BioMedical Engineering OnLine

Towards the Automation of Early-stage Human Embryo Development Detection

--Manuscript Draft--

Manuscript Number:	BMEO-D-19-00274R1
Full Title:	Towards the Automation of Early-stage Human Embryo Development Detection
Article Type:	Research
Abstract:	Background: Infertility and subfertility affect a significant proportion of humanity. The assisted reproduction technology proved itself capable for threatening the infertility issue. The in vitro fertilization is one option whose success highly depends on the selection of high quality embryo for transfer. Usually it is done manually by analyzing embryo under a microscope. However, the evidence shows that the success rate of manual selection remains low. The usage of new incubators with integrated time-lapse imaging system is providing new possibilities for embryo assessment. Therefore, we address this problem by proposing the approach based on deep learning for automated embryo quality evaluation by analyzing time-lapse images. Automatic embryo detection is complicated by the topological changes of tracked object. Besides, the algorithm should process a large number of image files of different quality in a reasonable amount of time. Methods: We propose the automated approach to detect the human embryo development stage during its incubation and to highlight the embryo with abnormal behavior by focusing on five different stages. The method encompasses two major steps. First, the location of embryo in the image is detected by employing Haar feature-based cascade classifier and leveraging the radiating lines. Then, the multi-class prediction model is developed to identify a total cell number in the embryo using the technique of deep learning. Results: The experimental results demonstrate that the proposed method achieves the accuracy of at least 90% in the detection of embryo location. The implemented deep learning approach to identify the early stages of embryo development resulted in overall accuracy above 92% using the selected CNNs architectures. The most problematic was the 3-cell stage, presumably because of its short duration during development. Conclusion: This paper contributes to the field by proposing a model to automate the monitoring of early-stage human embryo development. Unlike in other imaging fields, there exists only few attempts that publish the leveraging of deep learning in this field. Thereupon, the approach presented in this study could be used in the creation of novel algorithms integrated in the assisted reproductive technology used by the embryologists.
Response to Reviewers:	Response to the Comments of Reviewers Manuscript #: BMEO-D-19-00274 Title of the Paper: Towards the Automation of Early-stage Human Embryo Development Detection We appreciate the time and effort on the part of the editor and referees in reviewing this manuscript and providing constructive comments that helped to improve and clarify the manuscript considerably. We hope that the revised version of the manuscript answers their concerns. Below we included the comments of the reviewers and responded to them individually, indicating exactly how we addressed each concern and explaining the changes that we have made. These changes are marked in blue, and the revised manuscript is attached to this response. Reviewer Comments, Author Responses, and Manuscript Changes Reviewer 1 Overall comments: This work suggested a machine learning method to detect the human embryo development stage using microscopic imaging. Results showed that

the authors' method achieved an accuracy of more than 90% to detect the embryo location. They also addressed some issues in the machine learning such as the short duration of the 3-cell stage using the CNNs architectures. Overall, this manuscript is quite comprehensive and thorough. Machine learning on imaging is currently a timely topic. I only have some minor concerns in this work.
• Response: We thank the reviewer for the positive overall comment. We have addressed all concerns as follows.

1. Comment: Can the authors estimate the improvement of accuracy and analyzing time when their machine learning method is used to replace the manual method? So the reader will know how good is the machine learning algorithm.
• Response: The manual detection of embryo development stages is called annotation. The embryologist must identify the moment when the embryo cleaved from 1 cell to 2 cells, from 2 to 3 cells, and so on until the blastocyst stage is reached. This should be performed on the whole image sequence of embryo development. Usually, the development of an embryo is recorded for up to 5 days at 5-minute intervals. Modern time-lapse incubators, such as ESCO Miri TL, have optical microscopes that can capture the embryo in seven different focal planes for additional information. Manual annotation involves the analysis of over 10000 images per embryo; more specifically, 5(days)*24(hours)*60(minutes)*7(focal planes)/5(minute intervals). Less than 2 minutes is required for a skilled embryologist to annotate one embryo in the case of the embryo not having a high percentage of fragmentation. Usually, IVF patients have up to 5 or 10 embryos. Therefore, the manual annotation of all embryos for one patient can take up to 20 minutes. Notably, the automated annotation system can do the same work 10 times faster without human intervention. However, an accuracy comparison of manual and automated annotations is not straight forward. For example, the human factor (lack of attention, fatigue, stress, etc.) should be considered alongside how it correlates with manual annotation accuracy; though this topic could represent a separate research project. In response to this comment, we have added few statements in "Methods" section. Please, refer to Page 14, lines 14-20; Page 13, lines 23-26.
2. Comment: Did the authors also try other algorithms rather than the CNNs architectures? Basically, different algorithms should be examined to identify the best result when carrying out machine learning.
• Response: The proposed algorithm consists of two major steps: embryo localisation and cell stage classification. In the present study, the location of an embryo in the image is detected by employing a Haar feature-based cascade classifier and leveraging the radiating lines. Alternatively, this task could be solved using object detection methods such as Local Binary Patterns (LBP) or Histogram of Oriented Gradients (HOG). Both methods were tested, but the cascade classifier was selected for further development. The HOG and LBP methods lack localisation accuracy because they require a high contrast image, where the target object is captured with sharp edges. Moreover, these methods fail in detecting partially overlapped, noisy or blurred objects, as well as they are too sensitive to object rotation and the location of a region of the target object [Chen et al (2017); Ozturk and Akdemir (2018); Huang et al (2011)]. An embryo image captured using a time-lapse system is slightly blurry and the boundaries of the embryo are too fuzzy; therefore, methods that are able to generalise the results should be employed. These comments have been added in the revised manuscript. Please, refer to Page 17, lines 24-28; Page 18, lines 1-4.
Two different CNNs architectures (AlexNet and VGG16) were used in the cell stage classification. The first attempt to automatically detect the stage included the use of a cascade classifier, PCA and SVM. A cascade classifier was used to detect the location of the embryo in the image. PCA was used for the reduction of data dimensionality and feature extraction. SVM was trained to classify different cell stages based on PCA features. The current combination resulted in approximately 85% classification accuracy, which was significantly improved by exploiting deep learning. In response to this comment, we have added a paragraph regarding the use of alternative approach in "Methods" section. Please, refer to Page 18, lines 10-18.
3. Comment: More details about the source of data regarding machine learning should be given such as how the authors filter the bad data in this study.
• Response: The training and test data set was prepared by skilled embryologists. All images in the data set were carefully examined and labelled. In our case, poor data such as low-resolution images, images where the whole embryo is not seen, blurry images, and others were excluded from the analysis. In response to this comment, we

have added several sentences in "Results" section. Please, refer to Page 7, lines 15-17.
4. Comment: To avoid the issue of embryo location and so on, did the authors think to set up this method as machine learning-aided? That is to use human to finish the job that machine learning is difficult to complete, while machine learning performs the remaining which human cannot do better?
• Response: The localisation problem can be solved by placing the embryo into a culturing dish with narrower well, which limits the travel of embryo in growing media. Such a proposal requires a major technical revision of the time-lapse incubator. The current culturing dish has a well that is twice as large (diameter of 300 micrometres) as the early-stage embryo (diameter of 150 micrometres). Therefore, the localisation problem remains an existing issue. The authors of this work believe that a vast amount of annotation can be automatically performed using the proposed algorithm. Thus, only cases when the automated algorithm responds with low confidence must be analysed manually.
5. Comment: Figure 2 captions: Please give captions in the subfigures (a-d).
• Response: Corrected. Please, refer to Page 8, lines 16-18.
6. Comment: Figure 3: Please explain the regions outlined by the red and green lines.
• Response: The well-localised embryo is highlighted with a green circle. The green circle should fit the boundaries of the embryo. The red point illustrates a centre of detected region. Please, refer to Page 8, lines 27-28.

