# Handwritten Mathematical Expression Recognition from Whiteboards

## Introduction:

Handwritten mathematical expression recognition is a challenging problem due to the complicated two-dimensional structures, ambiguous handwriting input and variant scales of handwritten math symbols. In the first phase of our project, we will implement an attention based encoder-decoder model that recognizes mathematical expression images from two-dimensional layouts to one-dimensional LaTeX strings, as proposed in this [*paper*](https://arxiv.org/abs/1801.03530).

## Description of Components:

The paper proposes a novel attention based encoder-decoder model.

Compared with conventional approaches for handwritten mathematical expression recognition, the attention based encoder-decoder model possesses three distinctive properties:

1) It is end-to-end trainable

2) It is data-driven, in contrast to traditional systems that require a predefined math grammar

3) Symbol segmentation can be automatically performed through attention model

**Encoder:** It is a convolutional neural network (CNN) based on VGG architecture that maps images to high-level features. The encoder is based on DenseNet. The main idea of DenseNet is to use the concatenation of the output feature maps of preceding layers as the input of succeeding layers.

**Decoder:** The decoder is a recurrent neural network *(RNN)* with gated recurrent units *(GRU)* that converts these high-level features into output strings one symbol at a time. For each predicted symbol, an *attention model* built in the decoder scans the entire input expression image and chooses the most relevant region to describe a math symbol or an implicit spatial operator.

**Multi-Scale Attention with Dense Encoder:**

The idea behind using multi-scale attention in the encoder is to ensure that we do not miss miniscule features in the input due to pooling operations of CNN. This approach includes two branches in the encoder, one for capturing high resolution annotations and one for low resolution annotations each.

## Future Scope:

We aim to improve the performance (accuracy) of the above discussed model. Also, we would like our solution to cater to varying degrees of complicated mathematical expressions.

Moreover, the above mentioned solution focuses on offline expression recognition, our aim would be to work with online mode as well.