RESEARCH ARTICLE

INTERNATIONAL RESEARCH JOURNAL OF MULTIDISCIPLINARY TECHNOVATION

Asian Research Association

DOI: 10.54392/irjmt2636

A Computational Ensemble Framework for Multiclass Blastocyst Segmentation: Modelling Morphological Complexity in Human Embryos

R. Barkavi [a,*], G. Yamuna [a], C. Jayaram [b]

[a] Department of Electronics and Communication Engineering, Annamalai University, Chidambaram, India
[b] Embryologist, Faik IVF Fertility Al Ain, United Arab Emirates
* Corresponding Author Email: barkavi.radhakrishnan@gmail.com

DOI: https://doi.org/10.54392/irjmt2636
Received: 11-10-2025; Revised: 29-03-2026; Accepted: 20-04-2026; Published: 06-05-2026

Abstract: Artificial intelligence is evolving in the field of embryology, offering exciting possibilities for improved results in Assisted Reproductive Technologies such as In Vitro Fertilization. The morphological quality of blastocysts of day 5 human embryo is a vital factor for determining the success of In Vitro Fertilization Hence making accurate and automated analysis of embryonic structures essential. To achieve an automated assessment of human embryo quality on the basis of morphological image features, it is crucial to precisely segment the regions of the embryo. In this research, a comprehensive analysis of U-Net and its variants for the semantic segmentation of human embryo day 5 blastocyst images is performed. Based on this insights gained from the comparative analysis, a novel Ensemble segmentation model is proposed to exploit the complementary strengths of multiple models. The proposed ensemble approach demonstrates robust performance, achieving an overall segmentation accuracy of 98%, with an F1-score of 0.95081 and a Jaccard index of 0.90625, indicating high spatial agreement between predicted and ground-truth segmentations. The framework effectively addresses key challenges inherent to blastocyst imaging, including low-contrast boundaries, heterogeneous cellular organization, and limited annotated data. By enabling precise and reproducible segmentation of critical embryonic regions, the proposed method provides a reliable foundation for automated embryo quality assessment and grading systems. This work also contributes to the integration of biologically inspired computational models into clinical embryology and supports the broader adoption of AI-driven decision support tools in reproductive medicine.

Keywords: In Vitro Fertilization (IVF), Image Analysis, Artificial Intelligence, Deep Learning Neural Networks, Human Embryos, Day-5 Blastocysts.

1. Introduction

Infertility is a disease of the male or female reproductive system defined by the failure to achieve a pregnancy after 12 months or more of regular unprotected sexual intercourse. [1] That is it affects up to 15% of reproductive aged couples worldwide and estimate the overall prevalence of primary infertility in India is between 3.9 to 16.8% .A survey under phase 2 of National Family Health Survey (NFHS-5, 2019) released factsheets on India's Total Fertility Rate (TFR), the average number of children per woman, has further declined from 2.2 to 2 at the national level [2, 3]. This can be caused by a number of causes, including age, lifestyle choices, environmental variables, hormonal imbalances, and abnormalities of the reproductive system. Depending on this, infertility may be temporary or chronic and may necessitate medical intervention, including fertility treatments like Assisted Reproductive Technology (ART) [4]. It helps people with difficulties to conceive a child naturally by enhancing the probability of pregnancy. Among them IVF is the most common infertility treatment, whose success is predominately depends on the quality of the embryo. The main challenge for the embryologist is to identify this quality embryo to be transferred that is most likely to implant. This selection process should be accurate, non-invasive, inexpensive, reproducible, and available to IVF laboratories worldwide. [5] A single-embryo transfer is considered as priority as it helps to avoid high risks during the pregnancy [6].

Most commonly, the quality of an embryo is classified or graded by an embryologist. An embryologist can evaluate the viability of an embryo in the blastocyst stage via manual assessment of its morphological features, such as the zona pellucida (ZP), trophectoderm (TE), blastocoel (BL), and inner cell mass

(ICM) as illustrated in the Figure1. However, these conventional visual evaluations of embryo morphology have high intervariability and intravariability, and the lack of consistency in such assessments leads to unevenness in decision-making, which is detrimental to both patient care and the progress of the field overall. [7] Studies demonstrate that AI based methods decrease the variability among human observers by 40% while achieving equivalent accuracy. This reduction in inter-observer variability could significantly standardize embryo selection procedures across different clinical settings [8]. In this context, segmenting the regions of the embryo is indispensable, garnering significant attention as a demanding area of research.

Many studies are being conducted to use artificial intelligence in all feasible methods in order to create accurate, and effective technological inventions and healthcare is of no exception. [9]. AI also transforms embryology, with the ability to accommodate everything from embryo selection to child care. Deep learning based architectures have shown remarkable results in image segmentation, outperforming traditional machine learning approaches in terms of both accuracy and speed [10]. Among these architectures, U-Net has emerged as a seminal model for biomedical image segmentation due to its encoder–decoder design and efficient feature reuse through skip connections [11]. In embryology-related applications, U-Net-based models have been successfully applied [12]. However, blastocyst images pose unique challenges, including low-contrast boundaries, heterogeneous cellular organization, imaging artifacts, and limited availability of annotated datasets. Dilated convolutional networks have also demonstrated improved performance by capturing global contextual information, which is particularly important for segmenting spatially extended embryonic structures such as the blastocoel and zona pellucida [13]. Despite these advances, no single network architecture consistently achieves optimal performance across all segmentation targets. Models emphasizing fine boundary detail may underperform in capturing global context, while context-aware models may compromise local structural accuracy. Ensemble learning has emerged as a robust solution to this challenge by combining multiple complementary models to reduce prediction variance and improve generalization [14]. Ensemble-based segmentation frameworks have demonstrated superior robustness and accuracy across diverse biomedical imaging tasks compared to individual models [15].

Motivated by these developments, this study proposes a novel ensemble segmentation framework that integrates a modified U-Net and a modified Dilated U-Net for the semantic segmentation of day-5 human blastocyst images. The modified U-Net is designed to enhance fine-grained boundary delineation of critical embryonic regions, particularly the ICM and TE, through improved feature fusion. Concurrently, the modified Dilated U-Net employs atrous convolutions to effectively capture multiscale contextual information essential for accurate segmentation of the BL and ZP. The ensemble strategy fuses the complementary predictions of both architectures to mitigate individual model biases and improve segmentation robustness. The proposed framework is evaluated using standard segmentation metrics, including Dice similarity coefficient, Jaccard index, precision, recall, and F1-score, following established medical image analysis protocols [16]. By enabling accurate, reproducible, and anatomically consistent segmentation of blastocyst components, this work provides a technically sound foundation for downstream automated embryo grading systems and supports the clinical translation of AI-assisted decision support tools in assisted reproductive technology.

2. Background

The literature survey is the foundation of the research, the main aim to determine the key trends in segmentation and highlights the findings merits and demerits. [17] Proposed an Ensemble Deep Convolutional Neural Network (EDCNN) model which integrates the strengths of the MobileNet and Xception to produce an accuracy of 87.82% on Dataset1 and 85.69 on Dataset 2.

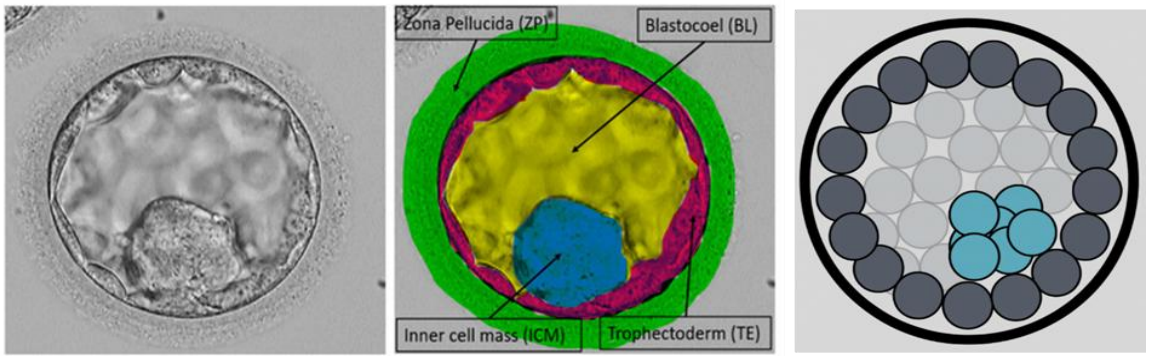

Figure 1. Blastocyst microscopic Image with colour annotation and schematic representation

[18] Proposed a semantic segmentation of human blastocyst components using atrous spatial pyramid pooling (ASPP) combined with dense progressive upsampling model to capture multi scale contextual and structural boundaries in blastocyst images. [19] Evaluated multiple neural network models for accurate semantic segmentation of human oocytes, enabling extraction of morphological features, based on these evaluation a tuned model is proposed for oocyte region extraction. [20] An automated approach to segment human embryo in early-stage development from a sequence of dark field microscopy images is introduced where segmentation is considered as an energy minimization problem, and solved efficiently via graph-cuts. [21] Introduced Ensemble Edge Fusion, a novel ensemble learning–based framework designed to improve semantic segmentation performance in microvascular decompression (MVD) imaging.