Reviewer 2
1. Comment: In the "Background" I propose to write shortly about other medical disciplines (specialties), where image data, their time-variability are taking into account, too. In this chapter Authors can take into account such papers, like:
- Lyssek-Boroń A, Wylęgała A, Polanowska K, et al. Longitudinal changes in retinal nerve fiber layer thickness evaluated using avanti Rtvue-XR optical coherence tomography after 23G vitrectomy for epiretinal membrane in patients with open-angle glaucoma. J Healthc Eng. 2017;2017: 4673714. doi./10.1155/2017/4673714
- Gao X, Boccalini S, Kitslaar PH, Budde RPJ, Tu S, Lelieveldt BPF, Dijkstra J, Reiber JHC. A novel software tool for semi-automatic quantification of thoracic aorta dilatation on baseline and follow-up computed tomography angiography. Int J Cardiovasc Imaging. 2019 Apr;35(4):711-723. doi: 10.1007/s10554-018-1488-9.
• Response: In response to this comment, a few sentences have been added in the "Background" section as suggested by a reviewer. Please, refer to Page 5, lines 1-7.
2. Comment: The current study has really great potential for further studies.
• Response: We would like to thank the reviewer for this positive feedback.

Reviewer 3
Overall comments: This paper contributes to the field by proposing a model to automate the monitoring of early-stage human embryo development. An automated detection system is presented in the paper, which consists of two main components: (1) the localization of embryo in the image and (2) the identification of embryo development stage, with the aim to identify the abnormal division patterns. However, some important revisions are required to improve its quality. Suggested revisions are given below.
• Response: Thank you for the suggestions. The revisions we made are as follows.
1. Comment: The paper should be re-organized.
• Response: While preparing the manuscript, we followed the order of sections provided in the submission guidelines (please, refer to https://biomedical-engineering-online, biomedcentral.com/submission-guidelines/preparing-your-manuscript/research-article). We have also contacted the Journal's Editorial Office regarding the structure of the paper and we were informed that the editor had no comments regarding the formatting.
2. Comment: Abbreviations should be defined and used carefully. For instance, the term CNN has been defined at page 17, but it has been used previous pages, such as at page 13.
• Response: We thank the reviewer for this comment. Actually, the definitions, assumptions and description of research methods used in the study are presented in the section "Methods". Since it is required in the submission guidelines, the "Methods" section should be provided nearly at the end of the paper. We have corrected the

abbreviation used for CNNs, which has not been defined in the text at first use ("Background" section). Also, the section "List of Abbreviations" is given in the paper.
3. Comment: The authors should explain clearly why deep learning is required in this work. For instance, they can add the following review type paper to indicate superiority of deep neural networks for medical image processing; "Deep Learning in Medical Image Analysis: Recent Advances and Future Trends".11th Int. Conf. on Computer Graphics, Visualization, Computer Vision and Image Processing (CGVCVIP 2017), pg.305-311, 20-23 July 2017, Lisbon, Portugal
• Response: The motivation of using deep learning is added in the "Background" section. The reference suggested by a reviewer is also added. Please, refer to Page 5, lines 12-19.
4. Comment: Recently Sobolev gradient based optimizers are used. The authors should add the following statement to inform the readers about sufficiency of the their optimizer. "Although Sobolev gradient based optimizers have been used in some works [R1-R3], in this work the proposed method is efficient the traditional optimizer."
- R1:"Diagnosis of Alzheimer's Disease with Sobolev Gradient Based Optimization and 3D Convolutional Neural Network", International Journal for Numerical Methods in Biomedical Engineering, in press, DOI:10.1002/cnm.3225, 2019
- R2:"Fully Automated Liver Segmentation Using Sobolev Gradient-Based Level Set Evolution", International Journal for Numerical Methods in Biomedical Engineering, 32(11), 2765, DOI: 10.1002/cnm.2765, 2016
- R3:"A Level Set Method with Sobolev Gradient and Haralick Edge Detection", The 4th World Conference on Information Technology (WCIT 2013), 5, 131-140, November 26-27, 2013, Brussels, Belgium
• Response: We are grateful for this comment. We added the suggested statement in the Subsection "Proposed Algorithm for the Detection of Embryo Location". Please, refer to Page 17, lines 14-16.
5. Comment: Images are usually noisy due to various reasons.The authors should emphasis the advantage of the proposed method by adding the following statement: "Images are usually noisy and there are different noise types. Therefore, in traditional approaches, noise reduction is applied according to determined noise types or levels [R1,R2]. The proposed approach can give successful results without an extra step for noise reduction or intensity normalization such as presented in [R3,R4]."
- R1:"Natural image noise level estimation based on local statistics for blind noise reduction", The Visual Computer, Vol.34, Issue 4, pp 575-587, 2018
- R2:"Quantitative Validation of Anti-PTBP1 Antibody for Diagnostic Neuropathology Use: Image Analysis Approach". Numerical Methods in Biomedical Engineering, 33(11), e2862, 2017
- R3:"Fully Automated and Adaptive Intensity Normalization Using Statistical Features for Brain MR Images", Celal Bayar University Journal of Science, 14(1), 125-134, DOI:10.18466/cbayarfbe.384729, 2018
- R4:"Intensity Normalization in Brain MR Images Using Spatially Varying Distribution Matching". 11th Int. Conf. on Computer Graphics, Visualization, Computer Vision and Image Processing (CGVCVIP 2017), pg.300-304, 20-23 Juy 2017, Lisbon, Portugal
• Response: In response to this comment, we extended the statement about the method's suitability for detecting objects in a noisy background (see the Subsection "Proposed Algorithm for the Detection of Embryo Location", Page 17, lines 17-20.)
6. Comment: Grammatical mistakes should be corrected.
• Response: We would like to thank the reviewer for their comments. The authors have sent the paper to a professional for editing and proofreading. The quality of the English should now be improved.

Reviewer 4
Overall comments: The manuscript proposed a method to detect the location of embryo using cascade classifier and identify the early-stage human embryo development using convolutional neural networks. The proposed method has high prediction accuracy in the multi-class prediction, and can potentially lower the burden of visual inspection. The method has some novelty. However, there are some problems to be solved.
• Response: Thank you for the constructive suggestions. Please refer to the responses given below.
1. Comment: The lining number is confusing, please don't duplicate it.
• Response: We do agree that the line numbering complicates the reading. When

