Article

Towards Metric-Driven Difference Detection between Receptive and Nonreceptive Endometrial Samples Using Automatic Histology Image Analysis

**Vidas Raudonis [1,*], Ruta Bartasiene [2] , Ave Minajeva [3], Merli Saare [4,5], Egle Drejeriene [2],
Agne Kozlovskaja-Gumbriene [2] and Andres Salumets [4,5]**

[1] Automation Department, Kaunas University of Technology, LT-44249 Kaunas, Lithuania
[2] Department of Obstetrics and Gynecology, Lithuanian University of Health Sciences, LT-44307 Kaunas,
    Lithuania; ruta.bartasiene@lsmu.lt (R.B.); egle.drejeriene@lsmu.lt (E.D.); agne.gumbriene@lsmuni.lt (A.K.-G.)
[3] Institute of Biomedicine and Translational Medicine, Tartu University, 51003 Tartu, Estonia;
    ave.minajeva@ut.ee
[4] Institute of Clinical Medicine, Department of Obstetrics and Gynecology, Tartu University,
    50406 Tartu, Estonia; merli.saare@ut.ee (M.S.); andres.salumets@ccht.ee (A.S.)
[5] Competence Centre on Health Technologies, 50411 Tartu, Estonia
*   Correspondence: vidas.raudonis@ktu.lt

Abstract: This paper presents a technique that can potentially help to determine the receptivity stage of the endometrium from histology images by automatically measuring the stromal nuclear changes. The presented technique is composed of an image segmentation model and the statistical evolution of segmented areas in hematoxylin and eosin (HE)-stained histology images. Three different endometrium receptivity stages, namely pre-receptive, post-receptive, and receptive, were compared. An ensemble-based AI model was proposed for histology image segmentation, which is based on individual UNet++, UNet, and ResNet34-UNet segmentation models. The performance of the ensemble-based technique was assessed using the Dice score and intersection over unit (IoU) values. In comparison to alternative segmentation architectures that were applied singly, the current ensemble-based method obtained higher Dice score (0.95) and IoU (0.90) values. The statistical comparison highlighted a noticeable difference in the number of nuclei and the size of the stroma tissue. The proposed technique demonstrated the positive potential for practical implementation for automatic endometrial tissue analysis.

Keywords: image segmentation; endometrial biopsy; histology image

1. Introduction

Infertility is the inability of a mature organism to produce offspring. The inability to become pregnant after 12 months or more of frequent, unprotected sexual activity defines this condition of the male or female reproductive system [1]. It is estimated that millions of people and couples worldwide struggle with infertility. According to estimates, one in six persons who are of reproductive age are impacted [1]. The prevalence of infertility accounts for 17.5% of the worldwide population [2]. Medically assisted reproduction (MAR), or in vitro fertilization (IVF), was developed to overcome the problem of infertility and was introduced into clinical practice more than 40 years ago. MAR technologies enable the fertilization of human oocytes and the creation of human embryos in vitro. Viable embryos can then be transferred into the woman's uterus, and the development of intrauterine pregnancy begins upon successful implantation. Since the beginning of successfully applying IVF in humans, these treatments have resulted in the birth of over 10 million children already. Several countries now see more than 3% of all births as a result of MAR [3]. However, the success of MAR remains limited despite the rapid development of related technologies [4–6].

Changes in the cellular or molecular composition of the innermost lining of the uterus, known as the endometrium, are critical for IVF success; its unique structure and receptivity allow the embryo to be implanted [7,8]. The highest receptivity of the endometrial tissue occurs during the so-called "window of implantation" (WOI) when optimal conditions for embryo implantation are established. Prior to the WOI or past this time, the chance of embryo implantation is scarce. Therefore, an endometrial receptivity assessment has proven to be advantageous for the success of IVF. Transcriptomic biomarkers are most commonly used to assess endometrial receptivity, as the expression levels of these biomarkers depend upon the receptivity status of the uterine tissue [9,10]. Hence, endometrial tissue transcriptome testing allows an optimal time for the personalized embryo transfer to be chosen [11,12]. In parallel, endometrial microbiota changes have also been shown to affect embryo implantation, which can be evaluated by tissue profiling during IVF [13,14]. However, the most common and widely used endometrial tissue evaluation is performed according to the Noyes dating criteria for histological sections [15]. These criteria are valid to date and are used in clinical pathology: when assessing the endometrial phase, both the glandular and stromal parameters must be evaluated [16,17]. Evaluating stromal cell changes, such as the extent of edema, is subjective, and very small-sized stromal cell nuclei cause further difficulties in their evaluation under standard light microscopy.

The clinical value of histological endometrial dating, which defines the individual WOI, has recently been shown to benefit patients with recurrent pregnancy loss [18]. Nevertheless, implementation of the latest technical breakthroughs in deep learning to establish AI-powered technologies to optimize histopathology assessments of endometrial tissue for IVF is still clearly lacking. Therefore, AI-driven clinical decision support tools for endometrial tissue assessments may offer efficient, metric-driven, objective, and accurate diagnoses for endometrial receptivity evaluation. The main aim of this research is to propose a computer-aided system for the evaluation of endometrial tissues from endometrial histological slides, which can be used as an automatic algorithm for endometrial receptivity assessments and predictions. This research paper is organized into seven main sections. The state-of-the-art in the current research topic is provided in the second section. Information about the research object, the proposed AI model, and its validation are given in the third, fourth, and fifth sections accordingly. The experimental setup and results are shown in the seventh section. The discussion, conclusions, and references are provided at the end of the research paper.

2. State-of-the-Art Review

Endometrial histologic dating criteria, such as thickness, glandular density, and maturation, are important in IVF, being crucial for the successful implantation of an embryo. These criteria are important predictors of the pregnancy outcome in IVF cycles. The endometrium is the innermost layer of the uterus and is the site of implantation of the fertilized egg during pregnancy. For the development of automated segmentation techniques for histologic images, several approaches merging computer vision and machine learning technologies have been put forth over the years. The automatic evaluation of histology images mainly depends on the target objects that have to be evaluated, such as cells, nuclei, glands, or tissue type. An automated segmentation algorithm based on morphology is first applied to light microscopy images stained with Hematoxylin and Eosin [19].

Image segmentation is the process of locating certain objects in an image or splitting it into several areas, each of which represents a different object or portion of the image. For the purpose of segmenting or dividing up images, deep learning techniques such as autoencoders like UNet are utilized. Segmentation is based on the transformation of a high-dimensional space to a lower-dimensional space [20]. Usually, two-dimensional data are down-sampled to a one-dimensional vector and again decoded to its original appearance. Researchers proposed several types of autoencoders, which are applied for image segmentation tasks, including plain fully convolutional autoencoders, multi-loss autoencoders, fully convolutional variational autoencoders, and repeated fully convolutional

autoencoders. Different autoencoders distinguish themselves with their own strengths and weaknesses. Therefore, the selection of a specific structure of the segmentation model depends on the task and image characteristics. Gondim, with a group of researchers, proposed an assisting analysis tool that is used to examine the vascular network of bone histologic images. The proposed automatic method is based on a convolutional neural network (CNN) and random forest classifier (RFC) [21]. Novel segmentation methods are applied in the detection of individual nuclei in stained breast histopathology images [22] and segment whole biopsy slides [23]. The latter method employs a U-Net-like neural structure that was trained on a large annotated dataset. Ali et al. proposed a multiscale attention learning network (MSAL-Net) to segment individual nuclei in histopathology images [24]. The proposed methods include mathematical morphology, convolutional neural networks, random forest classifiers, and multiscale attention learning networks. These methods have shown promising results in the automatic segmentation of histology images. Several variants of U-Net have been proposed, each with its own strengths and weaknesses. Many limitations can be overcome by combining different variants of U-Net models and their strengths.

