**Title of Research**

Machine Learning Methods for the Agtron of Coffee Beans and Ground Coffees

**Objective**

* To use machine learning model to give out accurate roast level of coffee from photos taken by users.
* Constructing a dataset that has photos of coffee beans and grounded coffee as data and Agtron as label.

**Background**

Roasting is a crucial process to the sensory quality of coffee. The color of coffee beans changes during roasting, and the color of roasted coffee is related to coffee roasting level (Yeager et al., 2022). Ontoum et al. developed a Mobilenet based classifier to determine the general coffee roasting level of coffee (light, medium, dark), while no Agtron is provided by their model. Agtron is the industry standard roasting level measurement measured by specific spectrometer. It can help Roasters determine how evenly the coffee is roasted, which significantly influence the flavor of coffee. Another significance of Agtron is that it helps Roasters to keep different batches of coffee beans consistent.

Traditional methods for Agtron calculation is based on near-infrared (NIR) spectroscopy (Pires et al., 2021). Most of such equipment (Agtron M − basic II, Lighttells CM – 200, DiFluid OMNI) are high cost for individual roasters. L\*a\*b\* measurements of coffee were studied in past studies, and a relationship between L\*a\*b\* and Agtron was discovered (Yeager et al., 2022). Yet no formula between L\*a\*b\* and Agtron is given, such relationship represent a probability for machine learning to output Agtron value from L\*a\*b\* value of the coffee. In this case, photos taken by users can be used to determine Agtron of coffee without specific equipment.

Yeager et al. (2022) carries out color analysis with a tristimulus colorimeter at a consistent angle and lighting environment, which is hard to achieve at inference level. As a result, the model for Agtron prediction based on photos should be robust to environment light.

No existing dataset of inconsistently photoed coffee on white paper with Agtron as label is available online, so ideally constructing such a dataset and publish it is also my goal.

**Methodology**

An DiFluid Omni is already purchased (back in China), and one of my friends, who is a roaster, regularly send me photos of coffee he roasted (intact and grounded) on a 70g double A print paper with Agtron value as label.

Only cutting and resizing would be applied to the photos as input, since these photos contains some noise that could also appear at inference level.

There are two models I would like to test. The first model directly takes photos as input and outputs Agtron. To make the model robust to environment light, photos under different lighting conditions of the same coffee could be used for training so the model would be insensitive to lighting condition.

The second model has two modules. Since coffees are photoed on double A print paper, and expect print paper is easy to acquire by users, the first module could standardize the L\*a\*b\* of coffee with respect to the print paper. Then the second module could map the L\*a\*b\* of coffee towards the Agtron of coffee.

**References**

Ontoum, S., Khemanantakul,T., Sroison, P., Triyason, T., and Watanapa, B. Coffee Roast Intelligence. *arXiv preprint arXiv:2206.01841,* 2022

Pires, F. C., Pereira, R. G. F. A., Baqueta, M. R., Valderrama, P., and Alves da Rocha, R. Near-infrared spectroscopy and multivariate calibration as an alternative to the Agtron to predict roasting degrees in coffee beans and ground coffees. *Food chemistry*, pp. 365, 130471. 2021

Yeager, S. E., Batali, M. E., Lim, L. X., Liang, J., Han, J., Thompson, A. N., Guinard, J.-X., and Ristenpart, W. D. Roast level and brew temperature significantly affect the color of brewed coffee. *J Food Sci*, pp. 87, 1837–1850. 2022