Towards Precision and Safety in regenerative Therapies: A Comprehensive pipeline for enhanced stem cell therapies

Yasmine Amor
DS Engineer
Esprit
Tunis, Tunisia
yasmine.amor@esprit.tn

Hadil Affes
AI Engineer
ISI
Tunis, Tunisia
affeshadil1@gmail.com

Jasser Dridi
ICT Student
ENIT
Tunis, Tunisia
jasser.dridi@etudiant-enit.utm.tn

Louai Azouni
AI Engineer
FST
Tunis, Tunisia
louai.azouni@etudiant-fst.utm.tn

Bilel Belhadjrhouma
AI Engineer
ENSI
Tunis, Tunisia
bilel.belhadjrhouma@ensi-uma.tn

Emna Ridene
Software Engineer
ISSAT
Tunis, Tunisia
Rideneemna88@gmail.com

Rabie Baccari
ICT Student
ENIT
Tunis, Tunisia
mohamedrabie.bakari@etudiant-enit.utm.tn

Safa Rafrafi
Software Engineer
Sesame
Tunis, Tunisia
safa.rafrafi@gmail.com

***Abstract*—Stem cells hold immense potential for regenerative medicine, offering the ability to repair or replace damaged tissues and organs, Drugs Toxicity tests on the cellular level, Or potentially Bio-printing specific cells, tissues and potentially organs. However, challenges such as controlling cell differentiation, The health of a stem cell colony, ensuring treatment safety,and optimizing cell production cycle limit their broader clinical application. In this paper, we present a novel approach that uses advanced artificial intelligence (AI) techniques to tackle such challenges, enhancing both the efficacy and safety of stem cell therapies.**

Our approach integrates Deep learning models to predict the health and the differentiation of stem cells, focusing on identifying key biomarkers that indicate successful lineage commitment. In This paper, we introduce an AI-driven framework that monitors and ensures the healthiness of stem cells prior to their application, thereby mitigating the risks of tumorigenesis and other complications.The next step is to optimize the production process of specific cell types, increasing yield and reducing variability meanwhile assuring the safety and the precision of the therapy.

The results of our study demonstrate the possibility of an improvement in the precision and reliability of stem cell-based treatments. The suggested not only streamline the creation of desired cell types but also offer a robust solution for advancing regenerative therapies toward clinical practice. This work lays the foundation for an all-purpose regenerative therapy platform, paving the way for safer, more efficient, and more personalized treatments in the future.

***Index Terms*—Bioinformatic, stem cells, RAG, RNA, Biomarker**

I. Introduction

The discovery of embryonic stem cells (ESCs) marked a pivotal moment in regenerative medicine, offering the unprecedented potential to differentiate into any cell type in the human body. This pluripotency made ESCs an attractive candidate for developing therapies to treat a myriad of diseases and injuries. However, the use of ESCs has been fraught with ethical concerns due to the need to destroy human embryos to obtain these cells, leading to significant debate and regulatory challenges. Additionally, the risk of immune rejection when ESC-derived cells are transplanted into patients has posed a substantial hurdle to their widespread therapeutic use.

In 2006, a groundbreaking advancement by Shinya Yamanaka and his team revolutionized the field: the development of induced pluripotent stem cells (iPSCs). By reprogramming adult somatic cells to a pluripotent state, iPSCs offered a solution to the ethical issues surrounding ESCs and reduced the risk of immune rejection since the cells could be derived from the patient's own tissues. This breakthrough opened up new possibilities for personalized regenerative medicine, where patient-specific iPSCs could be used to generate the needed cell types for therapy.

However, while iPSCs overcame some of the critical challenges associated with ESCs, they introduced new complexities. The reprogramming process is inefficient and can lead to

genetic and epigenetic abnormalities, raising concerns about the safety and stability of iPSC-derived cells. Moreover, the potential for tumorigenesis due to incomplete or improper reprogramming remains a significant barrier to the clinical application of iPSCs. These issues necessitate rigorous quality control and advanced techniques to ensure the safe and effective use of iPSCs in regenerative therapies.

