ImgTLClass R-package

Version 1.0-0

MANUEL UTILISATEUR

G. WACQUET

IFREMER
LABORATOIRE ENVIRONNEMENT RESSOURCES
CENTRE MANCHE MER DU NORD

TABLE DES MATIERES

INTRODUCTION	5
INSTALLATION ET EXECUTION	
UTILISATION DE L'INTERFACE GRAPHIQUE	
Bouton DATA SELECTION	
Bouton SETTINGS	
Bouton CLASSIFY	
Bouton VIEW	
Bouton MORE (+)	
DUULUH IVIUNE ITI	

INTRODUCTION

TO DO!

INSTALLATION FT EXECUTION

La version 1.0-0 de ImgTLClass nécessite une version récente de R (version 4.0.x ou plus). Elle peut être téléchargée directement sur le site du CRAN (http://cran.r-project.org).

En double-cliquant sur l'icône R sur le bureau, ou en sélectionnant R dans le menu de démarrage, une fenêtre apparaît à l'écran : il s'agit de la console R. Cette dernière permet de contrôler R directement par lignes de commande. Elle permet également d'afficher les principaux résultats et messages des actions effectuées avec ImgTLClass.

Les packages nécessaires à ImgTLClass :

- colorRamps
- ggplot2
- grid
- jpeg
- mapplots
- maps
- randomForest
- reticulate
- SDMTools
- shapefiles
- stringr
- svDialogs
- svMisc
- tcltk2
- tiff
- zooimage

peuvent être installés directement à partir de la console R, en tapant :

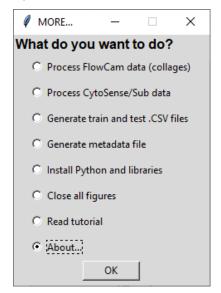
install.packages(c("colorRamps", "ggplot2", "grid", "jpeg", "mapplots", "maps", "randomForest", "reticulate", "SDMTools", "shapefiles", "stringr", "svDialogs", "svMisc", "tcltk2", "tiff", "zooimage"))

Choisir ensuite un miroir (par défaut : 0-cloud) pour démarrer les téléchargements et les installations.

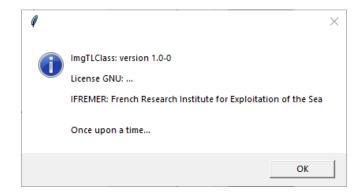
Une fois l'installation des packages terminée, il est possible de s'assurer du bon déroulement des étapes précédentes en vérifiant que la version installée est bien 1.0-0. Pour cela, dans un premier temps, taper dans la console R: require(ImgTLClass), pour charger le package, puis : ImgTLClass(), pour lancer l'Interface Graphique Utilisateur (GUI) :



Cliquer sur le bouton + (en bas, à droite). Une nouvelle fenêtre apparaît :



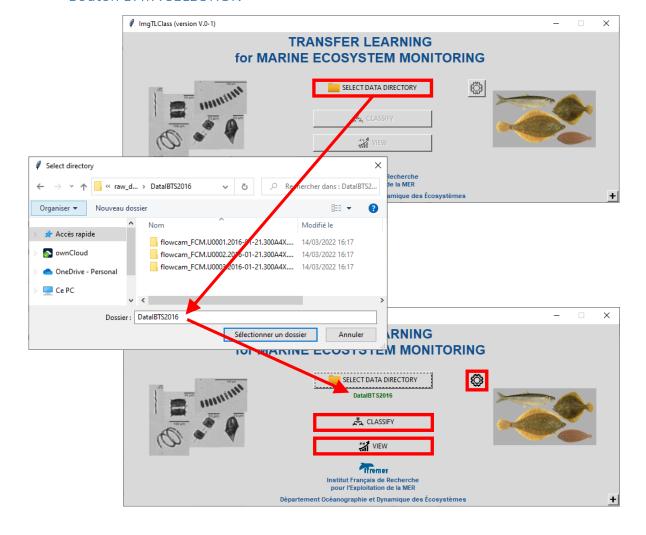
En sélectionnant "About...", une boîte de dialogue apparaît et informe l'utilisateur de la version de ImgTLClass en cours d'exécution.



UTILISATION DE L'INTERFACE GRAPHIQUE

Pour l'utilisation de ImgTLClass en routine, une interface graphique utilisateur ergonomique est disponible.