preparing the manuscript, we used the default options specified in the latex template and followed the recommendations provided in the submission guidelines (please, refer to https://biomedical-engineering-online.biomedcentral.com/, submission-guidelines/preparing-your-manuscript).
2. Comment: Page 2, line 20. Since the number of samples in 3-cell stage is small, data augmentation might be helpful for the imbalanced sample. I think it could be related with the duration of 3-cell stage. If the duration is short, the number of samples in certain stage will be small and imbalanced samples won't be fixed. Please discuss the feasibility of data augmentation.
• Response: The authors agree with the reviewer. The duration of the 3-cell stage is approximately 8–10 times shorter than, for example, the 2-cell stage; therefore, the number of samples of the 3-cell stage in the training image data set is smaller. The authors believe that simple data augmentation cannot solve the problem of unbalanced data set because the appearance of an embryo in the 3-cell stage is not merely limited in scale, translation or rotation. The variation is much greater and can vary in shape, form, the order of cells, and fragmentation, among other factors. The imbalanced data set issue is solved by limiting (randomly selecting) the number of samples of other cell stages (1, 2, 4, or more cells) to the number of 3-cell samples. All images representing the 3-cell stage were used in the experiment. In response to this comment, a few sentences have been added in the revised manuscript. Please, refer to Page 7, lines 18-21.
3. Comment: Page 6, line 19. What is the frame rate of the camera?
• Response: The images of early-stage embryos were taken using an ESCO Miri TL incubator. It has a camera that is capable of capturing 47 frames per second. However, the recording of development is done at 5-minute intervals because embryo development is a relatively slow process. This information is added in the revised manuscript. Please, refer to Page 7, lines 10-12.
4. Comment: Page 6, line 21. How did you select the training and testing samples? Randomly, or manually selected?
• Response: The data were split into training and testing sets at a ratio 70:30 by randomly selecting samples (images) of each cell stage specified in the paper. This information is added in the revised manuscript. Please, refer to Page 7, lines 22-23.
5. Comment: Page 6, line 23. Please specify the size of the pixel. So we can get the length of radiating line and cell size.
• Response: The time-lapse incubator has a camera with a 2.35 MP image sensor. The sensor has a 1936x1216 pixel resolution (one pixel = 5.86 micrometres). This information is added in the revised manuscript. Please, refer to Page 7, lines 8-10.
6. Comment: Page 7. I don't understand how to find the center of radiating lines.
• Response: The centre of radiating lines is detected using the algorithm presented in the "Proposed Algorithm for the Detection of Embryo Location" section (see, Page 15-17). The location of an embryo in the image is detected by employing a Haar feature-based cascade classifier. This classifier returns the rough location f the embryo in the form of a rectangular area. The geometrical centre of this area is used as the centre for the radiating lines.
7. Comment: Page 7, line 2. 'The proposed embryo location detection algorithm was regarded as successful ...'. The sentence is confusing.
• Response: We thank the reviewer for this comment. The sentence is corrected. Please, refer to Page 8, lines 3-4.
8. Comment: Page 8, Figure 5. (1) It would be better to have the range of y-axis to be 70 100%. It is not necessary to cover the range (0 100%). (2) Please explain the figure, is this only one sample or many samples? (3) line length 200px, it should be abbreviation of pixel, please note it at the first time 'px' was used. (4) I prefer that unit of x-axis in figure 5(b) would be um, not pixel.
• Response:
– (1) The range of y-axis was modified in response to this comment. Please, refer Page 10, Figure 5;
– (2) Figure 5 represents the average detection rate computed from many image samples;
– (3) The abbreviation of pixel "px" is inserted in the paper. Please, refer Page 7, line 10;
– (4) The size of one pixel in an image sensor is 5.86 um (this information is added in the revised manuscript (see Page 7, line 10.)) However, the results of experimental investigation are presented in pixels, because real distance measurements based on sensor's pixel size are not fully correct. The embryo in the image can appear much

smaller than the actual size due to several factors such as the scale of magnification (the type of lenses) and distance between embryo and camera, which can vary during cultivation.
9. Comment: Page 9, line 3. "Closer inspection", I can see the details. Please consider to modify the figure.
• Response: The sentence is corrected. The figure is modified. Please, refer to Page 9, lines 26.
10. Comment: Page 10, line 12. "As we more focus on the identification of actual positives, ..." I can not understand this sentence.
• Response: We are sorry for the confusing sentence. We removed it, since the sentence repeats the information that is already given in the paragraph.
11. Comment: Page 10, line 18. Please remove 'Closer inspection'.
• Response: Corrected. Please, refer to Page 11, lines 21.
12. Comment: Page 11, line 3. "KIDScore grading method". Please add the reference.
• Response: Added. Please, refer to Page 12, lines 6.
13. Comment: Page 11. Please compare with previous results, and discuss the advantage of the proposed method.
• Response: The first attempt to solve the given problem involved leveraging a cascade classifier, PCA and SVM. The cascade classifier was used to detect the location of embryo in the image. PCA was used for the reduction of data dimensionality and feature extraction. SVM was trained to classify different cell stages based on PCA features. The combination of a cascade classifier, PCA and SVM resulted in approximately 85% classification accuracy, which has now been significantly improved by using deep learning. This information has been added in the revised manuscript. Please, refer to Page 18, lines 10-18.
14. Comment: It would be interesting to report the accuracy of location detection, and the whole system performance. The whole system, including the location detection and multi-class identification, should work as a flow. If the location detection is not accurate, the multi-class identification won't make any sense.
• Response: Definitely. If the detection of embryo location is not accurate, then the proposed algorithm also fails. This was also explained in the paper (see Page 6, line 28, Page 7, lines 1-5; Page 14, lines 22-28). and was the reason why we devoted a significant period of time to develop it in the research. The process of measuring embryo localisation accuracy is not trivial. There is no objective way to compare the detected locations because even several skilled embryologists can highlight a slightly different location. The number of cells in the image is more of an objective criterion, on which all embryologists could agree. Therefore, the authors of this work decided to provide the overall classification accuracy of the proposed approach.
15. Comment: Page 12, line 10. There are 114,793 frames for 300 sequences. There seems to be 38 frames for each sequence if my understanding is correct. "It captures the images of embryo (In our case, each five minutes)". It will be only 38*5 mins = 190 mins = 3.2 hrs. I don't see the human embryo can develop so fast. It could be around 5 days. Right? Please check your parameters and give me a reasonable explanation.
• Response: Correct. However, not all available images were used in the experimental investigation. Only 114,739 images were selected from 300 different embryos that were cultivated in different TL incubators. All images represent different embryo stages such as 1-, 2-, 3-, 4- and >4-cell stages. Images with poor focus and empty images were not used. All 300 development sequences would have more than 500000 images. Notably, not all embryos were cultivated in the incubator for up to 5 days since embryologists can transfer embryos on the fourth day of development.
16. Comment: Page 15, line 2. "The more accurate location". I suggest to use 'a more accurate location'.
• Response: Corrected. Please, refer to Page 12, line 27; Page 16, line 13, Page 16, line 19.
17. Comment: Page 15, line 4. 'take as an input a gray-scaled image'. It could be 'take a gray-scaled image as an input '.
• Response: Corrected. Please, refer to Page 16, lines 15.
18. Comment: Page 15, line 14. I don't understand the histogram. Histogram of what values?
• Response: The histogram is computed from radiating lines (see example in Figure 4). Each line begins at the same starting point and extends in different directions for a certain length. The length of each line is measured in pixels and it represents the horizontal axis of the histogram. The vertical axis represents the accumulated greyness intensity of each image pixel over which each line travels at a certain

distance from the starting point.
19. Comment: Page 15. Is it possible to use deep learning method instead of cascade classifier?
• Response: Yes, it is possible to use deep learning-based localisation algorithms, such as YOLO (You-Only-Look-Once) [Redmon et al (2016); Cavaioni (2018)], Regional-proposal Convolutional Neural Network (RCNN) [Girshick et al (2014)] or Single Shot Multibox Detector (SSD) [Liu et al (2016)]. These object detection algorithms require larger computing power than the cascade-based object localisation algorithm. The cascade classifier proved that it is very useful in devices with limited computing resources (such as photo cameras for face detection). Therefore, the authors of the work investigated the possibility of applying a cascade classifier to the cell localisation problem since a similar resource issue is faced with incubators.
20. Comment: Page 18, line 8. What is the hardware? Please briefly introduce the time cost.
• Response: Experimental investigations were executed on a Windows 10 machine with 16.0 GB of RAM installed with an Intel Core i7-7700K 4.20GHz CPU. Less than 45 milliseconds were required to process one image and around 1 minute (depending on the number of incubating days) was required to analyse entire embryo development from the beginning to end. This information is added in the revised manuscript. Please, refer to Page 18, lines 24-28.
21. Comment: Page 18, line 16. 'more rich' could be 'richer'.
• Response: Corrected. Please, refer to Page 20, lines 23.

