Machine Learning Techniques to Improve the Success Rate in In-Vitro Fertilization (IVF) Procedure

Patil N Sujata¹, S M Madiwalar² and V M Aparanji³

¹,² Department of Electronics & Communication Engineering, KLE Dr M S Sheshgiri College of Engineering & Technology, Belgaum, Karnataka, India
³ Department of Electronics & Communication Engineering, Siddaganga Institute of Technology, Tumukur, Karnataka, India
¹sujata.patil@klescet.ac.in, ²shwetamadiwalar85@gmail.com

Abstract. In Vitro Fertilization (IVF) usually assesses the embryo quality by visual morphological methods to transfer the potential embryo. But the success rate of IVF still remains low because of variations in selection process. The main objective is to improve the implantation rate by predicting the quality of embryos transferred from Day-2 to Day-3. Here using the Machine Learning techniques, thousands of the images trained together for the Day-2, the selection of embryos to come for the further assessment i.e. for Day-3. This will assist the doctors to check for the quality embryo without human intervention. We have also compared the results obtained by our Artificial Intelligence methods precision of >0.98 and also generalizes the method for potential embryo selection. Around 3000 plus embryo images are trained by CNN based Azure model and the results were validated using the Machine Learning techniques. Potentially viable embryo will help improve the implantation and success rate.

Keywords: Artificial Intelligence (AI), Blastocyst, Convolutional Neural Network (CNN), Embryo, In Vitro Fertilization (IVF), Machine Learning.

1. Introduction
In recent years, deep learning has changed the way images are handled by computers. In fact, it has brought in quite a revolution in computer science itself. Deep learning systems mimic working of biological brain. These systems are not programmed but trained to understand and interpret data, much like biological structures. The implementation of the paper work is to improve the success rate by selecting a potentially good embryo. This will be done if a growth of Day-2 embryo is better and the morphological features are defined precisely. The percentage of embryos go to Day-3 depends on the False positive rate of Day-2. Machine Learning technique is used to assess the rate of True Positive and False positive rate [6]. The total number of fragments detected can be neglected when they are considered for Day-3 growth. Here the Deep Learning will help us to predict these good quality embryos.
[5]. Combinations of these basic elements in certain specific ways tend to enable them to understand images, predict events, etc. Here we have discussed more on how to improve the success rate by knowing the false positive rate. If the number of embryo images from Day -2 is more than the rejected then the success rate will be improved. Deep learning has changed that. In this research, I have used such systems to understand and interpret embryo images. I have trained these systems with embryo images collected during this research and tested the accuracy of evaluation. Surprisingly the results are exceptionally good [5].

Machine learning and deep learning are related but not same. Experts differ in their definition. So I will record my opinion based on the historic developments in this field. In data science, machine learning refers to certain statistical operations on data to understand complex and big data [6]. These methods include linear regression, random forests, k-nearest neighbours etc., which are essentially statistical procedures. These have been used in data mining, recommendation systems, etc. which have been very effective on the internet. On the other hand, artificial neural networks use brain like structures to interpret data, similar to the way brain interprets data. They have been around for more than 50 years but there was no supporting hardware platform to prove their capabilities. Mobile revolution brought in the change. Mobile devices needed highly number crunching capacity, multi-core parallel execution, and had to work with extremely low electric power. This made large computing capacities to be built at data centres for relatively low cost. This combined with developments like Convolution Neural Networks (CNN) made image recognition [13] an easy task for them. Currently, compute power required to process data on artificial neural networks is available over Internet. Web services have evolved to support image learning capabilities. Currently we have used the web services provided by Microsoft Azure [6].

2. Literature Review

Pegha Khorsavi, Eshan Kazemi April 2019, suggested that the use of Artificial Intelligence (AI) method being trained with thousands of embryo images will precisely predict the quality of embryo image without actually needs to look into the development stages. Here they have implemented an AI approach based on Deep Neural Networks (DNNs) to choose the most viable and good quality embryos using a huge collection of human embryo time-lapse images (about 50,000 images) from a high-volume fertility center in the United States. They developed a framework (STORK) based on Google’s Inception model. Here the STORK predicts blastocyst grade with an AUC of >0.98 and generalizes well to images from other clinics outside the US and outperforms individual embryologists [18].

