Exploring Curriculum Learning for Image Denoising

BY:
Dwijesh Jigarbhai Shah - ds1991
Haarika Ambati - ha539
Hrishikesh Sunil Salunkhe - hs1121
Jash Jayant Shah - js3297
Kush Jayank Pandya - kp1127
Shreya Rajendra Pai - sp2305

Abstract

Curriculum Learning is a training strategy that trains a Machine Learning model from more straightforward to more challenging data. The idea is to use the strategy of how humans learn, i.e., learning easier problems first and taking more complex problems afterward. It has shown its advantages by showing faster convergence, better generalization, and use less computation while training compared to other machine-learning models. In this project, we will use Curriculum Learning for Image Denoising, a process to remove corruption in an image caused by different factors. We aim to remove this type of corruption in an image with high computational and performance efficiency.

1 Introduction

For a better understanding of our project and our approach, we will discuss a few topics one by one:

1.1 Curriculum Learning

In the movie Iron Man[1], Tony Stark tells Jarvis-

"Sometimes you got to run before you walk."

This statement is valid very rarely. Usually, humans must learn the easy things first, clear their concepts and do complex tasks. For example, a child first learns basic arithmetic, then learns algebra before solving calculus problems. We take the idea of using a Curriculum for humans to learn and apply it to Machine Learning.

Curriculum Learning[2,3,4] is a training strategy that trains a Machine Learning model from more straightforward to more challenging data. The idea is to use the strategy of how humans learn, i.e., learning easier problems first and taking more complex problems afterward. It has shown its advantages by showing faster convergence, better generalization, and use less computation while training compared to other machine-learning models. *In this project, we will not only increase the difficulty level of the data but also modify our neural network architecture as we progress.*

The general idea of Curriculum Learning depends on two things:
1) Difficulty measurer
2) Training Scheduler (sometimes called scoring and pacing functions, respectively).
The difficulty measurer is a metric to define the difficulty level of the data. The training scheduler decides the sequence of data subsets throughout the training process based on the current model expertise.

1.2 Image Denoising

Image Denoising is a process to remove corruption in an image caused by different factors. Denoising an image is difficult since the noise is tied to the image's high-frequency content, i.e., the details. As a result, the goal is to strike a balance between suppressing noise as much as possible and not losing too much information.

Owing to different factors of environment, transmission channel, and other influences, images are inevitably contaminated by noise during capturing, compression, and transmission, leading to loss of information and other distortion. Due to this noise the further process of image processing , such as image analysis and tracking would be affected. Therefore, image denoising plays an important part image processing systems.

1.2.1 Types of Corruption

In this project, we are going to handle different types of frequent image noise[5], which are:-

1. Gaussian Noise: This is a statistical noise whose presence in statistical data is attributed to its probability density function (PDF), which corresponds to the normal distribution. This noise originates from camera electronic circuits and sensor noise due to poor illumination or high temperature.
2. Salt and Pepper Noise: Its common noise seen in photography, they are bright and black pixel/small spots which is caused by dead pixels or, error in data transmission and failure in memory cell.
3. Speckle Noise: Speckle noise is a distinct form of noise, unlike Gaussian or Salt and Pepper noise, that is multiplicative in nature. When it comes to diagnostic imaging, this noise can create images that have the appearance of a back scattered wave due to a large number of microscopic dispersed reflections that pass through internal organs. As a result, it becomes more challenging for the observer to discern fine details in the images. This noise is present in a diverse array of systems, including ultrasound imaging, synthetic aperture radar (SAR) images, and various others.
4. Poison Noise: Poisson noise is generated as a result of the nonlinear responses of image detectors and recorders. This noise is dependent on the image data and is created due to the arbitrary electron emission associated with detecting and recording procedures, which follows a Poisson distribution with a mean response value.

1.2.2 Types of Error Metrics

In our project, we compare our original image with the resultant denoised image based on the following comparison metrics[6]:

1. MAE/L1: The absolute difference between each predicted value and its associated actual value is added to determine the mean absolute error (MAE), then divided by the total number of data points. An indicator of the effectiveness of denoising in pictures is the average absolute difference, which compares the original image and the denoised image, wherein a lower MAE indicates better denoising performance.
2. MSE/L2: The average squared difference between the original and denoised images is measured by the mean squared error (MSE) or L2 loss in denoising images; a lower MSE suggests better denoising performance.
3. SSIM: A technique for forecasting the observed quality of digital television, cinematic images, and other types of digital images and videos is called the Structural Similarity Index Measure (SSIM). The distinction between these methods and others, like PSNR, is that they predict absolute errors. The concept of structural information that the pixels hold have strong interdependencies, particularly when they are spatially near one another. These dependencies carry important details about the structure of the items in the visual scene.