References

•Cavaioni M (2018) Deeplearning series: Object detection and localization — yolo algorithm, r-cnn. URL https://medium.com/machine-learning-bites/ deeplearning-series-objection-detection-and-localization-yolo-algorithm-r-cnn-71d4dfd07d5f
•Chen J, Zheng H, He N, Luo Z, Zhu R (2017) Object detection by learning oriented gradients. In: Zhao Y, Kong X, Taubman D (eds) Image and Graphics, Springer International Publishing, Cham, pp 86–96
•Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. IEEE Conference on Computer Vision and Pattern Recognition pp 1–8, DOI 10.1109/CVPR.2014.81
•Huang d, Shan C, Ardabilian M, Chen L (2011) Local binary patterns and its application to facial image analysis: A survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C 41:765–781, DOI 10.1109/TSMCC.2011.2118750
•Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, Berg AC (2016) Ssd: Single shot multibox detector. ECCV 2016 19(1):1 – 17, URL hhttps://www.cs.unc.edu/~wliu/papers/ssd.pdf
•Ozturk S, Akdemir B (2018) Comparison of hog, mser, sift, fast, lbp and canny features for cell detection in histopathological images. HELIX 8:3321–3325, DOI 10.29042/2018-3321-3325
•Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: Unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition pp 779–788, DOI 10.1109/CVPR. 2016.91, URL https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf

RESEARCH

Towards the Automation of Early-stage Human Embryo Development Detection

Vidas Raudonis[1†], Agne Paulauskaite-Taraseviciene[2†], Kristina Sutiene[3*†] and Domas Jonaitis[1†]

rrespondence:
tina.sutiene@ktu.lt
epartment of Mathematical
delling, Kaunas University of
hnology, 51368 Kaunas,
uania
list of author information is
ilable at the end of the article
qual contributor

Abstract

Background: Infertility and subfertility affect a significant proportion of humanity. Assisted reproductive technology has been proven capable of alleviating infertility issues. *In vitro* fertilisation is one such option whose success is highly dependent on the selection of a high-quality embryo for transfer. This is typically done manually by analysing embryos under a microscope. However, evidence has shown that the success rate of manual selection remains low. The use of new incubators with integrated time-lapse imaging system is providing new possibilities for embryo assessment. As such, we address this problem by proposing an approach based on deep learning for automated embryo quality evaluation through the analysis of time-lapse images. Automatic embryo detection is complicated by the topological changes of a tracked object. Moreover, the algorithm should process a large number of image files of different quality in a reasonable amount of time.

Methods: We propose an automated approach to detect human embryo development stages during incubation and to highlight embryos with abnormal behaviour by focusing on five different stages. This method encompasses two major steps. First, the location of an embryo in the image is detected by employing a Haar feature-based cascade classifier and leveraging the radiating lines. Then, a multi-class prediction model is developed to identify a total cell number in the embryo using the technique of deep learning.

Results: The experimental results demonstrate that the proposed method achieves an accuracy of at least 90% in the detection of embryo location. The implemented deep learning approach to identify the early stages of embryo development resulted in an overall accuracy of over 92% using the selected architectures of convolutional neural networks. The most problematic stage was the 3-cell stage, presumably due to its short duration during development.

Conclusion: This research contributes to the field by proposing a model to automate the monitoring of early-stage human embryo development. Unlike in other imaging fields, only a few published attempts have involved leveraging deep learning in this field. Therefore, the approach presented in this study could be used in the creation of novel algorithms integrated into the assisted reproductive technology used by embryologists.

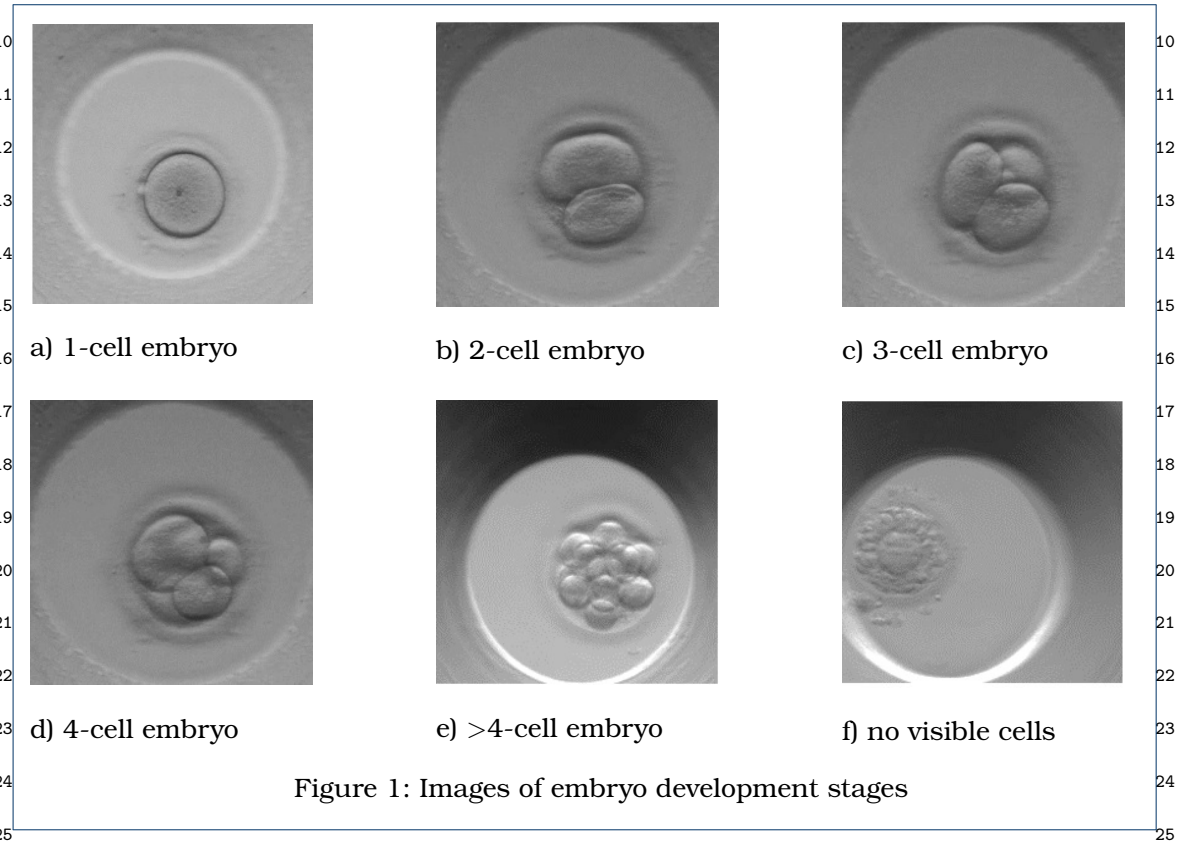
Keywords: Deep learning; Location detection; Embryo development; Image recognition; Multi-class prediction