Chavez Badiola, AFS Farias, September 2020 have developed a model comprising of various morphological feature extraction algorithms using digital micrographs, along with other parameters for assessment to predict the viable or the highest grade embryo to predict the implantation rate. Here the system is evaluated using five dissimilar classifiers: Probabilistic Bayesian, Support Vector Machines (SVM), Deep Neural Network (DNN), Decision Tree, and Random Forest (RF), using a K-fold cross validation to assess the model’s generalization competencies. In the database A, the SVM classifier achieved an F1 score of 0.74, and AUC of 0.77. In the database B the RF classifier obtained a F1 score of 0.71, and AUC of 0.75. Their results suggest that the system is able to predict a positive pregnancy test from a single digital image [19].

Human Embryo Image Generator Based on Generative Adversarial Networks; here the authors D Dirvanauskas, R Maskeliūnas, V Raudonis Sensors, 2019 used human embryo images obtained during cell development processes for training a deep neural network (DNN). The proposed algorithm used generative adversarial network (GAN) to generate one-, two-, and four-cell stage images [20]. We achieved a misclassification rate of 12.3% for the generated images, while the expert evaluation showed the true recognition rate (TRR) of 80.00% (for four-cell images), 86.8% (for two-cell images), and 96.2% (for one-cell images).

Also CL Curchoe, CL Bormann in January 2019 have discussed about the use of AI & ML methods will help to predict the most viable embryo image which will lead to implantation rate [21]. In this work each and every aspect of patient carefulness was scrutinized, including sperm morphology, sperm identification, identification of empty or oocyte containing follicles, predicting embryo cell stages,
predicting blastocyst formation from oocytes, assessing human blastocyst quality, predicting live birth from blastocysts, improving embryo selection, and for developing optimal IVF stimulation protocols [5].

Here our methods discuss about the how the machine learning approach will help to assess the quality of embryo images from Day-1 through Day-5. The selection of viable embryo in Day-1 will lead to healthy development in Day-2 and Day-2 to Day-3 so on. Further the growth will be till the blastocyst stage. The percentage of rejection will be more in Day-1 and as it goes further the rejection rate reduces, the blastocyst stage will have more potential embryo selection. This will help to improve the success rate [5]. Most of the literature review discussed above use the Blastocyst stage as the main criteria for assessing the implantation rate by using machine learning techniques. But here in this work we have considered the assessment analysis from Day-1 to Day-5. Here the false positive rate of the respective day embryo images are calculated prior feeding it to the machine learning system for day-2 training process. This procedure will help to achieve the good success rate.

**Cleavage Stage**

Morphological analysis of the embryo images at the cleavage stage is multivariate and also complex so it is important to characterise the different features of the Day-2 or Day-3 human embryo parameters. In this stage the percentage of fragments will be considered as the number of cell counts so the fragments need to be segmented by morphological assessment. The cleavage stage will include other categorical parameters such as shape of the blastomere cells, cell division and number of cells in a radius of Zona Pellucida (ZP) (i.e. the outer radius of the cell which includes the blastocysts). The rate at which the blastomere cells divide can be observed in association with the morphological cleavage of embryo growth. The rate of true positive cells identified with Day-2 will carry forward to the next stage i.e. Day-3, this will help to improve the success rate of implantation [13]. Also the Day-3 embryo images will have 8-cells and positive true values at this stage will lead to Day-5. Machine learning technique is applied with the attributes considered in all respects of its assessment parameters so that the correct classification is done. As indicated in the figure 1 each of the embryo parameters for the respective days is shown with cleavage and the number of blastomere cells from (a) to (e).

