Exploring the Weber-Fechner Law of Psychophysics in Deep Learning

Kelvin Kuang
CN510

Abstract

The Weber-Fechner law of psychophysics, relating to human perception, states that the just noticeable difference of an increase in a stimulus is proportional to the pre-existent stimulus. Fechner's law states that this intensity of our sensation increases as the logarithm of an increase in stimulus rather than as rapidly as the increase. This study investigates if the same properties hold in convolutional neural networks, which are a type of neural network based off of the human visual system.

1 Introduction

Convolutional neural networks (CNNs) are feed-forward networks based on the visual system deemed the gold standard for computer vision and image recognition tasks. CNNs, through their deep layers of convolutional filters, are proficient at extracting features from images, leading to significant advances in various applications such as object detection, facial recognition, and medical imaging. In psychophysics, Weber's law states that the just noticibale difference of an increase in a stimulus is proportional to the pre-existent stimulus [1]. Fechner's law states that this intensity of our sensation increases as the logarithm of an increase in stimulus rather than as rapidly as the increase [2]. This study investigates if the same properties hold in CNNs.

2 Methodology

2.1 Dataset
This study was conducted using the MNIST dataset, which consists of 80,000 28x28 images of handwritten digits. 70,000 were used for training, and the remaining 10,000 for testing. These images were normalized to [0, 1], and no image augmentations were used.

2.2 Model Training
A 8 layer CNN consisting of 2 convolutional layers and 1 fully connected layer were used for this study (Figure 1). The CNN was trained on the 60,000 images in the training set.

2.3 Evaluating Weber's Law
In order to evaluate Weber's law, the probability vectors were extracted from the softmax activation function for each image in the testing set. To alter the luminance of the images in the testing set, each pixel value was multiplied by a luminance factor (Lf): [1, 10, 100, 1000]. The mean luminance of the testing set was also used to evaluate Weber's law. To better visualize the shifts in the probability vectors and luminance vectors of the testing set, these vectors were fitted to Gaussians, and plotted on both linear and logarithm scales.

3 Results

3.1 Testing Accuracies
The baseline accuracy on the testing set with no altering of luminance was 0.9922. When the luminance factor was increased, the testing accuracy decreased initally by 0.004 for a luminance factor of 10, but there were smaller decreases in accuracy as Lf increased, -0.006 and -0.007 respectively (Figure 2).

3.2 Observing Weber's law
After extracting the probability vectors and luminance vectors for each test set with varying luminance, the vectors were fit to Gaussians. When plotting on the linear scale, the softmax curve has almost no change between the different luminance factors (Figure 2), with the means ranging from 0.9980

to 0.9864. The means of the Gaussians are also close, signifying that there is no major difference in the predictions with change in luminance factor. For the luminance curves, there is much more variance in the Gaussians, with the means ranging from 0.3491 to 0.3996. There is a more noticeable shift, but these shifts are not pure shifts (Figure 3). When we plot these Gaussians on a logarithmic scale, we see a similar pattern. In the softmax curves, there is a small divergence from the other curves when Lf = 100. In the luminance curves, the slope increases as Lf increases, but again, these shifts are not pure shifts as observed in the Weber-Fechner law (Figure 4).

4 Discussion

Since Weber's law can be derived from a shunting network, Weber's law in CNNs would be most affected by the loss function. In this model, categorical cross-entropy was the chosen loss function:

$$CE = -\sum_{i=1}^{i=N} y_{gti} \cdot \log(y_i)$$

where y_{gti} is the ground truth label for i, and y_i is the predicted label for i. This loss function essentially does not include shunting inhibition, which is necessary for Weber's law. Kausik [3] proposes a alteration to the cross-entropy loss function that can take into account Weber's law:

$$-w(\sum_{(x,l)\in T} (F(x)_l + K \log(1-\delta)) \log(E[F'(x)_l)) + (1-w)(-\sum_{i=1}^{i=N} \log(y_i))$$

The former portion of this loss function imposes the Weber-Fechner law, and the latter is a sparse cross-entropy function used for better training of the model. Although Kausik [3] only tested how this loss function affects accuracy, it would be interesting to see if the shift property holds with this new loss function.