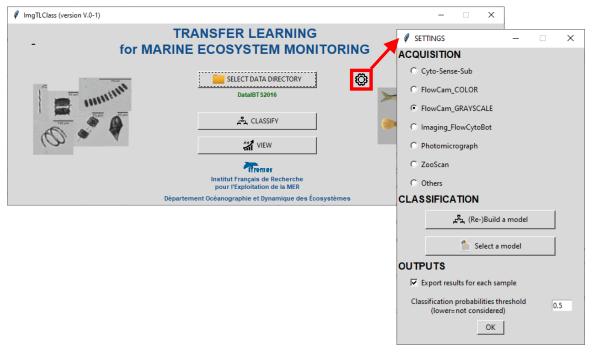
Bouton DATA SELECTION



Cliquer sur le bouton **SELECT DATA DIRECTORY**, sélectionner le répertoire contenant les images à classer automatiquement, puis valider en cliquant sur **OK** : le nom du répertoire sélectionné est alors affiché en dessous du bouton de sélection, et les boutons **SETTINGS**, **CLASSIFY** et **VIEW** deviennent alors actifs.

Bouton SETTINGS

Après sélection des données d'entrée, la fenêtre de paramétrage est automatiquement affichée (sans cliquer sur le bouton **SETTINGS**). Il est toutefois possible de réafficher cette fenêtre à tout moment en cliquant sur ce bouton.



Paramètres d'ACQUISITION

Choisir le type d'instrument utilisé pour l'acquisition des images.

<u>Remarque</u>: pour le FlowCam_COLOR et le FlowCam_GRAYSCALE, il est possible de mettre directement les données brutes en sortie des instruments (collages). Le package permet de créer directement les vignettes (1 image par particule) à partir des collages et du fichier lst.

Paramètres de CLASSIFICATION

(Re-)Build a model

Ce bouton permet de construire (ou reconstruire) un modèle de classification par Transfer Learning grâce à l'utilisation d'un jeu de données comprenant un training set (répertoire nommé **train**) pour l'apprentissage, et un test set (répertoire nommé **test**) pour la validation et l'évaluation.



Train/build a new model SELECT TRAIN/TEST DIRECTORIES SELECT THE MODEL TO BUILD: DenseNet201 **EPOCHS** 20 InceptionV3 BATCH SIZE 20 ResNet50 VGG16 IMAGE WIDTH (px) 224 VGG19 IMAGE HEIGHT (px) 224 4 I١ ▼ Use Data Augmentation Select direct Build on this computer Rechercher dans : data_flow. → ↑ « traini... > data_flowcam ∨ ∂ C Generate a .PY file Organiser • eau dossier == -BUILD THE MODEL Nom Modifié le Accès rapide test 14/03/2022 16:19 ownCloud 14/03/2022 16:20 train Train/build a new model OneDrive - Persona Ce PC SELECT TRAIN/TEST DIRECTORIES r: data_flowcam data flowcam Sélectionner un dossier Annuler SELECT THE MODEL TO BUILD: **EPOCHS** 20 InceptionV3 BATCH SIZE 20 ResNet50 VGG16 IMAGE WIDTH (px) 224 VGG19 IMAGE HEIGHT (px) 224 1 ✓ Use Data Augmentation Build on this computer C Generate a .PY file BUILD THE MODEL

En cliquant sur ce bouton, une nouvelle fenêtre apparaît.

Cliquer sur le bouton **SELECT TRAIN/TEST DIRECTORIES**, sélectionner le répertoire contenant les deux sous-dossiers **train** et **test**, puis valider en cliquant sur **OK** : le nom du répertoire sélectionné est alors affiché en dessous du bouton de sélection.

Il est ensuite possible de sélectionner différentes architectures de Réseaux de Neurones à Convolution (CNN pour Convolutional Neural Network) dans la liste **SELECT THE MODEL TO BUILD**, parmi :

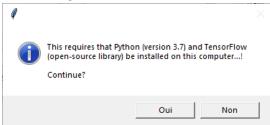
- DenseNet201
- InceptionV3
- ResNet50
- VGG16
- VGG19

et régler les paramètres liés aux images et à la phase d'apprentissage, en fixant des valeurs à **EPOCH** (par défaut=20), **BATCH SIZE** (par défaut=20), **IMAGE WIDTH** (par défaut=224) et **IMAGE WEIGHT** (par défaut=224), mais également choisir la possibilité d'utiliser la technique d'augmentation de données (**Use Data Augmentation**). Dans le cas d'un training set peu fourni en images, cette option permet de générer automatiquement des images complémentaires à partir des images du training set, en leur appliquant des transformations géométriques comme des rotations (par défaut, rotation_range=45) et des retournements horizontaux et verticaux (par défaut, horizontal_flip=True et vertical_flip=True).

