CS 199

Torchattacks and ResNet50:

A Comparative Analysis of Adversarial Attack Performance

ENCARNACION, Stephen
RAMIREZ, Dean
RECUERDO, Diane

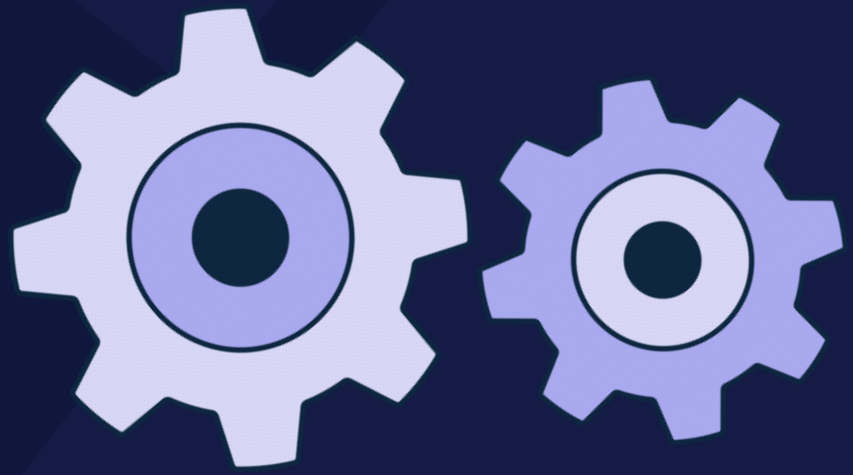
Background of the Study

BACKGROUND OF THE STUDY

data is EVERYWHERE

BACKGROUND OF THE STUDY

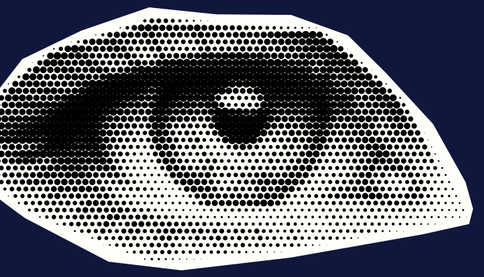
To process the sheer amount of information available, innovative solutions have emerged, with ***AUTOMATIONS*** taking center stage



BACKGROUND OF THE STUDY

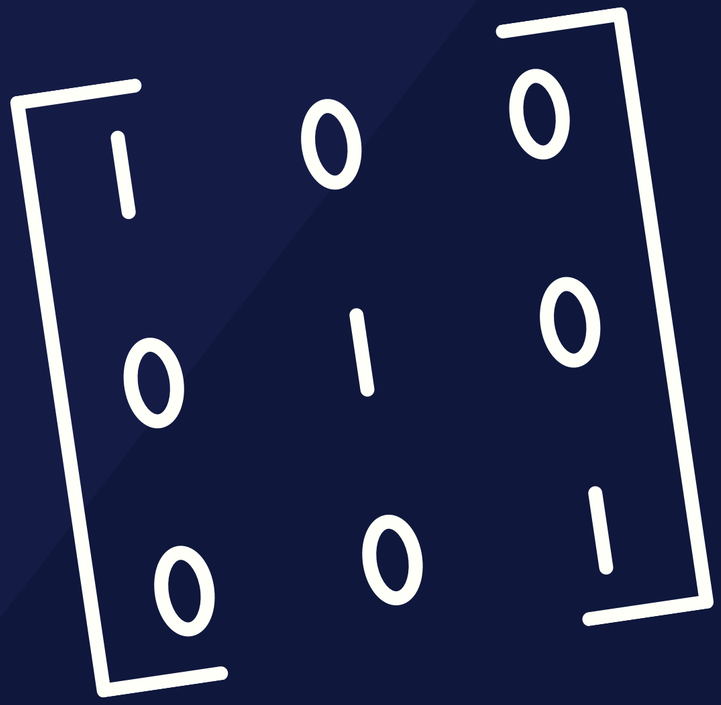
Among these innovations,
IMAGE CLASSIFICATION
stands out as a crucial component of computer vision, allowing for the automatic classification of images based on their visual content.

BACKGROUND OF THE STUDY



How do computers **see** images?

Computers interpret images through a **matrix of numbers**, with each number representing a specific color value of its corresponding pixel.



BACKGROUND OF THE STUDY

To enable a computer to make sense of these numerical values, the image must first undergo **preprocessing** to **improve the quality** of data and **standardize** the image.

RESIZING CROPPING ROTATION
NORMALIZATION

BACKGROUND OF THE STUDY

Afterward, the computer can start the process of ***FEATURE EXTRACTION***, wherein it can determine the image's edges, corners, textures, and other significant attributes.

The detected patterns are typically unique to certain classes of images.

BACKGROUND OF THE STUDY

Initially, **manual techniques** were employed for feature extraction. Researchers aimed to handcraft specific algorithms to detect and describe features within the images. While these manual techniques were effective to some extent, they had limitations in terms of flexibility and adaptability to different types of images.

BACKGROUND OF THE STUDY

To address these limitations, ***CONVOLUTIONAL NEURAL NETWORKS (CNNs)*** were developed. CNNs represent a significant advancement in the field of image recognition as they automate the process of feature extraction and classification. Instead of manually defining features, CNNs learn to identify them through training on very large datasets.

BACKGROUND OF THE STUDY

CNNs consist of multiple layers that progressively extract higher-level features from the raw pixel values. The initial layers might detect simple edges and textures, while deeper layers can identify more complex patterns, such as shapes or even specific objects like faces or cars. This hierarchical approach allows CNNs to build a detailed and abstract representation of the image, making them **highly effective** for tasks that involve the **classification of images.**

BACKGROUND OF THE STUDY

ResNet is a deep convolutional neural network (CNN) architecture that stands for "Residual Network," (Kundu, 2023). ResNet50 is a variant of the said architecture, with "50" in the name referring to the number of layers in the network. The researchers selected this model because it is accessible, easy to use as it is already pre-trained, and uses 1000 classifications, including those that are used by the image dataset from Kaggle. Mathworks (n. d.) mentioned, "You can load a pre-trained version of the neural network trained on more than a million images from the ImageNet database. The trained neural network can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals. As a result, the neural network has learned rich feature representations for a wide range of images."

