Recurrent R2U-Net for Video Segmentation

McKell Woodland (Rice University), Doug Poland (LLNL), Carmen Carrano (LLNL)

DATA SCIENCE SUMMER INSTITUTE

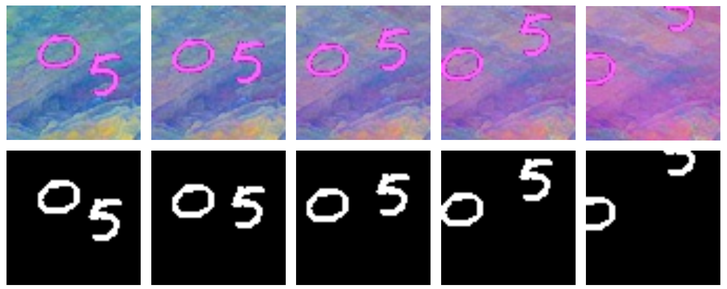
Lawrence Livermore National Laboratory

UNIVERSITY OF CALIFORNIA National Laboratories

We expand the R2U-Net to be recurrent across time by combining it with a convolutional LSTM. We use this new model to segment digits in the Moving Colorful MNIST dataset and hope to expand the use of the model to event detection applications.

DATASET

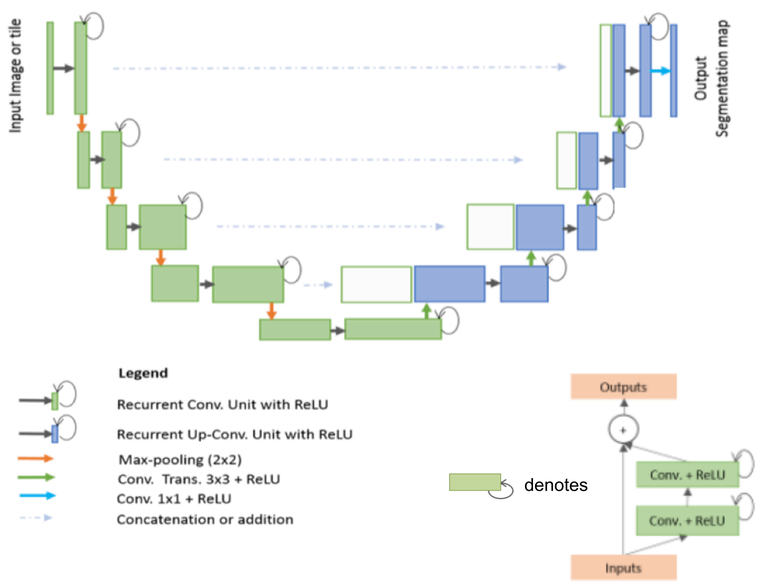
We generate our own dataset, called the Moving Colorful MNIST dataset. It consists of videos of two colored digits moving across a colored square. We also include the ground truth segmentation for every video.



Frames of a video in the Moving Colorful MNIST dataset.

R2U-NET ARCHITECTURE

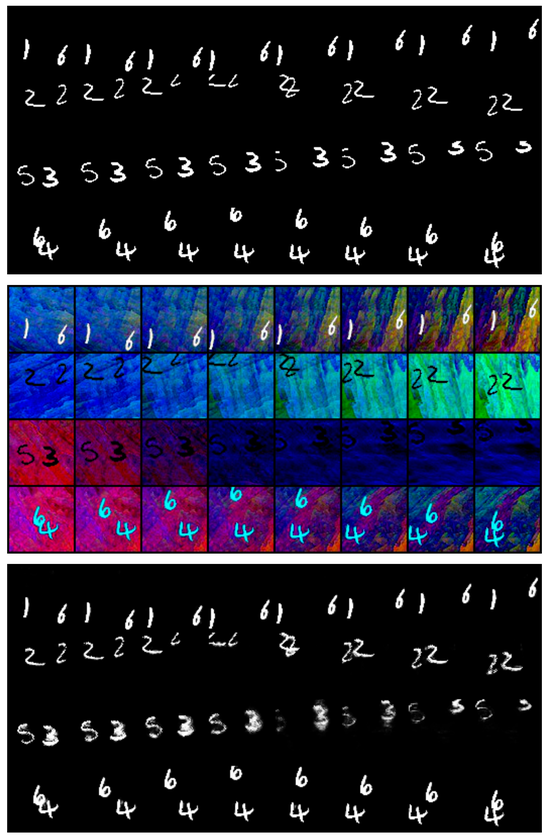
A R2U-Net is a residual and recurrent U-net created by Alom et al.[1]



The architecture of a R2U-Net.