La dernière étape consiste à choisir le matériel sur lequel construire et adapter le modèle de classification. Selon l'option sélectionnée, et après avoir cliqué sur le bouton **BUILD THE MODEL** :

- Build on this computer

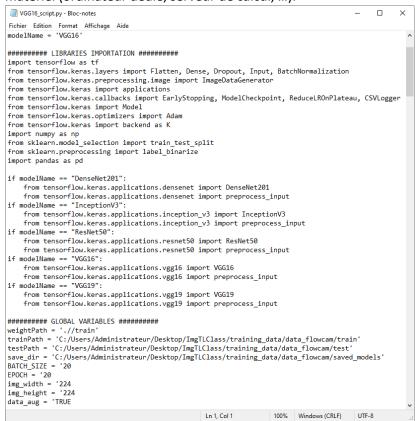
Une boîte de dialogue s'ouvre alors :



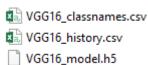
<u>Attention</u>: le temps d'apprentissage peut être long (plusieurs heures) selon le nombre d'images du training set et selon les paramètres définis lors de l'étape précédente.

- Generate a .PY file

Un script est créé automatiquement, et peut être exécuté sur un autre matériel (ordinateur dédié, serveur de calcul, ...).



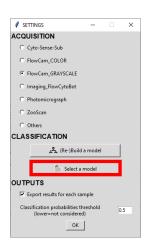
A la fin de l'exécution du script, un fichier .ZIP est créé. Ce dernier contient toutes les informations nécessaires pour la classification d'un nouveau jeu d'images.

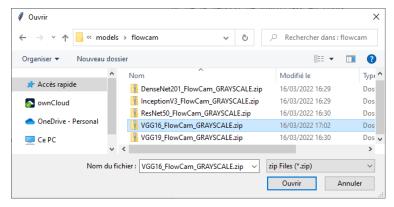


modelName_classNames.csv contient les noms des classes du training set. modelName_history.csv contient les scores d'évaluation du modèle créé. modelName_model.h5 est le modèle créé sous Python.

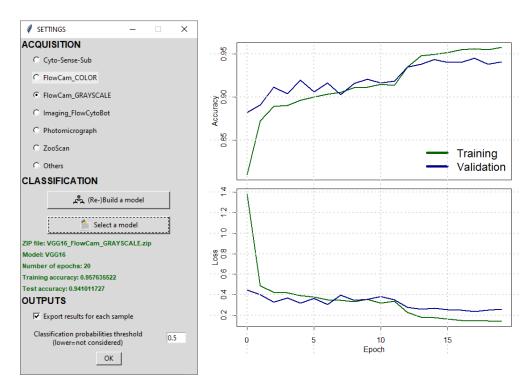
Select a model

Ce bouton permet de choisir un modèle pour la classification automatisée des images contenues dans le répertoire sélectionné lors de la première étape (cf. section « Sélection des données d'entrée »).





Sélectionner le fichier ZIP généré lors de l'étape précédente, puis valider en cliquant sur **Ouvrir**: des informations sur les performances du modèle (calculées lors des phases d'apprentissage et de validation) sont alors affichées en dessous du bouton de sélection, ainsi que les courbes d'**Accuracy** (≈ pourcentage de données correctement classifiées) et de **Loss** (≈ distance entre les données réelles et les données prédites).



❖ Paramètres d'OUTPUTS

Export results for each sample

Cette option permet d'exporter les résultats pour chaque sous-dossiers du répertoire sélectionné lors de la première étape (cf. section « Sélection des données d'entrée »).

