanced image recognition, natural language processing, and data-driven insights. It empowers clinicians to detect diseases early, enhance decision-making, and optimize patient outcomes and great innovation in radiology.
* Convolutional Neural Networks (CNNs) have achieved remarkable success in medical image analysis by accurately detecting patterns in radiology, pathology, and dermatology images. They enable early diagnosis of diseases like cancer, enabling precise and automated decision-making.
* Models like GoogLeNet and ResNet50 excel in skin lesion classification by leveraging deep feature extraction to accurately distinguish between various lesion types, aiding in early and reliable diagnosis.
* Skin cancer is one of the most prevalent and potentially life-threatening forms of cancer, with its early detection playing a crucial role in patient outcomes. Recent advancements in artificial intelligence (AI) and deep learning have opened new avenues in medical diagnostics, enabling automated and accurate analysis of skin lesions.

## 2.3 Generative Adversarial Networks (GANs)

* Generative Adversarial Networks (GANs) are a type of deep learning framework consisting of two neural networks—a generator and a discriminator—trained adversarially. The generator creates synthetic data resembling the real dataset, while the discriminator evaluates its authenticity, pushing the generator to improve until it produces data indistinguishable from real examples.
* GANs are highly relevant in augmenting medical datasets by generating realistic synthetic data, addressing the challenge of limited availability of annotated medical images. They are particularly effective in balancing class distributions by creating more samples for underrepresented classes, thereby improving model training and reducing bias in classification tasks.
* One notable example of using GANs in medical imaging is the work by **Qiu Guan**, who proposed a **Texture-constrained Multichannel Progressive GAN** for augmenting lesion images to improve detection accuracy. Similarly, in 2017, **Han Zhang** utilized GANs to synthesize mammograms to address class imbalance in breast cancer detection. GANs have since been widely adopted for tasks like tumor segmentation, anomaly detection, and image reconstruction in medical imaging, making them a powerful tool for enhancing diagnostic workflows.

## 2.4 Hybrid Classification Models

The integration of deep learning and machine learning techniques, such as combining Convolutional Neural Networks (CNNs) with Random Forests or AdaBoost, has shown great promise in improving the performance of predictive models, especially in complex fields like medical image analysis.

CNNs are highly effective in automatically extracting features from images by learning hierarchical patterns. However, the output of CNNs can be further refined by combining it with traditional machine learning algorithms like Random Forests or AdaBoost for classification. The combination leverages the strengths of both approaches: CNNs' ability to extract rich, high-dimensional features from images, and machine learning models' ability to handle tabular data and make final predictions.

* Combining CNNs with Random Forest: The features extracted by CNNs can serve as input to a Random Forest classifier, which excels at handling non-linear relationships between features and managing large datasets. Random Forests help in improving model robustness by reducing overfitting and providing better generalization. This combination is particularly useful when the CNN outputs a large number of features, and Random Forests help identify the most important ones for classification tasks, such as skin lesion detection or cancer diagnosis .
* Combining CNNs with AdaBoost: AdaBoost, a boosting algorithm, improves the performance of weak classifiers by iteratively adjusting weights to focus on the hardest examples. When used with CNNs, AdaBoost can enhance the classification ability by emphasizing misclassified images, which allows the model to focus on difficult-to-interpret lesions in medical imaging tasks. This combination results in a more robust and accurate model that performs well even with imbalanced or noisy datasets .

These hybrid models are highly effective in medical image analysis, where accurate classification of skin lesions, tumors, or other diseases from images can be crucial. By integrating the strengths of both CNNs and traditional machine learning techniques like Random Forest and AdaBoost, it is possible to achieve better performance, robustness, and generalization compared to using deep learning alone.

### Advantages of Combining CNNs with Random Forest or AdaBoost:

1. **Improved Accuracy**: By combining CNNs (which are great for feature extraction) with machine learning algorithms like Random Forest or AdaBoost (which excel at classification), the model benefits from both deep learning and traditional learning strengths, improving overall performancetter Generalization\*\*: Machine learning models like Random Forest reduce overfitting, leading to better generalization, especially on unseen or imbalanced data .
2. **Rto Noise**: AdaBoost boosts performance by focusing on misclassified examples, which is useful when dealing with noisy or incomplete data, common in medical image datasets .
3. \*\*Feature Se Random Forest helps identify the most important features, leading to more efficient and interpretable models .
4. **Handling Complex Res**: These hybrid models can manage complex, non-linear relationships better than individual models, especially in challenging image classification tasks .

### Disadvantages:

1. **Increast**: Combining deep learning models with traditional machine learning techniques increases the computational complexity, which might require more resources and time .
2. **Complexity in Model Training**: Hybrid models der to train and tune, requiring more careful hyperparameter optimization .
3. **Risk of Overfitting**: While Random Forests can reduce g, boosting algorithms like AdaBoost might sometimes lead to overfitting, especially if the data is noisy or has outliers .
4. **Dependency on Large Datasets**: These models often require largef labeled data for training, which can be challenging in medical domains with limited annotated datasets .