Our research investigates the potential of applying automatic histology image analysis for differentiating between receptive and nonreceptive endometrial samples using a metric-driven approach. Traditional assessment methods rely on a subjective evaluation by pathologists, introducing potential inconsistencies. Automatic image analysis offers an objective and quantitative alternative. This approach involves extracting the relevant metrics from histological images, which can then be used to distinguish between receptive and nonreceptive samples. This work explores the feasibility of this approach and paves the way for developing a reliable and automated system for endometrial receptivity assessment.

3. Research Object and Dataset

In this work, custom datasets of endometrial biopsy (these were stained with HE (Hematoxylin and Eosin)) histologic pictures in color were used. This dataset was produced as part of the ReadUteru initiative, an international effort that includes medical facilities in Lithuania and Estonia. The dataset included 150 whole-slide photos in color of endometrial biopsies (see example in Figure 1), together with manually segmented images (ground truth annotations). The endometrial samples from women receiving IVF therapy were divided into pre-receptive, receptive, and post-receptive groups using molecular profiling with the commercial beREADY test. Additionally, the majority of these biopsies were taken during the hormonal replacement treatment (HRT) IVF cycles.

Segmentation models are employed in order to divide the histological images into three different areas, such as endometrial glands, stroma, and other parts of the endometrial sample. The latter one is segmented based on the color, which is usually almost completely white. Areas with endometrial glands are detected using an AI model, and everything that is not glands or background is segmented as stroma. The endometrial glands' locations are indicated by the dark (grayish) regions (Figure 1 right). A team of pathologists from two countries created ground truth images (annotated photos). At various magnification settings, experts examined and annotated each whole-slide image. The original color whole-slide pictures are the same size as the ground truth segmentation maps. When using a $20\times$ magnification, the size of a gland is typically not greater than 350 pixels. This region is much smaller than the total image—more than 100 times smaller.

The whole-slide images of endometrial biopsy cannot be used as training inputs because of the limited computational resources. The appropriate image size is 576×576 pixels when taking into account the available computing power and segmentation accuracy. The choice of image size is influenced by the computational resources that are available for image processing. Smaller image sizes can be processed quicker and require less memory, making them more suitable for resource-constrained environments. Additionally, the selection of such size is influenced by the desired balance between the image resolution and the processing speed. Larger image sizes can provide higher resolution inputs but

at the cost of increased processing time and computational resources. The image parts that correlate with the gland position are also reduced up to 30 times due to the direct size reduction of the original whole-slide image to 576 × 576 pixels. It is challenging to segment such a small area because textural and structural information about the glands will be lost. We suggest using cropped regions of interest (RoI) of the whole-slide image instead of shrinking it, with each RoI being selected using an overlapping sliding window of size (576 × 576 pixels).
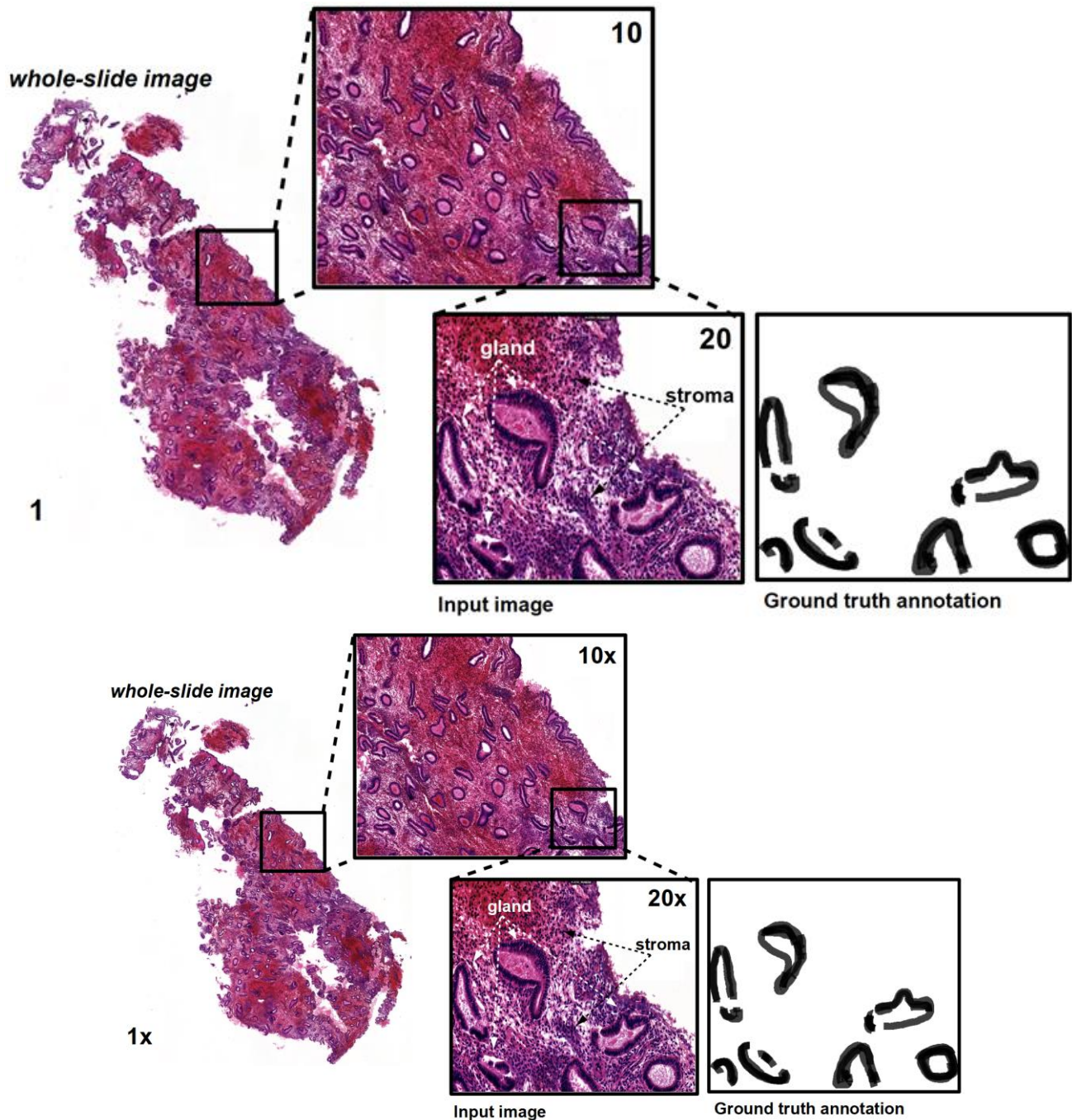
Figure 1. An example of a color, HE-stained whole-slide image of endometrial biopsy captured at different magnification levels 1, 10, and 20.

