Multi-label Classification of Crystallisation Outcomes from *In-Situ* Imaging

Parandeep Sandhu[1,2], Christopher Boyle[1,3], Christos Tachtatzis[2] and Javier Cardona[1,2,4]

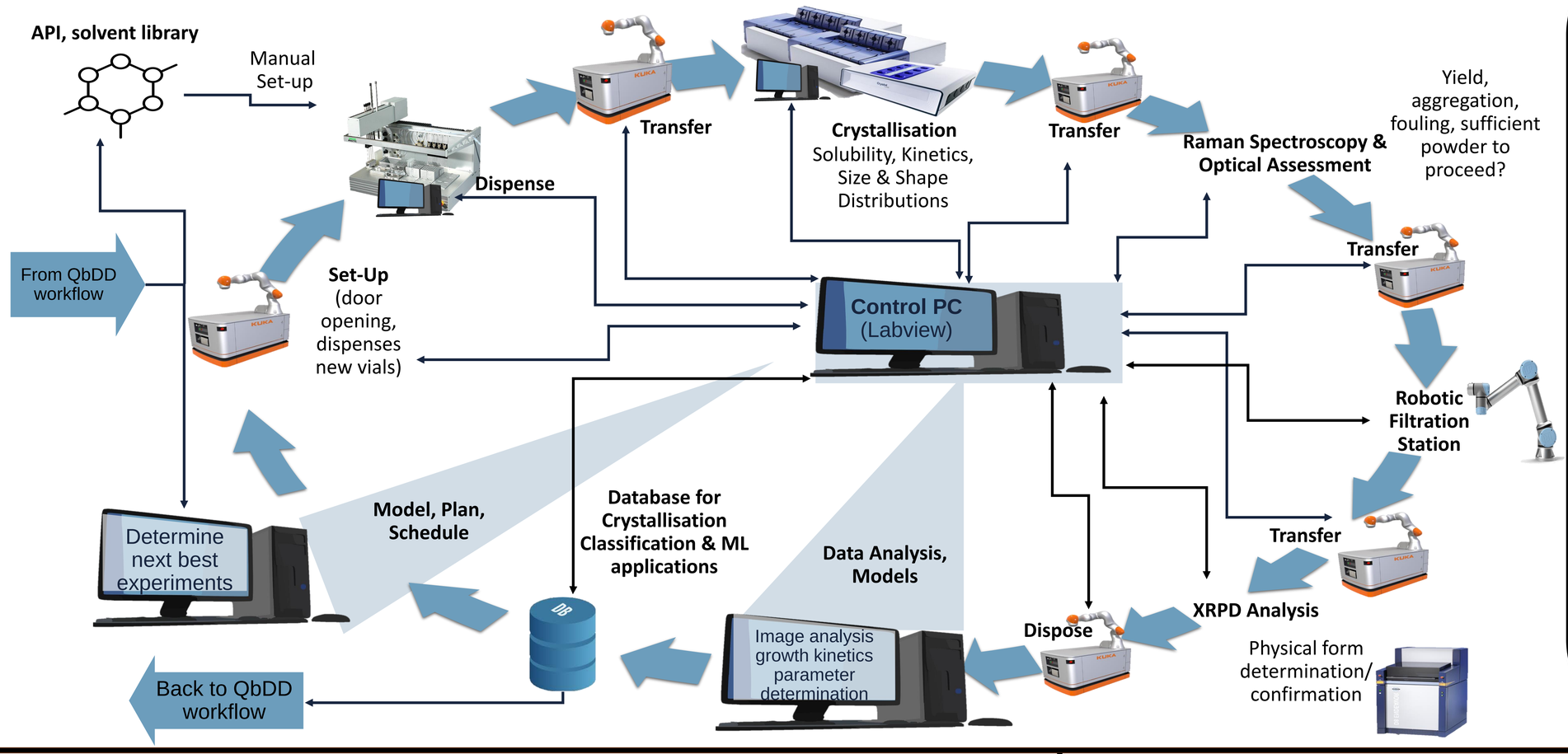
[1] EPSRC Future Manufacturing Research Hub for Continuous Manufacturing and Advanced Crystallisation (CMAC), Glasgow, UK.

[2] Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK.

