A Machine Learning Approach to Artefact Detection in Broadband Near-Infrared Spectroscopy (NIRS)

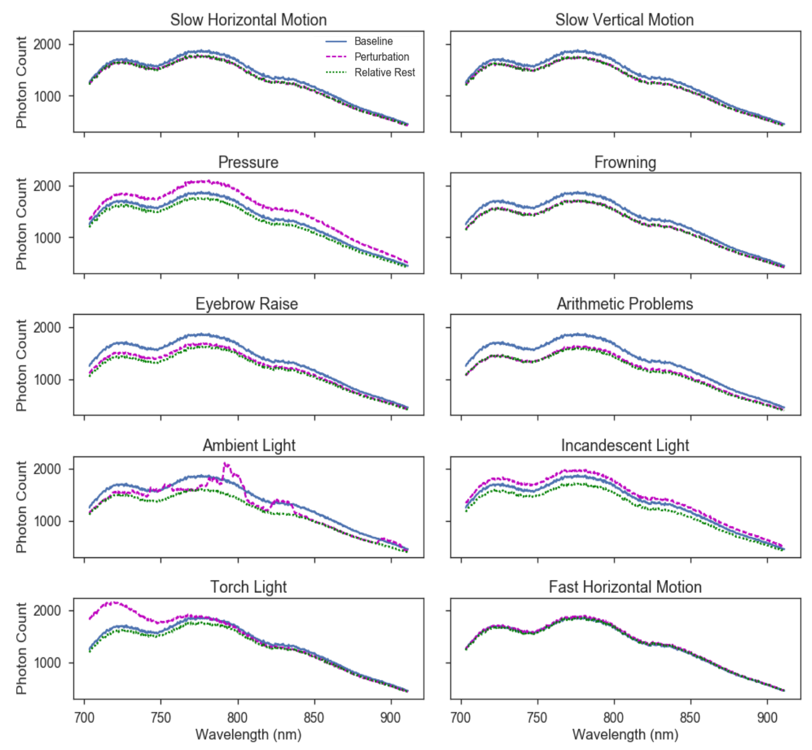
Dionna Jacobson[1*], Ilias Tachtsidis [1,2], & Subhabrata Mitra[2]
CoMPLEX[1], University College London, UK
Department of Medical Physics & Biomedical Engineering [2], University College London, UK
dionna.jacobson.17@ucl.ac.uk[*]

UCL

Introduction

- Broadband NIRS uses >100 wavelengths in the NIR range to resolve cerebral changes in light-absorbing chromophores
- Artefact contamination represents a source of noise in NIRS data, yet current detection methods isolate artefact signal based on faulty assumptions or by relying on external sensors
- Multivariate response or 'spectra' generated by broadband NIRS presents an opportunity for automatic artefact detection
- **Machine learning approaches were used on staged artefact data to:**