Table 1. Comparative Analysis of Ensemble semantic segmentation models

S.no.	Reference	Application	Method Used	Ensemble Type	Limitations
1.	Islam *et al.*, 2024 [17]	Breast cancer image analysis	Ensemble Deep CNN (MobileNet + Xception)	Model-level ensemble	Performance drops across datasets
2.	Bormann *et al.*, 2019 [18]	Blastocyst component segmentation	ASPP + Dense progressive upsampling (BLAST-NET)	Implicit multi-scale ensemble	Limited dataset size
3.	Targosz *et al.*, 2021 [19]	Human oocyte segmentation	CNN-based semantic segmentation	None	No ensemble strategy
4.	Khan *et al.*, 2016 [20]	Early-stage embryo segmentation	Graph-cut energy minimization	None	Limited to early-stage embryos
5.	Dhiyanesh *et al.*, 2025 [21]	Medical image segmentation (MVD imaging)	Ensemble Edge Fusion	Explicit ensemble	Domain-specific validation
6.	Diakogiannis *et al.*, 2020 [22]	Remote sensing segmentation	ResUNet-a	Multi-branch residual fusion	Not biomedical
7.	Huang *et al.*, 2020 [23]	Medical image segmentation (CT/MRI)	UNet 3+	Full-scale skip fusion	High memory usage
8.	Kamnitsas *et al.*, 2017 [24]	Biomedical image segmentation	Multilevel dilated residual network	Multi-scale dilation	No ensemble aggregation
9.	Saadati *et al.*, 2023 [25]	Medical image segmentation	Dilated Transformer + U-Net	Hybrid architecture	Limited clinical validation
10.	Nematzadeh *et al.*, 2023 [26]	Melanoma detection	Ensemble + GA-based explainer	Explicit ensemble with explainability	Computationally expensive
11.	Rad *et al.*, 2020 [27]	Trophectoderm segmentation in human embryo images	Inceptioned U-Net	Multi-branch feature learning (implicit)	High computational complexity
12.	Uysal *et al.*, 2022 [28]	Human embryo segmentation	U-Net variants	None	Absence of explicit ensemble strategy
13.	Arsalan *et al.*, 2022 [29]	Blastocyst component detection (ICM, TE, Blastocoel)	Multiscale aggregation semantic segmentation network	Implicit multi-scale fusion	Dataset-specific evaluation
14.	Wang *et al.*, 2021 [30]	Automatic blastocyst evaluation using multifocal images	Deep CNN framework	None	Increased computational cost

The proposed Ensemble Edge Fusion method achieved a superior MIoU of 77.73%, outperforming each individual architecture on the MVD dataset, and demonstrated enhanced segmentation accuracy and boundary delineation across multiple anatomical categories. [22] Proposed ResUNet-a, a residual U-Net–based architecture for segmentation in remote sensing imagery which integrates residual connections, atrous convolutions, and multi-scale feature fusion to improve segmentation accuracy It is an ensemble and multi-branch segmentation frameworks. [23] describes a U-Net architecture that aggregates multi-resolution features across all encoder and decoder stages which improves boundary structures and contextual awareness in medical image segmentation tasks such as CT and MRI analysis. [24] Proposed a multilevel dilated residual network capture both fine-grained local details and broader contextual information.

The residual learning framework improves feature propagation and convergence, resulting in enhanced segmentation accuracy. [25] Presented a hybrid segmentation architecture that integrates a Dilated Transformer with a U-Net backbone to enhance global contextual modelling and local feature extraction in medical image segmentation tasks. [26] Introduced an ensemble-based image segmentation framework integrated with a genetic algorithm (GA)–driven explainer for melanoma detection which combines multiple segmentation models and leverages evolutionary optimization to improve both prediction accuracy and interpretability. [27] proposed an inceptioned U-Net architecture for accurate trophectoderm segmentation in human embryo images. By integrating Inception modules into the U-Net framework, the model effectively captured multi-scale contextual features, resulting in improved boundary segmentation [28] conducted a comparative evaluation of multiple U-Net–based architectures for human embryo segmentation, demonstrating that architectural modifications such as depth variation and skip-connection refinement significantly influence segmentation accuracy and robustness. [29] Introduced a multiscale aggregation semantic segmentation network for detecting human blastocyst components, enabling precise localization of the inner cell mass (ICM), trophectoderm (TE), and blastocoel through hierarchical feature fusion. [30] presented a deep learning framework for automatic blastocyst evaluation using multifocal embryo images, where multi-focus feature integration improved segmentation reliability and supported downstream embryo quality assessment. Collectively, these studies highlight the effectiveness of multi-scale feature learning and U-Net–based architectures for Day-5 blastocyst segmentation.

2.1 Problem statement

The IVF has very low success rate, long-haul process, chances of multiple pregnancies (preterm labor, gestational issues), and high cost, which is unaffordable to common people.

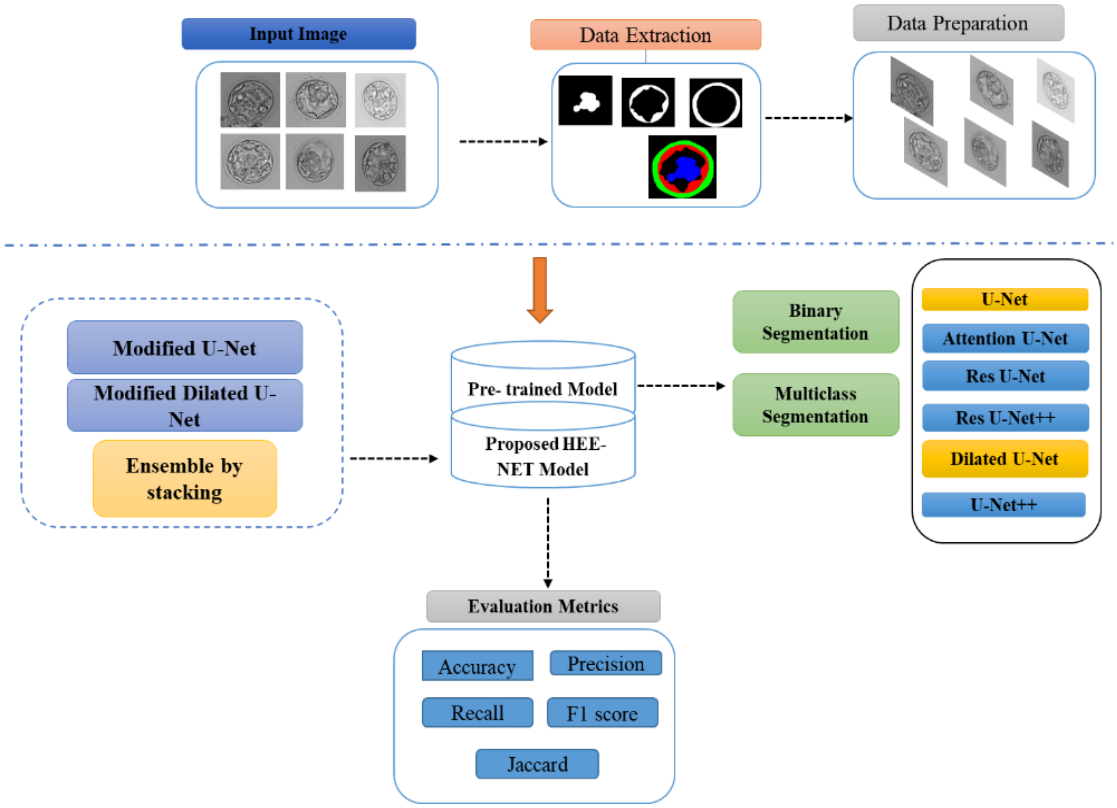

Figure 2. Overview of the methodology

As an emerging field, the lack of available medical experts, the potential for interobserver variability, and the laborious process of evaluation without a standard method underscore the need for an automated system. With the advancement of AI, a deep learning-based method leading to earlier detection and better patient outcomes is pivotal. However, segmenting the regions of the embryo is important, as it is the necessary first step for further processing. Semantic segmentation of human embryo images presents a unique challenge because of the complex shapes and textures involved. Single deep neural network architectures, such as standard U-Net or its variants, exhibit limitations in capturing both fine spatial details and long-range contextual information. Additionally many current segmentation approaches demonstrate limited robustness when applied to small, imbalanced datasets especially like of embryo imaging, whereby increasing the risk of overfitting and reduced generalizability in clinical settings. Therefore, it is necessary to bring lightweight, robust and reliable semantic segmentation framework.

The Figure 2 illustrates the proposed Ensemble based embryo image analysis pipeline. Raw embryo images are first acquired and subjected to data extraction and data preparation, including preprocessing and mask generation. The processed inputs are then fed into a pre-trained ensemble framework, where Modified U-Net and Modified Dilated U-Net models are combined using stacking to form the proposed Ensemble model. The system supports both binary and multiclass segmentation, leveraging multiple architectures (U-Net, Attention U-Net, ResU-Net, ResU-Net++, Dilated U-Net, and U-Net++). Model performance is quantitatively evaluated using Accuracy, Precision, Recall, F1-score, and Jaccard index.

The overview of the study is as follows:

1. ***Materials and methods:*** Examine briefly the dataset and its description.
2. ***Methodology:*** Explains the methods used in the experimental study and the proposed methodology.
3. ***Results and Discussion:*** Performance evaluation of the model is discussed.

3. Materials and Methods

3.1 Dataset Description

This study makes use of the sole publicly accessible human blastocyst dataset, which was first presented by [13]. The dataset includes 249 photos of embryos captured at the Pacific Centre for Reproductive Medicine (PCRM) in Canada using an Olympus IX71 inverted microscope fitted with Nomarski differential interference contrast (DIC) optics. Using the VGG image annotator (VIA), experts from Acıbadem Fulya Hospital IVF Center classified portions of the embryo in photos as either background or embryo. Using the OpenCV library, all of the embryo's closed border coordinates were entered into an annotated CSV file to create ground truth images. The input embryos (blastocysts) and the corresponding area masks are shown in Figure 3. The overlay mask that shows the regions' color mapping are shown below

3.2 Dataset pre-processing

Image pre-processing is the most essential step to obtain better performance from the deep learning model especially for smaller datasets. Data augmentation can prevent overfitting in training data and provide better generalizability to test data. To enhance training data, the samples are rotated, horizontally and vertically flipped, shifted, and zoomed 20 times to increase the dataset size to 5418 synthetic images. To further support early stopping based on validation loss and dropout layers were used within the network to prevent co-adaptation of features. The dataset [13] used for research consists of 249 images of size 512--512 that were resized to (128, 128). The images are electron microscopy images saved in PNG format. Though the images were augmented, the overfitting problem is still a concern in training the model to mitigate this concern, k-fold cross-validation was employed on the full dataset. A 5-fold cross-validation strategy was introduced where the dataset was divided into five mutually exclusive folds. In each iteration, four folds were used for training and one fold for validation, ensuring that each image contributed to validation exactly once as shown in Figure 4.