5 Conclusion

Overall, this study shows that the Weber-Fechner law is not apparent in CNNs, nor is the pure shifts in intensity. This may be due to the loss function chosen for most CNNs, categorical cross-entropy, as there is nothing similar to shunting inhibition, or terms relating the previous pre-existent stimulus to the change in stimulus. The study by Kausik [3] shows promising results, and investigating if the shift property is apparent with the modified loss function is a promising next step.

References

[1] J. L. Pardo-Vazquez et al., "The mechanistic foundation of Weber's Law," Nature News, https://www.nature.com/articles/s41593-019-0439-7 (accessed Dec. 16, 2023).

[2] R. A. Houstoun and J. F. Shearer, "Fechner's law," Nature News, https://www.nature.com/articles/125891b0 (accessed Dec. 16, 2023).

[3] B. N. Kausik, "Psychophysical Machine Learning," arXiv.org, https://arxiv.org/abs/2208.11236 (accessed Dec. 16, 2023).

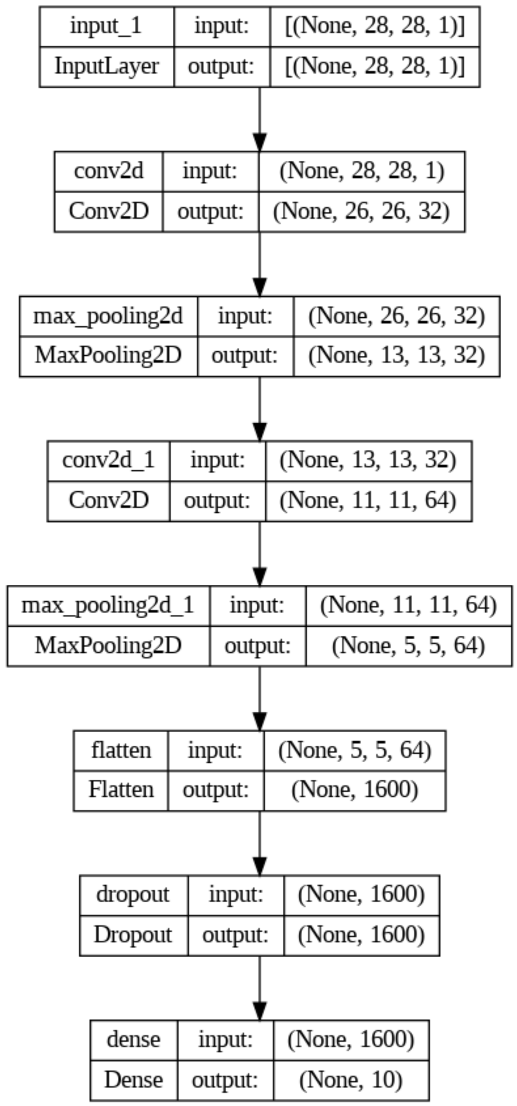

Figure 1: 8 Layer CNN. This model consists of 2 convolutional, each being fed to max pooling layers. Dropout with p=0.5 was implemented before the fully connected layer.

