Utrecht University

Literature review

Printing the Future:

On the Enhancement of Bioprinting Techniques Through Artificial Intelligence

Student: Anne Zijl, BSc
Student no.: 2817233
E-mail: a.zijl@students.uu.nl
Institution: Utrecht University

Examiner: Riccardo Levato, Dr. ir.
E-mail: r.levato@uu.nl

Daily supervisor: Gabriel Größbacher, MSc
E-mail: g.groessbacher@umcutrecht.nl

Second examiner: Sam van Beuningen
E-mail: s.vanbeuningen@umcutrecht.nl

Printing the Future: On the Enhancement of Bioprinting Techniques Through Artificial Intelligence

Abstract:

Interest in artificial intelligence has skyrocketed, being mentioned in over 2% of all publications on PubMed for the past two years. In this literature review, we illustrate the various ways in which artificial intelligence can support the field of biofabrication with a focus on bioprinting. Computer vision and machine learning are well-suited for various applications in bioprinting, ranging from 3D model design, to tuning printing parameters, to development of novel bioinks. The articles reviewed showed an overwhelmingly positive view of these and other possible utilisations of computer vision and machine learning, and the general view of the development of these technologies is one of excited optimism. That said, there are a number of issues, with data acquisition remaining the main hurdle within this field. Due to the nature of the bioprinting field, acquisition of data to such an extent that reliable ML models can be fashioned requires significant time and resources. This leads to many researchers running ML models with inadequate dataset sizes. A worrisome trend within the articles reviewed is that specifics such as number of datapoints, training : validation ratio, and validation methods are simply not mentioned. When these are mentioned, dataset sizes are often inadequate, and researchers often fail to note this fact. Instead, they opted to discuss the wider implications of their results without deliberating their credibility. This leads to difficulties when comparing results between research articles. This literature review posits that researchers should be required to make their ML model publicly available so that they may be scrutinised. In addition, we would like to impress that the importance of the creation of open-source databases containing data on bioink formulations, printing parameters, and print shape-fidelity would be an extremely useful tool for future research within this field. Furthermore, the research found has an understandable preference for extrusion printing, which has long been the most used printing method within biofabrication. However, the future of bioprinting is likely to feature printing methods with which higher resolutions are possible, such as volumetric bioprinting and 2-photon printing. Little research has been done on ML applications for these bioprinting technologies, although they are well-suited for similar applications as discussed in the reviewed literature. This would be an interesting avenue for novel research.

Keywords:

Artificial Intelligence, Biofabrication, Computer Vision, Machine Learning, Bioprinting

Layman's Summary:

Artificial intelligence, or AI, is a collection of techniques through which computers can learn to recognise complex patterns and draw conclusions from those patterns using machine learning (ML) models. A number of different variations and supporting technologies exist, which are expanded upon in this paper. These machine learning models are often supported through computer vision, which allows information to be extracted from visual data. ML capabilities are growing rapidly with the development of faster and more powerful hardware. This advent of AI technologies has a wide range of applications, including biofabrication.

Biofabrication is a branch of biotechnology that deals with the automation of tissue engineering, mainly focusing on bioprinting. The most widely used technology within biofabrication is extrusion bioprinting, a printing technique in which a printing nozzle extrudes the bioink (often a hydrogel) in layers to make a 3D object. Bioprinting, as essential as it is to biofabrication, faces a number of hurdles. Extrusion printing is a difficult process to get right, mainly because it is challenging to find the ideal printing parameters, bioink formulation, and the optimal printing path. This means that the printing process is often very slow, as researchers have to make their best guess at these parameters and often have to make multiple passes, before they achieve a result that is close to what they had in mind.

Because AI is so good at finding patterns in data, it is a logical solution to the problems that biofabrication is facing. As such, there have been a fair number of researchers who have investigated to what extent, and in which situations, AI can be integrated into bioprinting. This literature review takes a look at the existing research and aims to draw some conclusions on the applicability of AI within biofabrication overall. Generally, the outlook of AI within biofabrication is quite positive, and the realistic applications are wide. Although the results of the research reviewed were overwhelmingly positive, there are some issues to solve, with the largest being data acquisition. Almost all cases where ML was used, a significant positive effect was found on the aspect of bioprinting that the researchers were working on. There was no singular type of machine learning model which was found to be superior to any other between the articles reviewed, instead the selection of model type is highly specific to the application.

Unfortunately, the nature of biofabrication means that gathering data for AI to learn from has to be done manually, which is time consuming and expensive. This is mostly the case for applications such as the development of novel bioink formulations, as well as printing parameter optimalisation, and less so for cases such as optimising print trajectory. Some researchers made do with smaller datasets, expanding them by a prediction method called finite element simulation. A good solution in many cases would be open-source databases, so that researchers can combine their efforts. Although datasets used by researchers are often limited, researchers are clearly enthusiastic about the new possibilities AI brings them, and they are quick to judge their results as successful based on their models successfully finding patterns. However, it is often unclear how transferrable and replicable their results are due to the small dataset. This is often not discussed by researchers. A trend within the papers is that more time is spent on deliberating the biofabrication aspects of the research than on the specifics of the models utilised. This is unfortunate, as it decreases the replicability of the research, and prevents possible insights which may be gained from this information. In many cases, researchers fail to mention important information such as the number of datapoints used, the ratio of training to validation data, or data validation methods. This reduces the trustworthiness of their results. When dataset size *is* cited, researchers often fail to mention that the limited size of their dataset (usually below the standard threshold needed for the model to accurately reflect true patterns in data) affects the credibility of their results, and instead researchers seem more interested in discussing their implications without doubting their validity.

The research articles reviewed showed a bias towards extrusion printing, which is understandable considering this technology is used extensively within the biofabrication field. However, extrusion printing has limitations concerning print resolution, and technologies such as volumetric bioprinting and 2-photon printing are likely to make a come-uppance because of their superior printing resolution and a certain degree of freedom in print shape which extrusion printing lacks. Little to no research has been done as to the applicability of machine learning methods in supporting these types of bioprinting but they are expected to be well-suited for them. This could be an interesting avenue of future research.

Table of Contents

1. Introduction and background information:

Bioprinting is an innovative technology within the field of tissue engineering that enables the automated fabrication of three-dimensional structures. It holds great promise for various applications in regenerative medicine, including the eventual creation of functional tissues and organs for the purposed of transplantation, drug screening, and disease modelling (Matai et al., 2020).

Various 3D printing techniques are utilised for bioprinting, which address the complex requirements of tissue engineering. These include, among others: fusion deposition modelling (FDM), melt electro-writing (MEW), vat polymerisation-based printing techniques (stereolithography (SLA), digital light processing (DLP), and volumetric bioprinting(VBP)), ink jet printing, and powder bed fusion-based printing techniques (selective laser sintering (SLS), selective laser melting (SLM), and selective electron beam melting (SEBM; Bernal et al., 2019; Daghrery et al., 2023; Gupta & Meena, 2023; Sing et al., 2017; Ullah et al., 2023). Selection of 3D printing technique depends on the purpose, mechanical requirements, and material of the object to be fabricated. It is also important to consider whether cells should be added, as this reduces technique selection significantly as many techniques require high temperatures or exert excessive pressure on the bioink for cells to remain viable during/after printing. Currently, only inkjet, extrusion, MEW, and stereolithography-based bioprinting techniques have been successfully adapted to incorporate cells (Castilho et al., 2021; Kumar et al., 2021; Persaud et al., 2022).

Although bioprinting holds great promise with regards to regenerative medicine, the process through which these objects are currently produced is such that they are limited mostly to lab-based production, and the widespread production of i.e., personalised bioprinted wound dressings or replacement skin for burn victims in hospitals is not yet realistic (Shopova et al., 2023). This is likely due in part to the lack of standardised printing parameters and troubleshooting guides for various bioinks. 3D printing of non-cell-laden biocompatible materials has already been used to some extent in the creation of patient-specific implants for use in orthopaedics and maxillofacial surgery (Shopova et al., 2023). Each bioink composition and material combination may require specific printing parameters to achieve optimal results. Without standardised guidelines, the process of bioprinting becomes more time-consuming and requires extensive optimisation of printing parameters (Ruberu et al., 2021). The need for optimisation of printing parameters adds to the complexity and duration of the printing process. Researchers and scientists often engage in iterative experimentation and fine-tuning of parameters to achieve the desired print quality, cell viability, and functional outcomes. This optimisation phase can prolong the time required to produce bioprinted constructs, making large-scale production for clinical use challenging (Thattaruparambil Raveendran et al., 2019).

To enable the widespread production of bioprinted objects in hospitals and clinical settings, there is a need for the development of methods that may streamline the optimisation process in order to develop standardised protocols and printers that self-adjust their printing parameters based on material and environmental aspects. This would facilitate the translation of bioprinting technology into practical and accessible solution for personalised medicine and regenerative therapies.

In recent years, there has been a growing interest in leveraging artificial intelligence (AI), machine learning (ML), and computer vision (CV) techniques to enhance bioprinting processes (Ruberu et al., 2021). These advanced technologies offer new possibilities for improving the precision, efficiency, and outcomes of bioprinting, thereby accelerating the progress of regenerative medicine.

Machine vision, which involves the analysis and interpretation of visual data, can play a crucial role in monitoring and controlling the bioprinting process. By pairing imaging systems with competent algorithms, real-time feedback can be obtained on the deposition of bioink, cell alignment, and overall print quality. Machine vision enables the detection of errors, such as misalignment or inconsistent cell distribution, enabling automated adjustments or real-time feedback to optimize printing parameters (Liu, Yang, et al., 2022).