4. PSNR: Peak signal-to-noise ratio (PSNR) is an engineering term for the ratio of a signal's maximum possible power to the power of corrupting noise that compromises the representation of the signal's fidelity. Most frequently, it is used to measure the quality of the reconstruction of lossy compression codecs. When comparing compression codecs in image compression, PSNR approximates the human perception of reconstruction quality.

2 Methodology

2.1 Our Approach

Curriculum learning has great potential, but it comes with its own set of challenges. Most importantly, the overhead of classifying difficulty should be less than directly training the model. One reason that curriculum learning should be effective in image denoising is that its straightforward to classify data into difficulty levels. It is because we can easily control the amount of corruption in an image, thus setting the difficulty level of that image data.

The idea would be to start training the model with relatively low corrupted data (i.e., a low difficulty level), where the gradient descent will keep decreasing until it drops below a predefined threshold error value. It will signify that the model has gained expertise for that difficulty level and can be trained for higher difficulty levels. Thus, we introduce incrementally difficult data for the next iteration of training the model until the gradient descent drops below the threshold value. We repeat this iterative process until the gradient descent reaches a point where it becomes saturated above the threshold value.

Once the model reaches this stage, we introduce additional layers to the model architecture, thus increasing its complexity which would enable it to be trained on more difficult data. Doing this will initially exhibit a huge spike in the gradient descent since the newly added layers would have random weights initialized. However, we will observe a faster convergence of the gradient descent due to the existing knowledge of the previous layers. The expectation is that executing this process would finally enable the gradient descent to drop below the threshold value, which it could not overcome with the old set of layers.

We repeat this process of iteratively raising the difficulty of data and adding new layers to the model until we reach a point where either the gradient descent would not drop below the threshold value or we reach the highest difficulty level of the training data. It would indicate that the model will no longer benefit from adding new layers to the architecture. Thus, we will end the model training and achieve a model with optimal expertise in the data domain.

2.2 Comparison

For measuring the performance of the final model obtained using our Curriculum Learning approach, we will replicate the final model architecture into a newly generated model initialized with random weights. We would then train this new model using our entire training data without classifying it into different difficulty levels. Using the error metrics mentioned in section 1.2.2 and the execution time required for model training, we will compare the models obtained by the two approaches to understand the benefits, differences and potential drawbacks of our Curriculum Learning approach.

3 Expected Results

Having made the comparisons, we expect to see our Curriculum Learning approach outperforming the traditional model training approaches in terms of training time, computation power and accuracy.

References

[1]‘Sometimes you gotta run before you can walk’ | Iron Man flight scene | Iron Man (2008),” www.youtube.com. https://youtu.be/39yo2JjvtsI (accessed Mar. 06, 2023).

[2] X. Wang, Y. Chen, and W. Zhu, ‘A Survey on Curriculum Learning’. arXiv, 2020.

[3] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, ‘Curriculum Learning’, in Proceedings of the 26th Annual International Conference on Machine Learning, Montreal, Quebec, Canada, 2009, pp. 41–48.

[4] A. M. Hamdi, “Curriculum Learning in Deep,” Omdena, Feb. 11, 2022. https://medium.com/omdena/curriculum-learning-in-deep-d52fa1cbe188 (accessed Mar. 06, 2023).

[5] V. Lendave, “A Guide to Different Types of Noises and Image Denoising Methods,” Analytics India Magazine, Sep. 25, 2021. https://analyticsindiamag.com/a-guide-to-different-types-of-noises-and-image-denoising-methods/

[6] H. Zhao, O. Gallo, I. Frosio and J. Kautz, "Loss Functions for Image Restoration With Neural Networks," in IEEE Transactions on Computational Imaging, vol. 3, no. 1, pp. 47-57, March 2017, doi: 10.1109/TCI.2016.2644865.