Figure 3. a) Sample input image **b)** Ground truth of the zona pellucida**. c)** Ground truth of the trophectoderm**. d)** Ground truth of the inner cell mass**. e)** Color-mapped ground truth**.**

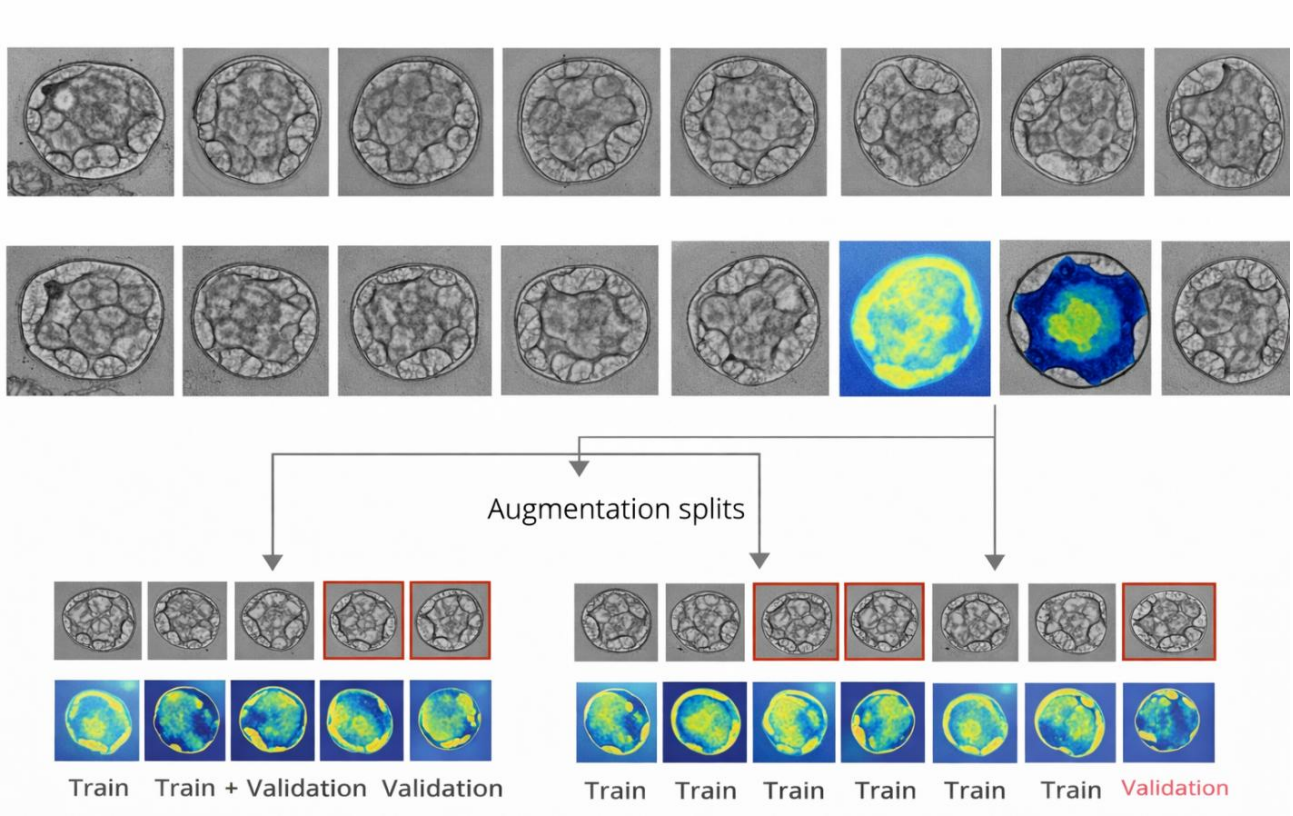

Figure 4. Augmentation methods and cross validation

Where augmented images derived from the same original embryo are grouped together before dataset partitioning. This ensures that augmented versions of a single image are not simultaneously present in both training and validation sets, thereby preventing data leakage. The bottom section depicts the 5-fold cross-validation protocol [31]. The augmented dataset is divided into five mutually exclusive folds. In each iteration, four folds are used for training, while the remaining fold, highlighted with red bounding boxes is used exclusively for validation. This process is repeated five times, allowing each fold to serve as the validation set once. The corresponding color-mapped outputs illustrate model predictions for training and validation samples. Performance metrics were averaged across folds to obtain a more robust and unbiased estimate of model performance. This approach reduces variance associated with a single train test split and improves confidence in the reported results, particularly for limited medical imaging datasets. Together, these measures help improve model generalization despite the limited dataset size.

4. Methodology

The U-Net convolutional neural network and its numerous variants are the pioneers in the state-of-the-art performance in medical image segmentation. This section explains U-Net and its variants, which are employed in this work to analyse the different artifacts of the blastocyst image. . Ultimately, an ensemble model with different strategies is proposed, and it outperforms all the other models.

4.1 U-Net and its Variants

The capacity of the U-Net architecture to precisely divide organs and other structures from medical pictures has made it well-liked in the medical imaging profession. Developed by Olaf Ronneberger [11], U-Net is a convolutional neural network (CNN) that excels at picture segmentation tasks. Although it was initially created for the segmentation of biomedical images, it has been used for a wide range of segmentation issues. The network's contracting path is made up of max pooling layers after a sequence of convolutional layers.

By gradually downsampling the input image and learning hierarchical features, it is able to capture its context. Convolutional layers come after a sequence of upsampling layers in the Expansive Path, which is symmetric to the encoder. Reconstructing the image's spatial dimensions from the encoder's learned compressed feature representations is the aim of the decoder. The appropriate encoder and decoder layers are directly connected via a skip link. These links give the reconstructed output fine-grained details and assist the network in recovering spatial information that may be lost during downsampling. In order to decrease the number of feature channels to the required number of output classes for segmentation, the last layer of the U-

Net is often a 1x1 convolution. Through the prediction of a binary mask for every pixel in the input image, this procedure allows the network to capture the image's high-level properties.

Numerous variations of U-Net have been developed, particularly for the segmentation of medical images. The diagrammatic depiction of U-Net and its variations covered in this work is displayed in Figure 5. Pay attention In order to improve its performance, U-Net [27] integrates attention mechanisms into the conventional U-Net architecture. This allows the model to focus on more pertinent features while suppressing irrelevant ones, further refining the feature maps by giving importance to various feature map regions. In order to provide a more resilient architecture for improved performance, the ResU-Net model [28] use residual units rather than simple neural units.The residual units used in the encoder, bridge, and decoder are made up of two convolution blocks and a skip link between the unit's input and output. A BN layer, a ReLU activation layer, and a convolutional layer are present in every convolution block.

An expansion of the original U-Net architecture is called Dilated U-Net [25]. The original U-Net's encoder-decoder design is usually maintained in this scenario, but both the contracting (encoder) and expanding (decoder) routes make use of dilated convolutions. This makes it possible for the network to efficiently extract both local and global information from the input image, leading to more precise segmentation outcomes, particularly when objects of interest have wide size variations. A stack of five dilated convolution layers with dilation rates of 1, 2, 4, 8, and 16 is part of the network's core bridge. Residual blocks, squeeze and excitation blocks, atrous spatial pyramid pooling, and attention blocks are all utilized in the ResUNet++ [22] architecture. UNet+++, [23], also known as UNet3+, is an enhancement over the original U-Net and UNet++. It aims to further improve the segmentation performance by integrating features from multiple levels of the encoder and decoder via deep supervision and full-scale skip connections.

4.2 Proposed Methodology

The robust ensemble model is proposed by leveraging the complementary strengths of U-Net and its variants to achieve higher accuracy and to enhance the overall predictive performance. The prediction result of a single model suffers from noise, bias and generalization problems [23]. The underlying principle is that by training a diverse set of learners and combining their predictions, the ensemble can capitalize on the unique insights and perspectives offered by each component model, leading to enhanced overall performance.