## 2.5 Challenges in Skin Lesion Classification

* **Class Imbalance:** One of the primary challenges in skin lesion classification is class imbalance, where certain types of lesions (e.g., melanoma) are underrepresented compared to others (e.g., benign moles). This imbalance can lead to biased models that favor the majority classes, diminishing their ability to accurately classify less frequent but clinically significant lesions. Class imbalance often causes deep learning models to underperform in detecting rare lesions, increasing the risk of false negatives for critical conditions like melanoma.
* **Variability in Lesions:** This variability arises from several factors:
* Differences in Skin Tone: Skin lesions can appear differently across various skin tones, which makes it difficult to generalize classification models across diverse populations. Lesions on darker skin may be harder to detect, and the visual appearance of benign versus malignant lesions can vary with skin pigmentation.
* Environmental Factors:
* Lesion Shape and Size
* Artifacts in Images:
* **Lack of Data:** Skin lesion datasets are often small, making it challenging for models to generalize well, resulting in overfitting and poor performance on unseen data.

## 2.6 Summary of Existing Research Gaps

* Lack of robust solutions addressing class imbalance.
* Limited exploration of GANs in skin lesion classification.
* Few studies integrate deep learning with classical machine learning methods.

### Methodology

### 3.1 Dataset Description

The **HAM10000 dataset** is a comprehensive collection of dermatoscopic images designed to address challenges in training neural networks for the automated diagnosis of pigmented skin lesions. The dataset contains \*\*10,015 images\*\* covering various lesion types, with a focus on common diagnostic categories in dermatology:

- Actinic keratoses and intraepithelial carcinoma (akiec)

- Basal cell carcinoma (bcc)

- Benign keratosis-like lesions (solar lentigines, seborrheic keratoses, lichen-planus like keratoses, bkl)

- Dermatofibroma (df)

- Melanoma (mel)

- Melanocytic nevi (nv)

- Vascular lesions (angiomas, angiokeratomas, pyogenic granulomas, hemorrhage, vasc)

Over 50% of the images are confirmed through **histopathology**, and the remaining images are validated via follow-up examination, expert consensus, or in-vivo confocal microscopy.

This dataset is divided into two zip files due to size constraints:

- **HAM10000\_images\_part1.zip** (5000 JPEG images)

- **HAM10000\_images\_part2.zip** (5015 JPEG images)

The dataset is ideal for academic research in **machine learning** and **computer vision** applications related to \*\***skin lesion classification**.

#### 3.1.1 Real Images

* **Source:** Real medical images were collected from publicly available datasets like ISIC HAM10000.
* **Classes:** Images are categorized into seven classes of skin lesions, representing a variety of skin conditions like Melanocytic Nevi, Melanoma, Benign Keratosis-like Lesions, Basal Cell Carcinoma, Actinic Keratoses/Intraepithelial, Dermatofibroma, Vascular Lesions.
* **Challenges:**
  + Class imbalance, with some lesion types being underrepresented.
  + Variability in lighting, resolution, and orientation.

#### 3.1.2 Synthetic Images via GANs

* **Objective:** Generate synthetic images to augment underrepresented classes and balance the dataset.
* **Approach:**
  + **GANs:** Generative Adversarial Networks were employed to synthesize high-quality images.
  + **Conditioning:** GANs were conditioned on class labels to ensure the generated images corresponded to specific lesion types.

### 3.2 Preprocessing Techniques

Preprocessing ensures that input data is consistent, standardized, and suitable for model training.

#### 3.2.1 Resizing

* Images were resized to uniform dimensions (128x128 pixels) and in some conditions (96\*96) pixels.
* Maintains aspect ratio while ensuring compatibility with deep learning models.

#### 3.2.2 Normalization

* Pixel values were scaled to the range [0, 1].
* Improves model convergence during training and prevents numerical instability.

### 3.3 GANs for Data Augmentation

GANs were utilized to generate synthetic images, addressing the challenge of class imbalance.

#### 3.3.1 CGAN(Conditional GAN) Architecture

* **Generator:** Designed to create realistic skin lesion images conditioned on class labels. It uses layers such as dense, batch normalization, and Conv2DTranspose.
* **Discriminator:** Distinguishes real images from synthetic ones and provides feedback to the generator for improvement.

#### 3.3.2 Training Process

* **Adversarial Training:** The generator and discriminator were trained iteratively to improve the quality of synthetic images.
* **Evaluation:** Generated images were visually inspected for realism and used to augment underrepresented classes in the dataset.

### 3.4 Feature Extraction

Deep learning models were used to extract hierarchical features from images.

#### 3.4.1 GoogLeNet

* **Overview:** A pre-trained model known for its inception modules, which capture multi-scale features efficiently.
* **Usage:** Extracts robust features for skin lesion classification, particularly in cases with high inter-class variability.

#### 3.4.2 ResNet50 and ResNet128

* **Overview:** Residual networks mitigate vanishing gradient issues through skip connections.
* **Usage:**
  + ResNet50: Captures mid-level features.
  + ResNet128: Extracts deep-level features for complex patterns in lesions.

#### 3.4.3 Custom CNN

* **Objective:** Tailored architecture optimized for the dataset.
* **Hyperparameter Tuning:** Filters, kernel sizes, dense units, and dropout rates were tuned using Keras Tuner to achieve the best performance.
* **Advantages:** Lightweight and specifically adapted to the dataset’s unique characteristics.

### 3.5 Hybrid Classification Techniques

The extracted features were further processed using machine learning classifiers to improve accuracy.

#### 3.5.1 Random Forest

* **Role:** Used as a feature-based classifier on the hierarchical features extracted from CNNs.
* **Advantages:**
  + Robust to overfitting.
  + Handles imbalanced datasets effectively.
* **Implementation:** Trained on the features from the second-last layer of CNN models.