![Figure 1:](image)

**Implantation Issues**

Implantation of the embryo at the women’s womb will involve the various physical parameters of the subject where it will be placed after fertilization. Figure 2 shows the embryo implantation procedure where the embryo will be placed into the women’s womb using a catheter. Usually the embryo is transferred in the blastocyst stage, but the observation for its development is done for all the five days. Figure 2 shows the stage wise process of transfer. In order to have at least one embryo to be implanted, generally the clinicians will transfer more than two embryos into the women’s womb. Usually the transferring of the embryo and how many healthy eggs to be transferred will be decided by the medical experts [6]. These decisions are based on patient's phase and other factors may be related to the fertility centre where the patient is of concern, the country, religious problems etc.
The transfer of multiple embryos will lead to complications for the mother as well as the baby. Sometimes it may present a greater risk for the women’s uterus. The health of the mother and baby may be seriously affected this causes the abnormal child, serious injury to the mother etc. So effective measures have to be taken by the infertility clinics to reduce the multiple transfers of the embryos. Automating the embryo selection process is essentially an image processing and image recognition task. It involves capturing a digital copy of the subject (fertilized egg, in our case) and applying various types of filters and feature extraction algorithms to understand the image. For example, we can use edge detection algorithms to highlight boundaries of organs in a biomedical image. Variation in the intensity and color of parts of image can help recognize organs. For example, the light intensity varies as a second order difference across a blood vessel but remains almost constant along the blood vessel. Algorithms to detect circles in an image are useful to detect organs with spherical, ovoid and ellipsoid shaped organs. In recent times, Machine Learning and Artificial Intelligence are changing the way we use computers for image processing.

IVF centre has supported with almost 500 embryo images comprising of around 100 images of each day. All the possible combination of the images is taken for the study of Day-2, Day-3 and Day-5. Normally the microscopic images have different magnification, illumination and also contain overlapping cells [5, 6]. These images were prepared for analysis by resampling and resizing to maintain some uniformity in the input data.

Images were divided into two datasets, one of which contained training dataset images. These images were chosen from all the class of sets containing the different category of images such as grades, fragments, sizes of blastomere, etc. The other set included the test dataset images which are used for
validating the software results. Around 70-80% of the dataset is taken for training the system and remaining 20-30% are for the test dataset.

Medical Image Analysis by using Machine Learning techniques are comparatively modern methods for the accurate grading of the embryos. To estimate the grading of the embryos basically these deep learning techniques make use of statistical properties of the data. The prediction rate achieved is more than 85% compared with existing methods of image classification [13]. We will get better result with the use of these statistical properties though there is a very minute difference in the cell boundaries. In many of the techniques discussed we find some error between expected output and applied input. The error is used to tune the machine learning system. Some of the methods popular methods include Regression, Support Vector Machine, Auto encoder, etc. Linear regression is useful for prediction as well as classification. The procedure involves identification of a geometric hyper-plane that separates the data into separate classes. Support vector machines also work on a similar principal but they calculate an exponent of the error so that the separation between the classes can be stretched for easy recognition. Auto encoders find a match between the applied input and existing templates. Implementation of a machine learning algorithm requires study of statistical methods for image processing.

Most of the classification algorithms are processed using the neural networks which help to improve the accuracy of the classifier. ANNs are mainly a pool of layered arrangements used as a scheme of consistent processing elements, intermittently called nodes, which resemble just as that of biological nerves or neurons by function [7]. The figure 4 shows the Neural Network system for classification of images. Here the embryo images are taken from the microscope, pre-processed to for magnification compensation, the training the dataset (i.e. 70% of total samples) with required number of samples then evaluation is done with training iteration. If the model is working for all the dataset, then freeze the network. The weights are being denoted by numerical values and by adjusting the values of these weights we are able to approximate the network with desired function. The weights can be adjusted systematically to draw the specific function [13]. The layers will be fully connected and the weights will be continued further to add on to the final result of the classifier.

![Figure 4: Neural network training system](image)

Here every node connection in the fully connected network proceeds with several different access points and then estimates the result based on a single neuron. Here the output of this node is fed to the next neuron for processing and repeating this procedure will give the final output. The participation layer obtains the estimated inputs and the output layer yields a desired output. Appropriately there is a distinction between each layers in the neural network, the middle layer are called hidden layers.