Background

Infertility is a growing problem worldwide. According to the World Health Organization, one in every six couples has issues leading to infertility problems. It has been noted that the global *in vitro* fertilization (IVF) market is expected to grow at an approximated 10% compound annual growth rate between 2018 and 2026 [1]. Geographically, Europe dominates the market by capturing the largest share, which is driven by low fertility rates, government financial support for the adoption of IVF and other fertility treatments, and the increasing success rate of IVF methods. According to the forecasts [2], the Asia-Pacific region is anticipated to demonstrate rapid growth in the foreseeable future. Causes of infertility are numerous, potentially including factors such as anatomical or genetic problems, physiological dysfunction, sexually transmitted diseases, endocrinological or immunological problems, and many more. Moreover, the rising trend towards delaying pregnancy due to career concerns, financial reasons or not finding the right partner has also increased the need for IVF services. The success of IVF procedures is closely linked to many biological and technical issues. The fertilisation and *in vitro* culturing of embryos are dependent upon an environment that should be stable and correct with respect to temperature, air quality, light, media pH and osmolality. After fertilisation, an embryo that develops normally will continue to divide, growing to the blastocyst stage by the fifth or sixth day; however, only one-third of all embryos are capable of reaching this stage [3]. The success of IVF resulting in a pregnancy is only around 52.3%. For this reason, more than one embryo is transferred, which subsequently increases the risk of multiple pregnancies. In fact, more than 30% of IVF-induced pregnancies are multiple-infant births. For this reason, embryo viability is monitored by an embryologist during the IVF procedure. Nevertheless, embryo assessment is subjective and based on limited observations if it is performed visually by placing the fertilised embryo under a microscope once to a few times per day.

Time-lapse (TL) systems developed over recent years (with or without computer algorithms) provide a massive number of digital images of embryos at frequent time inter-

vals, thus enabling embryologists to assess the quality of the embryos without physically removing them from their culture environment [4]. Embryos can be transferred to the uterus at the cleavage stage (Day 2 or 3, Fig. 1, b-e) or blastocyst stage (Day 5, Fig. 1, f). Transferring embryos at the blastocyst stage may increase the likelihood of selectively transferring viable and genetically normal embryos [5]. The correct identification of cell number creates presumptions for determining the timing parameters from time-lapse imaging, such as the duration between different stages, which was approved as being significant in the evaluation of embryo quality [6].



a) 1-cell embryo b) 2-cell embryo c) 3-cell embryo

d) 4-cell embryo e) >4-cell embryo f) no visible cells

Figure 1: Images of embryo development stages

Despite all of the recent advances in computer vision research, the automatic detection and tracking of cells remain challenging. This task is complicated by the topological changes of tracked objects (cell division) in addition to the possible presence of randomly

appearing noise in the images. In comparison, many other medical imaging applications exist, where the variability of relevant data, such as target object, surrounding structures or image acquisition parameters, have a large impact on the decisions made by domain experts. For example, a previous experiment [7] emphasised the need to study longitudinal retinal nerve fiber layer (RNFL) thickness changes in patients with open-angle glaucoma, while the need to develop a single software package to automatically determine differences in aortic diameter from multiple scans of the same patient was presented recently [8]. Moreover, the algorithm to be developed should process a large number of image data files of different quality in a reasonable amount of time. Unlike in other fields of image recognition, far too little attention has been paid to the use of artificial intelligence in the detection of human embryo quality development.

Deep learning is now a state-of-the-art artificial intelligence model across a variety of domains and is seen as a key technique for future human-support technologies. As indicated by previous studies [9, 10], deep learning methods – more specifically convolutional neural networks (CNNs) – hold huge potential for medical imaging technology, medical diagnostics and healthcare in general. Unlike conventional machine-learning techniques, deep neural networks simplify the feature engineering process, provide abstract learning through a hierarchical representation of the data, efficiently deal with vast amounts of data and demonstrate their superiority in detecting abnormalities in medical images. Recently, an approach named STORK was developed that can be used for unbiased and automated embryo assessment using TL images [11]. They formulated a binary classification problem focusing on good- and poor-quality embryo assessment, which was tackled using deep neural networks, more specifically Inception-V1 architecture. In their research, authors used a large collection of human embryo time-lapse images (approximately 50,000 images) from a high-volume fertility centre in the US. The authors highlighted that STORK was able to predict blastocyst quality with an area under curve (AUC) of > 0.98, which is a very promising result. In the same manner, Iwata et al. [12] examined the use of deep learning on images of human embryos for predicting good- and poor-quality embryos.

They also referred to other studies [13–15] that utilized artificial intelligence approaches for quality prediction or grade classification with varying degrees of success. Comparatively, in another study [16], the authors used a list of the main morphological features of a blastocyst with the aim of automating embryo grading using Support Vector Machine (SVM) classifiers. They reported accuracies ranging from 0.67 to 0.92 for embryo development classification. Overall, these studies represent attempts to develop reliable algorithms for the prediction of a two-class problem.

Notably, the application of artificial intelligence focusing on multi-class prediction remains scarce. The recent study proposed a standalone framework based on Inception-V3 CNNs as the core to classify individual TL images up to the 4-cell stage for mouse and human embryos, respectively [17]. In their work, 31,120 images of 100 mouse embryos and 661,060 images of 11,898 human embryos cultured in the TL monitoring system were analysed. The experimental study on the test set demonstrated an average classification accuracy of 90% when the model was applied to predict individual images up to the 4-cell stage, while accuracy of 82% was achieved when it was applied to identify embryos up to the 8-cell stage. In this context, a three-level four-class embryo stage classification method based on the Adaboost ensemble was proposed with the aim to identify the number of cells at every time point of a TL microscopy video, which resulted in an average accuracy of 87.92% for human embryos, but exhibited only 20.86% accuracy for 3-cell detection [18]. To the best of our knowledge, these are the few known works that have addressed the identification of early-stage embryo development by formulating a multi-class prediction problem.

In line with these findings, the present study contributes to this field by proposing a model to automate the monitoring of early-stage human embryo development by focusing on the prediction of the cell number during the division process for up to 5 days. This involves segmenting embryos from the image and then predicting defined number classes that relate to the embryo development stages (i.e. 1-cell, 2-cell, 3-cell, 4-cell and >4-cell; see Fig. 1) using CNNs. Whereas one of the key elements of the system is the detection

of embryo location in an image, the algorithm is proposed for this purpose. It first determines the rough embryo location using a Haar feature-based cascade classifier and then specifies its accurate location by means of the radiating lines. The use of this algorithm allowed us to achieve an accuracy of over 92% in predicting the early stages of embryo development.

Results

Images of early-stage embryo development were captured using a TL incubator system with an integrated camera, which has a 2.35-megapixel image sensor that provides a 1936x1216 pixels (px) resolution output (1 px = 5.86 μm). This camera is capable of capturing 47 frames per second. However, recording of the development process is performed at 5-minute intervals since embryo development is a relatively slow process. The experiment included 300 TL embryo development sequences for a total of 114,793 frames (18.73%, 25.45%, 9.35%, 20.65% and 25.82% of the data set for 1 to >4-cell stages, respectively). The image data set was carefully examined and labelled by a skilled embryologist. Poor data such as low-resolution images, images without an embryo or images with an occluded embryo with a material that does not belong to the embryo were excluded. The duration of the 3-cell stage is approximately 8–10 times shorter than, for example, the 2-cell stage; as such, the number of samples of the 3-cell stage in the image data set is smaller. Therefore, the number of samples at other cell stages (1-cell, 2-cell, 4-cell, or higher) was limited to the number of 3-cell samples. Finally, the training and testing image data set (in a ratio of 70:30) was formed by randomly selecting samples (images) for each of the previously specified cell stages.

