

## Quick manual making a database II by images

*Date* : *September 2020*

*Referencenumber* : *TRP 2020.025*

*by* : *Thomas Rutten*

## Making a database(.mat) II by images

EasyClus can use a database or library, which can be made on basis of own species. This database contains the combination of FCM-attributes data, the signals profiles and the images. In the previous manual (TRP 2020.023), the building of the database is described after clustering (method 1).

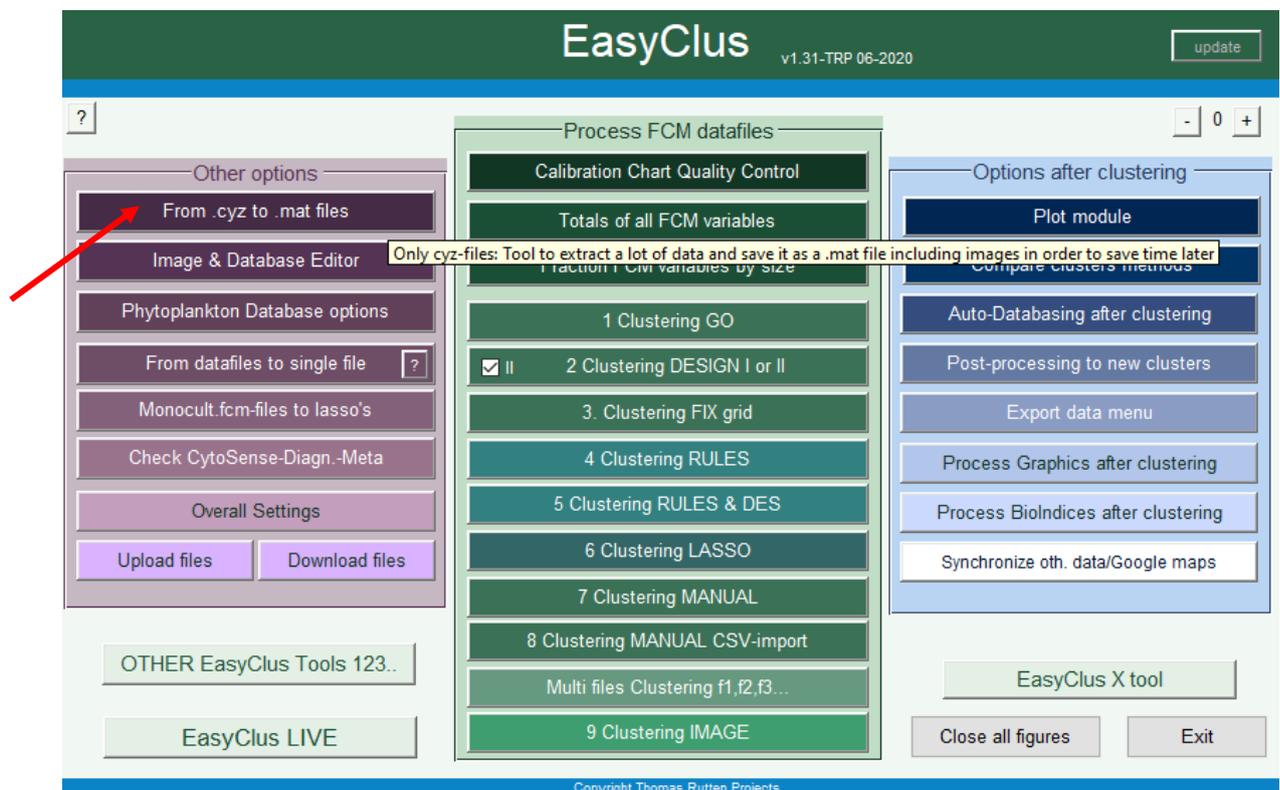
This manual describes the making of a database by moving images to species database folders (method 2). Only particles flowcytometric data and profiles with an image are used. This means (usually) that you need images from several files to build a database which is large enough (let's say at least 25 images for each species)

After the building of the database, clustering can be done by the RULES or LASSO method, the EasyClus supervised clustering methods. These clustering methods need a database (.mat) to assign individual particles into clusters.

