A TRANSFER AND MULTI-TASK LEARNING APPROACH TO CONSTRUCTING METABOLIC MODELS FOR BIOMANUFACTURING

Yen-An Lu, Department of Chemical Engineering and Materials Science, University of Minnesota, MN, USA
lu000285@umn.edu
Wei-Shou Hu, Department of Chemical Engineering and Materials Science, University of Minnesota, MN, USA
Qi Zhang, Department of Chemical Engineering and Materials Science, University of Minnesota, MN, USA

Key Words: Metabolism, Modeling, Cell Culture, Transfer Learning, Multi-Task Learning

Metabolism is the driver for environmental changes in cell culture processes and exerts profound effects on productivity and product quality [1]. Over time in culture, cells switch metabolic states with varying growth rates, nutrient, and metabolite levels. While cell lines derived from the same parental host cell bear similar metabolic networks, differences in their metabolic profiles are commonly observed. We hypothesize that those differences are the results of the subtle variations in metabolic machinery as well as the regulatory network, and that the global similarities and cell line differences can be captured by kinetic metabolic models.

A kinetic metabolic model that integrates the effect of growth control and metabolic allosteric regulations has been developed to capture the metabolic profiles of cells and facilitate the development of predictive tools for process enhancement [2]. While the model is largely generic to different cell lines, its application to specific cell lines requires a large amount of data for parameter estimation. However, the rich data sets are usually available only for well-established processes, whereas the need is greatest for the development of new cell lines, which often have sparse datasets. Two emerging machine learning techniques can be applied to address the limited-data issue: Transfer Learning (TL), which transfers learned knowledge from the existing model to a new one [3], and Multi-Task Learning (MTL), which allows parameter sharing while training the models for multiple processes or cell lines simultaneously [4]. In this work, we develop a TL-MTL-based algorithm that facilitates the data-driven construction of kinetic metabolic models for cell lines with limited data. The parameter estimation problem is formulated as a nonlinear optimization problem with two tailored regularization terms in the loss function, where one transfers certain learned knowledge from the previously developed model for generic cells while the other allows the simultaneous training of the models for multiple cell lines while accounting for shared model features as well as varying cell-line behaviors.

We collect sets of process and transcriptome data from different CHO cell lines over time in fed-batch culture. The process data are used as the training data for the proposed method while the transcriptome data are used to account for the differences as well as the dynamics of the cell lines' metabolic machinery. As a result, the proposed method fine-tunes the metabolic models for different cell lines, by which they recover the metabolic profiles of cell lines observed in the culture. The result highlights the efficacy of the approach for handling the limited-data scenario and indicates the significance of transcriptome variability to cell metabolism. Finally, model simulations are performed to distinguish the metabolic characteristics of different cell lines and identify factors that impact the metabolic shifts the most, providing insights into potential cell line improvement. We envision the algorithm can greatly improve the applicability of cell line-specific models that facilitate the new process development.

[1] H. Le *et al.*, "Multivariate analysis of cell culture bioprocess data—Lactate consumption as process indicator," *J. Biotechnol.*, vol. 162, no. 2–3, pp. 210–223, 2012, doi: 10.1016/j.jbiotec.2012.08.021.
[2] C. M. O. Brien, Q. Zhang, P. Daoutidis, and W. Hu, "A hybrid mechanistic-empirical model for in silico mammalian cell bioprocess simulation," *Metab. Eng.*, vol. 66, no. January, pp. 31–40, 2021, doi: 10.1016/j.ymben.2021.03.016.
[3] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, 2010, doi: 10.1109/TKDE.2009.191.
[4] M. Crawshaw, "Multi-Task Learning with Deep Neural Networks: A Survey," *arXiv*, 2020. doi: 10.48550/arXiv.2009.09796.