

07/09 – 11/09

OSCAR, YORAN, SAMIR, DANI, KLARA

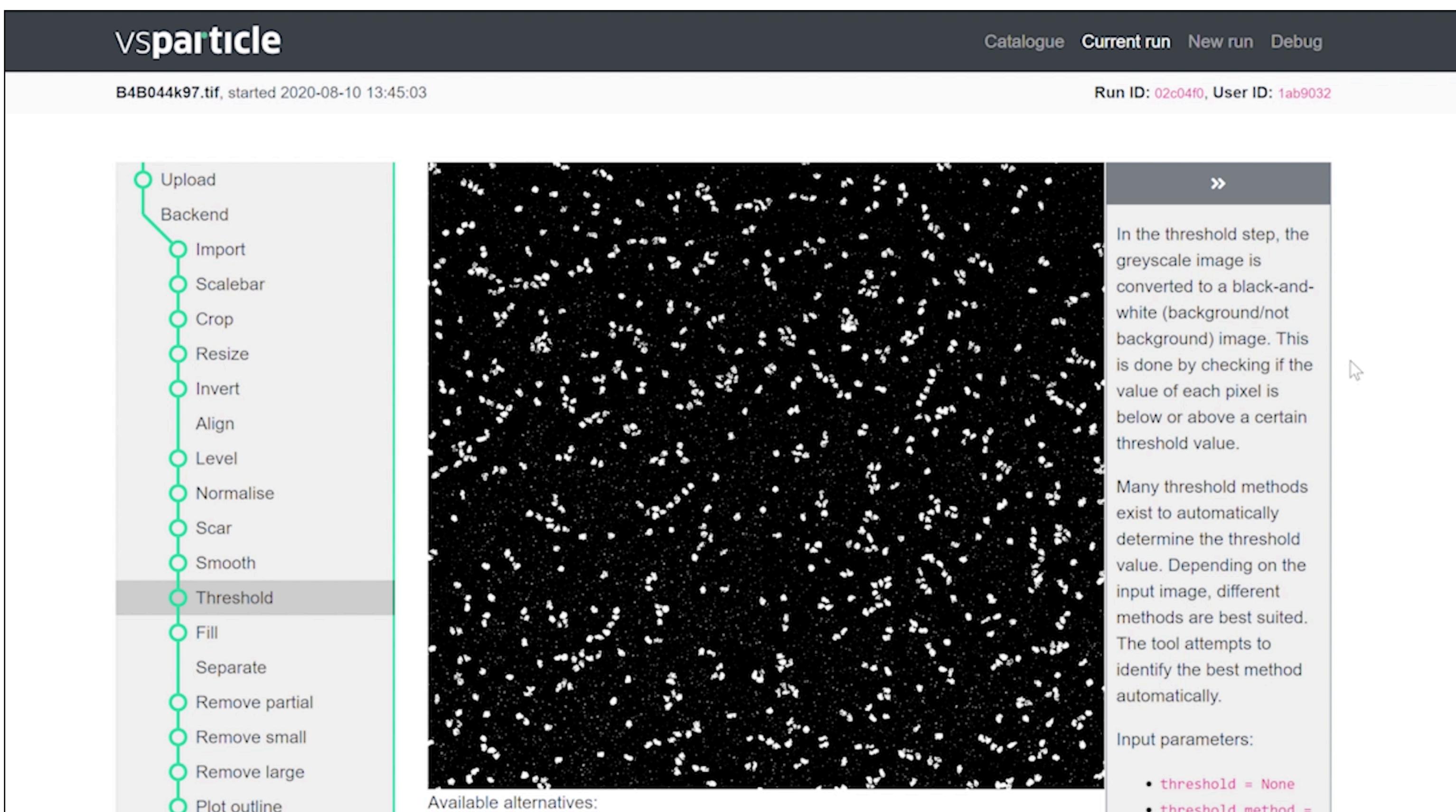
NANO - WEEK 2

OUR PROJECT

- ▶ VSParticle
- ▶ Predict the best thresholding algorithm for a specific image

GENERAL

- ▶ Demo from Tomas (Project owner)
- ▶ DataCamp
- ▶ Meetings with lecturers



WORKING WITH OUR DATASET

- ▶ Clean data
 - ▶ Figure out what parts of the data are useful for us
 - ▶ Script to remove any attributes we do not want/need
- ▶ Show records of dataset that have a relationship

The screenshot shows a MongoDB interface with three dataset documents listed:

- Top Document:** _id: ObjectId("5f31323ad63f6decc02c0499")
parameters: Object
logs: Array
metadata: Object
results: Object
- Middle Document:** _id: ObjectId("5f31333cd63f6decc02c04f0")
parameters: Object
logs: Array
metadata: Object
results: Object
scores: Object
 - > invert: Object
 - > smooth: Object
 - > threshold: Object
 - > isodata: Object
 - > li: Object
 - > otsu: Object
 - > triangle: Object
 - > yen: Object

total: 7
- Bottom Document:** _id: ObjectId("5f31338fd63f6decc02c0553")
parameters: Object
logs: Array
metadata: Object
results: Object
scores: Object

RESEARCH PAPER

- ▶ Find main research question

“HOW CAN A MACHINE LEARNING MODEL BEST PREDICT THE MOST
OPTIMAL THRESHOLDING ALGORITHM FOR A SPECIFIC IMAGE,
BY THE EXAMPLE OF PROCESSING IMAGES OF NANOPARTICLES DONE
BY THE COMPANY VSPARTICLE?”

DIFFICULTIES

- ▶ One algorithm (Yen) is used the majority of the runs
 - ▶ Not enough data samples from other algorithms
 - ▶ Focusing on when Yen is used/not used for now
- ▶ No access to the images that are being processed
- ▶ User score to evaluate how well the process worked is not linked to one specific step, but to the whole process

OUR NEXT STEPS

- ▶ Visualise data: When is Yen used/not used?
 - ▶ Scatter plot
 - ▶ t-SNE plot
- ▶ DataCamp
- ▶ Meetings

THANK YOU!

QUESTIONS?