Artificial intelligence and machine learning techniques can further enhance bioprinting by facilitating process optimization and predictive modelling. AI can analyse large datasets, including experimental results, imaging data, and printing parameters, to identify patterns, correlations, and optimal printing conditions. This information can be used to create predictive models that optimise printing parameters, improve cell viability, and enhance tissue functionality (Shin et al., 2022).

Furthermore, ML can assist in the design and fabrication of bioprinted structures. By analysing vast amounts of existing data on tissue characteristics, mechanical properties, and cellular behaviour, these technologies can generate optimised designs and scaffolds. Such designs can consider the specific requirements of different tissues and incorporate factors like cell types, spatial organisation, and vascularisation to create more biomimetic and functional constructs (An et al., 2021; Sun et al., 2022; Xu et al., 2022).

This literature review aims to provide a comprehensive analysis of the potential ways in which AI techniques can contribute to enhancing the efficacy and speed of bioprinting processes. By examining the existing body of research and scholarly articles, this review seeks to elucidate the current state of knowledge and identify potential future directions in this emerging field, contributing to the broader understanding of how these technologies can be harnessed to improve the efficacy, quality, speed, and overall outcomes of bioprinting processes in the field of regenerative medicine.

2. General Overview of AI Technologies

Over the past decades, progress in computational power, data availability, and machine learning has fuelled significant progress in the development of artificial intelligence technologies. At their core, they simulate human-like cognitive functions such as learning, reasoning, problem-solving, perception, and language understanding using computer models that act as intelligent agents. One of the key drivers behind the creation of sophisticated AI systems is the advent of machine learning algorithms (Ketkar & Moolayil, 2021).

Machine learning (ML) is a branch of AI that encompasses the development of algorithms and statistical models which allow computer systems to improve their performance on a specific task through learning from vast amounts of data (Shin et al., 2022; Zaman et al., 2021). ML has applications in various domains, including healthcare and bioprinting. In the context of bioprinting, ML can be incorporated at different stages, addressing various challenges, which will be discussed later in this review.

In addition to ML, the exponential growth in computational power has played a significant role in advancing AI technologies. There is a trend, called "Moore's Law", which says that the number of transistors on identically sized computer chips doubles every two years on average (Debenedictis et al., 2017; Moore, 1965). High-performance computing, advances in graphics processing units (GPUs), and specialised hardware such as tensor processing units (TPUs), have made it possible to train complex neural networks, leading to the development of deep learning models (IEEE Computer Society et al., 2022). This type of model simulates human learning processes through a many-layered neural network, which progressively extracts higher-level features from the raw input data. Numerous types of deep learning models exist, including convolutional neural networks, transformers, and deep belief networks, and they may be supervised, unsupervised, or a mixture of the two (Ketkar & Moolayil, 2021).

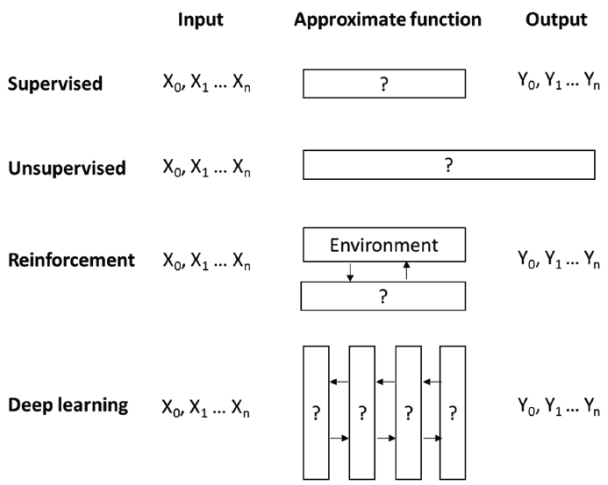

Figure 1. Methods of machine learning, taken from An et al. (2021). Each type of machine learning model takes input and transforms it into output. The process through which this transformation takes place depends on the type of model and typically involves a number of layers and functions.

The division between supervised and unsupervised learning is a basic but important distinction. Supervised learning models are trained on labelled datasets, where the labelled data acts as a teacher guiding the model's learning process. Classification and regression are important terms of prediction in the field of supervised learning and ML overall. Here, classifications predict data based on expected labels of the test data based on the labelled training data. On the other hand, regression makes predictions of validation data labels based on characteristics of the data (Shin et al., 2022). The model progressively adjusts its parameters in order to reduce the discrepancy between the predictions and the actual labelled output (Jo, 2021). Important to note is that the labelled data must be split into sets, one for training purposes, and one test set to assess the quality of the model's predictions. Without splitting the data

into these sets, it would simply learn the correct answers to each datapoint and may not yield adequate results when used on novel data. Performance of classification systems is usually measured by a confusion matrix (Figure 2), and high-performing classification systems feature relatively few false negatives/positives. Within the realm of biofabrication, classification systems have applications in post-processing analysis such as classifying tissues based on characteristics like cell type, density, and distribution (classification), or predicting and optimising parameters such as material viscosity, print speed, and temperature, in order to achieve desired results (regression).

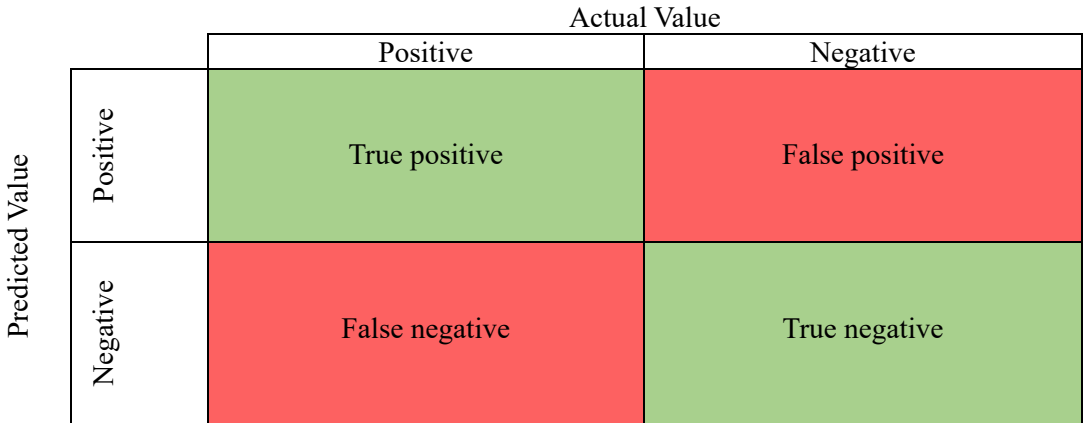

Figure 2. A confusion matrix, used to calculate model accuracy depending on the ratios between accurately and falsely predicted values. A highly accurate model will maximise true positives and negatives and minimise false positives and negatives. This type of matrix is also used when determining model sensitivity (the ability to find all positives).

Unsupervised learning, on the other hand, deals with unlabelled data, where the algorithm investigates relationships and structures within the input without any guidance. This is done to generate meaningful clusters and to reduce the dimensionality of the data (Razzaque, 2019). Unsupervised learning is already a common technique in bioinformatics. Principal component analysis (PCA), for example, has proven to be a useful analysis of genetic and metabolomic data, for example to visualise and identify patterns typical of certain cancers (Worley & Powers, 2013). This type of analysis is invaluable when working with complex data where explicit labels are unavailable, providing valuable tools for data exploration and pattern discovery (Jo, 2021). This may prove useful in the field of biofabrication in a number of ways, for example in the characterisation of tissue architecture, and the design of scaffolds based on these structures. It could also be useful in quality control and prediction of bioink, biomaterial, and cell behaviour.

Scoring or evaluating the performance of these ML models typically involves metrics such as precision, recall, F-1 score, mean squared error (MSE), and the area under the receiver operating characteristic curve (AUC-ROC), depending on the specific problem and learning type (Kamath & Liu, 2021). The formulas for these metrics can be found in figure 3. Important for the scoring of a ML model is that the model has successfully converged. This means as much as that the model has found its (local) optimum parameters, and subsequent iterations do not significantly change the outcomes of the model, the ML model has stabilised.

While supervised and unsupervised learning models are perhaps the most commonly used learning methods used in ML, other architectures exist. These include reinforcement learning (where the model learns through trial and error from interacting with its environment), semi-supervised learning (where the model generates a surrogate supervision signal without externally labelled data), transfer learning (training a model for one task, and leveraging that knowledge to improve

performance on a related task), and many others (Razzaque, 2019). The choice of the ML method and the stage at which it is incorporated in bioprinting depends on the specific goals of the project, the type of data available, and the challenges being addressed.

ML performance metrics	Equation
Precision	$\frac{TP}{TP + FP}$
Recall/true positive rate	$\frac{TP}{TP + FN}$
False positive rate	$\frac{FP}{FP + TN}$
Specificity	$\frac{TN}{TN + FP}$
F1-score	$\frac{2}{\frac{1}{Precision} * \frac{1}{Recall}}$
MSE	$\frac{\sum(y_i - p_i)^2}{n}$
AUC-ROC	Area under the ROC curve, which plots recall/true positive rate against the false positive rate.

Figure 3. Commonly used performance metrics for machine learning models and their corresponding equations. Here, TP = true positives, FP = false positives, TN = true negatives, FN = false negatives, y_i is the ith observed value, p_i is the corresponding predicted value for y_i, and n is the number of observations.