[3] Strathclyde Institute of Pharmacy & Biomedical Sciences (SIPBS), University of Strathclyde, Glasgow, UK.

[4] Department of Chemical and Process Engineering, University of Strathclyde, Glasgow, UK.

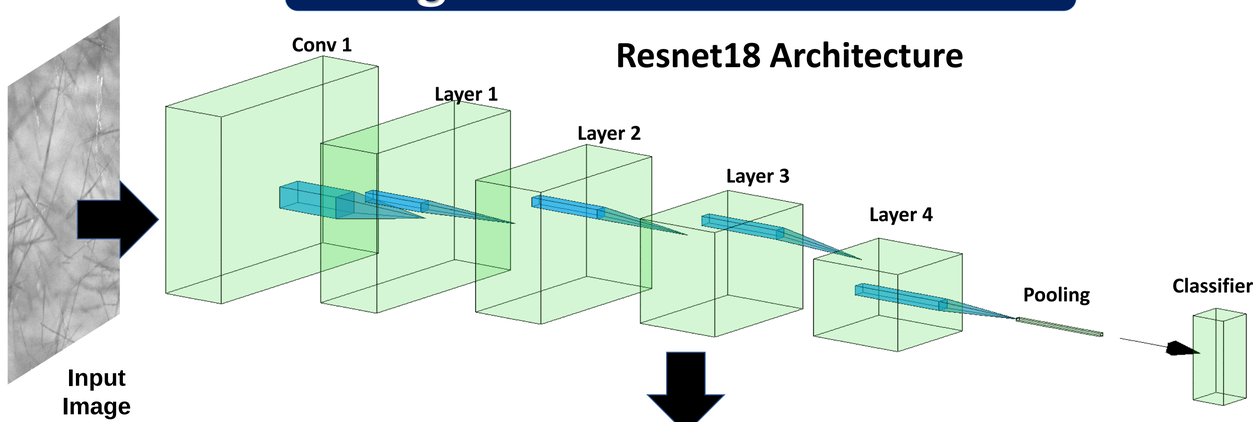
DataFactory

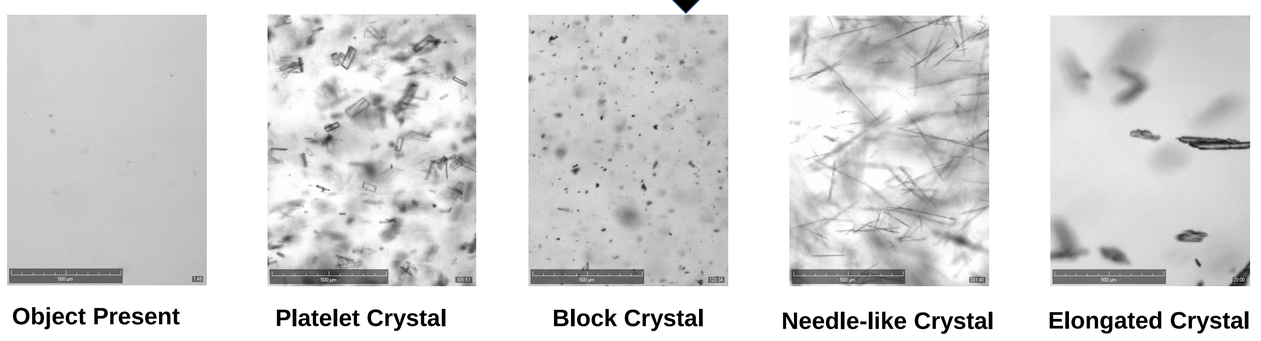


Motivation

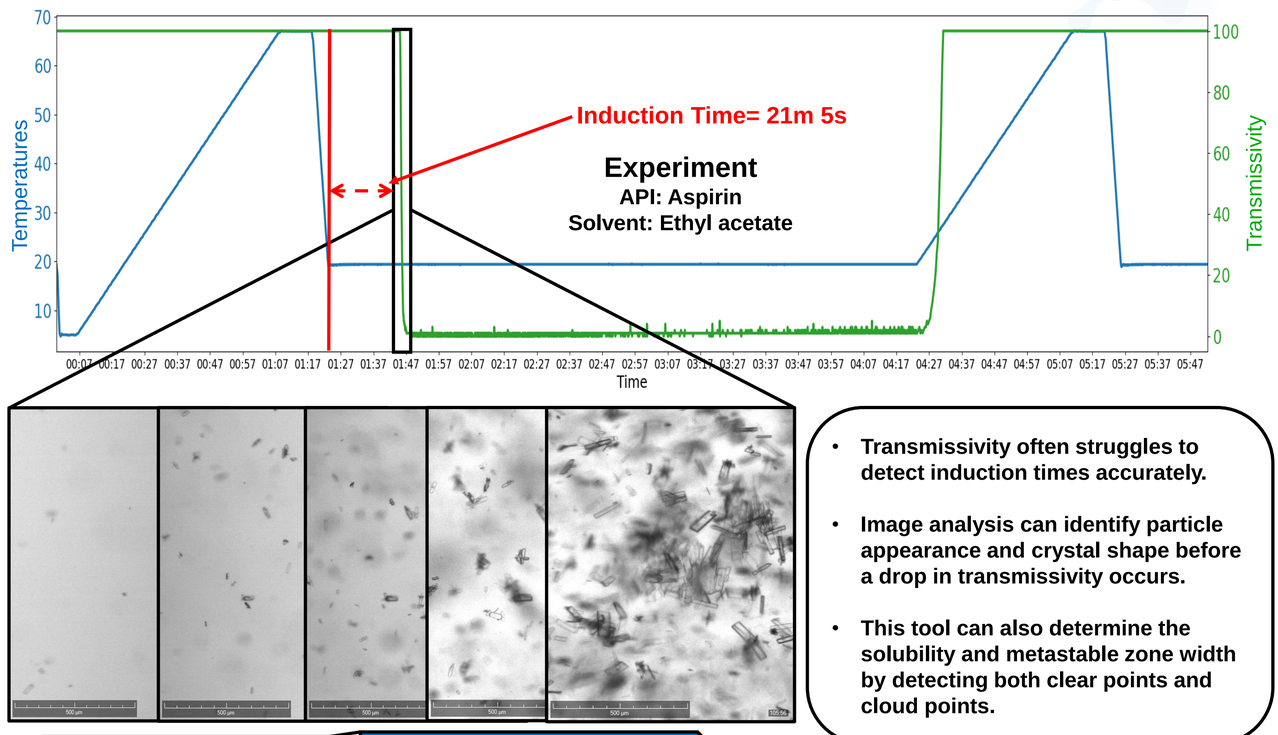
- The Autonomous Crystallisation Screening DataFactory streamlines and enhances the efficiency of crystallisation experiments conducted on the Technobis Crystalline platform.
- By automating these experiments, we aim to generate large, structured datasets that can be rigorously analysed through image analysis and machine learning techniques.
- The comprehensive exploration of the physicochemical phase space—from molecular interactions to solubility, kinetics, growth, agglomeration, and fouling—enables a deeper understanding of crystallisation processes.
- Our work on developing a multi-label classifier for crystalline images is a crucial component of this initiative, ensuring precise and efficient classification of crystallisation phenomena. This allows us to extract relevant features on crystallisation outcomes from *in-situ* images.

Image Classification Model

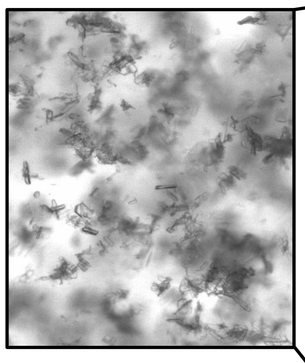




Application Kinetics Case Study



- Transmissivity often struggles to detect induction times accurately.
- Image analysis can identify particle appearance and crystal shape before a drop in transmissivity occurs.
- This tool can also determine the solubility and metastable zone width by detecting both clear points and cloud points.