BACKGROUND OF THE STUDY

As image classifiers become more prominent in various real-world applications, their security and robustness have to then be considered. One of the most significant threats to these models is **adversarial attacks**. These types of attacks introduce perturbed versions of the input image that seem normal to human observers but have the capacity to trick the classifier into labeling it as something else entirely.

BACKGROUND OF THE STUDY

RESEARCH OBJECTIVE

The study aims to perform a **comparative analysis** of the performance of the convolutional neural network (CNN) image classification model ResNet50 when classifying adversarial images generated using various adversarial attack types from the torchattacks library. The study uses **several metrics** such as **accuracy, precision, recall, F1 score,** and **attack execution time** to evaluate the model's performance, **gain insights into the robustness of ResNet50 against adversarial attacks**, and **identify potential areas for improving its resilience** against such attacks.

METHODOLOGY

METHODOLOGY

Adversarial-Attacks-PyTorch

Torchattacks is a PyTorch library that provides adversarial attacks to generate adversarial examples.

It contains *PyTorch-like* interface and functions that make it easier for PyTorch users to implement adversarial attacks.

Various digital attacks have already been created, with each algorithm differing from another. ***Torchattacks*** is a PyTorch library developed by Hoki Kim (2021) to generate adversarial examples from different types and test the robustness of deep learning models. It includes a variety of adversarial attack algorithms such as including Fast Gradient Sign Method (FGSM), Basic Iterative Method (BIM), Carlini & Wagner's (CW) attack, Random+FGSM (R+FGSM), Projected Gradient Descent (PGD), EOT+PGD (EOTPGD), PGD in TRADES (TPGD), FGSM in fast adversarial training (FFGSM), and MI-FGSM (MIFGSM), among others.

METHODOLOGY

SCOPE AND DELIMITATIONS

The researchers utilized the ResNet50 model, the torchattacks library, and nine adversarial attack types from the library: FGSM, PGD, CW, PGD L2, NIFGSM, Pixle, SINIFGSM, VMIFGSM, and VNIFGSM. These adversarial images were generated from a subset of a Kaggle dataset of fruits and vegetables), where 50 images were randomly selected (25 images of fruits and 25 images of vegetables). Thereafter, the performance of the model was assessed based on various metrics. This included measuring the model's accuracy, precision, recall, F1 score, and attack execution time.

METHODOLOGY

SCOPE AND DELIMITATIONS

As the researchers lack time and exceptional hardware resources, they decided to limit the number of images used to **50**, use only ResNet50 as the classification model, and utilize just torchattacks as the adversarial attack library. These delimitations should be kept in mind as they can affect the results and should be considered for future studies that may be similar to or based on this paper.

METHODOLOGY

DATA COLLECTION

The researchers utilized a dataset from Kaggle that contains **images of fruits and vegetables** used for training, testing, and validating machine learning models. The images in this dataset were scraped from Bing Image Search (Seth, 2020). The dataset contains 36 image classifications:

METHODOLOGY

DATA COLLECTION

Fruits: banana, apple, pear, grapes, orange, kiwi, watermelon, pomegranate, pineapple, and mango.

Vegetables: cucumber, carrot, capsicum, onion, potato, lemon, tomato, raddish, beetroot, cabbage, lettuce, spinach, soy bean, cauliflower, bell pepper, chili pepper, turnip, corn, sweetcorn, sweet potato, paprika, jalepeño, ginger, garlic, peas, and eggplant.

METHODOLOGY

DATA COLLECTION

The dataset contains three folders:

train

(100 images per classification)

test

(10 images per classification)

validation

(10 images per classification)

METHODOLOGY

PREPROCESSING

To ensure consistency in the input data, the sampled images were first converted to the RGB color space. Aside from converting the images, they were also resized to a standard size of 224x224 pixels. The images were also then converted to tensors (multi-dimensional arrays) as the model processes them in that form. Lastly, the images were normalized using predefined mean and standard deviation values for each color channel. This ensured that the input values (pixel intensities) have a similar data distribution, making the learning process more efficient.

METHODOLOGY

SAMPLING METHOD

This study only uses the **test folder** as it only requires a relatively smaller number of images due to previously discussed limitations. To avoid biases, the researchers use random sampling on the fruit and vegetable image dataset. The study uses **50 out of 359** test images, which consisted of **25 fruit images and 25 vegetable images**. These were stored in two folders (fruits and vegetables). Inside these folders are also folders that contain the test images and are named after the actual classification of the images inside them.