To address these challenges, we propose leveraging advanced artificial intelligence (AI) techniques to enhance the process of generating and using iPSCs for therapeutic purposes. By integrating AI into the stem cell workflow, we aim to improve the precision of cell differentiation, optimize the production of specific cell types, and ensure the safety of the resulting cells. Our approach includes the use of machine learning models to predict and guide the differentiation pathways of iPSCs, reducing the risk of undesirable outcomes such as tumorigenesis. Additionally, AI-driven analysis can help identify biomarkers indicative of successful reprogramming, facilitating the production of high-quality, stable iPSCs.

This paper presents a novel AI-enhanced framework for stem cell therapies, focusing on overcoming the limitations of current iPSC methodologies. Our solution not only addresses the safety and efficiency concerns associated with iPSCs but also paves the way for more reliable and scalable regenerative therapies. By harnessing the power of AI, we aim to advance the field of regenerative medicine, making personalized and safe stem cell treatments a reality.

II. Landscape : explaining how to obtain stem cells

A. Origins of stem cells

1) Embryonic Stem Cells (ESCs) : Embryonic stem cells (ESCs) are derived from the inner cell mass of a blastocyst-stage embryo, typically created through in vitro fertilization (IVF). These cells are pluripotent, meaning they have the potential to differentiate into any cell type in the body. The process of obtaining ESCs involves the destruction of the embryo, which has raised significant ethical concerns. Opponents argue that since the embryo has the potential to develop into a human being, its destruction is morally unacceptable. This ethical dilemma has led to stringent regulations surrounding ESC research in many countries. For instance, in the United States, federal funding for research involving new ESC lines created after August 9, 2001, is prohibited. Such restrictions have slowed the development of ESC-based clinical therapies, limiting the exploration of their full potential in regenerative medicine [1].

2) Adult Stem Cells : Adult stem cells, also known as somatic or tissue-specific stem cells, are found in various tissues throughout the body, including bone marrow, adipose tissue, and the liver. These cells are multipotent, meaning they can differentiate into a limited range of cell types related to their tissue of origin. For example, hematopoietic stem cells from bone marrow can give rise to various blood cell types. While adult stem cells are considered less controversial than ESCs due to their non-embryonic origin, they present their own challenges. They are often rare in mature tissues, making them difficult to isolate and expand in culture. Additionally, their limited differentiation potential restricts their application in regenerative therapies compared to the broader capabilities of ESCs [1].

3) Induced Pluripotent Stem Cells (iPSCs) : Induced pluripotent stem cells (iPSCs) represent a significant advancement in stem cell technology. They are generated by reprogramming adult somatic cells, such as skin fibroblasts, to revert to a pluripotent state similar to that of ESCs. This innovative approach circumvents the ethical issues associated with embryo destruction, as iPSCs can be derived from the patient's own cells, significantly reducing the risk of immune rejection upon transplantation. However, the generation of iPSCs is not without its challenges. The efficiency of reprogramming varies, and there are concerns about the potential for tumorigenicity due to the presence of undifferentiated iPSCs in transplanted cell populations. Furthermore, the current methods for generating iPSCs often involve viral vectors, which pose additional safety concerns and regulatory hurdles. As research progresses, developing more effective and safer methods for generating purified populations of differentiated iPSC-derived cells remains a critical goal in the field of regenerative medicine [1].

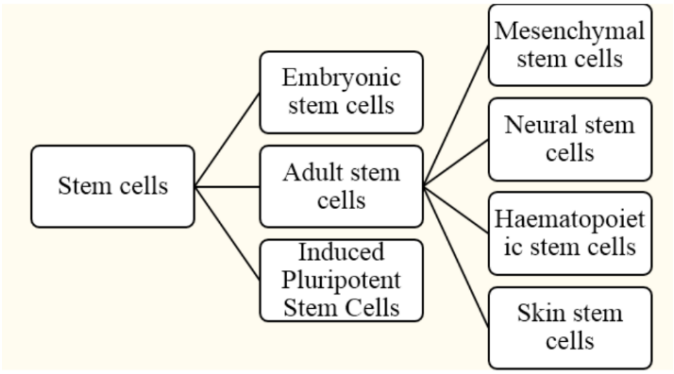


Fig. 1. **Classification of stems cells based on their origins**

B. classification of stem cells based on their potential

- Totipotent stem cells: These are the undifferentiated cells and are found in early development. These are omnipotent in nature and have the capacity to divide and differentiate into cells of the whole organism.
- Pluripotent stem cells (PSCs): Pluripotent stem cells are able to differentiate into cells that arise from the 3 germ layers — ectoderm, endoderm, and mesoderm – from which all tissues and organs develop. These form the cells of all germ layers, but not extra embryonic structures, such as the placenta.
- Multipotent stem cells: Multipotent stem cells are found in most tissues and differentiate into cells from a single germ layer. These can differentiate into specialized cells of specific cell lineages. For example- Hematopoietic

stem cells, which can develop into several types of blood cells.

