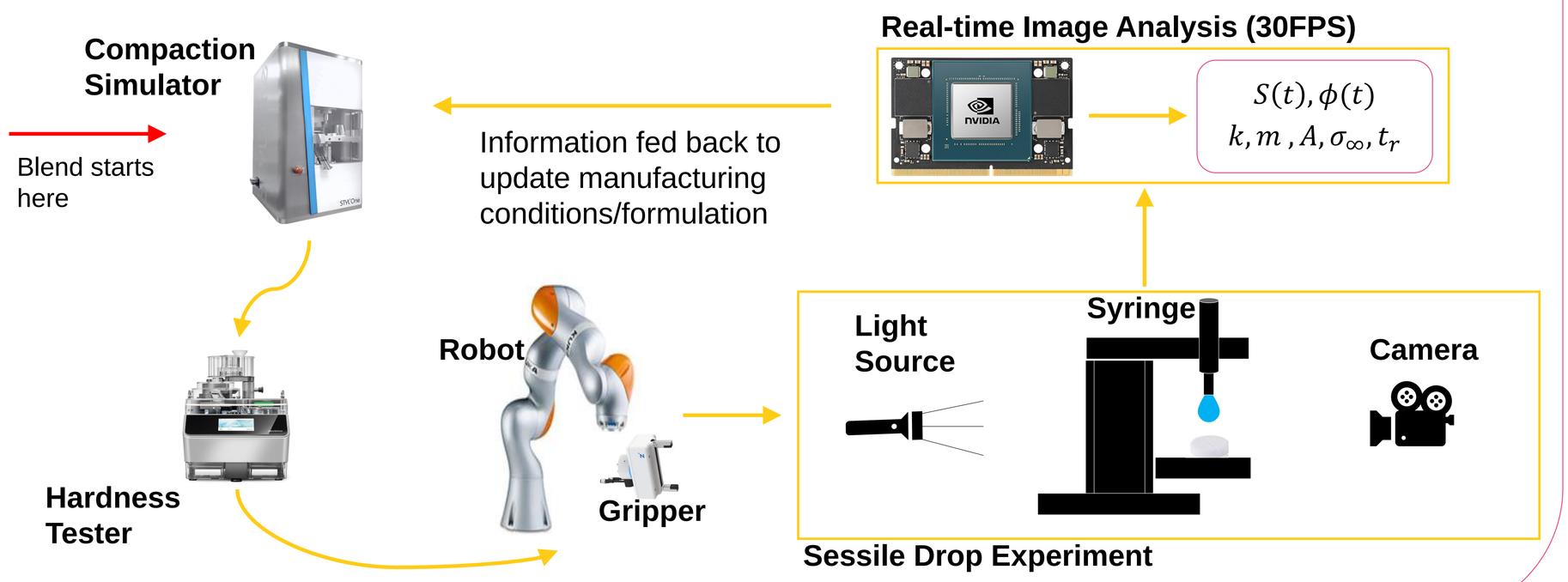
An End-to-end Machine Learning Framework for the Characterisation of Tablet Disintegration in Real-time

John A. Armstrong, Natalie Maclean, Faisal Abbas, Daniel Markl, Blair F. Johnston

Continuous Manufacturing and Advanced Crystallisation (CMAC), Strathclyde Institute of Pharmacy & Biomedical Sciences, University of Strathclyde, Technology and Innovation Centre, 99 George Street, Glasgow G1 1RD, UK.

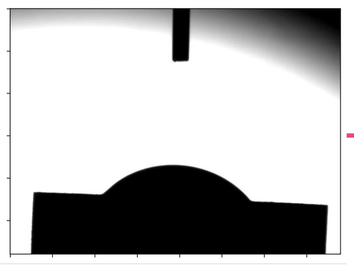
DM² Automated Sessile Drop Experiment Overview

- Tablet disintegration can be described as the combination of four interconnected processes: liquid absorption, swelling, the breaking of interparticle bonds and excipient particle dissolution.
- These mechanisms are related to the raw material properties of the tablet as well as its composition and manufacture.
- Understanding these relationships is important for optimising the formulation and manufacture of a tablet.
- The sessile drop experiment is a simple yet effective tool for testing the liquid absorption and swelling characteristics of a tablet. The following introduces a method for extracting and analysing the data obtained from the sessile drop experiments in real-time (building on the work of [1]) using a combination of feature extraction, classification and regime detection.