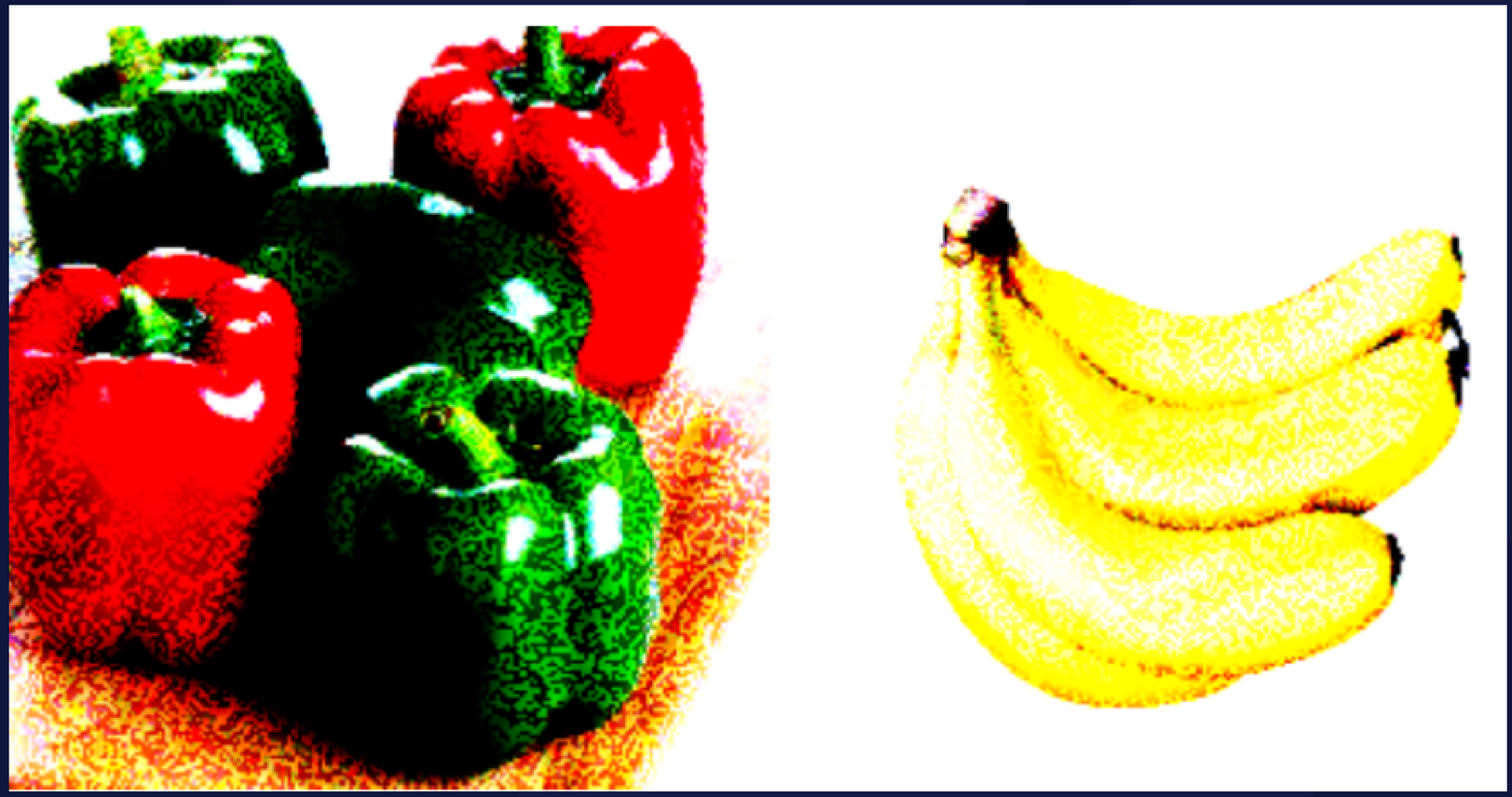
METHODOLOGY

IMAGE GENERATION

In order to choose which attack type to use, the researchers looked at the torchattack repository and searched for the types that were already studied and the most recently cited, specifically in the year 2022. Below are some untampered images from the sampled dataset:

METHODOLOGY

Fast Gradient Sign Method (FGSM)



This uses the neural network's gradients to create an adversarial sample. This method uses the gradients of the loss with respect to the image to make a new image that maximizes the loss.

METHODOLOGY

Projected Gradient Descent (PGD)

This is an iterative version of FGSM. Considered one of the strongest gradient-based attacks, this process consists of performing multiple FGSM attacks with small step sizes and clipping pixels that get modified beyond an epsilon ball.

METHODOLOGY

Carlini & Wagner (CW)

This generates adversarial images by formulating them as optimization problems. It seeks to discover the smallest perturbation to the image, causing a misclassification by the target model.

METHODOLOGY

Projected Gradient Descent with L2 Norm Constraint (PGD L2)

A variant of PGD that uses the L2 norm constraint on the adversarial perturbation. This generates examples using different Lp norms.

METHODOLOGY

Nesterov Iterative Fast Gradient Sign Method (NIFGSM)

The transferability of adversarial examples is improved by incorporating Nesterov's accelerated gradient into the iterative attacks. It looks ahead and enhances the transferability of examples.

METHODOLOGY

Pixle

A black-box attack that can correctly attack a high percentage of dataset samples. This is done by rearranging a small number of pixels within the input image.

METHODOLOGY

VMIFGSM

Enhances the class of iterative gradient-based attack methods and enhances their attack transferability. For every gradient calculation's iteration, it considers the variance of the previous iteration to determine the current gradient. This is opposite to attacks that directly use the current gradient momentum accumulation

METHODOLOGY

VNIFGSM

Similar to VMIFGSM. Unfortunately, the available information regarding this attack method is limited.

METHODOLOGY

IMAGE GENERATION

These attacks were used to generate adversarial images to be fed into the ResNet50 model. Thereafter, the model classified them while the researchers gathered information regarding their performance.

METHODOLOGY

IMAGE GENERATION

```
[ ]  attacks = {
        'FGSM': FGSM(model, eps=0.3),
        'PGD': PGD(model, eps=8/255, alpha=2/255, steps=2),
        'CW': CW(model, c=1, kappa=0, steps=2, lr=0.01),
        'PGDL2': PGDL2(model, eps=0.3, alpha=0.01, steps=2),
        'NIFGSM': NIFGSM(model, eps=0.2),
        'Pixle': Pixle(model, x_dimensions=(2, 10), y_dimensions=(2, 10),
                       pixel_mapping='random', restarts=20, max_iterations=10, update_each_iteration=False),
        'SINIFGSM': SINIFGSM(model, eps=8/255, alpha=2/255, steps=2, decay=1.0, m=5),
        'VMIFGSM': VMIFGSM(model, eps=8/255, alpha=2/255, steps=2, decay=1.0),
        'VNIFGSM': VNIFGSM(model, eps=8/255, alpha=2/255, steps=2, decay=1.0, N=5, beta=3/2)
    }
```

METHODOLOGY

EVALUATION METRICS

To effectively assess the performance of the ResNet50 model against adversarial attacks, several metrics were calculated. The researchers also considered the type of machine learning task, specifically multi-class classification, in the selection of the following parameters:

METHODOLOGY

EVALUATION METRICS

ACCURACY - Measures the ratio of correctly classified cases to the total number of objects in the dataset

PRECISION - In a multi-class classification problem, precision is the percentage of cases that, out of all those that the model predicted to belong to a given class, were properly classified as such.