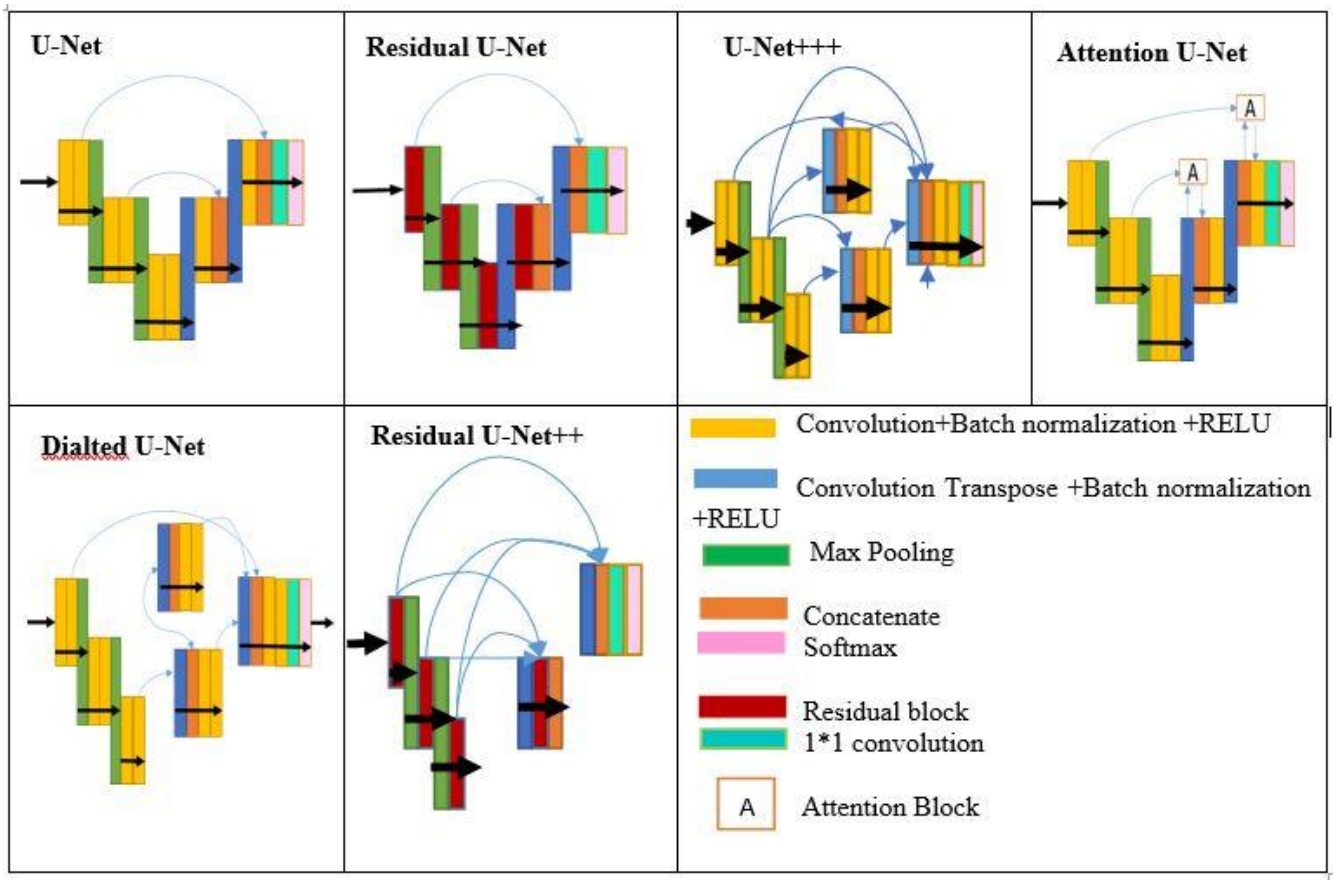

Figure 5. Summary of U-Net and its variant architecture

There are several types of Ensemble techniques available. The Ensemble strategy assay in this study is an ensemble method that involves stacking with a small customized neural network (CNN) as a meta-learner. In this research, the model selection rationale is based on earlier investigations, where attention U-Net yields stronger performance in binary segmentation, whereas dilated U-Net and U-Net perform exceptionally well in multiclass segmentation; therefore, both models are employed for further exploration with much improved fine-tuning. Figure 5 provides a comprehensive overview of the suggested ensemble model.

The architecture consists of a modified U-Net, which aims to address the issues of increasing precision while managing computational complexity in terms of parameters and FLOPs. In addition, the network's performance on image segmentation boundaries can be improved. The modified U-Net is incorporated as the first component in the ensemble. It focuses on preserving the spatial features while reducing overfitting and computational overhead.

4.2.1 Strided Convolutional Layers for Efficient Down sampling

The two consecutive 2D convolutional layers are engaged with 64 and 128 filters, respectively, a kernel size of 3×3, and ReLU activation to preserve spatial dimensions. A stride of 2×2 is applied in place of pooling to enable learnable downsampling, reducing feature map resolution while preserving contextual richness. The convolution operation is as follows:

$$Y = ReLU(W * X + b) \quad (1)$$

Where W and b are the learnable weight and bias parameter, respectively.

4.2.2 Batch normalization

To stabilize the training process by reducing the internal covariate shift and to reduce sensitivity to initialization, batch normalization is introduced after each convolution. For a batch $\boldsymbol{B = \{x_1 \ldots\ x_m\}}$, batch normalization is computed via the following formula in equation 2:

$$i = \frac{xi - \mu B}{\sqrt{\sigma B2 + \varepsilon}}, yi = \gamma x^i + \beta \quad (2)$$

Where μB and $\sigma B2$ are the batch mean and variance, respectively, and γ and β are trainable scaling factors.

4.2.3 Max Pooling

To allow the network to abstract semantic information, the max pooling operation with a 2*2 window is used. This operation also helps with translation invariance and reduces the spatial size of feature maps.

$$Yi, j = max_{(m,n)\in Ri,j}\, Xm, n \quad (3)$$

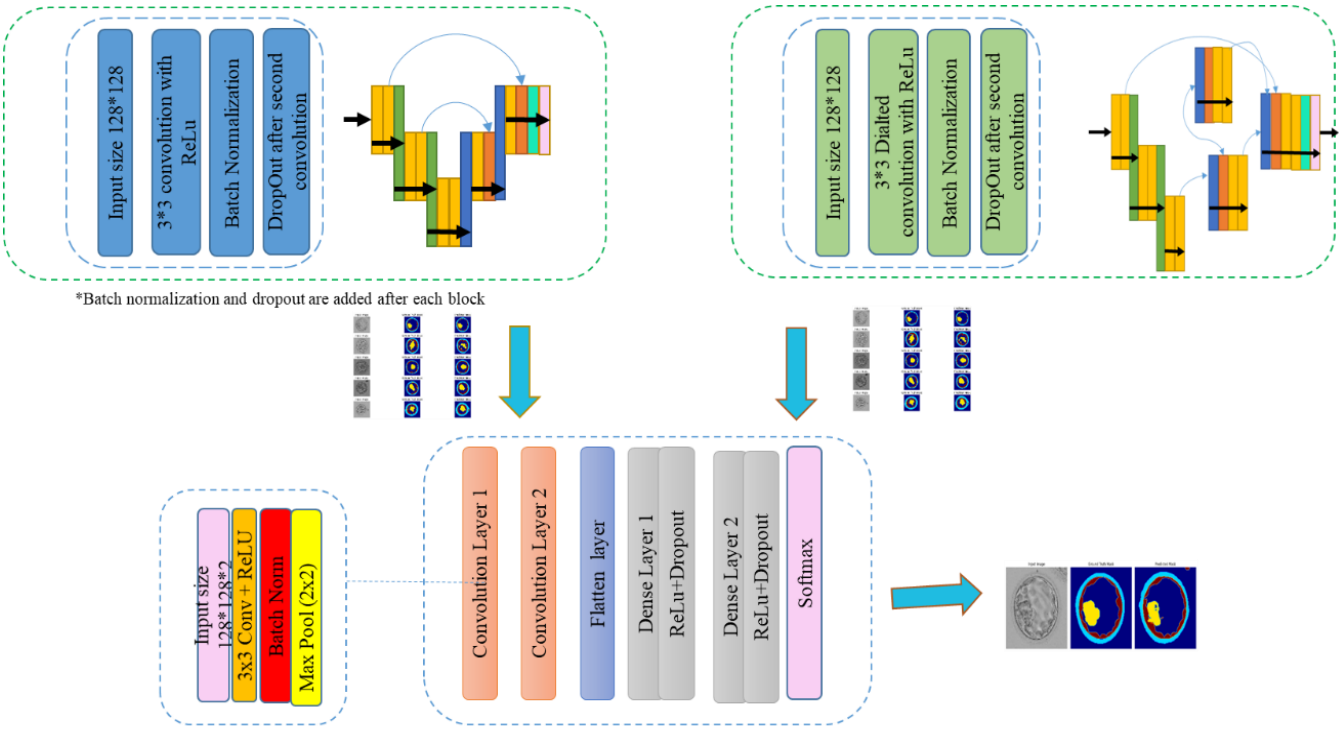

Figure 6. Overview of the proposed model methodology

4.3.4 Dropout Regularization

A dropout layer with a probability of p=0.3 is applied during training to randomly deactivate neurons, reducing overfitting and improving generalizability.

$$Xm\widetilde{x_{i=}}x_i z_i \quad zi \sim Bernoulli(1-p), n \qquad (4)$$

The second part is a modified dilated U-Net, which uses the capabilities of dilated convolutions to capture multiscale contextual information without increasing the number of parameters. This approach helps address the limitations of the classical U-Net in effectively extracting and integrating features at different scales. Like above, batch normalization is used consistently after each layer. Dropout after each pooling layer. The dilated convolution with a dilation rate of r=2 is used to exponentially increase the receptive field while maintaining resolution. The equation of operation is defined as

$$y[i] = k = 1\sum Kx[i + r.k].w[k] \qquad (5)$$

The outputs of both models are integrated via the ensemble stacking method with a small lightweight convolution neural network as a meta-learer. Inspired by the recent development of modifications in neural networks, it consists of a convolutional layer with 32 filters followed by a flattened layer, as described in Figure 6. A dense layer is added to predict better multiclass results. The activation function used here is the Softmax function. Here, both modified models are trained, and the predicted results are concatenated. Thus, this strategy works effectively, resulting in more fine-grained results

The pseudo code of the proposed model is of

Input:

$I \in R^{H\times W\times C}$ *: Raw blastocyst embryo image*

Output:

$Sfinal \in R^{H\prime\times W\prime\times Cs}$ *Final segmentation probability map*

Initialization:

Set input resolution $H' \times W' = 128 \times 128$

Set number of encoder–decoder levels $\mathcal{L}$

Set dilation rates $\mathcal{D}$ *= {2, 4,8}*

Set dropout rate r

Initialize learnable parameters for Modified U-Net

Initialize learnable parameters for Modified Dilated U-Net

Initialize fusion and refinement convolution parameters

Begin

Preprocessing :

Compute global mean

$\mu \leftarrow Mean(I)$

Compute standard deviation

$\sigma_l \leftarrow StdDev(I)$

Normalize image

$$In(x,y) = \frac{I(x,y) - \mu}{\sigma}$$

Resize normalized image

$Ir \leftarrow Resize(In, 128 \times 128)$

Apply data augmentation (rotations, flips, shifts, zoom)

Branch 1: Modified U-Net

Initialize encoder feature map set:

$E_U \leftarrow \emptyset$

Encoder Path:

$for\ l\ =\ 1\ to\ L\ do$

$Fl \leftarrow Conv_{3\times3}(I_r)\ +\ ReLU$

$Fl \leftarrow BatchNorm(F_l)$

$Fl \leftarrow Dropout(F_l, r)$

$Fl \leftarrow Conv_{3\times3}(F_l)\ +\ ReLU$

$Store\ F_l in\ E_u$

$Ir \leftarrow MaxPool_{2\times2}(F_l)$

end for

Bottleneck:

$BU \leftarrow Conv_{3x3}(Ir)$ + ReLU

Decoder Path:

for l = L down to 1 do

$Ul \leftarrow UpSample_{2\times2}(B_U)$

$U_l \leftarrow Concat(U_{l,} E_U[l])$

$BU \leftarrow Conv_{3x3}((U_l)\ +\ ReLU$

end for

Output Layer:

$S1 \leftarrow Conv_{1\times1}(B_U)$

Return S1

Branch 2: Modified Dilated U-Net

Initialize encoder feature maps $E_D \leftarrow \emptyset$

For level l*=1 to L do*

$Fl' \leftarrow Conv_{3x3}(Ir) + ReLU$

$Fl' \leftarrow BatchNorm(Fl')$

Store encoder feature maps

$Fl' \in ED$

Apply down-sampling

$I_r \leftarrow MaxPool_{2\times2(Fl')}$

Dilated Bottleneck

For each dilation rate $d \in D$*:*

$B_d \leftarrow DilatedConv_{3\times3(,)}$+$ReLU$

Multi-scale Bottleneck Feature Aggregation

$B_D \leftarrow (B_2, B_4, B_8)$

Decoder Path

For level l*=*$\mathcal{L}$ *down to 1:*

$Ul' \leftarrow UpSample_{2\times2}(B_D)$

$Ul' \leftarrow Concat(U_l' E_D[l]))$

$B_D \leftarrow Conv_{3\times3}(U_l')$+ReLU

Output Feature Map

$S_2 \leftarrow Conv_{1\times1}(B_D)$

Ensemble Fusion and Refinement

Feature-level Fusion

$F \leftarrow Concat(S1, S2)$

Refinement Convolution

$F \leftarrow Conv_{3\times3}(F)$+ReLU

$F \leftarrow BatchNorm(F)$

$F \leftarrow Dropout(F, r)$

Output Prediction

Softmax Classification
$\mathcal{S}_{final} \leftarrow Softmax(\mathcal{F})$
Return

5. Results and Discussion

5.1. Implementation

All experiments were conducted using Python 3.9.16 in the Google Colaboratory environment using an Acer Aspire 5 (Intel Core i3) personal computer where model training and evaluation were accelerated using cloud-based GPU resources provided by Google Colab. The original embryo (blastocyst) images of resolution 512 × 512 pixels were uniformly resized to 128 × 128 pixels to reduce computational complexity and ensure feasible training given the limited dataset size.

