

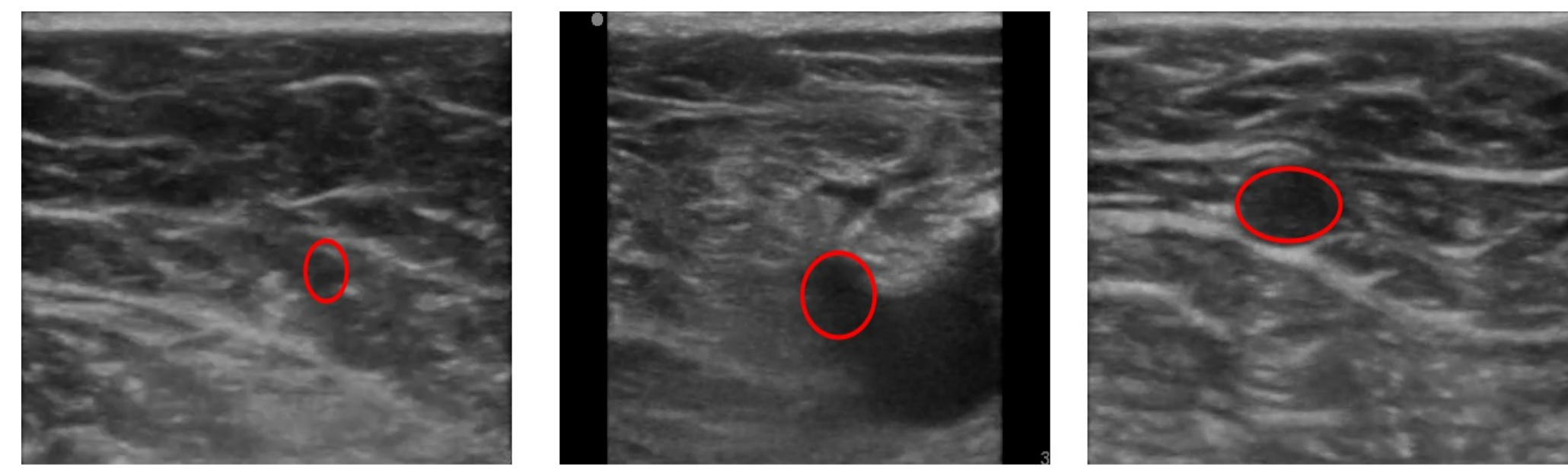
# LSTM-U-net for the robust segmentation of veins in ultrasound sequences

Daniel Mensing<sup>1,2</sup>, Johannes Gregori<sup>1</sup>, Jürgen Walter Jenne<sup>1,2,3</sup>, Michael Stritt<sup>1</sup>, Björn Gerold<sup>4</sup>, and Matthias Günther<sup>1,2,5</sup>

<sup>1</sup>mediri GmbH, 69115 Heidelberg, Germany, <sup>2</sup>Fraunhofer MEVIS, 28359 Bremen, Germany, <sup>3</sup>German Cancer Research Center, 69120 Heidelberg, Germany, <sup>4</sup>Theraclion S. A., 92240 Malakoff, France, <sup>5</sup>Universität Bremen, 28359 Bremen, Germany

## Introduction

- High Focused Ultrasound (HIFU) treatment for varicose veins is a non-invasive technique to ablate insufficient veins
- The therapy is monitored by diagnostic Ultrasound B-mode imaging
- The course of the targeted vein must be followed during the procedure in order to ablate the entire vessel



## Methods

We propose a LSTM-U-net architecture which is able to track and segment the vein in real-time in diagnostic ultrasound image series. We conducted multiple experiments:

### Experiment 1: Prediction of future frames

- Trained the model on ultrasound time series while replacing the last three frames with zeros
- Skipped frames randomly to account for rapid changes of the veins appearance

