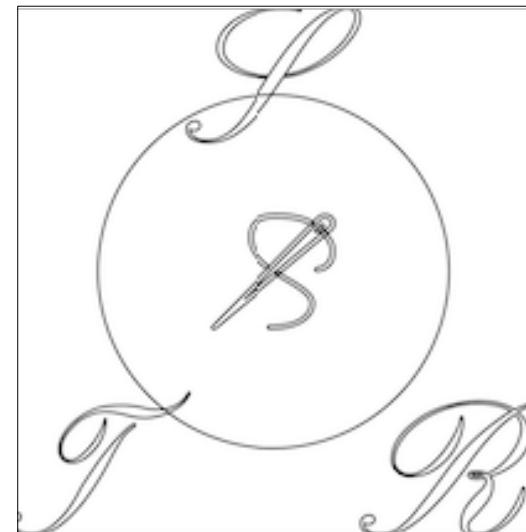
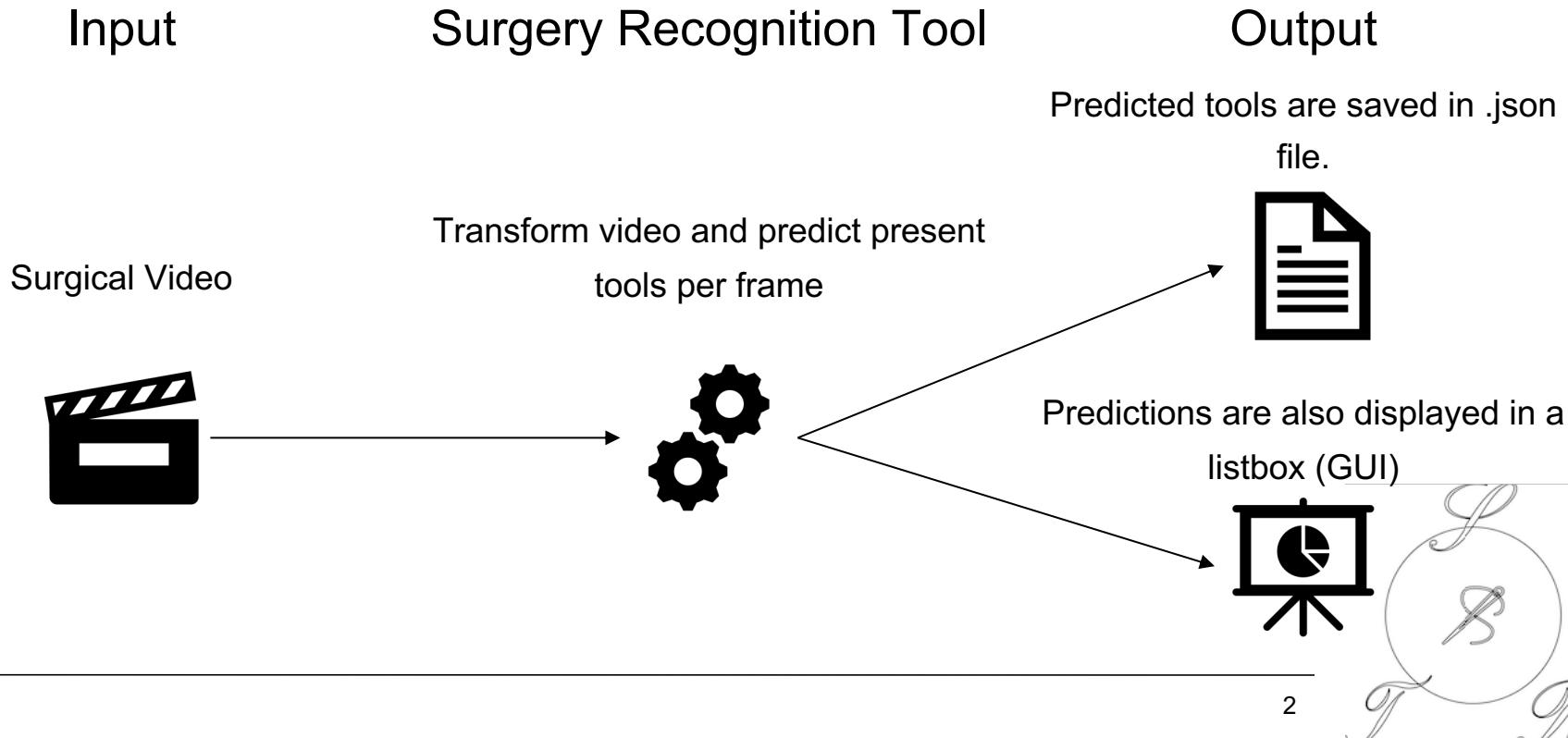