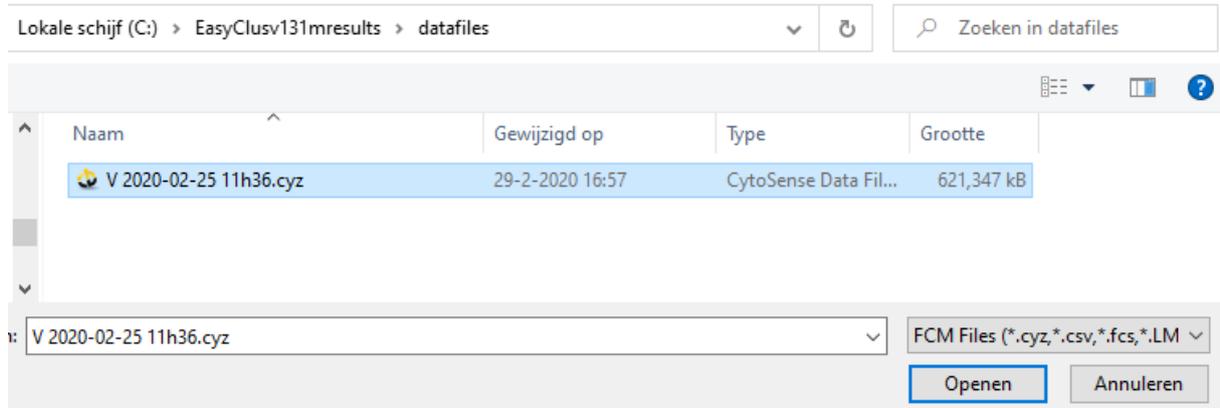
### Method 2: Images are available

- 2.1 Extract the images from the .cyz file into separate .jpg files in a separate folder
- 2.2 Make species folders in the database-folder
- 2.3 Copy images to the right species folder
- 2.4 Update the database
- 2.5 Clustering can be done by RULES or LASSO with this database

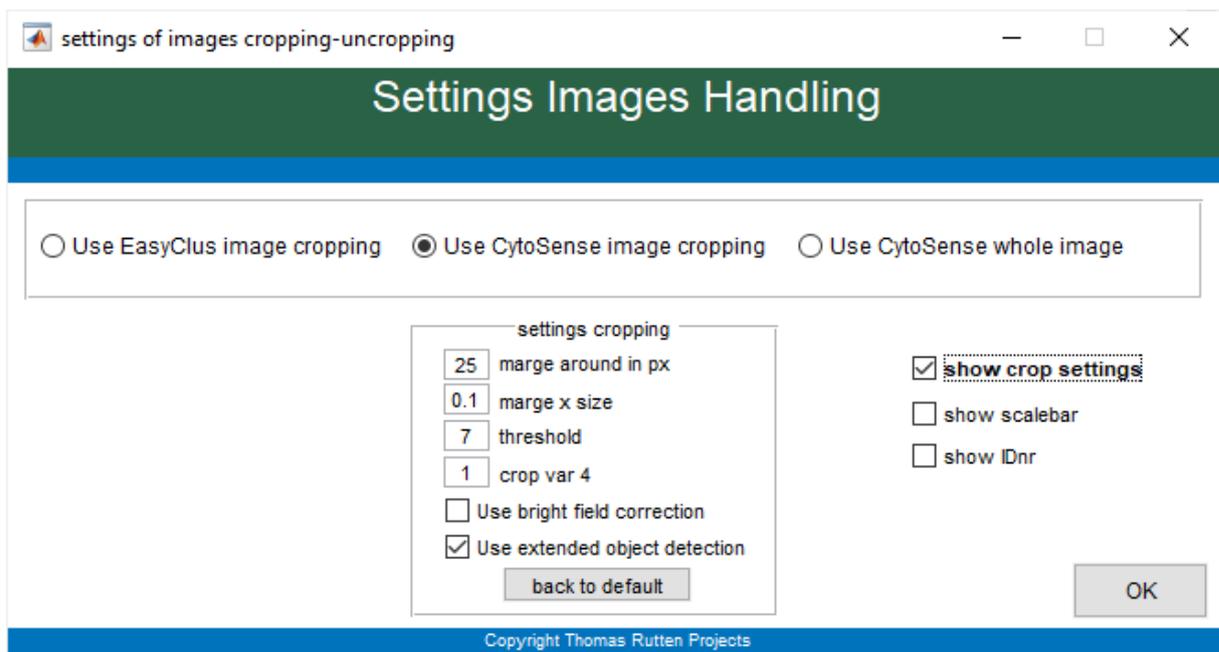
#### 2.1 Extract the images from the .cyz file into separate .jpg files in a separate folder