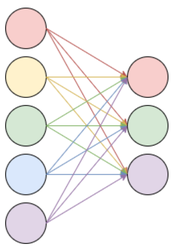


Experimental Data Extraction Using Machine Learning

Griseofulvin/MCC/Mannitol/MgSt



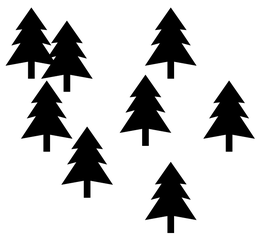
VGG-16 Feature Extraction [2]



Stopped Before First Maxpooling Layer

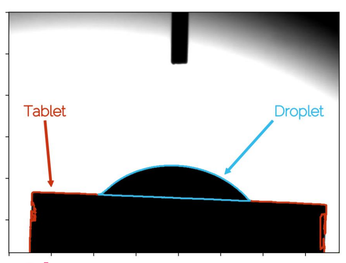
The early layers of a CNN focus on extracting edges and contrasts which makes these features useful for segmentation

Random Forest Classification

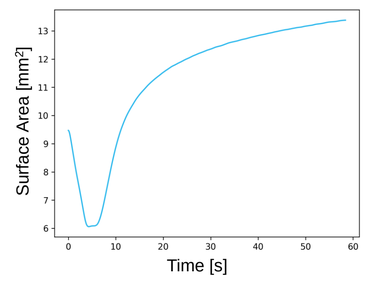


The RF classifies each pixel feature vector with a 1 if it is part of the tablet or droplet and 0 otherwise.

Image Segmentation Result

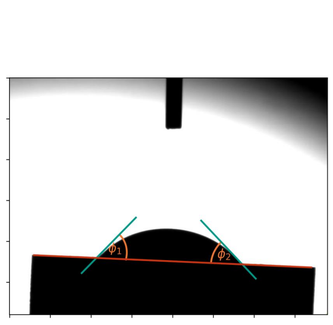


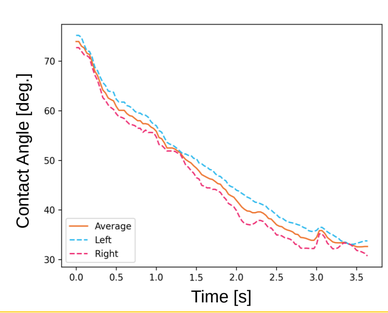
Surface Area, $S(t)$



The surface area of the droplet is just the sum of all pixels identified as being the droplet

Contact Angle, $\phi(t)$

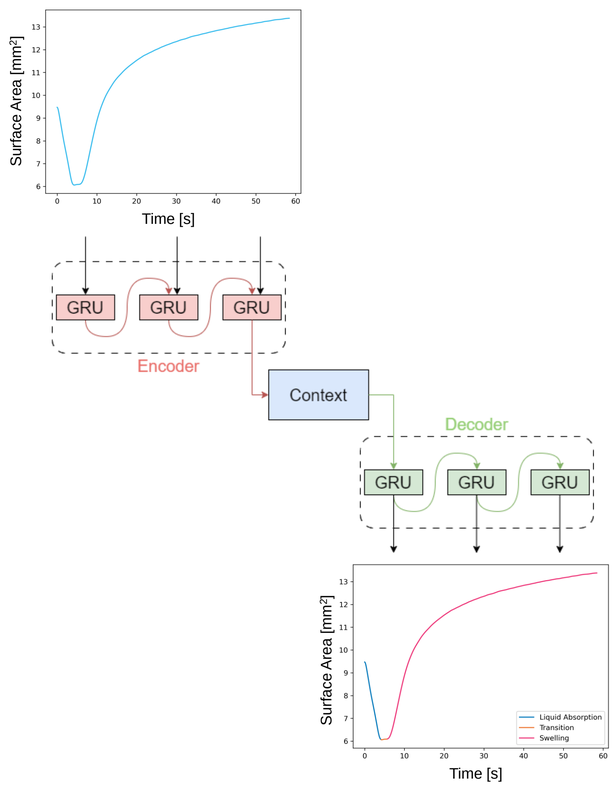




The contact angle is the angle between the tangent to a fitted ellipse and the tablet surface