# CAI Classification – Surgery Tool Recognition

Nil Crespo-Peiró, Paco Rahn, Amin Ranem

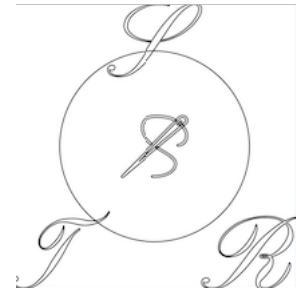


# Introduction





# DATASET



# Cholec80 Dataset

- Contains footage of 80 surgical operations at 25 fps each
- 7 different tools used (Grasper, Bipolar, Hook, Clipper, Scissors, Irrigator, Specimen Bag)



Grasper



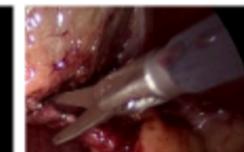
Bipolar



Hook



Clipper



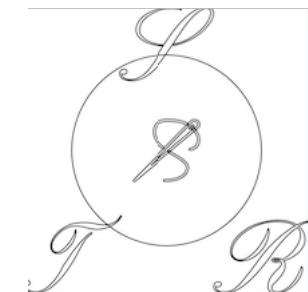
Scissors



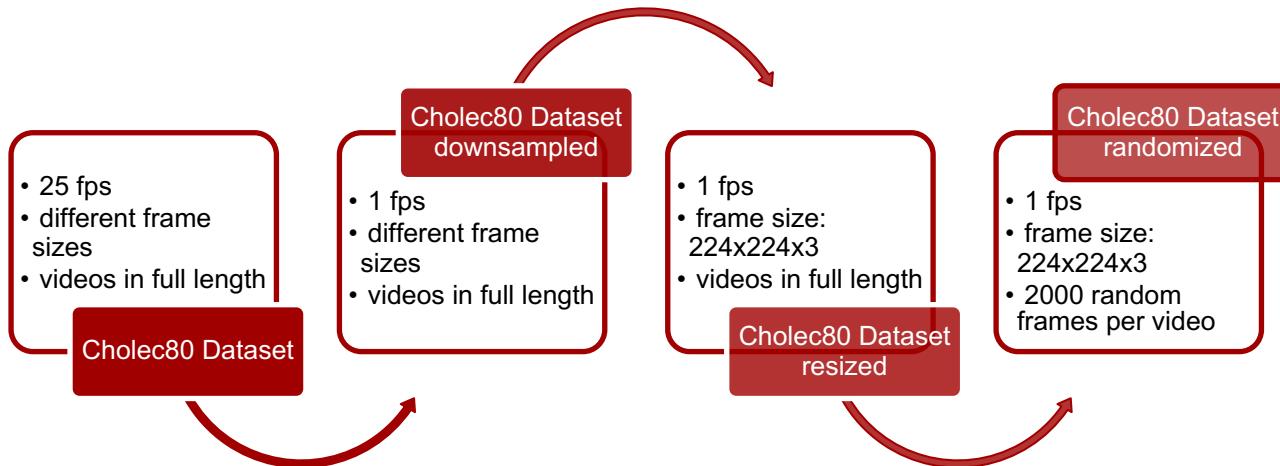
Irrigator



Specimen bag



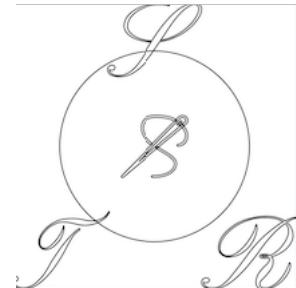
# Cholec80 Dataset Preprocessing





TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

# MODELS



# Models: Structure

## ResNet:

- Transfer learning using pre-trained ResNet50 model (PyTorch) by changing the number of output features from the last layer to 7 (number of tools).
- Idea based on the Cholecystectomy Cataract lecture from the Deep Learning for Medical Imaging (DLMI) course where the ZIB-Net was introduced.

