Data augmentation with Mobius transformations

Derek T. Brown[1] Dariya Mamyrbek[2] Hong Phuc Nguyen[3] Margo Thompson[4]

[1]Purdue University Fort Wayne

PURDUE UNIVERSITY FORT WAYNE

Introduction

In the real world, insufficient data is one of the common issues for machine learning. This problem could possibly lead to a model overfitting. One way to solve this problem is through data augmentation: Creating new data by transforming existing data. "Data Augmentation with Mobius transformations" [3], proposed a novel data augmentation technique using bijective conformal maps known as Mobius transformations, which preserve image-level labels by minimizing local distortions in an image. In this project, we are interested in testing the results obtained in the above-mentioned paper, thus, we expect that with the use of data augmentation, the testing error will decrease. Moreover, we aim to explore the correlation between the neural network's accuracy and the parameters used to perform Mobius transformation on the original images.

Data Sources and Software

The input images we will use in our data augmentation will be taken from the CIFAR-10 dataset. The CIFAR-10 is a collection of images collected by groups at MIT and NYU that are commonly used to train machine learning and computer vision algorithms [2]. It consists of 10 classes of images (airplane, automobile, bird, cat, deer, dog, frog, horse, ship, truck) with 6000 images per class. We will then apply Mobius transformations on a subset of these images to have a larger collection of photos to train a neural network, before testing it with some of the remaining images. The analysis will be performed using MATLAB and its Deep Learning Toolbox. The VGG-16 neural network architecture was used.

Mobius Transformations

Mobius transformations are a type of conformal mapping. A conformal mapping is a mapping that preserves local angles [3]. Mobius transformations also preserve the anharmonic ratio [3].The anharmonic ratio, also known as cross ratio has been used for identifying objects from different perspectives with a high accuracy [1]. The preservation of the this ratio will allow mobius transormations of an image to represent the subject of the image from multiple perspectives.
Mobius transformations can be defined as:

$$f(z) = \frac{az+b}{cz+d} : ad - bc \neq 0$$

Where a, b, c, d are complex numbers. Mobius transforms reflect around the unit circle which allows for different distortions in scale [3]. This allows for different representations of the image at different scales which will aid in training the neural network. One issue that comes up is balancing the correct amount of distortion on an image. Too much distortion will result in the image not representing the object it is classified as. Too little distortion will result in the image serving as a duplicate of the original which is not helpful. This is addressed by adding the following constraint as done by Sharon Zhou et al. [3]:

$$\frac{1}{M} < |f'| < M : M > 1$$

Cross Validation

If we augment our training dataset with the Mobius transformation and consider the few following facts:

- If we train the same network (that is, utilize the same network structure and configuration parameters) using the same training dataset, then the final network parameters will be the same, provided the training is done without validation or with fixed validation sets
- If the network has the same parameters, then its accuracy when testing on the same testing dataset will be the same.
- If we apply the same Mobius transformation (that is, Mobius transformations with the same parameters) to the same dataset, the new dataset that is generated will be the same.

It then follows that if we:

- Fix the initial training and testing datasets
- Fix the validation sets or ignore validation during training
- Fix the network structure and configurations

Then the network's accuracy on the testing dataset can be expressed as a function of the Mobius transformation parameters. We can then build a regression model using the network accuracy as our output and the Mobius transformation parameters as our explanatory variables.

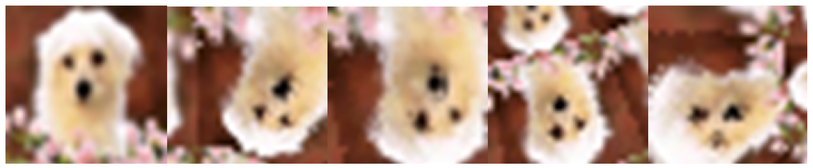

Fig. 1: Image from CIFAR-10 Data Set [2] (leftmost) with Augmentations

Analysis

1000 images were picked from CIFAR-10 with 100 images per class for training. For the testing dataset, 200 images were used with 20 images per class. All images from the training dataset were augmented using Mobius transformation with 32 different sets of parameters (a,b,c,d). The augmented data from each of these parameters set were added to the base training dataset, and each of these augmented dataset was used to train a different neural network. All these neural networks had the same architecture and training parameters. After training, we tested these neural networks on the same testing dataset and compared their accuracy. As discussed in section 2.3, this accuracy can be expressed as a function of the Mobius parameters, so we can analyze the correlation between the Mobius parameters and the according neural network's accuracy.

Results and Discussion

Each neural network trained with augmented images performed better than the control group trained with no additional augmented images. The average neural network accuracy was 27.65625% while the control group only had an accuracy of 5%. With a p-value less than $10^{-13}\%$ it can be concluded that the augmented images aided in training the neural network. We performed linear regression with 8 explanatory variables, that is, the real and imaginary parts of the Mobius transformation parameters. In order to do this, we first transformed the output variable - the neural network accuracy - from a sigmoid function used to express probabilities to a linear function using the formula

$$a = ln(\frac{y}{1-y})$$

and run the least-squares linear regression model. Note that this formula is only a change of coordinates so that our linear regression model yields plausible values (instead of being constrained from 0 to 1), not an attempt at logistic regression. To transform a prediction acquired from this linear model back to a predicted neural network accuracy value, we use the regular sigmoid function:

$$y = 1 - \frac{1}{exp(a)+1}$$

A linear regression model fitted to the 8 variables came out with a multiple R-squared of 0.2756 and an adjusted R-squared value of 0.02362. These values resulted p-value of 0.4024

Conclusion

Even though there was not much correlation between a neural network's accuracy and the Mobius parameters themselves, the accuracy for almost any case with data augmentation would be much higher than the control case - the case without any data augmentation. Hence, we can conclude that Mobius transformation is a helpful tool for data augmentation. With Mobius transformation for data augmentation, the neural network's accuracy in any case is also higher than random guessing. Thus, we can conclude that the data augmentation using Mobius transformations helped the neural network learn to better classify the images.

References

[1] John Fryer. "An object space technique independent of lens distortion for forensic videogrammetry". In: (Jan. 2000).

[2] Alex Krizhevsky, Geoffrey Hinton, et al. "Learning multiple layers of features from tiny images". In: (2009).

[3] Sharon Zhou et al. "Data augmentation with Mobius transformations". In: *Machine Learning: Science and Technology* 2.2 (2021), p. 025016. DOI: 10.1088/2632-2153/abd615.