### Experiment 2: Tracking of structures outside of the image edges on whole and on cropped image series

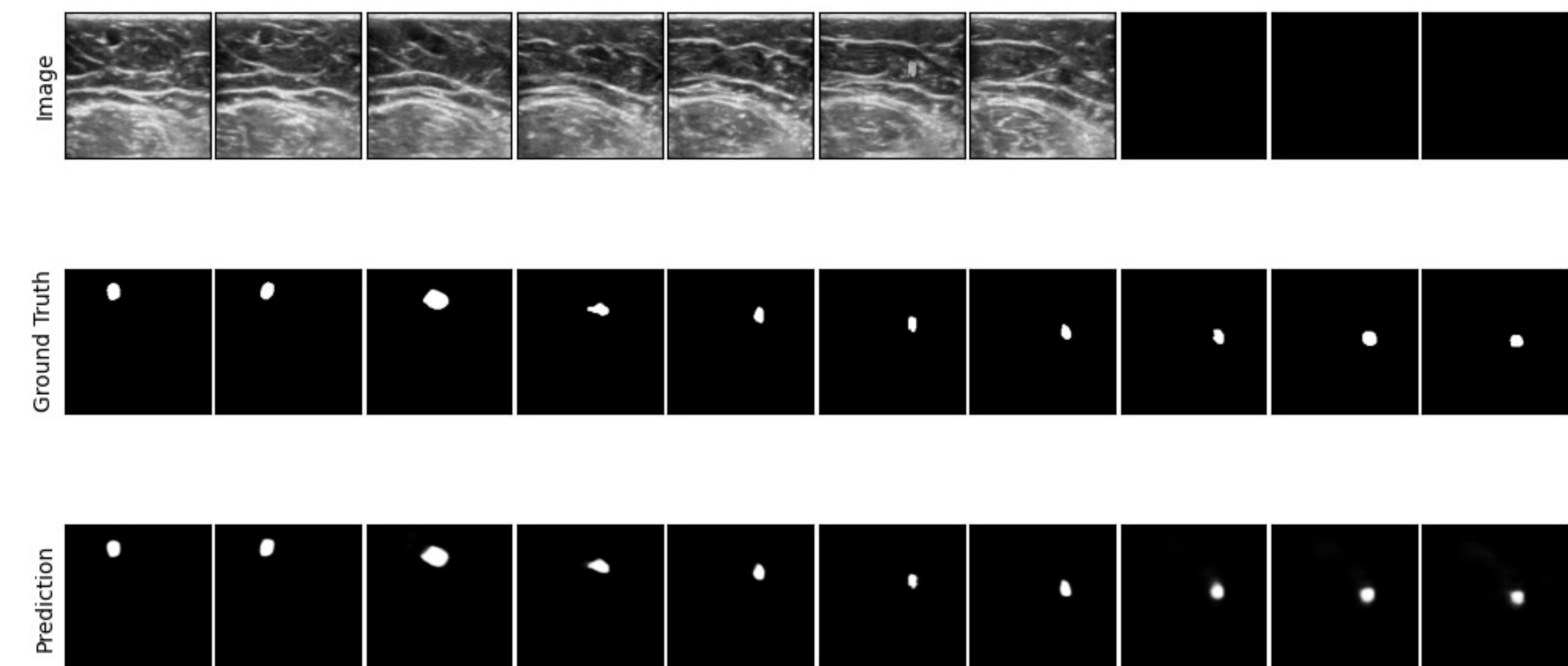
- Trained the model on images series in which the edges are replaced by zeros (See **Figure 2.**)
- Trained the model on patches of images series in which the edges are replaced by zeros