First, the automatic detection of embryo location in the image was performed using the cascade classifier. It was noted that mostly linear diagonal Haar-like features were leveraged by the algorithm (see Fig. 7a). Unfortunately, the location of the entire embryo was not always successfully detected, as illustrated in Fig. 2. For instance, (a) a wrong area of the entire embryo is determined; (b) the individual cells are detected but not the

entire embryo; (c) the empty areas are determined; or (d) the objects of no interest are also detected. Therefore, the algorithm developed by the authors was used for embryo location detection . The proposed embryo location detection algorithm was considered successful for a problem if the entire embryo and its fused membrane were correctly identified in the image. The thickness of the membrane, its brightness and the number of granules are among the top criteria for assessing the quality of an embryo. That is why their detection is a crucial step in the present research. In Fig. 3, a well-localised embryo is highlighted by a green circle.

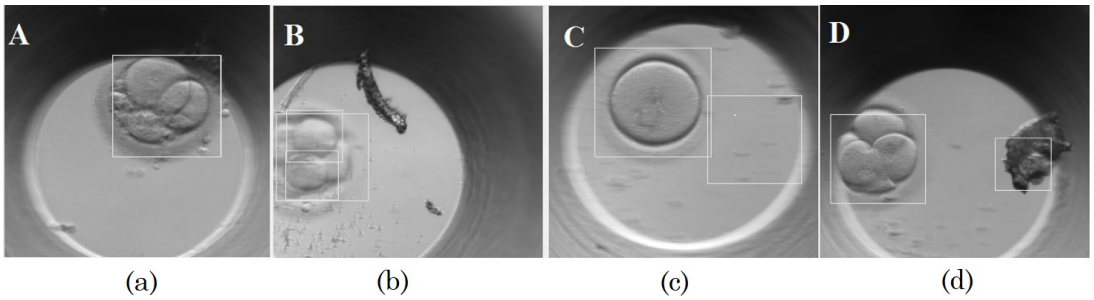


Figure 2: Illustration of unsuccessful identification of embryo location early-stage development: a) wrong area detected; b) individual cells determined; c) empty areas determined; d) not relevant objects detected

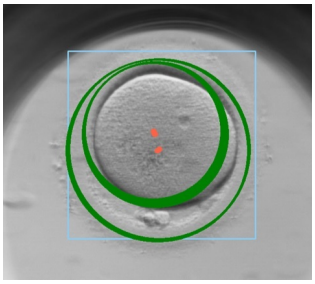
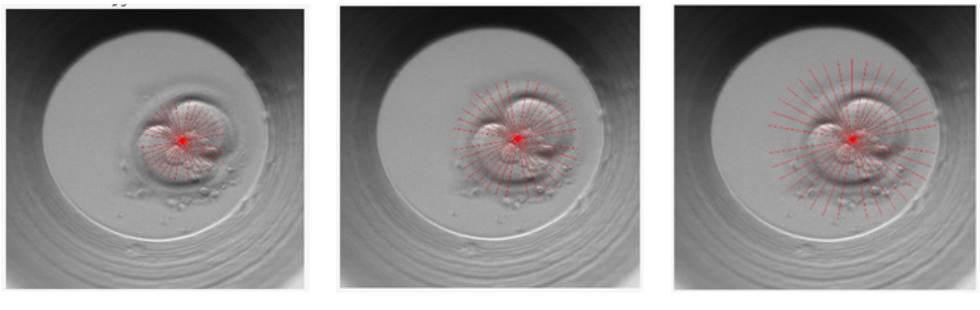


Figure 3: Case of well detected embryo location: the green circle should fit the boundaries of the embryo; the red point illustrates the centre of detected region

The algorithm proposed here includes the drawing of radiating lines, where the length of the line and the angle between lines are the main parameters to be considered. The change of line length affects the area of the image to be covered, while the change of angle between lines determines a different density to be explored in the image. Fig. 4 demonstrates the scattering of lines in the image for different lengths of radiating lines, given in pixels.



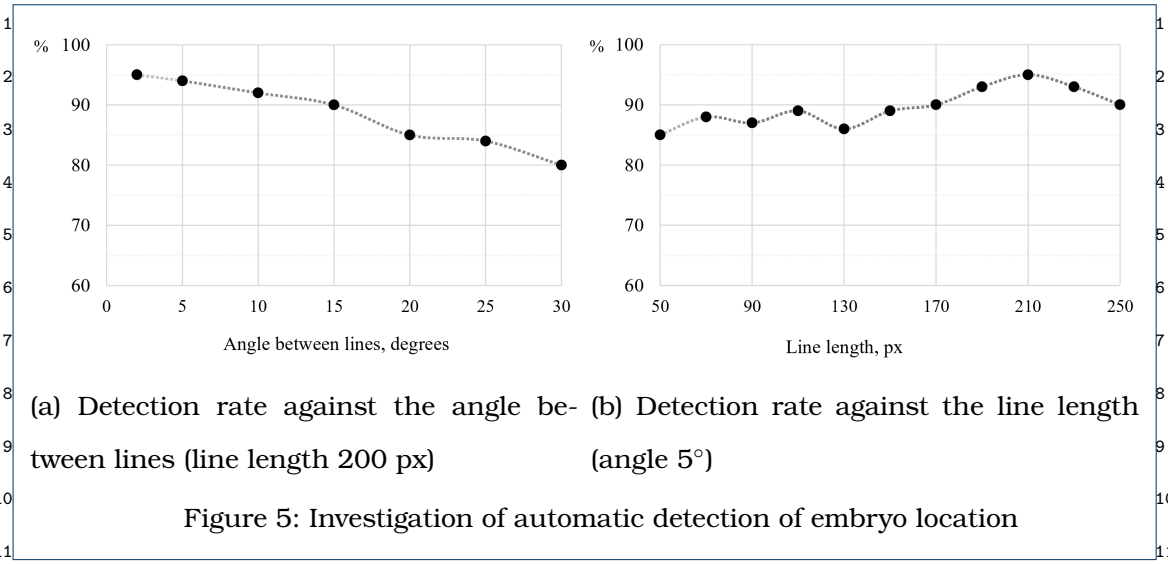
(a) (b) (c)

Figure 4: Radiating lines with lengths of 100, 150, 200 px

The ability of the proposed algorithm to correctly detect an entire embryo location is demonstrated in Fig. 5, where different radiating line lengths and the angle between them are investigated.

As illustrated in Fig. 5, the correct location detection rate for the entire embryo is rather high. However, the algorithm is more sensitive to changes in angle size between lines (see Fig. 5(a)). The increase of angle negatively impacts the detection quality. On the other hand, the number of points to be processed increases rapidly if the angle is decreased. Fig. 5(b) shows that the detection rate is above 90% if the line length is over 170 px when the angle is 5°. Typically, an embryo covers an area from 250x250 px up to 300x300 px.

Next, the classification of embryo development stages is explored. In the present research, five classes were specified in order to represent each early stage of embryo devel-



(a) Detection rate against the angle between lines (line length 200 px)

(b) Detection rate against the line length (angle 5°)

Figure 5: Investigation of automatic detection of embryo location

opment (i.e. 1-cell, 2-cell, 3-cell, 4-cell, >4-cell). The obtained confusion matrix for two CNNs architectures, such as AlexNet and VGG16, is presented in Table 1.

