NORMALIZATION ON VISUAL DATA FOR CATEGORIZATION

UNIVERSITY OF MIAMI
DEPARTMENT of COMPUTER SCIENCE

G. Hubbard[1,2], O. Schwartz[3], L. Sanchez Giraldo[3]

[1]REU: Computing for Structure, University of Miami, Coral Gables, FL, USA; [2]Old Dominion University, Norfolk, VA,USA;
[3] Department of Computer Science, University of Miami, Coral Gables, FL, USA

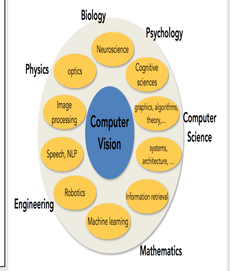
Introduction

Computer vision (CV) is the scientific discipline concerning techniques for automated extraction of information from visual digital media. Modern audits of internet transmission show that over 80% of all data processed is some form of visual data. This sheer magnitude has spearheaded both the academic and industrial interest in computer vision systems. CV technology is rapidly advancing as more capable hardware and techniques inspired by biological vision systems have been incorporated.

Goal:

- **Understand the effects of normalization on image data**
- **Classify the data using a Convolutional Neural Network (CNN)**
- **Compare accuracy of Categorization of normalized vs unaltered data**

Global internet traffic from videos will make up 82% of all consumer internet traffic by 2021
-Cisco

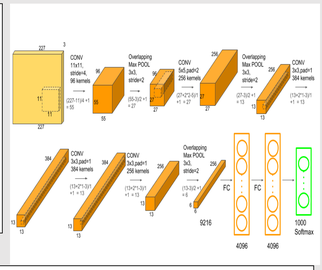


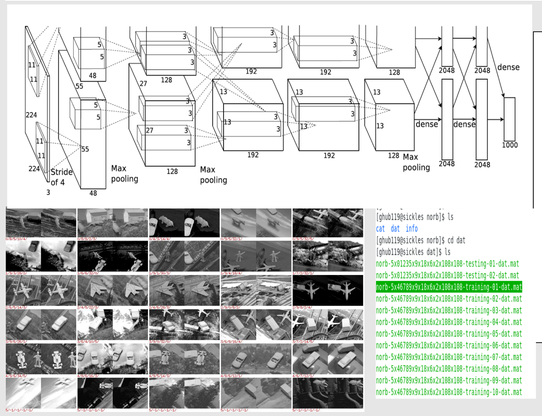
METHODS: An effective CNN not only depends on the architecture design, but careful consideration of input data format. In neural cortical processing a computation known as normalization has been deemed as a canonical computation in the brain as there is rich evidence of its occurrence at different stages and sensory modalities. CNNs have included simple forms of normalization.

Tools:

Convolutional Neural Networks

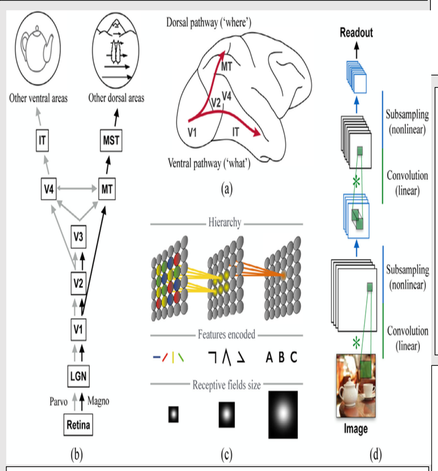
Artificial computer vision/image analysis technique that uses neural network inspired by the connectivity of neurons in the human visual cortex.
The CNN used for this experiment is AlexNet. It is composed of 5 convolutional layers followed by 3 fully connected layers, as depicted in below.





