BI-DIRECTIONAL CONVLSTM U-NET WITH DENSELY CONNECTED CONVOLUTIONS

Jahnvi Rohela 20D070040

Mohit Jindal 20D070052

CONTENTS

WHAT IS IMAGE SEGMENTATION?

- Technique to partition a given image into image segments
- Helps robustly analyze the image and extract meaningful data
- **Medical image segmentation** is used for diagnostics and aims to provide precise interpretation of medical images
- Helps in reduction of time, cost, and error of humanbased processing

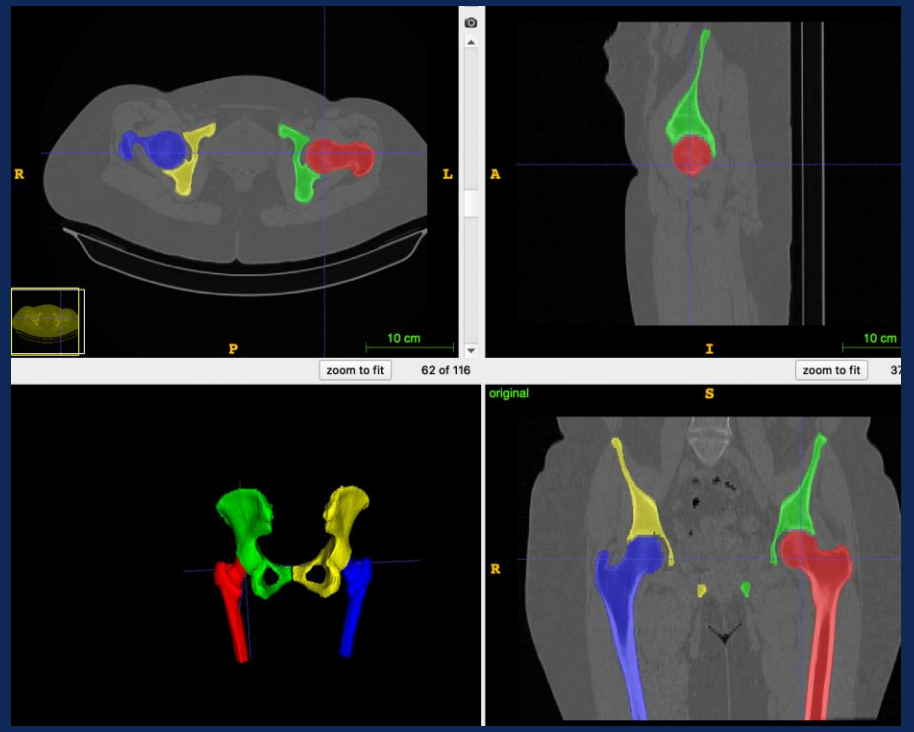


Fig. 1 Image segmentation of a hip joint

TRADITIONAL APPROACHES FOR IMAGE SEGMENTATION

FCN (FULLY CONVOLUTIONAL NEURAL NETWORK)

- One of the first deep networks applied to image segmentation
- Consist of a series of convolutional layers with pooling, followed by upsampling layers that increase the spatial resolution of the output
- In the final layer, a SoftMax function is applied to classify each pixel into a specific class
- Well-suited for tasks where the input size may vary