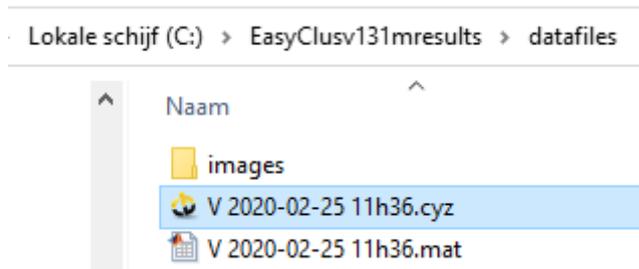
Select a .cyz file:



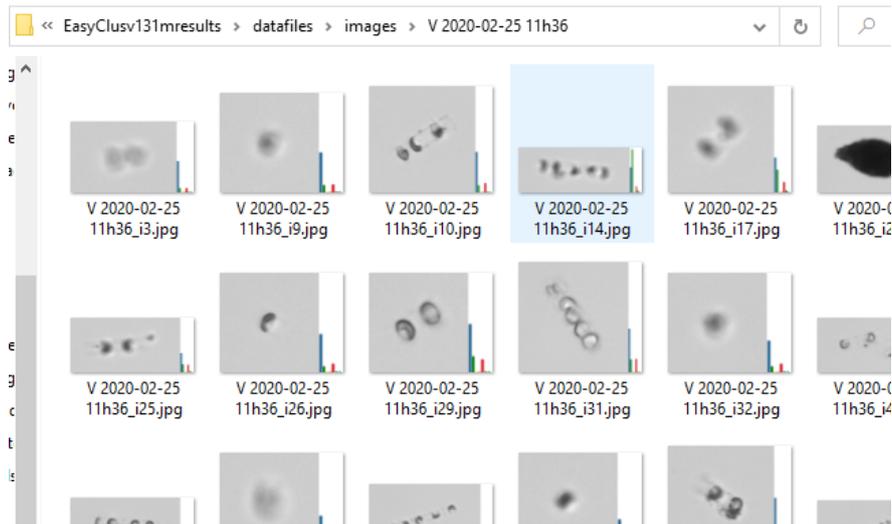
Use the cropping of the images settings below (fast and cropping usually okay):  
(if used before this is automatically stored and you will not see the schedule below, change by 'Overall Settings' button)



The .mat file (with FCM-data and profiles) is stored as well as the images in a separate folder / images / filename /



We have the images in a separate folder:



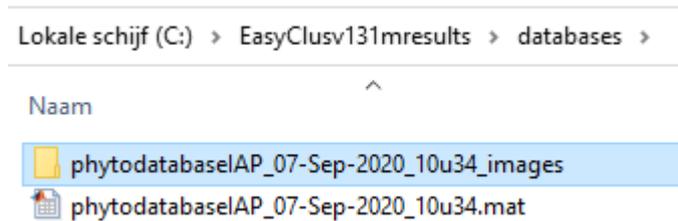
## 2.2 We define an empty database.mat

(if we already have a database, we don't have to start with a new empty database)

The screenshot shows the 'Image & Database Editor' software interface. On the left, a menu titled 'Other options' is visible, with a red arrow pointing to the 'Image & Database Editor' option. The main window has a green header and a list of actions on the left, each with a corresponding button on the right. A second red arrow points to the 'New empty database' button. On the right side, there is a 'From Images' section with a grid of image thumbnails and a green arrow pointing down to the text 'to a Phytodatabase'. A 'Back' button is located at the bottom right.

