Neural Network for Image Transformation

Brezuleanu Mihai Alexandru

17 decembrie 2023

1 Introduction

The primary objective of this assignment is centered on constructing a neural network capable of mimicking a composite of PyTorch transforms, including resize, flip, and grayscale operations. The goal is to optimize this neural network to surpass the efficiency of traditional transforms in terms of time complexity.

2 Implementation

2.1 Neural Network Arhitecture

My objective is to develop a neural network that can learn a combination of image transforms, including resizing, horizontal and vertical flipping, and grayscaling. The architecture is designed with simplicity in mind: the input layer accommodates images of size $3 \times 32 \times 32$, and the output layer produces transformed images with dimensions $1 \times 28 \times 28$.

2.2 Loss function

The chosen loss function, Mean Squared Error (MSE), serves a crucial role in the proposed method. MSE stands out as a fitting choice for generating transformed images owing to its effectiveness in calculating pixel-wise differences between the generated and target images. By computing the average squared difference between corresponding pixel values, MSE provides an intuitive measure of the overall reconstruction error.

Especially in transformations like rotation, resizing, or grayscale conversion, where preserving intricate details and precise pixel values is of utmost importance, MSE guides the model to minimize these differences. This, in turn, encourages the generation of outputs that closely align with the intended transformations. Furthermore, MSE's compatibility with optimization algorithms contributes to efficient convergence during training.

While alternative loss functions might be considered based on specific use cases or priorities, the straightforward nature of MSE, coupled with its alignment with the goal of faithful image reconstruction, establishes it as an effective choice across various image transformation scenarios.

2.3 Early Stopping

Early Stopping emerges as a valuable asset in the training of neural network models, playing a pivotal role in averting overfitting and optimizing the model's generalization performance. This mechanism, facilitated by the Early Stopping, monitors the validation loss throughout the training process. If the validation loss fails to exhibit improvement over a specified number of epochs, determined by the patience parameter, the training is halted.

Early Stopping prevents prolonged training, a circumstance that could lead to overfitting. The delta parameter establishes a minimum threshold for defining a meaningful improvement, ensuring that only substantial decreases in validation loss trigger the reset of the count.

Therefor, the Early Stopping contributes to the efficiency of the training process by intervening once the model's performance on the validation set plateaus. This intervention enhances the model's capability to generalize to new data, which also helps its overall predictive performance.

3 Performance and Results

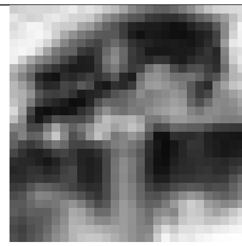
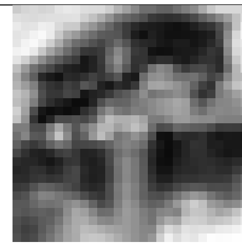
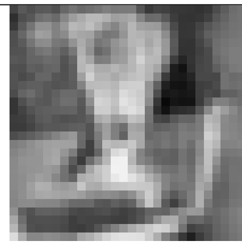
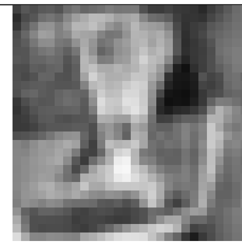
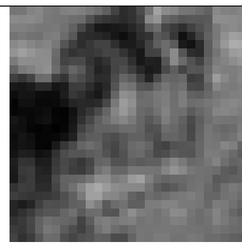
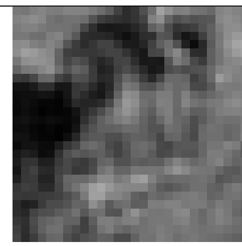
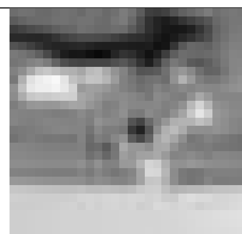
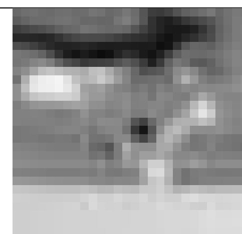
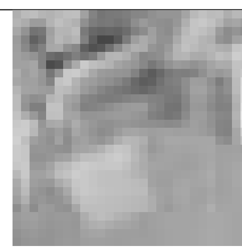
Input	Label	Model Output
		
		
		
		
		