R2U-NET RESULTS

We compare the R2U-Net to a vanilla U-Net with various parameters. Although the U-Net performs better, we justify using a R2U-Net as we would like to make our dataset more complex.



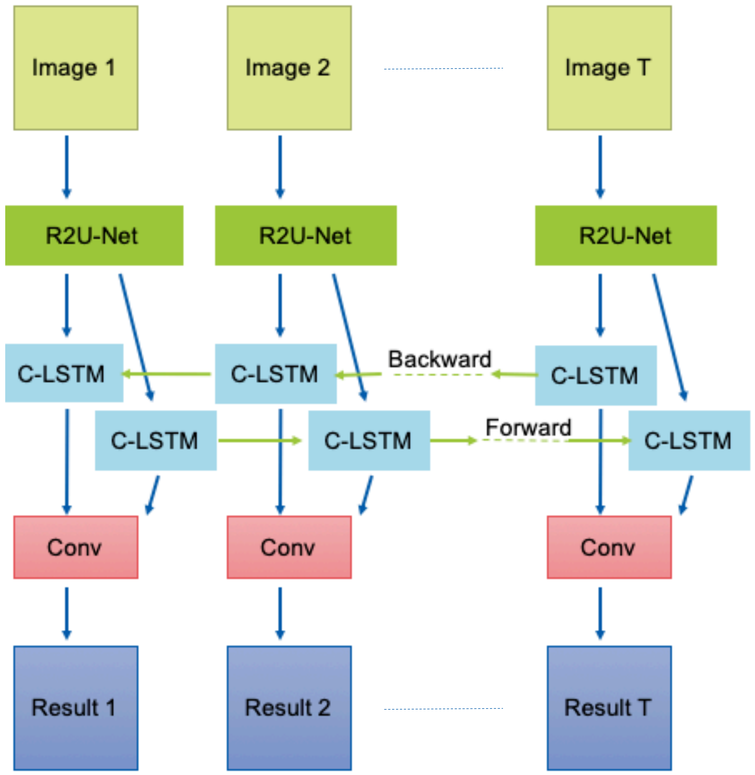
Visual results of our R2U-Net Small implementation on randomly selected still shots from the Moving Colorful MNIST test dataset.

	R2U-Net		U-Net	
	39M params.	609k params.	34M params.	537k params.
Accuracy	99.73%	99.46%	99.91%	99.87%
Sensitivity	97.97%	93.68%	99.19%	99.61%
Specificity	99.72%	99.76%	99.95%	99.94%
Precision	99.66%	95.46%	99.09%	98.84%
F1-Score	95.34%	94.51%	99.14%	98.72%
Jaccard Similarity	94.81%	89.75%	98.30%	97.49%
Dice Coefficient	97.29%	94.51%	99.14%	98.72%

Average of five empirical results of our U-Net implementations on randomly selected still shots from the Moving Colorful MNIST test dataset.

CONVOLUTIONAL LSTM

We hypothesize that motion aids in object recognition. As such, we attempt to improve upon the R2U-Net results by putting them through a bidirectional convolutional LSTM.[2]



The architecture of a bidirectional convolutional LSTM.

FUTURE WORK

We will bring the outputs of the R2U-Net into the Convolutional LSTM and test it on our dataset. We would also like to make our dataset harder for the model to learn when it is solely given still images (e.g. add in Gaussian noise, obscurations, etc.). We do this to test our hypothesis that motion aids in object recognition.

ACKNOWLEDGEMENTS

Thank you to Juanita Ordonez for providing the Moving Colorful MNIST data generator.

1. Alom, M.Z., Hasan, M., Yakopcic, C., Taha, T.M., & Taha, V.K. Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation. *CoRR* abs/1802.06955.

2. Shi, X., Chen, Z., Wang, H., Yeung, D.Y., Wong, W.K., & Woo, W.C. Convolutional LSTM Network: A Machine Learning approach for Precipitation Nowcasting. *CoRR* abs/1506.04214.

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.