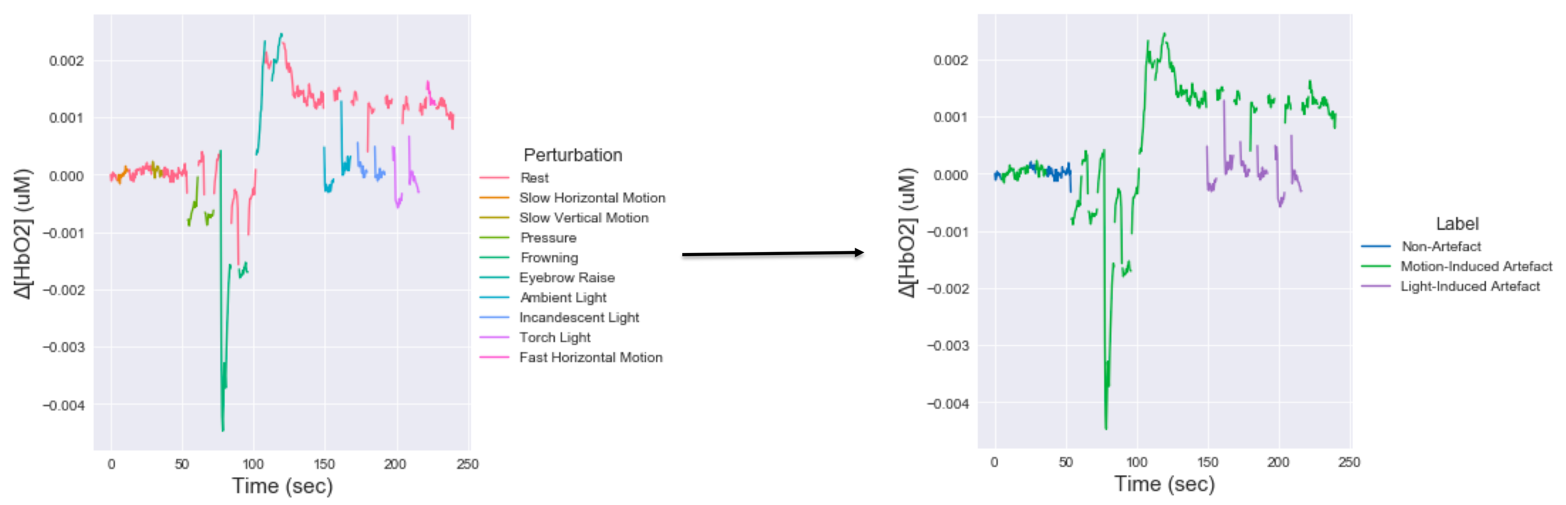
1. **Identify patterns distinguishing artefact spectra**
2. **Determine if artefact spectra can be detected**
3. **Compare performance of two ML classifiers for artefact prediction**
4. **Compare performance of ML-based methods to MARA - the standard detection algorithm**



Broadband spectra from various external perturbations

Artefacts in broadband NIRS spectra

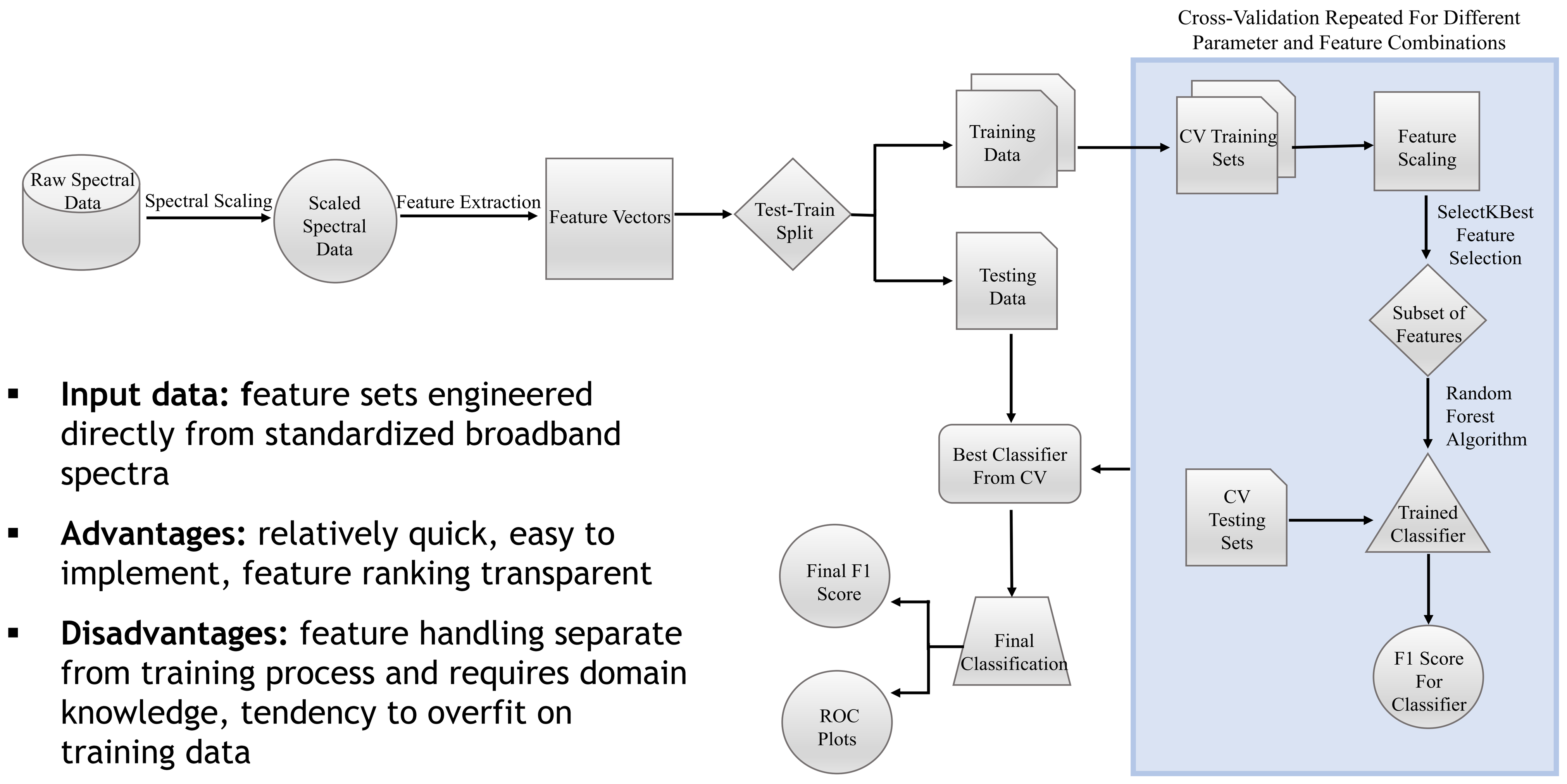
- Motion and light artefacts were staged in 16 healthy volunteers using broadband NIRS
- Spectra were re-labeled into 3 classes: **Non-Artefact, Motion-Induced Artefact, Light-Induced Artefact**



