From Pixels to Sequences: Unraveling the Power of CNN-LSTM Combination

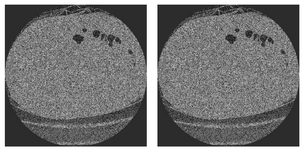
Mingyang Liu 530131573, Jingbo Zhang 480402996, Yuqing Yang 530194981

COMP5329 Deep Learning

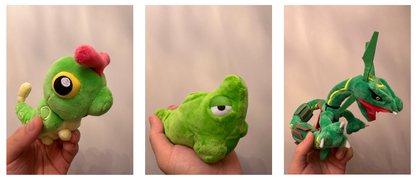
Research Problem

The objective of this study is to propose a method for sequential image analysis using a hybrid CNN-LSTM model for automated pneumonia detection. By connecting the two networks together, we aim to leverage the strengths of both CNN and LSTM in order to improve the accuracy of pneumonia detection.

Research Methodology

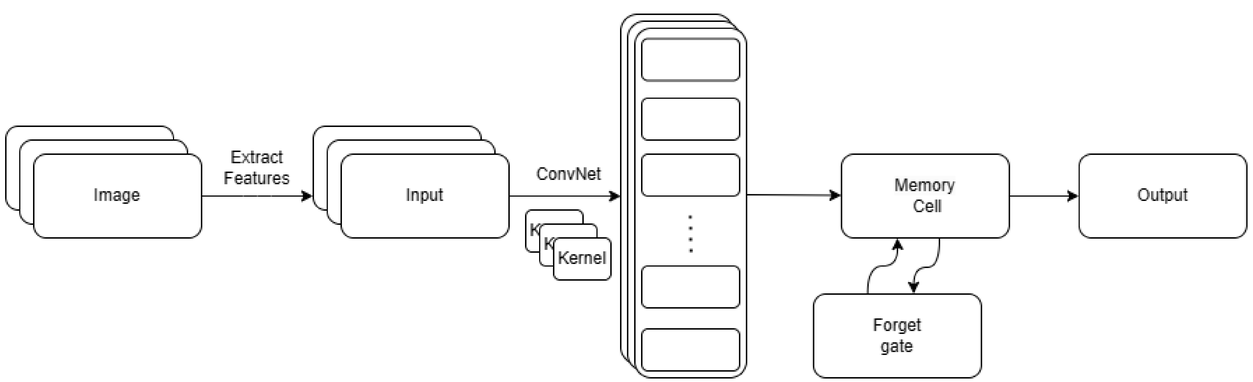


Ambiguous Sequential Image data



Obvious Sequential Image data

Our proposed model aims to reduce spectral variance by introducing input features and incorporating a time background in CNN. The output of the CNN layer is then fed into LSTM to mitigate temporal variations. This sequential image analysis model facilitates easier classification and disease detection



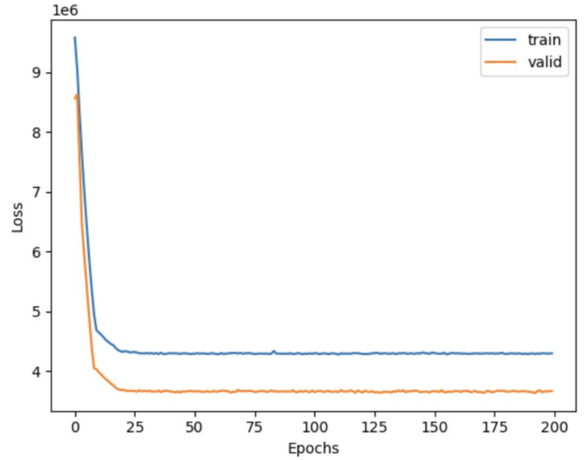
CNN-LSTM model Process

Experimental Results

· The introduction of LSTM layers in the CNN-LSTM architecture brings additional variables, making the SGD optimizer more convenient due to its better handling of these variables.

· The CNN-LSTM architecture allows for the extraction of spatial features from image sequences through CNN layers. The LSTM layer effectively models the temporal information, capturing correlations within the images. This combined approach is particularly advantageous for processing large volumes of continuous data, making it suitable for various disease prediction tasks in the medical field.

· In terms of future directions, there is an opportunity to enhance the interpretability and visualization capabilities of CNN-LSTM models in medical image analysis. Techniques such as attention mechanisms or saliency mapping can improve interpretability by identifying influential regions or features in the images. Additionally, visualizing and explaining the decision-making process of CNN-LSTM networks would provide valuable insights for medical professionals.



- SGD optimizer
- Learning rate scheduler
- Weight decay
- L2 normalization
- ReLU activation
- 3 hidden layer

Discussion

Our model still has a lot of room for improvement, as the dataset used is too small to be very convincing for the evaluation of the model. For future development EfficientNet can be called upon to increase the processing power of the model in order to enhance its generalisation and migration capabilities, and the model has a high accuracy rate and can be applied to scenarios of medical images.