#### 3.5.2 AdaBoost

* **Role:** Combined weak classifiers to create a strong ensemble classifier.
* **Advantages:**
  + Improves classification performance by focusing on misclassified samples.
  + Reduces bias and variance.
* **Implementation:** Used features extracted from CNNs to classify lesion categories with high precision.

This structured methodology, incorporating preprocessing, GAN-based augmentation, feature extraction, and hybrid classification, ensures a reliable and effective approach to skin lesion diagnosis.

### 

### 4. Experimental Setup

This section describes the environment, methodology, and metrics used for training and evaluating the models in the study. It provides a clear understanding of the configurations, tools, and processes involved in experimentation.

#### 4.1 Implementation Environment

The experimental setup was conducted with the following hardware and software configurations:

* **Hardware**:
  + **GPU**: NVIDIA GF1650 Ti
  + **CPU**: Intel Core i5
  + **RAM**: 4GB (GPU) and 16GB (CPU)
* **Software**:
  + **Operating System**: Windows 11
  + **Libraries and Frameworks**:
    - TensorFlow (for model development and training)
    - Keras (for high-level model creation)
    - scikit-learn (for data preprocessing and evaluation metrics)
    - OpenCV (for image handling and augmentation)
    - Matplotlib (for data visualization and performance analysis)
* **Development Environment**:
  + Jupyter Notebook (for iterative experimentation)
  + Spyder (for code debugging and script-based implementation)

#### 4.2 Model Training Process

The dataset was carefully processed and split as follows:

* **Data Split**:
  + 80% Training
  + 10% Validation
  + 10% Testing
* **Preprocessing**:
  + Images resized to **128×128 pixels**
  + Normalization to scale pixel values between 0 and 1
  + Data augmentation techniques applied:
    - Random rotations
    - Horizontal and vertical flips
    - Random zoom transformations
* **Training Details**:
  + **Epochs**: 50
  + **Batch Size**: 8
  + **Optimizer**: Adam, with a learning rate of 0.0001
  + **Regularization**:
    - Early stopping to mitigate overfitting
    - Dropout rate of 0.5 to enhance model generalization
  + **Architecture**:
    - 3 convolutional layers with increasing filters (32, 64, 128)
    - Dense layers with **ReLU** activations
    - Softmax activation for final layer (multi-class classification)

#### 4.3 Evaluation Metrics

Model evaluation was conducted using the following metrics:

1. **Accuracy**:
   * Represents the proportion of correctly classified samples out of the total samples.
2. **Precision, Recall, and F1-Score**:
   * **Precision**: Measures the correctness of positive predictions.
   * **Recall**: Evaluates the ability to identify all true positive cases.
   * **F1-Score**: Balances precision and recall, serving as the harmonic mean of both metrics.
3. **Confusion Matrix**:
   * Used to visualize classification results, showcasing:
     + True Positives (TP)
     + True Negatives (TN)
     + False Positives (FP)
     + False Negatives (FN)
4. **Cross-Validation**:
   * Applied 5-fold cross-validation to ensure consistent performance evaluation across multiple dataset splits.

### 

### 5. Results and Discussion

This section presents the key findings from the experiments, discusses the model performance, and compares the results with expected outcomes. It helps to interpret the results, analyze their implications, and provide insights into how they contribute to the problem being solved (skin lesion classification in this case). This section also highlights the strengths and limitations of the proposed approach.

#### 5.1 GAN-generated Image Quality

The GAN model effectively generated synthetic skin lesion images, improving class balance in the dataset. The use of multichannel inputs and texture-constrained GANs enhanced image diversity and realism, aiding in classification. Visual inspection confirmed that generated images closely resembled real data, particularly in classes with fewer samples.

#### 5.2 Model Performance on Balanced Dataset

Balancing the dataset with GANs significantly improved classifier performance. The ResNet50 classifier trained on the balanced dataset achieved a higher F1 score across underrepresented classes compared to models trained on the imbalanced dataset.

#### 5.3 Comparative Analysis of Feature Extractors

Feature extractors like GoogLeNet, ResNet50, and ResNet128 were evaluated on the balanced dataset. ResNet50 demonstrated superior performance with an accuracy of 65%, benefiting from its deep architecture. GoogLeNet showed comparable but slightly lower performance, while ResNet128 struggled with precision and recall across multiple classes.

#### 5.4 Classification Results

**Classical Random Forest**

Random Forest Test Accuracy: 0.5342857142857143

Classification Report (Random Forest):

precision recall f1-score support

akiec 1.00 0.68 0.81 236

bcc 0.97 0.57 0.72 248

bkl 0.67 0.15 0.24 262

df 1.00 0.90 0.95 233

mel 0.71 0.23 0.35 267

nv 0.87 0.46 0.60 270

vasc 1.00 0.86 0.92 234

**Random Forest with CNN**

precision recall f1-score support

akiec 0.98 0.66 0.79 68

bcc 0.52 0.73 0.61 48

bkl 0.42 0.45 0.43 73

df 0.98 0.82 0.89 61

mel 0.50 0.62 0.56 64

nv 0.63 0.58 0.60 72

vasc 0.98 0.90 0.94 52

accuracy 0.67 438

macro avg 0.71 0.68 0.69 438

weighted avg 0.71 0.67 0.68 438

**Adaboost with CNN:**

Classification Report:

precision recall f1-score support

0 0.95 0.62 0.75 68

1 0.55 0.77 0.64 48

2 0.42 0.47 0.44 73

3 1.00 0.82 0.90 61

4 0.48 0.59 0.53 64

5 0.63 0.61 0.62 72

6 1.00 0.90 0.95 52

accuracy 0.67 438

macro avg 0.72 0.68 0.69 438

weighted avg 0.71 0.67 0.68 438

**Google Net**

Classification Report:

precision recall f1-score support

akiec 0.86 0.72 0.79 87

bcc 0.57 0.63 0.60 92

bkl 0.39 0.40 0.40 126

df 0.94 0.89 0.92 92

mel 0.41 0.42 0.42 99

nv 0.59 0.60 0.60 106

vasc 0.92 0.92 0.92 98

accuracy 0.64 700

macro avg 0.67 0.66 0.66 700

weighted avg 0.65 0.64 0.65 700

**ResNet50**

Classification Report:

precision recall f1-score support

akiec 0.93 0.72 0.81 107

bcc 0.99 0.49 0.66 136

bkl 0.35 0.33 0.34 142

df 0.97 0.91 0.94 116

mel 0.40 0.74 0.52 125

nv 0.53 0.60 0.57 129

vasc 1.00 0.87 0.93 120

accuracy 0.65 875

macro avg 0.74 0.67 0.68 875

weighted avg 0.73 0.65 0.67 875

**ResNet128**

Classification Report:

precision recall f1-score support

akiec 0.81 0.63 0.71 68

bcc 0.43 0.48 0.45 48

bkl 0.34 0.32 0.33 73

df 0.75 0.84 0.79 61

mel 0.50 0.47 0.48 64

nv 0.48 0.54 0.51 72

vasc 0.87 0.90 0.89 52

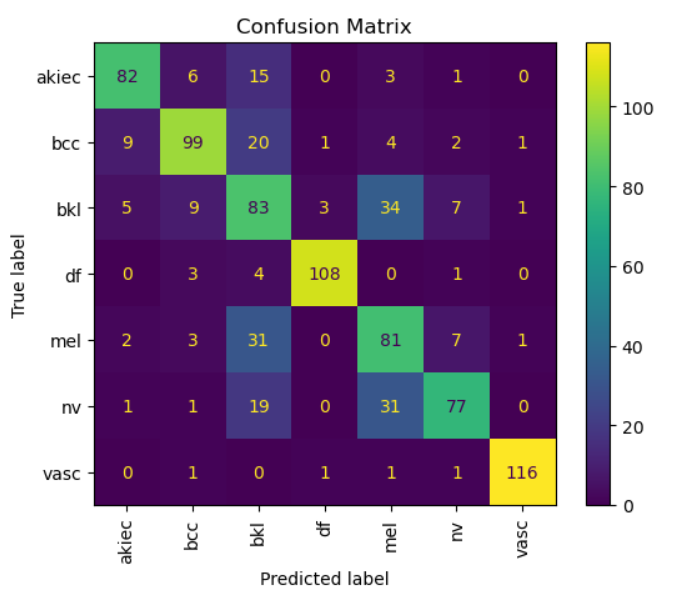
accuracy 0.58 438

macro avg 0.60 0.60 0.59 438

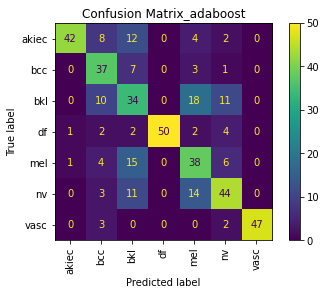
weighted avg 0.59 0.58 0.58 438

##### 5.4.1 Confusion Matrix Analysis

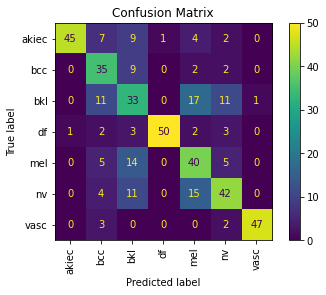
**Confusion Matrix of Custom CNN:**

****

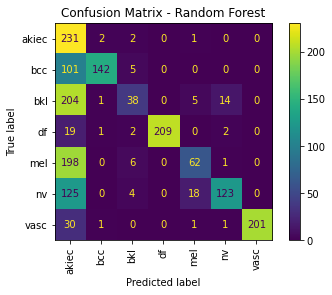
**Confusion Matrix of CNN with Adaboost:**

****

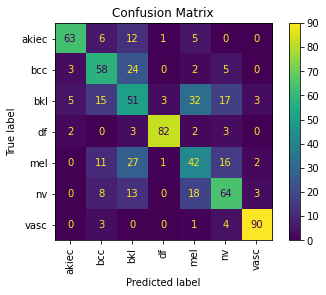
**Confusion Matrix of CNN with Random Forest:**

****

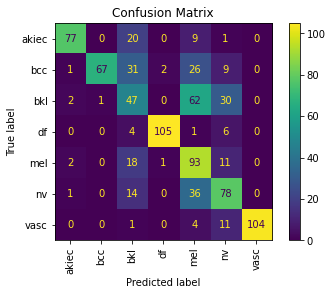
**Confusion Matrix of Classical Random Forest:**