### Experiment 3: Evaluation of usage of ConvLSTM layers in the U-net architecture

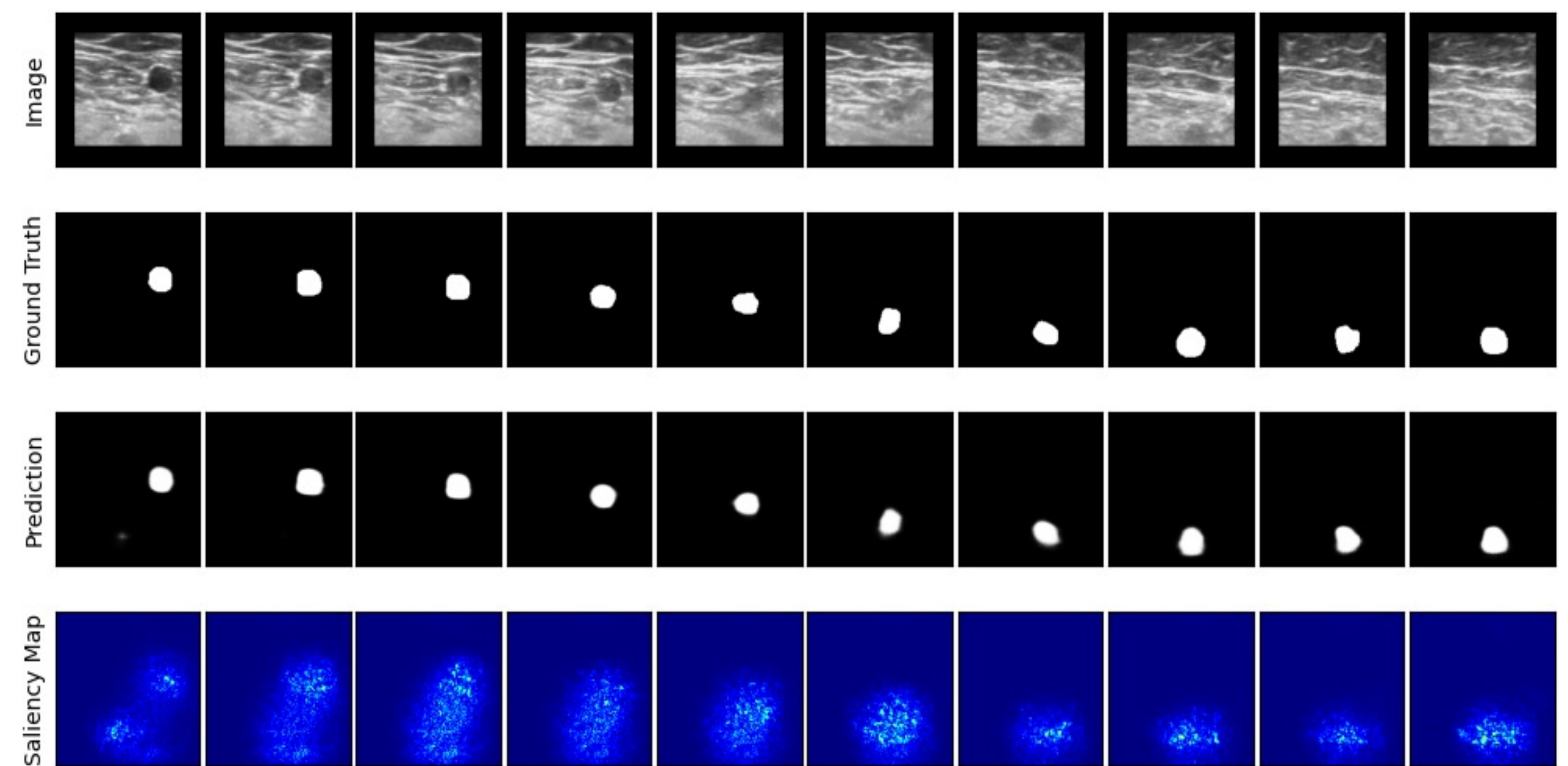
- Implemented two architectures derived from the LSTM-U-net
- The „Encoder“ architecture in which the ConvLSTM layers in the decoder part were replaced by regular 2D convolution layers
- The „Decoder“ architecture, which employs ConvLSTM layers in the decoder part and regular 2D convolution layers in the encoder
- Evaluated all three architectures with the same training setup (See **Table 1.**)

