*Convolutional Neural Network (CNN) for Image Classification*

Yibo Cheng  
*Department of Computer Science (The Jonsson School)*  
*The University of Texas at Dallas*Richardson, TX  
yxc190039@utdallas.edu  
Mo Han  
*Department of Computer Science (The Jonsson School)*  
*The University of Texas at Dallas*Richardson, TX  
mxh190001@utdallas.edu  
Minchao Zhu  
*Department of Computer Science (The Jonsson School)*  
*The University of Texas at Dallas*Richardson, TX  
mxz190002@utdallas.edu

*Abstract*—This electronic document is a research study on Convolutional Neural Network. This network technique will be used to solve image classification problem. Results will be presented in this document and conclusion will be made based upon experiment results.

Keywords—CNN, Image Classification, Tuning (key words)

# Introduction & background

Science and technology are developing in the fast pace. Especially equipped with advanced computing hardware, algorithm and theoretical foundation, technology starts to lead in people’s daily life and transforms the way it is. One of the common applications is using convolutional neural network to classify images or objects based upon their features.

The team’s research topic is to implement a classical deep learning algorithm, convolutional neutral network (CNN), to realize image classification and further increase accuracy by tuning parameters in the neutral network. Since 1943 when Warren McCulloch, a neurophysiologist, and a young mathematician, Walter Pitts, wrote a paper on how neutron might work, artificial neutral network has been studied. There are thousands of types of specific neural networks proposed by researchers as modifications or tweaks to existing model and even more are still emerging. These models can be categorized into three classes in general, Multilayer Perceptron (MLP), Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN). Each class has its own design purpose and focus. CNN’s ability to develop an internal representation of a two-dimensional image makes it the preferred network for image classification.

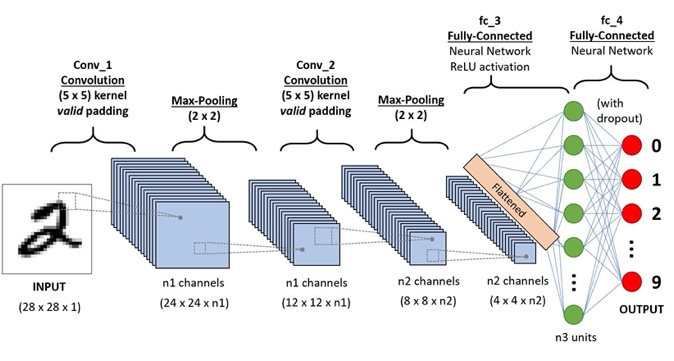
The convolutional neural network (CNN) was first proposed in 1960s. Hubel and Wiesel proposed the concept of “receptive field” which observed that neurons were sensitive to moving edge on visual cortex cells of cats. Later in the 1980s, based on concept of “receptive field”, Fukushima and Miyake proposed “neocognitron” which is regarded as the first implementations of CNNs. However, due to lack of proper learning algorithm, CNN was not the main focus in the network. After that, researchers started to use multilayer perceptron to learn features and incorporated backpropagation (BP) algorithm. However, since traditional BP neutral network would have series of problems requiring detailed study, the research on deep neutral network model was stopped. Until Hinton et al found that the artificial neural network with multiple hidden layers addresses those old issues and has great performance in feature learning, deep learning starts to re-gain attention and more and more sophisticated and accurate models were developed and used in daily practice, especially in the fields of OCR, autonomous drive, image recognition and analysis, social media, etc.

In the research, the team will build up a convolutional neutral network from the scratch. Stages include preprocessing images, convolutional layer, activation layer, pooling layer, fully connected layer and forward and backward propagation. Meanwhile, effects brought by parameters, like convolutional matrix, weights, will be studied and compared in terms of network efficiency, accuracy and generality.

# Technique & algorithm

## Convolutional Neutral Network (CNN)

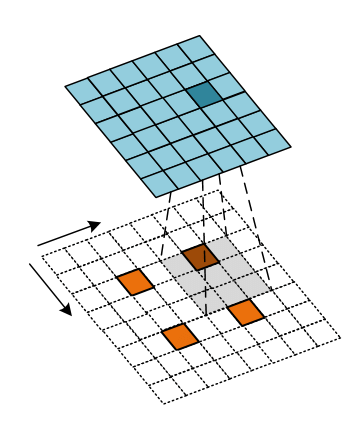
Compared to other image classification algorithms, CNNs do not require extra work on preprocessing images and this means that they can learn the filters that have to be hand-made in other algorithms. Also, CNNs have advantage in dimensional reduction without losing learning features. This could tremendously save computing time. layers including convolution, pooling, activation and fully connection constitute the network and the sequence of each operation could be shuffled according to the need.

