Comprehensive Framework for Clinical Data Fusion

Anam Haq, Szymon Wilk
Poznan University of Technology (PUT), Poznan, Poland

Alberto Albelló
Universitat Politcnica de Catalunya, Barcelona, Spain

Introduction

Research Goal

Development of a comprehensive framework for clinical data (Image and Non-Image) fusion

Motivation

Development of an accurate and reliable diagnostic and therapeutic decisions models using wealth of patients information.

Challenges

- Underutilization of clinical data
- Significant heterogeneity of data stored in clinical information systems
- Imbalance Data

Comprehensive Framework

Comprehensive Framework Key Components

- Pre-processing clinical and image data
- Control model construction to handle clinical data problems
- Implementing data fusion using General Fusion Framework

General Fusion Framework

Brings all data sources to a *common knowledge representation* before fusing them together.

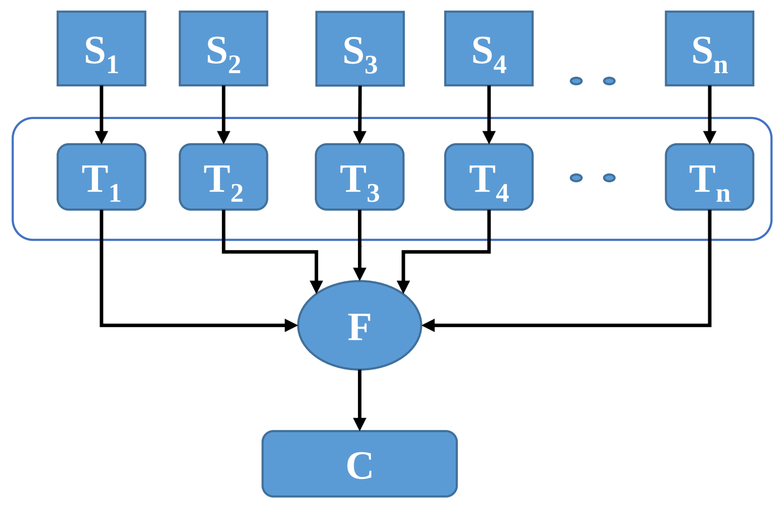


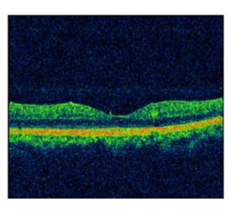
Image and Non-Image Data Analysis

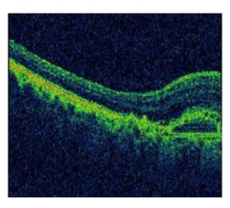
Image Data

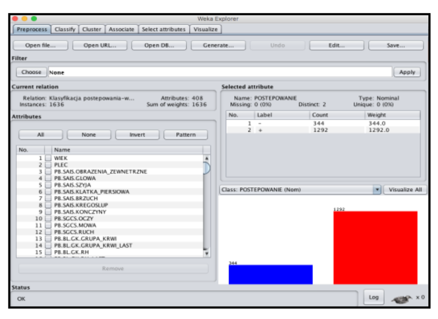
- Identify the most appropriate set of features

Non-Image Data (Clinical Records)

- Identifying appropriate pre-processing techniques for clinical data to resolve issues (imbalanced, noise and heterogeneous)





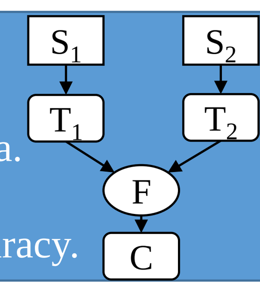


Developed Case Studies

1. Fusion of X-ray and Clinical Records (Prediction of Type of Treatment for Fractured Bones)

The case study follows the GFF model (combination of data approach):

- Transformation on X-ray images to extract attribute.
- Removing attributes having low or high variability from clinical data.
- Integrating both image and clinical attributes
- Building classifiers and evaluating performance with respect to accuracy.



