



HAL
open science

Otolith age estimation by machine learning approaches

Émilie Poisson Caillault, Florentin Vanhelst, Charlet Bastian, Kélig Mahé

► **To cite this version:**

Émilie Poisson Caillault, Florentin Vanhelst, Charlet Bastian, Kélig Mahé. Otolith age estimation by machine learning approaches. IOS2018 - 6th International Otolith Symposium 2018, Apr 2018, Keelung, Taiwan. hal-03444225

HAL Id: hal-03444225

<https://hal.science/hal-03444225v1>

Submitted on 3 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Otolith age estimation by machine learning approaches

Emilie Poisson Caillault¹, Florentin Vanhelst¹, Bastian Charlet¹, Kélig Mahé²

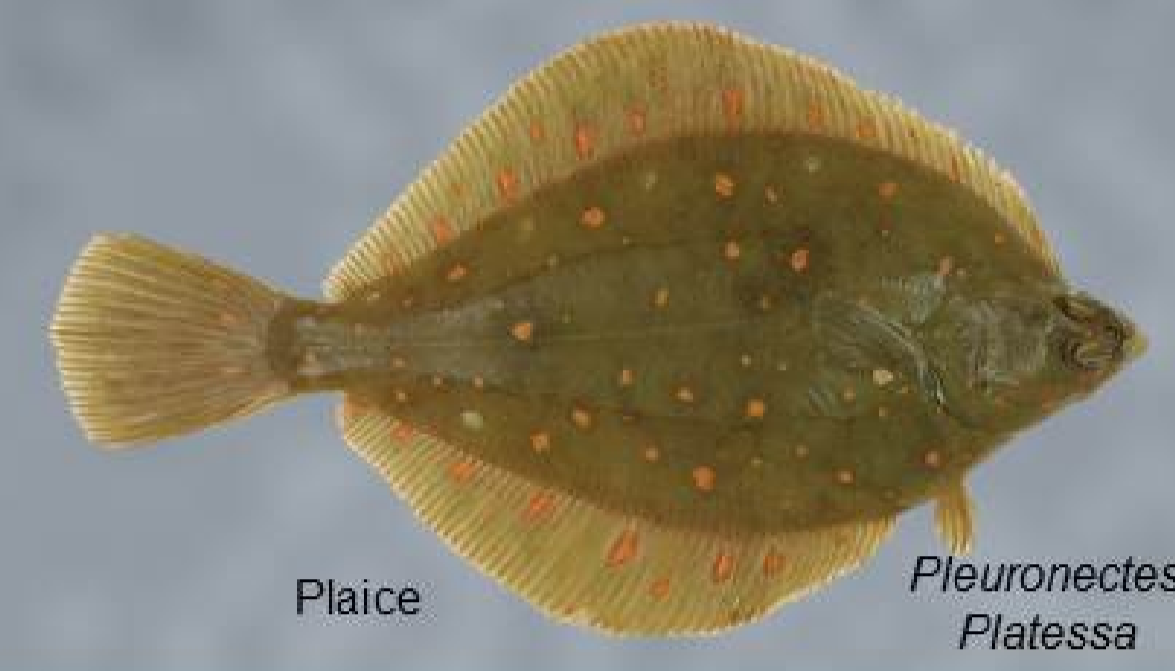
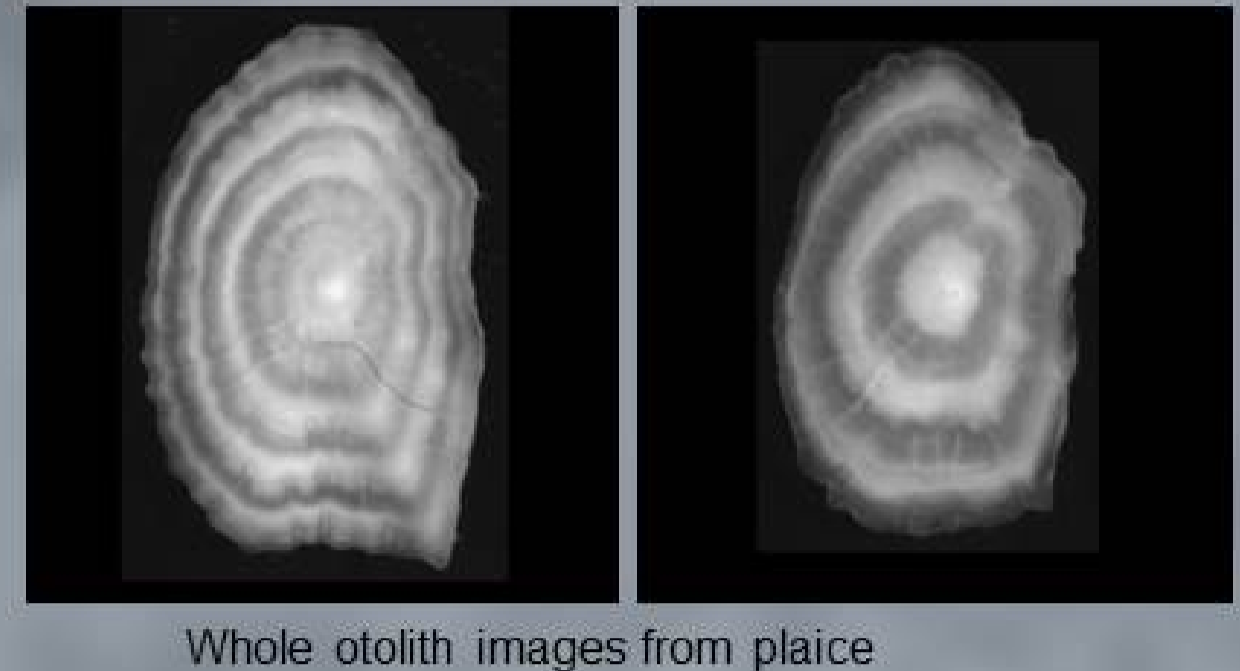