The Neural network system used to train here involves acquiring the microscopic images; pre-process those using the techniques described earlier [17]. Training process involves uploading of pre-processed images and trains them for little iteration. Then supervise and test it for the obtained dataset. Generally, the deep learning models will directly act on the unprocessed data so the convolutional neural networks help for feature extraction. The Machine Learning techniques will help to predict the grading and proving to be a highly accurate validation process for the prediction of the implantation rate. Figure 3
shows how the neural network training can be used as a multilayer perceptron (MLP) model which are designed to have minimum pre-processing. After the convolutional layer the pooling layer which is the maximum pooling layer, the fully connected will have all the activations connected to the previous layer [8-10].

Figure 5 shows the embryo images are given as a input to the CNN model that convolve in an effective manner to produce the relevant data set for the training as well as testing from the 525 embryo images. The result show how these CNN based classifiers will help to improve the success rate of implantation. Our model is able to predict the grading for the respective days of the embryo images with an accuracy precision of 88% and the network to be able to recall at the rate of 87% for the all the features of the embryo images considered.

As indicated in the figure 6 the position of the embryo cell is not regular it may be placed any corner of the zygote and also we observe that there exist undistinguishable intensity variations. In (b) as seen from the figure 6 sometimes there will be three cells instead of two for day 1.5 thus it may lead to misclassification. So in order to clearly distinguish between the viability of the embryo the false positive cells have to be eliminated at each stage so that it helps to reduce the chances of misclassification.

3. Application of AZURE API
Microsoft also provides an API for the analysis of images with our own attributes; here we can give our features of respective day embryo images for the CNN model. This API helps us to develop the system according to our requirement, test the model for all the available parameters and also we can validate the result with our Machine Learning algorithms developed. Microsoft AZURE uses the Representational State Transfer (REST) API to train the CNN model. Here we can train the model with
the embryo images day wise and validate the result. Before feeding it into the system the images are pre-processed and resized to 255x255 pixel size so as to increase the efficiency of the model. The quality of the embryo images are retained so that there is a clear distinction between the features. The features of the embryo images are provided with attributes for the day-1 it is the nucleoli and alignment of nucleoli at the pronuclear junction, day-2 or day-3 it is shapes, sizes and the number of fragments with respect to cleavage stage, for day-4 it is morula stage and finally for day-5 it is Trophectoderm (TE) membrane, Inner Cell Mass (ICM). Allowing these to assess the parameters in a appropriate way will help to improve the success rate of implantation.

4. Results & Discussion
Here the results show that refer Table 1, in case of Day-3 embryo images the acceptance percentage is 80.85% which is more than the traditional algorithmic approach. Also the embryo images considered from Day-2 to Day-3 is 76% which indicates that the embryo will be implanted in the blastocyst stage is 80.85%. This will help to assess the quality of embryo to be implanted in Day-5. Even the success rate is comparatively better with Machine Learning Technique. The system involves training with CNNs and Deep Neural Networks (DNN), as a result of which the building of the network starts with the interconnection of neurons as nodes. Training of the network is done rigorously by uploading the embryo images, assign the tags to them with given features, then look for the iteration to mark. Once the iteration is marked the network is ready to use for the validation of results. Work described here focuses on new techniques for human embryo image classification based on a training sequence for the convolutional neural network system.

The results clearly depict as to how the machine learning method for the assessment of the human embryo images helps to predict the grade of the embryo. The AI based CNN model is used to classify these embryos and is encouraging, in particular considering that they have been obtained using a small training set with very few positive samples. For pro-nucleated (initial stage of embryo development) oocytes and embryos, it is probable that other type of descriptors not specifically designed for prevalent textural images might be used. In fact, the alignment and the number of nucleoli or the number and size of blastomeres are not textural features. The two possibilities (dynamic and static observation) might be used together and integrated in a more thorough analysis for practical aims in a normal IVF clinical setting.

