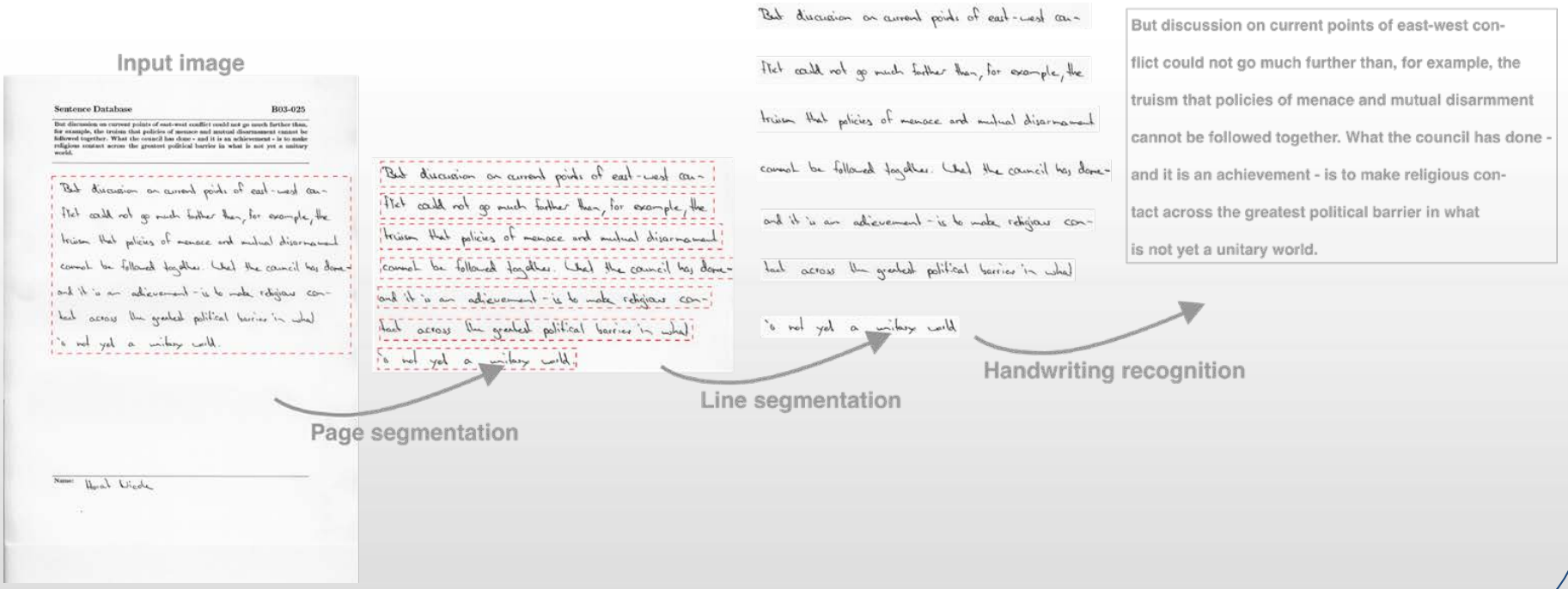
Handwriting Recognition

Problem Statement: Devising a deep learning approach detect handwritten text

Pipeline:

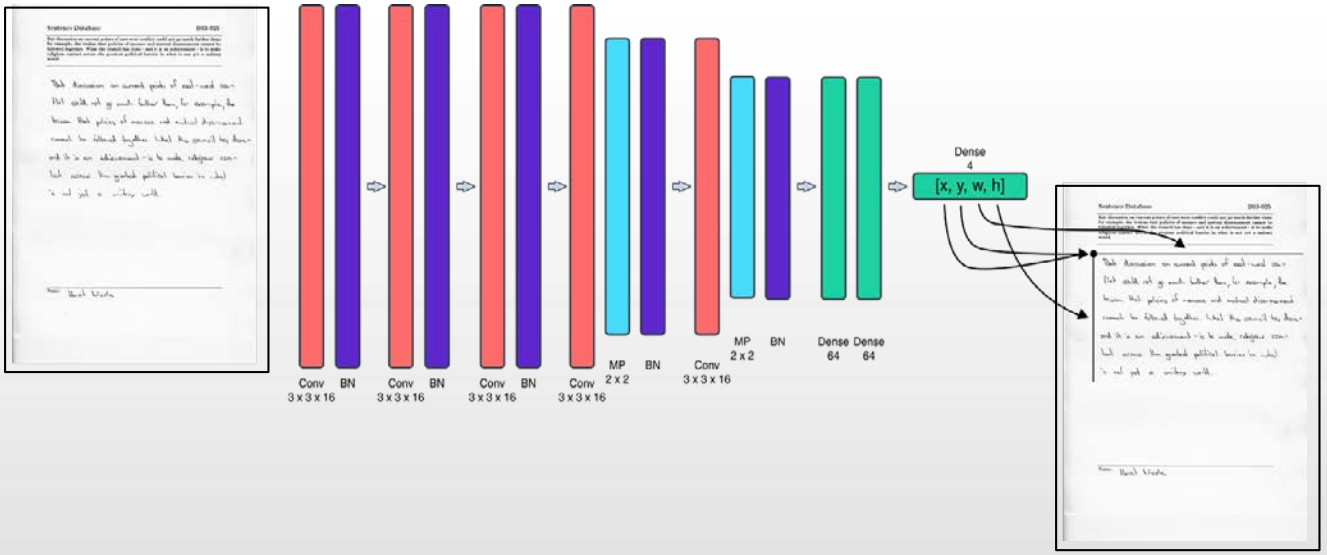


AlgoAnalytics

Network Architecture

Pipeline is composed of three steps:

1. Page Segmentation



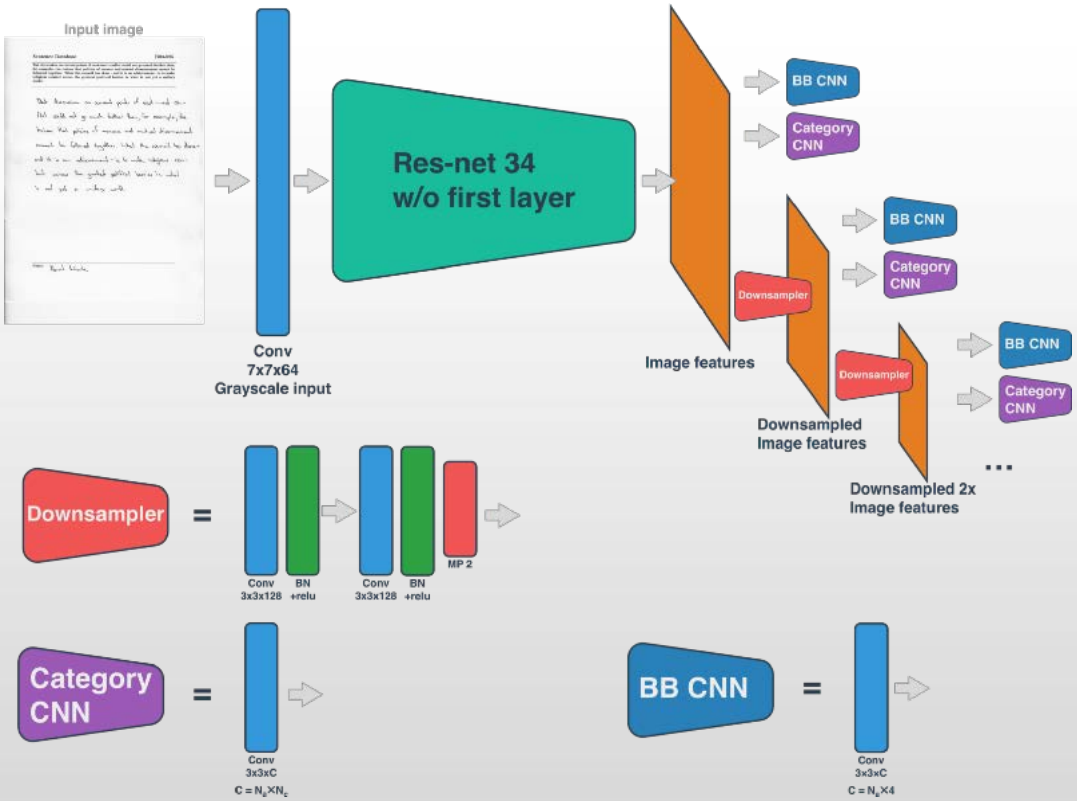
The deep CNN using Apache MXNet takes the IAM document as an input and predicts the bounding box of the handwritten passage.

