# Advancements in AI enhance analysis of biological data from individual cells



Researchers at the Technical University of Munich (TUM) and Helmholtz Munich have demonstrated significant advancements in the use of artificial intelligence, specifically self-supervised learning, to analyse vast quantities of biological data derived from individual cells. This innovative approach aims to deepen understanding of human cellular functions, particularly how variations occur between healthy individuals and those with diseases.

Our bodies are composed of approximately 75 billion cells, and discerning the specific roles of these cells is crucial for medical research and treatment development. The study's focus is on single-cell technology, which allows for the examination of tissues at the cellular level. This technology enables researchers to identify the distinct functions of various cell types and investigate the alterations caused by factors such as smoking, lung cancer, and COVID-19 on cellular structures.

The research tackles the challenge of interpreting the immense datasets generated through these investigations. Machine learning, particularly self-supervised learning, has emerged as a viable method for analysing unlabelled data—data that has not been pre-classified into specific categories. According to Fabian Theis, who holds the Chair of Mathematical Modelling of Biological Systems at TUM, self-supervised learning proves beneficial because it leverages the abundance of unlabelled data to derive meaningful insights without necessitating prior categorisation.

Two core techniques underpin self-supervised learning: masked learning, where parts of the input data are concealed and the model learns to reconstruct the omitted portions, and contrastive learning, which involves teaching the model to distinguish between similar and dissimilar data. In their recent study, published in Nature Machine Intelligence, the research team employed these techniques to analyse over 20 million individual cells, comparing the effectiveness of self-supervised learning against traditional learning methods.

Notably, the researchers found that self-supervised learning notably enhances performance in transfer tasks—where insights from larger datasets inform the analysis of smaller ones. Additionally, the study produced promising results when executing zero-shot cell predictions, which entail performing tasks without pre-training. The team concluded that masked learning is particularly well-suited for tackling large datasets associated with single-cell genomics.

With these findings, the researchers are progressing towards the development of virtual cells—sophisticated computer models that encapsulate the diversity of cells across various datasets. These models are anticipated to facilitate the analysis of cellular changes related to diseases, thereby offering significant potential for future biomedical research. The study provides critical insights into the efficient training and optimisation of these models, marking a noteworthy step forward in the intersection of artificial intelligence and biomedicine.

For more information, individuals can contact Professor Fabian Theis at the Technical University of Munich.

Source: [Noah Wire Services](https://www.noahwire.com)

## References

* <https://www.helmholtz-munich.de/en/newsroom/news-all/artikel/a-key-to-analyzing-millions-of-individual-cells> - This article supports the claim that researchers at Helmholtz Munich and TUM are using self-supervised learning to analyze millions of individual cells, focusing on single-cell technology to understand cellular functions and variations between healthy and diseased states.
* <https://www.helmholtz-munich.de/en/newsroom/news-all/artikel/a-key-to-analyzing-millions-of-individual-cells> - It corroborates the use of machine learning, particularly self-supervised learning, for analyzing unlabelled data in single-cell genomics.
* <https://www.actuia.com/english/munich-researchers-develop-ai-methods-for-next-generation-biomedical-research/> - This article highlights the development of AI methods by Munich researchers for next-generation biomedical research, including the use of single-cell genomics to understand disease origins.
* <https://www.helmholtz-munich.de/en/newsroom/news-all/artikel/single-cell-analysis-best-practice-guidelines> - It discusses the development of best practice guidelines for single-cell analysis by researchers at Helmholtz Munich and TUM, reflecting the complexity and rapid evolution of the field.
* <https://www.nature.com/nmi> - This is the website for Nature Machine Intelligence, where the study on self-supervised learning in single-cell genomics was published.
* <https://www.tum.de/en/> - The Technical University of Munich (TUM) is a key institution involved in the research on AI and single-cell genomics.
* <https://www.helmholtz-munich.de/en/> - Helmholtz Munich is another primary institution involved in the research, focusing on computational health and biological systems.
* <https://www.nature.com/articles/s41576-023-00586-w> - This article provides best practices for single-cell analysis across different modalities, reflecting the ongoing advancements in the field.
* <https://www.sc-best-practices.org> - This online book offers detailed guidelines and code examples for single-cell analysis, supporting the development of robust analytical methods.