The sliding window overlapping technique is a data augmentation method often used in image processing. This technique involves creating overlapping windows of the

original data to generate additional training samples for machine learning models. The overlapping windows help to capture more diverse patterns and variations in the data, thereby enhancing the model's ability to generate accurate segmentation predictions and partly solve overfitting problems. The sliding window with the overlap data augmentation technique involves using a fixed window length to extract samples from the original histology image with a certain degree of overlap with the consecutive window. We defined the 30% overlap between RoIs in our study. There are various benefits of using RoIs rather than whole-slide histology pictures. First, up to 15,000 pictures (RoIs with only black pixels are eliminated) are generated for training and testing, which is a 100 times increase compared to the required number of samples (from 150 whole-slide images). Second, there is a decrease in the demand for computing resources. Third, the endometrial glands retain their original size and shape. Overall, this technique contributes to improving the robustness and performance of machine learning models by enriching the training dataset with diverse samples.

The evaluation of the endometrium tissue is accomplished by estimating the size and density of the nuclei displacement. Firstly, we estimate the density of the nuclei next to the detected endometrium glands and, secondly, in the whole endometrium tissue (see Figure 2). Detection is conducted using an ensemble-based segmentation model that aggregates the outputs of three individual models (more info in Section 4). The resulting output is a segmentation map where the endometrium glands and all other tissues are segmented.

Metric-driven analysis is performed based on a segmentation map, which highlights the positions of the endometrium glands. The nucleus density is calculated using two grids of different densities, such as 30 and 50 μm (see Figure 3). Nuclei in the endometrium image are HE stained; therefore, they appear as blueish areas in the image, therefore, we estimate the size of a nucleus as a detected bluish area in the histologic image.
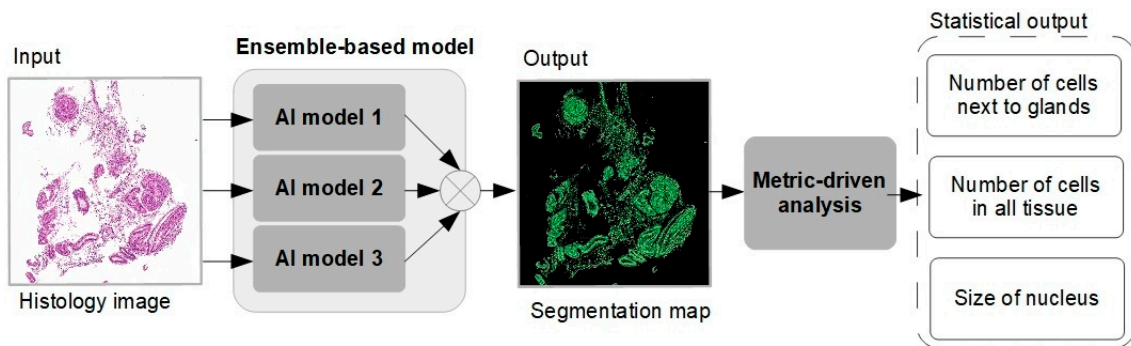

Figure 2. The block diagram of proposed evaluation method of endometrium tissue.
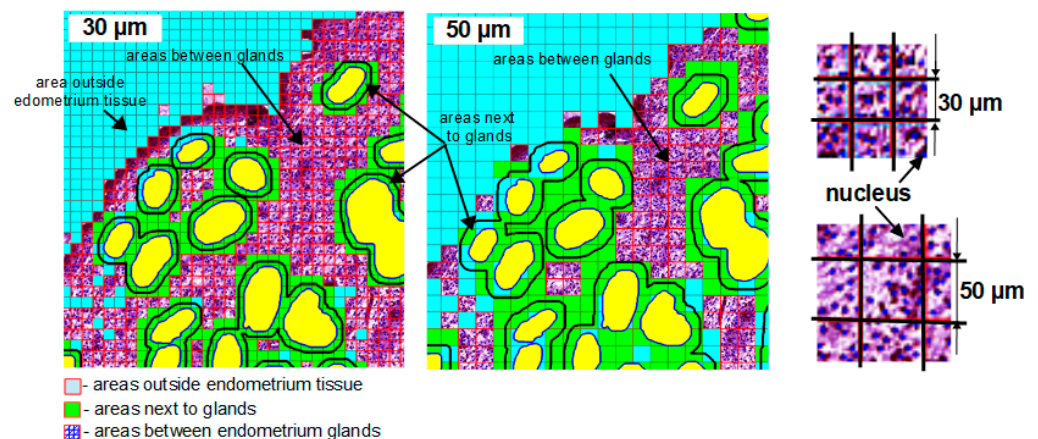

Figure 3. Grid-based estimation of nuclei density in the endometrial image.

4. Endometrial Tissue Segmentation Model

In the following section, the proposed gland detection method is described. The application of deep encoder–decoder CNN architectures of different depths and complexity was inspired by the complex appearance of the target object. The uterus endometrium has glands that highly vary in size, shape, and structure. These glands are tubular in shape and have ciliated columnar epithelium lining them. Due to the estrogen release and depending on the menstrual cycle, the appearance of the glands is changing. In this study, the ensemble-based detection model is investigated and built using segmentation model variants such as U-Net, U-Net++, and Residual U-Net. These models are selected due to their ability to capture intricate features and patterns within medical images, making them suitable for segmentation tasks. When used on various histology images, each segmentation model has a certain level of accuracy that it can achieve. These restrictions may manifest as slices of endometrial biopsies that are not of the same thickness, are not evenly lit, have varied image noise that may obscure key imaging findings or are simply taken at different times of the menstrual cycle. To increase accuracy and produce results that are reliable, all models that were trained on the same dataset are combined into one ensemble. The risk of complex model overfitting is managed through the size of the training dataset and the employment of regularization techniques. To mitigate overfitting, techniques such as dropout, batch normalization, and data augmentation were used. The resulting output, i.e., the predicted segmentation map of endometrial glands, is acquired using a weighted sum of all three individual segmentation maps. The weights are calculated by minimizing the criterion using the pseudo code given in Figure 4. The inverse of the intersection over unit (IoU) value is used as an optimization criterion in our work (1-IoU). The optimization starts from generating α and β values in the range from 0 to 1. The gamma (γ) weight is estimated as the difference between 1 and the sum of the α and β values. The weighted model output or prediction P is calculated in step 4, where $P_{\text{UNet++}}$ is the prediction map given by the UNet++ model, $P_{\text{Resnet34-UNet}}$ is the prediction map given by the residual model, and P_{UNet} is the predictions of the UNet segmentation model. The intersection over unit value is estimated between manually annotated mask M and weighted prediction P. Optimized weights are estimated when the criterion meets the condition given in step 5. Final weights α^*, β^*, and γ^* are given in step 6. The final prediction map is calculated using the given weights: $\alpha = 0.346$, $\beta = 0.250$, and $\gamma = 0.404$.

```
1. Generate weights in range from 0 to 1: α = [0..1], β = [0..1]
2. For i, j in range(α, β)
3.     γ_{i,j} = 1 − (α_{i,j} + β_{i,j})
4.     P_{i,j} = P_{UNet++} · α_{i,j} + P_{Resnet34−UNet} · β_{i,j} + P_{UNet} · γ_{i,j}
5.     If Λ > min(1 − IoU(M, P_{i,j})) Then
              i,j
6.              Λ = min(1 − IoU(M, P_{i,j})),  α* = α_{i,j},  β* = β_{i,j},  γ* = γ_{i,j}
                  i,j
```

Figure 4. Pseudo code for calculating impact weights of each segmentation model in ensemble-based approach.