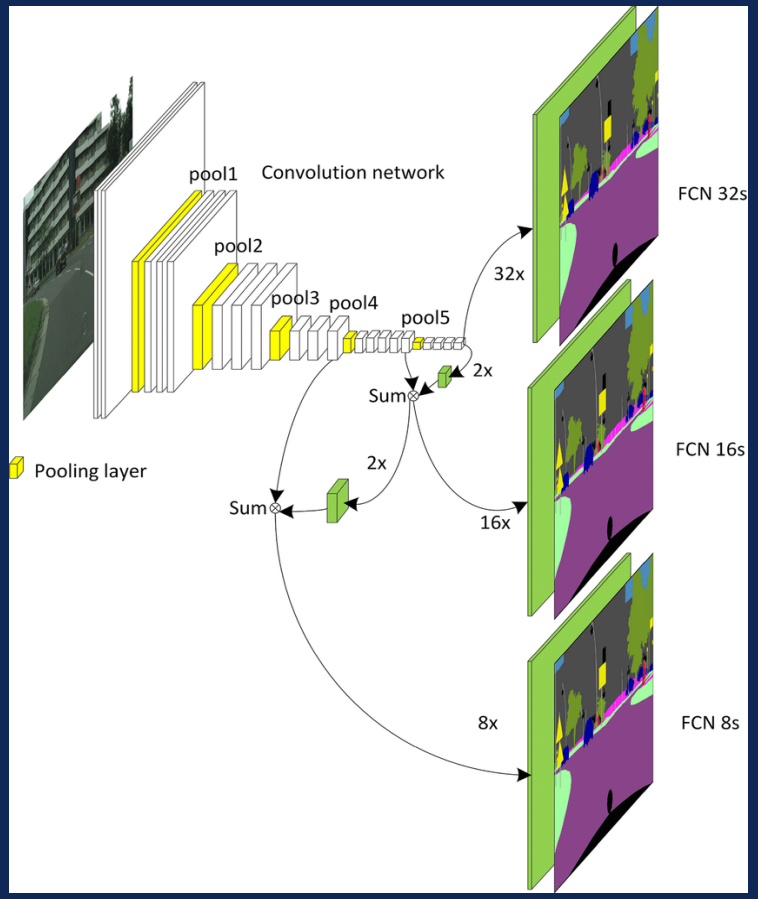


Fig. 2 FCN architecture

U-NET

- Extension of FCN
- Consists of encoding and decoding paths
- Encoding process involves extracting numerous feature maps with smaller dimensions from the input data
- Decoding process creates segmentation maps (which have the same size as the input) using up-convolutions

