COSE474-2023F: Final Project
"binary classification with noise condition"

Jaeung Yun

Abstract

Our society is accompanied by numerous video devices such as CCTV and black boxes. While advancements in video technology have alleviated many noise-related issues, potential challenges still persist. Therefore, in this project, we investigated how changes in Gaussian noise impact object classification. The results indicated that the performance does not strictly correlate with the magnitude of variance. These findings provide a substantial basis for in-depth discussions and open up avenues for further performance evaluations under diverse conditions. Building on this research, we anticipate that the insights gained could be effectively applied to enhance performance in tasks such as image denoising and data augmentation in the future.

1. Introduction

In contemporary society, the widespread adoption of black boxes and CCTV has made the acquisition of high-quality videos increasingly crucial. However, challenges arise in scenarios with low light conditions or excessive noise, making effective video capture difficult. This research focuses on exploring video and image restoration technologies in environments characterized by low light and high noise, aiming to overcome these constraints. The studies range from utilizing optical flow for restoration in noisy situations (Buades et al., 2016) to enhancing low-quality images by four times through SRGAN (Ledig et al., 2017) and using DCGAN for CCTV video restoration (Oh et al., 2021).
This research aims to explore the limitations of basic CNNs in noisy conditions, deviating slightly from past studies. The study is expected to contribute to applicable solutions for real-world problems by improving the image quality of black box and CCTV footage. The goal is to enhance the efficiency of black box and CCTV systems, thereby broadening their applicability in the fields of safety and security.

1.1. Image Denoising

Digital image noise has been a concern since the advent of digital images. Noise introduced during compression and transmission processes, arising from inherent errors in these procedures, has the potential to result in issues such as loss of clarity. Therefore, research aimed at effectively reducing such noise has been essential and garnered attention (Tian et al., 2020). Among the various types of noise, Gaussian noise stands out as one easily encountered in everyday life. Termed white noise, this noise is so named due to its adherence to a normal distribution.

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{1}$$

Various approaches have been undertaken to mitigate Gaussian noise, including methods utilizing local statistics (Nguyen et al., 2010) and those employing nonlinear algorithms (Russo, 2003). Furthermore, strategies have been proposed for situations involving a mixture of Gaussian and other types of noise (Noise, 2010).

1.2. Convolution Neural Network

Convolution Neural Network(CNN) is a deep learning technique widely employed across various domains at the present time. The primary mechanism of CNN involves partial computation of the entire image through kernels. Convolution operation is defined as follows, and it is computed using the tensors of the entire image and the kernel.

$$(f*g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)\,d\tau \tag{2}$$

The introduction of this approach can be traced back to LeNet in 1998(LeCun et al., 1998), and since then, it has been utilized as a backbone in various architectures, including res2net(Gao et al., 2019), cspnet(Wang et al., 2020), and others.

2. Problem definition & challenges

In this project, we will explore the classification performance by adding noise to images and conducting classification on the images in their noisy state. The motivation behind defining this problem is to investigate whether restoration is necessary when recognition is successful even in situations with significant noise. The variance will be divided into approximately three levels, ranging from low to high noise, to observe how the classification performance varies under different noise conditions.

3. Experiments

3.1. Data preprocessing

We fundamentally considered accuracy to be crucial in scenarios involving CCTV and black boxes. Therefore, we chose the "Bike and Motorbike" dataset on Kaggle for our analysis. The objective of this dataset is binary classification between bike and motorbike, consisting of 6046 bike images and 6078 motorbike images.
As these images vary in size, we will resize them to 64*64 dimensions for use in our data. As mentioned in the problem definition, we will introduce Gaussian noise with variances of 25, 50, and 100, respectively, and evaluate the loss for each scenario.At this point, the mean is set uniformly to zero, and we observe the results accordingly.

3.2. Model

I kept the model relatively simple as excessively complex models might not allow for a smooth measurement of differences across noise levels. I constructed the model with three sets of Conv2d layers, ReLU activation, and pooling, followed by a linear layer at the end. The model summary is as follows: This model architecture consists of three sets of Conv2d layers, each followed by ReLU activation and pooling, and finally a linear layer at the end. The total number of parameters is 64,554, all of which are trainable.

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 16, 62, 62]	448
ReLU-2	[-1, 16, 62, 62]	0
MaxPool2d-3	[-1, 16, 31, 31]	0
Conv2d-4	[-1, 32, 30, 30]	8,224
ReLU-5	[-1, 32, 30, 30]	0
MaxPool2d-6	[-1, 32, 15, 15]	0
Conv2d-7	[-1, 64, 12, 12]	32,832
ReLU-8	[-1, 64, 12, 12]	0
MaxPool2d-9	[-1, 64, 6, 6]	0
Linear-10	[-1, 10]	23,050

Table 1. Model Architecture Summary

Furthermore, throughout this project, I divided the dataset into training, validation, and test sets with proportions of 80/10/10 respectively. I utilized the PyTorch library in Python for implementation. With 5 epochs, cross-entropy loss, and the SGD (Stochastic Gradient Descent) optimization method, I conducted the training on a V100 machine in the Google Colab environment, yielding the results.