Tabela 1: Comparison between label and the NN's output

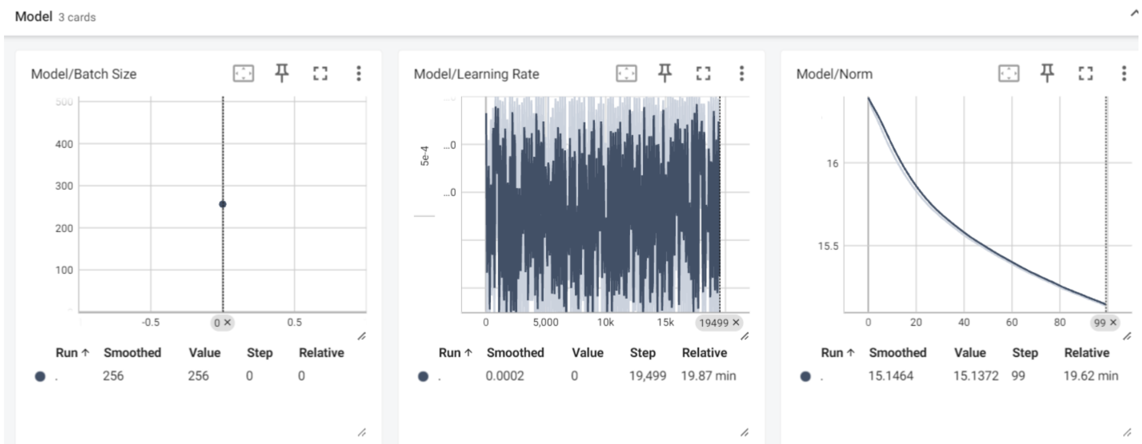


Figura 1: Model

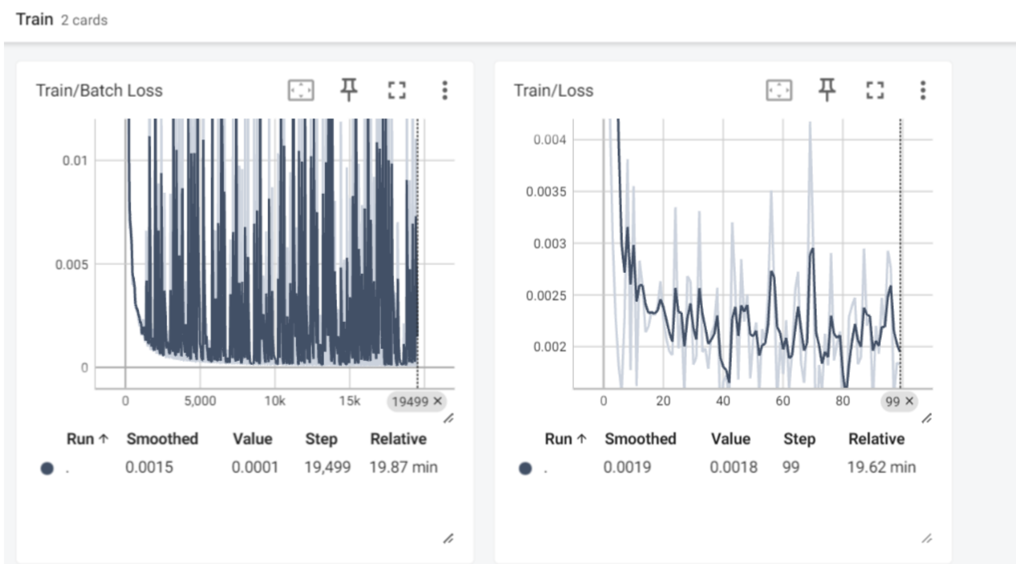


Figura 2: Training

Device	Sequential Transforms Time	NN Time
CPU	2.53408	0.608607
GPU	2.24578	0.472566

Tabela 2: Device Time Comparison

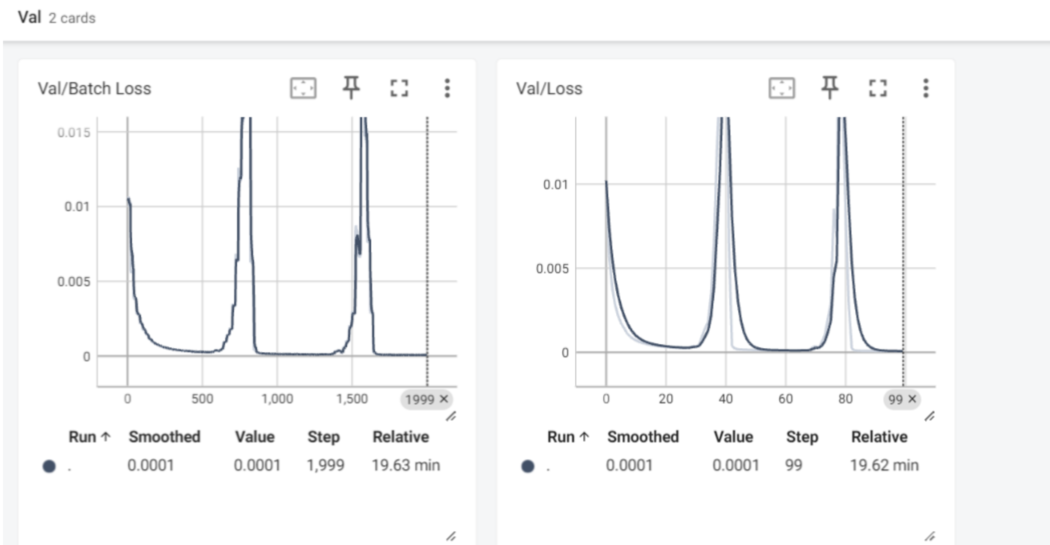


Figura 3: Validation

4 Conclusion

In summary, the executed neural network for this assignment has not only met but exceeded the anticipated objectives, delivering the sought-after results. The network's straightforward architecture alongside the training process not only live up to but also surpass the set expectations.