Acta of Turin Polytechnic University in Tashkent, 2022, 31, 24-27
Published Online March 2022 in Acta TTPU (http://www.acta.polito.uz/)

NEURAL NETWORKS PERFORMANCE IMPROVEMENT WITH GAUSSIAN WHITE NOISE AUGMENTATION

Abdurashitova Muniskhon
Department of Control and Computer Engineering, Turin Polytechnic University in Tashkent, Uzbekistan
Email: abdurashitova.munisxon@gmail.com

Abstract– Data augmentation is a strategy for creating synthetic data from existing data by adding slightly changed copies of current data to expand the amount of available data. When training a machine learning model, it functions as a regularizer and helps to reduce overfitting. This work will discuss the improved neural network performance with the data augmentation method employing Gaussian white noise.

Key words– Gaussian White Noise, Neural Networks, Data Augmentation Techniques

I OBJECTIVE

Training neural network models on additional data can enhance their capacity to generalize what they have learned to new data. Augmentation techniques can provide variations of datasets that can improve the fit models' ability to generalize what they've learned to new unseen data. We will use data augmentation strategies to improve model inference generalization robustness when training neural networks in this paper.

II INTRODUCTION

Data augmentation refers to a set of strategies for creating new training samples from existing ones by introducing random variations and disturbances, ensuring that the data is not destroyed. Our purpose is to boost the model generalizability using data augmentation. We discuss the generation of augmented sets with the addition of Gaussian white noise to original data. We are exploring human indoor localization in a restricted area because it is trendy in the fields like energy management, health monitoring, and security. The purpose is to explore cheap but efficient techniques for indoor localization because for outdoors person tracking, there exists GPS technology that can quickly determine a person location through the wearable tag or cell phone. At home or inside a room, a person is not always carrying a phone; that is why we should search approaches for tagless localization. Four sets of experimental data are collected in a 3x3 meter room for a short period. A 4x4 pixel Omron D6T-44L-06 thermopile infrared sensor is installed on the ceiling of a room, and the person reference location is collected with an ultrasound-based tag of the Marvelmind Starter Set HW v4.9. Each tuple of experimental data has 18 elements: 16 pixels of the infrared sensor (IR) plus the X and Y coordinates of the person representing the label.

III MAIN PART

The first set of experimental data is used for the model training in a 60/20/20 ratio representing training/validation/test sets. The other three sets are used only for inference because our purpose is to improve the model generalization, which refers to model performance for unseen data. Since the amount of training data is small, we generate synthetic data by adding Gaussian white noise. Noise has a normal distribution with zero mean and finite variance

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}. \quad (1)$$

Function *f(x)* is the probability density, σ is the standard deviation, μ is the mean.

Experimental analysis

For the Gaussian white noise, the standard deviation parameter allows modulating the noise amplitude. That is why the first experiment trains the model with the following values {0.01, 0.11, 0.21, 0.31, 0.41, 0.51, 0.61}. Table 1 represents the output of the model for each white noise parameter. For each white noise amplitude in the range {0.01, 0.11, 0.21, 0.31, 0.41, 0.51, 0.61}, based on the best training giving the smallest Overall_Performance, obtained MSE metrics for sets A, B, C and D are given in the table 1.

Now we can compare the model results trained on augmented data with our baseline. Let us represent the model

Noise Amplitude	MSE_A	MSE_B	MSE_C	MSE_D
Baseline	0.001565	0.040742	0.113454	0.080686
0.01	0.001904	0.037636	0.033084	0.062054
0.11	0.002187	0.055132	0.053950	0.068099
0.21	0.001810	0.052376	0.054861	0.075765
0.31	0.001577	0.055950	0.075206	0.090898
0.41	0.002161	0.070060	0.051307	0.077119
0.51	0.002728	0.068208	0.040365	0.085448
0.61	0.002427	0.076800	0.035568	0.093364