6th International Otolith Symposium
2018, 15-20 April
Keelung, Taiwan

¹ LISIC, ULCO EA 4491, Université du Littoral Côte d'Opale, 62228 Calais, France.
² Ifremer, Fisheries Laboratory, 62321 Boulogne sur Mer, France

Otolith sampling and age estimation by Human experts

- Otoliths are paired calcified structures found in both inner ears of bony fish.
- Otoliths are used to determinate fish ages with different growth periodicities ranging from daily to annual increment.
- Between 800,000 and 1,000,000 otoliths are sampled and analysed per year
- Annual cost of approximately \$8 million
- The longest step in this process is the preparation and reading of otoliths by the experts.

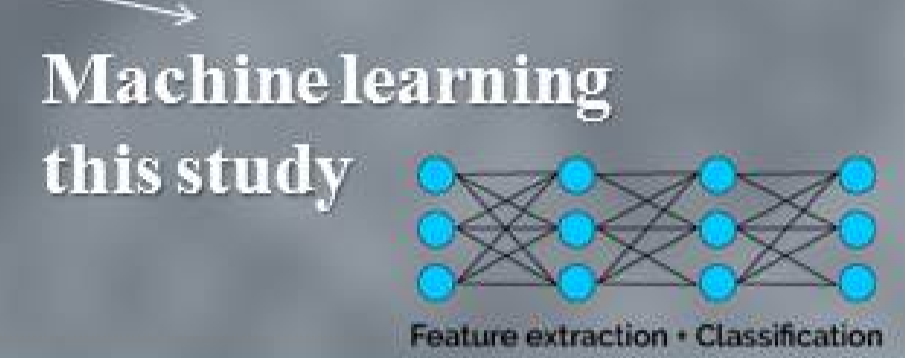
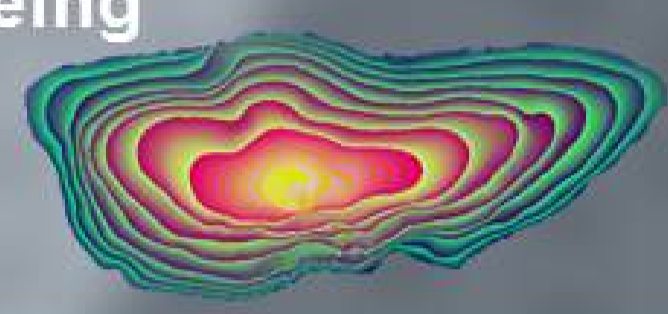


8578 sampled Fishes

From surveys and fishing markets
Period: 2010 - 2017
Only Quarters 3 & 4
Age: from 0 to 8+ year classes

