International Journal of Multidisciplinary Research and Technology
ISSN 2582-7359, Peer Reviewed Journal, Impact Factor 6.325
www.ijmrtjournal.com

DIGITAL BIOLAB SIMULATION WITH DIGITALCELL.AI

(A MODULE OF PROBIODIGITAL.AI)

Pramoda Kumar Sabat
Sterling, Virginia, USA

Abstruct:

A multi-scale, multi-modal large-neural-network-based model that can represent and simulate the behavior of molecules, cells, and tissues across diverse states.

DigitalCell.AI is allowing high-fidelity simulations, accelerating discoveries, and guiding experimental studies, offering new opportunities for understanding cellular functions and fostering interdisciplinary collaborations in open science.

Introduction:

ProBioDigital.AI constructs a digital cell models called DigitalCell.AI to simulate, predict, and steer cell behavior using computational model that simulates the biological functions and interactions of a cell.

It is a fully data-driven neural network-based representation of an AI based digital cell simulator that can accelerate research in biomedicine by enabling fast-paced *in silico* studies, as well as powerful bridges between computational methods and confirmatory wet-lab experimentation

A cell is a fundamental unit of life where every cell is a dynamic and adaptive system in which complex behavior emerges from a myriad of molecular interactions.

Bio cell simulation in physical lab are,

- Sensitive to even seemingly minor disruptions, such as a point mutation or an external factor that tips cells into dysfunction and disease.
- Remarkably robust **t**o alteration of the function of a biological system , such as the elimination of genes or their replacement with homologs from different species.
- With the availability of large preclinical datasets on cancer drug sensitivity and gene essentiality, computational biology models for predicting cancer sensitivity are gaining popularity. However, comparing these models proves to be a challenging task, as there are numerous published models and methods available, making it difficult to conduct meaningful comparisons without reproducing them on your own data.

Goal of Digital BioLab Simulation:

- DigitalCell.AI have the potential to revolutionize the scientific process, lead to the understanding of novel biological principles, and augment human intelligence to underpin future breakthroughs in programmable biology, drug discovery, and personalized medicine.
- Harmonize every bit of data and generate it where it really matters.
- Diverse data sources are boosted by an integrated lab-in-the loop system. Continuous iterations boost predictivity.
- Train parameters that govern signaling and are shared from cell lines to patients. This way our model generalizes to new experiments.
- Faster identification of novel targets, regardless of druggability or CRISPR-dependence with the help of AI based simulation using digital cell.
- Filtering causal hypotheses to enhance translatability of preclinical correlation to clinical efficacy and

accuracy.

- Build capacity to run drug modifier-like screens for every novel target candidate.
- Biomarker-driven target ID enhancing development potential and enabling target prioritization.
- Early identification of potential resistance mechanisms to reduce late-stage failure rates.
- It provides a benchmark that facilitates fair and meaningful comparisons of computational biology models in predicting cancer sensitivity.

The Literature Review:

Existing cell models are often rule-based and combine assumptions about the underlying biological mechanisms with parameters fit from observational data. Rely on explicitly defined mathematical or computational approaches, such as differential equations, stochastic simulations,10,11 or agent-based models.

Despite of various biological cell modeling system available, as of today the approaches has falled capturing many aspects of the operations of both bacterial and more complex systems, such as human cells.

The DititalCell.AI revolutions in science and technology, enable the construction of ditital cell models learned directly from data. A multi-scale, multi-modal, large-neural-network-based model that can represent and simulate the behavior of molecules, cells and tissues across diverse states. Simulated experiments are quick to design and easy to interpret.

DigitalCell.AI Methodology:

The AI models are self-consistent and propagate function across physical scales.Various modorn AI approaches are used for explaining model predictions, including causal modeling, sparse featurization, and counterfactual reasoning.

Create a detailed representation of each patient's cells by incorporating specific patient data, such as genetic sequences, single-cell profiles from blood, and tissue pathology images, along with additional clinical information from their health records. Consistent interactions between molecules measuring binding affinity, gene expression, cell-cell communication, or tissue organization .

It also have the capability to simulate the temporal evolution of alterations in cell states in response to both intrinsic and extrinsic factors, along with the resulting multicellular spatial arrangements.

1. **Collection of tokens:** A transformer neural network comprises of multiple transformer layers and the inputs are discrete pieces of information such as words, RNA molecules, or gene representations. Tokens use self-attention to integrate context from other tokens, enhancing their own representations, which are then processed through a feed forward network.
2. **Token Encoding:** Enable token encoding for transformers to process sequences, such as natural language, or biological sequences, such as DNA . This model predicts missing tokens in sequences, enhancing its understanding of contextual relationships within data.
3. **Image Analysis:** Use of deep learning model fora analyzing images. A multi layer that automatically and adaptively learn spatial hierarchies of features through backpropagation. Ability to detect complex patterns and structures, such as microscope images of cells and tissues. This technique is particularly useful in studying the complex interactions of different molecules or cell types within a heterogeneous tissue environment.
4. **Data Diffusion:** A generative deep learning models that generates high-quality, diverse samples across various domains. It operates by transforming images, text, cellular states through a process that mimics a physical diffusion process. A flow matching methods captures and generate sequences of data that reflect continuous transform which can be used to identify developmental stages of cells over time and space or the response of biological systems to treatments.

5. **Model Graphical Data:** In Graph Nurel Network, nodes are connected by edges which are representation of kind of biological data. Cells in a tissue form a graph and connected through edges of nodes represents how cells may pass chemical signals between one another.