## AlexNet (ToolNet):

- Transfer learning using pre-trained AlexNet model (PyTorch) by changing the number of output features from the last layer to 7 (number of tools).
- Idea based on official [Paper](#) (EndoNet -- ToolNet to detect tools in a surgical video)



# Models: Configuration

## ResNet:

- input\_shape: (224, 224, 3)
- lr: 0.001
- weight\_decay: 0.00005
- batch\_size: 62
- nr\_epochs: 40 (stopped at 25 → bad results)
- random\_frames: True
- nr\_video: 80
- nr\_frames: 2000
- val\_ratio: 0.2
- test\_ratio: 0.3

## AlexNet (ToolNet):

- input\_shape: (224, 224, 3)
- lr: 0.0001
- weight\_decay: 0.001
- batch\_size: 32
- nr\_epochs: 40
- random\_frames: True
- nr\_video: 80
- nr\_frames: 2000
- val\_ratio: 0.2
- test\_ratio: 0.3



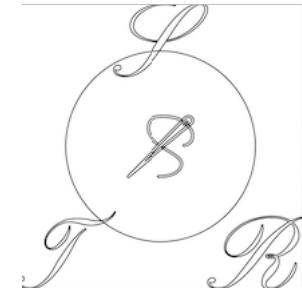
# Models: Results

ResNet:

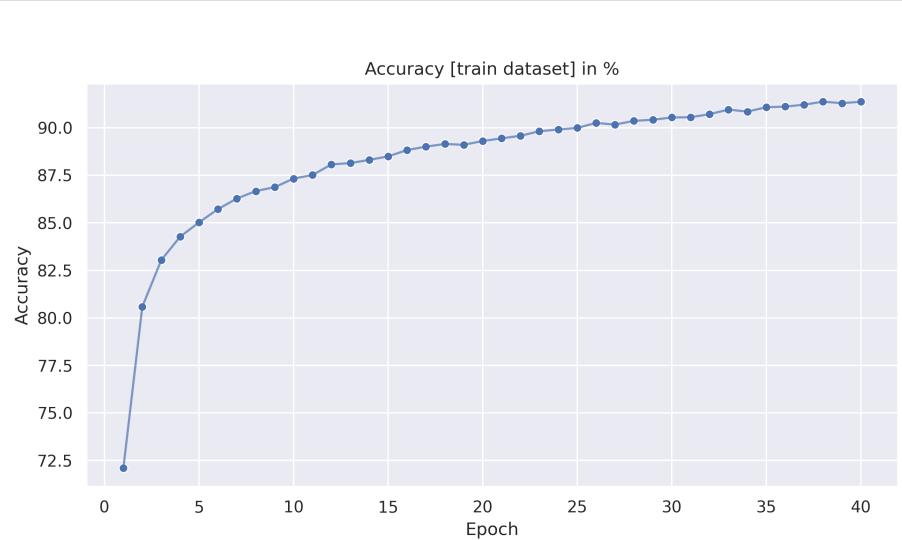
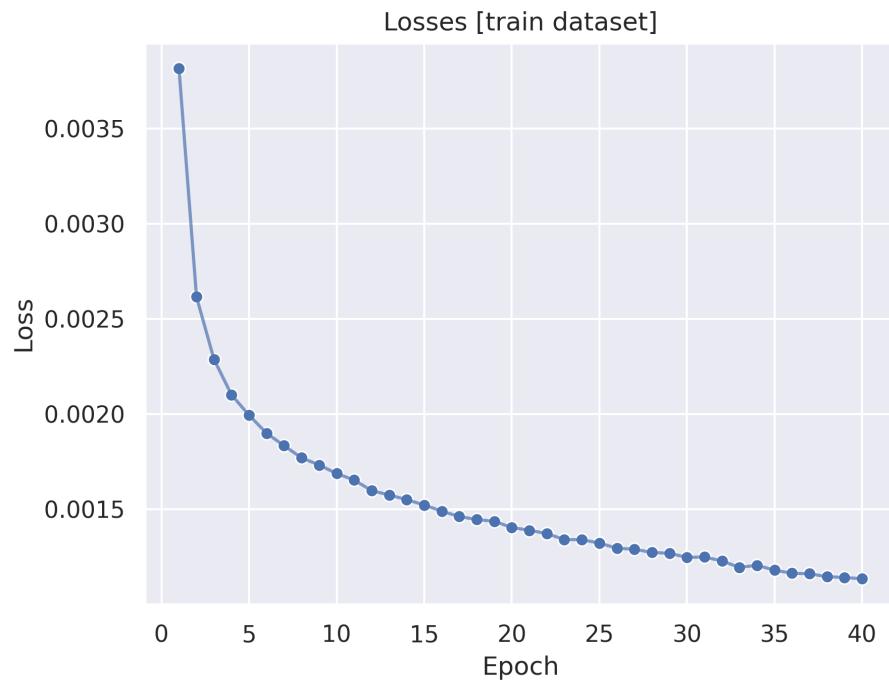
- Training of this approach has been stopped, since the training and validation accuracies were not as promised, although model overfitted on small dataset (see reports):
  - Epoch 10:
    - `train_acc = 30.37%`
    - `val_acc = 30.96%`
  - Epoch 20:
    - `train_acc = 30.37%`
    - `val_acc = 30.97%`

