Enhancing ResNet Robustness to Noisy Images Using GANs

Introduction

The goal of this project is to improve the robustness of ResNet when processing noisy images. Noise, such as Gaussian, salt-and-pepper, and Poisson noise, often appears in images due to low-quality cameras, poor conditions, or errors during transmission. Some data is easily lost during image transmission, which further reduces the quality of the images. This project aims to solve this problem by improving the robustness of the model.

Using a dataset of celebrity facial photos, we will compare the accuracy of facial recognition between a standard ResNet model and a ResNet+GAN model. The focus is to evaluate whether the ResNet+GAN architecture can handle noisy images better than the standard ResNet. By addressing this limitation, we aim to make ResNet more reliable in real-world scenarios. The goal is to optimize the original model ResNet50 to the ResNet + GAN model and compare it with ResNet + Mae.

Dataset

Original Dataset: This dataset is a collection of celebrity facial photos obtained from CARC. It contains images of the same celebrity taken at different ages, providing diverse facial features across time.

Noisy Dataset: The noisy dataset is created by adding noise to the original photos. To simulate real-world distortions, We combined three types of noise—Gaussian, salt-and-pepper, and Poisson— and applied them to the same image. This ensures that the dataset mimics the challenging conditions faced in practical scenarios.

Methodology

ResNet

ResNet Baseline: The standard ResNet model will be trained on the original dataset and tested both on the noisy and original datasets to establish baseline performance. In this part, we will use the implemented model to test the accuracy of the original image prediction and the accuracy of the image with noise.

ResNet + GAN

We implemented and integrated the GAN model into ResNet as a solution to improve robustness. The ResNet+GAN model is trained on the original dataset and tested on the noisy and original dataset to analyze its accuracy.

In the ResNet + GAN model. ResNet50 is used as the classification model, and the loss function is CrossEntropy Loss. For GAN, the loss function is BCE Loss. The Discriminator uses a convolutional neural network architecture to distinguish real and fake images. That is, multiple convolutional layers, Sigmoid activation functions, and batch normalization layers are selected. The Generator uses a convolutional neural network architecture and input to generate fake images. Multiple convolutional layers and Tanh activation functions are selected based on noise.

ResNet + MAE

The Masked Autoencoder (MAE) model will be trained on the original dataset and tested on both the noisy and original datasets to evaluate its performance. In this part, we will use MAE to test the accuracy of the original image prediction and the accuracy of the image with noise. As shown in Figure 1~5, we will walk through each processing step, including the original image, the masked image, the reconstructed image, and the final reconstructed image with the visible areas.

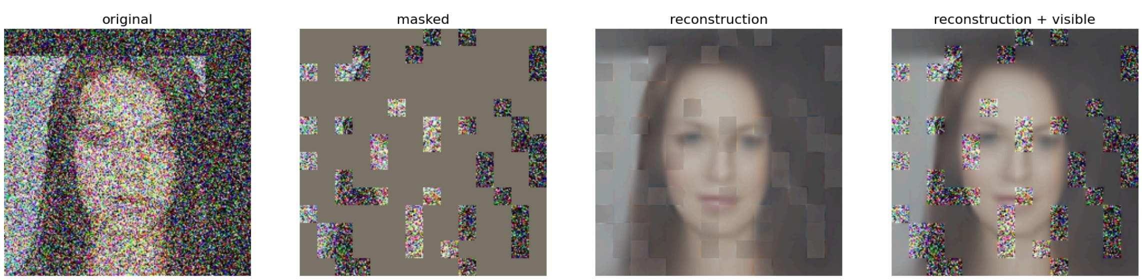

Figure 1. Reconstructed Picture

(gaussian_noise-standard deviation=0.7; salt_and_pepper_noise-probability=0.07)

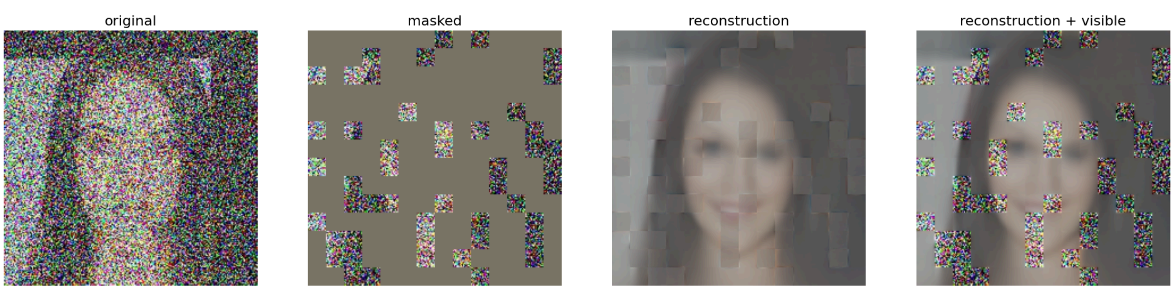

Figure 2. Reconstructed Picture

(gaussian_noise-standard deviation=0.7; salt_and_pepper_noise-probability=0.09)

