SCIENTIFIC TALK - MIC@DKFZ

Federated Learning with Kaapana

Name: Marius Kempf

Supervisors: Klaus Kades, Maximilian Zenk, Klaus Maier-Hein & Niklas Kühl (KIT)

Overview

1. Motivation
2. Gap to The State of The Art
3. Implementation
4. How to up an Experiment
5. Results
6. Outlook

Type text

Motivation

Medical data is rare, sensitive and precious

Need for Data

Deep Learning needs numerous samples to train models which generalize sufficiently

Expensive Labeling

Annotating medical image data requires experts, is time consuming, and is therefore expensive

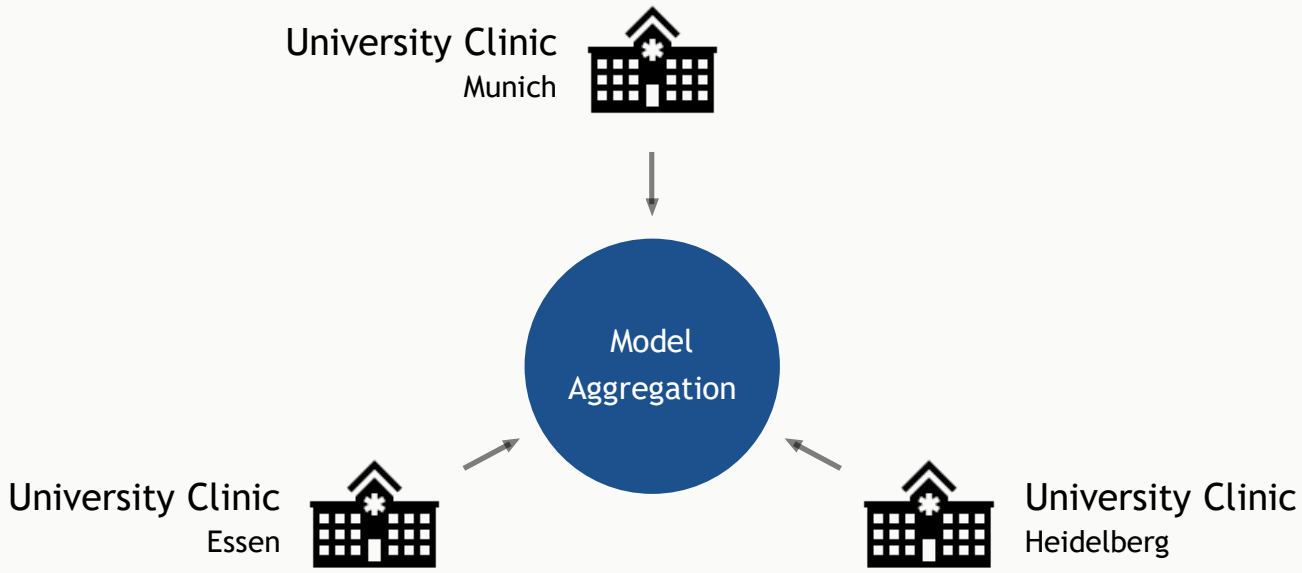
Maintaining Patients' Privacy

Even if annotated data exists locally, it is not permitted to simply share it with others