6. **Mollecular Scales in a Digital cell:** It represent DNA, RNA, and proteins as sequences of character (Nucleotides or amino acids).An AI method originally developed for natural language processing, such as large language models (LLMs). A primary building blocks of these entities are atoms, a neural network trained to model molecules at the atomic level.

7. **Multicellular scale in Digital cell:** Profile the spatial location of RNA and proteins in cells, along with various imaging methods for select molecular species

Challenges in developing AI based digital cell:

Publicly available data are still limited, we foresee a rapid development in this domain providing multi-omic 2D and 3D datasets.

Need of generalized data generation effort with open frameworks could greatly accelerate modeling at the multicellular scale.

DigitalCell.AI in Medical Biotechnology:

Our model learns fundamental rules of biology to predict events it hasn't seen. Drug assays in heterogeneous populations of cancer cell avatars with representation of alterations driving acquired resistance

Mechanistic link between drug and biomarkers to ensure that the therapeutic effects in preclinical models are more likely to be replicated in clinical settings.

Patient relevant testable and actionable biomarker hypotheses to guide the next stage of benchtop therapeutic research. Broader clinical search space by generating unique patient representative avatars.

Technology Used:

AWS Cloud Computing, Microservices, API, Event Driven System, Python, Java, Scala, Spark, ETL, Data warehouse, Data Lake, JavaScript, ReactJS, TypeScript, Docker, Kubernetes, CI/CD, GIT, Swagger, Machine Learning (ML), Artifical Inteligence.

Result:

- 9 drug with gene alteration combination tested in 5 patient derived in avatars.
- Predicted 3 significant disease subtype associated biomarkers.
- Outperforms best in class models in predicting cell behavior on the phenotypic and molecular levels in experimental conditions the model has never seen.
- From 1 million simulations to actionable recommendations in 4 weeks. Integrating DigitalCell.AI, partner delivered key data package 2X faster.
- DigitalCell.AI simulation guided data generation gains the same model performance from 2x fewer data points than high-throughput data generation.
- Predicting cytotoxic drug response.
- Predicting post-KO gene expression.
- AI implementation in biobanks can identify new biomarkers, develop diagnostic strategies, and provide support in the selection of targeted therapies, ultimately leading to environmentally friendly hospital care with reduced costs and improved efficiency
- 2x faster from asking questions to validation with 40% less $ spent on data.

Conclusion:

Ability to learn patterns and processes directly from data without needing explicit rules or human annotation. DigitalCell.AI modeling can be used in biomolecular realm for example - Predict three-dimensional (3D) molecular structures from sequences.

The AI modeling methodologies will provide inference tools making the system being predictive, generative, and queryable, which are key utilities for advancing biological research and understanding.

DigitalCell.AI factors on business growth and Public Health:

DigitalCell.AI will have vital role in medical biotechnolgoy research sector by 2027 enabling AI in BioCell formulation, reducing physical biolab simulation and accelarating business growth and public health.

Business Growth:

The digital bio cell will faciliate biomanufacturing market size is projected to reach $2.4 billion in the current year 2025. Driven by improved data analytics, better product yield, faster production timelines, access to real time operations and rising interest in digital bio simular in the manufacturing sector amongst innovators, lucrative opportunities are expected to emerge for players offering bioprocessing services. The market size is anticipated to grow to $12.0 billion by 2035, representing a CAGR of 18% during the forecast period.

Reference:

1. https://www.rootsanalysis.com/reports/digital-biomanufacturing-market.html#:~:text=Market%20Size%20Analysis:%20Digital%20Biomanufacturing,18%25%20during%20the%20forecast%20period.

Public Health:

Digital.AI will enable personalized and preventive medical approaches, addressing challenges such as demographic change, healthcare accessibility, and sustainability.

Integrating AI and digital pathology enhances the speed, accuracy, and remote capabilities of pathology diagnostics relying on novel small devices which measures for 340 milion population in USA.

Reference:

1. https://www.globenewswire.com/news-release/2025/02/10/3023708/28124/en/Digital-Biomanufacturing-Market-Industry-Trends-and-Global-Forecasts-to-2035-Digital-Twins-is-the-Fastest-Growing-Segment-of-the-Digital-Biomanufacturing-Market.html

DigitalCell.AI references:

1. https://www.rootsanalysis.com/reports/digital-biomanufacturing-market.htm
2. https://www.nature.com/articles/d43747-022-00108-3
3. https://chanzuckerberg.com/science/technology/virtual-cells/
4. https://www.statnews.com/sponsor/2024/07/30/ai-virtual-cells-and-a-revolutionary-platform-for-biology/
5. https://cshlpress.com/default.tpl?cart=174469067321058113&fromlink=T&linkaction=full&linksortby=oop_title&--eqSKUdatarq=1282
6. https://pmc.ncbi.nlm.nih.gov/articles/PMC3489704/
7. https://turbine.ai/approach/
8. https://news.stanford.edu/stories/2024/12/scientists-call-for-all-out-global-effort-to-create-an-ai-virtual-cell
9. https://www.cell.com/cell/fulltext/S0092-8674(24)01332-1
10. https://pmc.ncbi.nlm.nih.gov/articles/PMC2711391/pdf/nihms121244.pdf
11. https://www.sciencedirect.com/science/article/pii/S0006349523002369
12. https://arxiv.org/abs/2409.11654
13. https://pubmed.ncbi.nlm.nih.gov/39672099/
14. https://www.calicolabs.com/publication/how-to-build-the-virtual-cell-with-artificial-intelligence-priorities-and-opportunities/