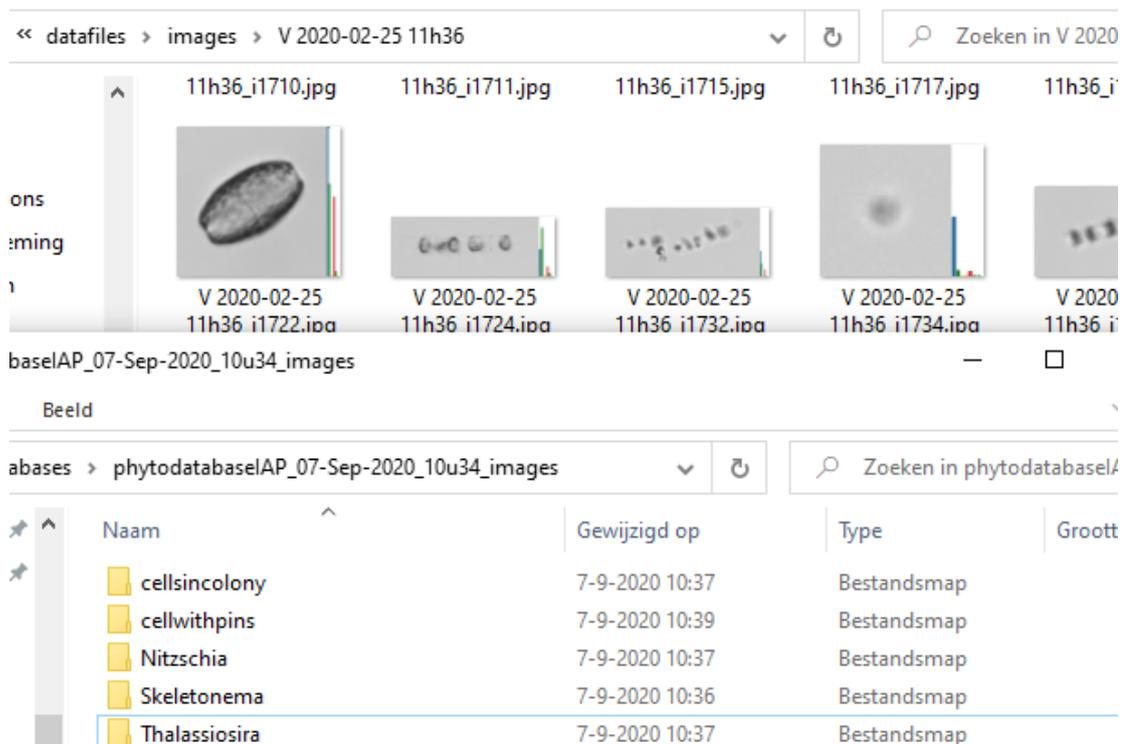
Action	Button
Make a NEW EMPTY database	New empty database
SWIPE IMG in/out FOLDER You have changed names, removed images, add new images in the dbs_image FOLDER directly: Update is database is only DONE by pressing this button!	Update DBS
COPY the chosen database	Copy database
MAKE species SUBFOLDERS deeper than \dbsname_images\ folder to swipe or copy images of .mat converted files to	Make species subfolders
MERGE DATABASES TOGETHER You want to phyto-.mat files with images together to create one big phytodata_*.mat file.	Merge databases
IMAGES SELECTION which need to be ADDED to another database	Add Images II to I
IMAGES SELECTION You want to use the IMAGE EDITOR to edit images names and/or add to existing or other database:	Image Editor
IMPORT PHYTO_*.mat You want to use the DATABASE EDITOR to edit events, which are in a stored database :	Database Editor
FILE IMGMOosaic You want to use the PROPOSAL IMAGE MOSAIC tool to edit in a FCM-file events for adding them to a database:	Proposal mosaic
TABLE.TXT IMPORT You want to the IMPORT.TXT EDITOR to edit events, which are in a stored database :	Import.txt editor
PRECLUSTER You want to use a PRECLUSTER tool to produce jpg-images in a folder with cluster images names :	Cluster file
GET IMAGES FROM CYZ-file You want to produce jpg-images of a FCM-file in a folder :	File to Images