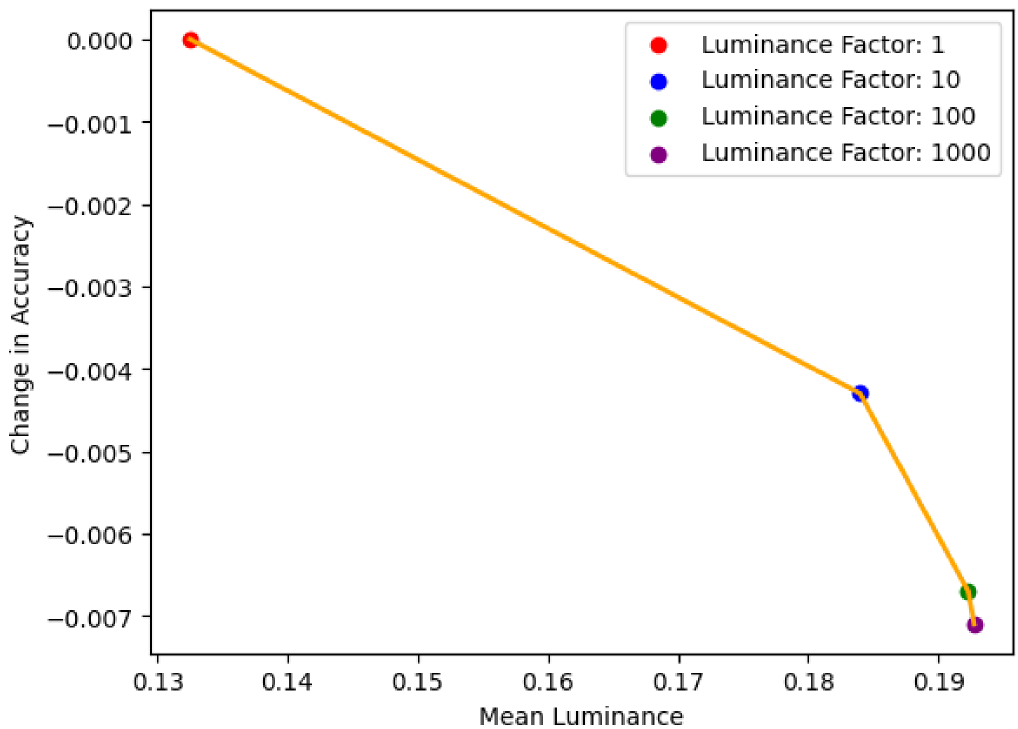

Figure 2: Change in accuracy vs mean luminance of testing set images. Sharp decrease in accuracy when Lf = 10, but has smaller decreases as Lf begins to reach its maximum value.