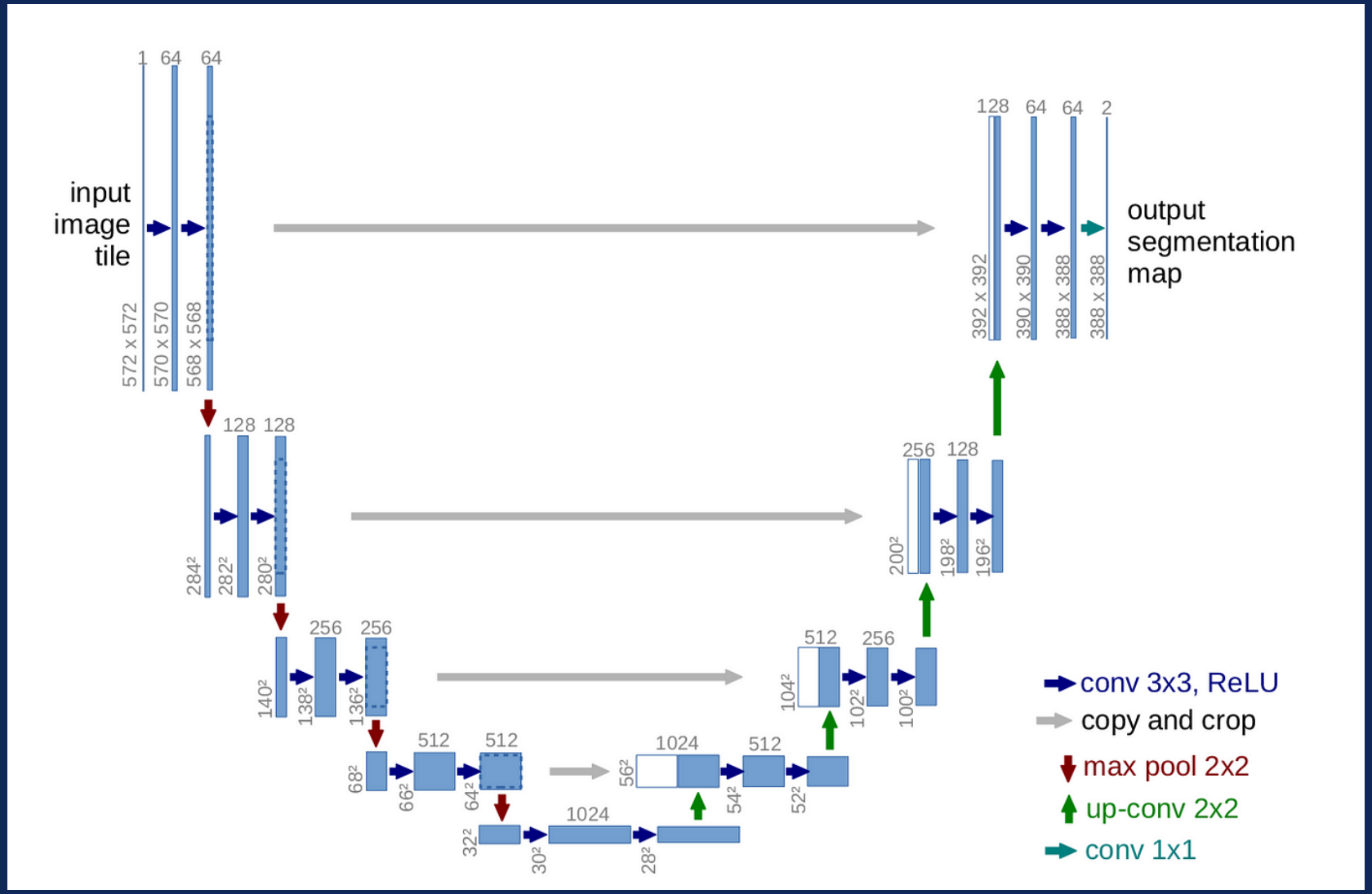


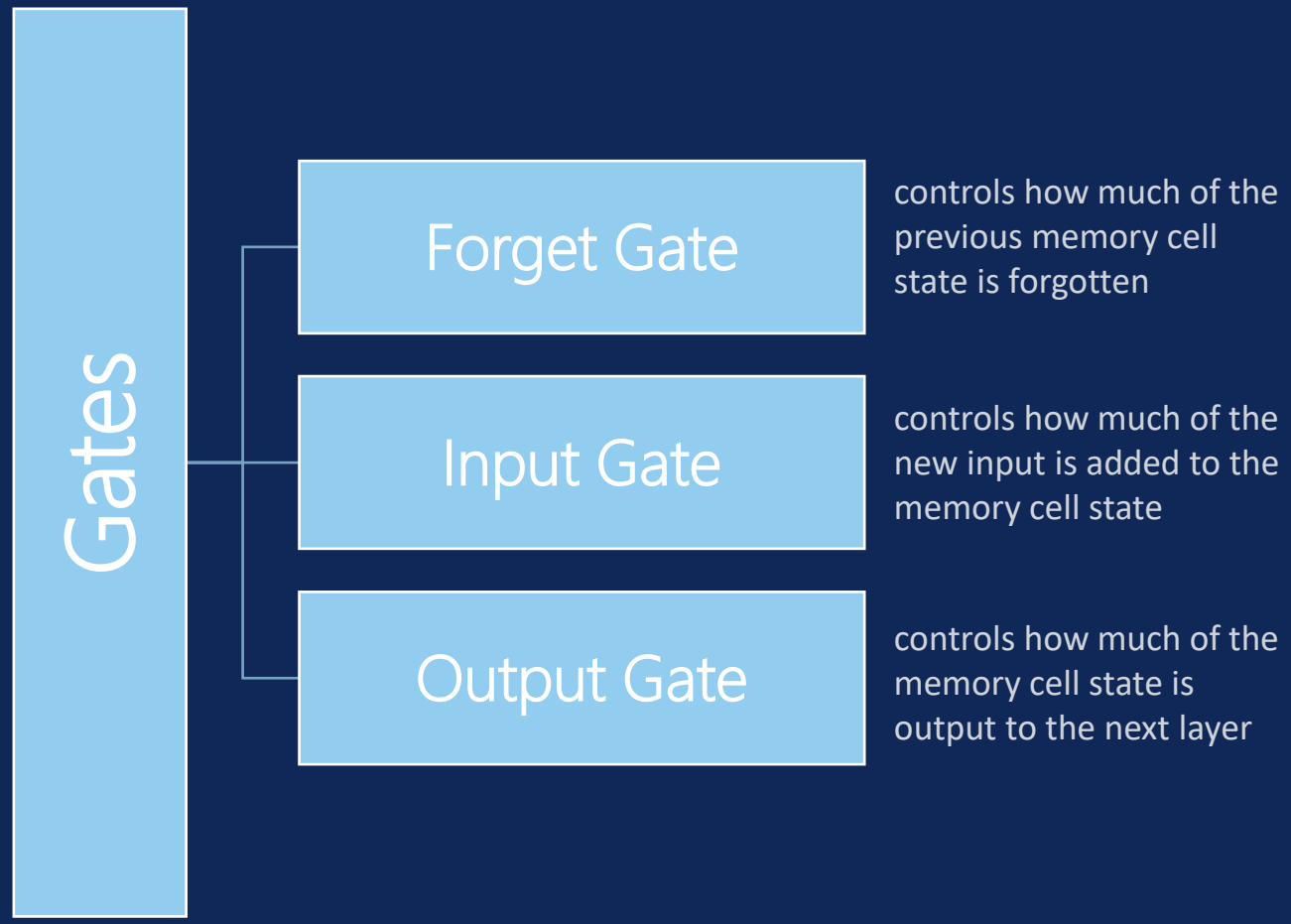
Fig. 3 UNet architecture

MODIFIED U-NET USING BI-DIRECTIONAL CONV**LSTM** AND DENSENET

BCDU-Net with Densely Connected Convolutions

LSTM(LONG SHORT-TERM MEMORY)

- A type of recurrent neural network (RNN) architecture, specifically designed to handle long-term dependencies between input sequences
- Use a memory cell to store information over time and a series of gates to control the flow of information in and out of the cell
- The ROI in medical images is usually distributed over multiple adjacent slices resulting in correlations in successive slices, LSTMs are able to extract inter-slice contexts from the input slices as a form of sequential data