5.2. Hyper Parameter Setting

Hyperparameter configuration plays a critical role in achieving stable training, efficient convergence, and optimal segmentation performance in deep learning based medical image analysis. In this study, the hyperparameters were carefully selected based on empirical validation, prior literature, and the constraints imposed by limited annotated day-5 blastocyst data. The chosen settings aim to balance model complexity and generalization while avoiding overfitting and excessive computational overhead. The comparison of models needs to be unbiased; hence, all the hyperparameters were kept the same for the models, as shown below.

Table 2. Hyper Parameter Settings

Hyperparameter	Settings
Activation	Softmax
Optimizer	Adam
Image size	128x128
Batch size	16
Epochs	100
Learning rate	0.0001
Loss	Dice Loss

The consequential step is to determine the loss function well suited for the embryo (blastocyst) images with class imbalance. [32] Typically, binary cross entropy and categorical cross entropy are used for semantic segmentation however, the Dice loss function outperforms other loss functions. Dice loss is especially effective in handling class imbalance, a common characteristic of embryo (blastocyst) segmentation where foreground regions such as the inner cell mass (ICM) and trophectoderm (TE) occupy relatively small portions of the image. The Dice Similarity Coefficient between a predicted segmentation mask and the ground truth mask

Is defined as:

$$DSC = (2 \sum_i y_i \hat{y}_i) / (\sum_i y_i + \sum_i \hat{y}_i) \quad (6)$$

During the training process, the validation loss function is monitored, and the learning rate is adjusted accordingly. The minimum learning rate is set to le-7, in which the learning rate is reduced from le-4 when the validation loss function is not improved after five consecutive epochs, which is the patience factor. This strategy helps prevent premature convergence and improves model generalization. With the reduced size of input resolution to 128 × 128 pixels and a batch size of 16, the average GPU memory consumption during training was approximately 6–8 GB, leaving sufficient overhead for model ensembling and intermediate feature storage. Each segmentation model was trained for 100 epochs, with an average training time of approximately, 35–45 minutes per model on the Tesla T4 GPUAn Adam optimizer [33] is used for an efficient and reliable learning process, combining the advantages of adaptive learning rate estimation and momentum-based optimization, making it well suited for deep neural network training in medical image segmentation tasks.

5.3 Model Evaluation Measures

Table 3 describes the parameters utilized for evaluation in this study. Pixels that are accurately identified as belonging to the object in both the anticipated and ground truth masks are known as true positives (TPs). Pixels that are mistakenly identified as belonging to the item in the anticipated mask but not in the ground truth are known as false positives, or FPs. Pixels that belong to the item in the ground truth but are absent from the anticipated mask are known as false negatives, or FNs. Recall, sometimes referred to as sensitivity or true positive rate, quantifies the percentage of real positives that the model accurately detects.

Table 3. Performance metrics

Metric	Description
Precision	$\frac{TP}{TP+FP}$
Recall	$\frac{TP}{TP+FN}$
Jaccard index	$\frac{TP}{TP+FN+FP}$
F1 Score	$2*\frac{Precision*Recall}{Precision+Recall}$

The percentage of expected positives that turn out to be positive is known as precision. By comparing the size of two sets' intersection and union, the Jaccard index (intersection over union) calculates how similar

two sets are. The F1 score is calculated by taking the harmonic mean of recall and precision.

5.4 Evaluation of U-Net and its Variants in the Binary Segmentation of Blastocyst Regions

In this section, a comprehensive evaluation is conducted to provide a complete analysis and a complex approach to the binary segmentation of several blastocyst regions utilizing U-Net and its variants. Their logs and other score values of the trainable parameters are hoarded, and their prediction is tested on the basis of the evaluation metrics mentioned in the Table 3. The results, summarized in Table 4, while Attention U-Net achieves superior validation performance across most metrics, the training accuracy and loss metrics in Table 5 indicate that U-Net and dilated U-Net yield more consistent training outcomes.

The main objective of binary segmentation is to analyse the performance of the model in each region, Figure 7 shows a qualitative study of U-Net and its variants along with its segmentation results. For consistency and better interpretation, a single representative image is used, accompanied by its ground truth and the corresponding predictions from each model. The results indicate that the TE and ZP regions are clearly defined; nevertheless, the ICM region needs to be improved in terms of precise segmentation. Figure 8 shows a graphical representation of the performance of the model in each region.

Table 6 describes the performance metric values of U-Net and its variants for multiclass segmentation. The metric values show the efficiency of each model. The best results are highlighted so that the model can be selected for further ensembling. Although U-Net and its variants have proven strong in medical image segmentation, they yield varied results for these specific embryo blastocyst images. Blastocyst images have larger backgrounds and distinct, varied complex regions.

Table 4. Performance Metrics of U-Net and Its Variants

Zona Pellucida Region					
S.No.	**Network Model**	**F1**	**Jaccard**	**Recall**	**Precision**
1.	U-Net	0.92901	0.8674317	0.891057	0.970341
2	**Attention U-Net**	**0.929716**	**0.8686627**	**0.905451**	**0.979739**
3	RES U-Net	0.843523	0.7293907	0.899565	0.789525
4	RES U-Net t++	0.826438	0.7042132	0.802548	0.851794
5	U-Net +++	0.893954	0.808243	0.884553	0.888412
6	Dilated U-Net	0.910193	0.8351867	0.854878	0.973161
Trophectoderm Region					
1	U-Net	0.883537	0.791371	0.87765	0.889503
2	**Attention U-Net**	**0.945081**	**0.895881**	**0.954878**	**0.935484**
3	RES U-Net	0.866285	0.764111	0.88878	0.8449
4	RES U-Net t++	0.934513	0.877076	0.92093	0.948503
5	U-Net +++	0.908208	0.831851	0.919371	0.897313
6	Dilated U-Net	0.874971	0.777732	0.903911	0.847826
Inner Cell Mass Region					
1	U-Net	0.936255	0.880150	0.956395	0.916945
2	**Attention U-Net**	**0.958452**	**0.920219**	**0.942608**	**0.981609**
3	RES U-Net	0.903478	0.901267	0.881450	0.899123
4	RES U-Net t++	0.911513	0.837413	0.899529	0.92382
5	U-Net +++	0.939322	0.885587	0.900529	0.974839
6	Dilated U-Net	0.828129	0.706673	0.903436	0.764411

Table 5. Training and validation performance metrics

Zona Pellucida Region					
S.No.	**Network Model**	**Accuracy**	**Loss**	**Val Accuracy**	**Val loss**
1	U-Net	0.99038	0.031209	0.972925	0.320827
2	Attention U-Net	0.839500	0.160435	0.714178	0.285822
3	RES U-Net	0.899873	0.100019	0.662009	0.337991
4	RES U-Net t++	0.805204	0.194421	0.817162	0.182832
5	U-Net +++	0.954340	0.045643	0.779875	0.220125
6	Dilated U-Net	0.920331	0.079832	0.685294	0.317380
Trophectoderm Region					
1	U-Net	0.970423	0.218177	0.926247	0.433125
2	Attention U-Net	0.992209	0.177186	0.968317	0.306837
3	RES U-Net	0.866555	0.133169	0.652977	0.347023
4	RES U-Net t++	0.879061	0.120828	0.648342	0.351658
5	U-Net +++	0.951005	0.048885	0.792539	0.207461
6	Dilated U-Net	0.939881	0.060095	0.557245	0.442755
Inner Cell Mass Region					
1	U-Net	0.980656	0.257634	0.930345	0.273103
2	Attention U-Net	0.844213	0.155725	0.71409	0.285910
3	RES U-Net	0.853224	0.146486	0.723139	0.276861
4	RES U-Net t++	0.906875	0.104950	0.773998	0.226002
5	U-Net +++	0.919196	0.080970	0.679980	0.318154
6	Dilated U-Net	0.980656	0.257634	0.930345	0.27310 3

