Transfer Learning to detect falls

From young volunteers to elderly

Ludovic MINVIELLE[1,2], Mounir ATIQ[1,2], Sergio PEIGNIER[1], Mathilde MOUGEOT[3], Nicolas VAYATIS[1]

[1]CMLA, ENS Paris-Saclay, CNRS UMR 8536, Universit Paris-Saclay, France

[2]Tarkett, Paris, France

[3]ENSIIE, LPSM, UMR 8001, Paris, France

Abstract

Tarkett recently set up a fall detection monitoring system for nursing homes based on a sensitive floor sensor and the Random Forest algorithm. The random forest model, was trained using data recorded by volunteers, that simulated falls. However, young volunteers simulated fall signals that are likely to be different from elderly real fall ones. In order to cope with this problem we use Transfer Learning over Random Forest. We present here preliminary work using algorithms from [4] and we provide their results on classical datasets and on our own data.

Motivation

Fast and reliable fall detection systems :

- Improve chances of surviving the accident
- Cope with physical and psychological consequences

However, the training data was obtained using signals from volunteers, which makes it different from real world data and can thus give us lower performance.

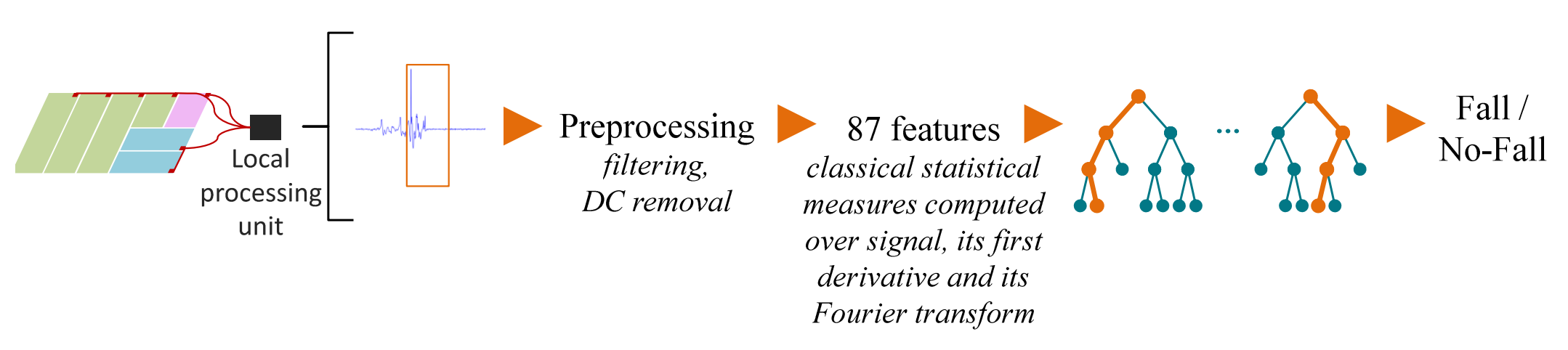

Local processing unit
Preprocessing
filtering, DC removal
87 features
classical statistical measures computed over signal, its first derivative and its Fourier transform
...
Fall / No-Fall