METHODOLOGY

EVALUATION METRICS

RECALL - In multi-class classification, recall refers to the percentage of a class's instances that the model successfully classified out of all the class's occurrences. Ultimately, recall measures the model's ability to identify all instances of a particular class

F1-SCORE - This is a measure of a model's performance, representing the harmonic mean of precision and recall. By giving both precision and recall equal weight, the balance between these two can be indicated. A higher F1 score generally signifies a better balance between the two metrics, denoting better model effectiveness

METHODOLOGY

EVALUATION METRICS

ATTACK EXECUTION TIME - To measure how fast and efficiently adversarial attacks generate images, the attack execution time is measured.

METHODOLOGY

ANALYSIS METHODS

Various measures of center **(mean, median, and mode)** and a measure of variation **(standard deviation)** were used to interpret the results. This gave the researchers a sense of the central tendency and variability of each metric. Aside from calculating the descriptive statistics, the attacks were ranked based on each performance metric to help identify which attacks perform best and worst on average.

RESULTS & DISCUSSION

RESULTS AND DISCUSSION

PERFORMANCE ANALYSIS

Attack type	Accuracy	Precision	Recall	F1-score	Time
No attack	72.0	48.53	49.16	48.57	0.00
FGSM	14.0	18.18	7.32	9.55	38.37
PGD	52.0	23.21	18.49	20.09	44.32
CW	52.0	23.21	17.98	17.98	73.67
PGD (L2)	52.0	23.21	17.98	17.98	65.92
NIFGSM	54.0	23.81	20.11	21.00	236.56
Pixle	74.0	51.56	53.12	52.08	31.72
SINIFGSM	52.0	23.21	18.49	20.09	229.59
VMIFGSM	52.0	23.21	18.49	20.09	267.68
VNIFGSM	54.0	23.57	19.39	20.68	283.14

No Attack: As expected, without any adversarial attack, the model performs the best with an accuracy of 72.00%, precision of 48.53%, recall of 49.16%, and F1 score of 48.57%.

Fast Gradient Sign Method (FGSM): FGSM significantly reduces the model's performance, lowering the accuracy to 14.00%. This shows the effectiveness of FGSM in creating adversarial examples that can mislead the model.

RESULTS AND DISCUSSION

PERFORMANCE ANALYSIS

Attack type	Accuracy	Precision	Recall	F1-score	Time
No attack	72.0	48.53	49.16	48.57	0.00
FGSM	14.0	18.18	7.32	9.55	38.37
PGD	52.0	23.21	18.49	20.09	44.32
CW	52.0	23.21	17.98	17.98	73.67
PGD (L2)	52.0	23.21	17.98	17.98	65.92
NIFGSM	54.0	23.81	20.11	21.00	236.56
Pixle	74.0	51.56	53.12	52.08	31.72
SINIFGSM	52.0	23.21	18.49	20.09	229.59
VMIFGSM	52.0	23.21	18.49	20.09	267.68
VNIFGSM	54.0	23.57	19.39	20.68	283.14

Projected Gradient Descent (PGD): PGD also effectively reduces the model's performance, with an accuracy of 52.00%. It takes slightly longer than FGSM, indicating a trade-off between attack effectiveness and computational cost.

Carlini & Wagner (CW) Attack: The CW attack decreases the model's accuracy to 52.00%. It takes more time than both FGSM and PGD, suggesting that it might be a more complex attack method.

RESULTS AND DISCUSSION

PERFORMANCE ANALYSIS

Attack type	Accuracy	Precision	Recall	F1-score	Time
No attack	72.0	48.53	49.16	48.57	0.00
FGSM	14.0	18.18	7.32	9.55	38.37
PGD	52.0	23.21	18.49	20.09	44.32
CW	52.0	23.21	17.98	17.98	73.67
PGD (L2)	52.0	23.21	17.98	17.98	65.92
NIFGSM	54.0	23.81	20.11	21.00	236.56
Pixle	74.0	51.56	53.12	52.08	31.72
SINIFGSM	52.0	23.21	18.49	20.09	229.59
VMIFGSM	52.0	23.21	18.49	20.09	267.68
VNIFGSM	54.0	23.57	19.39	20.68	283.14

Projected Gradient Descent with L2 Norm Constraint (PGD L2): This variant of PGD has the same effect on the model's performance as the standard PGD and CW attack, but it takes slightly more time.

Nesterov Iterative Fast Gradient Sign Method (NIFGSM): NIFGSM slightly improves the model's accuracy compared to PGD, CW, and PGD L2, but it takes significantly more time, indicating a higher computational cost.

RESULTS AND DISCUSSION

PERFORMANCE ANALYSIS

Attack type	Accuracy	Precision	Recall	F1-score	Time
No attack	72.0	48.53	49.16	48.57	0.00
FGSM	14.0	18.18	7.32	9.55	38.37
PGD	52.0	23.21	18.49	20.09	44.32
CW	52.0	23.21	17.98	17.98	73.67
PGD (L2)	52.0	23.21	17.98	17.98	65.92
NIFGSM	54.0	23.81	20.11	21.00	236.56
Pixle	74.0	51.56	53.12	52.08	31.72
SINIFGSM	52.0	23.21	18.49	20.09	229.59
VMIFGSM	52.0	23.21	18.49	20.09	267.68
VNIFGSM	54.0	23.57	19.39	20.68	283.14

Pixle: Pixle seems to be the least effective adversarial attack in this list, as the model's performance is only slightly reduced compared to the case where no attack was used. However, it takes less time than most other attacks.

Scale-Invariant Nesterov Iterative Fast Gradient Sign Method (SINIFGSM): SINIFGSM has the same effect on the model's performance as PGD, CW, and PGD L2, but it takes significantly more time.