[](https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53)

## Convolutional Layer

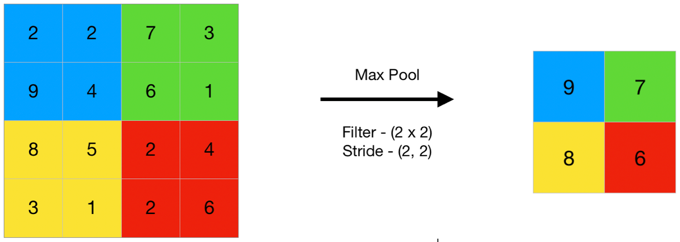
Identify applicable funding agency here. If none, delete this text box.

First layer of CNN normally starts with a convolutional layer. A convolutional layer contains a set of filters whose parameters including height, weight, number of filters are determined through inputs. In general, the height and weight of the filters are smaller than those of the input volume (in research, input is a 2D matrix image with 1 channel, gray image). Each filter is convolved with the input volume to compute an activation map. In specific, the filter is slid across the width and height of input with appointed stride and the dot products between the input matrix and filter are computed at every spatial position. After number of filter’s iterations, the output of convolutional layer is obtained by stacking the activation maps of all filters along the depth dimension. For example, in the research, given the situation that input volume is a uniform a\*a matrix, convolutional filter is chosen to be b\*b matrix and each stride is assigned as c, the final output’s shape after convolutional layer will be (a + c – b) \* (a + c – b).

[](https://www.researchgate.net/figure/Convolution-process-of-transposed-convolution-layer_fig4_340020703)

## Pooling Layer

The purpose of pooling layer in the network is to progressively reduce the spatial size of representation to reduce the number of parameters and computation in the network. Normally, pooling layer will exist in-between convolutional layers and it takes a series input parameter including input volume, number of filters, height, width, stride and padding. Two main techniques in pooling are popular and widely used, average pooling and maximum pooling. In the research, maximum pooling is implemented for down sampling input volume. Maxpooling ensures a lower resolution version of an input signal is created and still contains large portion of key features, without fine detail that might not be as useful to the task.

[](https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/)

## Activation Layer

The function of activation layer in the network is to transform linear output from either convolutional layer or pooling layer to non-linear output so that the network is able to learn complex patterns in the data. Many activation functions are available and in the research, rectified linear unit (ReLU) is implemented, which simply computes the function: f(x) = max (0, x).

[](https://medium.com/@kanchansarkar/relu-not-a-differentiable-function-why-used-in-gradient-based-optimization-7fef3a4cecec)

## Fully Connected Layer

The objective of fully connected layer in the network is used to classify the image into a target label. At this layer, it will receive output from either convolutional layer or pooling layer. The input will then be further flattened into a single vector of values, each representing a probability that a certain feature belongs to a label.

[Diagram

Description automatically generated](https://www.superdatascience.com/blogs/convolutional-neural-networks-cnn-step-4-full-connection)

# Result and Analysis

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections A-D below for more information on proofreading, spelling and grammar.

# Conclusion

Our results show that a complete and comprehensive convolutional neutral network (CNN) is able to be reproduced without using any official libraries in either TensorFlow or Python machine learning. The team follows few steps to accomplish the image classification tasks. First, the team preprocessed image datasets including 4000 images on cat and dog each. Preprocessing includes converting RGB images into gray images and utilizing Gaussian filtered edge detection technique prior to the CNN. Then, the team implemented convolutional layer, max-pooling layer, activation layer and fully connected layer algorithms along with forward and backward propagation. Experiments on tuning parameters in each layer and sequencing layers were conducted to explore potential opportunity in improving result accuracy. Even though, the best image classification accuracy the team obtained is around 30%, it is still acceptable as there are few issues the team does not intend to tackle down at this time. Open issues include overfitting and underfitting, activation function, image resolution loss, etc. Potential improvement on classification accuracy could be realized if these problems were addressed and will be left for the future tasks.

##### Acknowledgment *(Heading 5)*

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

##### References

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

1. G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*
2. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
3. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
4. K. Elissa, “Title of paper if known,” unpublished.
5. R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
6. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
7. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**