- Oligopotent stem cells: Oligopotent stem cells are able to self-renew and form 2 or more lineages within a specific tissue. For example- A neural stem cell that can create a subset of neurons in the brain.
- Unipotent stem cells: These cells are only able to differentiate into one cell type. For exampleSpermatogonial stem cells [1].

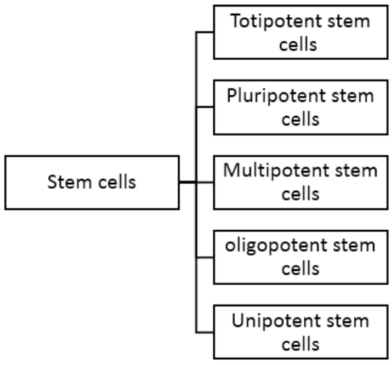


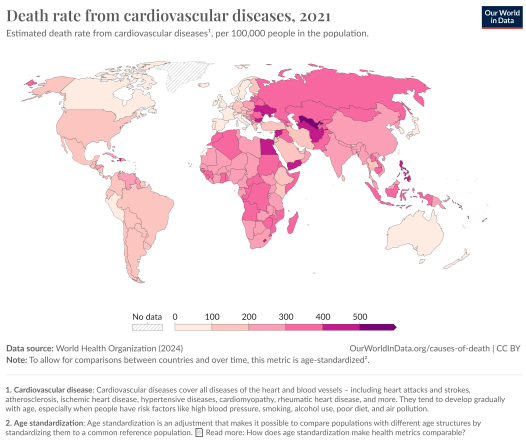
Fig. 2. **Classification of stem cells based on their potential**

III. Diseases related to the degenerative feature of organs

Stem cell therapy holds significant promise for treating a range of diseases, including heart failure, Parkinson's disease, type 1 and type 2 diabetes, and stroke, by addressing underlying pathologies that current treatments cannot fully resolve.

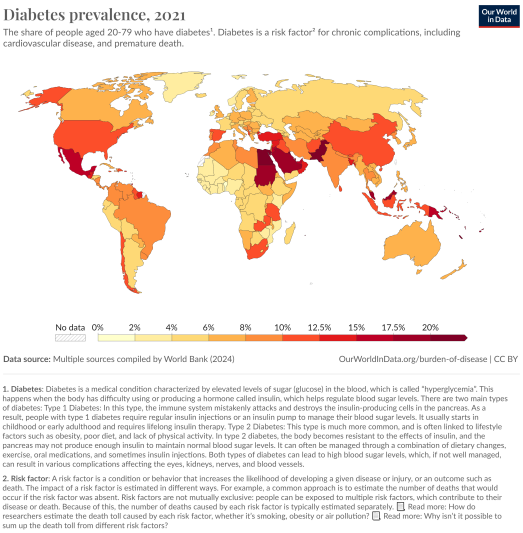
Cardiovascular Diseases and Stem Cell Therapy: Cardiovascular diseases, including heart failure and coronary artery disease, represent a leading cause of morbidity and mortality worldwide, affecting over 64 million people. Heart failure, often resulting from myocardial infarction, is characterized by the loss of viable heart muscle, leading to reduced left ventricular (LV) size and impaired systolic function. This deterioration is marked by a decrease in ejection fraction, sometimes dropping below 40%, significantly impairing the heart's ability to pump blood effectively. Additionally, reduced myocardial perfusion, or blood flow through the heart muscle, further exacerbates tissue damage and increases the risk of heart failure.

Stem cell therapy, particularly using mesenchymal stem cells (MSCs) and cardiac progenitor cells, offers a promising approach to regenerate damaged heart tissue, improve LV function, and enhance myocardial perfusion. Clinical trials have demonstrated that patients receiving stem cell therapy can experience a 5-10% improvement in LV ejection fraction and a reduction in infarct size by up to 30%. Furthermore, stem cell treatment has been shown to increase myocardial perfusion by 15-20%, which is crucial for maintaining the viability of heart tissue and preventing further deterioration. These improvements are associated with better clinical outcomes, including reduced symptoms of heart failure and enhanced overall cardiac function, offering a potential lifeline for millions of patients suffering from cardiovascular diseases [2].

