EECS 453 Project Proposal

Joseph Lynch (lynchjo, 11442632)
November 10th, 2024

1. Introduction

Data augmentation is common practice for training image classification or segmentation machine learning models. Augmentation is used to increase the size and or diversity of a training dataset using a bevy of transformation to training images that include geometric operations, introduction of noise, and combining multiple samples into a single image. Data augmentation is widely agreed to improve training efficacy and model generalizability [1], but which transformation, or combination of transformations is most effective for image classification models? The aim of this project is to explore the implementation of modern and traditional augmentation methods and investigate which of these methods are best suited for training an image classification model.

2. Methods

There are two parts to this project: the creation of an augmentation library and the evaluation of different augmentation compositions. In order to gain a better understanding of the image transformations that are used in modern machine learning, I will create a custom Python library including the following common augmentations:

1. Horizontal Flip
2. Vertical Flip
3. Rotation
4. Perspective Shift
5. Crop
6. Crop and Paste
7. Gaussian Blur
8. Gaussian Noise
9. Adjust Sharpness
10. Adjust Brightness

Next, I will construct a naïve UNET[2] model using the PyTorch library and train it to classify images in the MNIST public dataset[3]. Using different combinations of the above augmentations, I will train the model and evaluate its performance on a segregated test set.

3. Related Work

The performance of both individual and entire categories of data augmentation methods have been extensively explored. Both traditional augmentation techniques, such as rotations or flips, and more modern methods, such as cropping and pasting between samples, have been shown to be effective in improving model generalizability and reducing overfitting[4,5]. However, there have been fewer investigations into the optimal combinations of the existing augmentations; this is where this project will focus.

4. Project Plan

All work will be done in Python. To construct the augmentations from scratch, the NumPy library will be primarily used. To build the naïve UNET model, the PyTorch library will be used. Efficacy of a given composition of augmentations will be determined by the model's loss on a segregated test set of MNIST data after reaching a validation loss of 0.97. Turning to logistics, the primary anticipated challenge will be allocating the required time for model training and evaluation. Regarding project milestones, the augmentation library has a project completion date of December 2nd with training and evaluations to be completed by December 9th.

5. References

1. Shorten, C., Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. J Big Data 6, 60 (2019). https://doi.org/10.1186/s40537-019-0197-0
2. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241). Springer International Publishing.
3. Deng, L. (2012). The mnist database of handwritten digit images for machine learning research. IEEE Signal Processing Magazine, 29(6), 141–142.
4. Kiran Maharana, Surajit Mondal, Bhushankumar Nemade, A review: Data pre-processing and data augmentation techniques, Global Transitions Proceedings, Volume 3, Issue 1, 2022, Pages 91-99, ISSN 2666-285X, https://doi.org/10.1016/j.gltp.2022.04.020.
5. A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," 2018 International Interdisciplinary PhD Workshop (IIPhDW), Świnoujście, Poland, 2018, pp. 117-122, doi: 10.1109/IIPHDW.2018.8388338. keywords: {Image color analysis;Machine learning;Lesions;Image classification;Neural networks;Cancer;Task analysis;Machine learning;style transfer;data augmentation;deep learning;medical imaging},