New labeling scheme in staged artefact data

- Two classifiers, Random Forests and Convolutional Neural Networks, were trained on staged data to perform supervised classification

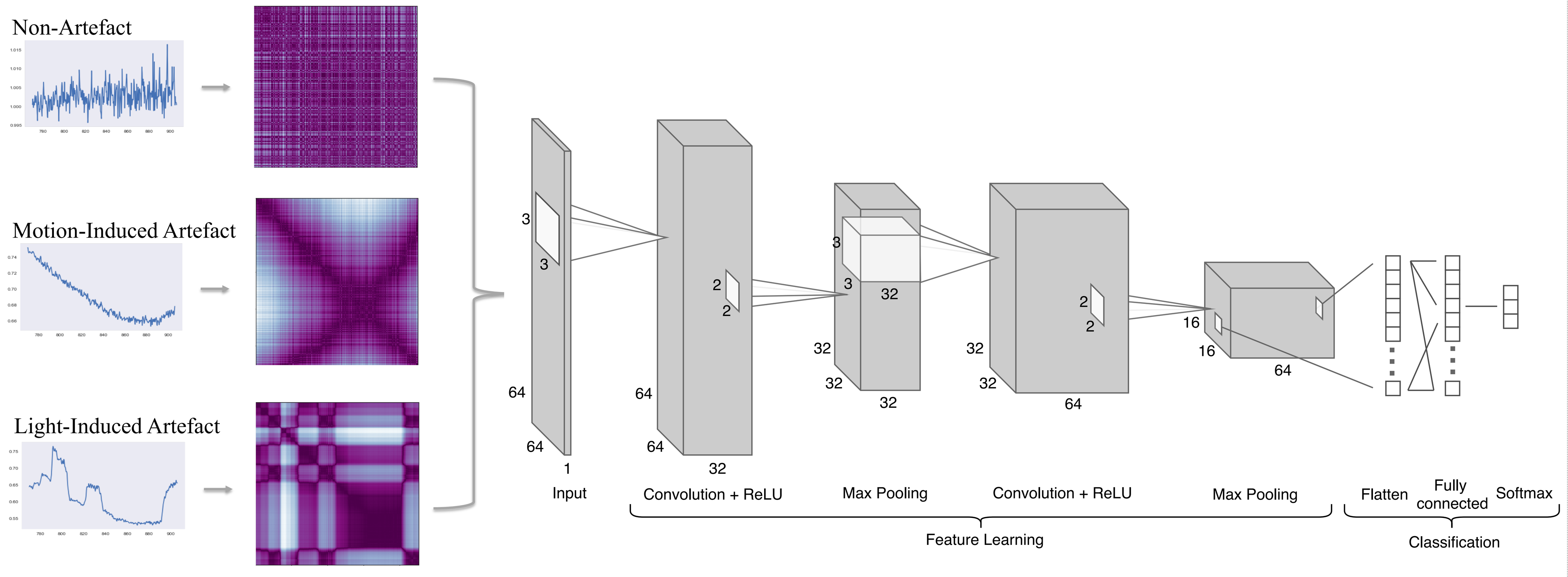
Detection System One: Random Forests (RF)