TABLE 1: MODEL TRAINING RESULTS AT DIFFERENT WHITE NOISE AMPLITUDES ARE COMPARED TO THE BASELINE. FOR EACH NOISE PARAMETER CORRESPONDING MSE METRICS FOR SETS A, B, C, D ARE SHOWN.

results in a graphical way so that we can easily extract the features. Figure 1 shows that for set A, the model behavior fluctuates in an interval. For smaller values of the noise amplitude, the model output seems to have smaller MSE values, while for higher noise values, MSE is also increasing. Overall results show that model training with augmented sets does not improve with respect to the baseline. The model shows a more coherent behavior for the set B. The MSE is steadily rising for increasing amounts of noise, showing improvements in a neighborhood of 0.01 noise standard deviation. Set C is the only one fully improved by the training with augmented data. It has smaller MSE values compared to baseline over the whole noise range. Finally, set D expresses better model improvement up to noise standard deviation 0.5 except for the peak around a point 0.3 and increasing MSE for higher noise standard deviations. Since we are interested in overall model generalization quality, we should merge all independent results of the four sets. In the noise amplitude under the exploration, all sets show different characteristics, one showing very significant improvement while others without any improvements. When we sum MSE metrics, significant improvements of one or two sets could exceed losses in other sets, so it could be a misleading point to decide as a better model generalization. To avoid these kinds of "misleading" points after generating the overall sum of MSEs, we extract the areas with smaller MSE sum values with respect to the baseline(obtained without augmentation). We should further inspect extracted areas for each set independently. If at least three sets have improvements in these intervals, we can decide that the model has improved generalization for all sets. Suppose less than three sets show more extensive improvements that surpass the unsatisfactory results of other sets. In that case, these noise intervals are not considered as a noise range with better generalization quality. Going back to our results in the figure 2, we have overall model results for all sets. It is showing improvement in a total interval of investigated noise standard deviation. However, it is not enough to decide on the improved generalization. We go back to independent plots of sets and search for the areas with more than two sets showing smaller MSEs than the baseline. It gives the range roughly around zero.

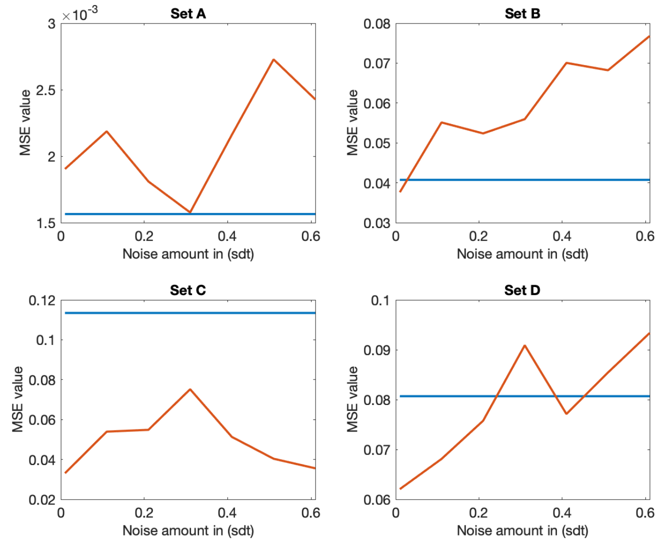

Fig. 1: Model output MSE values(red line) obtained by augmentation with white noise amplitude range [0; 0.6], for data sets A, B, C and D. Blue line represents the model MSE baseline obtained without augmented data.

We perform another experiment in the same interval, but with a higher resolution. White noise standard deviations, which are directly proportional to white noise amplitudes, are chosen as follows: {0.01 0.06 0.11 0.16 0.21 0.26 0.31 0.36 0.41 0.46 0.51 0.56 0.61}. When we analyze the results for every 30 trainings best model is chosen near to 200 epochs, which means that the model was still able to learn. According to the second experiment, model training with specified noise parameters (by performing 30 trainings for each noise value and choosing the best one) generates the following results given in Table 2. Since the resolution of the given interval is higher than the previous experiment, we can precisely extract the sections of the range with better model generalization. Let us visually analyze the model output with graphs to determine the most promising areas in Figure 3.