RESULTS AND DISCUSSION

PERFORMANCE ANALYSIS

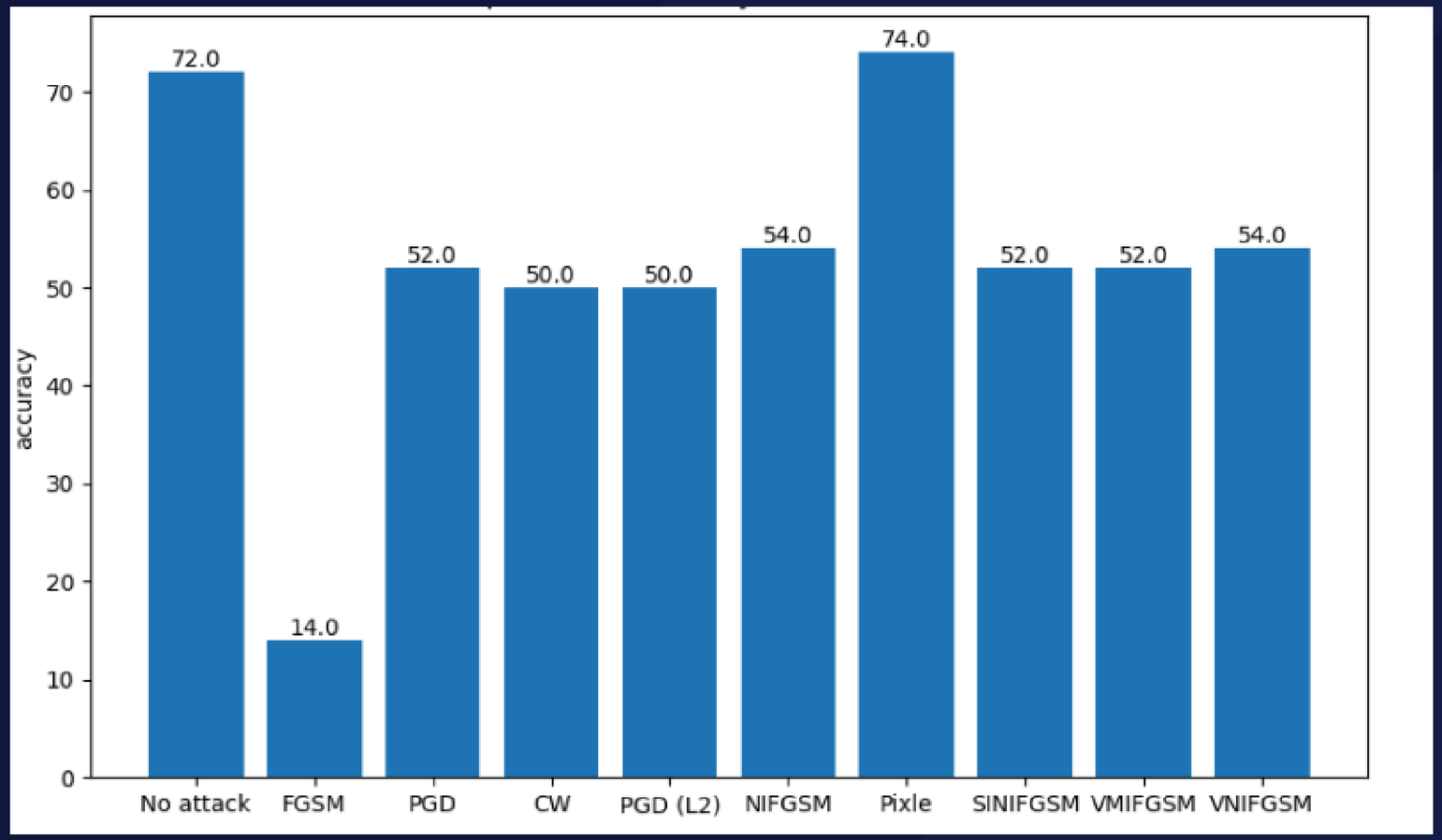
Attack type	Accuracy	Precision	Recall	F1-score	Time
No attack	72.0	48.53	49.16	48.57	0.00
FGSM	14.0	18.18	7.32	9.55	38.37
PGD	52.0	23.21	18.49	20.09	44.32
CW	52.0	23.21	17.98	17.98	73.67
PGD (L2)	52.0	23.21	17.98	17.98	65.92
NIFGSM	54.0	23.81	20.11	21.00	236.56
Pixle	74.0	51.56	53.12	52.08	31.72
SINIFGSM	52.0	23.21	18.49	20.09	229.59
VMIFGSM	52.0	23.21	18.49	20.09	267.68
VNIFGSM	54.0	23.57	19.39	20.68	283.14

VMIFGSM: VMIFGSM has the same effect on the model's performance as PGD, CW, and PGD L2, but it takes even more time than SINIFGSM.

VNIFGSM: VNIFGSM slightly improves the model's accuracy compared to PGD, CW, PGD L2, SINIFGSM, and VMIFGSM, but it takes the most time among all the attacks.

