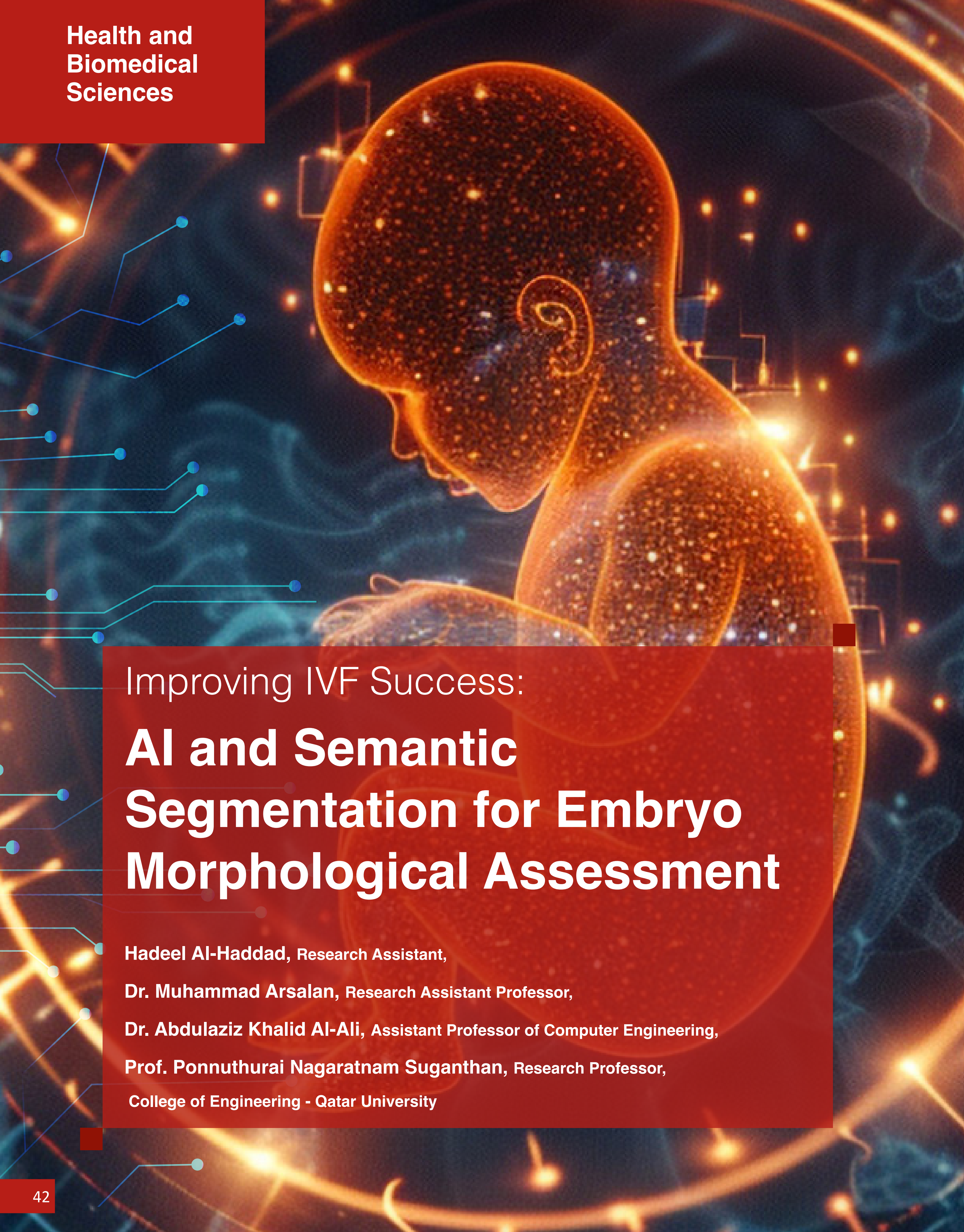

Improving IVF Success:

AI and Semantic Segmentation for Embryo Morphological Assessment

Hadeel Al-Haddad, Research Assistant,

Dr. Muhammad Arsalan, Research Assistant Professor,

Dr. Abdulaziz Khalid Al-Ali, Assistant Professor of Computer Engineering,

Prof. Ponnuthurai Nagaratnam Suganthan, Research Professor,

College of Engineering - Qatar University

Infertility, defined as the inability to achieve pregnancy after a year of unprotected sexual intercourse, is a growing global health concern. According to the World Health Organization (WHO), approximately 17.5% of the adult population worldwide will experience infertility at some point in their lives. This highlights the significance of addressing infertility and exploring effective solutions. Among the various assisted reproductive technologies (ART), in-vitro fertilization (IVF) stands out as a leading option for addressing infertility. IVF involves the fertilization of eggs with sperm outside the body in a controlled laboratory setting. Embryos are cultured until they reach the blastocyst stage before being transferred back into the uterus. Despite its effectiveness, IVF treatment can be financially burdensome for many couples, especially considering that multiple attempts may be needed to achieve a successful pregnancy.