Pour chaque échantillon traité, trois fichiers CSV sont alors créés :

- sampleName.csv

Sample	Group	Relative	Count	Date	Volume	Prob_threshold	Percent_used
flowcam_FCM.U0003.2016-01-21.300A4X.01	Asterionellopsis	0	0		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	C_curvisetus	8.474576271	5		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	C_danicus	15.25423729	9		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	C_socialis	1.694915254	1		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	Ciliophora	8.474576271	5		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	Dactyliosolen	3.389830508	2		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	Dytilum	6.779661017	4		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	G_flaccida	1.694915254	1		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	G_striata	1.694915254	1		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	Gymnodinium	3.389830508	2		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	Lauderia	1.694915254	1		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	Leptocylindrus	1.694915254	1		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	Odontella	3.389830508	2		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	P_globosa	0	0		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	Pleuro_Gyrosigma	3.389830508	2		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	Prorocentrum	11.86440678	7		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	PseudoNitzschia	3.389830508	2		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	Rhizosolenia	16.94915254	10		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	Thalassionema	3.389830508	2		1000	0.75	92.15500945
flowcam_FCM.U0003.2016-01-21.300A4X.01	Thalassiosira	3.389830508	2		1000	0.75	92.15500945

sampleName_CLASSIF.csv

. –		
Filename		Class
unknown\flowcam_FCN	И.U0003.2016-01-21.300A4X.01_1.jpg	dark
unknown\flowcam_FCN	И.U0003.2016-01-21.300A4X.01_10.jpg	dark
unknown\flowcam_FCN	M.U0003.2016-01-21.300A4X.01_100.jpg	dark
unknown\flowcam_FCN	N.U0003.2016-01-21.300A4X.01_1000.jpg	dark
unknown\flowcam_FCN	И.U0003.2016-01-21.300A4X.01_1001.jpg	dark
unknown\flowcam_FCN	N.U0003.2016-01-21.300A4X.01_1002.jpg	dark
unknown\flowcam_FCN	M.U0003.2016-01-21.300A4X.01_1003.jpg	dark
unknown\flowcam_FCN	M.U0003.2016-01-21.300A4X.01_1004.jpg	dark
unknown\flowcam_FCN	M.U0003.2016-01-21.300A4X.01_1005.jpg	dark
unknown\flowcam_FCN	M.U0003.2016-01-21.300A4X.01_1006.jpg	Ciliophora
unknown\flowcam_FCN	M.U0003.2016-01-21.300A4X.01_1007.jpg	Odontella
unknown\flowcam_FCN	M.U0003.2016-01-21.300A4X.01_1008.jpg	dark
unknown\flowcam_FCN	M.U0003.2016-01-21.300A4X.01_1009.jpg	fiber

sampleName_PRED.csv

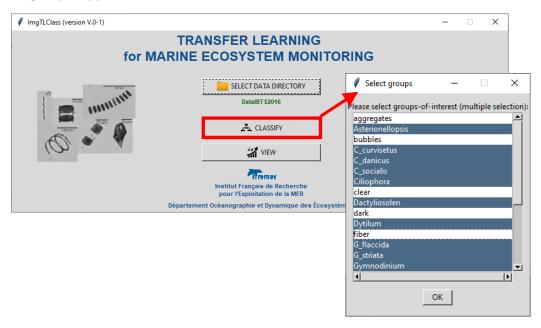
Asterionellopsis	C_curvisetus	C_danicus	C_socialis	Ciliophora	Dactyliosolen	Dytilum	G_flaccida
3.52E-09	1.24E-08	4.22E-10	3.65E-11	0.000285448	1.65E-06	9.95E-09	7.16E-10
1.34E-11	2.78E-11	7.02E-14	1.16E-16	4.84E-10	5.95E-14	3.10E-13	3.12E-16
1.86E-12	2.97E-09	1.78E-13	8.65E-14	8.47E-05	1.23E-07	4.80E-09	1.05E-11
9.69E-17	1.54E-11	3.29E-18	7.69E-20	3.97E-08	1.30E-10	1.50E-12	2.34E-17
6.27E-08	5.92E-07	3.55E-09	2.57E-09	7.62E-06	1.90E-06	1.73E-07	1.10E-07
1.42E-08	1.77E-08	5.07E-08	7.74E-11	1.00E-05	1.52E-07	1.58E-08	1.10E-11
1.01E-24	1.02E-15	9.25E-27	4.89E-27	3.53E-14	3.10E-17	2.96E-18	2.73E-22
3.96E-15	1.95E-12	1.79E-16	2.53E-18	6.86E-06	1.27E-08	2.07E-13	3.37E-14
6.43E-15	2.69E-09	8.91E-14	5.48E-17	0.001769048	2.66E-08	1.20E-12	4.52E-14

Classification probability threshold

La valeur de seuil définie à cette étape permet de ne prendre en compte que les particules ayant une probabilité de « bonne » classification supérieure à ce seuil. Pour prendre en considération la totalité des particules, cette valeur doit être fixée à 0.