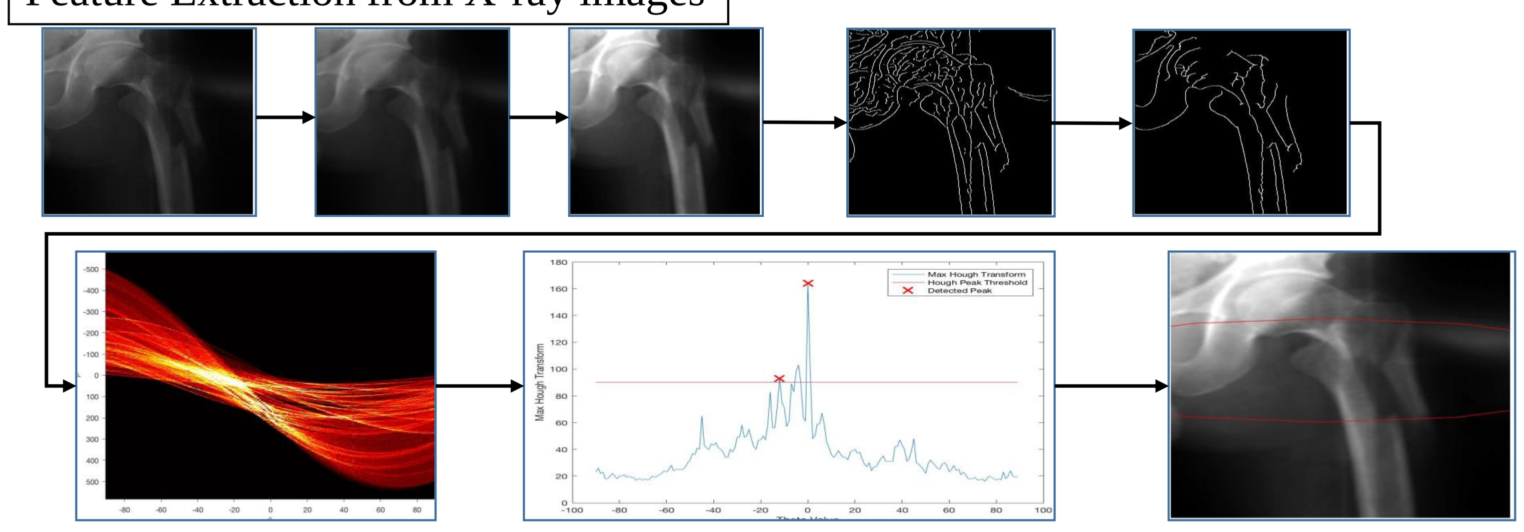
2. Selection of Appropriate Attribute from OCT Images

This case study aims towards the analysis of OCT images:

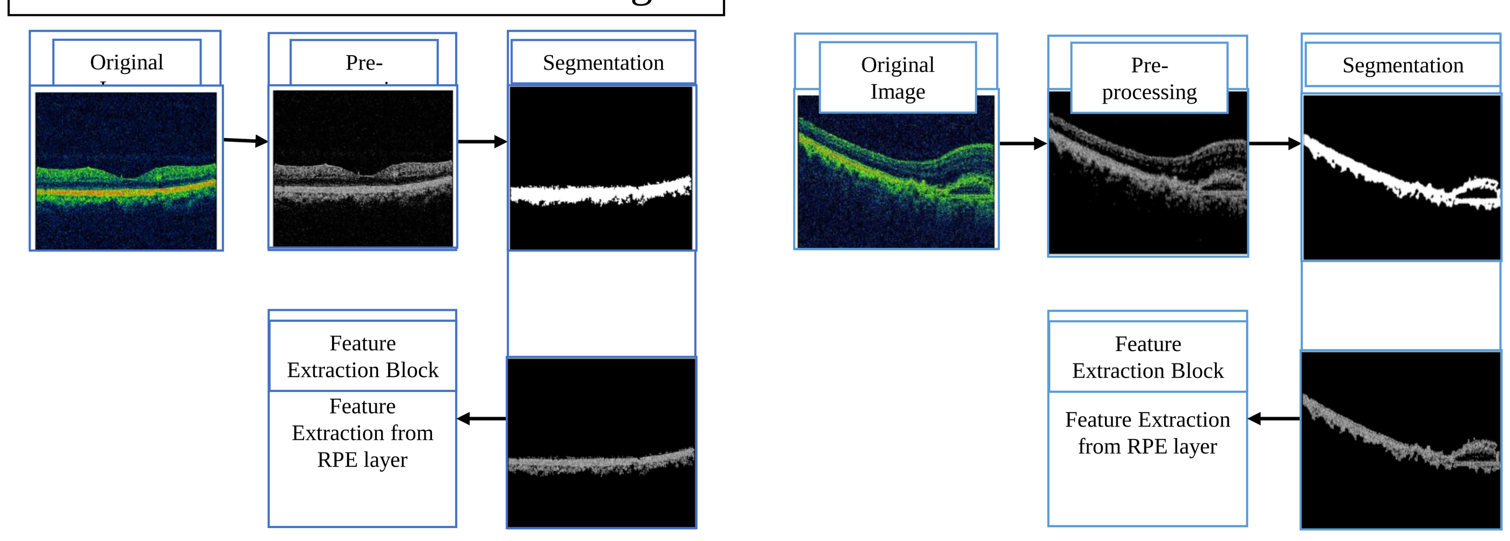
- Extracting structural and textural properties of RPE Layer
- Building classification models for different attribute sets and evaluating their performances based on accuracy, sensitivity, specificity and geometric mean.