Another challenge is that the IVF procedures involve the transfer of multiple embryos to increase the chances of pregnancy. However, this practice has led to a higher incidence of multiple pregnancies, which come with increased risks for both mothers and babies. Gestational complications, such as preterm birth and low birth weight, are more common in multiple pregnancies.

Objective

To address the risks associated with multiple pregnancies, recent research suggests a shift

Hadeel Al-Haddad

towards transferring a single viable blastocyst during IVF procedures. This approach minimizes the likelihood of multiple pregnancies while maintaining reasonable success rates. By selecting the most viable blastocyst for transfer, the risk of gestational complications can be significantly reduced.

The aim of this research is to improve embryo selection method by incorporating Artificial Intelligence for morphological analysis. The intelligent scheme will reduce human error and provide faster morphometric insight for different embryo components.

Suggested Solution

AI is increasingly utilized across various sectors,

From the left:
Dr. Muhammad Arsalan,
Dr. Abdulaziz Khalid Al-Ali,
and Prof. Ponnuthurai
Nagaratnam Suganthan.

including the medical field, to offer innovative solutions. Semantic segmentation is a branch of deep learning that deals with images by providing pixel-level labels for each pixel in an image. This approach is particularly suitable for our problem for several reasons:

- Provide an aid to the embryologist to analyse the inner structure and the components of the blastocyst.

- Formation of some components like ZP/ICM represents that the embryos are converted into blastocysts, and it is ready to transfer.

- Semantic segmentation helps to identify each component of the blastocyst (pixel-wise) and provide an individual mask for each class (ICM/TE/ZP/BL). Subsequently, these masks can be used to analyse the morphology of these components.

.**Figure 1**: Blastocyst components.

Methodology

Dataset

For this research, a publicly available dataset containing 235 images of human blastocysts with pixel-level annotations of TE (Trophectoderm), ZP (Zona Pellucida), ICM (Inner Cell Mass), and BL (Blastocoel) was used. As shown in Figure 1, these annotations were provided by an expert embryologist. The dataset was divided into a training set comprising 85% (200 images) and a testing set comprising 15% (35 images). To have sufficient training, image augmentation was used to create 6400 images from 200 training images by general image transformations (flipping, rotations, etc.).

Model Building Training

The Figure 2 presents the example of conventional classification network. The proposed approach involved converting a classification network into a semantic segmentation network through the following steps:

- Removing the fully connected layers to adapt to the semantic segmentation model nature.

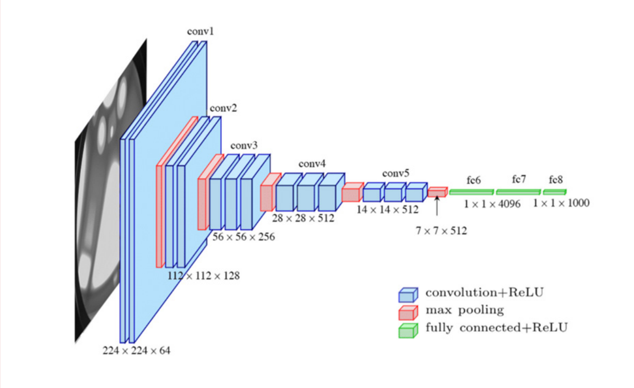

Figure 2: Typical classification model architecture.

- Adding up-sampling layer and ensuring that the output size matches the original image size.

- Adding pixel classification layer to the network in conjunction with a suitable loss function for semantic segmentation tasks.

After experimenting with various networks and architectures such as VGG16 and Shuffle-Net, the proposed network for this task was DarkNet-53. This choice was based on the following reasons:

- It is commonly used in object detection, which proved its suitability for semantic segmentation tasks like blastocyst component segmentation.

- The use of residual skip connections, which preserve the information through the network and help avoid the vanishing gradient problem.

- There is no use of pooling layers, which can help preserve spatial information and prevent loss of detail during feature extraction.

- Using Leaky ReLU as an activation function, which helps avoid information loss by allowing gradients to propagate even for negative inputs, enhancing the network's ability to capture subtle features.

Model Modifications

After testing and experimentation, as shown in Figure3, several modifications have been made to the network architecture:

- The network has been resized to stop at a feature map size of 50x50. This adjustment helps in reducing computational complexity while still preserving essential features.