NORB
The NORB (NYU Object Recognition Benchmark) image Data set contains 194,400 images of toys under the 5 categories of : four-legged animals, human figures, airplanes, trucks, and cars. The NORB data set also contains Right/Left 'eye' image pairs that attempt to emulate data for visual field processing

History



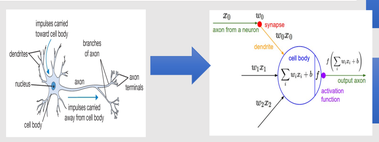
"The eye is like a mirror, and the visible object is like the thing reflected in the mirror."
-- Avicenna, early 11th century

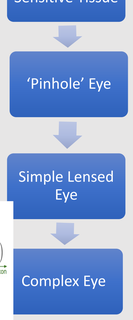
Human Interests:

Since antiquity, early philosophers and scientists studied the eye. Early Greek philosophers and scientists, such as Plato, Euclid, and Ptolemy believed in extramission (theory that the eye actively produced rays). Kepler's studies in optics allowed him to refute the etramission theory claiming that eyes resemble organic optical systems.

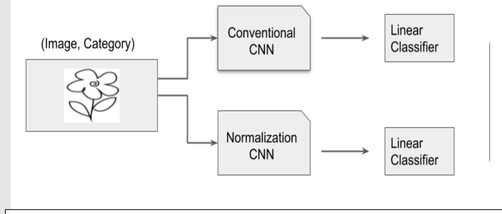
PAST TEN YEARS:
Since the 1960s, the theory of computer vision has been progressing with few major breakthroughs. The use of CNNs for image classification has provided major progress and renewed interest in computer vision

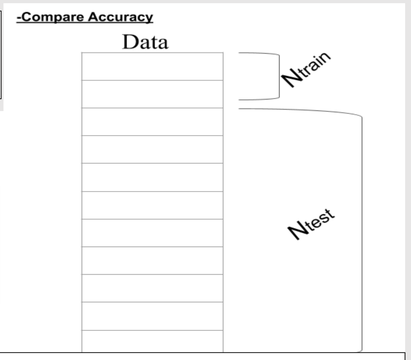
IN BIOLOGY:
The first biological eyes evolved around 540 million years ago. Vision's ancestry predates the appearance of identifiable eyes.





Method:





NORB dataset consists of labeled data whose categorical prediction has already been defined. To initially train the network, ten data sets are iterated in batches. After the training set is computed through the network, adjustments are made to the weights of the network. The NORB dataset contains two test sets to categorize (new) untrained data. The performance on the test set is very important in the assessment of the network's accuracy.

Results:

```
epoch 0 accuracy train 0.787071, test 0.785768:
epoch 1 accuracy train 0.783711, test 0.782956:
epoch 2 accuracy train 0.738237, test 0.709002:
epoch 3 accuracy train 0.520027, test 0.521742:
epoch 4 accuracy train 0.787620, test 0.784294:
epoch 5 accuracy train 0.732819, test 0.730195:
epoch 6 accuracy train 0.791564, test 0.789780:
epoch 7 accuracy train 0.792627, test 0.785923:
epoch 8 accuracy train 0.752092, test 0.746571:
epoch 9 accuracy train 0.759911, test 0.739129:
epoch 10 accuracy train 0.693999, test 0.684757:
epoch 11 accuracy train 0.763340, test 0.748251:
epoch 12 accuracy train 0.786831, test 0.779801:
epoch 13 accuracy train 0.698765, test 0.679733:
epoch 14 accuracy train 0.767318, test 0.750686:
```

Data	Accuracy
Training	79%
Testing	78%
Normalized Training	81%
Normalized Testing	78%

Given that the testing data and the training data are relatively close in accuracy, the model offers good performance on unfamiliar (toy) visual data. Note that the normalized data offered better accuracy on categorizing visual data

Conclusion:

This report focused on the practice of normalization incorporated to the intermediate stages of CNNs. In particular, we focused on the problems of object recognition in digital visual processing. We expected that normalization of visual data would help reduce "clutter" in raw image data. The results of our experiment show that normalization of the NORB data set offer increased accuracy in the metric of categorizing data. This work can help future studies in understanding the role of normalization at intermediate stages of the visual cortex as well as building better artificial systems.

Acknowledgements
This material is based upon work supported by the National Science Foundation under Grant No. CNS-1659144.

References
https://web.stanford.edu/class/history13/earlysciencelab/body/eyespages/eye.html

https://www.cs.toronto.edu/~lczhang/360/

https://www.researchgate.net/figure/The-classical-view-of-hierarchical-feed-forward-processing-a-The-two-visual-pathways_fig1_301756727

https://cs.nyu.edu/~ylclab/data/norb-v1.0/