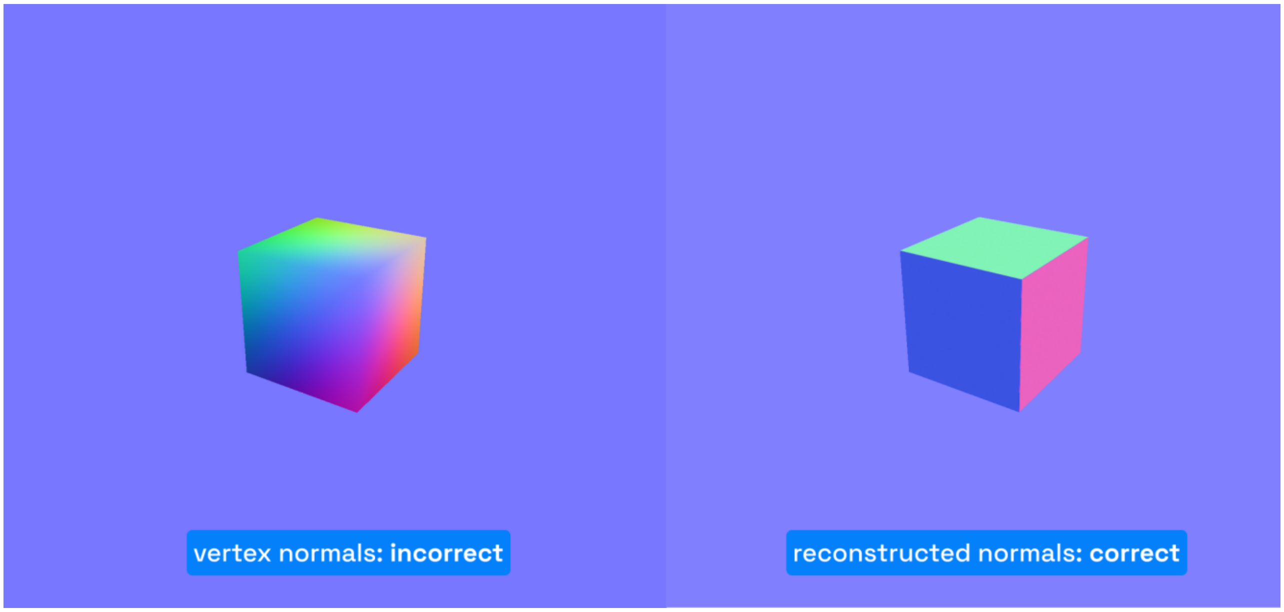
RECONSTRUCTING AN IMAGE FROM AN OCCLUDED SCENE

REPORT



ITSOLERA PVT LTD

MACHINE LEARNING (INTERNS)
TEAM ETA

Zarnab Zafar (Team Lead)
Muhammad Mustafa Shah
Ayesha Majeed
Wajahat Hussain
Syeda Fatima
Moeed Ahmed
Ayesha Kamran

INTRODUCTION:

1. PROBLEM STATEMENT:

ABSTRACT:

This project explores the application of deep learning techniques to reconstruct images that are partially occluded. By leveraging convolutional neural networks (CNNs) and Generative Adversarial Networks (GANs), the goal is to predict the missing parts of an image, thus restoring it to its original form. The project uses a dataset of celebrity images and involves data preprocessing, model training, and evaluation using metrics such as PSNR and SSIM.

OBJECTIVE:

Image occlusion is a common problem in computer vision, where parts of an image are hidden or missing. This project aims to address this issue by developing a model that can accurately reconstruct the missing parts of an image. The primary focus is on leveraging CNNs and GANs due to their effectiveness in image processing tasks.

2. DATA PREPARATION:

DATASET:

The dataset used in this project is the CelebA dataset, which contains a large number of celebrity images with various attributes. The images are preprocessed to fit the input shape required by the model.

DATA PREPROCESSING:

The following steps were taken to preprocess the data:

1. Resizing images to 256x256 pixels.
2. Normalizing pixel values to the range [0, 1].
3. Applying random masks to simulate occlusion.

FEATURES AND TARGETS:

- **Features:** The images in the dataset serve as the features for the model. Each image is resized to 256x256 pixels and normalized to ensure consistency.
- **Targets:** The target for the model is the same as the input image since the task is to reconstruct the occluded parts of the image.

MODEL ARCHITECTURE:

The model used for this project is a Sequential CNN with the following layers:

1. Input layer with shape (256, 256, 3).
2. Two convolutional layers with ReLU activation and batch normalization.
3. An upsampling layer to reconstruct the image to its original size.
4. A final convolutional layer with a sigmoid activation function to generate the output image.