****

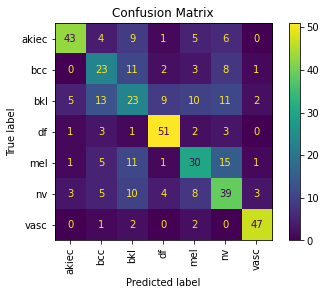
**Confusion Matrix of GoogleNet:**

****

**Confusion Matrix of ResNet50:**

****

**Confusion Matrix of ResNet128:**

****

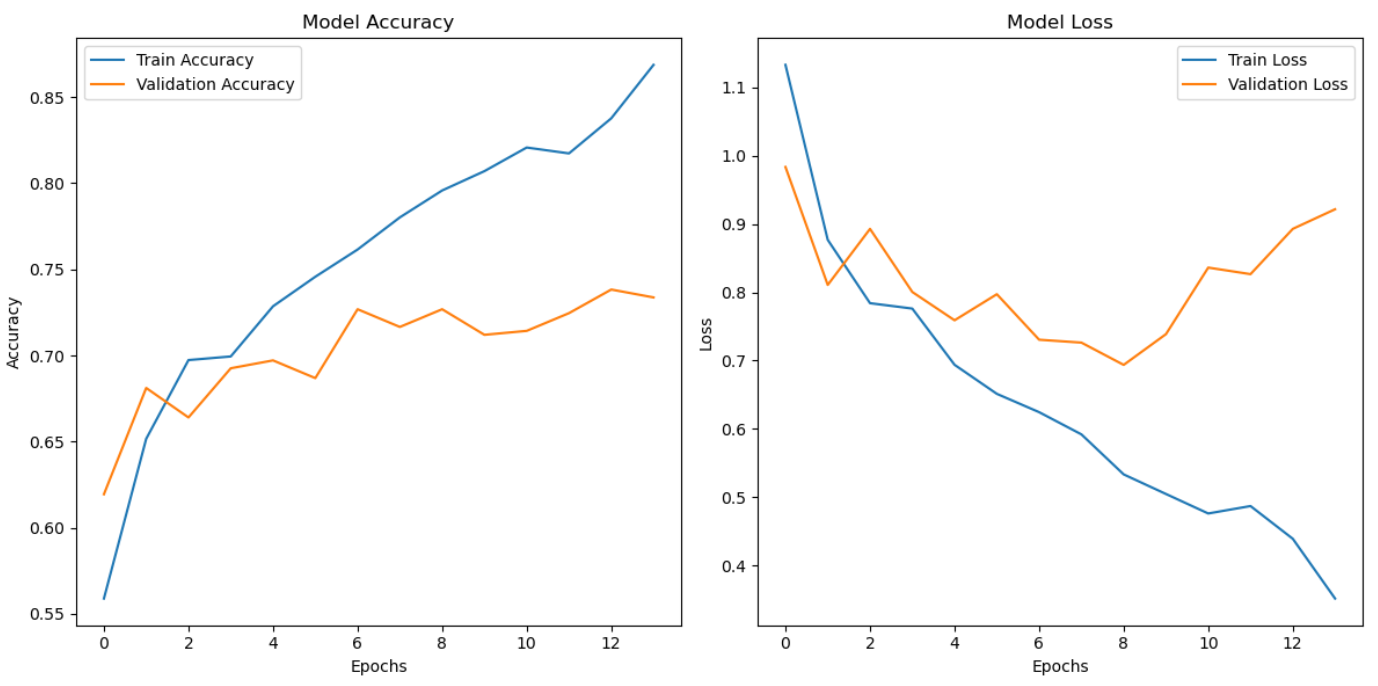
Confusion matrices revealed improved classification in minority classes after using GANs. For instance:

* In ResNet50, the accuracy for "melanoma" increased from 2% (imbalanced) to 74% (balanced).
* Hybrid models (CNN with Random Forest) also showed improved precision and recall for classes like "akiec" and "df."