Here is the empty database, with an empty database\_images folder

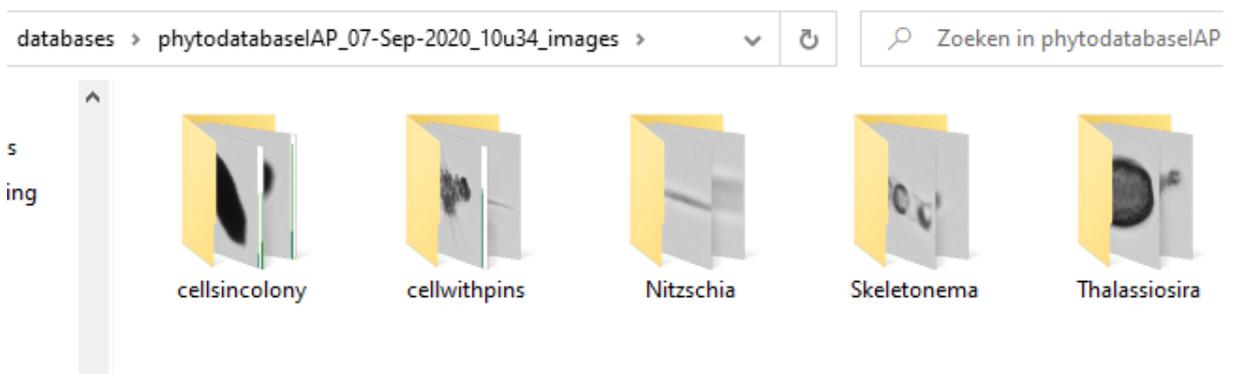


### 2.3 Copy images to the right species folders

We make separate folders with the species name in the database\_images folder:



We copy the separate .jpg images to the species folders.