4. Result

The following are the results of checking the loss for each epoch with variances of 25, 50, and 100. The results for variance 25 are represented in blue, variance 50 in orange, and variance 100 in gray. Through these results, we discovered that the reduction in loss is similar between variance 25 and 50, contrary to our hypothesis before the project. Initially, we hypothesized that the differences in loss would be proportional, with a factor of 1x, 2x, and 4x for variances 25, 50, and 100, respectively. However, this assumption was overturned based on our observations. Furthermore, the following displays a portion of the validation results, specifically when the variance is set to 100. Despite the considerable noise, there were instances where it was challenging for the human eye to identify regions with background. However, it is noteworthy that the CNN demonstrated proficiency in distinguishing and classifying such regions accurately.

Epoch / Standard Deviation	25	50	100
1	0.7103	0.7262	0.7406
2	0.4823	0.5130	0.6047
3	0.3864	0.3816	0.5226
4	0.3155	0.3291	0.4140
5	0.3039	0.3003	0.3661

Table 2. Loss values for each standard deviation and epoch

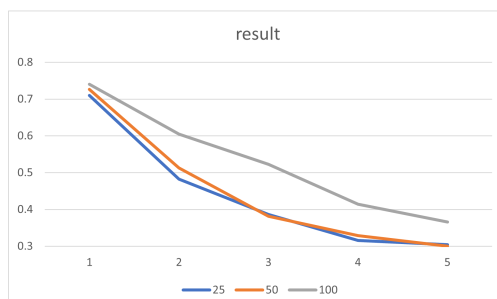

Figure 1. loss result

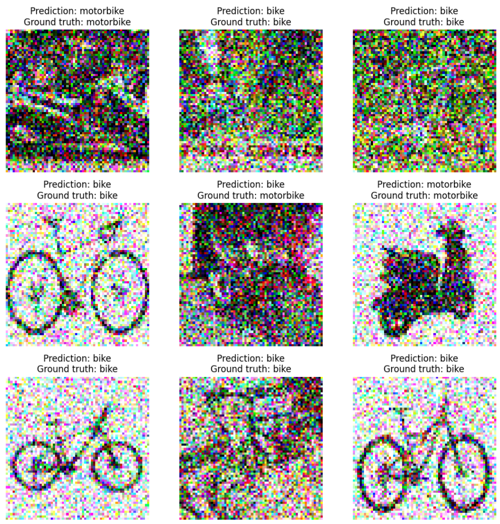

Figure 2. std=100: Validation

5. Conclusion

Through this project, I implemented the basic operations of a CNN from scratch, and the results turned out to be informative, albeit different from what was initially anticipated. However, the unexpected outcomes have opened up new challenges that we should discuss together in the future. Despite the noise following a consistent normal distribution, the uniformity in the magnitude of the loss discrepancy as the variance increased was not observed. It might be worthwhile to explore whether there is any underlying pattern, or if the increase in variance exhibits nonlinear characteristics rather than linear ones. Additionally, it could be interesting to investigate how the difference in loss behaves based on the degree of obscuration in the image, as opposed to Gaussian noise. If bike and motorbike coexist in the absence of localization and share the same size, considering which object is more likely to be classified may be an intriguing aspect to explore.

References

Buades, A., Lisani, J.-L., and Miladinović, M. Patch-based video denoising with optical flow estimation. *IEEE Transactions on Image Processing*, 25(6):2573–2586, 2016. doi: 10.1109/TIP.2016.2551639.

Gao, S.-H., Cheng, M.-M., Zhao, K., Zhang, X.-Y., Yang, M.-H., and Torr, P. Res2net: A new multi-scale backbone architecture. *IEEE transactions on pattern analysis and machine intelligence*, 43(2):652–662, 2019.

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., and Shi, W. Photo-realistic single image super-resolution using a generative adversarial network, 2017.

Nguyen, T.-A., Kim, M.-J., and Hong, M.-C. Fast and efficient gaussian noise image restoration algorithm by spatially adaptive filtering. In *28th Picture Coding Symposium*, pp. 122–125, 2010. doi: 10.1109/PCS.2010.5702438.

Noise, M. P.-G. Image denoising in mixed poisson-gaussian noise. 2010.

Oh, G., Lee, J., and Jeon, B. Color noise detection and image restoration for low illumination environment. *Journal of Broadcast Engineering*, 26(1):88–98, 2021.

Russo, F. A method for estimation and filtering of gaussian noise in images. *IEEE Transactions on Instrumentation and Measurement*, 52(4):1148–1154, 2003.

Tian, C., Fei, L., Zheng, W., Xu, Y., Zuo, W., and Lin, C.-W. Deep learning on image denoising: An overview. *Neural Networks*, 131:251–275, 2020.

Wang, C.-Y., Liao, H.-Y. M., Wu, Y.-H., Chen, P.-Y., Hsieh, J.-W., and Yeh, I.-H. Cspnet: A new backbone that can enhance learning capability of cnn. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pp. 390–391, 2020.