Figure 5 displays the proposed ensemble's functional diagram. The ensemble comprises the three deep neural network models ResNet34-UNet, UNet, and UNet++. The segmentation model processes the input image, and as a result, an output segmented image of the same as the input size is produced [25], called a segmentation map. The deep neural network encodes input data (color image) through the down-sampling path, decodes a segmentation map through the up-sampling path, and skips connections. A max-pooling layer and a sequence of convolutional layers with decreased size are used to build the down-sampling route. The down-sampling route effectively reduces the high-resolution image to low-resolution. The contextual information is extracted in the left part of the network (down-sampling path), and the spatial information is reconstructed in the right part of the network (up-sampling path). The output map, or segmentation map of the endometrial glands, is produced using the U-Net network's up-sampling route and is the same size

as the input picture. A high-resolution picture was created in the up-sampling step by decoding the low-resolution image (or features). The color images of size 576×576 pixels are used as inputs. The size of the input images is gradually reduced; depth is gradually increased until the size of the features reaches $18 \times 18 \times 512$ pixels. The middle convolutional layer is used as a bottleneck, and it is positioned between the encoder and decoder paths. The "bottleneck" consists of convolutional and batch-normalization layers. Through the concatenation of feature maps from the down-sampling side to the up-sampling side, skip connections are employed to transmit local features. The final output layer consists of a convolutional layer with a softmax activation function and a convolutional filter size of 1×1. The resulting output segmentation map consists of grayish blobs on a white background, which correspond to the area of the endometrial glands in a colored, HE-stained histology image. Adam optimizer and binary cross-entropy (BCE) loss function were used to train a U-Net network with 7.7 trainable parameters.
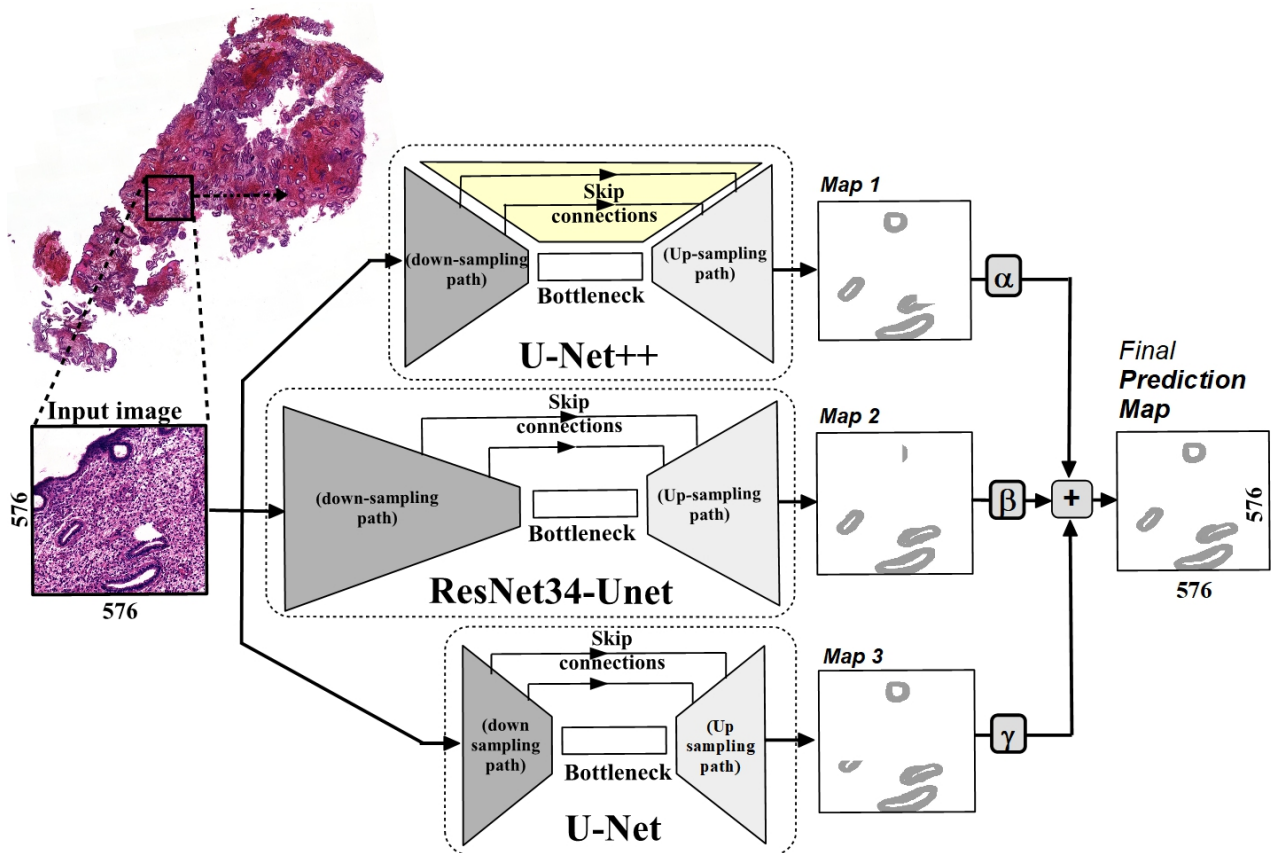
Figure 5. An illustration of the proposed ensemble-based segmentation model's functionality.

A U-Net++ type network serves as the second deep neural network in the ensemble of models. Nested and thick linkages serve as the model's foundation. It has skip connections that incorporate convolutional layers for passing semantic information from the encoder (down-sampling) to the decoder (up-sampling) paths. The dense skip connections are used to improve the gradient flow, i.e., training [26]. The U-Net++ model is efficient at collecting the fine features of 2D pictures used to define endometrial glands. The number of convolutional blocks is controlled by deep supervision. It prunes and adjusts the complexity of the network, keeping high performance (segmentation accuracy) and inference time. A segmentation map with a resolution of 576×576 pixels is the output of the U-Net++ model. This model was trained using the Adam optimizer with the loss function BCE and has around 9.1 million trainable parameters. The third pretrained ResNet34-Unet network is used as a part of the ensemble-based method. The ResNet34-based network is used

as the semantic features encoding part of the segmentation model (the down-sampling path). A residual network is built from serially connected residual blocks that have skip connections. The enhanced attention module and a decoder block (the up-sampling route) make up the network's decoding section. There are two connections from the input for each residual block. Convolutional layers, batch normalization, and linear functions are applied to the initial input. The second input omits the further addition of the convolutions and functions. In order to preserve segmentation accuracy and resilience, residual networks assist in resolving the vanishing gradient problem that is typical in big and complicated networks. The residual U-Net model's up-sampling path resembles that of U-Net from the preceding description. Through skip connections, the up-sampling component receives the local data and concatenating features. This model was trained using the Adam optimizer with the loss function BCE and has around 21.6 million trainable parameters.