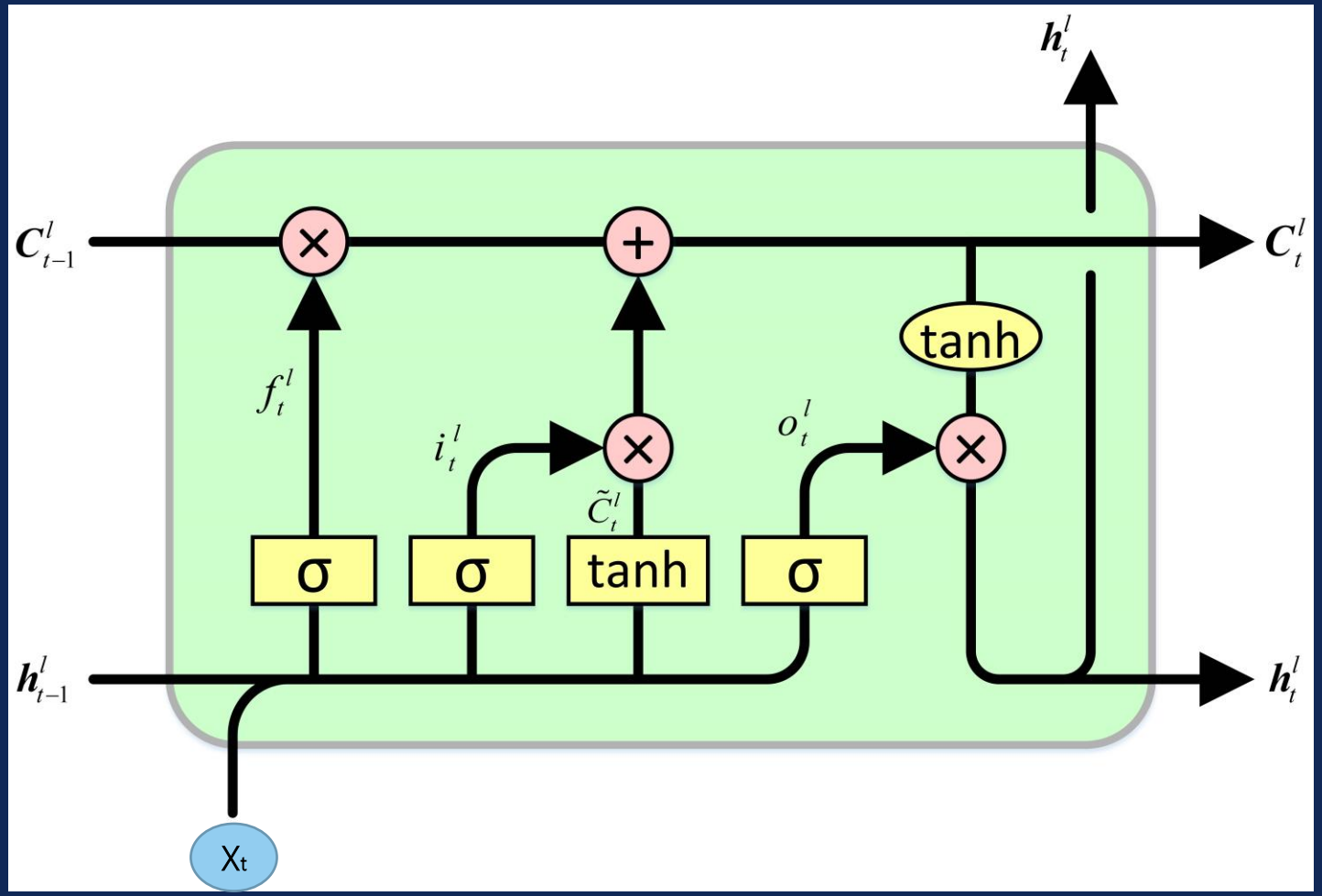


Fig. 4 LSTM Architecture

BI-CONVOLUTIONAL LSTM

- Convolutional LSTM is a type of recurrent neural network architecture that combines the spatial processing capabilities of convolutional networks with the temporal processing abilities of LSTMs
- BiConvLSTM is a model that uses two ConvLSTMs to analyze input data in both forward and backward directions.
- Helps to make more informed decisions by taking into account data dependencies from both directions