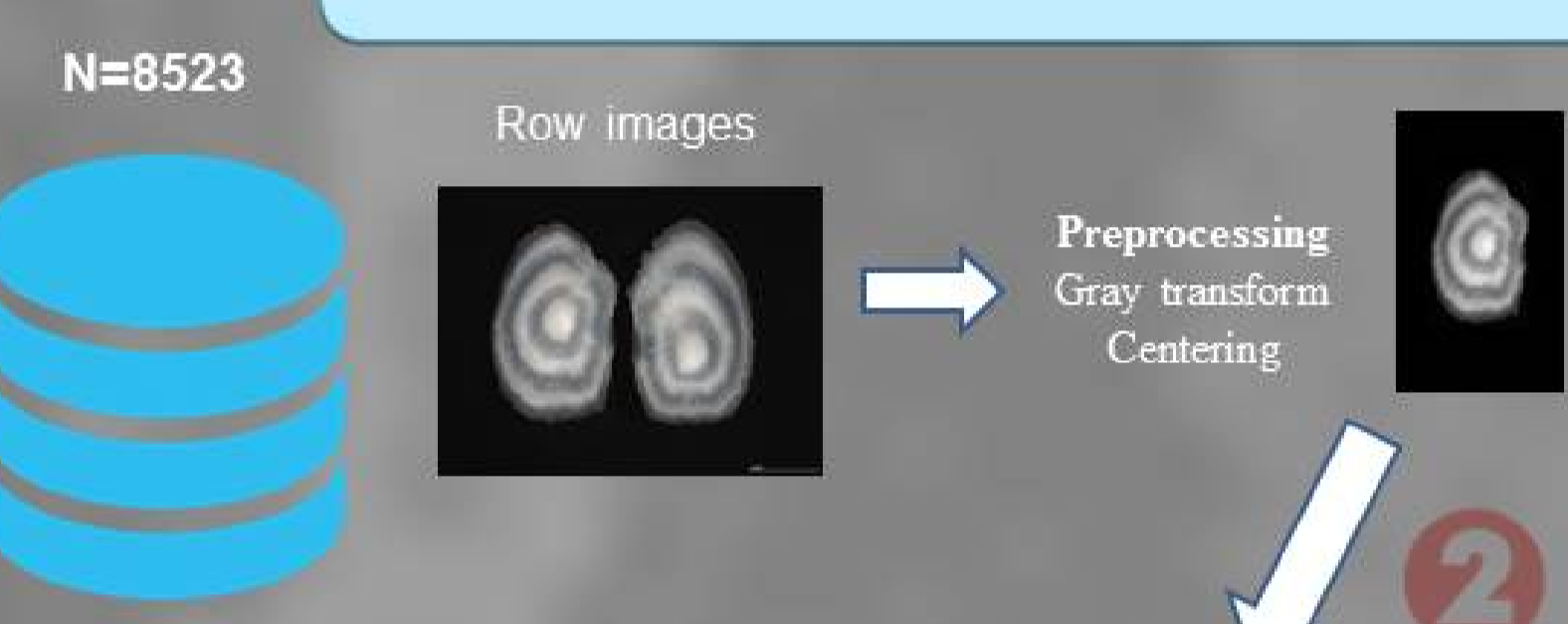


Automated FISH Ageing EU project (AFISA)
Mahé, 2009

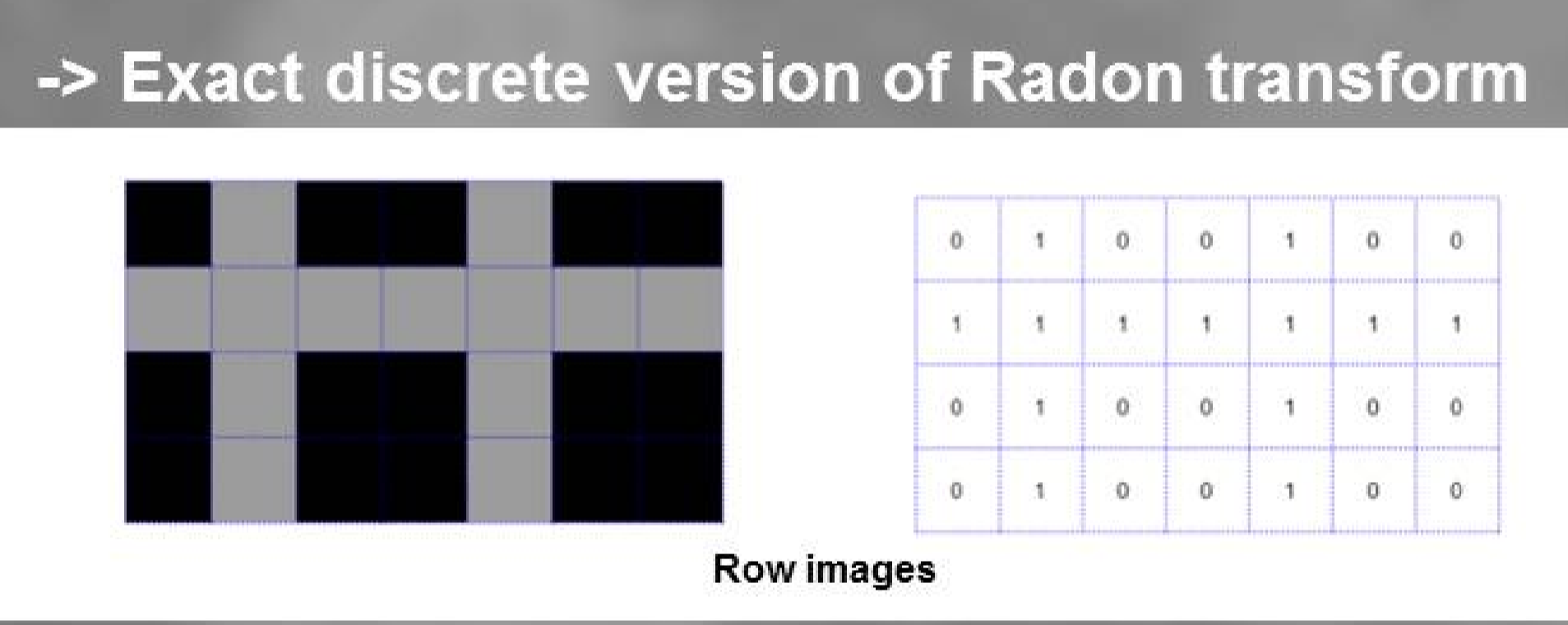


Automation of age reading ?