Bouton CLASSIFY

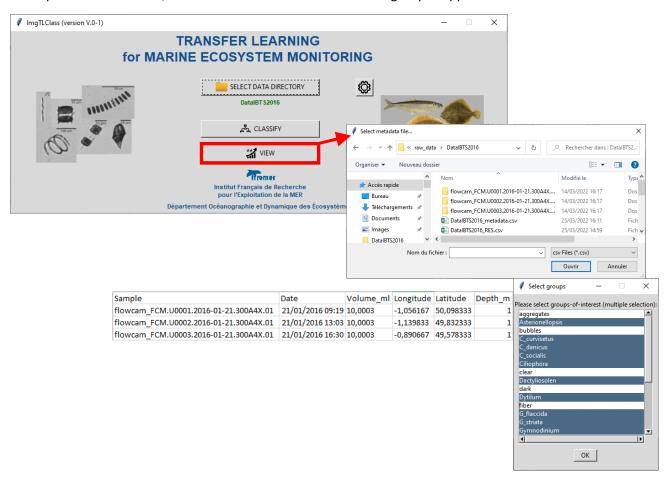
Pour classer de nouvelles images, cliquer sur le bouton **CLASSIFY**. Une nouvelle fenêtre de sélection de groupes apparaît.

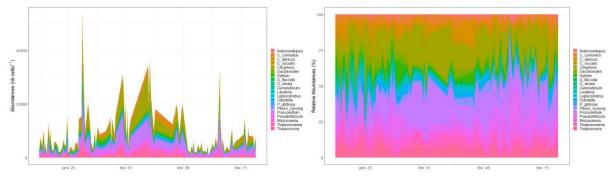


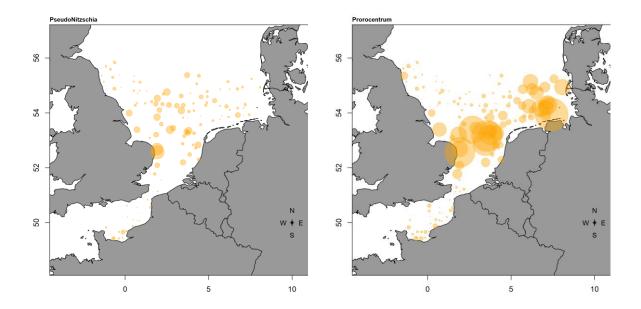
Sample	Group		Count	Prob_threshold	
flowcam_FCM.U0001.2016-01-21.300A4X.01	Asterionellopsis	2.941176471	1	0.75	89.88476312
flowcam_FCM.U0001.2016-01-21.300A4X.01	C_curvisetus	5.882352941	2	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	C_danious	8.823529412	3	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	C_socialis	0	0	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	Ciliophora	5.882352941	2	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	Dactyliosolen	8.823529412	3	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	Dytilum	8.823529412	3	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	G_flaccida	8.823529412	3	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	G_striata	5.882352941	2	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	Gymnodinium	2.941176471	1	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	Lauderia	0	0	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	Leptocylindrus	0	0	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	Odontella	0	0	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	P_globosa	0	0	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	Pleuro_Gyrosigma	5.882352941	2	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	Prorocentrum	14.70588235	5	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	PseudoNitzschia	2.941176471	1	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	Rhizosolenia	11.76470588	4	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	Thalassionema	2.941176471	1	0.75	89.8847631
flowcam_FCM.U0001.2016-01-21.300A4X.01	Thalassiosira	2.941176471	1	0.75	89.8847631
flowcam_FCM.U0002.2016-01-21.300A4X.01	Asterionellopsis	1.388888889	1	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	C_ourvisetus	1.388888889	1	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	C_danious	12.5	9	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	C_socialis	0	0	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	Ciliophora	6.94444444	5	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	Dactyliosolen	5.55555556	4	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	Dytilum	6.94444444	5	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	G_flaccida	4.166666667	3	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	G_striata	5.55555556	4	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	Gymnodinium	5.55555556	4	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	Lauderia	0	0	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	Leptocylindrus	0	0	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	Odontella	4.166666667	3	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	P_globosa	0	0	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	Pleuro_Gyrosigma	4.166666667	3	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	Prorocentrum	15.27777778	11	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	PseudoNitzschia	8.333333333	6	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	Rhizosolenia	6.94444444	5	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	Thalassionema	5.55555556	4	0.75	90.7581227
flowcam_FCM.U0002.2016-01-21.300A4X.01	Thalassiosira	5.55555556	4	0.75	90.7581227