In general, we can divide the applications of the ML methods in the biofabrication process up into three categories: pre-processing, in-processing, and post-processing (see figure 4). An example of how these categories might be applied to real-life research can be found in figure 5. Here, pre-processing accounts for everything that happens before the bioprinter is turned on: from preprocessing data (cleaning up and preparing data used for bioprinting, ensuring data quality, normalisation, and standardisation) to 3D model generation. In-processing entails all methods applied while the bioprinter is on, including bioprinter optimisation (optimising printing parameters, such as print speed, nozzle temperature, and material deposition, to improve print quality and efficiency), automated printer calibration, and real-time monitoring (surveying print quality and detecting issues such as nozzle clogs or material inconsistencies to adjust printing parameters in real-time and to alert the researcher of any issues). Post processing applications mainly relate to quality control, analysing the printed tissue for structural defects, cell viability, and overall quality, as well as analysing and characterising the printed tissue's mechanical properties, cellular behaviour, and compatibility with the desired application. Dividing a ML-model up into these pre- in- and post-processing steps helps give an easily digestible overview of the research being done, an example of this can be found in figure 5.

Pre-processing	Bioimaging	Computer aided design	Cell selection	Biomaterial selection
Bioprinting	Extrusion bioprinting	MEW	Inkjet bioprinting	SLA
Post-processing	Maturation of bioprinted cells	Submerge method	Real-time monitoring	Bioreactor

Figure 4. A selection of factors that may be present in the different stages of bioprinting, adapted from Shin et al. (2022).

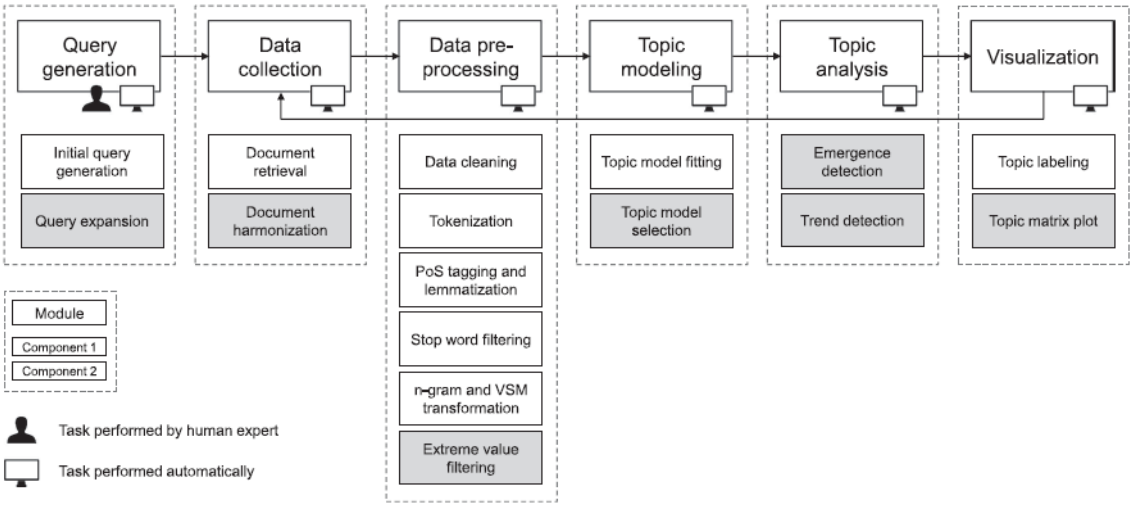

Figure 5. An overview of the steps used by the machine learning model developed by Muhlroth & Grottke, 2022.

Most biofabrication applications mentioned above rely at least in part on visual data. This is acquired by the system through a field known as computer vision (CV), which enables computers to extract meaningful observations from images and/or videos in a manner similar to humans (Liu, Liu, et al., 2022). This process starts with the acquisition of visual data by a sensor (i.e., a camera) which is then processed to improve its quality or to extract specific features (i.e., by utilising techniques and filters such as noise reduction, contrast adjustment, and edge detection). Afterwards, the CV extracts relevant features from the image that may be used for further analysis, these include features such as edges, shapes, colours, and patterns. Further analysis can be done through deep learning models such as convolutional neural networks (CNNs), which can identify objects, scenes, or patterns, and draw meaningful conclusions (Liu, Wang, et al., 2022).

CNNs are a class of deep learning algorithms designed specifically for structured grid data, such as images and videos. Characteristic of this type of neural network is the use of many layers, which automatically and adaptively learn to represent the input data in a hierarchical structure based on increasingly complex features and patterns (Karim et al., 2018). CNNs typically use three types of layers: convolutional layers which extract features using learnable features/kernels and represent them as a stack of feature maps (where higher layers represent features pulled from a wider context window), pooling layers which help reduce dimensionality and retain only the most important data from the feature maps (thereby preventing overfitting), and fully connected layers which process the extracted features and use them to make decisions based on the data (Karim et al., 2018).

In addition to CNNs, another type of neural network highly applicable to bioprinting is the Recurrent Neural Network (RNN), which is designed to process sequential data, where information is retained between cycles and past patterns are remembered by the system. This is unlike the most basic neural network architectures such as Feedforward Neural Networks (FNN) in which the information travels through the layers in only one direction (input layer → hidden layer(s) → output layer), where each layer is made up of nodes (neurons) connected to nodes in the subsequent layer (Svozil et al., 1997). In RNNs the hidden state is instead updated based on the previous hidden state(s) and the current data. This type of neural network is particularly useful in biofabrication applications that are

time-sensitive, such as real-time quality assessment and dynamic adaptation during the printing process.

As with learning types, there are many variations upon neural network architectures in addition to the widely used RNN, FNN, and CNN types. These include Long Short-Term Memory Networks (LSTMN, a variation on RNNs), Gated Recurrent Units (GRU, a simpler version of the LSTMN), autoencoders (unsupervised network used for dimensionality reduction), and many others (Karim et al., 2018; Mirzaei et al., 2022). However, in order to limit the scope of this review, we will limit ourselves to the neural network architectures most applicable to 3D bioprinting, focusing on print quality control and optimising printing parameters.

3. Scope and Methodology:

As the fields of regenerative medicine, biofabrication, and artificial intelligence are developing rapidly and evolving constantly, the publication range of articles used was limited to articles published within the last five years (2017-2023), only including peer-reviewed articles from before 2017 that were deemed exceptionally interesting or important to the field. In addition to evolving rapidly, these fields also have a wide range of applications. In order to reduce the scope, articles on CV, AI, and ML had to also mention applications in biofabrication to be included, unless the article was chosen to help illustrate a specific technique or concept.

Biofabrication and regenerative medicine utilise a large number of techniques for the automated production of tissues, including but not limited to de/re-cellularisation, microfluidic devices, self-assembly, electrospinning, and bioreactor systems. To reduce the scope of this literature review, and also because few other appropriate topics have been investigated in depth so far, topics discussed were limited to aspects of 3D printing with bioinks, as this technique is widely used in biofabrication applications and is thus highly relevant. In addition, while there are many techniques that fall under the 3D printing umbrella, the AI technologies discussed (such as those used for quality control) can be applied in many types of 3D bioprinting. While techniques such as electrospinning and electro melt writing may benefit from AI-enhanced quality control, the focus of the literature review is limited mostly to 3D printing methods which utilise bioinks, as these can have a high degree of variance in composition and behaviour and are thus more prone to irregularities that could be identified by AI.

Our search strategy involved using a combination of the following key terms, which were adapted and combined as needed to yield the most relevant results:

- Biofabrication
- Machine learning
- Computer vision
- Artificial intelligence
- 3D
- Bioprinting
- Bioink
- Quality control
- Cell viability
- Tissue engineering
- Bioprinter optimalisation
- Tissue constructs
- Scaffold design
- AI-enhanced tissue engineering
- Automated tissue production
- AI in tissue regeneration
- Deep learning

Our search encompassed variations and combinations of these terms, ensuring a thorough exploration of the literature. The search was conducted on the WorldCat catalogue, focusing on articles published within the specified time frame (2017-2023) and related to the fields of regenerative medicine, biofabrication, and the application of AI technologies in this context.

4. Review of Literature:

4.1 Interest in topics over time

Interest into AI and biofabrication is pictured in the form of PubMed citations in figures 6, 7, and 8. Data from before 1946 is sparse. This doesn't affect our graphs however, since the first hit for any of our search terms occurs in 1951 for artificial intelligence, with the other AI-related technologies all following within around a decade. The first hit for biofabrication occurs in 1985, and for bioprinting in 2000. By this time, papers discussing artificial intelligence and/or neural networks are in the hundreds per year, and it is no surprise that the first hit which mentions both an AI technology and either biofabrication or bioprinting occurs in 2003, only 3 years after the first mention of bioprinting. The first mention of 3D bioprinting by the Gartner Hype Cycle (a graphical presentation by the information technology firm Gartner, which represents the progress of emerging technologies) occurred in 2011, with the topic seeming to have been emerging since 2009, although the number of papers on the topic was fairly low (Muhlroth & Grottke, 2022).