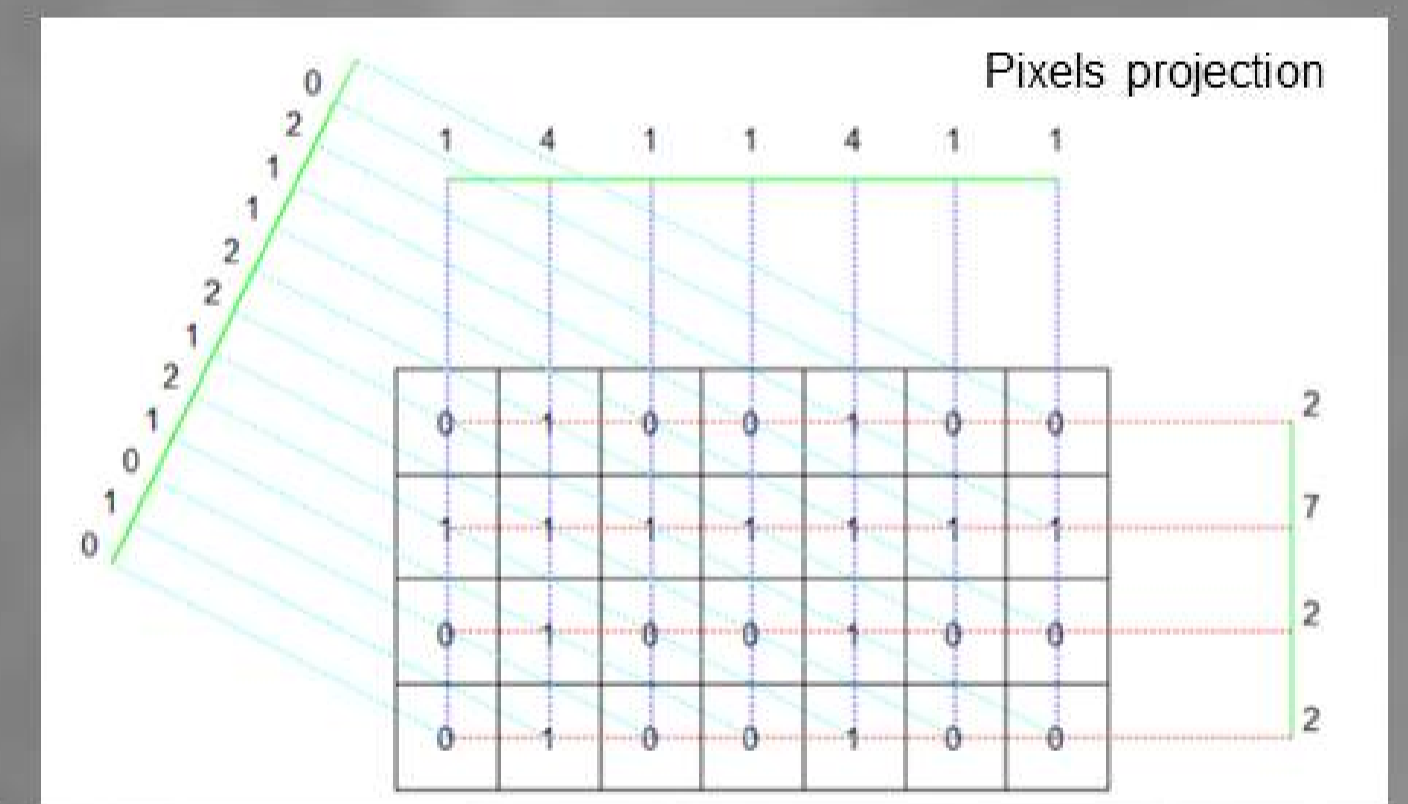
Age estimation by Machine



Mojette Transform



-> To reduce the information by recording only few projections that allow to reconstruct original image