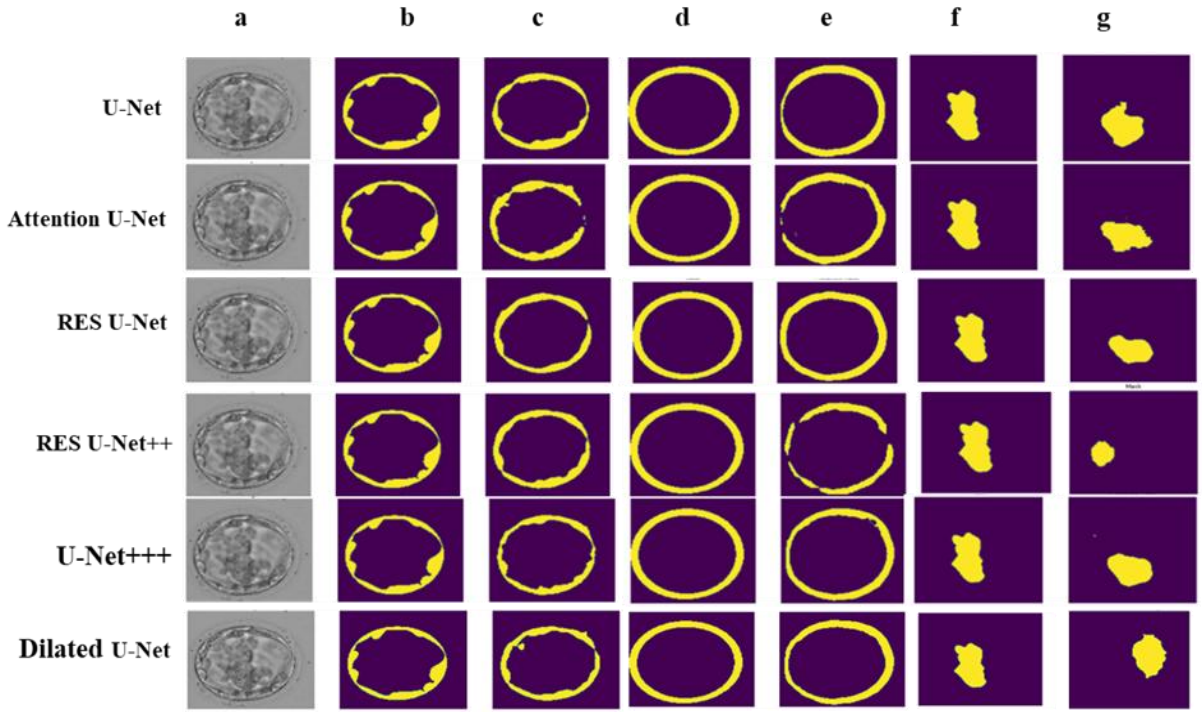

Figure 7. **a)** Input image **b)** Ground truth of TE, **c)** Predicted mask of TE, **d)** Ground truth of ZP, **e)** Predicted mask of ZP, **f)** Ground truth of ICM, **g)** Predicted mask of ICM

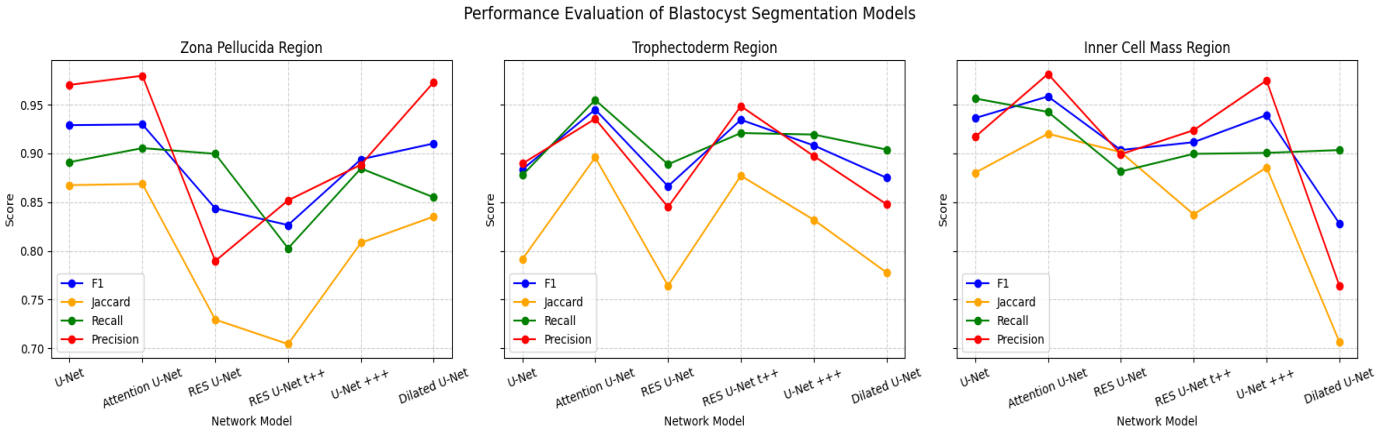

Figure 8. Graphical representation of the performance evaluation model.

Table 6. Comparative experimental blastocyst segmentation results.

Training and validation evaluation for multiclass segmentation of Blastocysts					
S.No.	**Network Model**	**Accuracy**	**Loss**	**Validation Accuracy**	**Validation loss**
1	**U-Net**	**0.9732**	**0.0416**	**0.8966**	**0.4157**
2	Attention U-Net	0.9431	0.1357	0.8978	0.2765
3	RES U-Net	0.8974	0.2443	0.8781	0.2923
4	RES U-Net ++	0.9687	0.0524	0.8951	0.3797
5	U-Net +++	0.9664	0.0798	0.8939	0.4556
6	**Dilated U-Net**	**0.9707**	**0.0380**	**0.8935**	**0.7260**

Performance evaluation for multiclass segmentation of Blastocysts					
S.No.	**Network Model**	**F1**	**Jaccard**	**Recall**	**Precision**
1	**U-Net**	**0.946478**	**0.898394**	**0.930404**	**0.963118**
2	Attention U-Net	0.842430	0.727758	0.929545	0.770245
3	RES U-Net	0.894158	0.808576	0.893742	0.894574
4	RES U-Net ++	0.932363	0.873296	0.955822	0.910028
5	U-Net +++	0.896975	0.813196	0.916465	0.878297
6	**Dilated U-Net**	**0.945488**	**0.896611**	**0.912409**	**0.981055**

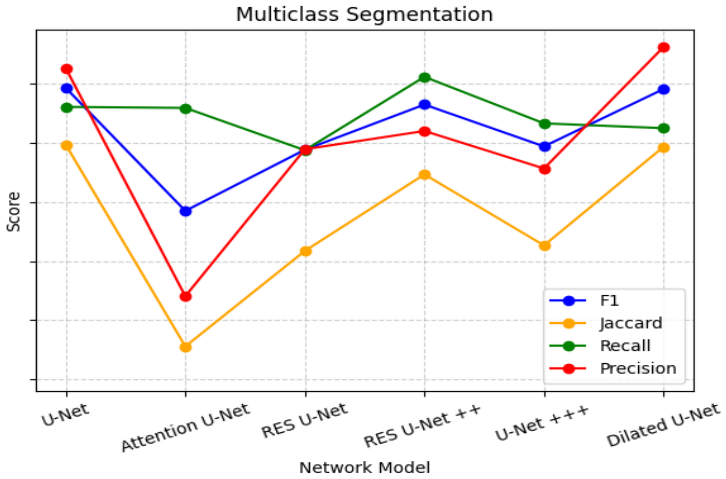

Figure 9. Graphical representation of the performance evaluation of the multiclass segmentation model.

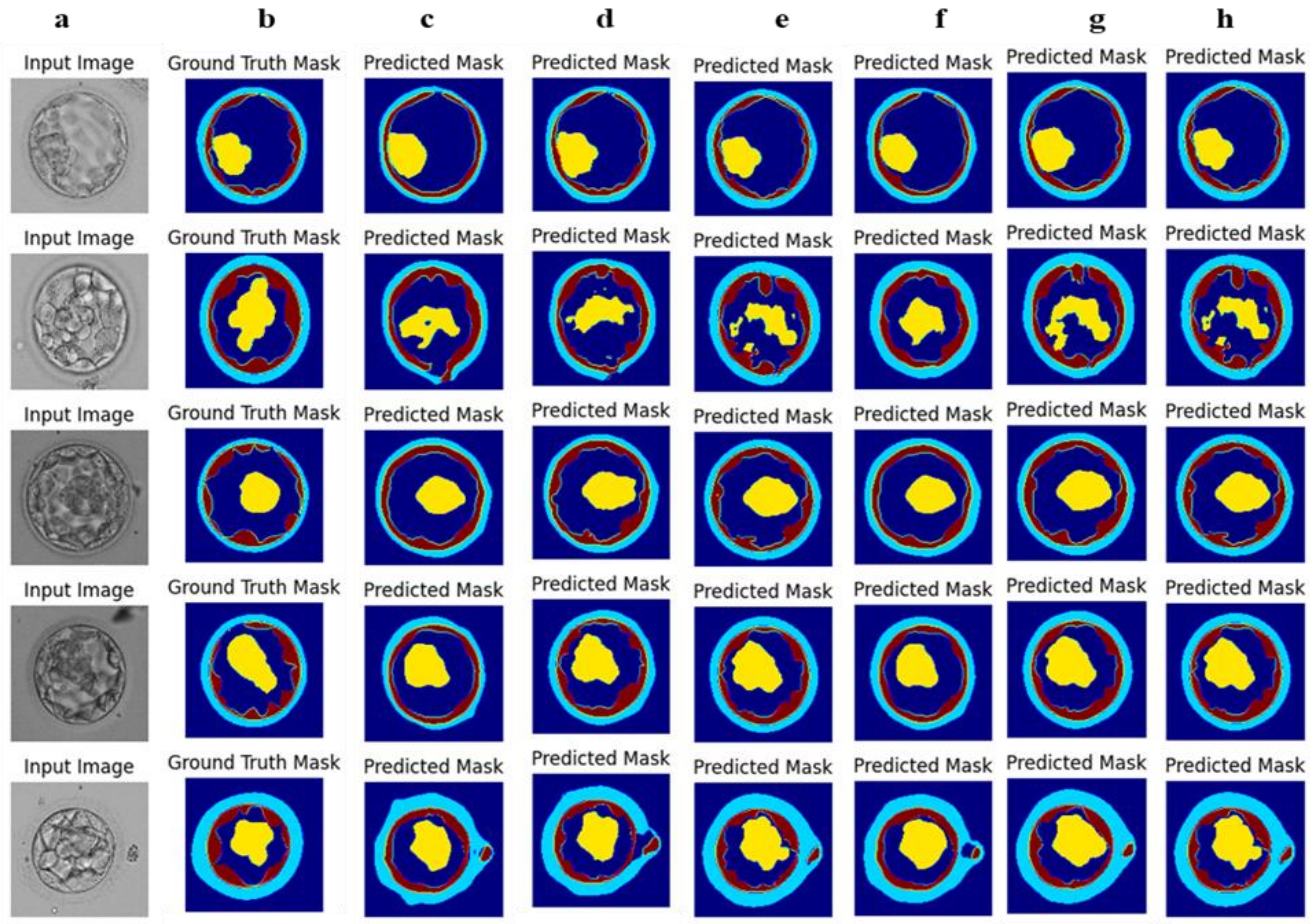

Figure 10. **a)** Input image **b)** Ground truth **c)** Predicted mask of U-Net**, d)** Predicted mask of Attention U-Net**, e)** Predicted mask of RES U-Net**, f)** Predicted mask RES U-Net ++, **g)** Predicted mask of U-Net +++, **h**) Predicted mask of Dilated U-Net