Identifying Absorption and Swelling

- Surface area curve can be divided into three regimes: liquid absorption, transition and swelling identified by considering the change in surface area [1]
- The classification into the different regimes is done by training a Seq2Seq model [3]
- Often used for language translation, Seq2Seq models encode one sequence into a **context vector** which is then decoded into a set-length sequence of value
- In this case, the surface area at each time is the sequence which is encoded and the decoded sequence identifies the liquid absorption, transition and swelling phases of the sessile drop experiment
- The encoding and decoding is achieved through the use of gated recurrent unit (GRU) networks which use information from previous timepoints as well as information given about the current time to make predictions about the data



The main output of the Seq2Seq model are the time where absorption ends, $t_{a,f}$, and the time where swelling begins, $t_{s,i}$. This allows us to split the surface area into these two processes and fit mechanistic models to them.

Model Fitting

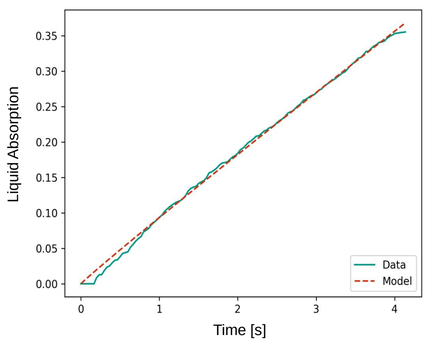
Liquid absorption can then be formally characterised by:

$$y(t) = \left| \frac{S(t) - S(t_{a,i})}{S(t_{a,i})} \right|$$

and modelled by a power law:

$$y(t) = kt^m$$

where $t_{a,i}$ is the absorption start time, and $k,\ m$ are fitting parameters determining the rate of liquid absorption and the linearity of the process, respectively.

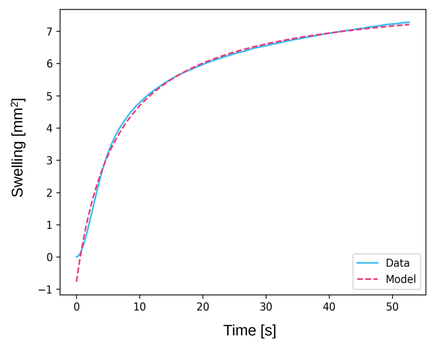


Similarly, swelling is given by:

$$\sigma(t) = S(t) - S(t_{s,i})$$

and modelled by the Schott model [4]:

$$\sigma(t) = \frac{t - t_r}{A + {(t - t_r)}/{\sigma_\infty}}$$



where A is the multiplicative reciprocal of the initial swelling rate, σ_∞ is the swelling rate at infinite time and t_r is the time between the start of swelling and when the swelling becomes uniform.

- This method is 180x faster than not using machine learning allowing for deployment in the automated lab setup.
- [1] showed that tablet disintegration time negatively correlates with both liquid absorption fitting parameters and the initial swelling rate ($1/A$), while being positively correlated with the start time of the swelling, $t_{s,i}$. The efficiency of this framework will allow rapid analysis of historical data stores to uncover more correlations between model and disintegration parameters.
- Once the relationships between model and disintegration parameters is better understood, they can be used to update the manufacturing conditions in real-time to produce tablets with optimal disintegration conditions as part of the automated tabletting setup.

References

[1] Markl, D., Maclean, N., Mann, J., Williams, H., Abbott, A., Mead, H., and Khadra, I. (2021). *International Journal of Pharmaceutics*, 601:120382.
[2] Simonyan, K. and Zisserman, A. (2014). *arXiv:1409.1556v6 [cs]*.
[3] Sutskever, I., Vinyals, O. and Le, Q.V. (2014). *arXiv:1409.3215 [cs].*
[4] Schott, H. (1992). *Journal of Pharmaceutical Sciences*, 81(5):467–470.

Acknowledgment: The authors would like to thank the Digital Medicines Manufacturing (DM²) Research Centre (Grant Ref: EP/V062077/1) for funding this work. DM² is co-funded by the Made Smarter Innovation challenge at UK Research and Innovation, and partner organisations from the medicines manufacturing sector. For more information, visit cmac.ac.uk/dm2-home

DM² website