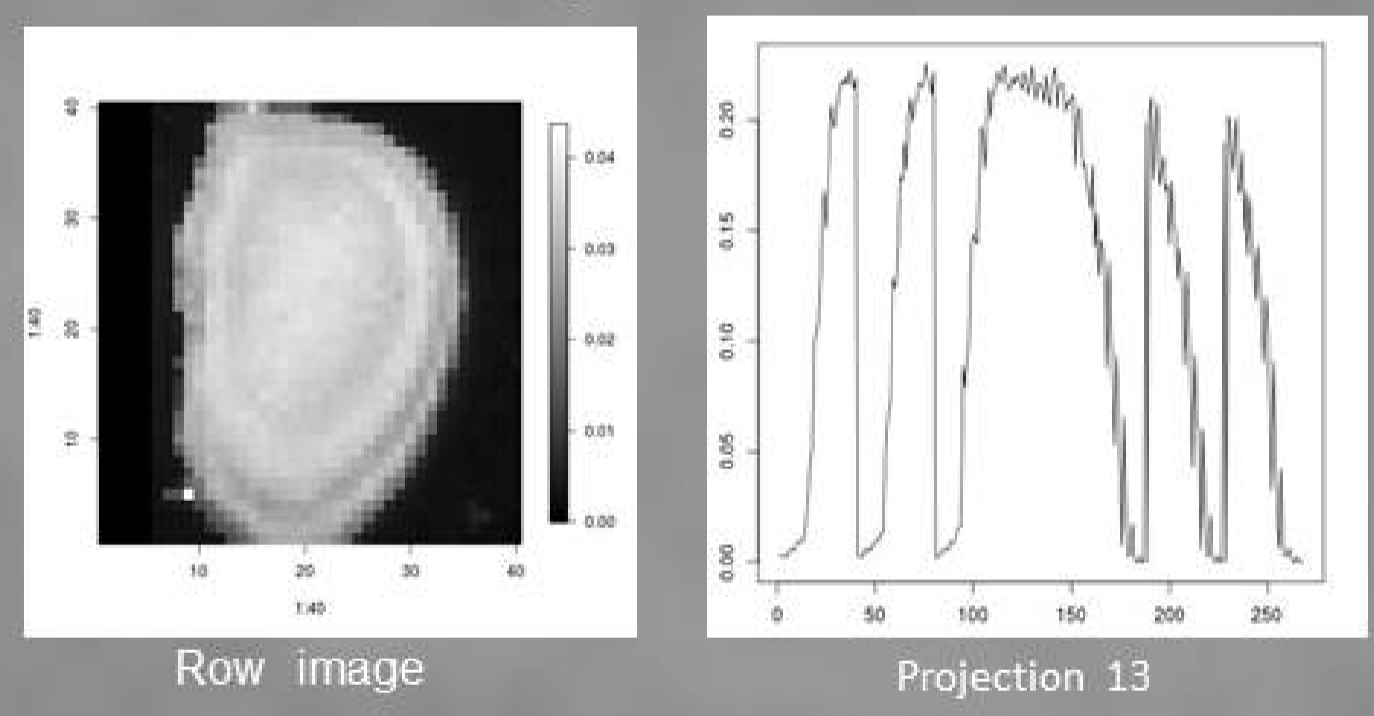


Each projection depends of its angle defined by (p,q) and it consists in a set of bin value M with b the index number of this bin.
All angles are built according to the Farey suite with n order and symmetry
A bin value M is the sum of pixel value f(i,j) according the direction characterized by (p,q).

Katz Criterion to select the number of projection :
sum(p) > P width or sum(q) > Q height.

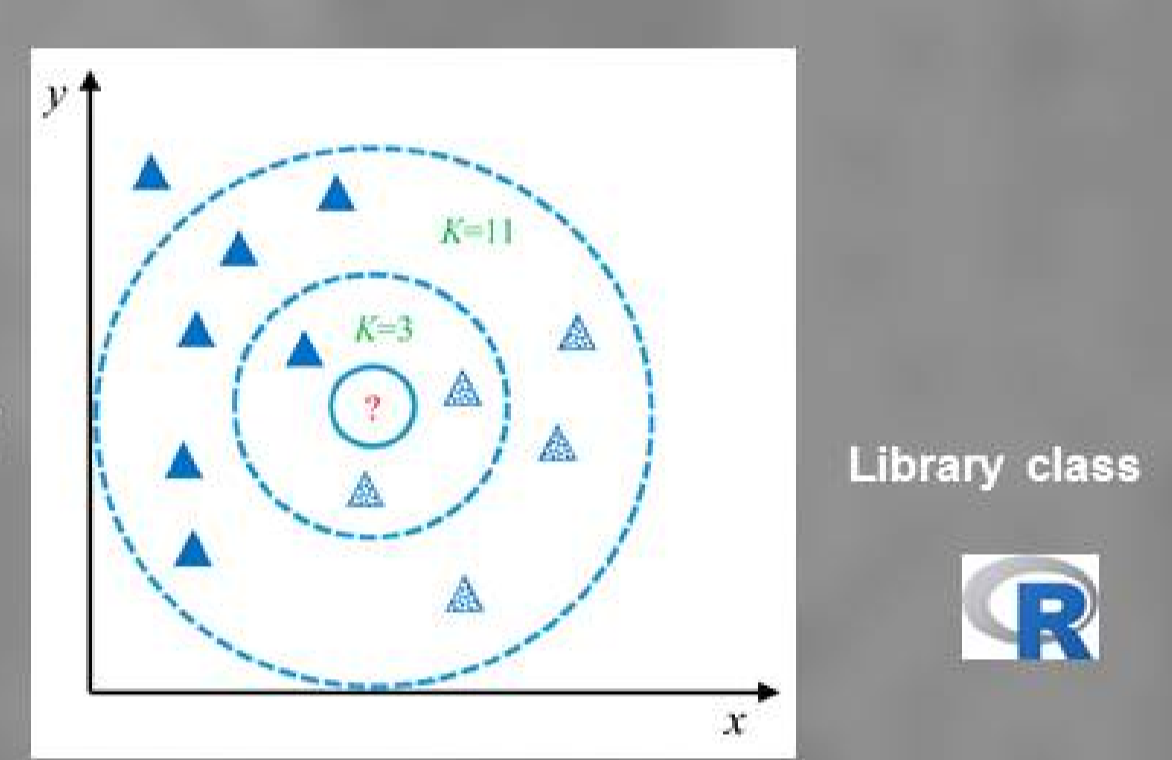
$$M(b,p,q) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} f(i,j) \times \Delta(b-p \times i + q \times j).$$

