

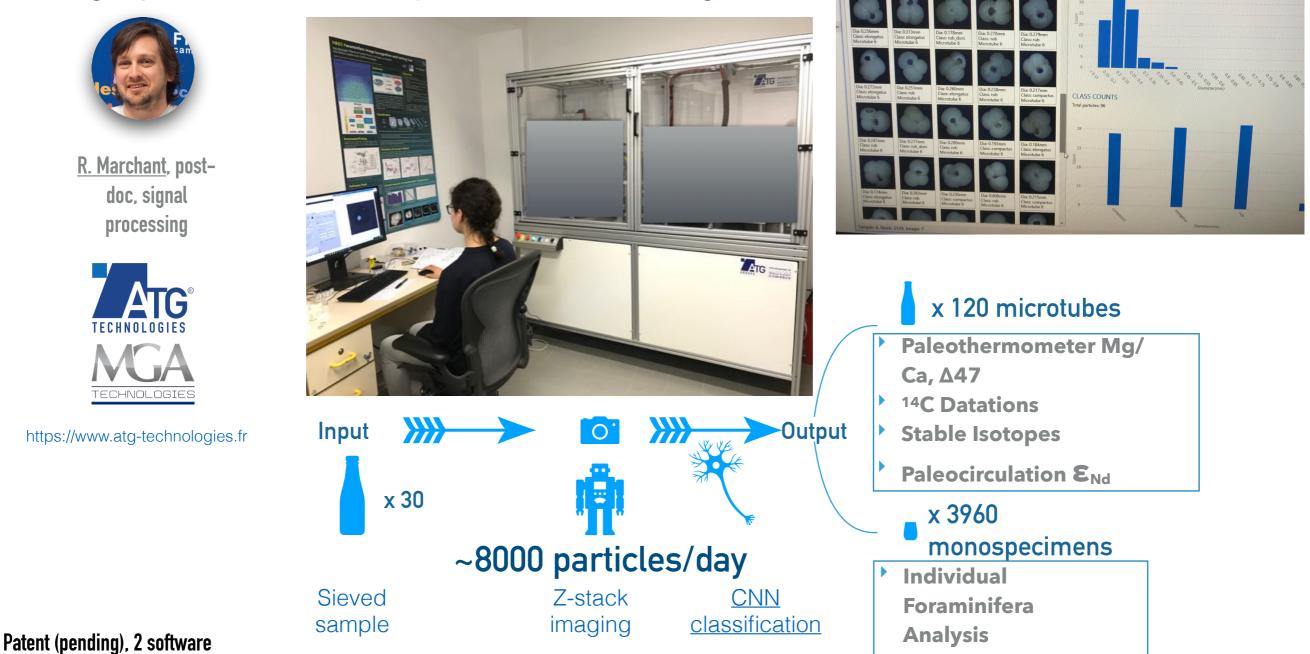
# AUTOMATED RECOGNITION AND PICKING OF FORAMINIFERA USING THE MISO (MICROFOSSIL SORTER) PROTOTYPE

THIBAULT DE GARIDEL-THORON, ROSS MARCHANT\*, MARTIN TETARD, MICHAEL ADEBAYO, AND YVES GALLY | CEREGE, AIX EN PROVENCE

#### @TdeGaridel

### MISO – MICROFOSSIL SORTER

In collaboration with ATG Technologies, we designed and built a fully automated machine able to image and pick single particles in the 100  $\mu$ m to 1mm size range.



Particles are recognized using a <u>Convolutional Neural Network classifier</u> and separated in microtubes or micro slides. Machine is cleaned between each sample.

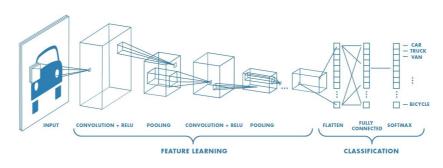
#### Control software (MiSo)

DIAMETER HISTOGRAM

## **CONVOLUTIONAL NEURAL NETWORKS : A FORAM CNN**

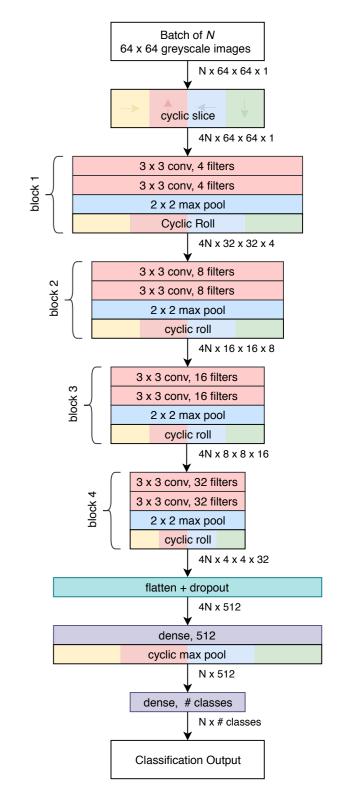
« In the area of computer vision (...) deep artificial neural networks have reached superhuman capabilities on a wide range of visual recognition problems »

#### Serre, Ann. Rev. Vision Sci., 2019



- MiSO CNN topology adjusts to input image dimensions
  - Uses cyclic layers for rotation invariance
  - Trained with augmentations common to microscope images

We tested a variety of CNN setups and are now using ResNet and BaseCyclic as in the preprint linked. The steps from the labeling to the training of the CNN on foraminifera images is achieved on a dedicated user friendly software ParticleTrieur.