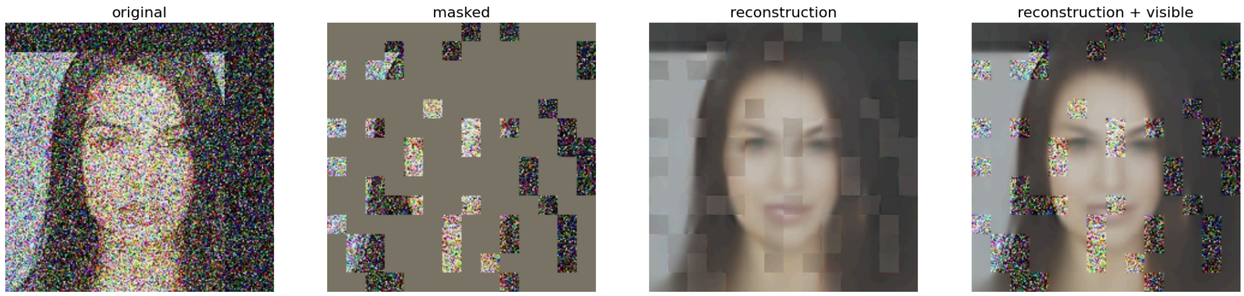

Figure 3. Reconstructed Picture

(gaussian_noise-standard deviation=0.5; salt_and_pepper_noise-probability=0.07)

Figure 4. Reconstructed Picture

(gaussian_noise-standard deviation=0.6 ;salt_and_pepper_noise-probability=0.08)

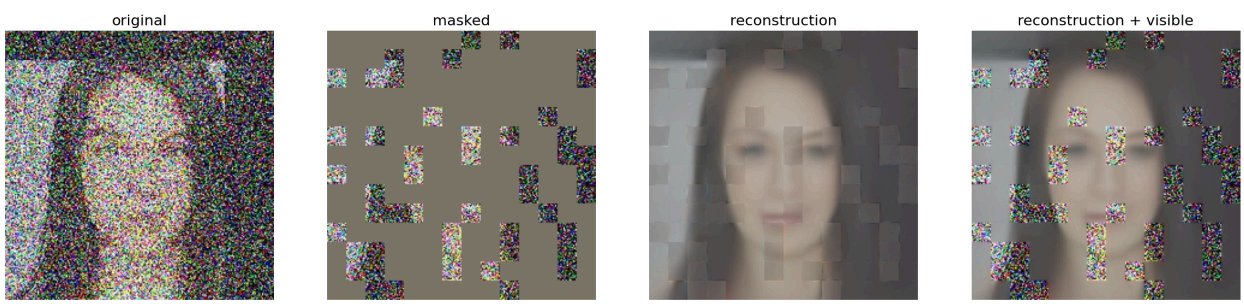

Figure 5. Reconstructed Picture

(gaussian_noise-standard deviation=0.7 ;salt_and_pepper_noise-probability=0.09)

The training loss is 0.4063, indicating the model has converged significantly compared to the initial loss of 0.655. The learning rate is extremely low (1.01e-05), suggesting the model is in its fine-tuning phase with minimal parameter updates. This epoch reflects stable training, with diminishing returns in further optimization, making it a good point to evaluate validation performance.

```
"train_lr": 1.0097998225769366e-05, "train_loss": 0.4063461707267433
```

Figure 5

We have evaluated the performance of our model in terms of image quality reconstruction. The average MSE is 0.0213, showing that we have effectively minimized errors. The average PSNR is 16.54 dB, reflecting the quality of the reconstructed outputs.

```
Average MSE: 0.021264372367391572
Average PSNR: 16.54205106820968 dB
```

Figure 6

We use MAE's output as the input to ResNet for further processing and compare its performance with our GAN+ResNet. Through hyperparameter tuning for our GAN+ResNet, the results indicate that the performance of GAN+ResNet is slightly better than that of MAE+ResNet in handling noisy datasets.

Validation

Models will be evaluated both on the noisy and original datasets. The key comparison will be the accuracy change before and after adding noise to the model. The ResNet+GAN model needs to show improvement (has a better accuracy) in handling the noisy images compared to the baseline ResNet to prove the improvement of model robustness. And it needs to be equivalent to or even better than Mae.

