**Blending Mode Data Augmentation in CNN Classification**

**Abstract**

Model Architecture, Dataset size directly impacts the ability of a Convolutional Neural Network to learn generalized patterns contained in training data. Overfitting occurs when a network learns too specifically to the training data and not general enough for the true scope of test data. A limited dataset in both size and breath produces overfitting in the model. Data Augmentation is used to create new training data from a limited dataset. Classical techniques such as flipping, rotating, translating, and transforming color channels are ways the data can be augmented while not effecting the data’s corresponding label. In this paper I propose a technique utilizing blending modes to highlight features within the dynamic range of image information. The Multiply and Screen blending modes are used to focus the data on the shadows (dark areas with Multiply) and highlights (light areas with Screen) respectively. Blending modes are used along in conjunction with classic data augmentation techniques to increase classification accuracy and reduce overfitting.

**Acknowledgments**

Thank you to my advisor, Professor Min Chen, for her guidance and friendship.

For Damien & Colette.

**Introduction**

1. CNN

Convolutional Neural Networks have revolution

1. Data Augmentation

Data Augmentation is the technique by which training data is manipulated to produce new examples for the network. Data Augmentation is a, “data-space solution to a limited data set”[1].

*In classic discriminative examples such as cat versus dog, the image recognition*

*software must overcome issues of viewpoint, lighting, occlusion, background, scale, and*

*more. The task of Data Augmentation is to bake these translational invariances into the*

*dataset such that the resulting models will perform well despite these challenges.*

When decided on which Data Augmentation techniques to employ you must decide if the augmentation is “safe” for the training data. An augmentation is considered safe if it transforms the data without changing the corresponding label. An example of a safe augmentation would be horizontally flipping and image of a dog. An unsafe augmentation would entail flipping the digits in the Minsit Dataset, for example flipping a “6” would change it’s corresponding label. The safety of an augmentation is domain dependent and must be chose specially for the dataset it is used on.

Geometic transformations build the base of Data Augmentation techniques. Flipping an image on the horizontal axis is one of the most common augmentation methods. In most real world images, the horizontal flip retains the image’s corresponding label and is easily implemented.

Along with flipping, cropping, rotating, and translating image data also provide geometric augmentation. Cropping an image reduces the size of the input (for example 200 x 200 -> 100 x 100) but allows the network to train on specific areas of the image. The idea is to crop in on the signal in the image while removing the noisy information that surrounds it. Rotating and translating retain the image size and perform a similar feature removal and attention process as cropping.

[1] Survey of Data Aug

1. Description of Problem
2. Thesis Outline

**Background**

1. Model: Densenet
2. Dataset: Tiny Imagenet
3. Augmentation: Blending Modes

**Proposed Approach**

1. Blending Mode Data Augmentation layer
2. Full Data Augmentation Pipeline

**Experiments & Results**

1. CNN Accuracy for different Data Augmentation layers
2. Training time.

**Conclusion & Future Work**

Since the advent of photography practitioners have been searching for process to maximize the detail in their images. The photographic image is inherently a limited representation of our visual reality. Each image sacrifices certain elements to produce a generalized view of the scene.

Photography in its earliest form was a practice in capturing brightness values. It's invention in the 1800's as a Black and White medium was our first semi-permanent (all physical prints fade over time) process to capture our visual existence. Photography democratized the image creation process away from the artist and allows all of us the ability to curate our visual world.

The digitalization of the image created the opportunity to . . The discrete nature limits the dynamic range and compresses the visual relationships.

Overfitting is a major issue with a limited dataset. The best CNN models come from big data. The more images available the better the ability of the model to form a more generalized view of the relationships in the data.

Image issues: Limited size, lighting, exposure, viewpoint, occlusion, background, scale, \dots

My Thesis will focus on lighting and exposure issues.

Maximize the information in the dataset by creating a more generalized representation by training on the full dynamic range of the image.