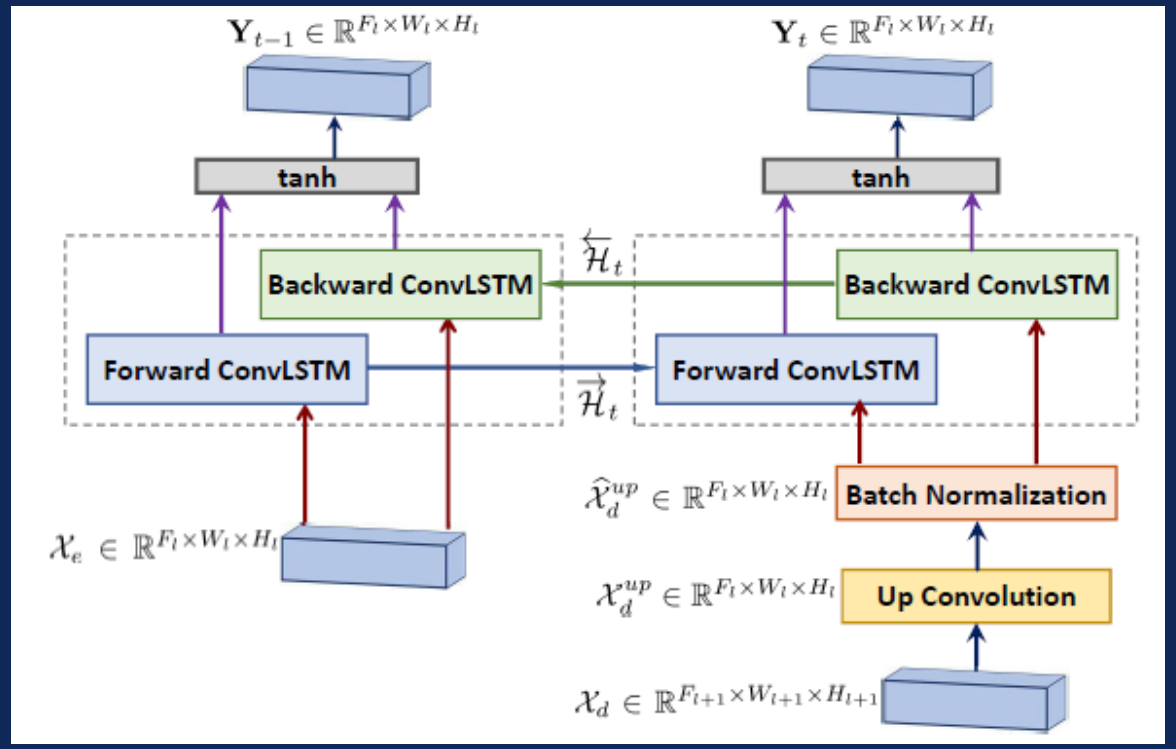


Fig. 5 Bidirectional-Convolutional LSTM Architecture

DENSELY CONNECTED CONVOLUTIONS (DENSENET)

- To mitigate the problem of the network learning redundant features, Densely Connected Convolutions are used
- Each layer is connected to every other layer in a feed-forward fashion
- Enable feature reuse, which can lead to more efficient and effective training
- Helps to combat the vanishing gradient problem