Table 1: Confusion matrices: each column shows the reference, while numbers running diagonally show the percentage of correct classification for every class considered in the experimental study

VGG architecture

		Reference				
		1	2	3	4	>4
Prediction	1	0.930	0	0	0	0
	2	0.023	0.943	0.036	0.019	0.027
	3	0.022	0.029	0.932	0.022	0.027
	4	0.025	0.027	0.032	0.957	0.028
	>4	0	0	0	0.001	0.917

Alex architecture

		Reference				
		1	2	3	4	>4
Prediction	1	0.910	0	0	0	0
	2	0.030	0.944	0.039	0.030	0.026
	3	0.029	0.029	0.920	0.032	0.026
	4	0.031	0.027	0.041	0.937	0.031
	>4	0	0	0	0.001	0.917

It can be seen that the classification performance is generally quite high. The comparison of two classifiers was performed by computing the confusion matrix-based performance measures [19, 20].

Table 2: Overall performance measures

	VGG architecture	Alex architecture
Macro Accuracy	0.936	0.927
Micro Accuracy	0.935	0.925
Macro F1	0.926	0.919
Micro F1	0.963	0.952

Table 2 highlights that the overall performance in terms of selected measures using the AlexNet architecture is slightly worse when compared to results from using the VGG architecture. It is evident that no difference exists between micro-accuracy and macro-accuracy. Compared to a macro F1 score, micro F1 obtains larger values for both CNNs architectures used in the experiment. Since F1 score is a balance between precision and recall, Table 3 was created to reveal the classifier performance by class to address these measures.

Table 3: Class-specific performance measures

VGG architecture

	Precision	Recall
1	1.000	0.930
2	0.927	0.943
3	0.790	0.932
4	0.901	0.957
5	0.999	0.917

Alex architecture

	Precision	Recall
1	1.000	0.910
2	0.915	0.944
3	0.767	0.920
4	0.888	0.937
5	0.999	0.917

Table 3 shows that precision is rather low for the third class, which defines the embryo stage as having three cells. Since micro-averaging favours classes with a larger number of instances, the final estimate was influenced by good performance for the classification of the other classes.

Discussion

The evaluation of early-stage embryo quality has been a matter of debate for many years. Using novel computer vision algorithms, various techniques have been developed to max-

imise the effectiveness of assisted reproductive technology. The use of TL imaging might increase the IVF success rate since this new approach allows the detection of abnormal behaviour in developing embryos.

TL imaging enhanced the selection criteria of the transferable embryo since the development of the embryos is observed to be more accurate. The quality of an embryo can be described by the KIDScore grading method [21]. It demonstrates that the embryo transition or cleavage from one stage to another has a certain optimal time. If an embryo cleaves from one cell to more cells too quickly or too slowly, then the embryo has a low probability for transfer. The authors of this paper aim to evaluate the embryo development with the use of deep learning techniques in order to automate the assessment of embryo quality at early development stages. The proposed method consists of two major steps: the embryo localisation into 2D image space and embryo stage classification.

Notably, the proposed approach has certain limitations. A deep learning-based method is only as smart and accurate as the data provided in training. For this research, the model was trained using TL images from a private IVF clinic. Therefore, the authors believe that the training database used to construct a decision-making core could be expanded by capturing all possible variations of different embryos. This could be the next step towards being able to build a fully automatic monitoring system for evaluating embryo quality.

Conclusion

The present study has reported the problems and suggested methods to automate early-stage human embryo detection. The proposed algorithm consists of two components, namely embryo localisation in the image and classification of embryo development stage. The detection of embryo location has been succeeded by using the improved object detection algorithm. First, the rough centre of the embryo is identified using Haar-like features. Then, a more accurate location of the embryo is computed by leveraging the radiating lines. The experimental investigation showed that detection accuracy of at least 90% was

reached using radiating lines of length 200 px placed at every 5 degrees. It was also determined that 200 px is the optimal line length (radius detected from the rough centre of an embryo), which is sufficient to wrap the entire embryo in the image. Embryo stage classification performance had an overall accuracy above 92%, which was achieved for both CNN architectures considered in the paper. The most problematic was the third class, which defines the 3-cell stage. This might have been caused by this stage usually being short compared to the other classes defined in the paper.

Methods

The TL system is part of the IVF incubator, which is used to register embryo development during its cultivation (see Fig. 6). It captures images of an embryo at certain time intervals (in our case, every 5 minutes) and stores the images. Typically, such a system consists of three main components: 1) a light source, 2) microscope optics and 3) a video camera. Usually, red light at 650 nm is used to illuminate an embryo, which is cultivated in a specially designed culturing dish, called a culture coin. Microscope optics magnify the embryo cells by 20 times. The TL system is equipped with a 2-megapixels video camera that allows the capture of an embryo in a 300x300 px area. The TL system uses a special mirror (prism) that concentrates light and directs it to the embryo and camera sensor.

Embryo assessment is based on the time intervals between cell cleavages, which are visually registered. The embryo is considered of high quality when the cleavage time intervals fit the normative data. Intervals that are too short or too long between cleavages signal the abnormal development of an embryo, which might lead to pregnancy failure. The TL system facilitates the recording of embryo development for up to 5 days at 5-minute intervals to create the sequence of images. Modern time-lapse incubators such as ESCO Miri TL have optical microscopes with which is possible to capture a human embryo at seven different focal planes for more information. Now, embryologists must evaluate each individual image in the sequence and decide which embryo is suitable for transfer. It is a complicated task not only because the embryo can behave unexpectedly

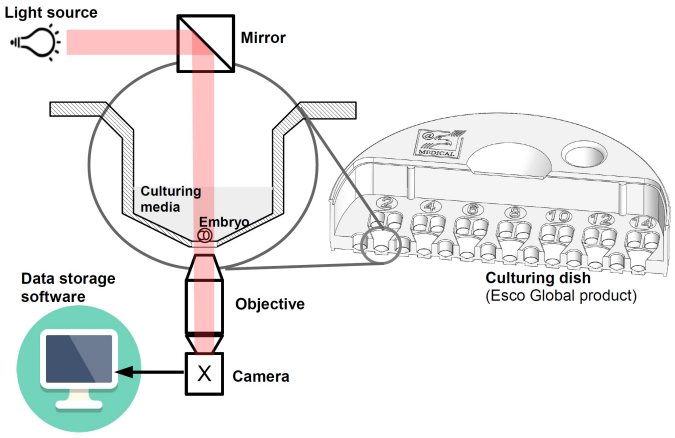


Figure 6: Scheme of time-lapse system

during its development, but also because of the massive image data set that includes over 10000 images per embryo, which must be manually assessed. A skilled embryologist requires less than 2 minutes to annotate one embryo in the case where embryos do not have a high percentage of fragmentation. Usually, IVF patients have up to 5 or 10 embryos. Henceforth, the manual annotation of all embryos for one patient can take up to 20 minutes. The automated annotation system can do the same work 10 times faster and without human intervention.

Therefore, an automated detection system of embryo development is presented in the paper that consists of two main components: (1) the localisation of an embryo in an image and (2) the identification of embryo development stages with the aim to identify abnormal division patterns. Since the detection of an embryo localisation in an image is a crucial step, the algorithm is proposed that uses a Haar feature-based cascade classifier to determine the rough embryo location and specify the accurate location with the help of the radiating lines.

Automatic Detection of Embryo Location

Cascade Classifier

One of the main steps in this research is to automatically determine embryo location. IVF embryos usually have a round shape with brighter edges. A cascade classifier was trained on a sample containing images with the target object labelled as positives, with negative images containing none of these objects. After the classifier is trained, it can be applied to identify targets in the image. In order to investigate the entire frame, the search window is moved across the image. The search window of a classifier can be easily changed when the size of the target object is unknown. In this case, the search should be performed several times using all possible search window sizes, which are placed on all possible locations in the image [22–24].