RF pipeline for artefact detection

- **Input data:** feature sets engineered directly from standardized broadband spectra
- **Advantages:** relatively quick, easy to implement, feature ranking transparent
- **Disadvantages:** feature handling separate from training process and requires domain knowledge, tendency to overfit on training data

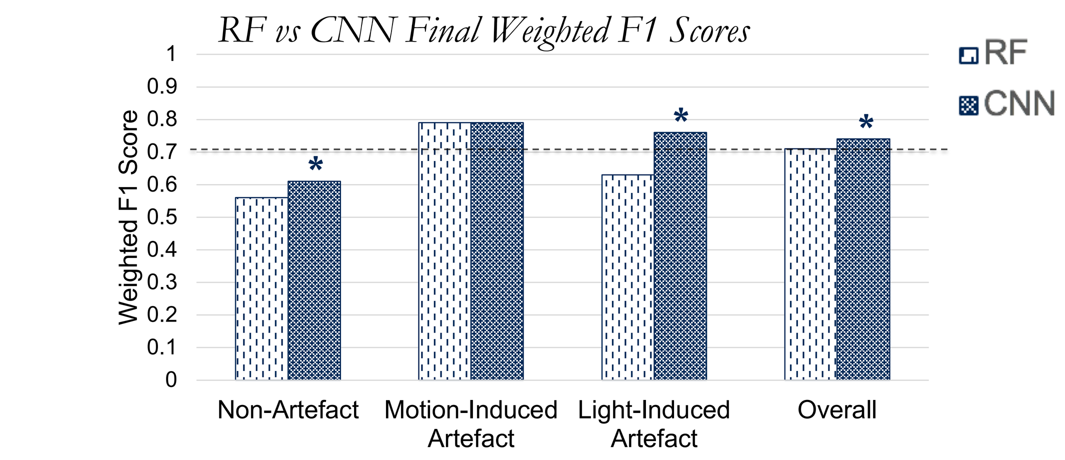
Detection System Two: Convolutional Neural Networks (CNN)

