2D classification results - unclassified particles

Thu, 05/31/2018 - 03:56 #1

Lubomir

2D classification results - unclassified particles

Dear Tim,

I ran a couple of 2D classifications and looked for class occupancies in the database using the sqlite commands. To mu surprise, there are a lot of particles (~30%) that aren't classified. They show up like this:

1 0.0 0.0 0.0 0 10.0 0.0
2 0.0 0.0 0.0 0 10.0 0.0
3 0.0 0.0 0.0 0 10.0 0.0
4 0.0 0.0 0.0 0 10.0 0.0
5 0.0 0.0 0.0 0 10.0 0.0
6 106.76000213623 2.29583621025085 -2.78013229370117 75 26.3810520172119 -12987.1298828125

...

So the first five aren't assigned to any class and the 6th one is. What happened to the unassigned particles? Are they excluded as somehow bad?

I had about 120000 particles in total, and only 84000 are non-zero in the database. The "header" programme from imod shows 120000 particles in the input stack, and there don't seem to be any wrong particles in the stack either. The number of non-zero particles corresponds to the number that cisTEM shows on the 2D results page.

Best,

Lubomir

timgrant

Hi Lubomir,

Hi Lubomir,

The 0 particles will not be included in the class averages. They should only be 0 if you are not using 100% of the data (this would happen by default in the early rounds of processing), or if you have Exclude blank edges set to true (in which case, particles which were close to the edge of the micrograph will be excluded).

Are you using less than 100%, or do you have exclude blank edges set to true?

Cheers,

Tim

Thu, 05/31/2018 - 15:38 (Reply to #2)

Lubomir

Hi Tim,

Hi Tim,

thanks for a quick response! I did use "Exclude blank edges" and "Auto percentage"of particles. So I ran a test without the "exclude blank edges" and then there weren't any excluded particles. How wide is that stripe? It cut out 40 000 particles!

Best,

Lubomir

timgrant

Hi Lubomir,

Hi Lubomir,

I think that more or less - if you have any edge at all, it will be removed. It should be off by default though?

Tim

Lubomir

Hi Tim,

Hi Tim,

maybe not, to go the safe way. Btw. the results of that classification were nice! It is a negative stain dataset so maybe that's the reason why so many edges were found.

L.