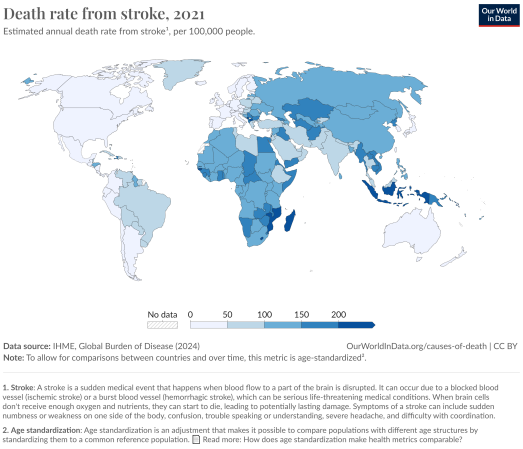


Diabetes (Type 1 and Type 2): Diabetes is a global epidemic, with over 537 million adults affected in 2021, and this number is expected to rise to 783 million by 2045. Type 1 diabetes is an autoimmune condition where the immune system destroys insulin-producing β cells in the pancreas. Without sufficient insulin, blood glucose levels rise, leading to a host of complications, including cardiovascular disease, kidney failure, and neuropathy. Type 2 diabetes, which is more common, is characterized by insulin resistance and eventual β cell dysfunction.

Stem cell therapy offers a potential solution by regenerating β cells and restoring insulin production. For example, iPSC-derived β cells have shown promise in early studies, with treated animals displaying normal glucose levels and C-peptide production—a marker of insulin synthesis. However, challenges remain, such as ensuring the complete differentiation of stem cells into functional β cells and preventing immune rejection. Clinical trials are ongoing, but early results suggest that stem cell therapy could significantly improve the quality of life for diabetes patients and reduce the burden of long-term complications [2].



Stroke: Stroke is the second leading cause of death worldwide, responsible for approximately 11% of all deaths, with 12.2 million new strokes occurring each year. A stroke occurs when blood flow to the brain is interrupted, either by a blood clot (ischemic stroke) or a burst blood vessel (hemorrhagic stroke), leading to significant neuronal damage. The brain's inability to repair itself after a stroke often results in long-term disability, with only about 10% of stroke survivors making a full recovery. Stem cell therapy is being explored as a way to repair this damage by promoting neurogenesis (the formation of new neurons) and angiogenesis (the formation of new blood vessels). Clinical trials have shown that stem cell therapy can improve recovery outcomes, with patients showing up to a 30% improvement in neurological function (as measured by the Modified Rankin Scale) compared to controls. Additionally, imaging studies have indicated increased perfusion in the affected areas of the brain, suggesting that stem cells may help restore blood flow and support brain tissue survival [2].



IV. Methodology

A. Health of stem cell colony

In the paper "Evaluating Cell Processes, Quality, and Biomarkers in Pluripotent Stem Cells Using Video Bioinformatics," the researchers utilized StemCellQC, a sophisticated video bioinformatics software tool, designed for the quantitative analysis of human pluripotent stem cell (hPSC) colonies. The software allowed for the extraction and evaluation of 24 distinct morphological and dynamic features from time-lapse phase-contrast videos of human embryonic stem cell (hESC) colonies over a 48-hour period. By continuously monitoring the colonies, the toolkit provided valuable insights into various key parameters such as growth patterns, cellular motility, and cell death events. These features were crucial in distinguishing between healthy, proliferating colonies and those that were unhealthy or on the verge of apoptosis. Specifically, the analysis revealed that healthy colonies exhibited consistent and robust growth rates, well-coordinated cell movements, and minimal cell death, while unhealthy or dying colonies showed irregular growth, aberrant motility, and increased rates of cell death. This kind of in-depth, real-time analysis offers a powerful method for assessing the viability and quality of stem cell colonies, which is essential for their use in regenerative medicine and research applications.