All mentioned models and ensemble-based segmentation methods segment only a small portion of the whole biopsy image. In order to reconstruct all the segmentation maps that correspond to the whole biopsy image, several processing steps have to be undertaken. There are three basic processes in this reconstruction flow: Each patch is stitched back onto the canvas in the same way that it was cropped from the original picture, and average prediction values are generated in the areas where patches overlap. The predictions are obtained from each image patch using the proposed segmentation model. Finally, a threshold value is used to determine the segmentation map of the expected glands. For purposes of paper representation, white was used to designate the pixels for which predictions were higher than 0.5; otherwise, white was used to mark the pixels.

5. Validation Metrics of Tissue Segmentation Model

The proposed segmentation model is evaluated using intersection over union (IoU) and Dice coefficient criteria. IoU is a performance metric often used in object detection and image segmentation applications to measure the accuracy of predicted bounding boxes. IoU estimates the intersection rate between predicted areas (segmentation maps produced by neural networks) and ground truth (provided by manual labeling). The range of intersection over union values is 0 to 1, with 0 denoting no overlap and 1 denoting a perfect overlap. In the case when comparative pixel-by-pixel analysis is performed, IoU is the first metric in the list, and it is defined using Formula (1).

$$IoU = \frac{TP}{(TP + FP + FN)}, \tag{1}$$

where *TP* is a true positive, *FP* is a false positive, and *FN* is a false negative (see Table 1 for a deeper explanation). The *TP*, *FP*, and *FN* terms are used to evaluate the object detection results because these values correspond to image areas (or the number of pixels).

Table 1. Model performance metrics.

True Positive (TP)	$TP = GT \wedge S$	It is a region of intersection between the manually labeled ground truth picture (GT) and the predicted segmentation map (S), which is the model output.
False Positive (FP)	$FP = (GT \vee S) - GT$	False positive refers to the anticipated area in the final segmentation map that is outside of the experts' provided ground truth areas.
False Negative (FN)	$FN = (GT \vee S) - S$	False negative refers to the quantity of pixels in the segmentation model's unpredicted area of the ground truth.

The Dice coefficient (DC) is a metric used in image segmentation and object detection tasks that measure the similarity between predicted and ground truth maps (the output of the models). The DC value ranges between 0 and 1, where 0 indicates the overlap between GT and S, and 1 indicates a perfect overlap (100% similarity). The Dice scores can be expressed through definitions of precision and recall (see Table 2).

Table 2. Description precision and recall terms.

precision	$\frac{TP}{TP+FP}$	Precision in object detection is the fraction of true positive detection among all positive detections.
recall	$\frac{TP}{TP+FN}$	Recall refers to the fraction of relevant instances that were successfully retrieved by the object detection system.

A high degree of accuracy means that the detection model produces fewer false positive detections than real positive detections. Recall gauges how well the endometrial glands were all properly identified in the histology picture using the glands identification model. Using Formula (2), the Dice coefficient is calculated. As recall and accuracy are averaged, *DC* reveals the model's overall performance or its capacity to identify and accurately categorize items in an image.

$$DC = \frac{2}{1/_{recall} + 1/_{precision}} \tag{2}$$

6. Experimental Setup and Results

In total, 15,000 annotated image sections were created by cropping whole-slide images with fixed-size sliding windows (see Section 4). All image samples of whole-slide histology images were taken from different menstrual cycle phases. The prepared training dataset is used to train the model to segment endometrial glands, irrespective of the menstrual cycle. The whole-slide picture parts from the custom dataset were split into the following groups at random: 70% of the data were provided for model training, 10% for model validation, and the remaining 20% were utilized for model testing. An Intel i7 CPU and an Nvidia RTX4090 GPU graphics card were used in a desktop PC for all computing, training, and testing. The segmentation results of the endometrial glands are shown in Figure 6. Several random image regions are shown that are taken from the test dataset. Additionally, from left to right, color image regions of a real histology image, ground truth annotation, and the model outputs are given (the U-Net, ResNet34-UNet, U-Net++, and ensemble-based methods) accordingly. In the given figure, endometrial glands are highlighted with dark-color pixels and the surrounding tissues with white-color pixels.

6.1. Histology Image Segmentation Results

The experimental findings show that various segmentation models provide various prediction outcomes or anticipated segmentation maps. In some instances, endometrial glands are appropriately and reliably segmented; however, in other instances, studied models were unable to detect fuzzy endometrial gland borders. On the other hand, combining the weighted predictions from three independent models produces outcomes that are superior to those of the individual models (see Figure 6).

The average IoU was 0.90 according to the ensemble-based technique. Among the three segmentation networks, this result is the best. When all three separate models are unable to reliably identify the endometrial glands from surrounding tissues, the IoU value of the ensemble-based model plummets dramatically. A Dice score of 0.95 on average was shown by the ensemble-based model. Comparing the ensemble-based approach to the other three deep neural networks, the given Dice score of the ensemble-based technique is the best (Table 3). When none of the three models successfully separate endometrial glands from their surroundings, such as other tissues, the Dice score of the ensemble-based technique suffers significantly. Table 3's bolded text highlights the examined approaches' best segmentation performance. The suggested ensemble-based strategy produced the best segmentation results, with IoU and Dice scores of 0.90 and 0.95, respectively. On the customized dataset, the U-Net deep neural network had the lowest accuracy, with IoU = 0.77 and Dice = 0.85 as a result.
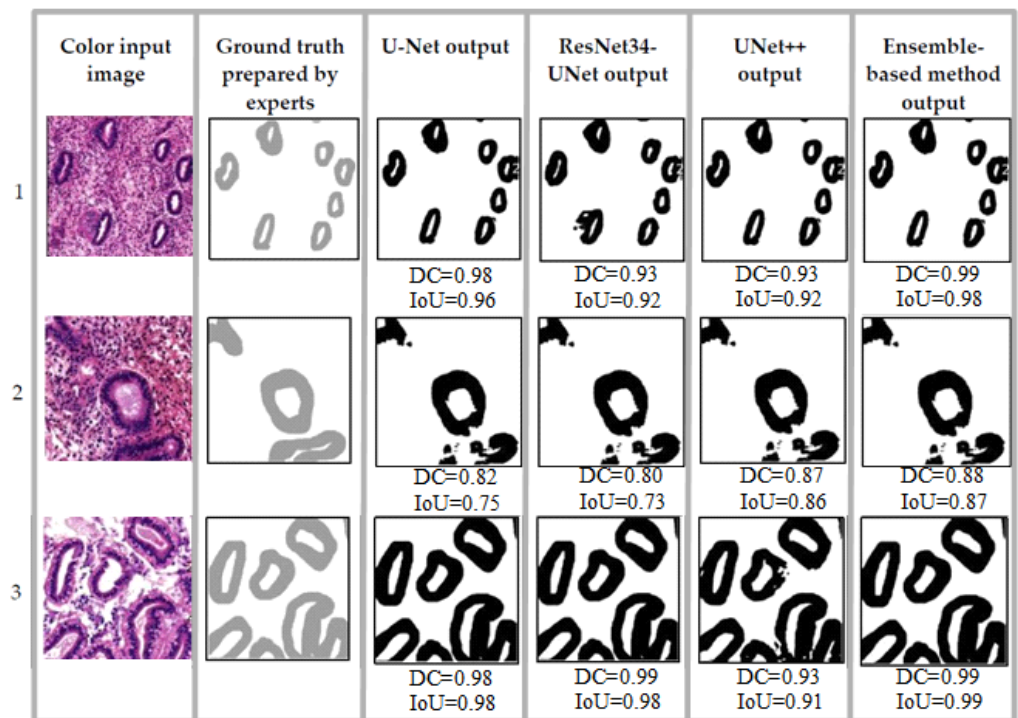
Figure 6. (From left to right): Original image section, corresponding ground truth annotation, endometrial glands segmentation results using the U-Net model, endometrial glands segmentation results using the residual U-Net model, endometrial glands segmentation results using the U-Net++ model, and endometrial glands results using the ensemble-based method.