Bouton VIEW

En cliquant sur ce bouton, une nouvelle fenêtre de sélection de groupes apparaît.

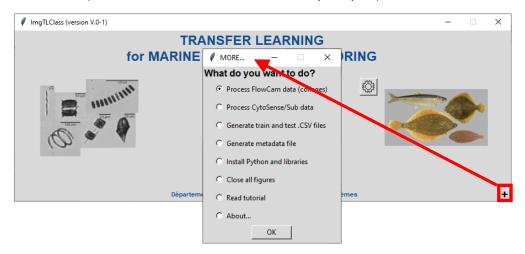






Bouton MORE... (+)

Pour aider l'utilisateur dans les différentes étapes de mise en forme des données pour l'analyse des images, plusieurs options sont disponibles. Pour visualiser la liste de ces outils additionnels, cliquer sur le bouton + (MORE..., en bas à droite de la fenêtre principale).



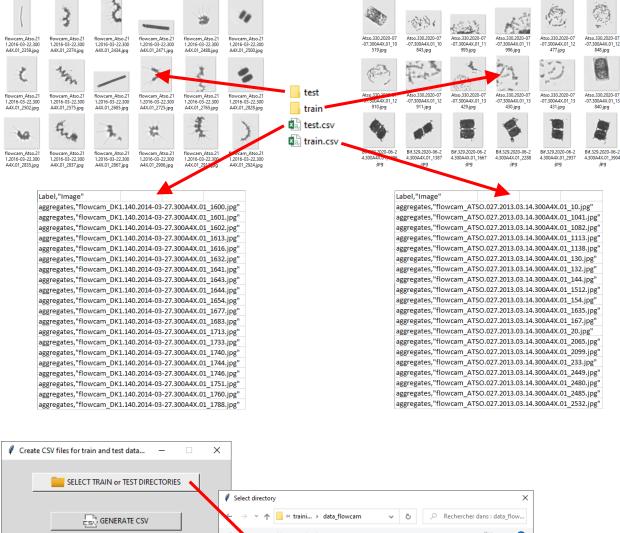
Process FlowCam data (collages)

Les données brutes issues du FlowCam sont souvent présentées sous forme de collages (un fichier avec plusieurs images de particules). Il est alors possible de découper et sauvegarder les vignettes (un fichier image par particule) à partir de ces collages. Pour cela, choisir **Process FlowCam data (collages)**.

Process CytoSense/Sub data

TO DO!

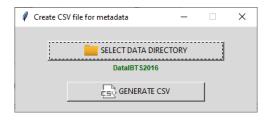
Generate train and test .CSV files



Nouveau dossier Nom Modifié le 14/03/2022 16:19 est test ownCloud 14/03/2022 16:20 ___ train OneDrive - Perso Create CSV files for train and test data... Ce PC data_flowcam SELECT TRAIN or TEST DIRECTORIES Sélectionner un dossier data_flowcam GENERATE CSV

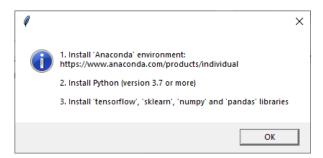
Generate metadata file

Template pour le fichier CSV.



Sample	Date	Volume_ml	Longitude	Latitude	Depth_m
flowcam_FCM.U0001.2016-01-21.300A4X.01	21/01/2016 09:19	10,0003	-1,056167	50,098333	1
flowcam_FCM.U0002.2016-01-21.300A4X.01	21/01/2016 13:03	10,0003	-1,139833	49,832333	1
flowcam_FCM.U0003.2016-01-21.300A4X.01	21/01/2016 16:30	10,0003	-0,890667	49,578333	1

Python and libraries requirements



Close all figures

TO DO!

Read tutorial

PDF file.

