Data Challenge

Hand Gesture Recognition Using 3D Skeletal Dataset

Anas ESSOUNAINI
Ali MOURTADA
Imad AOUALI
Abdellah RAHMANI

09 March 2020

Content

- Convolutional architecture
- Reccurent architecture
- Aggregating multiple neural networks
- Results

Convolutional architecture

A first approach

Architecture : main elements

- Input shape : (171,22,2)
- 4 sequential convolutional layers for features extraction
- 3 fully connected layers
- High penalty for regularization

Commentary

- Sequential deep convolutional layers allow us to extract higher-level features

Reccurent architecture

A second approach : GRU based

Architecture : main elements

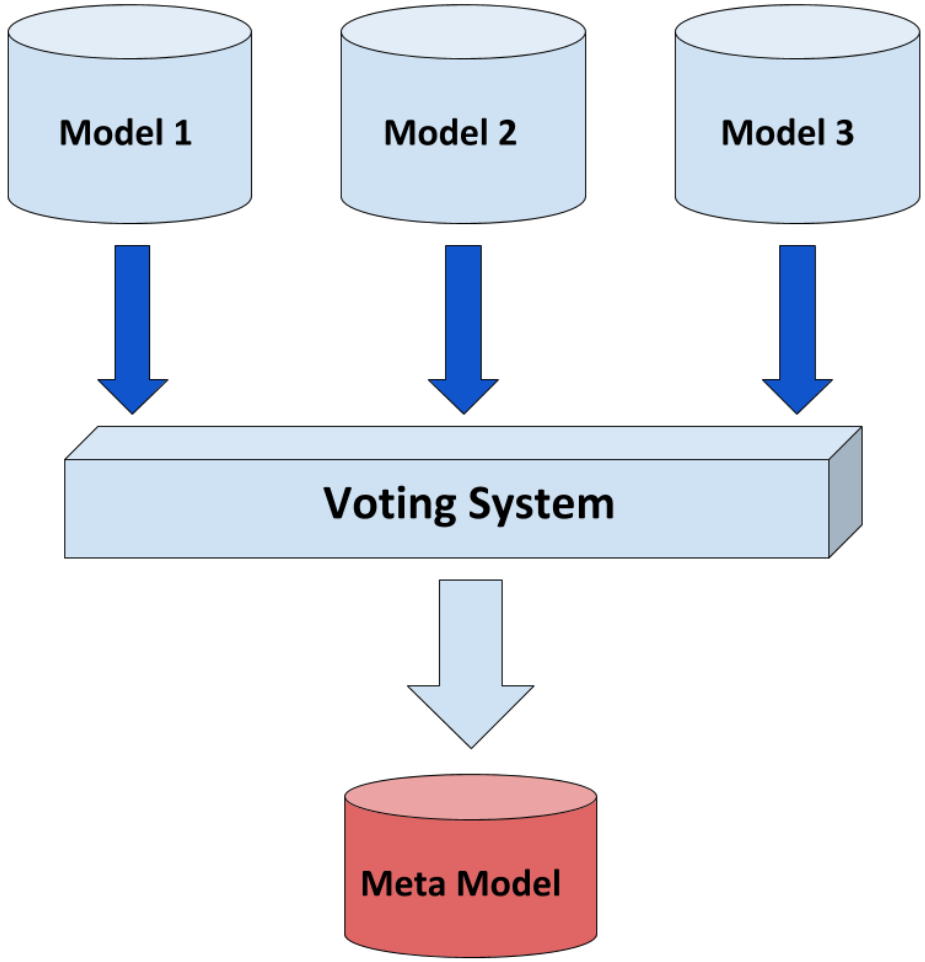
- Input shape : $(171, 66)$
- 3 sequential one dimensional convolution layers for features extraction
- 2 GRU layer
- 1 fully connected layer

Commentary

- GRU works well on small datasets
- more important variance compared to convolutional architecture

Aggregating Neural Networks

A final approach: Ensemble method



Aggregating Neural Networks

A final approach : Ensemble method

Majority Vote Ensemble

- Let f_{ens} denote the aggregated predictor
- The aggregation rule reads

$$f_{\text{ens}}(\mathrm{x}) = \arg\max_{y \in \mathcal{C}} \sum_{m=1}^{M} \mathbb{1}_y \left(f_m(\mathrm{x})\right)$$

- M : The number of models
- $\mathcal{C}$: The set of classes

Results

- CNN gives an accuracy of 83% on validation data.
- GRU is unstable with accuracies ranging from 81% to 85%.
- A combination of these models allowed us to reach 91% in accuracy.

Thank You