Figure 1: Fall detection workflow

Problem statement

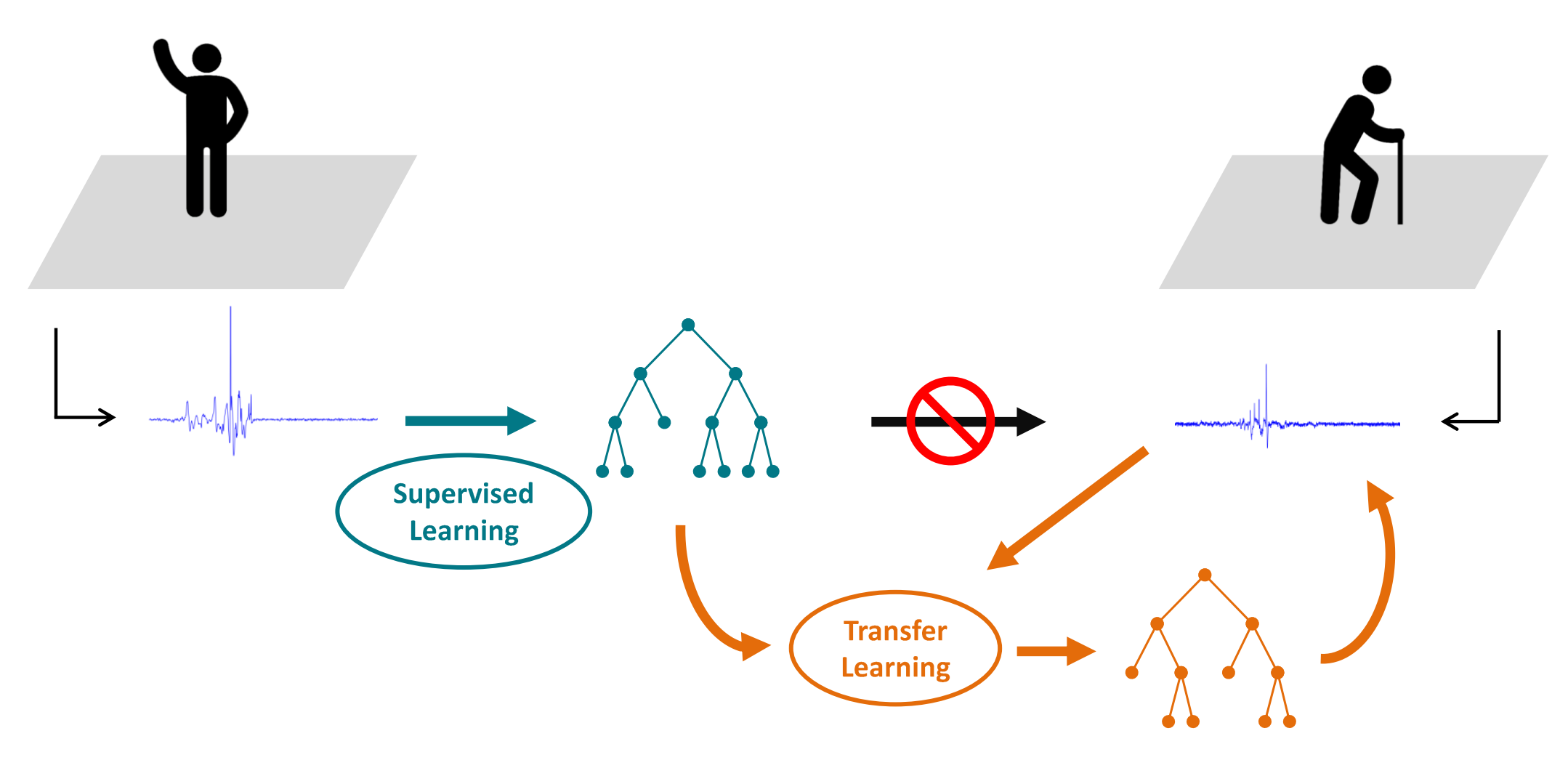


Figure 2: Transfer learning with few labeled target data and the RF trained on source data data

We denote a domain $\mathcal{D} = \{\mathcal{X}, P(X)\}$ where $\mathcal{X}$ is a feature space and $P(X)$ is a marginal probability distribution. We also denote a task $\mathcal{T} = \{\mathcal{Y}, P(Y|X)\}$ where $\mathcal{Y}$ is a label space and $P(Y|X)$ a conditional probability distribution.
We can define a source domain $\mathcal{D}_S = \{(x_{S1}, y_{S1})..., (x_{Sn}, y_{Sn})\}$ where $x_{Si} \in \mathcal{X}_S$ and $y_{Si} \in \mathcal{Y}_S$ and a target domain $\mathcal{D}_T = \{(x_{T1}, y_{T1})..., (x_{Tn}, y_{Tn})\}$ where $x_{Ti} \in \mathcal{X}_T$ and $y_{Ti} \in \mathcal{Y}_T$. In the same way, the source task is denoted $\mathcal{T}_S$ and the target task $\mathcal{T}_T$.
In our case:

- $\mathcal{X}_S = \mathcal{X}_T$
- $P(X_S) \neq P(X_T)$
- $\mathcal{Y}_S = \mathcal{Y}_T$
- $P(Y_S|X_S) \neq P(Y_T|X_T)$

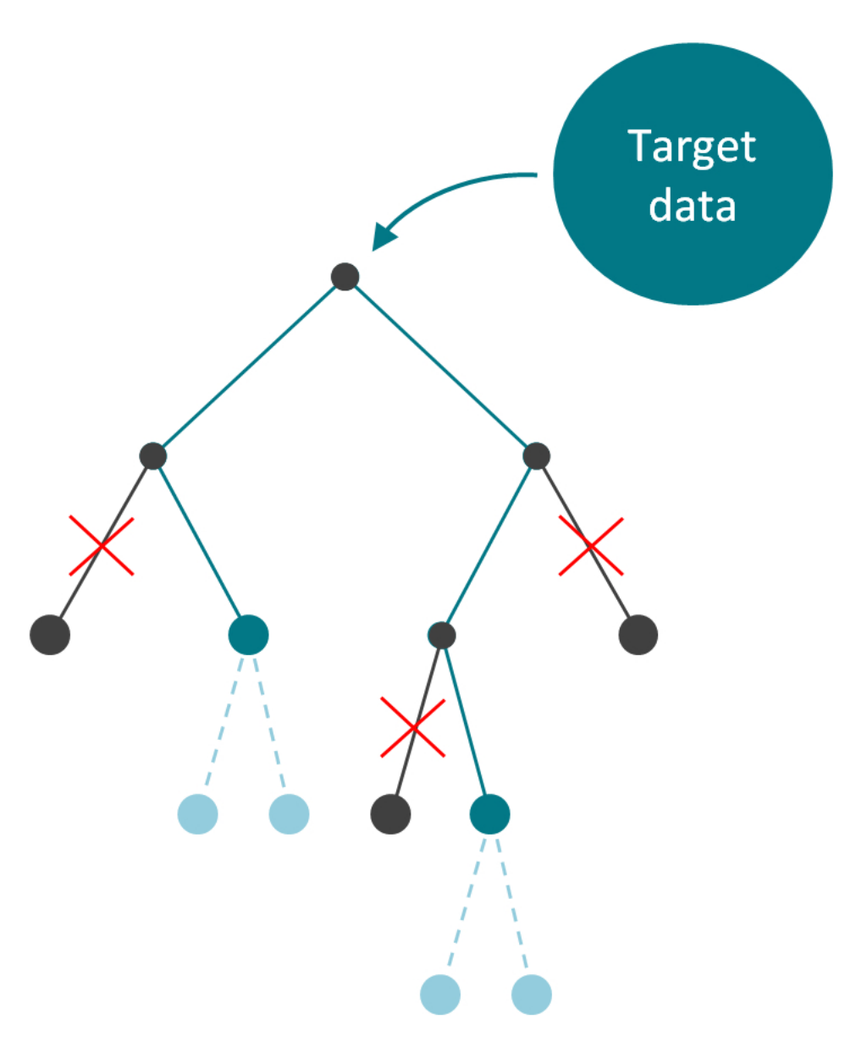
Method

Base model

Model-based Transfer Learning algorithms for decision trees (DTs) [4]:

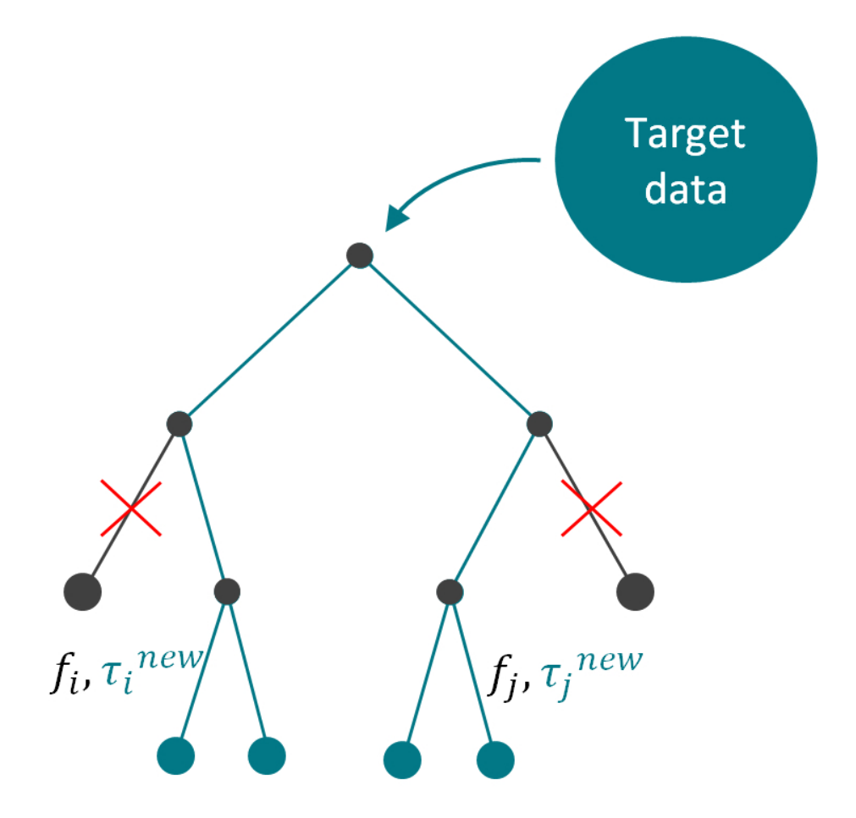
Structure Expansion Reduction (SER)

- Expand leaves using $\mathcal{D}_T$
- Compute subtree errors
- Prune low quality subtrees (w.r.t. error)



Structure Transfer (STRUT)

- Prune unreachable trees
- Recompute thresholds (keep structure)



Given a RF trained on $\mathcal{D}_S$:

- **SER :** apply SER on all its DTs
- **STRUT :** apply STRUT on all its DTs
- **MIX :** apply SER on 50% DTs, STRUT on the other 50%

Variant

As labelling data is costly (few labelled *Fall* events with a lot of available non-fall events), we also try two variations of SER:

- **SER_nosercl:** we do not apply the SER algorithm on *Fall* class leaves of the initial DT
- **SER_noredcl:** we do not apply Reduction step of SER on *Fall* class leaves of the initial DT

As few target data are available we hope that it will prevent from useful *Fall* leaves from the initial DT to be lost.

Experiments

Public datasets [2] [1]

Letter

Source and target separated along the median value of feat. *x2bar*

Wine

Source data is made of white wines and target data of red wines

Digits

Target data : random part of the dataset with inverted pixel values

Tarkett datasets

Simulated

28 volunteers aged from 25 to 45 years old. 742 acquisitions including 409 falls (55%).

Real world

Recorded signals from various nursing homes. 98 falls and 1200 non fall events.

Algorithms

- Comparison with model trained only on source dataset (**SourceOnly**), and model trained only on target dataset (**TargetOnly**)
- Tests with different maximum allowed depths for RF trained on $\mathcal{D}_S$, going from 1 to 15 and None (i.e. no maximum depth, the tree is *full*)
- Target dataset divided in training and testing with $|\text{Test}| = 0.05|\text{Target}|$.
- Tarkett data (imbalanced dataset) : folds such that $|\text{Test}| = 0.2|\text{Target}|$
- Mean scores of Accuracy, Sensitivity (i.e. True Positive Rate) and Specificity (i.e. True Negative Rate)

Results and conclusions

Results are given on Fig. 3.