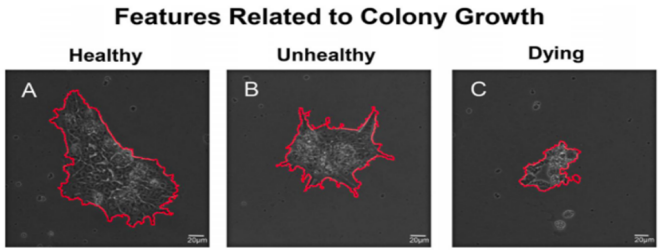


Fig. 3. features related to hESC colony growth.[3]

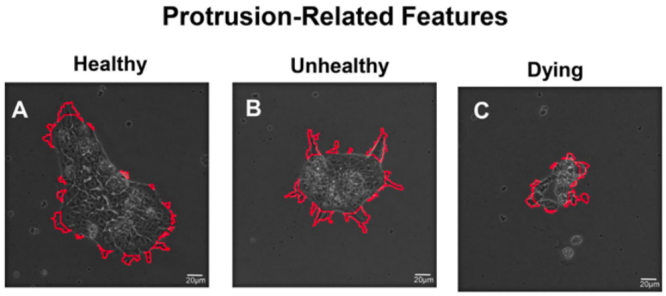


Fig. 4. Surface protrusions on colonies can be used to study cell morphology and growth[3]

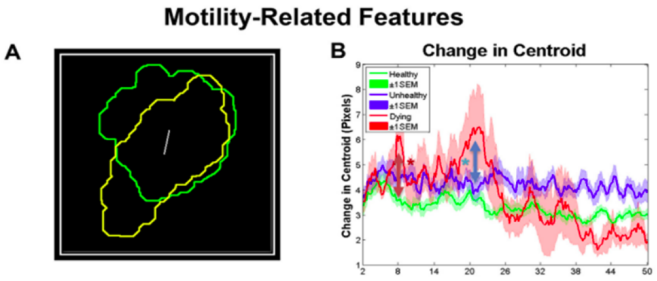


Fig. 5. Features related to hESC colony motility[3]

B. Biomarkers Indicating iPSC Differentiation

A critical aspect of understanding and controlling the differentiation of induced pluripotent stem cells (iPSCs) lies in the identification and characterization of specific biomarkers that reliably signal their future lineage commitments. The ability to accurately predict and influence the differentiation pathways of iPSCs is essential for advancing regenerative medicine and developing targeted therapies. Through our research, we have identified several key proteins that serve as robust indicators of iPSC differentiation pathways, providing insights into their molecular mechanisms and potential clinical applications.

One of the most significant discoveries in our study is the identification of **ROR2** (Receptor Tyrosine Kinase-like Orphan Receptor 2) as a pivotal biomarker. ROR2 has emerged as a critical indicator of iPSC differentiation into dopaminergic neurons, which are essential for neurogenic processes, particularly in the context of treating neurodegenerative diseases such as Parkinson's disease. Our findings demonstrate a strong correlation between the expression of ROR2 and the commitment of iPSCs to a dopaminergic neuronal fate, highlighting its potential as a target for enhancing neural differentiation protocols.

In addition to ROR2, our research has identified a range of other biomarkers, each associated with distinct differentiation outcomes:

- **OCT4** (Octamer-binding Transcription Factor 4): OCT4 is a master regulator of pluripotency in iPSCs. It plays a critical role in maintaining the undifferentiated state of stem cells. The downregulation of OCT4 marks the initiation of differentiation, as cells begin to exit the pluripotent state and commit to specific lineages. Monitoring OCT4 levels is crucial for understanding the timing and progression of differentiation, making it an essential marker in stem cell research.
- **SOX2** (Sex Determining Region Y-Box 2): SOX2 is another key transcription factor that works in tandem with OCT4 to maintain pluripotency. However, SOX2 also has a dual role in guiding iPSCs toward neural lineage differentiation. Its continued expression during differentiation directs iPSCs toward neural progenitor cells, making it a valuable marker for neural differentiation studies.
- **NANOG**: Like OCT4 and SOX2, NANOG is integral to maintaining the pluripotent state of iPSCs. The expression of NANOG diminishes as differentiation progresses, particularly as cells commit to early embryonic lineages. The downregulation of NANOG is a hallmark of the transition from pluripotency to differentiation, providing a temporal marker for tracking the differentiation process.
- **GATA4** (GATA Binding Protein 4): GATA4 is a transcription factor that drives the differentiation of iPSCs into cardiac myocytes and endodermal cells. It plays a crucial role in the development of the heart and pancreas, making it a key marker for studies focused on cardiac regeneration and endodermal lineage commitment. The presence of GATA4 indicates the activation of pathways leading to cardiac and endodermal differentiation, offering insights into the early stages of organogenesis.
- **PAX6** (Paired Box 6): PAX6 is a key regulator of ocular and neural development. Its expression in iPSCs is indicative of differentiation toward retinal and neural fates. PAX6 is particularly important in directing iPSCs to form specialized neural cells and retinal cells, making it a valuable marker for research into vision restoration and neural repair.
- **Nestin**: Nestin is an intermediate filament protein that serves as a marker for neurogenesis. Its expression in iPSCs signals the differentiation of these cells into neural progenitor cells. Nestin is widely used in neurobiology as an indicator of neural lineage commitment, providing a tool for tracking the development of neural tissues in vitro.