AlexNet (ToolNet):

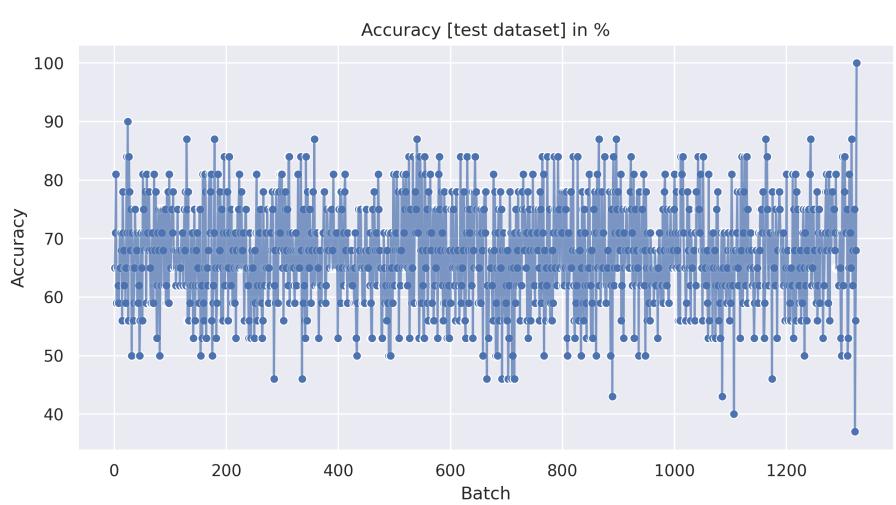
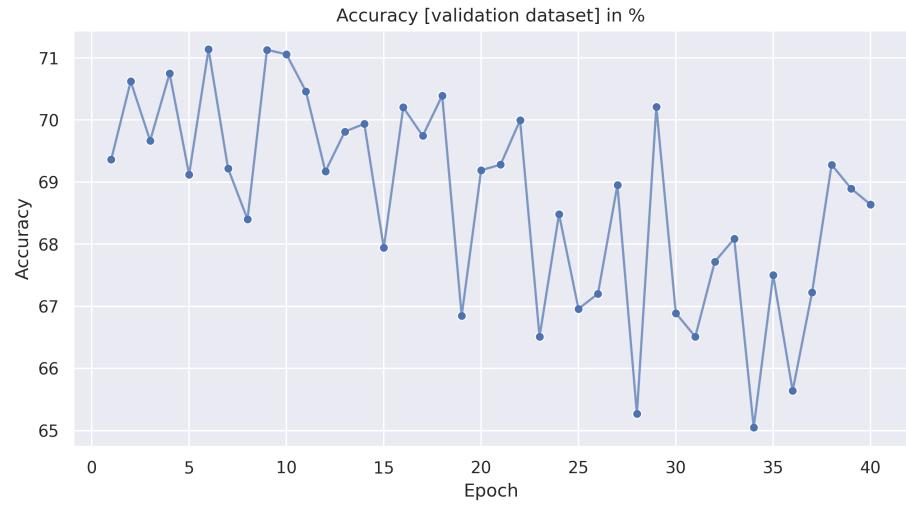
- Training after 40 epochs resulted in an averaged (test) accuracy of 67%
- See corresponding plots on following slide