RESULTS AND DISCUSSION

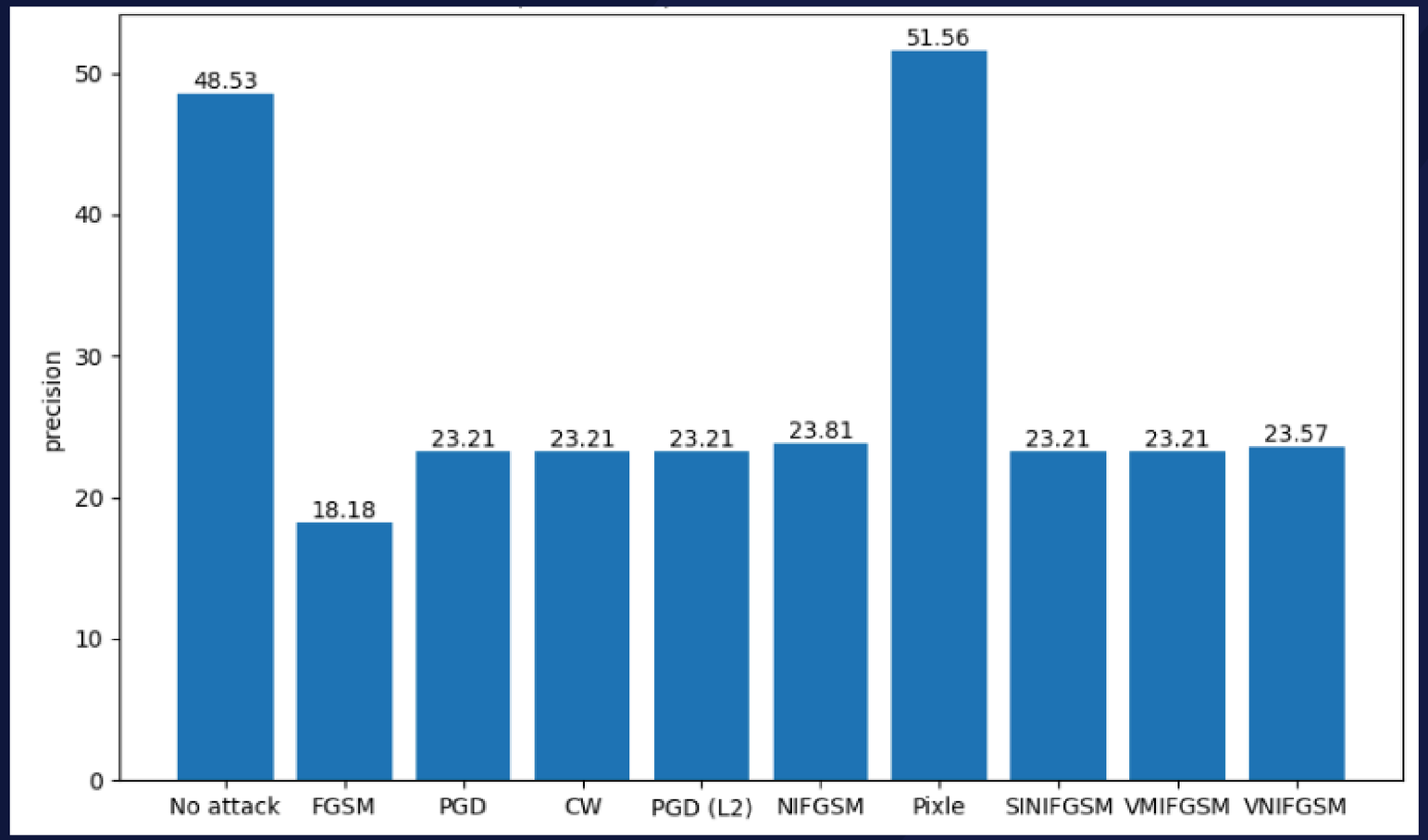
ACCURACY COMPARISON



The cases where Pixle was used and no adversarial attack was used yielded the highest accuracy, with scores of 74.00% and 72.00%, respectively. This indicates that the model performs best when there is no adversarial attack or when the Pixle attack is used. On the other hand, the FGSM attack results in the lowest accuracy with a score of 14.00%, suggesting that this attack significantly impairs the model's performance.