## Results

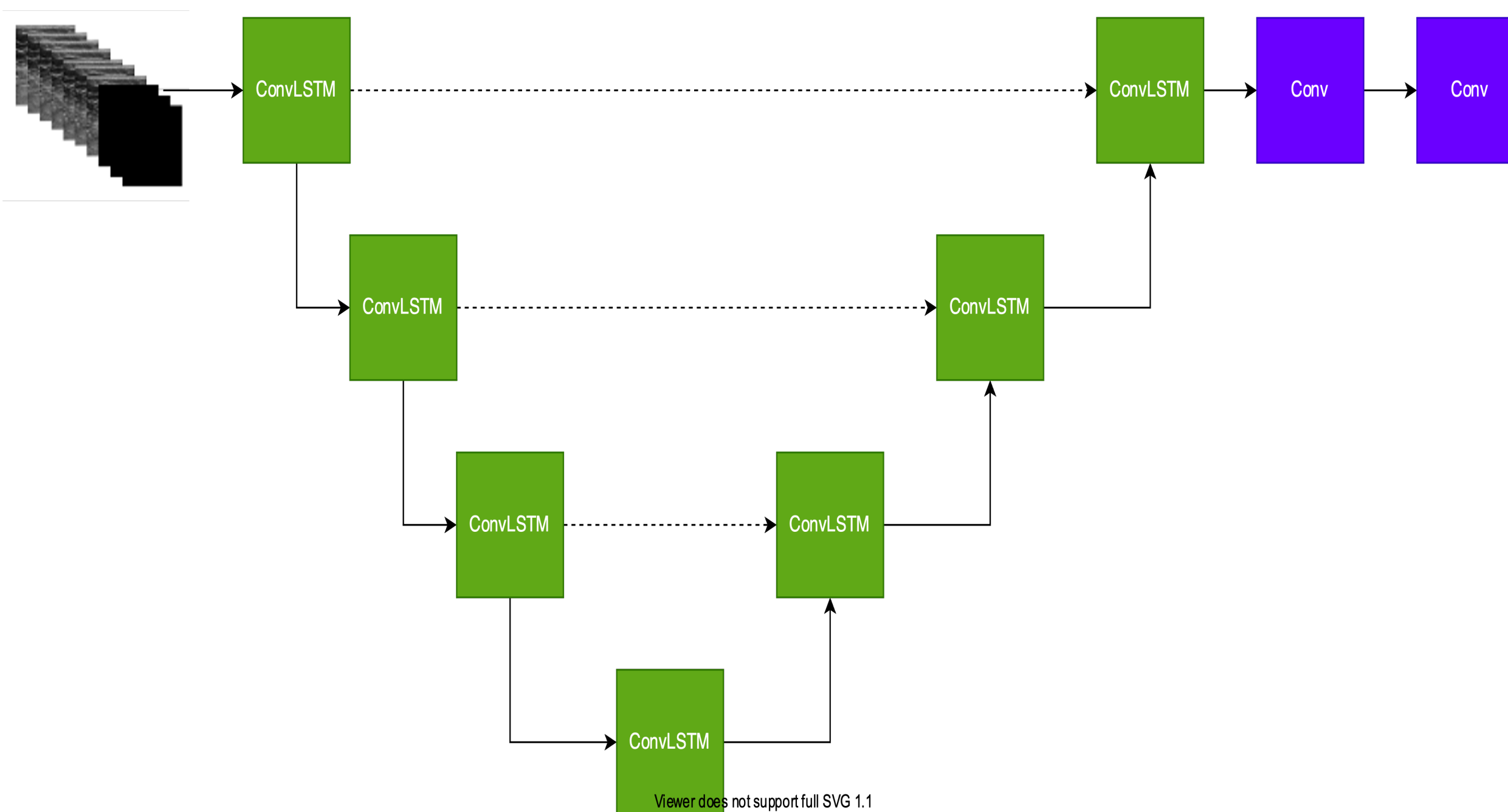
- Model based on **LSTM-U-net** architecture is well suited for the **robust tracking of veins** in ultrasound series in **real-time**
- ConvLSTM layers are more important in the encoder part of a U-net in term of DSC and can be omitted in the decoder in order to make the model smaller for a little reduction in the DSC



**Figure 1. Result of Experiment 1:** Segmentation of the vein over several time steps with predicted position for the last three time steps



**Figure 2. Result of Experiment 2:** Segmentation of the vein if it moves outside of the visible image



**Figure 3.** Diagram of the proposed LSTM-U-net architecture

Architecture	Dice Coefficient	Epochs until conergence	Trainable Parameters
Both	0.959	73	1,859,658
Encoder	0.94	46	1,520,634
Decoder	0.90	40	676,650

**Table 1.** Comparison of different architectures from **Experiment 3**

## Acknowledgement

The CURE-OP (01QE1848B E! 12491) project has received funding from the Eurostars-2 joint programme with co-funding from the European Union Horizon 2020 research and innovation programme.

The German partners were funded by the German Federal Ministry of Education and Research.