These biomarkers not only deepen our understanding of the molecular mechanisms that govern iPSC differentiation but also have practical implications for the development of targeted therapies. By leveraging these markers, we can refine iPSC-based interventions, ensuring that the cells differentiate into the desired cell types with greater accuracy and efficiency. This precision is crucial for the successful application of iPSCs in regenerative medicine, where the generation of specific cell types is necessary for treating a wide range of diseases and injuries. The ability to control and predict the differentiation of iPSCs paves the way for more effective and reliable therapeutic strategies, ultimately improving patient outcomes [4].

V. Pipeline

A. Classification model

1) Dataset and Feature Extraction: The dataset comprises time-lapse videos capturing the morphological and dynamical evolution of induced pluripotent stem cells (iPSCs) over a 48-hour period. Using QC stem cells software, we extracted 24 distinct features from these videos, which reflect the characteristics of individual stem cells within the colony at four specific intervals: 12, 24, 36, and 48 hours. The extracted data were structured into CSV files, with each file containing feature values corresponding to the individual cells at each time point [5].

2) Feature Selection and Classification: To evaluate the predictive power of these features, we constructed comparative tables for each time period. We selected features based on predefined criteria for each interval and subjected them to classification using Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Naive Bayes algorithms. The highest classification accuracy was achieved using the SVM classifier, reaching 82% after 48 hours of observation. The most discriminative features identified were the Number of Protrusions, Minimum Intensity, Area, and Change in Centroid.

Table 1. Classification Results Using 48 Hours of Video.

48 Hours	*Classification Techniques		
Single Features	**SVM**	**K-NN, k = 3**	**Naïve Bayes**
1) Area	94.12 ± 0.00	94.12 ± 0.00	94.00 ± 0.91
2) Number of Protrusions	90.71 ± 1.35	96.06 ± 1.32	91.24 ± 0.65
3) Total Distance Travelled	84.24 ± 1.20	74.06 ± 1.48	84.88 ± 1.01
Combination of Features			
1) Area, Orientation, Num. of Protrusions	94.12 ± 0.00	94.71 ± 1.15	94.12 ± 0.00
2) Num. of Protrusions, Min. Intensity	**97.06 ± 0.00**	**97.06 ± 0.00**	**96.47 ± 1.15**
3) Major Axis, Minor Axis, Change in Centroid	93.53 ± 1.57	92.94 ± 1.62	90.00 ± 1.27
Feature Selection Methods			
**CFS	91.76 ± 1.32	96.47 ± 1.32	91.76 ± 1.32
***Chi Square	91.76 ± 0.00	91.76 ± 0.00	95.29 ± 1.61
****QPFS	91.76 ± 1.32	94.12 ± 3.60	91.76 ± 2.46

*Classification of colonies as healthy or unhealthy using three different classification techniques: SVM, KNN, and Naive Bayes.
**CFS selected the following features: Area, Number of Protrusions, and Change in Area.
***ChiSquare selected the following features: Area, Number of Protrusions, and Major Axis Length
****Quadratic Programming Feature Selection selected the following features: Total Distance Travelled, Major Axis Length, Minimum Radius.