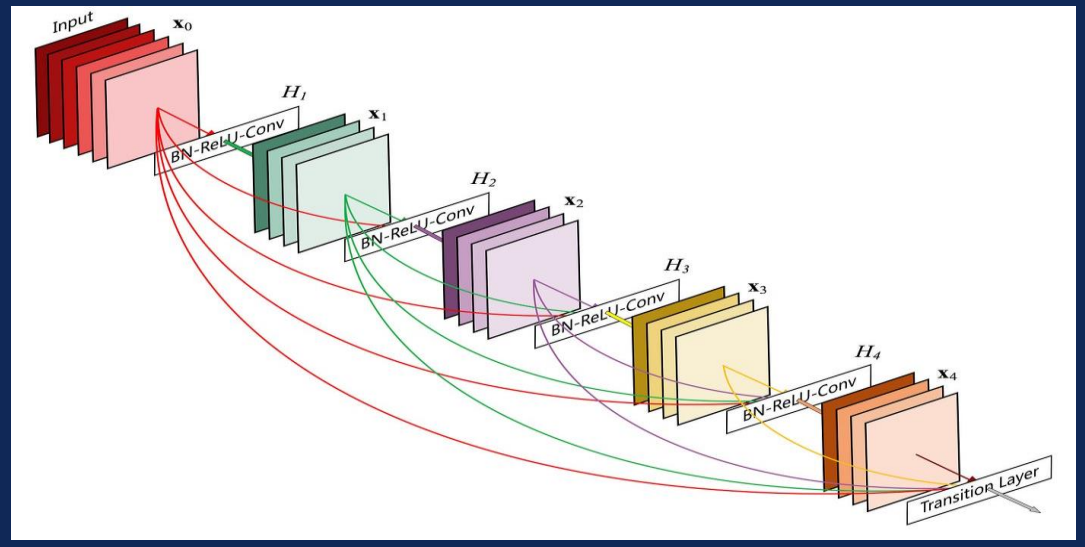


Fig. 6 DenseNet Architecture

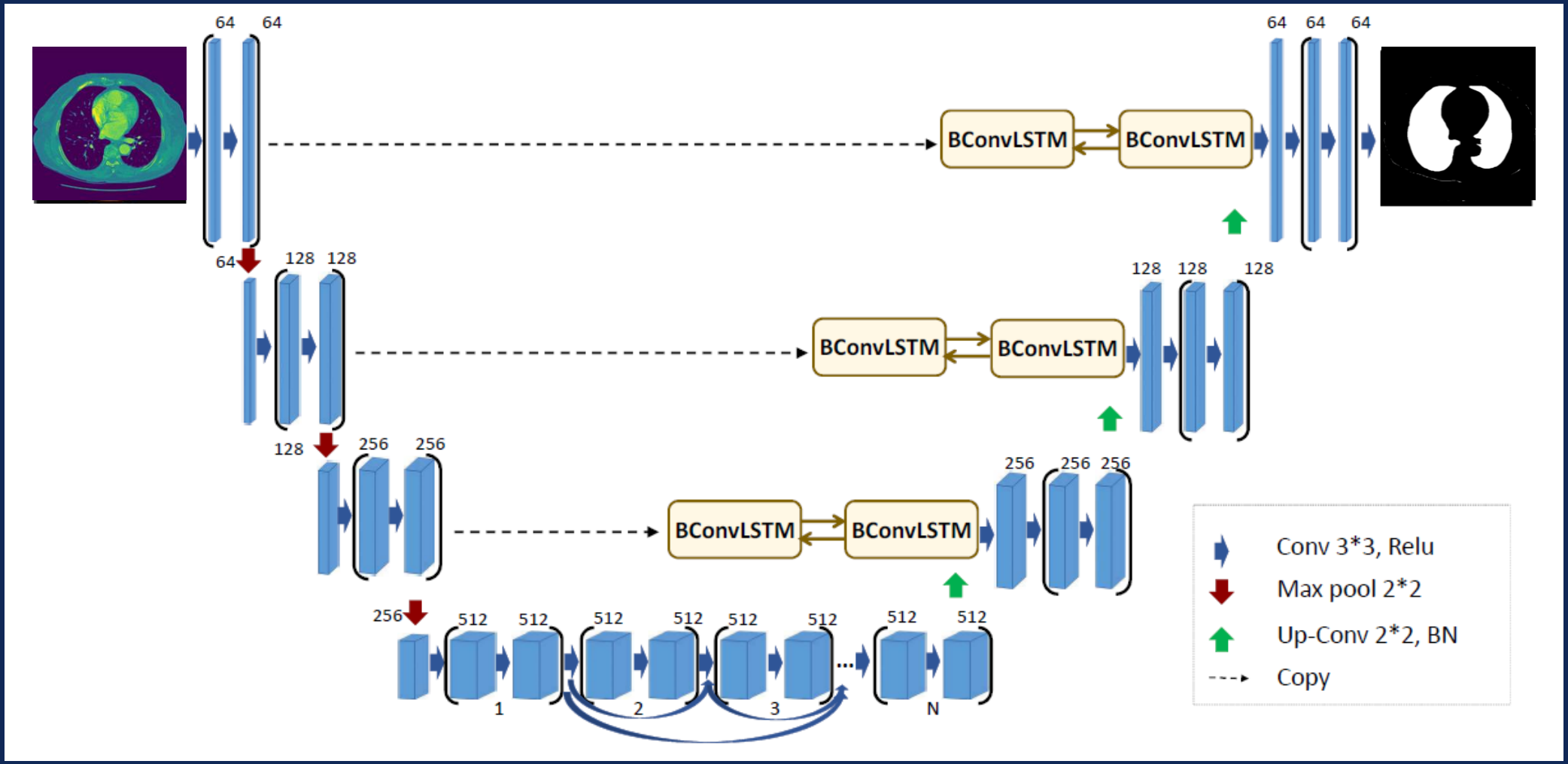


Fig. 7 BCDU-Net with bi-directional ConvLSTM in the skip connections and densely connected convolution

TIME AND ACCURACY ANALYSIS

	F1-Score	Accuracy	Recall	Precision	Time/Epoch
UNet	0.985086	0.996167	0.986691	0.983487	119s
BiConvLSTM (d=1)	0.986074	0.996420	0.988028	0.984128	369s
BiConvLSTM (d=3)	0.98713	0.9966	0.98982	0.98446	380s

Fig. 8 Performance comparison on the LUNA16 dataset (https://luna16.grand-challenge.org/)

CONCLUSION

- BiConvLSTMs are able to leverage the correlation between adjacent slices to perform better segmentation results but in turn increase the network complexity
- Dense layers strengthen feature propagation and also reduce the number of trainable parameters to the model
- Thus, a trade-off is made between improving segmentation accuracy and decreasing the computation time

THANK YOU