RESULTS AND DISCUSSION

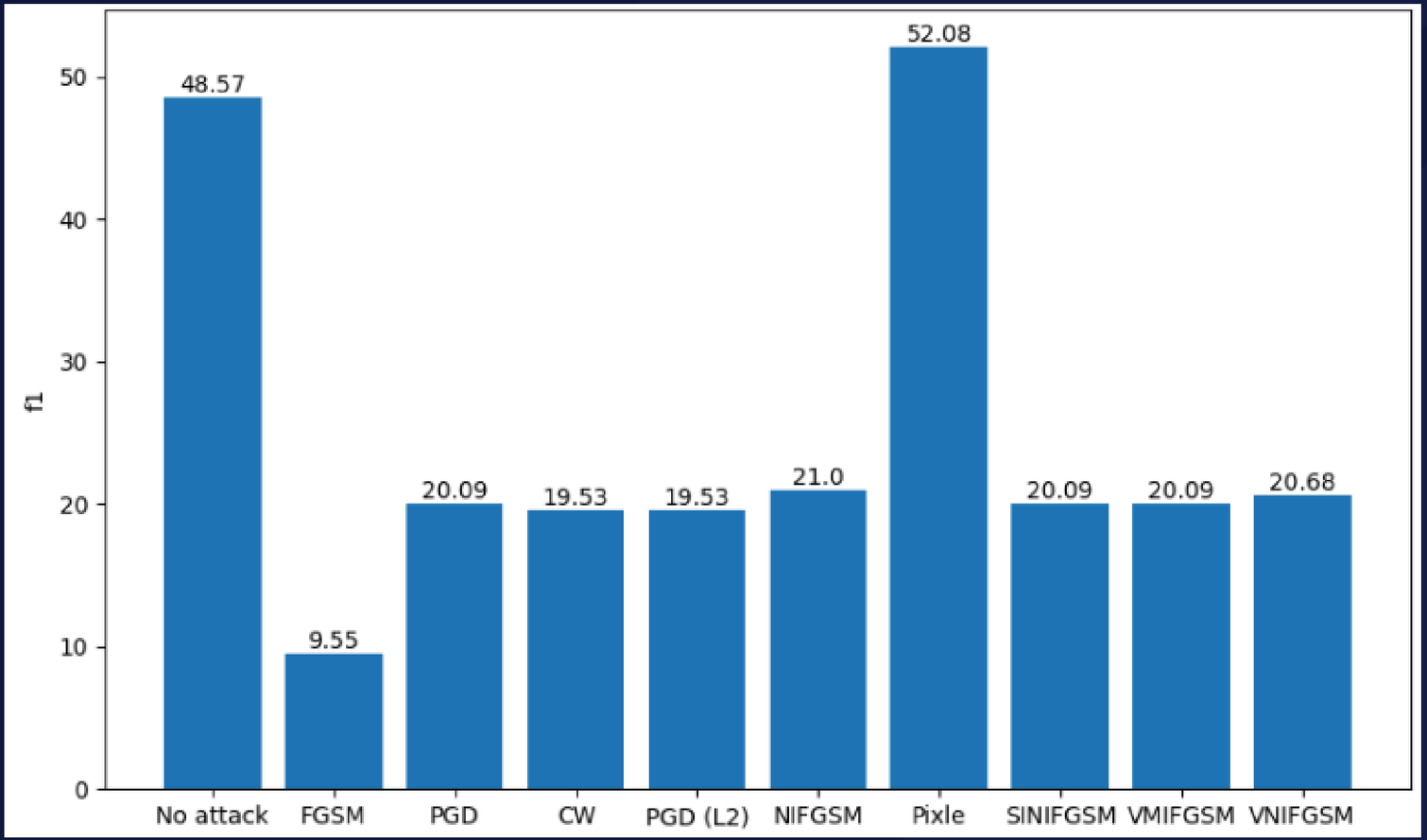
PRECISION COMPARISON



Again, the cases where Pixle was used and no adversarial attack was used yielded the highest precision, with scores of 51.56% and 48.53%, respectively. This indicates that the model's positive predictions are most reliable in these scenarios. On the other hand, the FGSM attack results in the lowest precision (18.18%), suggesting that this attack causes the model to make a large number of false positive predictions.

RESULTS AND DISCUSSION

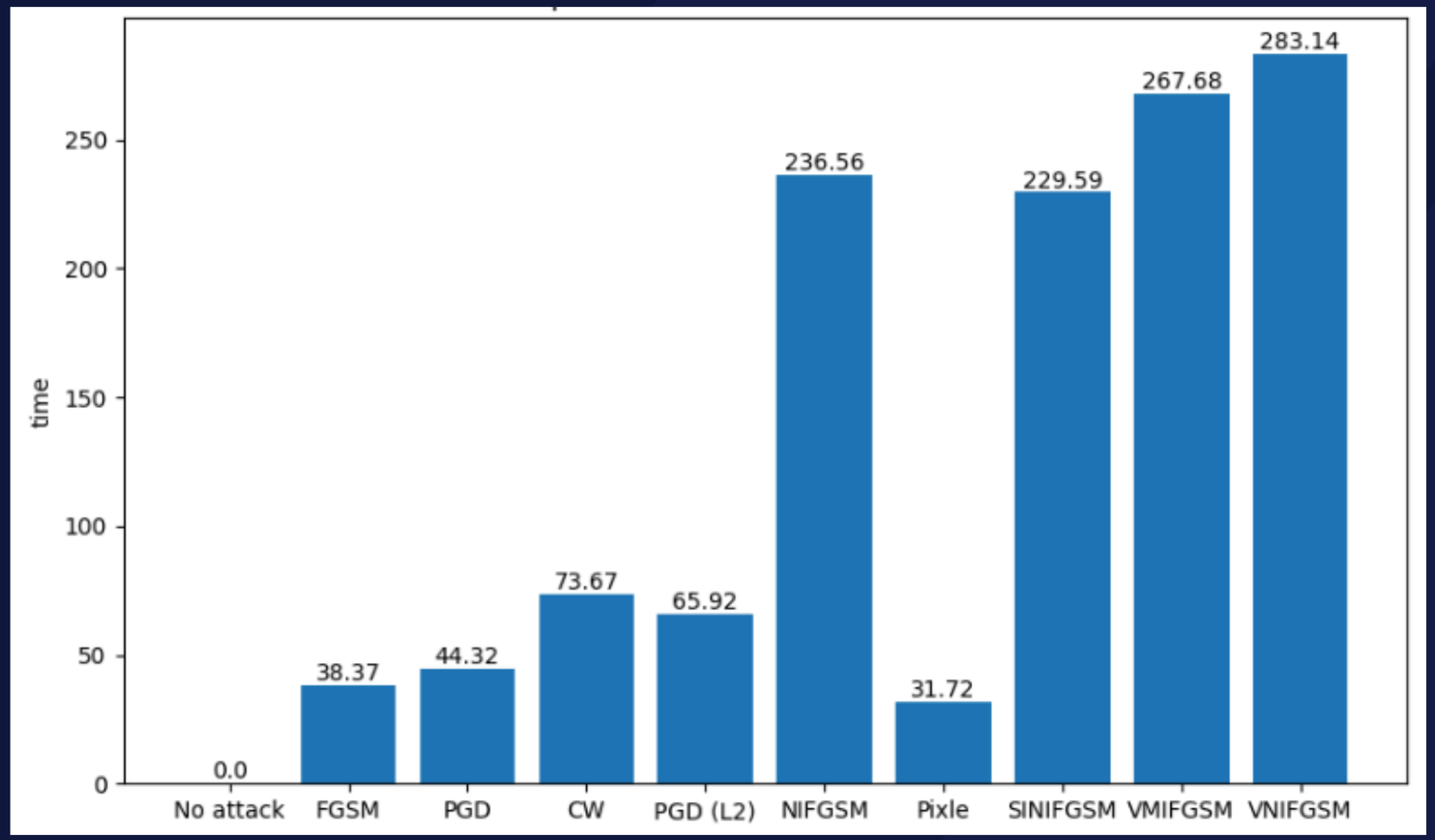
F1-SCORE COMPARISON



The cases where Pixle was used and no adversarial attack was used yielded the highest F1 with scores of 52.08% and 48.57%, respectively. This indicates that the model achieves the best balance between precision and recall in these scenarios. On the other hand, the FGSM attack results in the lowest F1 score (9.55%), suggesting that this attack significantly disrupts the balance between precision and recall.

RESULTS AND DISCUSSION

EXECUTION TIME COMPARISON



Pixle takes the least time of 31.72 seconds (excluding the case where no adversarial attack was used), while the VNIFGSM attack takes the most time of attack execution at 283.14 seconds. This suggests that some adversarial attacks may be more computationally intensive than others.