Table 3. The ensemble-based model and its average output from three neural networks.

Performance Metric	U-Net	RestNet34-UNet	U-Net++	Ensemble
IoU	0.77	0.79	0.85	0.90
Dice Score	0.85	0.88	0.89	0.95

Figure 7 shows the detected positions of endometrial glands on whole-slide images obtained using the ensemble-based technique (the histology image is zoomed in for clarity). Above a colored, HE-stained histology picture, blue contour color patches indicate the probable sites of the endometrial glands.
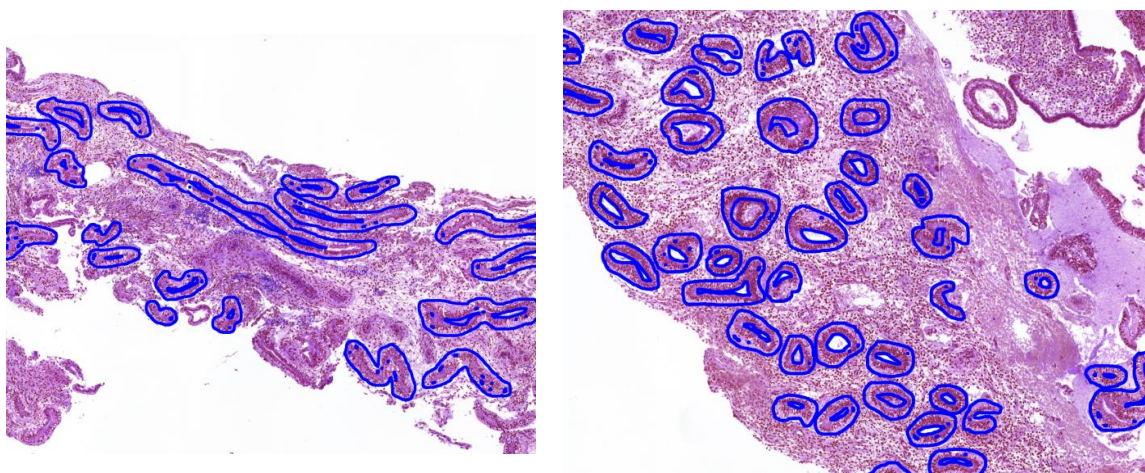
Figure 7. Example of histology image with detected endometrial glands.

6.2. Metric-Based Evaluation Results of Endometrial Sample

The size, density of the nucleus, and histograms of these measurements in the endometrial sample were chosen as objective metrics, which can be applied for the receptivity assessment. The size of the nucleus in the stroma tissue is equal to the blueish area of the HE-stained image. Squared pixel values are translated to squared micro-meters. Our investigation has shown that the stromal cell nucleus in the endometrial sample is of a size ranging from 0.005 to 0.68 μm^2 in the post-receptive stage. The nucleus size from the pre-receptive stage varies from 0.005–0.90 μm^2, and in the receptive stage, from 0.005–0.67 μm^2. The histogram of the nucleus size is shown in Figure 8, where the rational number of nuclei is given on the vertical axis, and the size of the nucleus is set aside on the horizontal axis. The black color curve represents the receptive stage, the red dashed line represents the pre-receptive values, and the post-receptive stage values are presented with the blue curve. The histogram suggests that the nucleus is bigger in the post-receptive stage when they are compared with the nuclei from other stages.

Figure 9 represents the rational histograms of the stromal cell's nucleus derived using (a) a 30 μm grid, (c) a 50 μm grid, (b) at the 30 μm range from endometrial glands, and (d) at the 50 μm range from endometrial glands. The number of grid cells is given on the vertical axis, and the number of stromal nuclei is shown on the horizontal axis. The histograms show that the highest nucleus density is higher in the receptive samples. Moreover, more nuclei can be detected next to the endometrial glands (these being the nuclei of stromal cells): for instance, on average, up to 12 nuclei can be detected next to the glands and, on average, 4 in all tissue areas.
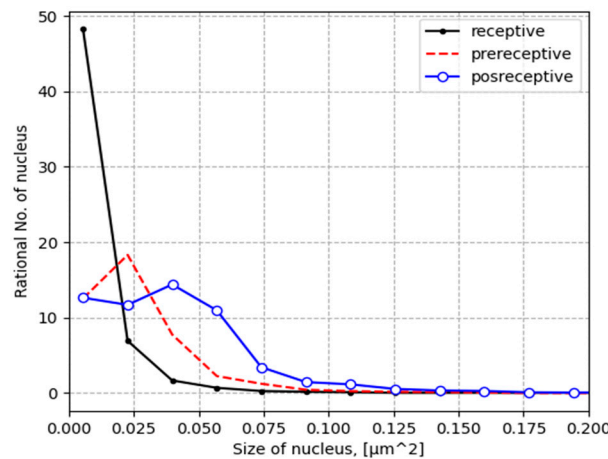

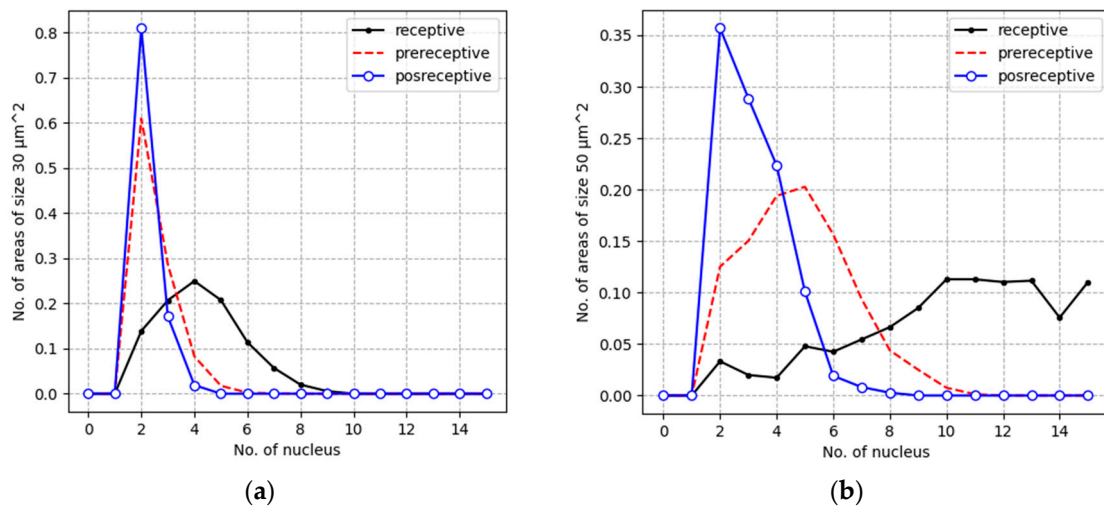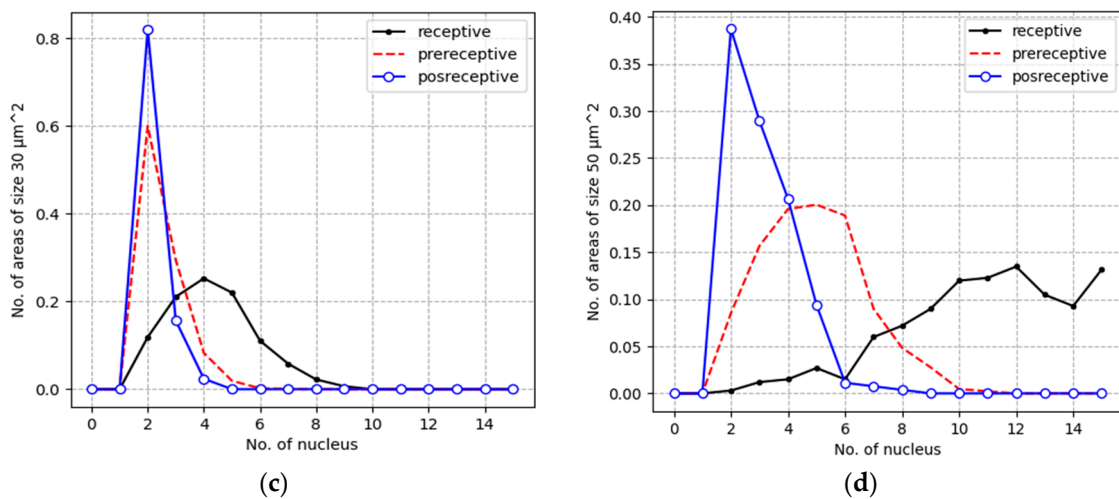**Figure 8.** Histogram of the stromal cell nucleus size.