In a recent paper discussed by Sarah Armstrong, Priya Bhide that the Time Lapse Imaging Systems will assist the embryologists to assess the quality in a better manner [1]. The Time Lapse Imaging Systems will give the thousands of embryo images with each day assessment as a measuring parameter. But they are much expensive and prone to error as there is no control on assessment. Hence it’s better to select the viable embryo by predicting the growth of Day-2 embryo images by discarding the false positive cells.

5. Conclusion
Here the machine learning techniques used for the embryo image classification to eliminate the false positive samples of each day is evaluated and it is seen that it helps to select only the most viable embryo images for the further days. Hence the overall efficiency of the classifier is improved with proposed method. In this paper the machine learning algorithm is trained with more than 100 plus images for each day and the rejection rate for day-1 was 28%, for day-2 it was 17%, day-3 12% and for the final stage blastocyst stage is was found to be less than 5%. Measurement of the accuracy is evaluated by the ration of number of true cells identified to the total number of cells fed to the machine learning model. In the literature review mentioned in section 2 most of the research is done on the blastocyst stage where the embryo features are clearly defined but whereas in the proposed method we have discussed the machine learning algorithm which will assess the quality of the embryo images for the next day to be evaluated. Here the accuracy of the system is increased and also the success rate of the embryo images by more than 85%.
The above table 1 shows percentage of the embryo images accepted and rejected according to the respective days. Above results shows that in case of Day-1 embryo images the acceptance rate is high because the features are not defined clearly. As it goes for the further development the acceptance rate reduced as number of fragments with respect to the abnormal embryo images will be more. Here each fragment is considered as a cell and it will be counted for further assessment. The sample result for the Day-3 and Day-5 embryo images is as shown in the table above. The blastocyst stage is completely developed and it gives a better acceptance rate. Cleavage stage of the embryo will also be considered for the success rate as it gives us the total number of healthy embryos transferred to Day-3. As shown in the graph below it gives the overall assessment of the single embryo selection or a potential embryo implantation [5].

The comparison of these individual days result shows that if a single viable embryo is selected and transferred to women’s womb which leads to the implantation definitely the success rate is going to improve. These Machine learning techniques [6] helps us to do these in an easier way. The selection criteria are more flexible and once the training is done it’s very easy to implement it on the setup.

Using these statistics, the precision, recall and F-measure can be obtained from the following formula. Here, precision represents the correction rate in detected boundary pixels. Recall represents the detected boundary pixels in the ground truth pixels. F-measure is the harmonic mean of precision and recall and
is used as the overall measure of the detector performance. We express the precision, recall and overall measure as:

\[
\text{Precision} = \frac{TP}{TP+FP} \quad (1)
\]

\[
\text{Recall} = \frac{TP}{TP+FN} \quad (2)
\]

\[
\text{F-Measure} = \frac{(1+\beta^2) \cdot \text{Recall} \cdot \text{Precision}}{\beta^2 \cdot \text{Recall} + \text{Precision}} \quad (3)
\]

\[
\text{F-Measure} = \frac{2TP}{TP+FP+FN} \quad (4)
\]

For \(\beta=1\), the accuracy of the blastomere cell is calculated as:

\[
\text{Accuracy} = \frac{TP+FP}{\text{Total number of cells detected}} \quad (5)
\]

Blastomere ratio is calculated by the number of true cells detected with the actual input cells in blastomere.

\[
\text{Blastomere ratio} = \frac{\text{True cells detected}}{\text{Total cells in Blastomere}} \quad (6)
\]

The use of Machine Learning algorithm gave excellent results in the analysis of the healthy development of the embryo selection from Day-2 to Day-3 based on the embryo features. Considering the example where all positive inputs are given, learning the concept of “fragments, nucleoli, grade, fertilization count, and number of cells in blastomere, also embryo features: embryo cell, radius, fragments specified, fertilization count and finally the grade.