Mean square error loss function values for set A show smoother behavior in this experiment. MSE value is slightly higher than the baseline at the beginning of the noise range. Then it gradually decreases below the baseline, demonstrating model improvement until around 0.3. After that, it continuously increases with increasing noise amplitude. It is evident that more significant noise amounts does not help the training generate improvements for set A. Analyzing the plot of set B, we can say that set B is less resistant to noisy data

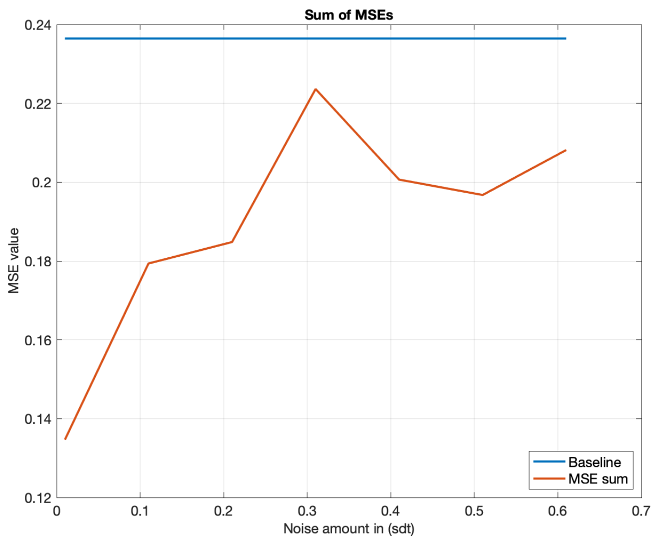

Fig. 2: The model overall inference generalization obtained by white noise augmentation.

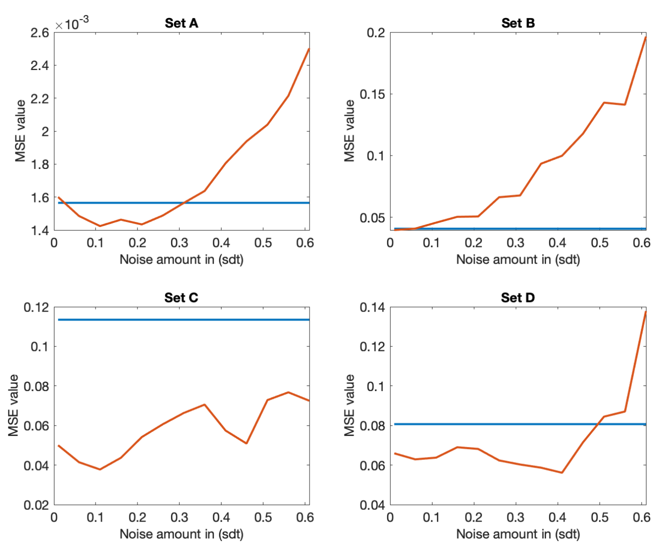

Fig. 3: The second experiment. Model output MSE values(red line) with augmented data generated in white noise amplitude range [0; 0.6], for data sets A, B, C and D. Blue line represents the model MSE baseline obtained without augmented data.

Noise Amplitude	MSE_A	MSE_B	MSE_C	MSE_D
Baseline	0.001565	0.040742	0.113454	0.080686
0.01	0.0016	0.039653	0.049996	0.065936
0.06	0.001485	0.040832	0.041435	0.062892
0.11	0.001424	0.04568	0.037765	0.063757
0.16	0.001463	0.050361	0.043676	0.069003
0.21	0.001434	0.05065	0.05417	0.068153
0.26	0.001488	0.066244	0.060627	0.062341
0.31	0.001564	0.067676	0.066316	0.06029
0.36	0.001637	0.093435	0.070558	0.05868
0.41	0.001805	0.099847	0.057439	0.056119
0.46	0.001938	0.117797	0.050876	0.071523
0.51	0.002039	0.142877	0.072843	0.084473
0.56	0.002214	0.141222	0.076763	0.087068
0.61	0.002502	0.196399	0.072507	0.137806