(a)   (b)

Figure 9. *Cont.*

Figure 9. The rational histograms of the stromal cell's nucleus are derived using (**a**) a 30 μm grid, (**c**) a 50 μm grid, (**b**) at the 30 μm range from endometrial glands, and (**d**) at the 50 μm range from endometrial glands, where the black color curve represents the receptive stage, the red dashed line represents pre-receptive values, and the post-receptive stage values are presented with the blue curve.

7. Discussion

For the purpose of the assessment of receptive and nonreceptive endometrial stages, we proposed a metric-driven approach, which is based on an image segmentation model and evaluation of nucleus size and density. Utilizing a customized dataset obtained through collaboration with Tartu University and the Lithuanian University of Health Sciences, the proposed technique and three additional deep architectures were tested. There are several issues that can induce misdetection and misclassification. The quality of the image collection used to segment the endometrial glands has a direct influence on the accuracy of the segmentation. It should be noted that various specialists may annotate the identical whole-slide picture of an endometrial biopsy differently due to differences in opinion on the borders of the glands. The accurate segmentation of endometrial tissue is a difficult task that requires several years of practice to master if it is conducted manually. Furthermore, histological image aberrations, uneven biopsy slide lighting, and image noise can all contribute to erroneous segmentation of endometrial glands.

The minimum DPI (dots per inch) for input images is important for our endometrium segmentation model's accuracy. Higher resolution generally leads to better results. However, the computational resources should be considered. The selection of image size is influenced by the desired balance between image resolution and processing speed. In our research, we selected an image resolution of 4 pixels per 1 μm. The ensemble-based method's attained detection accuracy comes at the expense of processing time. Since the ensemble technique is based on three deep neural networks, it takes three times longer to provide predictions. On a GPU graphics card, three separate segmentation models were trained and tested. The Nvidia RTX4090 with 24 GB of internal RAM was utilized. The identical custom training dataset and hyperparameters were used to train all three neural networks. To train the U-Net autoencoder that consists of 7.7 trainable parameters, less than 22 h is needed. The training of residual U-Net (21.6 trainable parameters) required up to 50 h. On average, 29 h are needed to train the U-Net++ (9.1 trainable parameters) segmentation network. In total, it took 105 h (less than 5 days) to train all ensemble-based segmentation models. Several strategies can be employed to potentially improve the overall performance when all three models demonstrate poor outcomes. The proposed endometrial assessment method can be used for a deeper analysis of images. By using the metric-driven approach, we could objectively demonstrate that the receptive endometrium had a higher density of stromal cell nuclei next to the endometrial glands compared to the pre- or post-receptive phases. This is in accordance with the generally applied pathohistological dating

criteria that describe increasing stromal edema in the post-receptive endometrium as the fluids push the stromal cells further apart. An interesting additional finding is the smaller stromal cell size in the receptive stage compared to the pre- and post-receptive stages. The pathohistologist is normally incapable of routinely measuring the stromal cell sizes, which are very small. Ensemble methods enhance histology image segmentation by combining predictions from multiple models. This combination leads to higher accuracy, capturing intricate tissue details. It also boosts robustness, making segmentation more reliable despite image noise or staining variations. The ensemble further controls overfitting to specific training data, enhancing its ability to handle new histology images. These advantages make ensemble segmentation a positive choice for digital pathology tasks like cell classification, tumor analysis, and tissue composition studies. The next step would be the application of an automatic algorithm that is capable of evaluating the differences between endometrial samples that are in the receptive stage and not. The endometrial gland's structure and size changes during the menstrual cycle and glands from post-receptive samples are different from pre-receptive glands. A computer-aided receptiveness forecasting method should be developed for further use in this practical application.

