

Report:

Change normalization parameter on images refers to altering the parameters used in the normalization process of an image to intentionally induce misclassification in a machine learning model.

Normalization is a pre-processing step commonly applied to images before training a deep learning model. It involves transforming the pixel values of an image to a standardized scale, typically with a mean of zero and standard deviation of one. This helps the model learn more efficiently and effectively by reducing the impact of variations in lighting and color on the image.

To induce misclassification, one could intentionally alter the normalization parameters of an image in a way that causes the model to misinterpret the image. For example, changing the mean value of a set of images of dogs to that of cats could cause a model trained on cat images to misclassify a dog image as a cat. Similarly, changing the standard deviation of the pixel values could also cause a misclassification.

This technique of manipulating normalization parameters to induce misclassification is a type of adversarial attack, and it can be used to test the robustness of machine learning models or to explore the vulnerabilities of image classifiers. It is important to note that while these attacks can be useful for testing, they can also be maliciously used to deceive and exploit machine learning systems.

I use 4 different sets of values for the mean and std parameters and check the misclassified images. I plotted GradCAM for visualization of these effects.

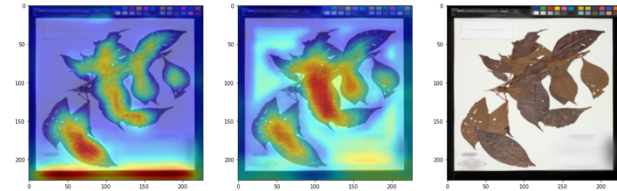
- 1- With normalization value mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225], testing acc = 75.5
- 2- With normalization value mean=[0.8, 0.8, 0.8], std=[0.335, 0.345, 0.355] , testing acc = 69.38
- 3- With normalization value mean=[0.9, 0.9, 0.9], std=[0.1, 0.1, 0.1], testing acc = 81.63
- 4- With normalization value mean=[0.2, 0.2, 0.2], std=[0.3, 0.3, 0.3], testing acc = 81.63

There are some images that are misclassified always. For example, you can see one of those:

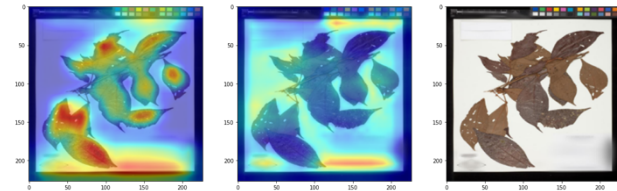
Note: The first column is predicted class GradCAM, the second one is real class GradCAM, and 3rd column is original image.

Real class name: Miconia approximata Gamba & Almeda:

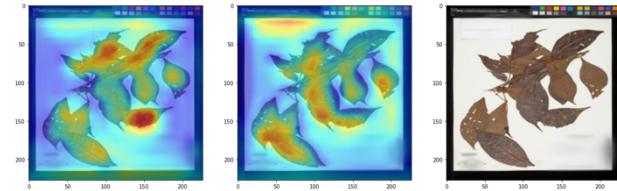
Detecting for class Miconia ibaguensis (Bonpl.) Triana with 0.867 for first value set



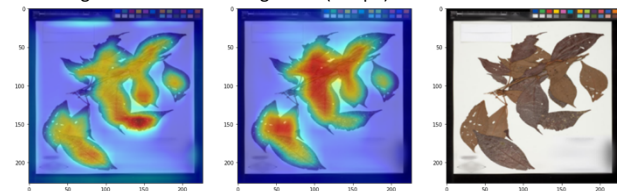
Detecting for class Miconia racemosa (Aubl.) DC. With 1.0 for 2nd value set



Detecting for class Miconia ibaguensis (Bonpl.) Triana with 0.999 for 3rd value set



Detecting for class Miconia ibaguensis (Bonpl.) Triana with 0.999 for 4rd value set

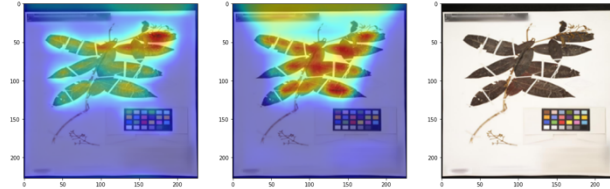


Some images are misclassified with the unusual normalization value that is classified correctly with the normal parameters, like:

This image was misclassified with the 2nd of normalization value set:

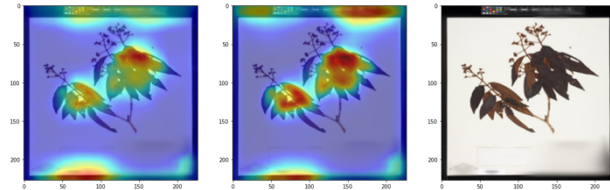
real class name: *Miconia ibaguensis* (Bonpl.) Triana

image 00000.jpg Detecting for class *Miconia racemosa* (Aubl.) DC. With 0.958



real class name: *Miconia ibaguensis* (Bonpl.) Triana

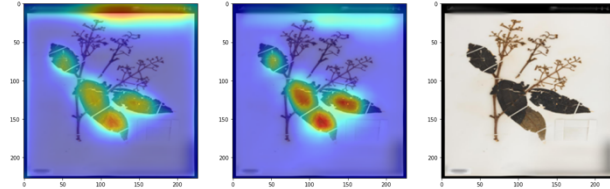
image 00003.jpg Detecting for class *Miconia racemosa* (Aubl.) DC. with 0.801