RESULTS AND DISCUSSION

CENTRAL TENDENCY EVALUATION

Accuracy: The mean, median, and mode of this metric are almost equal to each other (52.4, 52.0, and 52.0). This could signify that the distribution is nearly perfectly symmetrical and non-skewed (Statsdirect, n. d.). The mean of this metric shows that, on average, the model classifies approximately half of the adversarial examples correctly. This is 27.22% less than the model's accuracy in the 'no attack' case at 72%. On the other hand, despite having the second highest standard deviation among all of the metrics at 16.13, this indicates a moderate spread of scores.

Metric	Mean	Median	Mode	Standard Deviation
accuracy	52.40	52.00	52.000	16.13
precision	28.17	23.21	23.21	11.66
recall	24.05	18.49	18.49	14.75
f1-score	25.12	20.09	20.09	13.72
attack execution time	13.72	69.795	0.00	112.14

RESULTS AND DISCUSSION

CENTRAL TENDENCY EVALUATION

Precision: On average, the model's positive predictions are 28.17% correct. This is slightly higher than the median and mode, indicating a slight skew in the data. Furthermore, the precision's mean is 41.95% less than the model's precision in the 'no attack' case at 48.53%. The median and mode being equal at 23.21% suggests that these are the common values across different adversarial attacks. Having a standard deviation of 11.66 indicates a moderate spread of scores.

Metric	Mean	Median	Mode	Standard Deviation
accuracy	52.40	52.00	52.000	16.13
precision	28.17	23.21	23.21	11.66
recall	24.05	18.49	18.49	14.75
f1-score	25.12	20.09	20.09	13.72
attack execution time	13.72	69.795	0.00	112.14

RESULTS AND DISCUSSION

CENTRAL TENDENCY EVALUATION

Recall: On average, the model correctly identifies 24.05% of all actual positive cases. This is slightly higher than the median and mode, indicating a slight skew in the data. Furthermore, the recall's mean is 51.08% less than the model's precision in the 'no attack' case at 49.16%. The median and mode being equal at 18.49% suggests that these are the common values across different adversarial attacks. Having a standard deviation of 14.75 indicates a moderate spread of scores.

Metric	Mean	Median	Mode	Standard Deviation
accuracy	52.40	52.00	52.000	16.13
precision	28.17	23.21	23.21	11.66
recall	24.05	18.49	18.49	14.75
f1-score	25.12	20.09	20.09	13.72
attack execution time	13.72	69.795	0.00	112.14

RESULTS AND DISCUSSION

CENTRAL TENDENCY EVALUATION

F1-score: An F1 score is considered perfect at 100%, and worst at 0%. The mean F1 score of 25.12% suggests that the model has low effectiveness in classifying adversarial examples. Furthermore, the F1 score's mean is 48.28% less than the model's precision in the 'no attack' case at 48.57%. The median and mode being equal at 20.09% suggests that these are the common values across different adversarial attacks. Having a standard deviation of 13.72 indicates a moderate spread of scores.

Metric	Mean	Median	Mode	Standard Deviation
accuracy	52.40	52.00	52.000	16.13
precision	28.17	23.21	23.21	11.66
recall	24.05	18.49	18.49	14.75
f1-score	25.12	20.09	20.09	13.72
attack execution time	13.72	69.795	0.00	112.14

RESULTS AND DISCUSSION

CENTRAL TENDENCY EVALUATION

Attack execution time: On average, the model identifies all adversarial images, regardless of whether they are correct or wrong, within 127.10 seconds. This is slightly higher than the median and mode, indicating a slight skew in the data. The standard deviation of 112.14 indicates a high variation in prediction time, which could be due to image complexity or hardware performance.

Metric	Mean	Median	Mode	Standard Deviation
accuracy	52.40	52.00	52.000	16.13
precision	28.17	23.21	23.21	11.66
recall	24.05	18.49	18.49	14.75
f1-score	25.12	20.09	20.09	13.72
attack execution time	13.72	69.795	0.00	112.14

CONCLUSION & RECOMMENDATION

Fast Gradient Sign Method (FGSM) was able to fool the model the most by achieving the lowest accuracy for all adversarial attacks. On the other hand, the Pixle attack also displayed fascinating results by being the least effective, even displaying a slight performance improvement compared to when no attack was applied to the model. The study also showed potential for a trade-off between attack effectiveness and computation cost. A few adversarial attacks were relatively time-consuming, e.g. VNIFGSM took more time to create an adversarial example and marginally enhanced the accuracy results slightly compared to some other attacks.

CONCLUSION & RECOMMENDATION

The FGSM and Pixel Attack displaying varying degrees of effectiveness show that it is incorrect to say there is a one-size-fits-all approach for securing our models. Each attack has its challenges and requires unique countermeasures. Furthermore, this dilemma is compounded by the trade-off between the effect of an attack and computational cost. Some attacks might be better but could also consume more resources which may not be available in some cases.

In summary, this work is another reflection of the continuing race between improving model performance on the one hand and ensuring model security on the other. Continuous research and development will be crucial as we move ahead to stay well ahead of potential threats and make our machine-learning models trustworthy.

CONCLUSION & RECOMMENDATION

Based on the results of the study, the researchers propose to consider several recommendations for future research. The main constraints of this study include the need for more time and limitations in hardware resource adequacy. Hence, allotting more time and optimizing hardware resources should allow for more extensive testing. This includes expanding the dataset, utilizing other adversarial attack libraries, considering different sampling methods, and investigating additional performance metrics. This, in turn, provides a more comprehensive assessment of the model's performance against adversarial attacks.

CONCLUSION & RECOMMENDATION

Furthermore, future studies can also employ multiple models, aside from the ResNet50 model, in order to compare various models' levels of robustness and accuracy against adversarial attacks. Different models are trained on different datasets, resulting in varied classification specializations. Lastly, new research can focus on training the model(s) using the images generated by multiple adversarial attacks. Doing so explores the possibilities of improving the robustness and accuracy of the model(s) against significant perturbations.