Image Analysis Flow Diagram

Feature Extraction from X-ray images



Feature Extraction from OCT images



Case Studies Results

Performance Evaluation of Image and Clinical Data (Fractured Bones)

Classifier	Features	Overall Acc (%)	Non-Surgical (%)	Surgical (%)	G-Mean (%)
Naïve Bayes)	Image	78.64	**77.50**	79.40	78.44
	Clinical	78.64	62.50	88.90	74.54
	Fused	**83.50**	67.50	**93.70**	**79.53**
Support Vector Machine	Image	79.61	**80.00**	79.40	79.70
	Clinical	67.96	62.50	71.40	66.81
	Fused	**80.58**	77.50	**82.50**	**80.00**
Decision Tree	Image	78.64	**82.50**	76.20	**79.29**
	Clinical	66.20	60.00	69.80	64.71
	Fused	**79.61**	75.00	**82.50**	78.66
Random Forest	Image	70.87	65.00	74.60	69.60
	Clinical	75.72	57.50	**87.30**	70.85
	Fused	**79.61**	**82.76**	78.40	**80.55**

Performance Evaluation of Fusion of Structural and Textural Features(OCT IMAGES)

Feature set	Classifier	Acc [\%]	Sens [\%]	Spec [\%]	G-Mean
Abnormalities (S+T)	3-Nearest Neighbour	84.3	84.2	84.4	84.3
RPE Layer (S)	3-Nearest Neighbour	92.2	94.7	90.6	92.6
RPE Layer (S+T)	**3-Nearest Neighbour**	**94.1**	**95.0**	**93.5**	**94.3**
Abnormalities (S+T)	Naïve Bayes	88.2	86.0	90.0	87.8
RPE Layer (S)	Naïve Bayes	94.1	95.0	93.0	94.3
RPE Layer (S+T)	**Naïve Bayes**	**96.1**	**91.3**	**100.0**	**95.5**
Abnormalities (S+T)	PART	88.2	85.7	90.0	87.8
RPE Layer (S)	PART	92.2	90.5	93.3	91.9
RPE Layer (S+T)	**PART**	**94.2**	**91.9**	**96.5**	**93.7**
Abnormalities (S+T)	Support Vector Machine	86.3	88.9	84.8	86.8
RPE Layer (S)	Support Vector Machine	78.4	75.0	80.6	77.8
RPE Layer (S+T)	**Support Vector Machine**	**94.1**	**100.0**	**93.7**	**96.8**

Conclusion

- Appropriate and more reliable systems are developed using both clinical and image information.
- Data fusion improves the system accuracy.
- Selection of appropriate features from images is of great importance.

Future Work

- Evaluating the performance of other data fusion methods.
- Development of a “Control Model”
- Development of a comprehensive clinical fusion framework that can handle associated challenges with the clinical data.