Read the preprint of Marchant et al., with the description of the tests for the CNN here

#### MiSo CNN

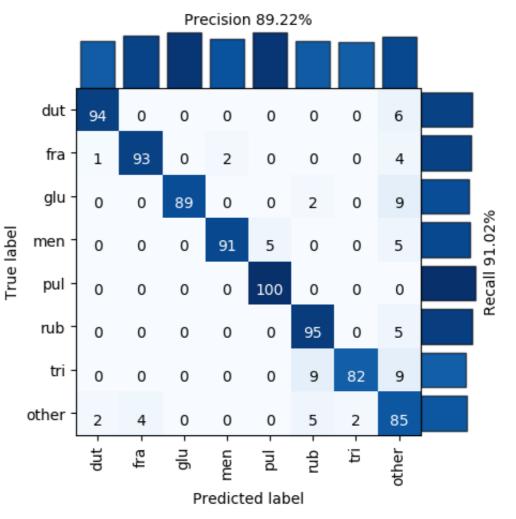
### OUR WORKFLOW

(1) Scan representative samplesfrom the sediment core using MiSo (n>10,000 to 20,000 images)

(2) Label a training database using ParticleTrieur

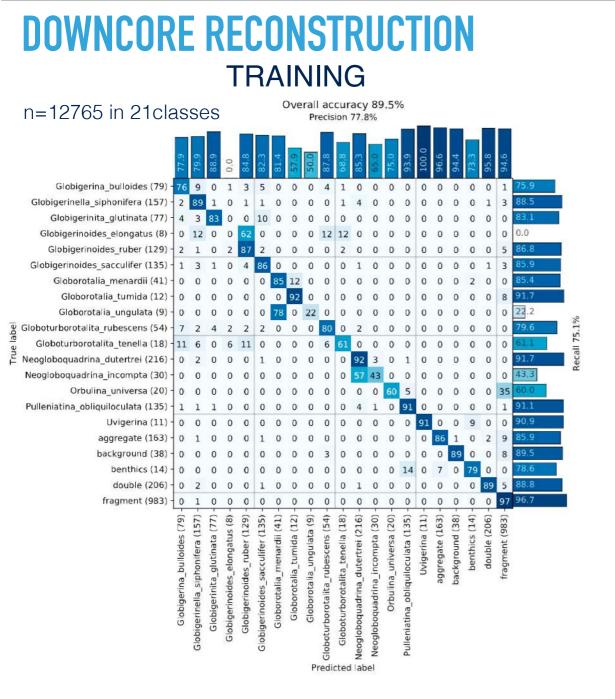
New Oper	n Save Add	- L Remove Edit	k-NN CNN E	t ∿ xport Train	All 27729 vectors calculated	
Filter list			»	Classification 12 Similarity ## Processing		
# Ima	ige Sample	Label/Tag	Filename	Image #9415	Labels Auto-advance 🥥	
410	unknown [1] 0.0 [2] 0.0	Globigerinoi auto	09410-Globigerinoid es_ruber.jpg	Globigerinoides_ruber	Beella, digitata Candeina, nitida Globigerina, bulloides Globigerina falconensis Globigerinella, adamsi   Globigerinella, calida Globigerinella, siphonifera Globigerinita, glutinata Globigerinita, suvula	
411	unknown [1] 0.0 [2] 0.0	Globigerinoi auto	09411-Globigerinoid es_ruber.jpg		Globigerinoides_conglobatus Globigerinoides_elongatus Globigerinoides_nuber Globigerinoides_acculifer Globoquadrina_conglomerata Globorotalia_crassaformis Globorotalia_hirsuta Globorotalia_inflata	
412	unknown [1] 0.0 [2] 0.0	Globigerinoi auto	09412-Globigerinoid es_ruber.jpg		Globorotalia_menardii Globorotalia_schula Globorotalia_truncatulinoides Globorotalia_tumida Globorotalia_ungulata Globorotalides_hexagonus Globoturborotalita_rubescens Globoturborotalita_neella	
413	unknown [1] 0.0 [2] 0.0	Globigerinoi auto	09413-Globigerinoid es_ruber.jpg		Hastigerina_pelagica Neogloboquadrina_dutertrei Neogloboquadrina_incompta   Neogloboquadrina_pachyderma Orbulina_universa Pulleniatina_obliquiloculata Sphaeroidinella_dehiscens	
14	unknown [1] 0.0 [2] 0.0	Globigerinoi auto	09414-Globigerinoid es_ruber.jpg	Label: Globigerinoides_ruber Tags: auto Sample: unknown Filenam: 09415-Globigerinoides_ruber.jpg Path: Cullers?vossm/Documents/Data/foraminifera/EndlessFo rams/border_removed/endless_forams_20190914_16534 3VGlobigerinoides_ruber Info: 43 4 x434 pixels GUID: 75jsOVIgoMMCK/L62bW9Mn	Tenuitella_jota Turborotalita_humilis Turborotalita_quinqueloba + Add	
115	unknown [1] 0.0 [2] 0.0	Globigerinoi auto	09415-Globigerinoid es_ruber.jpg		Tags duplicate auto + Add	
116	unknown [1] 0.0 [2] 0.0	Globigerinoi auto	09416-Globigerinoid es_ruber.jpg		BB k-NN Vector Globigerinoides ruber 월 Predict	
117	unknown [1] 0.0 [2] 0.0	Globigerinoi auto	09417-Globigerinoid es_ruber.jpg		Score: Threshold: 0.8 + Deep CNN Labet Globigerinoides_ruber @ Predict Pre-process	
418	unknown [1] 0.0 [2] 0.0	Globigerinoi auto	09418-Globigerinoid es_ruber.jpg			
119	unknown [1] 0.0 [2] 0.0	Globigerinoi auto	09419-Globigerinoid es_ruber.jpg		Score: Threshold: 0.8 -	
420	unknown [1] 0.0	Globigerinoi	09420-Globigerinoid es ruber.ipa			