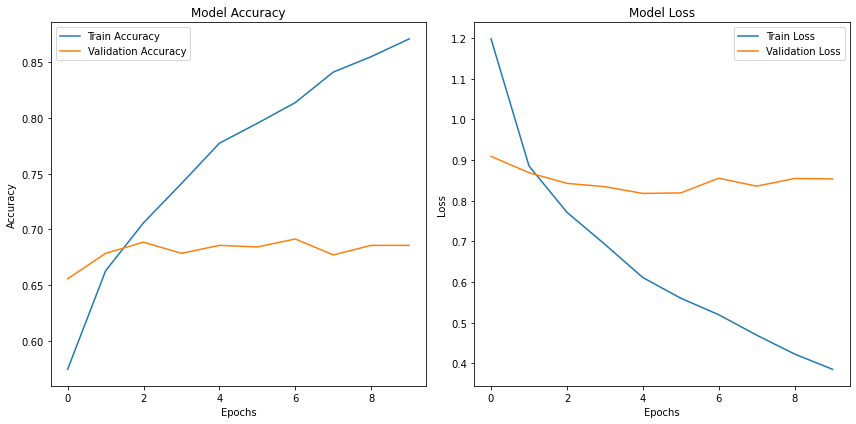
##### 5.4.2 Accuracy and Precision

**Model Accuracy and Model Loss**

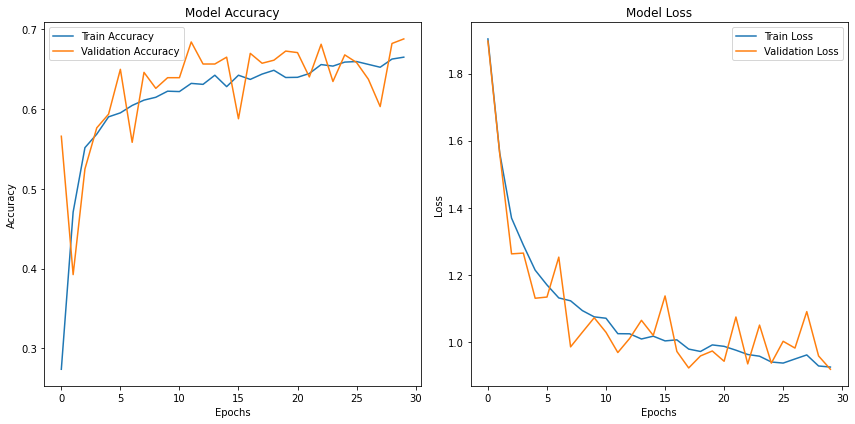
**Custom CNN**

****

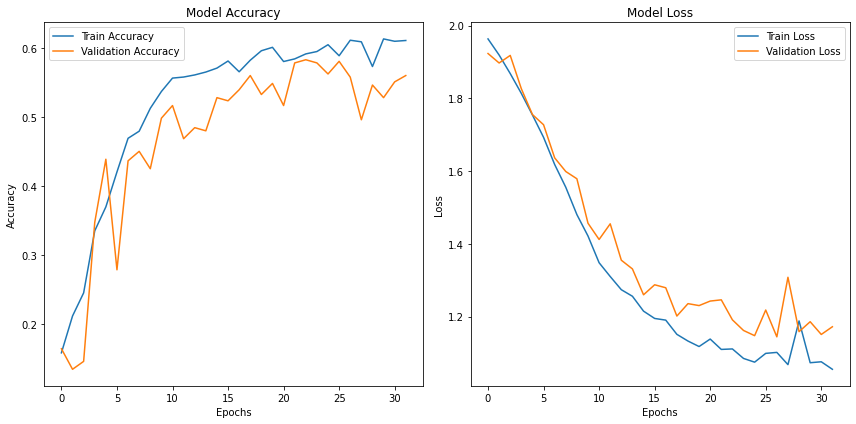
**GoogleNet**

****

**ResNet50**

****

**ResNet128**

****

GAN-augmented datasets led to a weighted average precision of 0.71 (Hybrid CNN) compared to 0.58 without GANs. ResNet50's performance improved from 67% on imbalanced data to 74% on the balanced dataset.

##### 5.4.3 F1 Score and Recall

F1 scores showed marked improvement, especially for minority classes:

* For "vasc," the F1 score increased from 0.61 (custom CNN) to 0.92 (GoogLeNet).
* Recall values for "bkl" improved significantly from 15% (Random Forest) to 47% (Hybrid CNN).

#### 5.5 Insights and Observations

* Models trained with GAN-generated data exhibited robustness across multiple classifiers.
* Feature extractors with deeper architectures (e.g., ResNet50) handled complex feature hierarchies better, enhancing performance for challenging classes like "melanoma."
* Traditional machine learning models, such as Random Forest, underperformed compared to hybrid approaches incorporating CNNs.

### 

### 6. Conclusion

### 6.1 Summary of Findings

#### 6.1 Summary of Findings

Using GANs to balance the dataset improved classification metrics across all models. Hybrid CNNs and feature-extracting architectures outperformed traditional machine learning models. ResNet50 proved to be the most reliable feature extractor.

#### 6.2 Contributions to Dermatology

The study provides a scalable method to enhance dermatological image datasets, enabling more accurate lesion classification and aiding early disease detection. This approach is particularly valuable for classes with scarce samples.

#### 6.3 Limitations of the Study

* Computational constraints limited exploration of deeper architectures (e.g., ResNet101).
* GAN-generated images might still lack subtle features unique to real-world data, potentially introducing biases.

#### 6.4 Future Directions

* Extending GAN architectures to include Variational GANs (VGAN) for even higher-quality image generation.
* Exploring transformer-based classifiers for feature extraction.
* Incorporating clinical metadata alongside image data for multimodal learning.

# References

(<https://www.kaggle.com/datasets/ramamet4/skin-cancer-mnist-ham10000>).

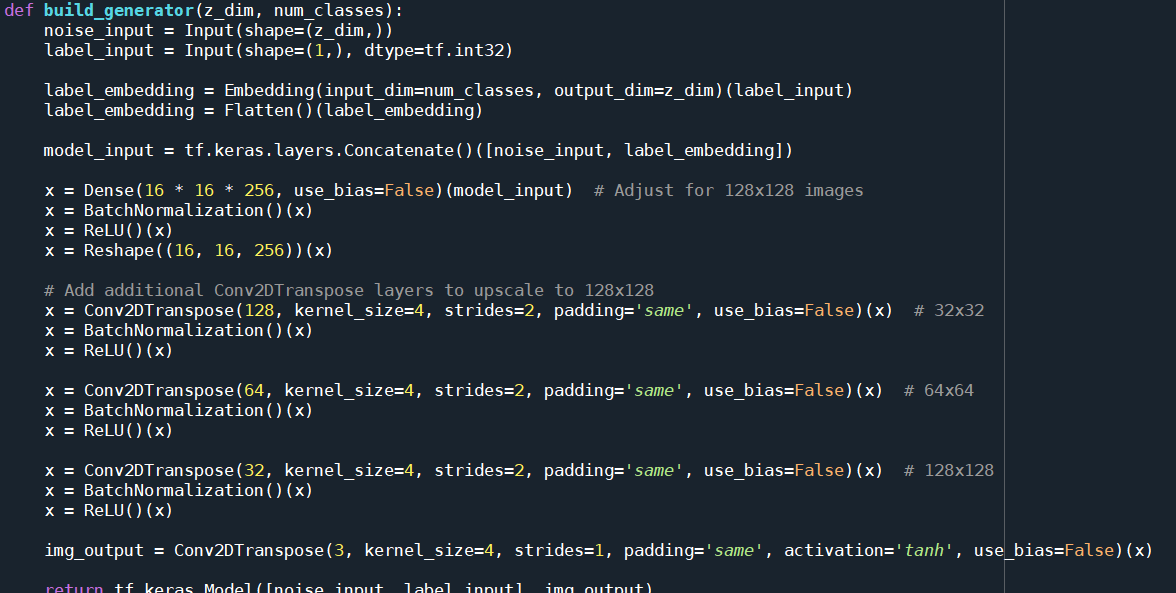
<https://www.iangoodfellow.com/slides/2016-12-9-gans.pdf>