Fig. 6. **[1]:** comparative table of different machine learning models

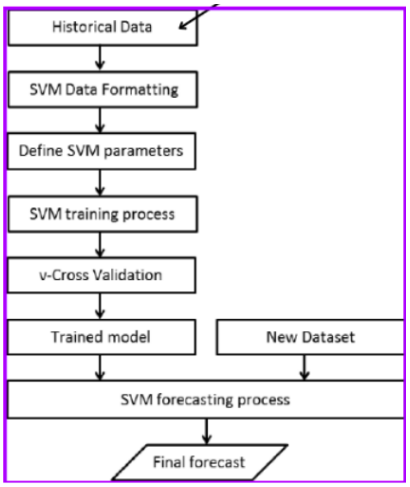


Fig. 7. **SVM model pipeline**

```
Accuracy: 0.8242271746944644
Precision: 0.829002514668902
Recall: 0.7762951334379906
F1 Score: 0.8017835427644914
```

Fig. 8. **SVM evaluation**

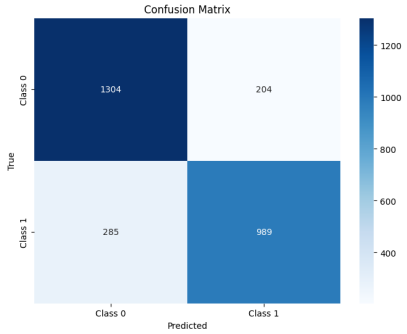


Fig. 9. **C**omfusion Matrix

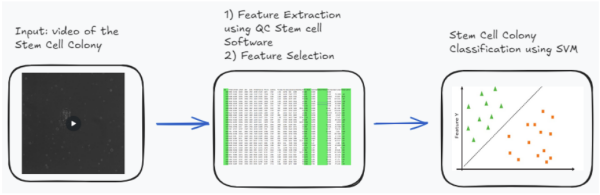


Fig. 10. **Stem cell health assessment pipeline**

B. Stem cell differentiation model

Predicting the differentiation of stem cells into specific cell types is a challenging task. This is why we've developed a Retrieval-Augmented Generation (RAG) model to predict the differentiation path of each stem cell. The process begins with converting RNA sequences into amino acid sequences using a Python script. Next, an advanced RAG model is implemented to perform protein matching based on sequence similarity.

Our stem cell differentiation model is built on a specialized embedding technique tailored for amino acid sequences. Inspired by the BLAST algorithm[6], which tokenizes sequences by sliding a window of three letters incrementing by one letter at a time, we developed an embedding model that processes amino acid sequences in a similar fashion. By breaking down the sequences into these overlapping triplets, our model captures the contextual relationships between amino acids more effectively. Each triplet is then embedded into a high-dimensional space, creating a unique vector representation for every sequence.

These embeddings are stored in a vector database, forming a comprehensive library of amino acid sequences associated with various cell types. When a new stem cell's RNA sequence, translated into its corresponding amino acid sequence, is queried, the model performs a similarity search within this vector database. The search identifies the top three most similar sequences, providing a shortlist of potential differentiation pathways for the stem cell.

To refine the predictions, these top matches are then passed to a large language model (LLM), which integrates additional biological knowledge to output a final result. This result includes a detailed analysis of the likelihood percentages for each potential differentiation pathway. By combining special-

ized embeddings, vector search, and LLM capabilities, our model enhances the precision and reliability of predicting stem cell differentiation, a crucial step in advancing regenerative medicine [7].

For example, consider the ROR2 protein, which plays a crucial role in skeletal development. When the amino acid sequence for a stem cell is input into our system, the RAG model might match it with the ROR2 sequence from our database. This match would indicate a high likelihood that the stem cell is differentiating into a bone cell, specifically contributing to skeletal development.