Farey order = 5

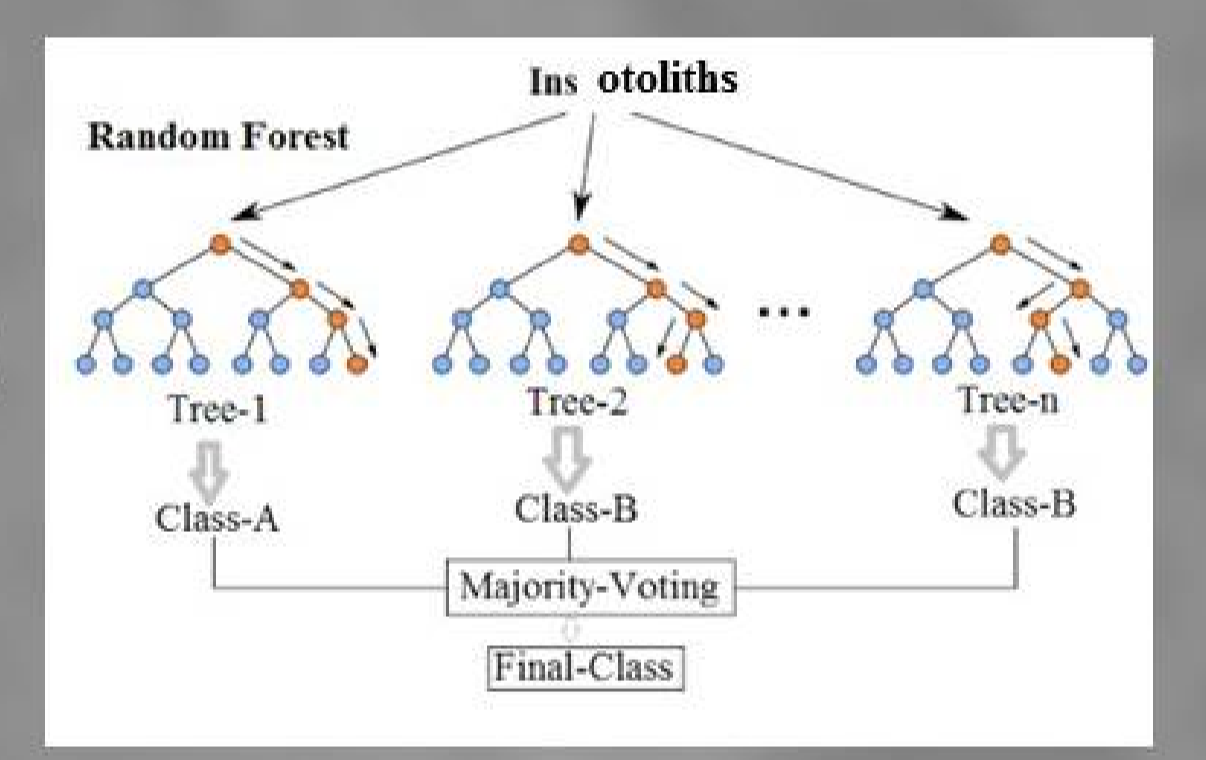


Machine Learning : 3 classifiers

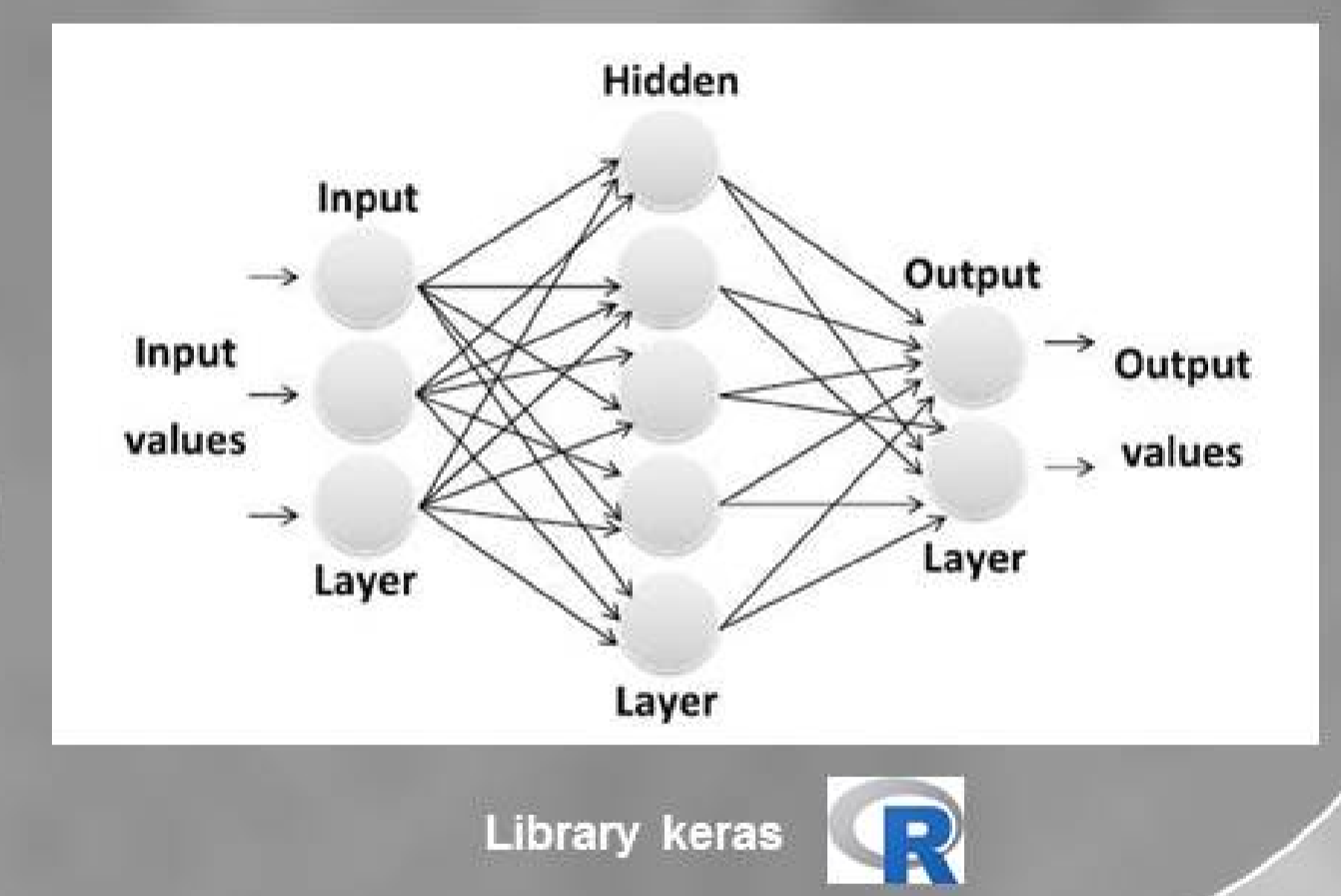
k-NN
k-Nearest Neighbors



Random Forest



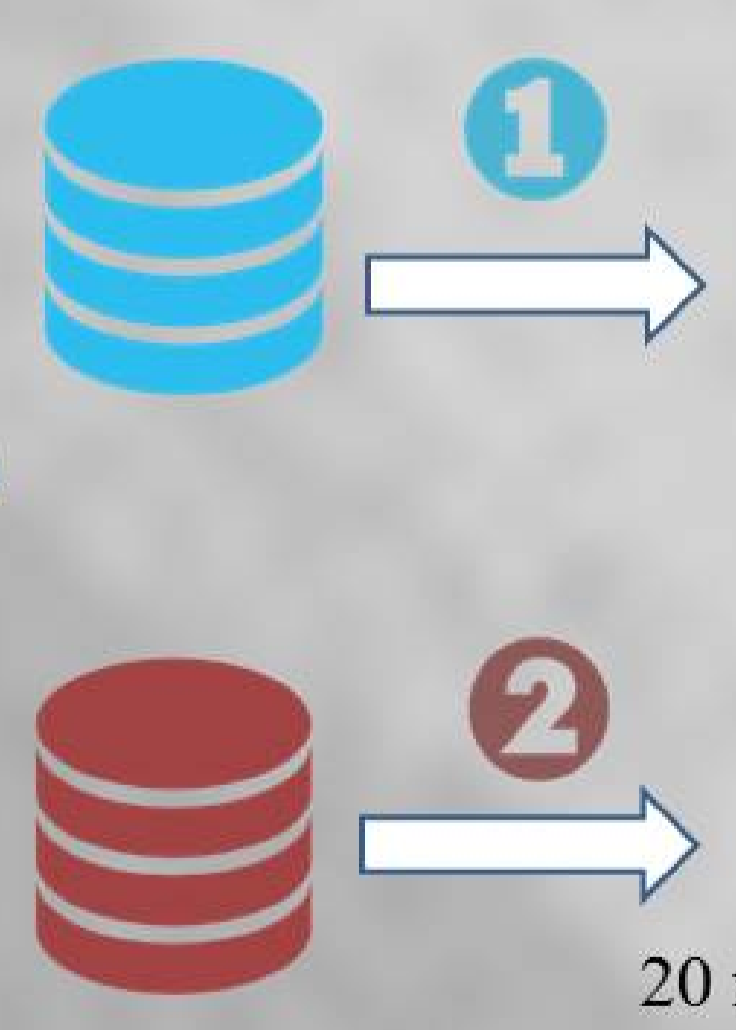
MLP
Multi-Layer Perceptron



Results

Human Expert

Age class	Training set	Test set
Age 0	116	49
Age 1	843	406
Age 2	1397	685
Age 3	1468	782
Age 4	1134	523
Age 5	419	237
Age 6	180	92
Age 7	130	54
Age 8+	71	37