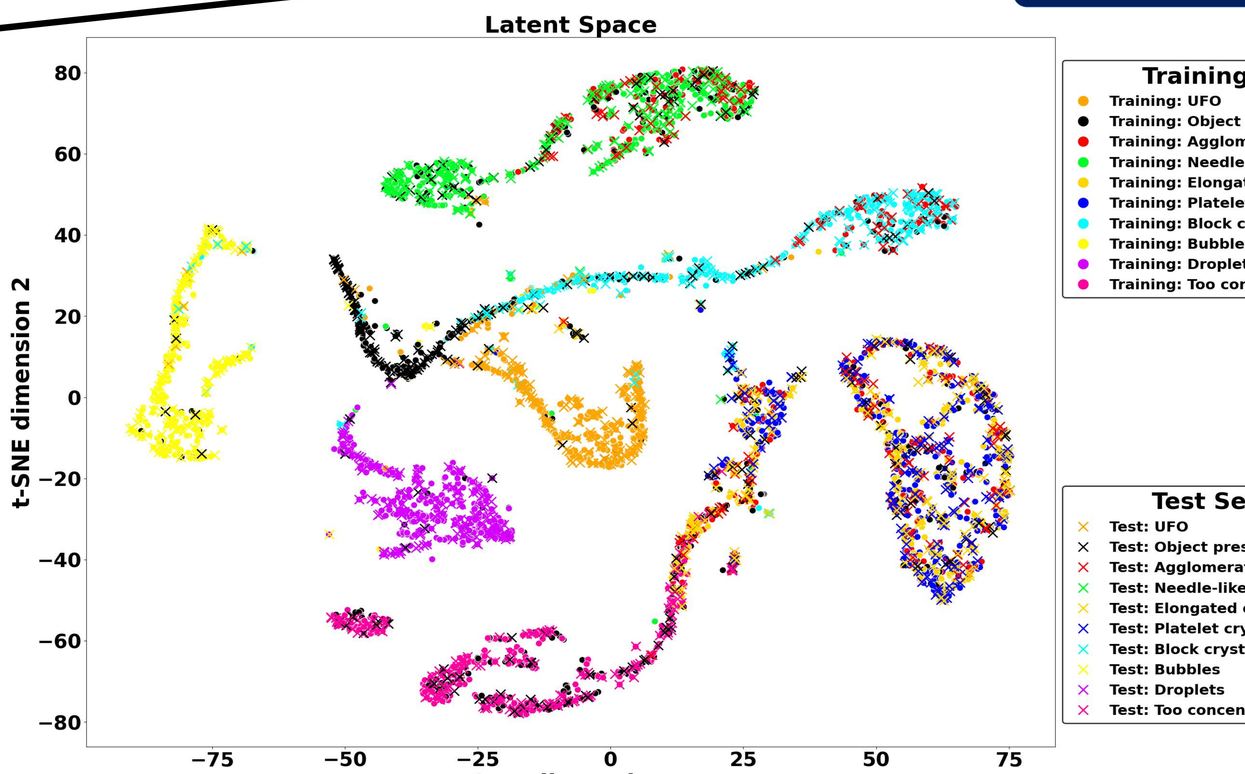
Confidence Scores

Object Present	98%
Platelet Crystal	89%
Block Crystal	25%
Needle Crystal	15%
Elongated Crystal	48%
Agglomerated Crystals	88%
Droplets	2%
Bubble	3%
UFO	10%
Too Concentrated	20%

Model Predictions

- Our model classifies images based on the confidence score for each label. If a label's confidence score exceeds a certain threshold, the image is classified accordingly.
- For a given experiment, the model processes images sequentially. By aggregating the classification results of all images, we can determine the overall outcome of the experiment.

Model Evaluation



Labels	Precision	Recall	F1-Score
Object Present	97.9%	98.4%	98.1%
Platelet Crystal	95.2%	96.7%	91.9%
Block Crystal	86.7%	86.3%	86.5%
Needle-like Crystal	98.5%	98.2%	98.4%
Elongated Crystal	90.5%	93.4%	91.9%
Agglomerated Crystals	96.8%	96.5%	96.7%
UFO	73.3%	66.5%	69.7%
Bubble	94.4%	94.0%	94.2%
Droplets	97.7%	97.4%	97.5%
Too Concentrated	94.7%	96.7%	95.7%

Conclusion

- The multi-label classifier has achieved precision, recall, and F1-score greater than 90% for the majority of labels.
- Moving forward, the Datafactory aims to use this model to classify further experiments, automating the capture and analysis of crystallisation outcomes.
- This approach is not limited to crystalline experiments but can be applied to any sensor capable of taking microscopic images of crystals.