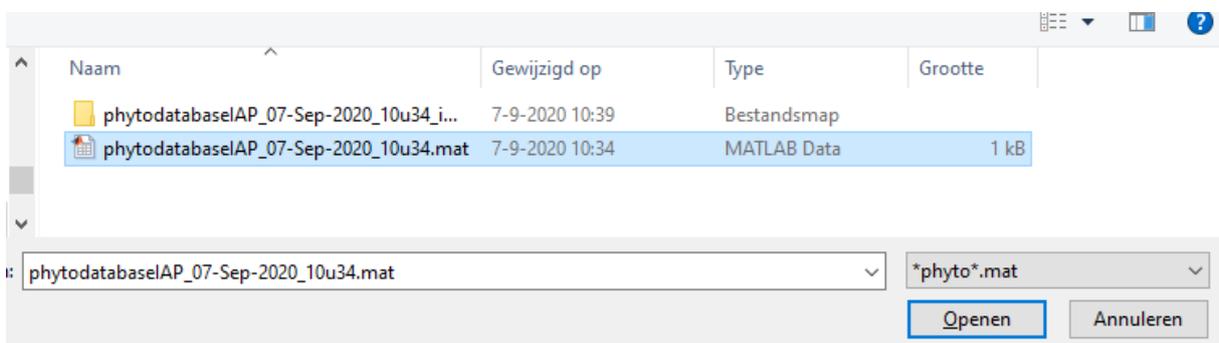


## 2.4 Update the database

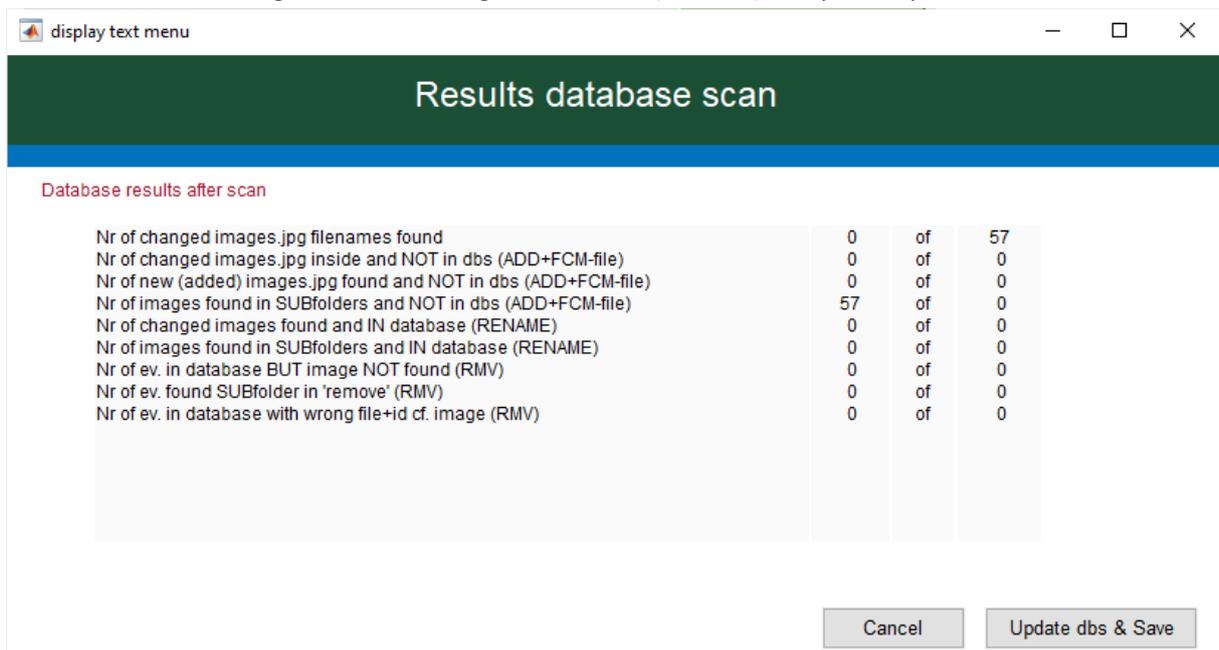
While updating the database, the FCM-data file is needed in the background for importing the FCM-data and profiles belonging the images. This file should be in de \...results\datafiles\ folder!