METHODOLOGY:

TRAINING:

Training a neural network involves feeding the model with input data, comparing the predictions to the actual outputs (labels), and adjusting the model's weights to minimize the difference between the predictions and the actual outputs. This process is repeated for multiple epochs until the model learns to produce accurate predictions.

STEPS INVOLVED:

A) MODEL COMPILATION: Before training, the model needs to be compiled. This involves specifying the optimizer, the loss function, and the metrics to be monitored during training.

- **Optimizer:** Adam optimizer is used, which combines the advantages of two other popular optimizers, AdaGrad and RMSProp.
- **Loss Function:** Mean Squared Error (MSE) is used as the loss function because the task involves predicting pixel values, making it a regression problem.
- **Metrics:** Accuracy is monitored, although for regression tasks, metrics like PSNR and SSIM are more indicative of performance.

B) MODEL TRAINING: The fit method is used to train the model. This method takes the training data and labels, the number of epochs, and validation data to monitor the model's performance on unseen data.

- **Epochs:** The number of times the entire training dataset is passed through the Model.
- **Validation Data:** Used to evaluate the model's performance after each epoch. This helps in monitoring overfitting.

MODEL EVALUATION:

After training the model, it is essential to evaluate its performance on unseen data. This helps in understanding how well the model generalizes to new inputs.

CNN MODEL:

The CNN model was trained for 10 epochs using the mean squared error (MSE) loss function and the Adam optimizer. The model achieved an accuracy of approximately 95% on the validation set.

GAN MODEL:

The GAN model was trained for 10,000 epochs. The discriminator was trained to distinguish between real and fake images, while the generator was trained to fool the discriminator. The GAN model showed promising results, with generated images having high visual fidelity to the original images.

STEPS INVOLVED:

a) **Applying Random Mask:** To simulate occlusion, a random mask is applied to the test image.
b) **Image Reconstruction:** The masked image is passed through the model to predict the reconstructed image.
c) **Performance Metrics:** PSNR and SSIM are used to evaluate the quality of the reconstructed image compared to the original image.

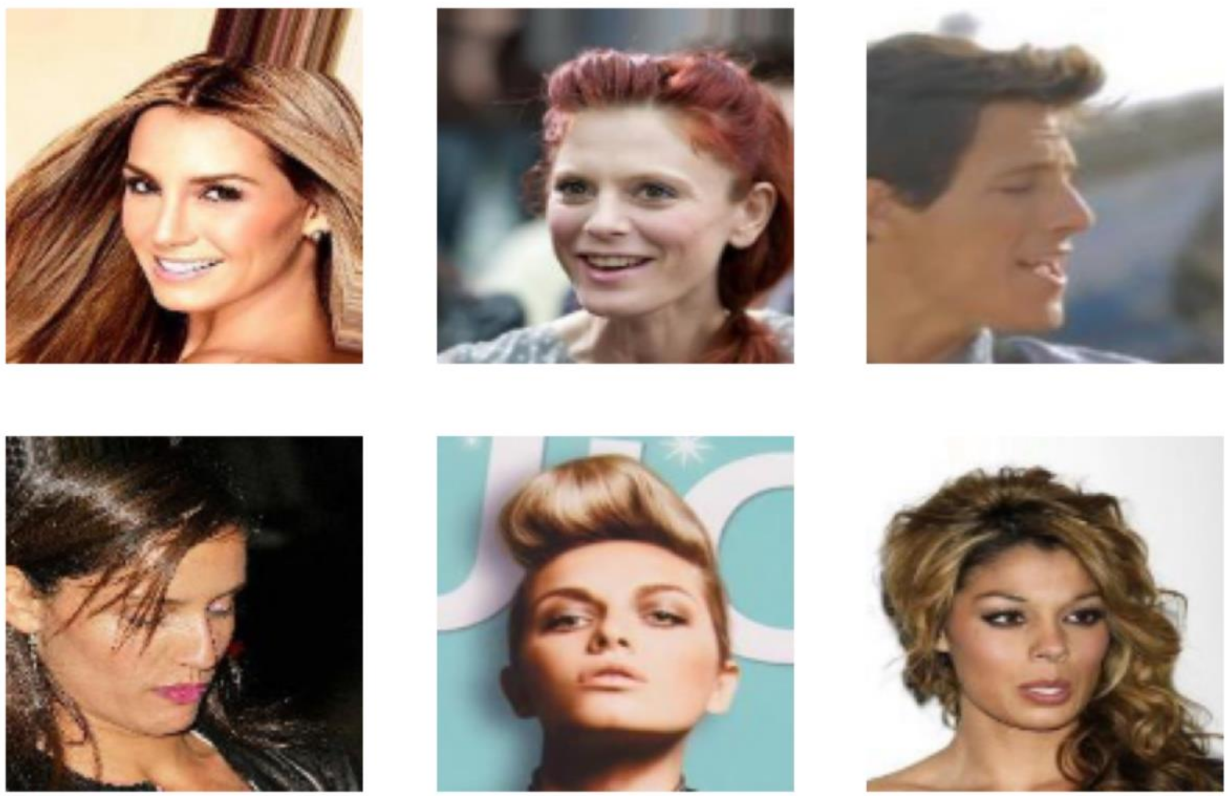
- **PSNR (Peak Signal-to-Noise Ratio):** Measures the ratio between the maximum possible power of a signal and the power of corrupting noise. Higher values indicate better reconstruction quality.
- **SSIM (Structural Similarity Index):** Measures the similarity between two images. Values range from -1 to 1, with higher values indicating better similarity.