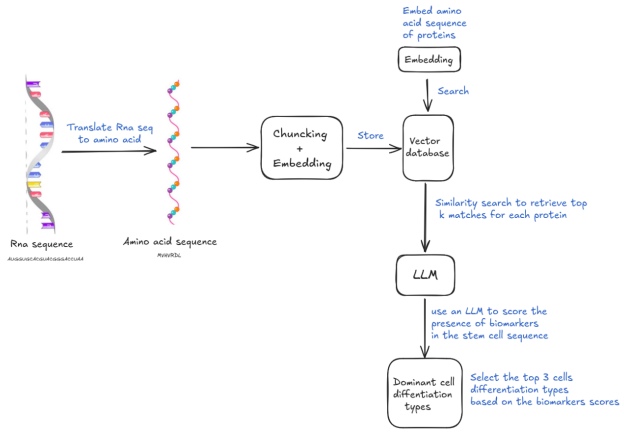


Fig. 11. **Differentiation model architecture**

Conclusion

In this study, we introduced an advanced artificial intelligence (AI) framework designed to enhance the precision and safety of stem cell therapies. By integrating deep learning models, we achieved accurate predictions of stem cell health and differentiation pathways. The AI-driven approach, exemplified by the StemCellQC toolkit, automated the assessment of pluripotent stem cell quality and lineage commitment, significantly reducing human error and improving the reliability of regenerative therapies. This is especially relevant in addressing challenges related to induced pluripotent stem cells (iPSCs) in clinical applications, such as optimizing cell production processes and identifying key biomarkers. One critical area where these advancements show promise is in the regeneration of pancreatic tissues, a vital step in combating type 1 and type 2 diabetes. For example, researchers have successfully generated insulin-producing pancreatic -cells from stem cells, offering a potential cure for type 1 diabetes by restoring insulin production (Pagliuca et al., 2014) [8] . Furthermore, innovations in 3D bioprinting have enabled the fabrication of pancreatic tissue structures using bioinks that support cell growth and differentiation (Ozbolat et al., 2017) [9]. These bioinks, composed of hydrogels and extracellular matrix components, mimic the pancreatic microenvironment and are used to print functional organoids, which can be implanted into patients or serve as platforms for drug testing. However, challenges remain, such as achieving long-term viability, preventing immune rejection, and scaling these technologies for widespread clinical use (Shah et al., 2020) [10]. AI can help overcome these barriers by optimizing bioprinting processes and ensuring that bioprinted tissues meet the necessary functional standards. In conclusion, the convergence of AI, stem cell therapy, and bioprinting represents a transformative approach to regenerative medicine, particularly in treating diabetes. As research progresses, these technologies will likely lead to more effective, personalized treatments that improve patient outcomes (Shah et al., 2020) [10].

Acknowledgment

This work was led in collaboration with Talan Innovation Factory, the RD Department of Talan Tunisia, in the frame of Talan SummerCamp2024 program which theme was, leveraging technological innovation to empower scientific research in biology.

References

[1] Panigrahy S, Sinha N, Patil S, Chaitanya TS, Kesigan JU, Tandon GD. Bioethical issues in stem cell research. *Glob Bioeth Enq*. 2019;8(1):23-30. https://doi.org/10.38020/GBE.8.1.2020.23-30.

[2] Nguyen PK, Nag D, Wu JC. Methods to assess stem cell lineage, fate, and function. *Adv Drug Deliv Rev*. 2010;62(12):1175-1186. https://doi.org/10.1016/j.addr.2010.08.008.

[3] Zahedi A, On V, Lin SC, Bays BC, Omaiye E, Bhanu B, Talbot P. Evaluating cell processes, quality, and biomarkers in pluripotent stem cells using video bioinformatics. *PLoS One*. 2016;11(2):e0148642. https://doi.org/10.1371/journal.pone.0148642.

[4] Schmidt M, Zeevaert K, Elsafi Mabrouk MH, Goetzke R, Wagner W. Epigenetic biomarkers to track differentiation of pluripotent stem cells. *Stem Cell Reports*. 2023;18(1):145-158. https://doi.org/10.1016/j.stemcr.2022.11.001.

[5] Visualization and Intelligent Systems Laboratory. Videos Dataset, 2017. https://www.vislab.ucr.edu/SOFTWARE/software.php.

[6] European Molecular Biology Laboratory. EMBL-EBI, Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK, 2023. https://www.ebi.ac.uk/ena/browser/view/ERR914288.

[7] UniProt Database. Protein sequences, 2024. https://www.uniprot.org/.

[8] Pagliuca FW, Millman JR, Gürtler M, et al. Generation of functional human pancreatic cells in vitro. *Cell*. 2014;159(2):428-439. https://doi.org/10.1016/j.cell.2014.09.040.

[9] Ozbolat IT, Hospodiuk M. Current advances and future perspectives in extrusion-based bioprinting. *Biomaterials*. 2017;76:321-343. https://doi.org/10.1016/j.biomaterials.2015.10.076.

[10] Shah K, Mrazek J, Hosseinkhani M, et al. 3D bioprinting and its applications in regenerative medicine. *Int J Bioprint*. 2020;6(2):257-268. https://doi.org/10.18063/ijb.v6i2.257.