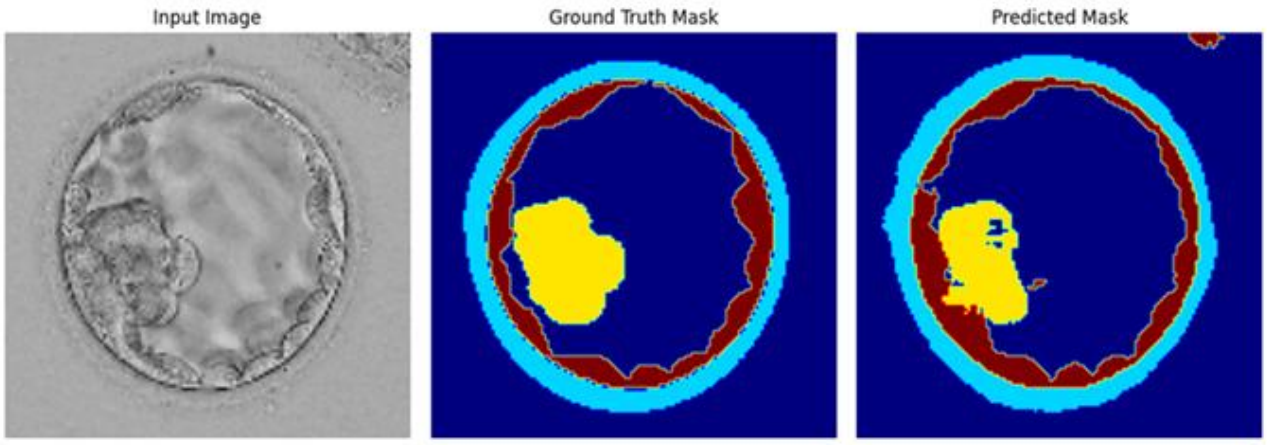

Figure 11. The input image, Ground truth and Predicted mask of the proposed model

The U-Net and dilated U-Net outperform the other methods in terms of accuracy and precision. Figure 9 represents the graphical representation of the values of the metrics for the multiclass segmentation models. Figure 10 displays the output of each model on multiclass segmentation, where each color represents a different region. In this study, to derive a better fine-tuned result, an Ensemble model is proposed. Figure 11 shows the output of the proposed model. From that visual description, it is clear that each region in the image is clearly segmented into its respective class, using distinct colors corresponding to the different areas. The model effectively captures fine details and sharp boundaries, particularly in the inner cell mass region, whereas earlier models fail to achieve accuracy. Moreover, very few artefacts or misclassifications are

visible, indicating the performance and robustness of the proposed model.

Table 6 explains that the proposed model shows evident performance in terms of accuracy and other metrics. However, the data focus more on the performance of the model, and visualization makes it easier to understand. The qualitative analysis is described in Figure 12, which clearly differentiate the performance of each model.

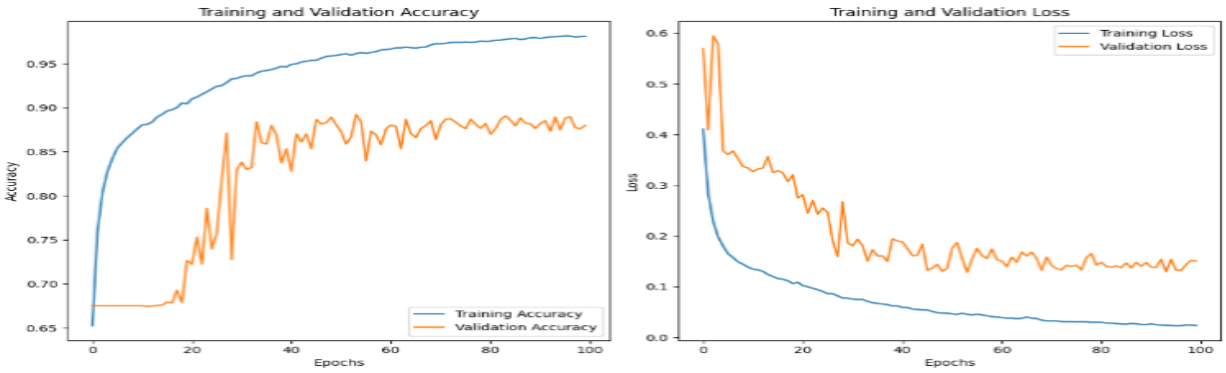

Figure 12. Training log graph of the proposed model displaying accuracy and loss

Table 6. Comparative experimental blastocyst segmentation results.

Performance evaluation of Ensemble methods proposed					
S.No.	**Network Model**	**Accuracy**	**Loss**	**Val Accuracy**	**Val loss**
1.	Modified U-Net	0.9720	0.0528	0.8340	0.1661
2.	Modified Dilated U-Net	0.9787	0.0524	0.8951	0.3797
3.	Proposed Ensemble model	0.9818	0.0229	0.8869	0.1339
S.No.	**Network Model**	**F1**	**Jaccard**	**Recall**	**Precision**
1.	Modified U-Net	0.93003	0.86921	0.91806	0.94230
2.	Modified Dilated U-Net	0.91710	0.84709	0.95791	0.87976
3.	Proposed Ensemble model	0.95081	0.90625	0.92197	0.98151

Table 7. Comparative Performance of Multiclass Segmentation State of art Methods for Day-5 Human Blastocyst Images

Ref	Dataset	Segmented Classes (Day-5 Blastocyst)	Method	Dice / F1 Score	Jaccard (IoU)
Saeedi *et al.*, 2017 [13]	Static embryo images	ICM, TE, Blastocoel	Texture + CNN	0.85	–
Khosravi *et al.*, 2019 [8]	Static blastocyst images	ICM, TE, Blastocoel	BLAST-NET	0.90	0.82
Rad *et al.*, 2020 [27]	Human embryo images	Trophectoderm, ICM	Inceptioned U-Net	0.91	0.84
Targosz *et al.*, 2021 [19]	Microscopy images	Oocyte regions	CNN-based semantic segmentation	0.88	–
Uysal *et al.*, 2022 [28]	Static embryo images	ICM, TE, Background	U-Net variants	0.86–0.92	–
Arsalan *et al.*, 2022 [29]	Blastocyst images	ICM, TE, Blastocoel	Multiscale aggregation network	0.93	0.87
Wang *et al.*, 2021 [30]	Multifocal embryo images	ICM, TE, Blastocoel	Deep CNN framework	0.92	–
Proposed work	Day-5 static blastocyst images	ICM, TE, Blastocoel, Background	Ensemble of U-Net variants	0.9508	0.9063

This table 7 presents a comparative summary of state-of-the-art multiclass segmentation approaches for Day-5 human blastocyst images, focusing on the accurate segmentation of key embryological structures. The datasets predominantly consist of static or microscopy-based embryo images, with segmentation targets including the inner cell mass (ICM), trophectoderm (TE), blastocoel, and background regions. Various methodological paradigms are reported, ranging from traditional texture-based CNNs and U-Net variants to advanced architectures such as BLAST-NET, Inception-based U-Net, and multiscale aggregation networks. Performance is evaluated using standard overlap metrics, namely the Dice/F1 score and Jaccard index (IoU), enabling direct comparison across studies. The proposed ensemble-based approach demonstrates superior segmentation performance on Day-5 static blastocyst images, achieving the highest reported Dice and IoU scores, indicating improved boundary delineation and class separation relative to existing methods.

6. Conclusion and Future Works

This paper focuses on a method of semantic segmentation of human embryos by means of deep neural networks and an ensemble approach. A performance comparison of different types of convolutional neural networks for semantic embryo segmentation was carried out. The merits and limitations of the selected deep neural networks are discussed. On this basis, an ensemble model is proposed that combines standard U-Net and dilated U-Net architectures with a CNN combiner, which demonstrates superior performance in segmentation tasks. By leveraging the strengths of both convolutional strategies, the model achieves improved class accuracy of 98%, precise boundary detection, and reduced misclassification. The results, summarized in Table 7, indicate that single models such as Inceptioned U-Net] and multiscale aggregation networks achieve Dice/F1 scores in the range of 0.85–0.93, with moderate Jaccard indices. In contrast, the proposed ensemble of U-Net variants achieves superior performance, with a Dice/F1 score of 0.9508 and a Jaccard index of 0.9063, demonstrating improved boundary delineation, reduced misclassification, and robust segmentation across all classes, including the more challenging background regions. The quantitative results show consistent improvements across key metrics such as the Jaccard index and F1 score, with significant gains in challenging regions of the images. Additionally, qualitative comparisons reveal that the model effectively captures intricate details, maintaining clarity and robustness against noise. This approach highlights the potential of ensemble methods in addressing complex segmentation challenges, making it a promising solution for applications requiring high precision and reliability. As a result, the proposed approach can be used to create deep neural network models for semantic embryo segmentation with high accuracy.

Despite the promising results, this study has some limitation like small dataset, even though augmentation and regularization techniques were employed evaluation on larger and multi-centre datasets is necessary to establish broader robustness. Future work could explore optimizing computational efficiency and extending the method to larger and more diverse datasets. Research in this area is ongoing, with scientists developing novel deep learning architectures and image analysis techniques to improve the accuracy and efficiency of human embryo segmentation. As the field progresses, we expect more robust and reliable automated tools that will revolutionize reproductive medicine and enhance our understanding of the earliest stages of human life.

References

[1] World Health Organization. (2025) WHO Fact Sheet, Infertility. https://www.who.int/news-room/fact-sheets/detail/infertility

[2] N. Purkayastha, H. Sharma, Prevalence and Potential Determinants of Primary Infertility in India: Evidence from Indian Demographic Health Survey. Clinical Epidemiology and Global Health, 9, (2021) 162-170. https://doi.org/10.1016/j.cegh.2020.08.008

[3] Ministry of Health and Family Welfare, Government of India. (2021) National Family Health Survey (NFHS-5), 2019–21: India Fact Sheet. International Institute for Population Sciences (IIPS), Mumbai, India.