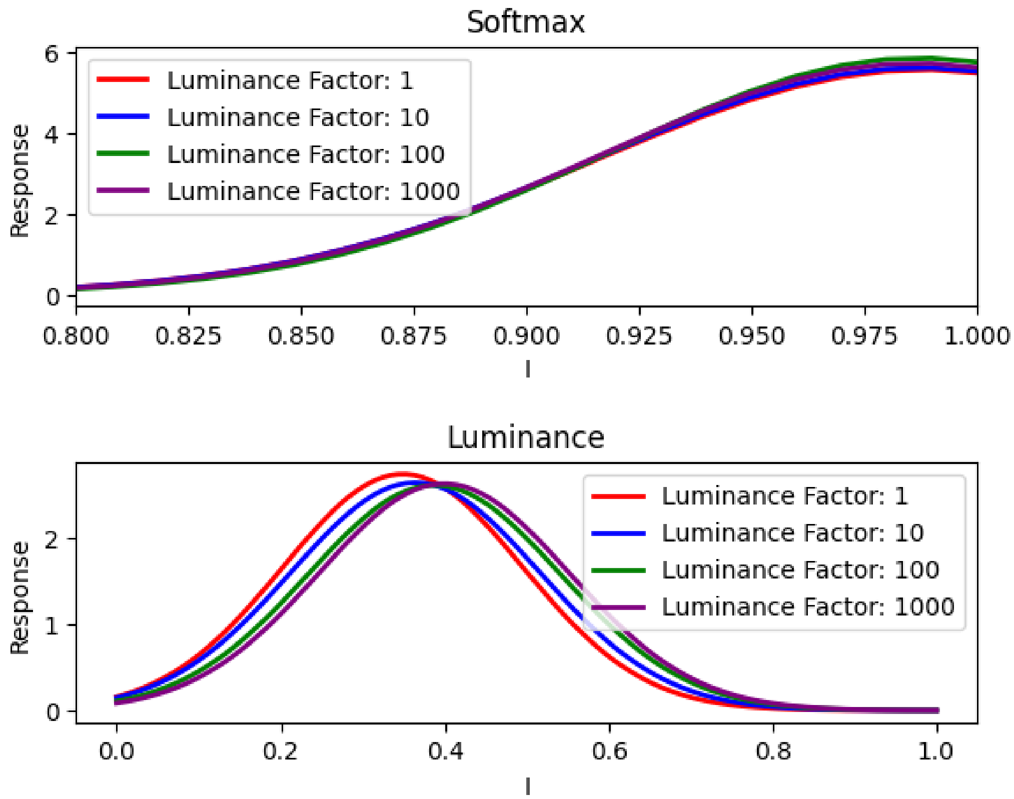

Figure 3: Probability and luminance vectors fitted to Gaussians on a linear scale. There is little to no shift for the probability vectors, but a more apprent shift is visible for the luminance vectors. The means of the Gaussians for the probability vectors are: [0.9880175, 0.9872369, 0.987095, 0.98644274], while the means for the luminance vectors are [0.07168832, 0.07101539, 0.06808409, 0.06963587]. The luminance curves show a larger change when the Lf is increased.

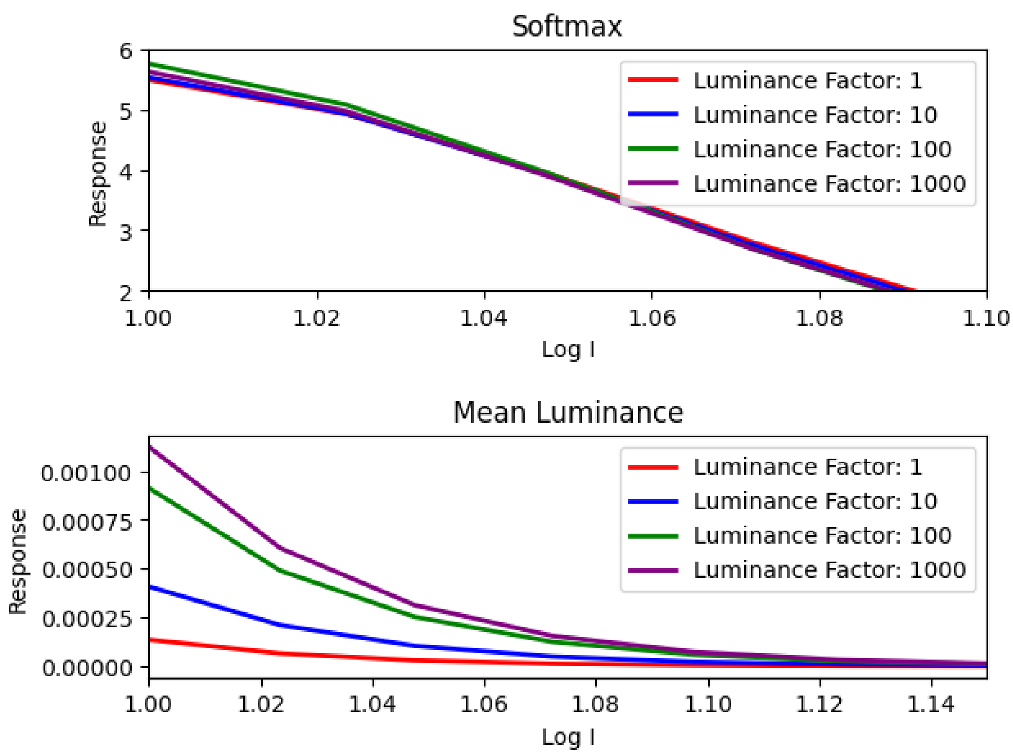

Figure 4: Probability and luminance vectors fitted to Gaussians on a logarithmic scale. Again, there is little change in the probability curves, but as the Lf increases, there is a greater corresponding shift in the luminance curves. These are not pure shifts as described in the Weber-Fechner law.