This image was misclassified with the 3rd of normalization value set:

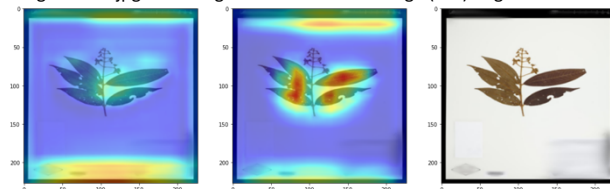
real class name: *Miconia ibaguensis* (Bonpl.) Triana

image 00010.jpg Detecting for class *Miconia racemosa* (Aubl.) DC. with 0.784



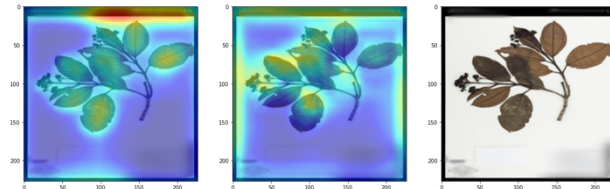
real class name: *Miconia ibaguensis* (Bonpl.) Triana

image 00013.jpg Detecting for class *Leandra nianga* (DC.) Cogn. with 0.631



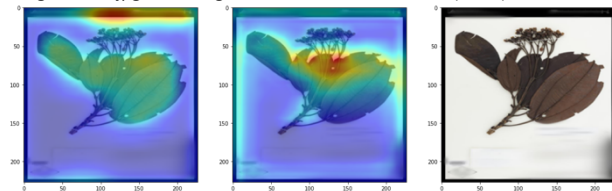
real class name: *Miconia corymbiformis* Cogn. (all of images in these class are misclassified)

image 00000.jpg Detecting for class *Miconia racemosa* (Aubl.) DC. with 0.876



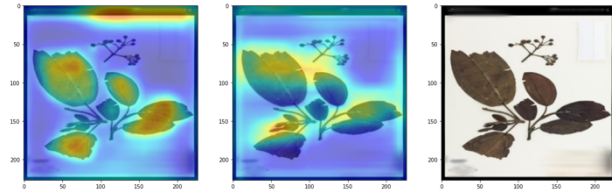
real class name: *Miconia corymbiformis* Cogn.

image 00001.jpg Detecting for class *Miconia racemosa* (Aubl.) DC. with 0.559



real class name: *Miconia corymbiformis* Cogn.

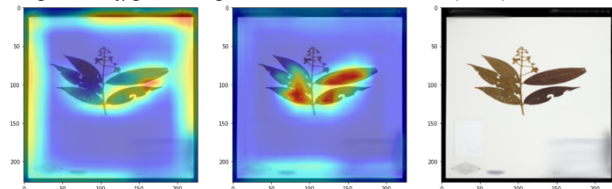
image 00002.jpg Detecting for class *Miconia racemosa* (Aubl.) DC. with 0.997



This image was misclassified with the 4rd of normalization value set:

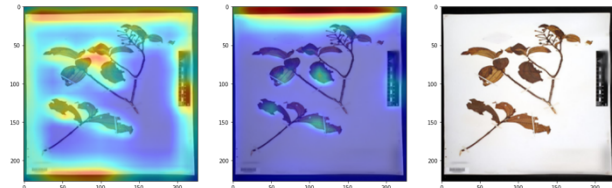
real class name: *Miconia ibaguensis* (Bonpl.) Triana

image 00001.jpg Detecting for class *Miconia racemosa* (Aubl.) DC. with 0.999



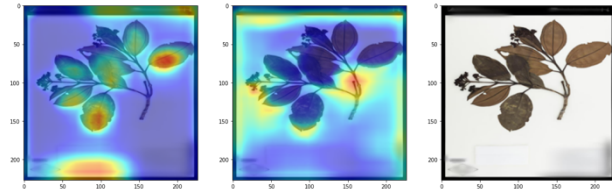
real class name: *Miconia racemosa* (Aubl.) DC.

image 00006.jpg Detecting for class *Mouriri myrtilloides* (Sw.) Poir. with 0.998



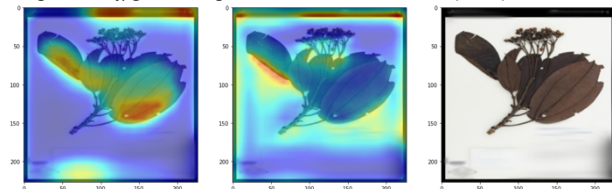
real class name: *Miconia corymbiformis* Cogn.

image 00000.jpg Detecting for class *Miconia racemosa* (Aubl.) DC. with 0.9872



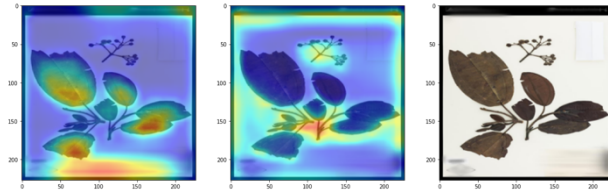
real class name: *Miconia corymbiformis* Cogn.

image 00001.jpg Detecting for class *Miconia racemosa* (Aubl.) DC. with 0.999



real class name: *Miconia corymbiformis* Cogn.

image 00002.jpg Detecting for class *Miconia racemosa* (Aubl.) DC. with 0.960



Note: There is also some misclassification with the normal value set that is classified correctly with the unusual value set.