- Three up-sampling layers have been added to the

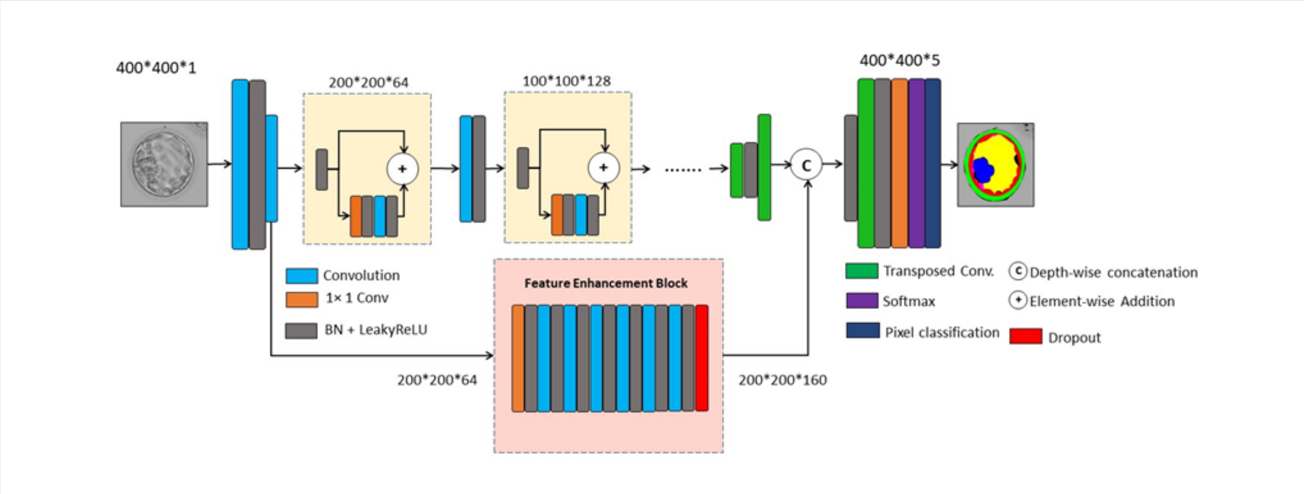

Figure 3: Modified Darknet Architecture.

network to retrieve the original image size.

- Adding Feature Booster Block (FBB), which is designed to retrieve information that may be lost through the down-sampling layers by employing a series of convolutional layers without any pooling layer. This mechanism enhances the network's ability to capture intricate details and improve accuracy.

Evaluation Metrics

The evaluation metric used is the Jaccard Index (JI) with true positive (TP), false negative (FN), and false positive (FN) pixels.

Jaccard Index (IoU) = $\frac{TP}{TP+FN+FP}$

Results and conclusion

Darknet outperforms the mentioned networks, boasting an accuracy of 87.7%. Leveraging Leaky ReLU as the activation function enhances performance by effectively capturing intricate features. Additionally, adding FBB (Feature Poster Block) into the network further enhances results and facilitates the accurate detection of minor classes.

Table 1 shows the accuracy of the proposed AI model's classifications (below) and compares this to other common methods in the field, where numbers closer to 1 are considered more accurate than numbers closer to zero.

The visual results:

The visual results of the proposed method are shown in Figure 4, where red, green, blue, yellow color represent TE, ZP, ICM, BL, respectively.

In conclusion, while our research showed promising insights, it's crucial to conduct further extensive experiments to firmly validate these findings. In the future, collaborations with healthcare institutions in Qatar will provide a great opportunity to acquire a larger and more comprehensive dataset. This step is important to ensure the reliability and effectiveness of our proposed approach.

Table 1: Comparison of the classification accuracy of the proposed artificial intelligence model in the research with previous proposals.

Method	TE	ZP	ICM	BL	BG	Avg
VGG16	0.709	0.810	0.827	0.861	0.951	0.831
VGG16 with skip	0.758	0.827	0.809	0.874	0.954	0.844
ShuffleNet	0.780	0.819	0.856	0.877	0.951	0.857
(ShuffleNet with skip (Two depth encoder	0.779	0.848	0.861	0.884	0.955	0.865
(ShuffleNet with skip (1 residual, 1 skip	0.790	0.841	0.862	0.884	0.951	0.866
ShuffleNet with FBB	0.792	0.836	0.869	0.887	0.951	0.867
DarkNet	0.783	0.845	0.849	0.880	0.955	0.862
Darknet with FBB	**0.808**	**0.855**	**0.876**	**0.900**	**0.960**	**0.880**

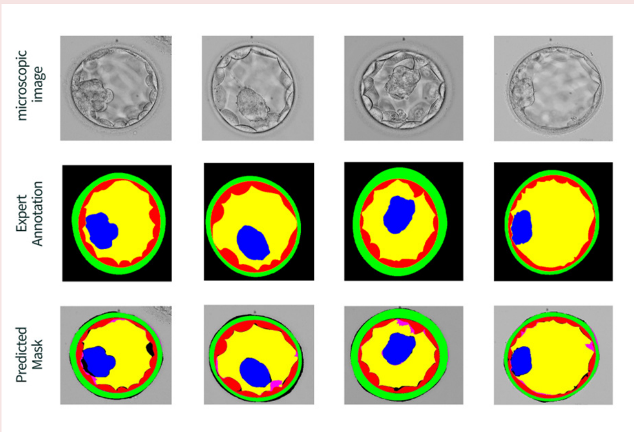

Figure 4: Visual results (comparison between the output of the embryo expert classifications and the output of the artificial intelligence model classifications).