From these graphs it is clear that interest in AI technologies has skyrocketed and has entered an exponential increase in citations since approximately 2016-2018, where AI in particular is mentioned in over 2% of all publications on PubMed for the past two years. The year 2022 was chosen as a cutoff point for the charts, as a flattening in the curve beyond this point may indicate that not all publications for the following two years may have been accessible for the tool or may not all have been published on PubMed (data was gathered for this graph in March of 2024), and as such it would be remiss to conclude that there would be a stabilisation in interest in certain topics. While at a much smaller scale, interest in biofabrication and bioprinting has been increasing as well, and bioprinting seems to have especially grasped the interest of researchers, showing a significant upturn in citations since 2015 (possibly related to the discovery of induced pluripotent stem cells in 2012 and their implications for the field). So too, there is an increase in interest in combining biofabrication/bioprinting with AI technologies, which roughly follows the shape of the bioprinting graph. This is logical, as bioprinting is a branch of biofabrication which more easily lends itself to optimalisation and pattern recognition through AI technologies. As an aside, we have added graphs showing only biofabrication/biofabrication and/or the combined search terms in order to better visualise trends in their data due to the stark difference in proportions compared to AI technologies.

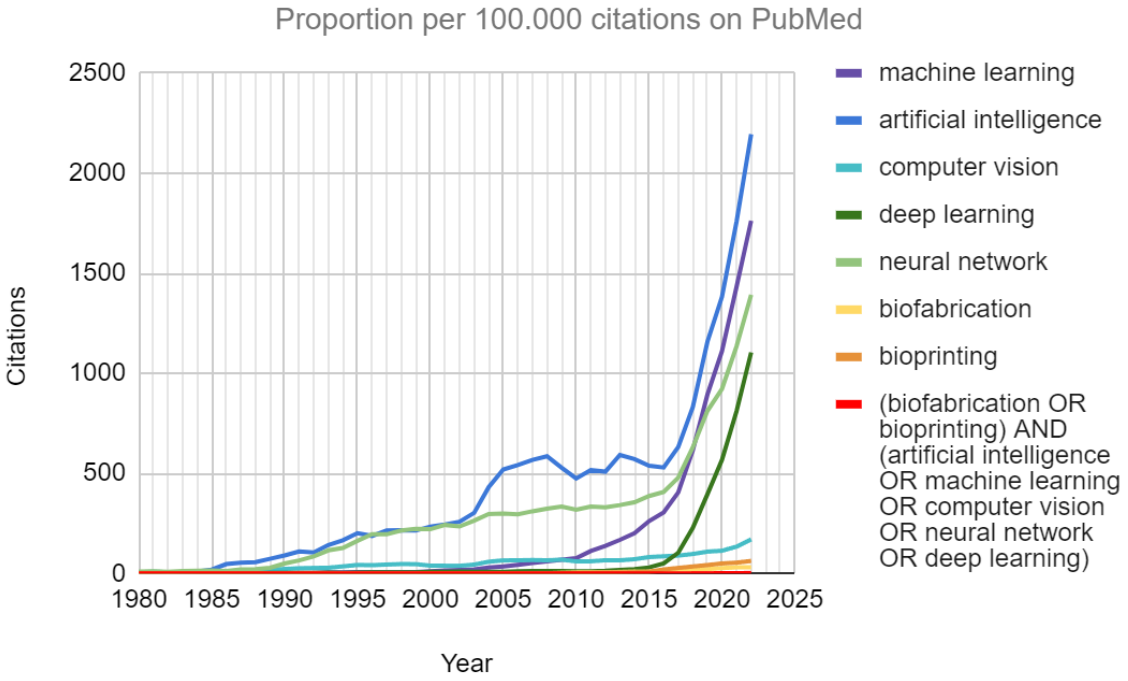

Figure 6. Graph developed through the tool "PubMed by year" showing the trends in citations of biofabrication and AI-related search terms. This graph and the ones in figures 7 and 8 were made using the tool "PubMed by Year", developed by Ed Sperr using open-source data from NCBI's PubMed, made available via GitHub (Sperr, 2016). They show the number of citations in the database that mention given search terms as a proportion per 100.000 citations, to generate a visual illustration of change in interest in given topics. Search terms were biofabrication (yellow), bioprinting (orange), AI-technologies (greens, blues, and purple), and papers combining either biofabrication or bioprinting with at least one AI technology (red). The database used by the tool is PubMed, which contains citations from approximately 26,000 journals, including articles from 1966 onwards (through the MEDLINE database) and from 1946 to 1966 (through the OLDMEDLINE database). For our graphs, we elected to have the time axis start from 1980 to limit the part of graph which would be barely distinguishable from 0.

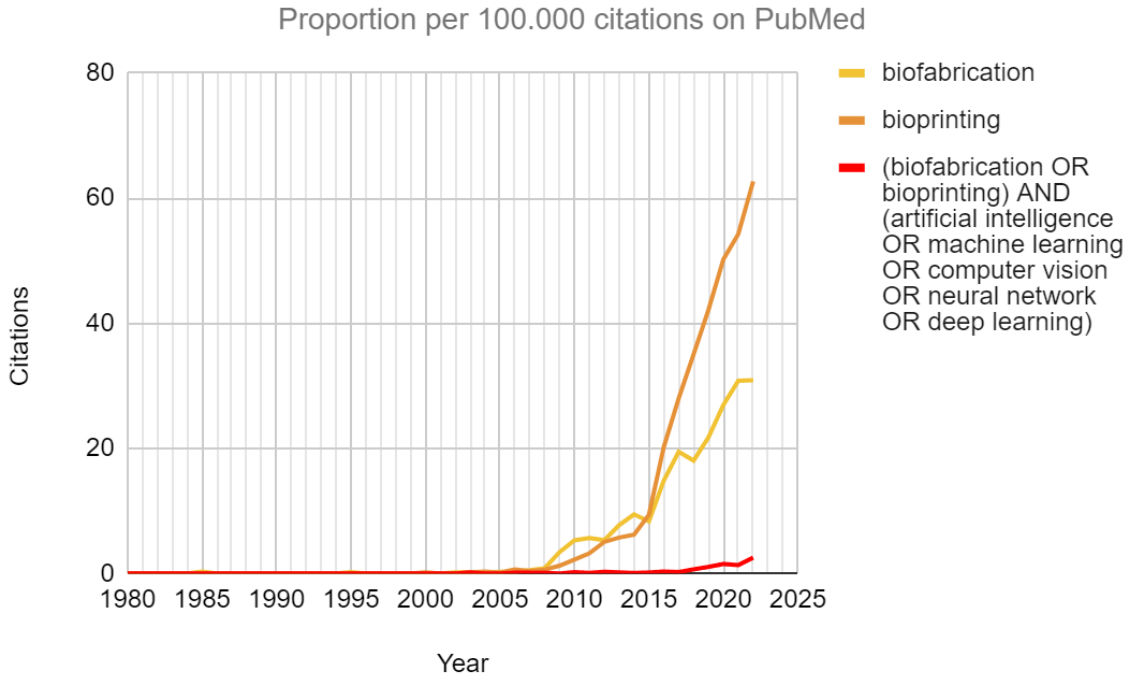

Figure 7. Graph developed through the tool "PubMed by year" showing trends in citations of biofabrication search terms as well as a combined search term attempting to show citations of papers which combine the ideas of biofabrication and an AI technology.

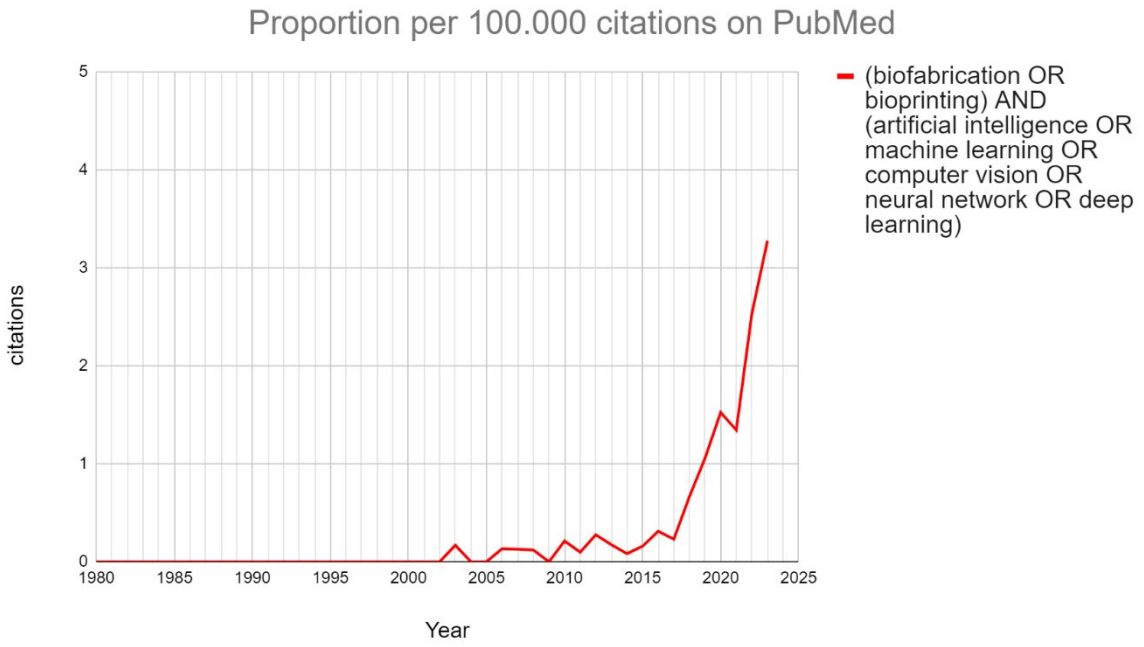

Figure 8. Graph showing solely the trend in citations using a combined search term attempting to find papers which combine the ideas of biofabrication and an AI technology.

4.2 Review

Please find a summarising table of the reviewed research articles in Appendix A for more information.