Update – Choose the database



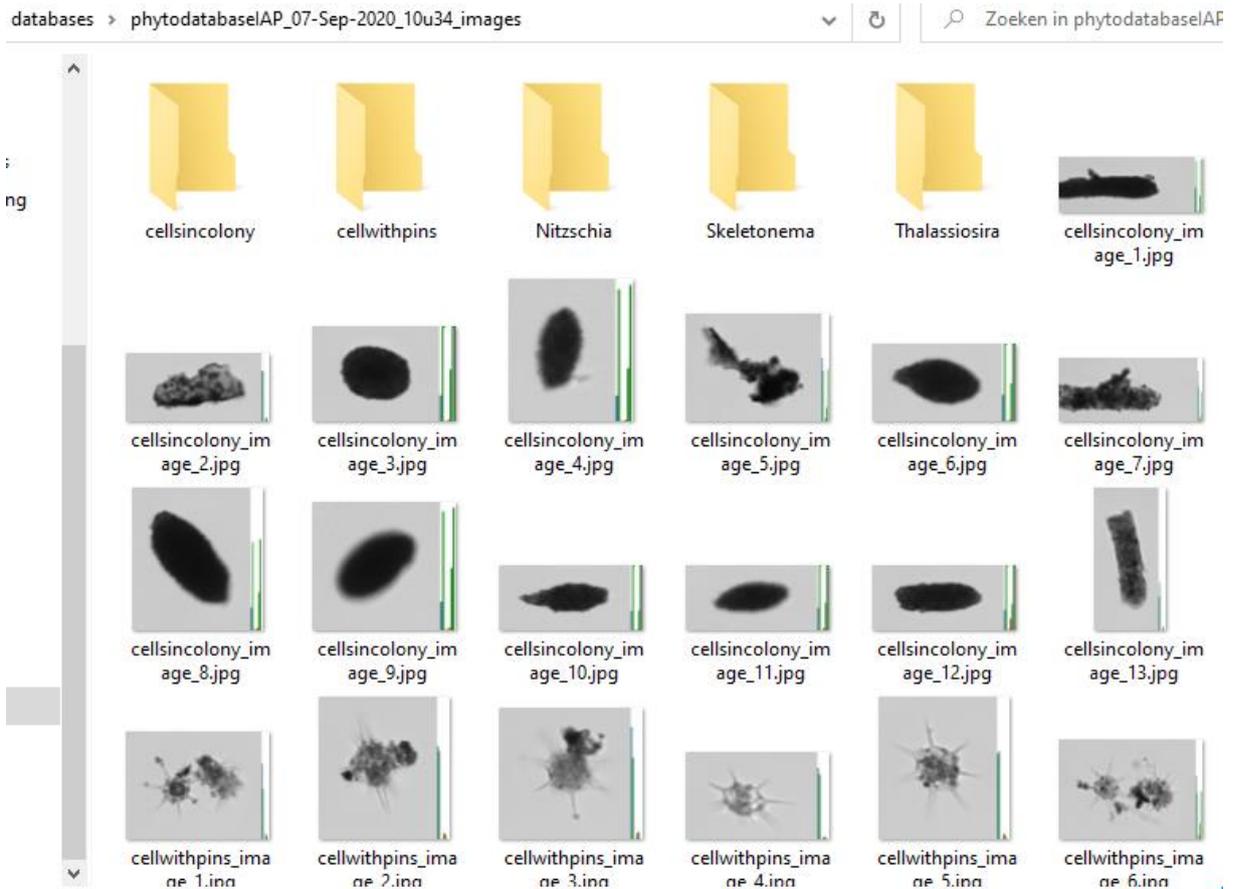
The software recognizes if new images are added (here 57) and press Update dbs & Save:



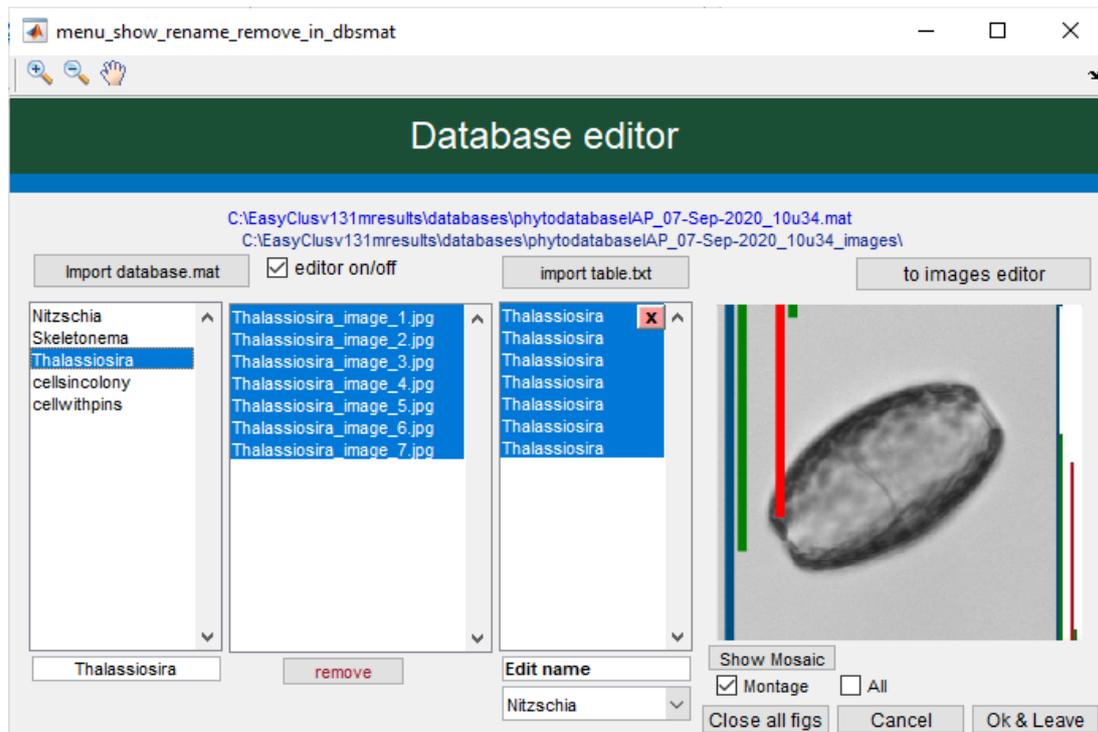
They are added to the database automatically.

```
! -> Nitzschia <- : NEW species added to C:\EasyClusv131mresults\databases\phytodatabaseIAP_07-Sep-2020_10u34.mat
! -> Skeletonema <- : NEW species added to C:\EasyClusv131mresults\databases\phytodatabaseIAP_07-Sep-2020_10u34.mat
! -> Thalassiosira <- : NEW species added to C:\EasyClusv131mresults\databases\phytodatabaseIAP_07-Sep-2020_10u34.mat
! -> cellsincolony <- : NEW species added to C:\EasyClusv131mresults\databases\phytodatabaseIAP_07-Sep-2020_10u34.mat
! -> cellwithpins <- : NEW species added to C:\EasyClusv131mresults\databases\phytodatabaseIAP_07-Sep-2020_10u34.mat
```