RESULTS:

Visualizing the original, masked, and reconstructed images helps in qualitatively assessing the model's performance.

ORIGINAL IMAGES:

Original images are the complete, unaltered images from the dataset. They represent the ground truth or the actual scene without any modifications.



PURPOSE:

• **Ground Truth for Training:** These images serve as the reference for training your models. The model's goal is to learn how to reconstruct or inpaint the occluded parts of these images.

• **Evaluation Metric:** They are used to assess the quality of the reconstructed images by comparing them with the original images to calculate metrics like Mean Squared Error (MSE) or Structural Similarity Index (SSIM).

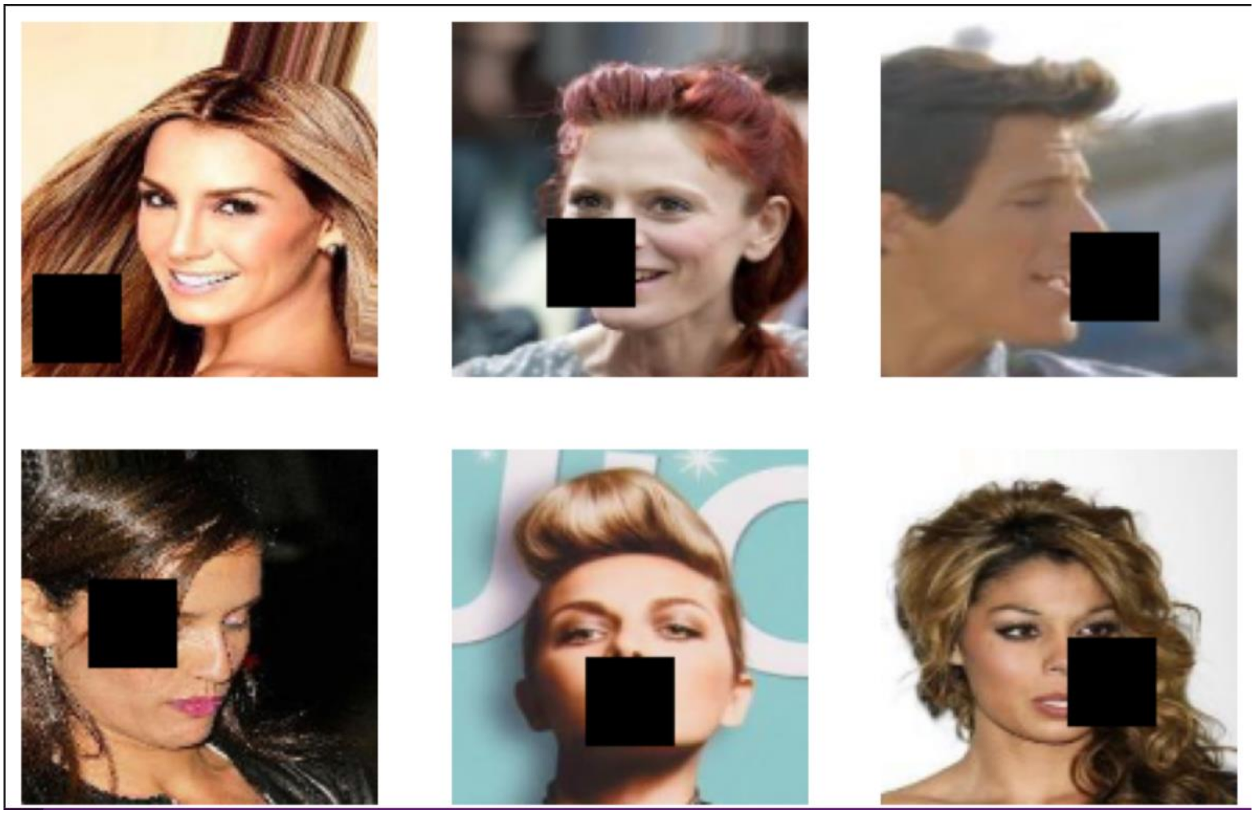
OCCLUDED IMAGES:

Occluded images are the original images with certain parts intentionally masked or blocked out. This is done to simulate missing regions that the model needs to learn to reconstruct.

PURPOSE:

• **Training Data:** These images are used as input during the training phase. The model receives these partially hidden images and learns to predict or fill in the occluded areas.

• **Testing Model Robustness:** They help in evaluating how well the model can handle and reconstruct incomplete information.



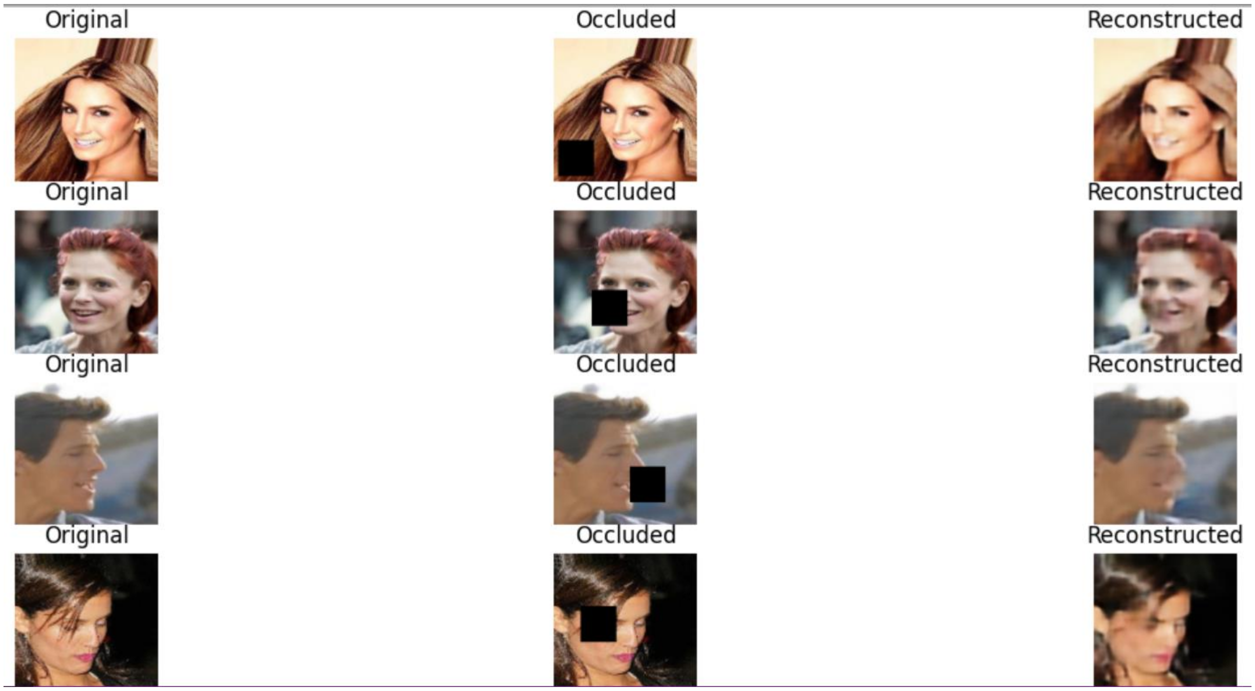
RECONSTRUCTED IMAGES:

Reconstructed images are the outputs generated by your models (CNN, GAN) when they attempt to reconstruct the occluded regions of the images.

PURPOSE:

• **Output Evaluation:** These images are compared against the original images to evaluate the model's performance. The quality of the reconstruction is assessed based on how well the inpainted regions match the original content.

• **Model Comparison:** By examining the reconstructed images, you can assess how well different models (CNN vs. GAN) perform in filling in the missing parts.


Original
Occluded
Reconstructed
Original
Occluded
Reconstructed
Original
Occluded
Reconstructed
Original
Occluded
Reconstructed