Main hurdle:

Many of the articles reviewed posited that the acquisition of data, whether that be the manual gathering of data or acquiring data via open-source databases, is the main hurdle that the field of biofabrication is currently facing (An et al., 2021). Another factor that was often cited as being one of the greatest challenges in 3D bioprinting, especially when concerning extrusion printing, is that of printing resolution and shape fidelity (Arjoca et al., 2023; Guan et al., 2022; Liu, Liu, et al., 2022; Liu, Yang, et al., 2022; Sun et al., 2022). As such, a majority of the research found was focused around integrating ML and CV methods into biofabrication to improve these two factors. Additional hurdles named were that of achieving optimal bioink formulation and optimising printing parameters (Chen et al., 2022; Freeman et al., 2022; Lee et al., 2020; Shin et al., 2022). The main solutions to relieving printing errors had to do with optimising print trajectories by adjusting print-axis corrections (Liu, Liu, et al., 2022; Liu, Yang, et al., 2022). Data acquisition for print trajectory was generally less fraught, whereas research that aimed to optimise bioink formulation and printing parameters generally found data acquisition more challenging. As a result, ML models used to optimise print trajectories more often had sufficient data to support complex ML models, whereas research that had issues gathering enough data (concerning bioink formulation and optimising printing parameters) such as that of Nadernezhad & Groll (2022) and that of (Xu et al., 2022), more often stuck to simpler models.

Success of research:

There is a wide variety in the ML methods used between the articles reviewed, and there doesn't seem to be one type of neural network (deep learning or otherwise) or other kind of ML model that is best suited for biofabrication applications in general. Instead, the choice in ML model to use is highly specific to the project and the individual needs and intricacies of the research.

Although the overwhelming majority of research reviewed found positive results using their ML method of choice, and researchers generally held a shared belief in the potential of ML to revolutionise biofabrication processes, there is a question of reliability of their results. Not all research had large sets of data to use. As a general rule of thumb, the ANN community requires the dataset to have at least 50-1000 times the number of prediction classes versus the sample size, for the sample size to be at minimum 10-100 times the number of features, or for the sample size to be at least 10 times larger than the number of weights in the network. Even so, the "factor 10" rule has been found to be lacking in conservativity, and a factor 50 rule would enable more sufficient performances (Alwosheel et al., 2018). This usually leads to models which use thousands of datapoints, and a few hundred datapoints is generally thought of as a small dataset, which has been shown to result in scaled errors of up to 15% if not accounted for, whereas datasets of at least a thousand datapoints showed scaled errors of 5% and below (Zhang & Ling, 2018). Unfortunately, many of the research articles which were reviewed did not reach an adequate dataset size. As can be garnered from the table in Appendix A, four articles failed to specify their dataset size, five had datasets below 500, and only four articles featured datasets of at least a thousand datapoints (of which one was achieved through finite element simulation). As such, while these ML models may have converged, there is a chance that patterns found by these models may not have been representative of the population. This is difficult to ascertain without access to the data and models used, but it is important to note that a ML model successfully finding patterns is not necessarily indicative of the model's success in accurately representing the data. This is not generally mentioned in the research articles reviewed which had to

deal with small datasets, although some (such as Xu et al., 2022) do mention wanting to increase their datasets through increased sampling in future research. The importance of this becomes clear when trying to replicate results in another research. If datasets are too small, patterns found may only be representative of patterns within that specific dataset, and not representative of a whole population (e.g., an AI might draw the conclusion that humans should not consume dairy if their sample contained a disproportionate number of persons with lactose intolerance). Several of the papers reviewed increased their dataset sizes through finite element simulation, which can be an appropriate and cost effective way to handle small dataset sizes if there is a careful validation process. Performance metrics such as precision, recall, F1-score, and AUC-ROC are invaluable in determining whether dataset size is sufficient. In addition, validation curves showing model performance (e.g., accuracy or error) can be used to determine whether additional data might be necessary, as a plateau in the validation curve may indicate that further data additions no longer significantly increase the model's performance.

Data:

A common issue among the literature reviewed was the challenge of acquiring sufficient data for the training and validation sets in order to generate reliable results from the ML model. This was even touted as often being one of the largest challenges when developing a ML for biofabrication purposes (Shin et al., 2022). A large dataset, generally containing at least 1000 data points, is crucial for a ML model to draw meaningful conclusions about patterns and variations present in the data (Shin et al., 2022). In addition, a large dataset is necessary to help prevent overfitting and ensure the model can perform well using novel data. Data was generally found difficult for a number of reasons, gathering visual data from bioprinting takes a significant amount of time and resources in the preparation of bioink, the printing itself, and the processing of the data (including computation time in the case of big data). Gathering data on biomaterial properties for the design of novel bioinks is also highly time-consuming and expensive, requiring the generation of large tables based on numerous rheological tests.

While some researchers worked with smaller datasets, some made use of simulated data. One example is the group of Sun et al. (2022), who generated a dataset with 6000 data points based on finite element simulations, although 2000 points were found to be sufficient to result in convergence around the 55th epoch of their ML models (a recurrent neural network and a convolutional network). Another group which went the simulation route was that of Guan et al. (2022), which generated simulated data for their research on compensating for the light scattering caused by cell-laden bioinks in digital light processing. They initially gathered data through fluorescent staining and imaging of printed scaffolds, and combined the data gathered with that of the masks used to generate simulated augmentation data to be used by their deep neural network. Both studies considered that simulating data significantly reduces time and resources spent on data gathering, without compromising on ML performance.

Other researchers didn't go this route, instead generating and compiling data manually, making do with smaller data sets. Xu et al. (2022), opted to combine results from a number of simpler ML systems to account for their lack of data (a mere 405 data points) in order to yield the most reliable results. Shohan et al. (2022), on the other hand, found a dataset of 660 data points to be sufficient to run a number of ML methods. However, this data was split 1:1 into training and validation sets, which would be deemed uncommon by Shin et al. (2022), who claim that large databases will usually dedicate 80% of data to the training set whereas smaller datasets might push towards 60-70% of the data to be used for the training set. The research group, however, deems the results of their research were positive as all models used were able to spot patterns in the data. Similar research methods were

used by the same group previously, to the same conclusion (Shohan et al., 2021). However, taking into account the small dataset and the method of splitting data 1:1, and given that ML methods will attempt to find patterns in data regardless of their existence, it is unclear whether the results found by this research group are reliable. The same goes for the research by Nadernezhad & Groll (2022), who generated a library of rheological data and printability scores for 180 different hydrogel formulations with just 13 rheological and printing parameter measures per hydrogel formation.

The group of An et al. (2021) states that it is unclear how ML will affect 3D bioprinting, and that two major factors will be Big Data and Digital Twin. ML models generally require large datasets, and complex ML models even more so than simpler ones, and these datasets are often not as readily available as they are for other applications of ML. One example is image-generation MLs such as Dall·E and Midjourney, which are both AI systems that can generate works of art and realistic images based on text-based prompts, using the vast number of images readily available on the internet. Vast databases such as these are not yet available for many biomedical applications and would help immensely in providing reliable data sources for researchers. The beginnings of such databases are already being worked on. Shin et al. (2022) make a particular note of the advantages of open source, and its benefits to the biomedicine research community. One recent example is a publicly available web-based nanomaterial database, which consists of hundreds of unique nanomaterials with annotated nanostructures providing nano-descriptors for use in ML. In addition, Digital Twins, being a cell-by-cell digital copy and model of human organs, could provide a significant support for ML applications in biofabrication research. The group of An et al. (2021) envisions a future of 3D bioprinting which is supported by ML and big data, through digital twinning and open-source databases. A graphical view of such a future as envisioned by their group can be seen in figure 9.

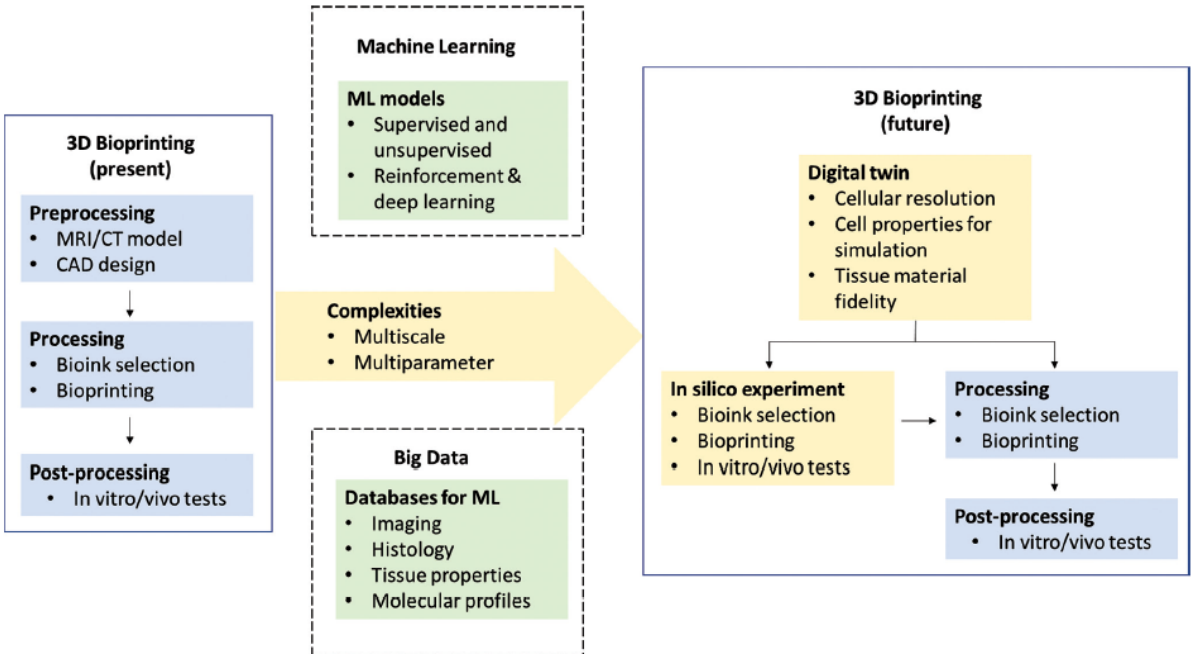

Figure 9. A future of 3D bioprinting, as envisioned by An et al. (2021). *Common aspects of 3D bioprinting research in the manner in which it is done currently are broken up into pre-processing, processing, and post-processing steps. Through the integration of machine learning and big data, and the various components thereof, An et al. (2021) envision a future of 3D bioprinting where pre-processing is fully supported by digital twins, and a sidestep may be made for in silico experiments, overall streamlining the process through which bioprinted models would be made.*