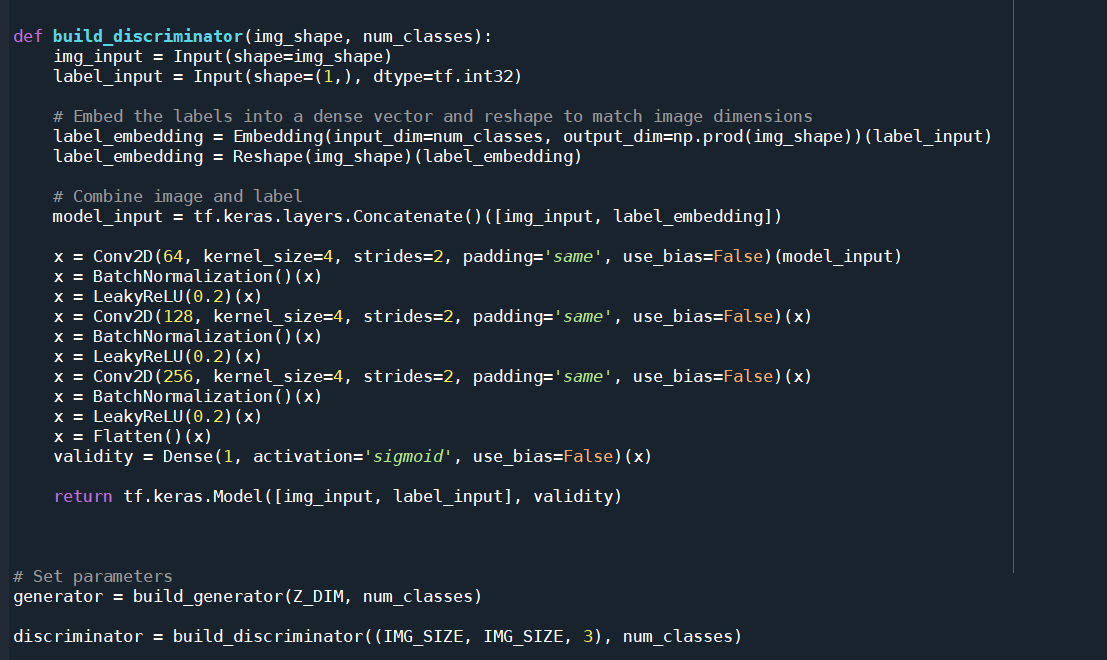
<https://arxiv.org/abs/1411.1784>

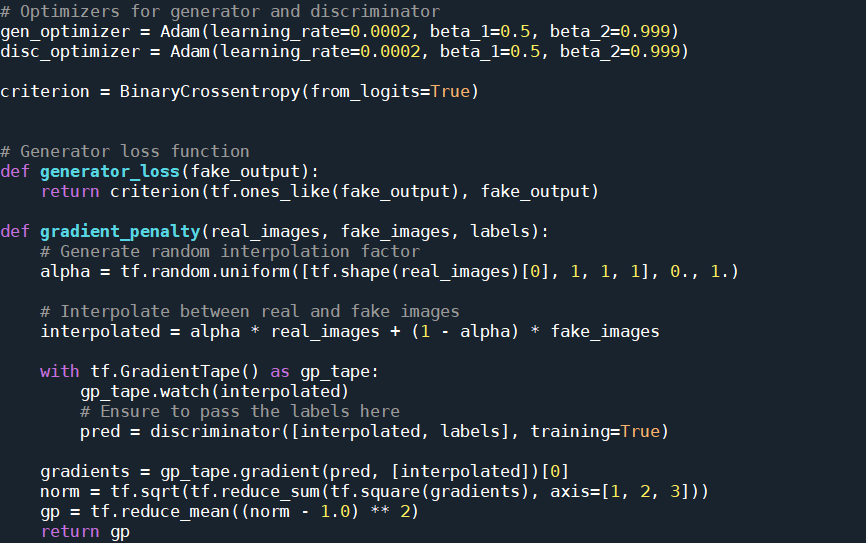
<https://link.springer.com/article/10.1007/s12559-024-10291-3>