[4] C.L. Bormann, M.K. Kanakasabapathy, P. Thirumalaraju, R. Gupta, R. Pooniwala, H. Kandula, E. Hariton, I. Souter, I. Dimitriadis, L.B. Ramirez, C.L. Curchoe, J. Swain, L.M. Boehnlein, H. Shafiee, Performance of a Deep Learning-Based Neural Network in the Selection of Human Blastocysts for Implantation. Elife, 9, (2020) e55301. https://doi.org/10.7554/eLife.55301

[5] L.L. Veeck, R.G. Gosden, (1999). An atlas of human gametes and conceptuses: An Illustrated Reference for Assisted Reproductive Technology, Parthenon Publishing Group, New York

[6] A. Pandian, A. Bhattacharya, E. Ozturk, S. Templeton, (2013) Number of Embryos for Transfer Following IVF or ICSI. Cochrane Database of Systematic Reviews. https://doi.org/10.1002/14651858.CD003416.pub4

[7] D.K. Gardner, M. Lane, J. Stevens, T. Schlenker, W.B. Schoolcraft, Blastocyst Score affects implantation and pregnancy outcome: towards a Single Blastocyst Transfer. Fertility and Sterility, 73(6), (2000) 1155–1158. https://doi.org/10.1016/S0015-0282(00)00518-5

[8] P. Khosravi, E. Kazemi, Q. Zhan, J.E. Malmsten, M. Toschi, P. Zisimopoulos, A. Sigaras, S. Lavery, L.A. Cooper, C. Hickman, M. Meseguer, Deep Learning Enables Robust Assessment and Selection of Human Blastocysts after in Vitro Fertilization. NPJ Digital Medicine, 2(1), (2019) 21. https://doi.org/10.1038/s41746-019-0096-y

[9] E.J. Topol, High-Performance Medicine: the Convergence of human and Artificial Intelligence. Nature Medicine, 25, (2019) 44–56. https://doi.org/10.1038/s41591-018-0300-7

[10] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A.W.M. van der Laak, B. van Ginneken, C.I. Sánchez, A survey on Deep Learning in Medical Image Analysis. Medical Image Analysis, 42, (2017) 60–88. https://doi.org/10.1016/j.media.2017.07.005

[11] O. Ronneberger, P. Fischer, T. Brox, U-Net: convolutional networks for biomedical image segmentation, in: N. Navab, J. Hornegger, W. Wells, A. Frangi (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Lecture Notes in Computer Science, Springer, Cham, 9351, (2015) 234-241. https://doi.org/10.1007/978-3-319-24574-4_28

[12] S.E. Roshan, J. Tanha, M. Zarrin, A.F. Babaei, H. Nikkhah, Z. Jafari, A Deep Ensemble Medical Image Segmentation. Computers in Biology and Medicine, 172, (2024) 108305. https://doi.org/10.1016/j.compbiomed.2024.108305

[13] P. Saeedi, D. Yee, J. Au, J. Havelock, Automatic Identification of Human Blastocyst Components Via Texture. IEEE Transactions on Biomedical Engineering, IEEE, 64(12), (2017) 2968–2978. https://doi.org/10.1109/TBME.2017.2664817

[14] T.G. Dietterich, Ensemble Methods in Machine learning, in: Multiple Classifier Systems, MCS 2000, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 1857, (2000) 1-15. https://doi.org/10.1007/3-540-45014-9_1

[15] R. Barkavi, G. Yamuna, C. Jayaram, (2023) Artificial Intelligence: Revolution in Assisted Reproductive Technology, in: S. Kumar, S. Hiranwal, S. Purohit, M. Prasad (Eds.), Proceedings of International Conference on Communication and Computational Technologies, ICCCT 2023, Algorithms for Intelligent Systems, Springer, Singapore. https://doi.org/10.1007/978-981-99-3485-0_76

[16] A.A. Taha, A. Hanbury, Metrics for Evaluating 3D Medical Image Segmentation: Analysis, Selection, and Tool. BMC Medical Imaging, 15(1), (2015) 29. https://doi.org/10.1186/s12880-015-0068-x

[17] M.R. Islam, M.M. Rahman, M.S. Ali, A.A.N. Nafi, M.S. Alam, T.K. Godder, M.S. Miah, M.K. Islam, Enhancing Breast Cancer Segmentation and Classification: An Ensemble Deep Convolutional Neural Network and U-Net Approach on Ultrasound Images, Machine Learning with Applications, 16, (2024) 100555.. https://doi.org/10.1016/j.mlwa.2024.100555

[18] P. Thirumalaraju, M.K. Kanakasabapathy, C.L. Bormann, R. Gupta, R. Pooniwala, H. Kandula, I. Souter, I. Dimitriadis, H. Shafiee, Evaluation of Deep Convolutional Neural Networks in Classifying Human Embryo Images based on their Morphological Quality. Heliyon, 7(2), (2021) e06298. https://doi.org/10.1016/j.heliyon.2021.e06298

[19] A. Targosz, P. Przystałka, R. Wiaderkiewicz, G. Mrugacz, Semantic Segmentation of Human Oocyte Images using Deep Neural Networks. BioMedical Engineering OnLine, 20(1), (2021) 40.https://doi.org/10.1186/s12938-021-00864-w

[20] A. Khan, S. Gould, M. Salzmann, (2016) Segmentation of Developing Human Embryo in Time-Lapse Microscopy. IEEE 13th International Symposium on Biomedical Imaging (ISBI), IEEE, Prague, Czech Republic. https://doi.org/10.1109/ISBI.2016.7493417

[21] B. Dhiyanesh, M. Vijayalakshmi, P. Saranya, D. Viji, EnsembleEdgeFusion: Advancing Semantic Segmentation in Microvascular Decompression Imaging with Innovative Ensemble Techniques, Scientific Reports, 15(1), (2025) 17892.https://doi.org/10.1038/s41598-025-02470-5

[22] F.I. Diakogiannis, F. Waldner, P. Caccetta, C. Wu, ResUNet-a: A Deep Learning Framework for Semantic Segmentation of Remotely Sensed Data. ISPRS Journal of Photogrammetry and Remote Sensing, 162, (2020) 94–114.https://doi.org/10.1016/j.isprsjprs.2020.01.013

[23] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.W. Chen, J. Wu, (2020) UNet 3+: a Full-Scale Connected UNet for Medical Image Segmentation. In: ICASSP 2020

– 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, Barcelona, Spain. https://doi.org/10.1109/ICASSP40776.2020.9053405

[24] K. Kamnitsas, C. Ledig, V.F. Newcombe, J.P. Simpson, A.D. Kane, D.K. Menon, D. Rueckert, B. Glocker, Efficient Multi-Scale 3D CNN with fully Connected CRF for Accurate Brain Lesion Segmentation. Medical Image Analysis, 36, (2017) 61–78. https://doi.org/10.1016/j.media.2016.10.004

[25] D. Saadati, O.N. Manzari, S. Mirzakuchaki, (2023) Dilated-UNet: a Fast and Accurate Medical Image Segmentation Approach using a Dilated Transformer and U-Net architecture, arXiv preprint arXiv:2304.11450. https://doi.org/10.48550/arXiv.2304.11450

[26] H. Nematzadeh, J. García-Nieto, I. Navas-Delgado, J.F. Aldana-Montes, Ensemble-Based Genetic Algorithm Explainer with Automized Image Segmentation: a Case Study on Melanoma Detection Dataset. Computers in Biology and Medicine, 155, (2023) 106613. https://doi.org/10.1016/j.compbiomed.2023.106613

[27] R.M. Rad, P. Saeedi, J. Au, J. Havelock, Trophectoderm Segmentation in Human Embryo Images via Inceptioned U-Net. Medical Image Analysis, 62, (2020) 101612. https://doi.org/10.1016/j.media.2019.101612

[28] N. Uysal, T.K. Yozgatlı, E.N. Yıldızcan, E. Kar, M. Gezer , E. Baştu, Comparison of U-Net based Models for Human Embryo Segmentation. Journal of Information Technologies, 15(1), (2022) 35–44. https://doi.org/10.17671/gazibtd.949430

[29] M. Arsalan, A. Haider, S.-W. Cho, Y.-H. Kim, K.R. Park, Human Blastocyst Components Detection using Multiscale Aggregation Semantic Segmentation Network for Embryonic analysis. Biomedicines, 10(7), (2022) 1717. https://doi.org/10.3390/biomedicines10071717

[30] S. Wang, C. Zhou, D. Zhang, L. Chen, H. Sun, A Deep Learning Framework Design for Automatic Blastocyst Evaluation with Multifocal Images. IEEE Access, IEEE, 9, (2021) 18927–18934. https://doi.org/10.1109/ACCESS.2021.3053098

[31] J.M. Gorriz, F. Segovia, J. Ramirez, A. Ortiz, J. Suckling, Is K-fold Cross Validation the Best Model Selection Method for Machine Learning?, arXiv preprint, (2024). https://doi.org/10.48550/arXiv.2401.16407

[32] R. Azad, M. Heidari, K. Yilmaz, M. Hüttemann, S. Karimijafarbigloo, Y. Wu, A. Schmeink, D. Merhof, (2024) Loss Functions in the Era of Semantic Segmentation: A Survey and Outlook. arXiv:2312.05391. https://doi.org/10.48550/arXiv.2312.05391

[33] D.P. Kingma, J. Ba, (2014) Adam: A Method for Stochastic Optimization. Proceedings of the 3rd International Conference on Learning Representations (ICLR), Banff, 14-16.

Authors Contribution Statement

R. Barkavi: Conceptualization, Methodology, Resources, Writing -Original Draft, Formal analysis. G. Yamuna: Supervision, Writing - Review & Editing, Project administration. C. Jayaram: Supervision, Resources, Data Curation, Writing - Review & Editing. All Authors have agreed to the published version of the manuscript.

Funding

The authors declare that no funds, grants or any other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Data Availability

The data supporting the findings of this study can be obtained from the corresponding author upon reasonable request.

Has this article screened for similarity?

Yes