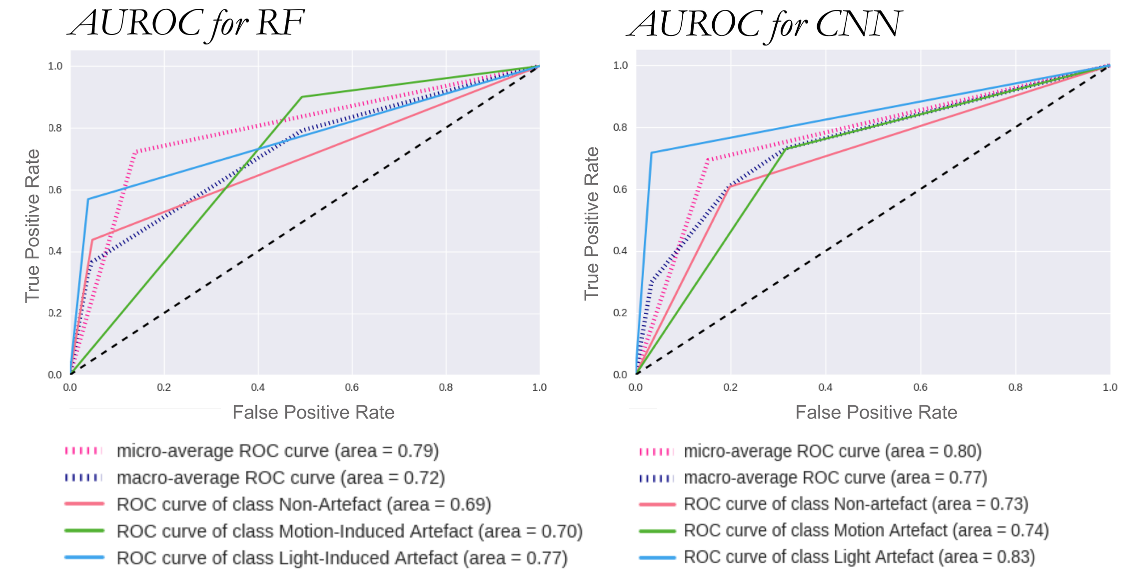


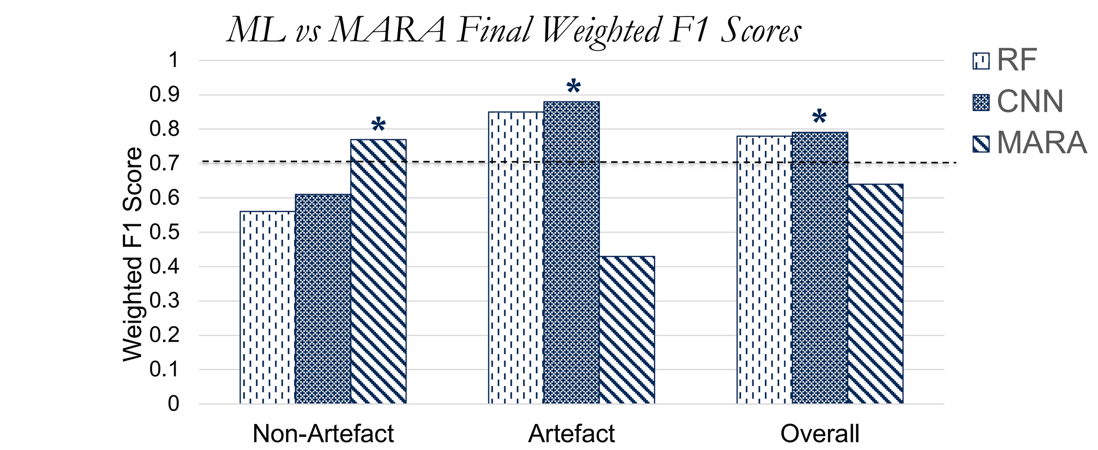
2-stage deep CNN pipeline for artefact detection

- **Input data:** 2D Recurrence Plot transformations on 1D standardized broadband spectra
- **Advantages:** image representation introduces new feature types, automated feature learning incorporated into the training process, weighted loss function to handle class imbalance
- **Disadvantages:** computational complexity, difficult to interpret

ML-based detection outperformed the standard detection method







Conclusions & Future Work

- **A pipeline based on machine learning approaches is proposed for detection of artefacts in broadband NIRS staged data**
- ML-based methods superior performance to standard methods highlight its potential to improve current noise reduction systems applied to NIRS datasets
- Future work is needed to:
  1. Enhance classifier performance
  2. Demonstrate clinical application
  3. Combine ML and MARA to handle 'passive' artefacts and incorporate artefact correction