8. Conclusions

For computer vision and AI researchers, the automated segmentation, recognition, and assessment of endometrial tissues remains a difficult task. In this work, we suggest using a metric-driven approach based on ensembles to automatically differentiate endometrial glands using group predictions. In order to create a training and testing dataset that is more statistically significant and to lower the computing resource needs, we also suggest an image processing phase. We used a dataset amassed during actual clinical practice to construct our experimental research. The suggested ensemble-based technique showed comparatively high endometrial gland identification accuracy. The ensemble-based model outperformed other designs, including U-Net, ResNet34-UNet, and U-Net++, with higher Dice score (0.95) and IoU (0.90) values. The metric-driven approach highlighted that the nuclei density is higher in the receptive samples. Endometrial stromal cell nuclear size and cell-to-cell distance are potential candidate criteria for the metric-driven automated evaluation of the endometrial phase. Moreover, more nuclei can be detected next to endometrial glands than in other parts of the tissue. Future studies will focus on the segmentation of more endometrial cells and/or structures, such as stroma and blood vessels. The current dataset will also be augmented with new whole-slide images of endometrial biopsies and expert-based annotations in order to build a better and more precise segmentation algorithm.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1.   World Health Organization (WHO). *International Classification of Diseases, 11th Revision (ICD-11)*; WHO: Geneva, Switzerland, 2018.
2.   Cox, C.M.; Thoma, M.E.; Tchangalova, N.; Mburu, G.; Bornstein, M.J.; Johnson, C.L.; Kiarie, J. Infertility prevalence and the methods of estimation from 1990 to 2021: A systematic review and meta-analysis. *Hum. Reprod. Open* **2022**, *2022*, hoac051. [CrossRef] [PubMed]
3.   Wyns, C.; De Geyter, C.; Calhaz-Jorge, C.; Kupka, M.S.; Motrenko, T.; Smeenk, J.; Bergh, C.; Tandler-Schneider, A.; Rugescu, I.A.; Goossens, V. ART in Europe, 2018: Results generated from European registries by ESHRE. European IVF Monitoring Consortium (EIM), for the European Society of Human Reproduction and Embryology (ESHRE). *Hum. Reprod. Open* **2022**, *2022*, hoac022. [CrossRef] [PubMed]
4.   Raudonis, V.; Paulauskaite-Taraseviciene, A.; Sutiene, K.; Jonaitis, D. Towards the automation of early-stage human embryo development detection. *BioMed Eng. OnLine* **2019**, *18*, 120. [CrossRef] [PubMed]
5.   Giménez, C.; Conversa, L.; Murria, L.; Meseguer, M. Time-Lapse Imaging. Morphokinetic analysis of In Vitro Fertilization outcomes. *Fertil. Steril.* **2023**, *120*, 218–227. [CrossRef] [PubMed]
6.   Doroftei, B.; Ilie, O.-D.; Anton, N.; Armeanu, T.; Ilea, C. A Mini-Review Regarding the Clinical Outcomes of In Vitro Fertilization (IVF) Following Pre-Implantation Genetic Testing (PGT)-Next Generation Sequencing (NGS) Approach. *Diagnostics* **2022**, *12*, 1911. [CrossRef] [PubMed]
7.   Wang, W.; Vilella, F.; Alama, P.; Moreno, I.; Mignardi, M.; Isakova, A.; Pan, W.; Simon, C.; Quake, S.R. Single-cell transcriptomic atlas of the human endometrium during the menstrual cycle. *Nat. Med.* **2020**, *26*, 1644–1653. [CrossRef] [PubMed]
8.   Giudice, L.C. Multidimensional transcriptomic mapping of human endometrium at single-cell resolution. *Nat. Med.* **2020**, *26*, 1513–1514. [CrossRef] [PubMed]
9.   Suhorutshenko, M.; Kukushkina, V.; Velthut-Meikas, A.; Altmäe, S.; Peters, M.; Mägi, R.; Krjutškov, K.; Koel, M.; Codoñer, F.M.; Martinez-Blanch, J.F.; et al. Endometrial receptivity revisited: Endometrial transcriptome adjusted for tissue cellular heterogeneity. *Hum. Reprod.* **2018**, *33*, 2074–2086. [CrossRef] [PubMed]
10. Altmäe, S.; Koel, M.; Võsa, U.; Adler, P.; Suhorutšenko, M.; Laisk-Podar, T.; Kukushkina, V.; Saare, M.; Velthut-Meikas, A.; Krjutškov, K.; et al. Meta-signature of human endometrial receptivity: A meta-analysis and validation study of transcriptomic biomarkers. *Sci. Rep.* **2017**, *7*, 10077. [CrossRef]
11. Koel, M.; Krjutškov, K.; Saare, M.; Samuel, K.; Lubenets, D.; Katayama, S.; Einarsdottir, E.; Vargas, E.; Sola-Leyva, A.; Lalitkumar, P.G.; et al. Human endometrial cell-type-specific RNA sequencing provides new insights into the embryo-endometrium interplay. *Hum. Reprod. Open* **2022**, *2022*, hoac043. [CrossRef]
12. Simón, C.; Gómez, C.; Cabanillas, S.; Vladimirov, I.; Castillón, G.; Giles, J.; Boynukalin, K.; Findikli, N.; Bahçeci, M.; Ortega, I.; et al. A 5-year multicentre randomized controlled trial comparing personalized, frozen and fresh blastocyst transfer in IVF. *Reprod. Biomed. Online* **2020**, *41*, 402–415. [CrossRef]
13. Toson, B.; Simon, C.; Moreno, I. The Endometrial Microbiome and Its Impact on Human Conception. *Int. J. Mol. Sci.* **2022**, *23*, 485. [CrossRef] [PubMed]
14. Moreno, I.; Garcia-Grau, I.; Perez-Villaroya, D.; Gonzalez-Monfort, M.; Bahçeci, M.; Barrionuevo, M.J.; Taguchi, S.; Puente, E.; Dimattina, M.; Lim, M.W.; et al. Endometrial microbiota composition is associated with reproductive outcome in infertile patients. *Microbiome* **2022**, *10*, 1. [CrossRef] [PubMed]
15. Noyes, R.W.; Hertig, A.T.; Rock, J. Dating the endometrial biopsy. *Obstet. Gynecol. Surv.* **1950**, *5*, 561–564. [CrossRef]
16. Murdoch, T.A.; Veras, E.F.T.; Mazur, T.M.; Kurman, R.J. *Diagnosis of Endometrial Biopsies and Curettings. A Practical Approach*, 3rd ed.; Springer: Berlin/Heidelberg, Germany, 2019.
17. Alfer, J.; Fattahi, A.; Bleisinger, N.; Krieg, J.; Behrens, R.; Dittrich, R.; Beckmann, M.W.; Hartmann, A.; Classen-Linke, I.; Popovici, R.M. Endometrial Dating Method Detects Individual Maturation Sequences During the Secretory Phase. *Vivo* **2020**, *34*, 1951–1963. [CrossRef] [PubMed]
18. Chen, G.D.; Pang, T.T.; Li, P.S.; Zhou, Z.X.; Lin, D.X.; Fan, D.Z.; Guo, X.L.; Liu, Z.P. Early pregnancy vitamin D and the risk of adverse maternal and infant outcomes: A retrospective cohort study. *BMC Pregnancy Childbirth* **2020**, *20*, 465. [CrossRef]
19. Fouad, S.; Randell, D.; Galton, A.; Mehanna, H.; Landini, G. Unsupervised morphological segmentation of tissue compartments in histopathological images. *PLoS ONE* **2017**, *12*, e0188717. [CrossRef]
20. Haque, I.R.I.; Neubert, J. Deep learning approaches to biomedical image segmentation. *Inform. Med. Unlocked* **2020**, *18*, 100297. [CrossRef]
21. Gondim, P.H.C.C.; Limirio, P.H.J.O.; Rocha, F.S.; Batista, J.D.; Dechichi, P.; Travençolo, B.A.N.; Backes, A.R. Automatic Segmentation of Bone Canals in Histological Images. *J. Digit. Imaging* **2021**, *34*, 678–690. [CrossRef]
22. Graham, S.; Vu, Q.D.; Raza, S.E.A.; Azam, A.; Tsang, Y.W.; Kwak, J.T.; Rajpoot, N. Hover-Net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Med. Image Anal.* **2019**, *58*, 101563. [CrossRef]
23. Priego-Torres, B.M.; Sanchez-Morillo, D.; Fernandez-Granero, M.A.; Garcia-Rojo, M. Automatic segmentation of whole-slide H&E stained breast histopathology images using a deep convolutional neural network architecture. *Expert Syst. Appl.* **2020**, *151*, 113387. [CrossRef]

24. Ali, H.; Haq, I.U.; Cui, L.; Feng, J. MSAL-Net: Improve accurate segmentation of nuclei in histopathology images by multiscale attention learning network. *BMC Med. Inform. Decis. Mak.* **2022**, *22*, 90. [CrossRef] [PubMed]

25. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, Munich, Germany, 5–9 October 2015; Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F., Eds.; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241.

26. Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. *IEEE Trans. Med. Imaging* **2019**, *39*, 1856–1867. [CrossRef] [PubMed]