Results

We expect transfer algorithms to give better performance than SourceOnly. However if TargetOnly is still better, then it is *negative* transfer.

- Depending on the dataset, transfer is not always necessary
- Wine : training only on source data with a short max depth (here 3) can be as good as training on target data
- As depth goes higher, STRUT is better
- Tarkett data :
  - Sensitivity : *negative* transfer ?
  - Specificity : transfer does not beat the SourceOnly model
- Not applying SER or its Reduction step on *Fall* class does not change significantly SER algorithm

Future work

- Compare with simple pruning
- Combine with sampling methods for imbalance problem
- Adapt to the case where $\mathcal{X}_S \neq \mathcal{X}_T$
- Compare to domain adaptation techniques (model independent)

References

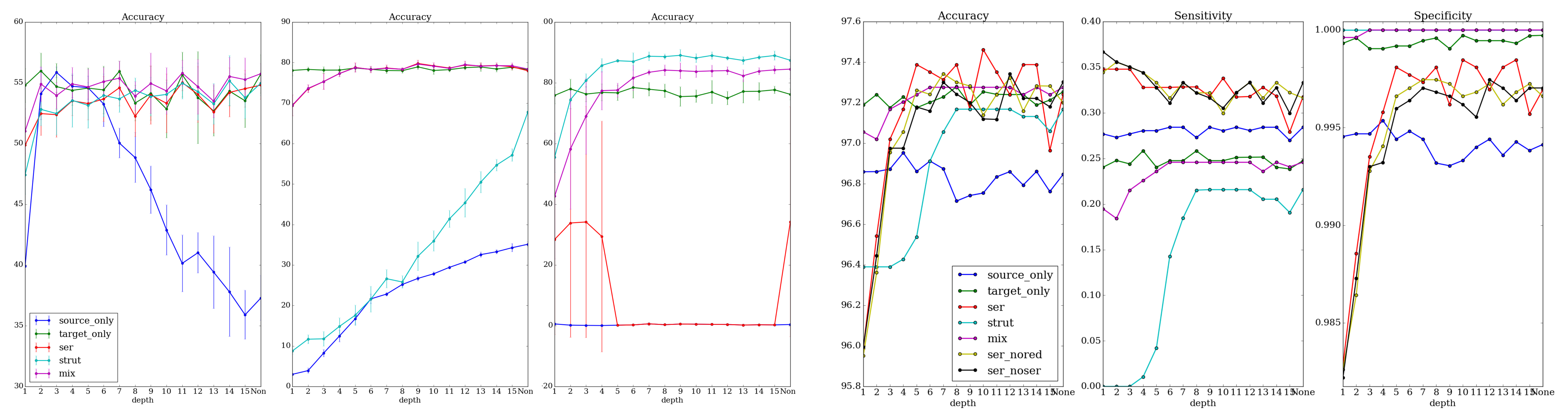
[1] Paulo Cortez, António Cerdeira, Fernando Almeida, Telmo Matos, and José Reis. Modeling wine preferences by data mining from physicochemical properties. *Decision Support Systems*, 47(4):547–553, 2009.

[2] Dua Dheeru and Efi Karra Taniskidou. UCI machine learning repository, 2017.

[3] L. Minvielle, M. Atiq, R. Serra, M. Mougeot, and N. Vayatis. Fall detection using smart floor sensor and supervised learning. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3445–3448, July 2017.

[4] Noam Segev, Maayan Harel, Shie Mannor, Koby Crammer, and Ran El-Yaniv. Learn on Source, Refine on Target: A Model Transfer Learning Framework with Random Forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(9):1811–1824, 2017.

[5] Karl Weiss, Taghi M. Khoshgoftaar, and DingDing Wang. *A survey of transfer learning*, volume 3. Springer International Publishing, 2016.



(a) Wine

(b) Letter

(c) Digits

(d) Tarkett

Figure 3: Scores on public datasets ((a), (b), (c), and on our data (d)