References

[1] Sarah A, Priya B, Vanessa J 2019 Time-Lapse systems for Embryo incubation and assessment in assisted Reproduction Cochrane Database of Systematic Reviews 5

[2] Meyers S, Burrue V 2019 Equine non-invasive time-lapse imaging and embryo development Reproduction, Fertility and Development 31 1874-84

[3] McLennan HJ and Saini A 2020 Embryo evaluation by AI and multi-spectral auto-fluorescence imaging: livestock embryology needs to catch-up to clinical practice Theriogenology 150 255-262

[4] Bovin J, Bunting L, Collins JA, Nygren KG 2007 Human Reproduction (Oxford, London): International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care PubMed Journal 22 2800.

[5] Sujata NP, Wali UV, Swamy MK 2016 Automation Technique for classification of Human In-Vitro Fertilized (IVF) Embryos using Digital Image Processing Techniques International Journal of Technology and Science 3 46-49

[6] Sujata NP, Wali UV, Swamy MK 2018 Deep Learning Techniques for Automatic Classification and Analysis of Human In Vitro Fertilized Embryos Journal of Emerging Technologies and Innovative Research (JETIR)

[7] Johansson M, Berg M 2005 Women's experiences of childlessness 2 years after the end of in vitro fertilization treatment PubMed Journal 19 58-63

[8] Emre B, Shevlin M, Sutcliffe A 2010 Reproductive Biomedicine Online: Growth of children conceived by IVF and ICSI up to 12 years of age. Pediatrics 126 270

[9] Robert GE, Johnson MH 2011 Reproductive Biomedicine Online: Robert Edwards: The path to IVF 23 245–262
[10] Malchoff CD and Kobolt U 1991 Clinical Chemistry, Principles, Techniques and Correlations China: Springer 55 625
[11] Callum FR, Hall MI and Heesy CP 2006 Were Basal Primates Nocturnal? Evidence from Eye and Orbit Shape. Ross Lab University of Chicago 67 695-707.
[12] Psaty BM, Heckbert SR, Koepsell TD, Siscovick DS 1995 The risk of myocardial infarction associated with antihypertensive drug therapies University of Washington, Seattle 274 620-5.
[13] Sujata NP, Uday VW, Swamy MK 2016 Application of Vessel Enhancement Filtering for Automated Classification of Human In-Vitro Fertilized (IVF) Images. International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), 27978-1
[14] Saritha KR, Bongso A 2001 Comparative Evaluation of Fresh and Washed Human Sperm Cryopreserved in Vapor and Liquid Phases of Liquid Nitrogen Journal of Andrology 22 857-862
[15] Abdel-Salam AH, Abdel-Bakey NF 2001 Life table and biological studies of Harmonia axyridis Pallas (Col., Coccinellidae) reared on the grain moth eggs of Sitotroga cerealella Olivier (Lep, Gelechiidae) Journal of Applied Entomology 125 455-462
[16] Hussain J, Salam A and Gohar 2001 A Study on the Cryopreservation of Stallion Semen with Alpha Lipoic Acid. International Journal of Pharmaceuticls 01
[17] Samaneh H, Farsi MM, Khafri S 2016 Should We Perform Semen Analysis, DNA Fragmentation, and Hypo-osmotic Swelling Tests together? International Journal of Molecular and Cellular Medicine 5 246-254
[18] Khosravi P, Kazemi E, Zhan Q, Malmsten JE 2019 Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization NPJ digital medicine 2 1-9
[19] Chavez-Badiola AFS, Farias G, Ruiz M 2020 Predicting pregnancy test results after embryo transfer by image feature extraction and analysis using machine learning Scientific Reports 10 1-6
[20] Dirvanauskas D, Maskeliūnas R, Raudonis V 2019 Hemigen: human embryo image generator based on generative adversarial networks Sensors 19 3578
[21] Curchoe CL, Bormann CL 2019 Artificial intelligence and machine learning for human reproduction and embryology presented at ASRM and ESHRE Journal of Assisted Reproduction and Genetics 36 591-600