Results:

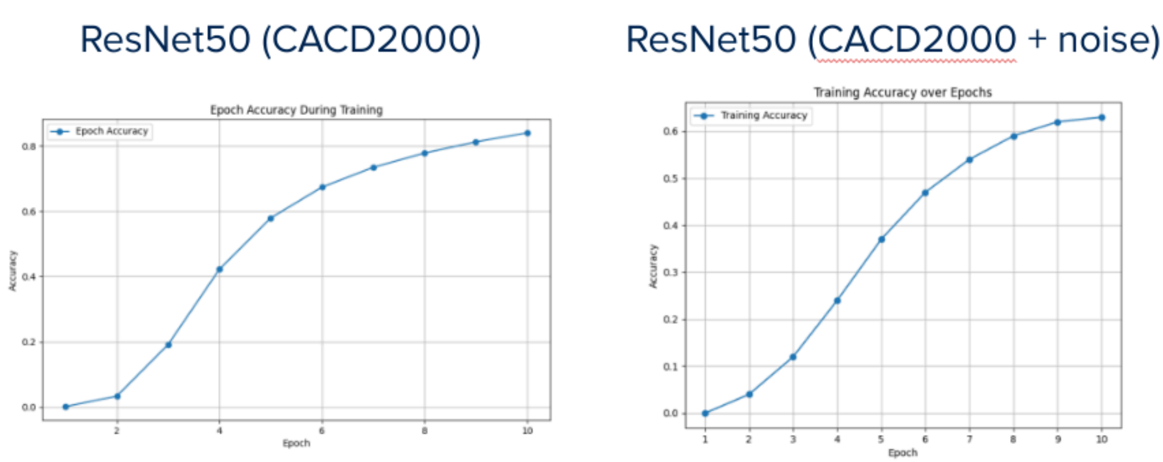

Figure 7. Accuracy Plottings of ResNet50 on CACD2000 and CACD2000+noise

Figure 7 compares the training performance of ResNet50 on the CACD2000 dataset, with and without added noise. On the original dataset (left), the model achieves a higher final accuracy, nearing 0.8 by the 10th epoch, with a smooth and stable progression. In contrast, training on the noisy dataset (right) leads to slower initial improvements and a lower plateau around 0.6–0.65 with a noticeable flattening in later epochs, reflecting the increased complexity caused by noise. This demonstrates the model's sensitivity to noisy inputs, where the disrupted patterns hinder its ability to generalize effectively.

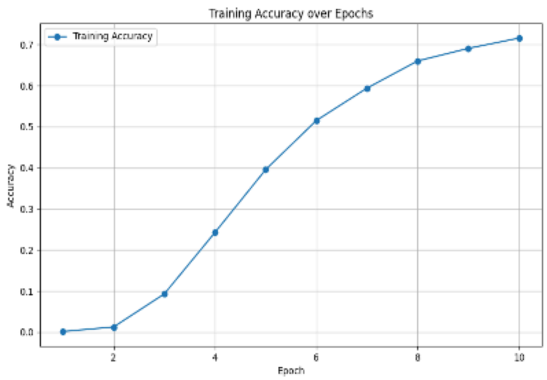

Figure 8. Accuracy Plottings of ResNet50+GAN and ResNet50+MAE

Figure 8 compares the training performance of ResNet50 combined with two different methods—GAN and MAE—on the noisy CACD2000 dataset. Both approaches lead to a final accuracy of approximately 0.7 by the 10th epoch, but their learning behaviors differ. The ResNet50 + GAN (left) demonstrates a gradual increase in accuracy, with slower initial learning during the first four epochs, followed by steady improvements. In contrast, the ResNet50 + MAE (right) achieves a faster initial accuracy increase, showing significant gains within the first five epochs before the curve flattens and progresses more slowly.

Conclusion

The experiments demonstrate that integrating GAN with ResNet significantly enhances the model's robustness to noisy images compared to the baseline ResNet. While both ResNet+GAN and ResNet+MAE achieve comparable final accuracies on the noisy dataset, ResNet+GAN shows steadier and more consistent learning, making it more reliable for long-term improvements.

Reference

ResNet50: https://github.com/KaihuaTang/ResNet50-Pytorch-Face-Recognition

Original dataset: https://bcsiriuschen.github.io/CARC/

Chen, B.-C., Chen, C.-S., & Hsu, W. H. (2015). Face recognition using cross-age reference coding with cross-age celebrity dataset. *IEEE Transactions on Multimedia*.

M. MohammadAmini, D. Matrouf, J. -F. Bonatsre, S. Dowerah, R. Serizel and D. Jouvet, "A Comprehensive Exploration of Noise Robustness and Noise Compensation in ResNet and TDNN-based

Speaker Recognition Systems," 2022 30th European Signal Processing Conference (EUSIPCO), Belgrade, Serbia, 2022, pp. 364-368

Verma, R., & Ali, J. (2013). A comparative study of various types of image noise and efficient noise removal techniques. *International Journal of Advanced Research in Computer Science and Software Engineering, 3*(10), 2277–128X.

Ma, L., Shuai, R., Ran, X. *et al.* Combining DC-GAN with ResNet for blood cell image classification. *Med Biol Eng Comput* **58**, 1251–1264 (2020).

Hu, K., Liu, Y., Nie, J., Zheng, X., Zhang, W., Liu, Y., & Xie, T. (2023). Rice pest identification based on multi-scale double-branch GAN-ResNet. *Frontiers in Plant Science, 14*.

Roles and responsibilities

Qianqian Tan - ResNet + Gan
Yanfeng Ma - ResNet + Mae and dataset
Jiatai Huang - Model training and adjustment