Choice of ML model

There did not seem to be a trend in which types of ML models were selected for the purposes of the different studies, although many of the researchers opted to compare between different types. In research which compared the performance of different types of ML model, there was no type of model which performed significantly better over different studies. This can mean that the optimal choice of ML model is highly dependent on the specifics of its implementation. Of course, this makes sense, as certain types of models have been developed to be specialised for certain types of data and certain types of patterns within that data. Even so, some of the studies we reviewed fed similar data to their models for similar purposes (increasing print shape fidelity) and found that different types of ML model performed best. In those cases, it is difficult to ascertain why one ML model performs better than the other when the type of data and purpose is similar between studies. It may be that this is due to the often-small datasets showing study-specific trends not reflective of the data as a whole. It may be that there are study-specific quirks that were not discussed as having an influence on the type of ML model that would be most applicable. It may also be that for certain ML model types, the choice between them doesn't matter all that much, or there are settings and variations to the model which may have varied between results (which weren't discussed). It isn't possible to tell for certain from the data examined alone and would require further research with larger datasets to verify.

5. Discussion and synthesis

Several papers discussed that the main problem with developing ML for biofabrication is the gathering of sufficient amounts of data. It has been proposed that open-source databases will be highly beneficial in this regard, as they have been in other applications of ML. What these papers do not discuss, however, is the likelihood of applicability of data between bioprinters. As bioprinters can vary greatly from machine to machine, it may be difficult to ascertain to what extent data is dependent on the specific bioprinter's features, and which parameters of the bioprinter have to be reported in order for data to be replicable. As such, there is a need for standardisation. This is also important when considering the replicability of rheological data of hydrogel formulations, although these have been more widely discussed in the past. When taking silk hydrogels as an example, not only does the processing of the silk (i.e., length of boiling time) have a great effect on the materials rheological and mechanical properties, so too does the choice of supplier. Biofabrication, by nature, is largely reliant on substances derived from biological origins, and as such there is great variation between samples and suppliers. This would be important to note in any open-source database.

Perhaps logically, there seems to be a trend in the papers analysed, that the focus of the research articles lies more on the biofabrication aspect and the results of the ML methods than on discussing the specifics of their ML methods. Researchers seem to be much more interested in relationships they find in their data and explaining and using those relationships for the development of improved biofabrication techniques. Many of the papers discussed only spend the bare minimum on discussing the specifics of their ML methods, 6/15 papers did not appropriately describe the number of datapoints used, 8/15 neglected to mention the ratio of training : validation data, and 8/15 did not use appropriate methods to validate the data (instead, 6 of these papers relied solely on assessing print quality). This is worrisome, as without this information it is difficult to assess the accuracy and trustworthiness of their results. Many of the papers which *do* cite the size of their datasets used a worryingly small amount of data, often much less than the standard minimum of 1000 datapoints that the average ML model needs in order to provide results which are reflective of true patterns in the data. This is understandable in part due to the biological nature of the data and the often difficult, time consuming, and expensive process of gathering data for the ML model to use, but the problematic nature of this lack of data is often not mentioned by the researchers as a deficiency in their research. This begs the question whether many researchers do not fully understand the importance of a large dataset in ML, and the implications on their results of using an inadequately sized dataset. In addition, the lack of exposition on the specifics of their ML models does not contribute to a greater understanding of the ways in which AI methods can be applied in biofabrication and doesn't allow researchers to find patterns in what may or may not work for their intended research. As such, although many of the reviewed papers discussed the benefits of an open-source database of rheological and printability data, we would argue that it may also be beneficial to the scientific community if researchers were to publish their models (via GitHub or another resource) to contribute towards a greater understanding of their research for readers, as well as a good jumping-off point for researchers wishing to further research the various ways in which AI technologies can contribute to biofabrication methods.

Another understandable but limiting focus of the reviewed research is that of printing method. The most used printing method within biofabrication has long been extrusion printing, and the table of articles in the review portion of this paper reflects that. It is difficult to find research that investigates the applications of AI technologies in biofabrication that does not make use of extrusion printing.

Some were found which discussed MEW, and a solitary article was found which discussed SLA and drop-on-demand printing, but these papers are in a very small minority. This is understandable, as the simplicity of extrusion printing and the issue of printing resolution mean that AI methods can be clearly and easily implemented in order to increase print quality. Because of the widespread use of extrusion printing and its simplicity, selecting this printing method for their research of applications of ML and computer vision in biofabrication is logical and a good proof-of concept, but extrusion printing is not necessarily expected to continue being the optimal method of 3D bioprinting. There are inherent limitations with the method, including its printing resolution. As such researchers have been adapting novel printing methods for biology such as volumetric bioprinting and 2-photon polymerisation (Bernal et al., 2019; van Altena & Accardo, 2023). As a proof-of-concept, the research reviews have been successful at showing that AI technologies can be implemented to spot patterns and improve print quality in extrusion printing. With regards to the future of biofabrication it would be interesting and valuable to show whether these technologies are applicable in other forms of bioprinting.

6. Conclusions and Future Research

In conclusion, many researchers view the future of biofabrication as a bright one with a myriad of possibilities. The main issue with biofabrication cited by many researchers in extrusion printing is that of printing resolution and print shape fidelity, which is dependent on many factors including printing parameters and bioink formulation. The traditional approach to addressing these challenges involved extensive manual fine-tuning, a process both time-consuming and resource intensive. The studies reviewed have shown that this can be alleviated, at least in part, through the integration of AI technologies, including CV and ML, in the optimalisation process.

The articles reviewed showed that researchers have an optimistic outlook on the use of AI for optimalisation processes, and initial research has shown positive results. However, due to the nature of biofabrication and the requirements of ML models, it remains difficult to acquire sufficient data to generate replicable results. In order to alleviate this hurdle, it would be beneficial to future researchers to set up open-source databases containing information pertaining to specific bioinks and their related print shape fidelity (including visual data), rheometric data, and printing parameters used. Of course, this type of database requires structured standardisation to make sure data between research is comparable. The generation of such a database would enable researchers in the field to collaboratively contribute and access valuable datasets, which would foster a culture of knowledge-sharing and accelerate progress in the optimisation of biofabrication processes. This collaborative effort could considerably mitigate the data acquisition challenge and propel the field towards more robust and universally applicable AI-driven solutions, which could conceivably allow greater scalability of biofabrication applications in order to bring them from the lab into the clinic.

In addition, although research thus far has logically focused on extrusion-based bioprinting, many bioprinting techniques exist which could benefit from AI integration. For instance, volumetric bioprinting utilised techniques such as holographic patterning and light-induced methods to print structures within a volume rather than layer by layer. This approach allows for faster printing speeds, higher printing resolutions, scalability, and scaffold-free printing. Similarly, 2-photon printing leverages precise laser-based techniques to achieve exceptionally high-resolution prints by activating photo-responsive materials at specific points within a three-dimensional space. Integration of AI within these bioprinting techniques could have similar applications as have been researched for extrusion printing, including optimalisation of the 3D designing process, optimising printing parameters, and analysing and identifying promising bioink formulations. The superior printing resolutions, printing speeds, and the option for scaffold-free printing provided by these techniques make them more future proof than extrusion printing, and it is likely that such techniques will become more widely used as the field of biofabrication evolves into the creation of more complex structures which require high resolution, multi-material printing, and complex architecture. As such, it would be beneficial for future research to investigate the possibility of AI integration into different bioprinting techniques.