They are cut out of the folders and put to the main database\_images directory:



You can check and edit it by the database editor (but this is not needed now):



## 2.5 Clustering can be done by this database

EasyClus RULES supervised clustering using phytoXX.mat

**Database & Matching species settings**

use database    use sel.species   Sel.   Browse dbs

C:\EasyClusv131mresults\databases\phytodatabase\AP\_07-Sep-2020\_10u34.ms

% Critical level (100%=default)

less critical    %   default critical

Increase parameters level

  most critical

use weighing

**Scatterplots selection**

Choose X-axis   Choose Y-axis

TOF  
Length FWS  
Length SWS HS  
Length FL Yellow HS  
Length FL Orange HS  
Length FL Red HS  
Length SWS 2A HS  
Length SWS 2B HS  
Total FWS  
Total SWS HS

TOF  
Length FWS  
Length SWS HS  
Length FL Yellow HS  
Length FL Orange HS  
Length FL Red HS  
Length SWS 2A HS  
Length SWS 2B HS  
Total FWS  
Total SWS HS

Ratio   Add below    Ratio

Import scatterplot combination

**Scatter plots used for visualization**

Total SWS HS vs Total FL Red HS  
Total FL Red HS vs Total FL Orange HS  
Total FL Red HS vs Total FL Yellow HS

Show Sc.plots all clusters (sn2)  
 Show Sc.plots merged clusters (sn3)  
 Size of symbol in scatterplot (2,3,4,5,...):   
 Only visualize the first X number of clusters :   
 show 3D-plots  
 logarithmic    linear  
 Save and Use latest settings auto  
 Save idnrs & clusters .txt   Use preprocessing

Cancel   Cluster

Figure 2: B:Total FL Red HS-Total FL Orange HS

nr.clus all 6 |Shown in plot 100.00%

Total FL Orange HS

Total FL Red HS

## 2.6 Let the datafiles images be sorted by an expert elsewhere is possible with this method 2!

Imagine that you have a lot of FCM files with images.

1. It is possible now that you send these images to elsewhere e.g. to an expert, who can do the sorting of the images by putting them in the species folders for you.
2. These species folders with images are sent back to you.
3. You put the folders in the database\_images folders
4. You ensure that the original FCM-datafiles (.mat) where the images have come from are in de \...\results\datafiles\ folder
5. You update the database and everything will be added to your database!

## 2.7 Be aware

The database uses FCM-data, so be aware that your CytoSense instrument is stable. E.g. check that beads cluster is at the same position as before. The database is your training set and it expects that the data that needs to be clustered will be around the same data values too.