# Models: Results (AlexNet)

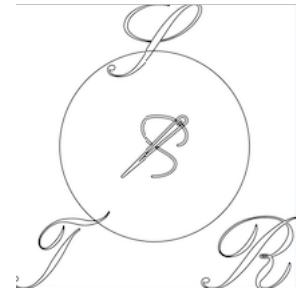


# Models: Results (AlexNet)





# GRAPHICAL USER INTERFACE



# Graphical User Interface: Demo Video

- Used for making predictions with our pre-trained models

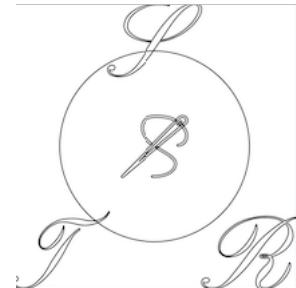
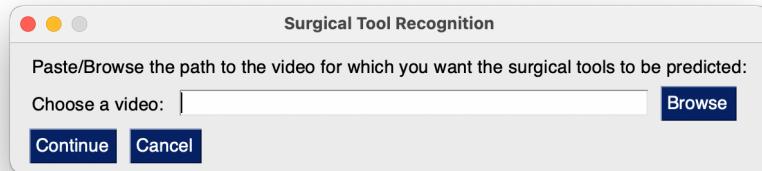


# Graphical User Interface: WelcomeWindow

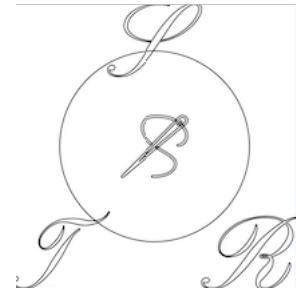
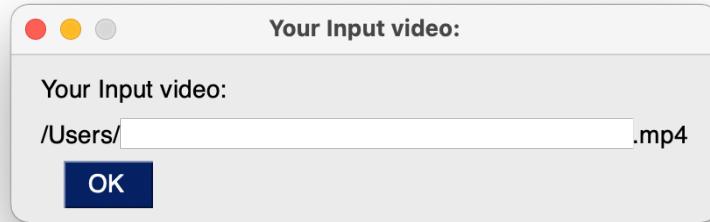