7. References

Alwosheel, A., van Cranenburgh, S., & Chorus, C. G. (2018). Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. *Journal of Choice Modelling*, *28*, 167–182. https://doi.org/10.1016/j.jocm.2018.07.002

An, J., Chua, C. K., & Mironov, V. (2021). Application of Machine Learning in 3D Bioprinting: Focus on Development of Big Data and Digital Twin. *International Journal of Bioprinting*, *7*(1), 1–6. https://doi.org/10.18063/ijb.v7i1.342

Arjoca, S., Robu, A., Neagu, M., & Neagu, A. (2023). Mathematical and computational models in spheroid-based biofabrication. In *Acta Biomaterialia* (Vol. 165, pp. 125–139). Acta Materialia Inc. https://doi.org/10.1016/j.actbio.2022.07.024

Bernal, P. N., Delrot, P., Loterie, D., Li, Y., Malda, J., Moser, C., & Levato, R. (2019). Volumetric Bioprinting of Complex Living-Tissue Constructs within Seconds. *Advanced Materials*, *31*(42). https://doi.org/10.1002/adma.201904209

Castilho, M., Levato, R., Bernal, P. N., De Ruijter, M., Sheng, C. Y., Van Duijn, J., Piluso, S., Ito, K., & Malda, J. (2021). Hydrogel-Based Bioinks for Cell Electrowriting of Well-Organized Living Structures with Micrometer-Scale Resolution. *Biomacromolecules*, *22*(2), 855–866. https://doi.org/10.1021/acs.biomac.0c01577

Chen, B., Dong, J., Ruelas, M., Ye, X., He, J., Yao, R., Fu, Y., Liu, Y., Hu, J., Wu, T., Zhou, C., Li, Y., Huang, L., Zhang, Y. S., & Zhou, J. (2022). Artificial Intelligence-Assisted High-Throughput Screening of Printing Conditions of Hydrogel Architectures for Accelerated Diabetic Wound Healing. *Advanced Functional Materials*. https://doi.org/10.1002/adfm.202201843

Daghrery, A., de Souza Araújo, I. J., Castilho, M., Malda, J., & Bottino, M. C. (2023). Unveiling the potential of melt electrowriting in regenerative dental medicine. In *Acta Biomaterialia* (Vol. 156, pp. 88–109). Acta Materialia Inc. https://doi.org/10.1016/j.actbio.2022.01.010

Debenedictis, E. P., Badaroglu, M., Conte, T. M., & Tech, G. (2017). *REBOOTING COMPUTING Sustaining Moore's Law with 3D Chips*. https://doi.org/https://doi.org/10.1109/MC.2017.3001236

Freeman, S., Calabro, S., Williams, R., Jin, S., & Ye, K. (2022). Bioink Formulation and Machine Learning-Empowered Bioprinting Optimization. In *Frontiers in Bioengineering and Biotechnology* (Vol. 10). Frontiers Media S.A. https://doi.org/10.3389/fbioe.2022.913579

Guan, J., You, S., Xiang, Y., Schimelman, J., Alido, J., Ma, X., Tang, M., & Chen, S. (2022). Compensating the cell-induced light scattering effect in light-based bioprinting using deep learning. *Biofabrication*, *14*(1). https://doi.org/10.1088/1758-5090/ac3b92

Gupta, K., & Meena, K. (2023). Artificial bone scaffolds and bone joints by additive manufacturing: A review. In *Bioprinting* (Vol. 31). Elsevier B.V. https://doi.org/10.1016/j.bprint.2023.e00268

IEEE Computer Society, IEEE Computer Society. Technical Council on Test Technology, IEEE Solid-State Circuits Society, & EDA Association. (2022). *Proceedings of the 2022 Design, Automation & Test in Europe Conference & Exhibition (DATE 2022) : 14 - 23 March 2022, online, virtual platform*.

Jo, T. (2021). Machine learning foundations: Supervised, unsupervised, and advanced learning. In *Machine Learning Foundations: Supervised, Unsupervised, and Advanced Learning*. Springer International Publishing. https://doi.org/10.1007/978-3-030-65900-4

Kamath, U., & Liu, J. (2021). Explainable artificial intelligence: An introduction to interpretable machine learning. In *Explainable Artificial Intelligence: An Introduction to Interpretable Machine Learning*. Springer International Publishing. https://doi.org/10.1007/978-3-030-83356-5

Karim, Md. R., Sewak, Mohit., & Pujari, Pradeep. (2018). *Practical Convolutional Neural Networks : Implement advanced deep learning models using Python*.

Ketkar, N., & Moolayil, J. (2021). Deep learning with python: Learn Best Practices of Deep Learning Models with PyTorch. In *Deep Learning with Python: Learn Best Practices of Deep Learning Models with PyTorch*. Apress Media LLC. https://doi.org/10.1007/978-1-4842-5364-9

Kumar, P., Ebbens, S., & Zhao, X. (2021). Inkjet printing of mammalian cells – Theory and applications. In *Bioprinting* (Vol. 23). Elsevier B.V. https://doi.org/10.1016/j.bprint.2021.e00157

Lee, J., Oh, S. J., An, S. H., Kim, W. D., Kim, S. H., & Kim, S. H. (2020). Machine learning-based design strategy for 3D printable bioink: Elastic modulus and yield stress determine printability. *Biofabrication*, *12*(3). https://doi.org/10.1088/1758-5090/ab8707

Liu, C., Liu, J., Yang, C., Tang, Y., Lin, Z., Li, L., Liang, H., Lu, W., & Wang, L. (2022). Computer Vision-Aided 2D Error Assessment and Correction for Helix Bioprinting. *International Journal of Bioprinting*, *8*(2), 174–186. https://doi.org/10.18063/ijb.v8i2.547

Liu, C., Wang, L., Lu, W., Liu, J., Yang, C., Fan, C., Li, Q., & Tang, Y. (2022). Computer vision-aided bioprinting for bone research. In *Bone Research* (Vol. 10, Issue 1). Springer Nature. https://doi.org/10.1038/s41413-022-00192-2

Liu, C., Yang, C., Liu, J., Tang, Y., Lin, Z., Li, L., Liang, H., Lu, W., & Wang, L. (2022). Error correction based on computer vision method in extrusion-based bioprinting. *Materials Today: Proceedings*, *70*, 328–332. https://doi.org/10.1016/j.matpr.2022.09.246

Matai, I., Kaur, G., Seyedsalehi, A., McClinton, A., & Laurencin, C. T. (2020). Progress in 3D bioprinting technology for tissue/organ regenerative engineering. In *Biomaterials* (Vol. 226). Elsevier Ltd. https://doi.org/10.1016/j.biomaterials.2019.119536

Mirzaei, S., Kang, J. L., & Chu, K. Y. (2022). A comparative study on long short-term memory and gated recurrent unit neural networks in fault diagnosis for chemical processes using visualization. *Journal of the Taiwan Institute of Chemical Engineers*, *130*. https://doi.org/10.1016/j.jtice.2021.08.016

Moore, G. E. (1965). Cramming more components onto integrated circuits. *Electronics*, *38*(8).

Muhlroth, C., & Grottke, M. (2022). Artificial Intelligence in Innovation: How to Spot Emerging Trends and Technologies. *IEEE Transactions on Engineering Management*, *69*(2), 493–510. https://doi.org/10.1109/TEM.2020.2989214

Nadernezhad, A., & Groll, J. (2022). Machine Learning Reveals a General Understanding of Printability in Formulations Based on Rheology Additives. *Advanced Science*, *9*(29). https://doi.org/10.1002/advs.202202638

Persaud, A., Maus, A., Strait, L., & Zhu, D. (2022). 3D Bioprinting with Live Cells. In *Engineered Regeneration* (Vol. 3, Issue 3, pp. 292–309). KeAi Communications Co. https://doi.org/10.1016/j.engreg.2022.07.002

Razzaque, M. A. , & K. M. R. (2019). *Hands-on deep learning for iot : train neural network models to develop intelligent iot applications*. Packt Publishing, Limited.

Ruberu, K., Senadeera, M., Rana, S., Gupta, S., Chung, J., Yue, Z., Venkatesh, S., & Wallace, G. (2021). Coupling machine learning with 3D bioprinting to fast track optimisation of extrusion printing. *Applied Materials Today*, *22*. https://doi.org/10.1016/j.apmt.2020.100914

Shin, J., Lee, Y., Li, Z., Hu, J., Park, S. S., & Kim, K. (2022). Optimized 3D Bioprinting Technology Based on Machine Learning: A Review of Recent Trends and Advances. In *Micromachines* (Vol. 13, Issue 3). MDPI. https://doi.org/10.3390/mi13030363

Shohan, S., Harm, J., Hasan, M., Starly, B., & Shirwaiker, R. (2021). Non-destructive quality monitoring of 3D printed tissue scaffolds via dielectric impedance spectroscopy and supervised machine learning. *Procedia Manufacturing*, *53*, 636–643. https://doi.org/10.1016/j.promfg.2021.06.063

Shohan, S., Hasan, M., Starly, B., Shirwaiker, R., & Fitts, E. P. (2022). *Manufacturing Letters Investigating Autoregressive and Machine Learning-based Time Series Modeling with Dielectric Spectroscopy for Predicting Quality of Biofabricated Constructs*. www.sciencedirect.com

Shopova, D., Yaneva, A., Bakova, D., Mihaylova, A., Kasnakova, P., Hristozova, M., Sbirkov, Y., Sarafian, V., & Semerdzhieva, M. (2023). (Bio)printing in Personalized Medicine—Opportunities and Potential Benefits. In *Bioengineering* (Vol. 10, Issue 3). MDPI. https://doi.org/10.3390/bioengineering10030287

Sing, S. L., Wang, S., Agarwala, S., Wiria, F. E., Ha, T. M. H., & Yeong, W. Y. (2017). Fabrication of titanium based biphasic scaffold using selective laser melting and collagen immersion. *International Journal of Bioprinting*, *3*(1), 65–71. https://doi.org/10.18063/IJB.2017.01.007

Sperr, E. (2016). *PubMed by Year*. http://esperr.github.io/pubmed-by-year/

Sun, X., Yue, L., Yu, L., Shao, H., Peng, X., Zhou, K., Demoly, F., Zhao, R., & Qi, H. J. (2022). Machine Learning-Evolutionary Algorithm Enabled Design for 4D-Printed Active Composite Structures. *Advanced Functional Materials*, *32*(10). https://doi.org/10.1002/adfm.202109805

Svozil, D., Kvasnieka, V., & Pospichal, J. (1997). Chemometrics and intelligent laboratory systems Introduction to multi-layer feed-forward neural networks. In *Chemometrics and Intelligent Laboratory Systems* (Vol. 39).