#### (3) Train and evaluate a CNN classifier



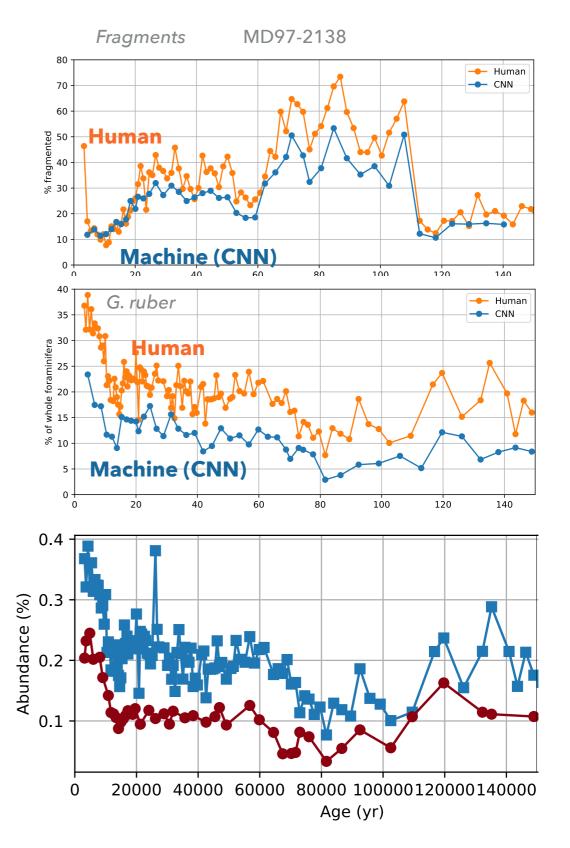
- Example of a simple training set (picking oriented) : each class includes different subclasses with different orientations
  - Typical precision & recall ~90±5%
  - Possible to cross-check later with different taxonomical frameworks
- Training set should be close to the working set
- Image pre-processing have to be very rigorously similar

For picking, when the number of false positives is more critical than false negatives, simple training as shown can learn to pick the most abundant classes with more >90% precision and recall



In the core MD97-2138, from the Western Pacific Warm Pool, we compare results from human counting with a CNN trained on this core. Fragment counts and the main species are close to the human counts. Yet, artefacts as clays infillings, dissolution artefacts can cause some misclassifications as in marine isotopic stage 5.

#### COMPARISON WITH HUMAN COUNTING



### CONCLUSIONS

- Automatic imaging & recognition performed 24/7 routinely at CEREGE using CNNs
- Dedicated workflow for foraminifera from image acquisition to specimen handling
- Image preprocessing critical for image recognition
- Software packages (x-platform) userfriendly
- Ongoing developments : depthreconstruction, size variations,
- Biometrics

- Very sensitive to acquisition method
  - Microscope lighting
  - Image resolution
- Data driven
  - Training set must span all variations
- Unknown forams can be incorrectly classified with high probability
- Size information is lost
- Misclassification causes signal offset and reduced range

Demo on request ! And if you're interested in crunching millions of images of sediments particles, get in touch.

What's next ? : PhD Michael Adebayo : Indian Ocean paleoceanography based on MiSo