AlgoAnalytics

Network Architecture (contd)

2. Line segmentation: Handwritten text is segmented line by line so that each line can be used for handwriting recognition.

- Single Shot multibox Detector (SSD) architecture is used to detect the positions of each line of the passage.
- The SSD architecture takes image features and repeatedly downsamples the features.
- At each downsample step, the features are fed into two CNNs: one to estimate the locations of bounding boxes relative to anchor points (BB CNN), and one to estimate the probability of the bounding box encompassing an object (Category CNN).
- Data augmentation and Non maximum suppression methods are used to obtain meaningful results

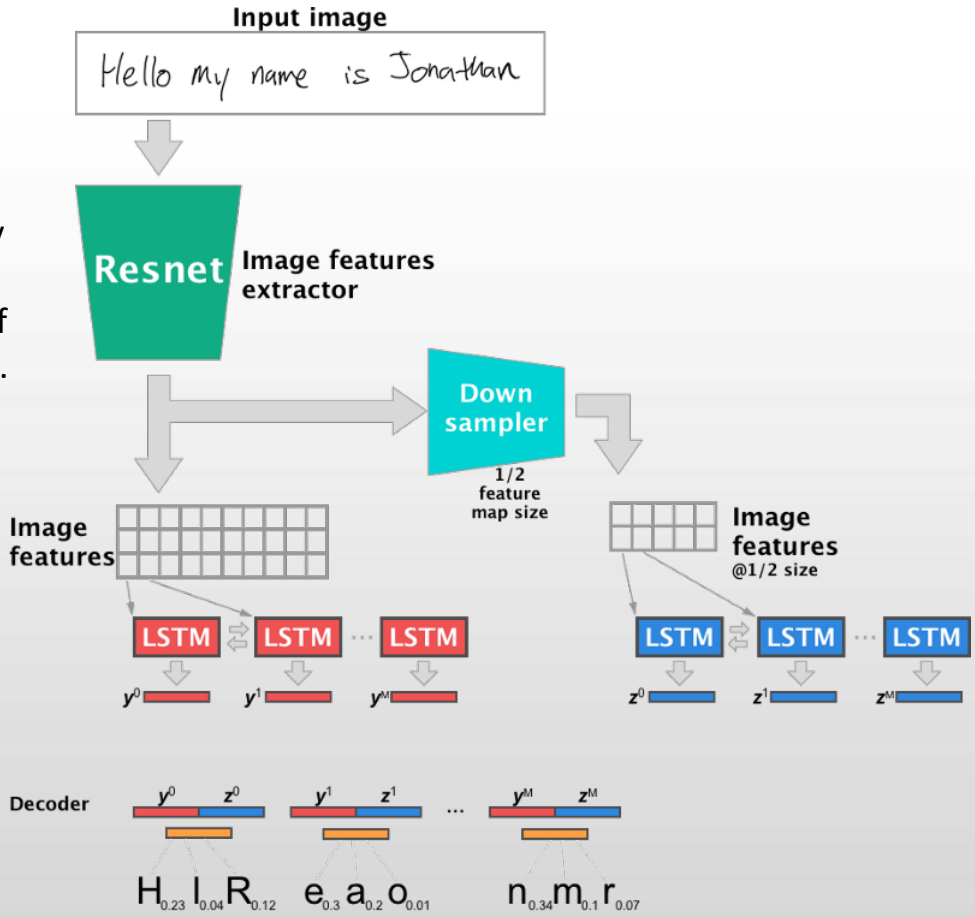


AlgoAnalytics

Network Architecture (contd)

3. Handwriting recognition and language modeling

- The CNN generates image features that are spatially aligned to the input image.
- Image features are then sliced along the direction of the text and sequentially fed into a biLSTM network.
- The output of the biLSTM network is fed into a decoder to predict probability distribution over the characters for each vertical slice of the image
- The network is trained to optimize Connectionist Temporal Classification (CTC) loss
- Beam search algorithm is used to extract the most probably sentence from the matrix



AlgoAnalytics