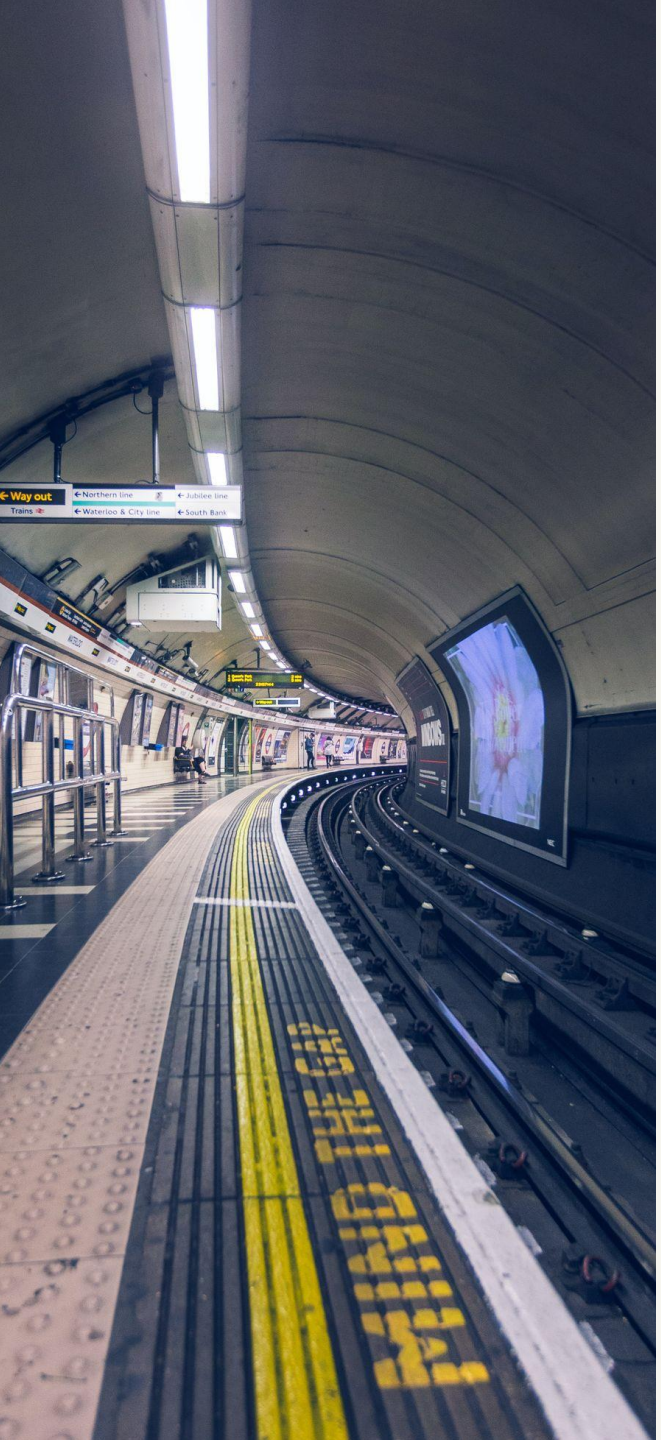
Motivation - Federated Learning

Federate Learning to share model updates instead of raw and sensitive data!

How to close the gap between simulation and real application?



Providing a technical solution to actually conduct federated deep learning experiments across medical institutions



Gap to The State of The Art

Technical solutions to bridge the gap from simulated FL to its application in medical institutions

Few real-world scenarios using medical images

8 scientific publications could be identified

Technical Solutions for FL exist

i.e. TensorFlow Federated, PySyft by Openmined, NVIDIA Clara Federated

Technical Solution for FL in medical Environment

NIVIDIA Clara Federated provides powerful features, but lacks of flexibility and openness

KAAPANA FEDERATED

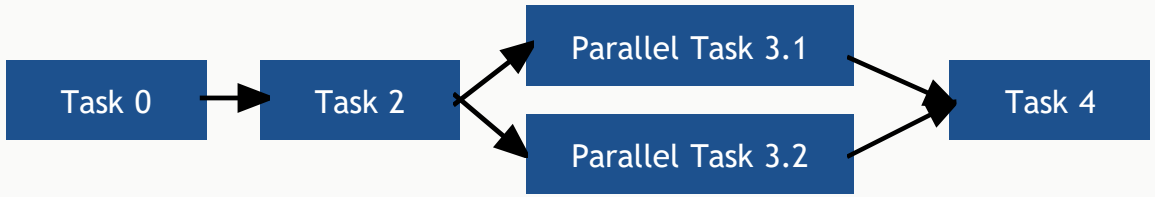
Can we extend Kaapana in such way, that Federated Learning is possible across actual distributed clinics?

Implementation - Apache Airflow & MinIO

Apache Airflow

- Scheduling & monitoring of workflows
- Accessible via its API
- Workflows configured as Directed Acyclic Graphs consisting of multiple operators
- An operator performs exactly one task

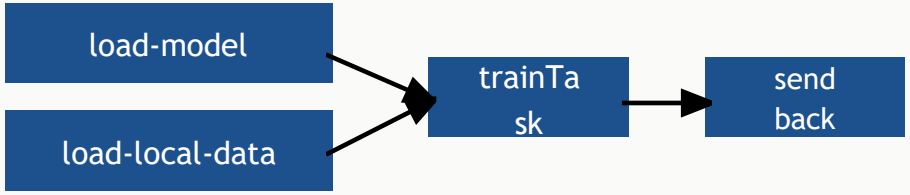
Exemplary Directed Acyclic Graph (DAG):



MinIO

- Object store serving as accessible 'file-system'
- Provides a powerful API
  - Check for objects & files
  - Move, copy, delete, ...
  - ...
  - Used to store model checkpoints, log-files, and test data, ...