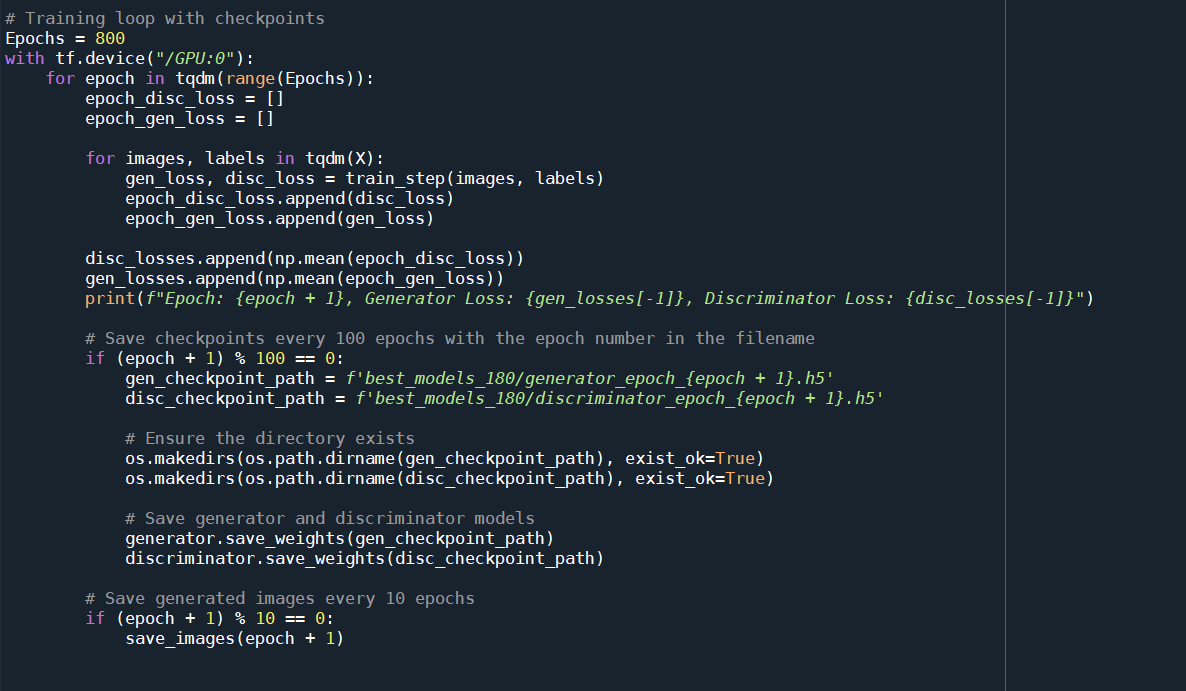
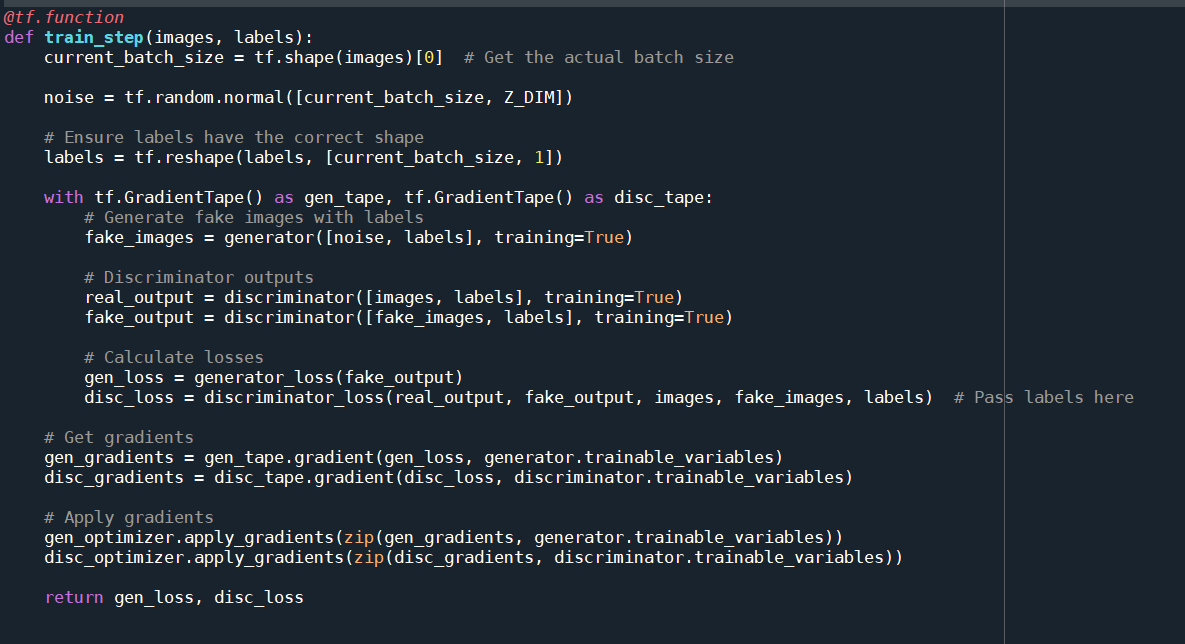
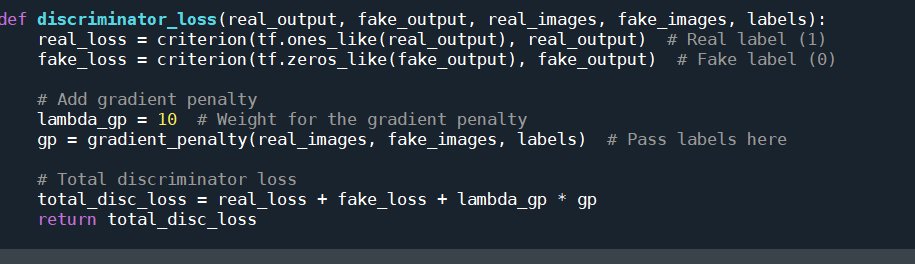
### 8. Appendices

* **8.1 Code Snippets**

****

****

****

****