Table 2: Result of the model trainings at different white noise amplitudes. For each noise parameter corresponding MSE metrics for sets A, B, C, D and best training are shown.

showing slight improvement of the model for small amounts of noise approximately up to 0.05. After that point, the model MSE value starts to increase gradually as for set A. Set C has total improvement in this range, as in the previous experiment. Considering the graph of set D, we can say that we could generate more stable and understandable behavior with respect to the previous experiment. It shows model improvement in a noise range of [0; 0.5], which is the same as the initial results. After noise amplitude 0.5, the MSE starts to increase abruptly.

Now we can generate a plot to check model overall improvement. The sum of the MSE values for all sets is represented on the vertical axis in Figure 4, while on the horizontal axis, noise amplitudes are described. Overall model behavior shows a gradual MSE increase for increasing amounts of noise amplitude. Up to 0.45, MSE is below the baseline, which means the model has overall improvement. However, Figure 4 is not enough factor to decide on model generalization quality because we should further check the plots of each set to avoid points in which improvements of one or two sets are big so that it surpasses the destructive results of other sets. At the upper end of the noise interval, the overall model characteristic shows higher MSE values than the baseline, which means the model has poor performance.

IV Conclusion

The model overall generalization quality is determined both by the improved MSE sum of all sets and individual improvements of at least three sets. Based on Figure 4, it is ob-

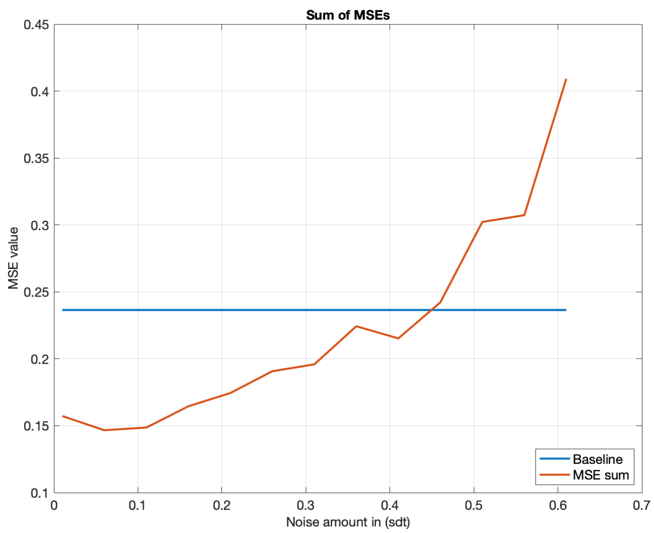

Fig. 4: The model overall inference generalization obtained by white noise augmentation (the second experiment).

vious that MSE sum is below the baseline in a noise range of [0; 0.45]. In this range sets C and D have total improvement. Fields of improvement for sets A and D are smaller than [0; 0.45]. Set A starts to go above the baseline after point 0.3, which means that we should reduce our initial range from [0; 0.45] to [0; 0.3] to have at least three sets with individual improvements. For set B, noise amounts in the range [0; 0.3] could result in MSE values above and below the baseline. However, we can accept this range as a reasonable interval giving better model generalization quality since at least sets A, C, D are performing adequately. The last experiment is accomplished to precisely determine the noise amplitude. All the sets have simultaneous improvement concerning the corresponding baselines and improved overall model generalization. We will not explicitly state the resulting table and plots here. The noise amplitude in the interval [0; 0.046] gives the best model inference generalization quality with all sets individual improvements.

References

1. Tariq, Osama Bin and Lazarescu, Mihai Teodor and Lavagno, Luciano(2021). Neural networks for indoor person tracking with infrared sensors

2. Carter, Mancini, Bruce, Ron (2009). Op Amps for Everyone. Texas Instruments. pp. 10–11

3. https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780471679370.app2

4. https://en.wikipedia.org/wiki/Whitenoise