Implementation - Kaapana Federated

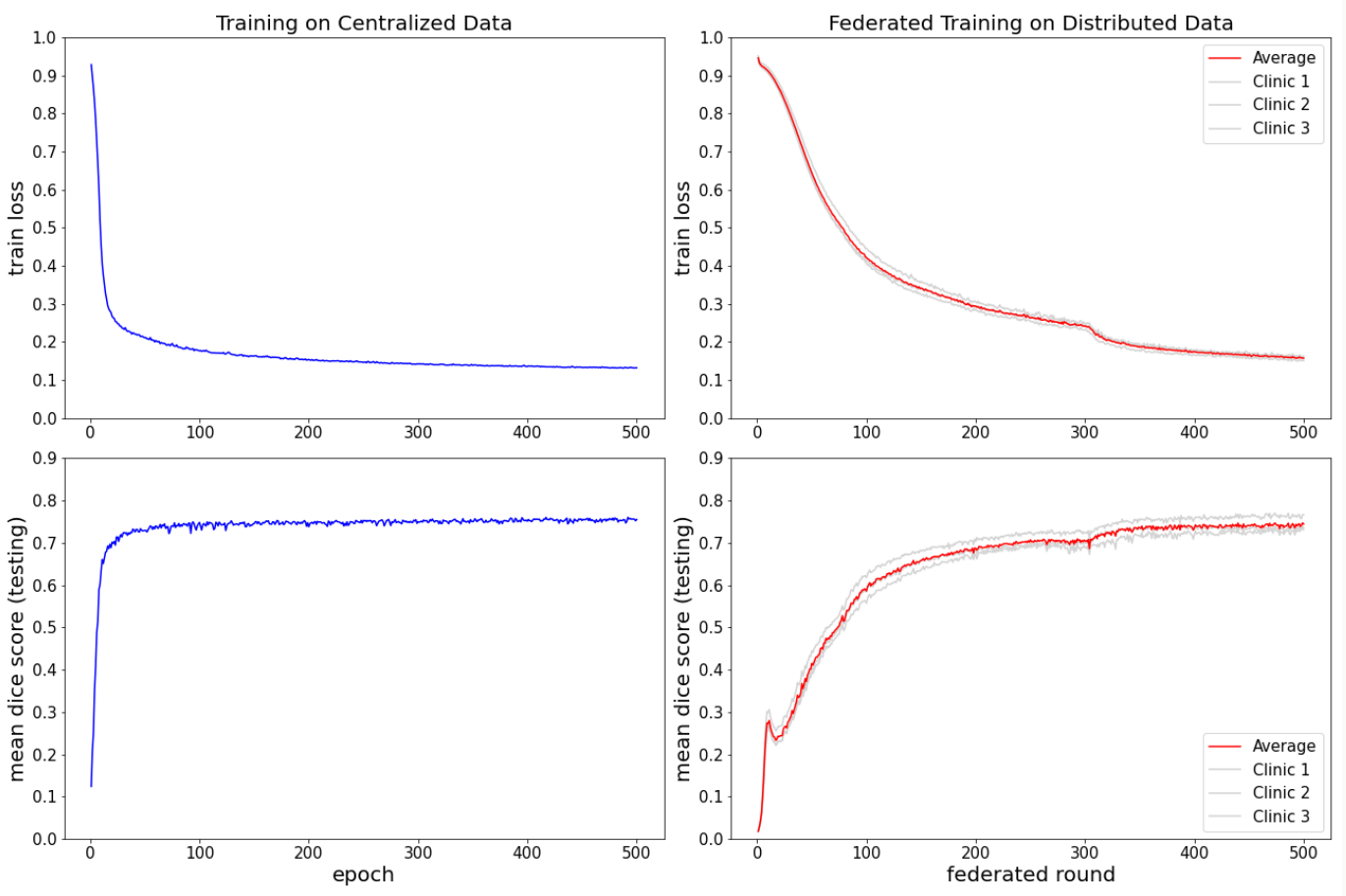


Results - Exemplary Segmentation Experiment using Kaapana Federated

- Brain MRI Scans from Medical Decathlon (BraTS)
- Equally distributed across 3 Kaapana Instances
- Trained with Federated Averaging
- Comparison with training behavior on centralized / pooled data

Training parameters:

- Adam optimizer (treatment strategy: resetting)
  - Learning rate: 0.001
  - & weight decay:
- Batch size: 2
- Epochs / Federated rounds: 500



Outlook

Add privacy mechanisms

Conduct “real” experiments across partner institutions

From prototype to Kaapana extension ...

Thanks a lot four your attention!