Machine Learning

- k-nn : k=1/5/10
- Random Forest : mtry=16 ntree=50/100/500
- MLP : linear/1hidden layer 20 neurons/sigmoid function output :softmax activation 9 classes

Projections 12,13,15 & 16
30 first Projections

		k-NN			Random Forest			MLP	
		1-nn	5-nn	10-nn	RF-50	RF-100	RF-500	Linear-MLP	20-MLP
Database	Precision	98,5	64,3	57,9	98,4	98,5	98,5	46,1	51,7
Training	± 0 year	99,5	88,6	87,1	99,5	99,5	99,5	84,3	87,8
Training	± 1 year	42,8	47,3	48,0	51,6	51,3	52,4	42,5	45,7
Test	± 0 year	79,6	84,2	85,2	88,7	88,9	89,2	83,0	85,5
Test	± 1 year	42,8	45,7	45,9	49,6	50,8	51,3	40,6	41,5
Database	Precision	98,5	61,6	55,7	98,5	98,5	98,5	45,8	43,5
Training	± 0 year	99,5	86,1	85,5	99,5	99,5	99,5	82,3	83,8
Training	± 1 year	42,8	45,7	45,9	49,6	50,8	51,3	40,6	41,5
Test	± 0 year	78,8	82,8	84,4	87,4	87,7	88,2	82,1	83,4
Test	± 1 year	43,7	45,5	47,6	50,7	51,7	51,9	-	44,6
Database	Precision	98,5	63,5	57,1	98,4	98,5	98,5	-	47,7
Training	± 0 year	99,5	88,6	87,3	99,5	99,5	99,5	-	87,4
Training	± 1 year	43,7	45,5	47,6	50,7	51,7	51,9	-	44,6
Test	± 0 year	81,4	84,3	85,6	88,3	88,6	88,9	-	85,9
Test	± 1 year	-	-	-	-	-	-	-	85,9

30 first projections
Random Forest Classifier
Test dataset : 2865 otoliths

51,9% of Correct Classification

Predicted age (year)

Age group	Age 0	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8+
Age 0	22	23	4	0	0	0	0	0	0
Age 1	4	303	84	9	6	0	0	0	0
Age 2	2	90	408	159	33	1	0	0	0
Age 3	3	24	169	439	145	2	0	0	0
Age 4	0	7	38	187	277	13	1	0	0
Age 5	0	1	9	33	152	35	2	5	0
Age 6	0	0	1	13	54	10	5	5	0
Age 7	0	1	0	8	29	4	11	1	0
Age 8+	0	0	0	1	7	16	3	7	0

Legend: Red = ± 0 year, Blue = ± 1 year

Contacts

Emilie Poisson Caillault
LISIC, ULCO EA 4491, Université du Littoral Côte d'Opale, Calais, France
Email: poisson@univ-littoral.fr

Kélig Mahé
Ifremer, Fisheries Laboratory,
Boulogne sur mer, France
Email: kelig.mahé@ifremer.fr

Conclusion & Prospects

- Mojette transform results are very close to raw image analysis.
- k-sensibility in k-nn were correlated with the high intra-class variability and close inter-class size.
- Database size is not sufficient to build an efficient neuronal network (MLP).
- Random Forest seemed to be the best classifier according to raw image or Mojette bins.
- The database was built from all otolith images (n=8578) used for stock assessment without prior filtering or images cleaning -> image quality (broken otoliths, dirty otoliths) impacts the results and must be evaluated.
- These results could be improved by optimizing machine learning parameters and by selecting discriminant projections.

References

Mahé, K. (2009). Project no. 044132 Automated FISH Ageing (AFISA): Final Activity 831 Report. <http://cordis.europa.eu/documents/documentlibrary/124722831EN6.pdf>
 Breiman, L., Ghahramani, Z. (2004). Consistency for a simple model of random forests. Statistical Department, University of California at Berkeley. Technical Report (670).
<https://cran.r-project.org/web/packages/randomForest/index.html>
 Cuvier, J. (2009). The Mojette transform. Theory and applications. ISTE-WILEY.



Financial Support



Experiments presented in this study were carried out using the CALCULCO computing platform, supported by SCOS/ULCO (Service Commun du Système d'Information de l'Université du Littoral Côte d'Opale)