# Graphical User Interface: StartWindow

- User needs to specify path to desired video he/she wants to use

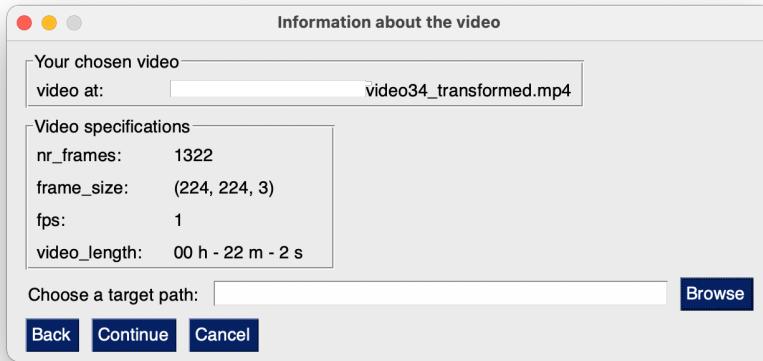


# Graphical User Interface: ConfirmationPopUp

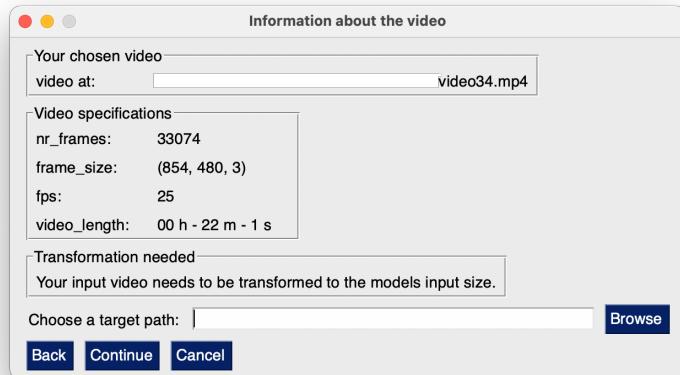


# Graphical User Interface: TransformVideo

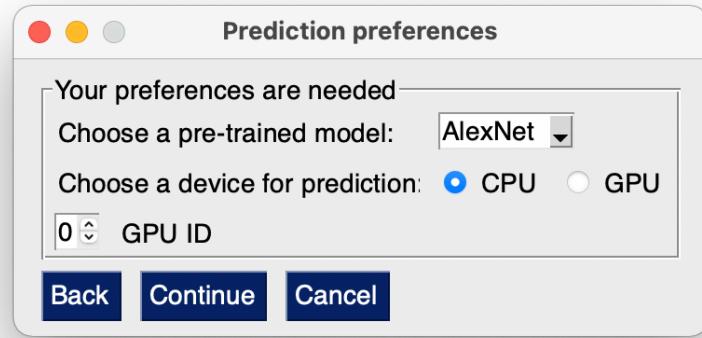
- Video is in desired format



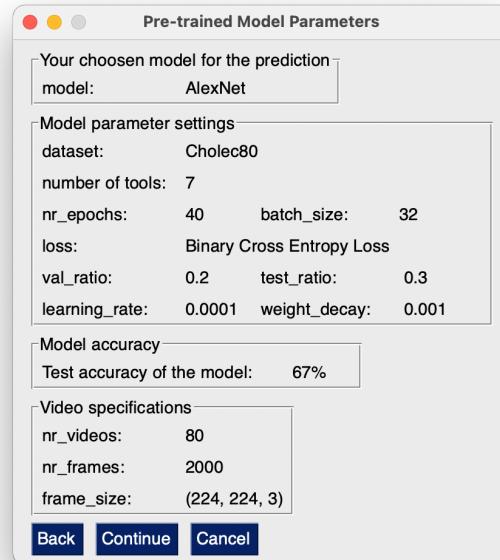
- Video needs to be transformed



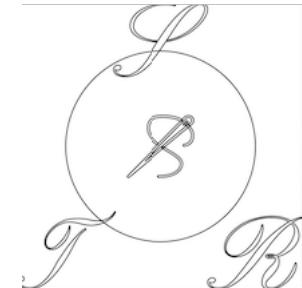
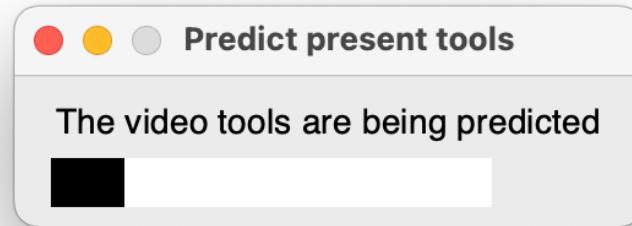
# Graphical User Interface: ChooseModelAndDevice



# Graphical User Interface: ModelSpecs



# Graphical User Interface: PredictVideoTools



# Graphical User Interface: ResultWindow