Thattaruparambil Raveendran, N., Vaquette, C., Meinert, C., Samuel Ipe, D., & Ivanovski, S. (2019). Optimization of 3D bioprinting of periodontal ligament cells. *Dental Materials*, *35*(12), 1683–1694. https://doi.org/10.1016/j.dental.2019.08.114

Ullah, M., Wahab, A., Khan, S. U., Naeem, M., ur Rehman, K., Ali, H., Ullah, A., Khan, A., Khan, N. R., Rizg, W. Y., Hosny, K. M., Alissa, M., Badr, M. Y., & Alkhalidi, H. M. (2023). 3D printing technology: A new approach for the fabrication of personalized and customized pharmaceuticals. *European Polymer Journal*, *195*, 112240. https://doi.org/10.1016/j.eurpolymj.2023.112240

van Altena, P. F. J., & Accardo, A. (2023). Micro 3D Printing Elastomeric IP-PDMS Using Two-Photon Polymerisation: A Comparative Analysis of Mechanical and Feature Resolution Properties. *Polymers*, *15*(8). https://doi.org/10.3390/polym15081816

Worley, B., & Powers, R. (2013). Multivariate Analysis in Metabolomics. *Current Metabolomics*, *1*(1), 92–107. https://doi.org/10.2174/2213235x11301010092

Xu, H., Liu, Q., Casillas, J., Mcanally, M., Mubtasim, N., Gollahon, L. S., Wu, D., & Xu, C. (2022). Prediction of cell viability in dynamic optical projection stereolithography-based bioprinting using machine learning. *Journal of Intelligent Manufacturing*, *33*(4), 995–1005. https://doi.org/10.1007/s10845-020-01708-5

Zaman, W. S. W. K., Karman, S. B., Ramlan, E. I., Tukimin, S. N. B., & Ahmad, M. Y. B. (2021). Machine learning in stem cells research: Application for biosafety and bioefficacy assessment. *IEEE Access*, *9*, 25926–25945. https://doi.org/10.1109/ACCESS.2021.3056553

Zhang, Y., & Ling, C. (2018). A strategy to apply machine learning to small datasets in materials science. *Npj Computational Materials*, *4*(1). https://doi.org/10.1038/s41524-018-0081-z

Appendices

A. Table summarising literature

Paper	Summary	Type of 3D printing	Processing focus	Main hurdle	Success	Data-set size	Training : validation data	Data validation	Techniques	ML model
Abdullah et al., 2023	Research investigating ML-based image analysis to improve MEW print quality, with a focus on printed line width.	MEW	In	Extremely limited dataset, which affected ML model performance and applicability of results.	Success was limited. Models often didn't converge, and validation of models failed due to limited dataset. Maximum ML model accuracy was 93%	168	80:20	MSE, RMSE, and R-squared. Models were optimised with Bayesian optimiser.	Computer vision, ML-models	Linear, vector suppose machine (VSM), decision trees, neural network, gaussian process
Chen et al., 2022	This research paper investigated an AI-assisted high-throughput printing-condition-screening system.	Extrusion printing	In	Optimising printing parameters.	The optimised conditions devised by the AI system resulted in the generation of higher-quality constructs.	280 images	Not specified	Fourfold cross-validation, precision calculation. The model was ran four times using different subsets for training and validation.	Computer vision, ML (deep learning)	Neural network
Guan et al., 2022	Investigated the applicability of ML predicting and accounting for light scattering to optimize masks in DLP.	Digital light processing (DLP)	Pre	Light scattering in DLP, which makes printer optimisation difficult. Limited training set..	ML models (inc. deep learning) could greatly improve shape fidelity in DLP. In addition, the limited dataset was expanded using calibrated simulation	4000 data pairs from 32 printing samples	90:10	Printing quality comparison	ML, calibrated simulation (to generate more data)	NN

Huang et al., 2023	Research using ML to predict the number of cells in thermal inkjet bioprinted droplets.	Thermal inkjet bioprinting	Post	Acquisition of usable data, limited dataset.	Good accuracy. Random forest regression was best at detecting presence/absence of cells, extra tree regressor was best at predicting cell count.	156 droplets with 0-2 cells	60:40	Linear regression was used as a control	ML, computer vision	Random forest model, linear regression, support vector regression, decision tree, extra tree.
Lee et al., 2022	Research using ML models to design novel bioinks based on rheological and printability characteristics.	Extrusion printing	In	Design of biocompatible 3D printable bioinks.	ML techniques were successfully implemented in order to predict bioink predictability based on rheological factors.	25 hydrogel samples	74:26	A novel bioink was devised based on ML model results and assessed for printability.	ML	Inductive logic programming methodology often used for classification problem, supported by multiple regression.

Liu, Liu, et al., 2022	Research which developed a method of error detection in extrusion bioprinting to increase resolution.	Extrusion printing	In	Printing resolution, due to a lack of process control. This leads to inconsistent cell classification and survival rates.	Computer vision was successfully implemented in order to reduce print errors.	6 printed helices, each with an associated point cloud of data with unspecified size.	n/a	Print quality comparison	Computer vision	n/a
Liu, Yang, et al., 2022	Research which developed an algorithm which analysed errors in extrusion print trajectories using computer vision	Extrusion printing	In	Printing resolution. In addition, there was a waste in resources due to the lack of a NN in their experimental set-up.	The method showed a significant increase in print resolution.	Not specified	n/a	Print quality comparison	Computer vision, automated image analysis	n/a
Mieszczanek et al., 2021	Research using computer vision to optimise print settings for MEW bioprinting.	MEW	In	The technique used was not applicable for multi-layer scaffolds.	Research resulted in the successful identification of optimal printing parameters and printing of high-accuracy MEW scaffolds.	14580 data points per print, print count unclear.	n/a	Prints made with varied printing settings were compared to control group.	Computer vision, automated image analysis	n/a
Muhlroth &	Research which devised a	n/a	n/a	The main hurdle	3D bioprinting was deemed a particularly	770 papers	Not specified	Sensitivity analysis	AI-enabled	Supervised ML model

Grottke, 2022	method to assess the current status of an AI technique within the scientific community using an AI-enabled data mining model.			identified was the lack of data-driven support and automation in analysing changes in interest in subjects in the scientific community.	promising application of AI. The ML model was deemed successful.	and 434 patents (in the case of 3D-bioprinting)			data mining model, ML	
Nadernezhad & Groll, 2022	Research using a ML system to predict extrusion printability of novel bioink formulations based on a library of rheological data and printability scores.	Extrusion printing	In	Global criteria that predict printability may not exist, limiting the use of an open-source database for the formulation of novel bioinks.	The ML was considered highly accurate and successful. Patterns were found in rheological data which enabled printability prediction.	180 different hydrogel formulations, each with 13 rheological measures.	Not specified	F-score (formula calculating precision)	ML	Random forest model
Ruberu et al., 2021	Research investigating the applicability of ML as a tool for printability assessment and optimisation.	Extrusion printing	In	Predicting and optimising bioink formulation printability. Resource availability. Used visual	The method used was successfully implemented in order to find optimal printing parameters	Not specified	Not specified	Visual assessment and comparison of print quality.	ML	Bayesian optimisation framework

				assessment of images to score printability.						
Shohan et al., 2022	Research which used a number of different ML models to predict quality of extrusion printed GelMA constructs.	Extrusion printing	In	Progress in biofabrication methods needs to be supported by non-destructive quality engineering techniques	Previous research by Shohan et al. (2021) showed that ANNs could extract patterns in relative permittivity of bioprinted scaffolds subjected to dielectric impedance spectroscropy. This research showed that time series modeling based on ML-models give accurate predictions.	660	50:50	MSE	CV, ML	SVM, ANN, CNN, LSTM (LSTM performed best)
Sun et al., 2022	Research investigating applicability of ML and evolutionary algorithm-based approaches in 4D print design.	4D extrusion printing	Pre	Acquiring data from prints is difficult, so this research used finite element simulations to predict forward shape-change.	The ML model was found to be extremely successful at spotting patterns (R^2>0.999) and can be applied to different target shapes without having to retrain.	6000 data points	70% training, 15% validation, and 15% testing.	Compared RNN to CNN performance. Accuracy, predicting speed, MSE, RMS, and R^2.	Finite element simulation, ML	Recurrent neural network and convolutional neural network.
Venkata Krishna & Ravi Sankar, 2023	Research which investigates the application of ML in the biofabrication of personalised nerve guide	Extrusion printing	In	Inadequate data lead to inaccurate predictions.	ML techniques successfully found optimum composition of materials, fibre diameter, and neurotoxicity of additives.	Not specified	Not specified	Unclear. MSE and accuracy were mentioned once and twice, respectively. For the random forest model, it	Finite element simulation, ML	Logistic regression, ANN, kernel ridge regression, SVM, lasso regression

	conduits, with a focus on biomaterials.							appears that predicted results were compared to experimental results, but the researchers could have been more clear.		
Xu et al., 2022	Research using ML-models to predict cell viability in bioprinted GelMA	Drop on demand bioprinting and SLA.	Post	Physics-based models are unable to accurately predict cell viability.	The research used ensemble learning, combining different types of ML-models, to account for the limited dataset to successfully predict cell viability.	405	70:30, 80:20, and 90:10 were all evaluated	R-squared, relative error (RE), RMSE	ML, ensemble learning	NN, random forest, k-nearest neighbours, ridge regression