Cascading is a particular case of ensemble model that is built from several classifiers that are sequentially connected. Learning is a multi-stage process where an extension of the original data by the insertion of new attributes is performed in each step. This process accelerates image processing multiple times, as there is no need to check all of the features that are already learned. Haar-like features (see Fig. 7) are usually used as inputs to the basic classifiers.

As seen in Fig. 7, Haar-like features are extracted from adjacent rectangular regions at a specific location in a search window. Then, the difference between the sums of the pixel intensities in each region is computed. Usually, a large number of Haar-like features must be retrieved to describe the target object with sufficient accuracy. Therefore, these features are fed into a cascade classifier to construct a strong learner.

Proposed Algorithm for the Detection of Embryo Location

By default, a cascade classifier allows us quickly to determine the approximate location of an embryo; however, this is not sufficient for solving our problem. Therefore, the embryo location detection algorithm is developed (see Algorithm 1). Embryo detection consists of two main processing steps. The first step involves the application of a cascade classifier for

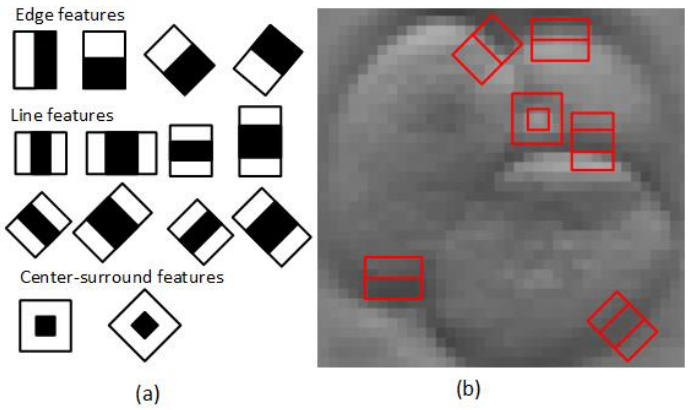


Figure 7: Graphical representation of Haar-like features: (a) templates of different Haar features; (b) a sub-image of embryo with different feature templates

the detection of rough location. A more accurate location of the embryo is then estimated in the next step using the radiating lines over the image filtered by a Sobel filter. The proposed algorithm uses a grayscale image as an input. The rectangular region of interest (ROI) is returned after the execution of the algorithm. The input image is processed in different scales in order to locate an embryo of the correct size (steps 3–10). If all Haar-like features are applied to satisfy the condition in step 7, then the rough location of the embryo is detected (step 8). A more accurate location (ROI*) of the embryo is estimated in steps 11–15. Sobel filter [25] is used to find the approximate gradient magnitude at each point in the gray-scale image at the ROI (step 11). The radiating lines at each point of the detected square are drawn based on the given parameters, such as line length and the angle between lines. For this purpose, Bresenham's Line-drawing Algorithm [26] is applied (step 13). The sum of gradient magnitude for each concentric circle is determined at each point lying in the lines. The result of this step is a histogram of obtained values. The point estimate is computed by determining the maximal value in the histogram and its distance from the centre (step 14).

Algorithm 1. Embryo location detection pseudo-algorithm

```
Input: gray-scale image
Output: ROI of the detected embryo in the image
foreach pyramid scale of the image do
        Down sample the image to create image_i
        Apply HAAR features on image_i sub-window
        Apply cascade classifier on image_i sub-window
        if sub-window passed all per-stage checks then
              Accepted ROI as rough location of the embryo
        end if
end foreach
Apply Sobel filter to ROI of the image
foreach radiating line in ROI do
        Find the key points of the embryo boundary
        Estimate the accurate embryo location ROI*

end foreach
```

The advantage of the proposed algorithm is the ability to strengthen edges at a substantially equal distance from the central point. Although Sobolev gradient-based optimisers have been used in some previous studies [27–29], the method proposed in this work efficient uses the traditional optimiser. In addition, the proposed approach is suitable to detect weak and round curves in a noisy background since it provides successful results without an extra step for noise reduction or intensity normalisation, as seen in previous studies [30, 31]. In comparison, noise reduction is usually applied based on the determined noise types or levels while using traditional methods [32, 33]. For the further processing of images, it is important that the entire embryo is correctly cropped, which is the basis for the determining the cell size, monitoring embryo development stages and then classifying them into defined classes.

Alternatively, this task could be solved using object detection methods such as Local Binary Patterns (LBP) or Histogram of Oriented Gradients (HOG). Both methods were tested, but the cascade classifier was selected for further development. The HOG and LBP methods lack localisation accuracy because they require a high contrast image, where the target object is captured with sharp edges. Moreover, these methods fail in detecting partially

overlapped, noisy or blurred objects, as well as they are too sensitive to object rotation and the location of a region of the target object [34–37]. An embryo image captured using a time-lapse system is slightly blurry and the boundaries of the embryo are too fuzzy; therefore, methods that are able to generalise the results should be employed.

Identification of Embryo Development Stage by Developing a Convolutional Neural Network-based Classification System

The identification of early-stage embryo development is formulated as a multi-class prediction problem with the aim to identify the cell number during the division process until day 5 of embryo development. The first attempt to solve the given problem incorporated the use of Principal Component Analysis (PCA) and SVM. A cascade classifier was used to detect the location of the embryo in the image. PCA was for the reduction of data dimensionality and feature extraction. SVM was trained to classify different cell stages based on PCA features. The combination of a cascade classifier, PCA and SVM gave approximately 85% classification accuracy. Therefore, we employed CNNs to construct an embryo cell classification system, since CNNs have become one of the most widely used models of deep learning and demonstrate high accuracy performance results in various image recognition tasks [38, 39]. A general CNNs architecture consists of several convolutions, pooling, and fully connected layers. A convolutional layer computes the output of neurons that are connected to the local regions in the input. A pooling layer reduces the spatial size of the representation in order to minimise the number of parameters and computations in the network. These layers are followed by fully connected layers leading to the softmax layer, which is the final classifier. Two popular architectures, AlexNet and VGG16, were selected for the present experiments (see Fig. 8). Experimental investigations were executed on a Windows 10 machine with 16.0 GB of RAM installed with an Intel Core i7-7700K 4.20GHz CPU. Less than 45 milliseconds were required to process one image and around 1 minute (depending on the number of incubating days) was required to analyse entire embryo development from the beginning to end.

AlexNet demonstrates high classification results in different types of applications while retaining a simple and clear structure [40]. As a result, the network of this architecture is easy to implement. The small number of parameters does not require large computational and memory resources. This architecture consists of five convolutional layers and three fully connected layers. AlexNet includes max pooling and makes use of a rectified linear unit (ReLU) nonlinearity which allows training of the network much faster compared to using a common activation function (e.g. tanh or sigmoid) together with data augmentation and dropout regularisation in order to avoid overfitting.

VGG16 network [41] is an improvement over AlexNet by providing the deeper architecture. A total of 16 layers exist in this architecture, including 13 convolutional layers and 3 fully connected (FC) layers followed by a softmax classifier. In VGG16, large kernel-sized filters in the first convolutional layers (11×11, 5×5) are replaced with multiple 3×3 filters that are used in all 13 convolutional layers. Max pooling layers use only a 2×2 px window with stride of 2. For all convolutional layers, the stride and padding are set to 1 px.

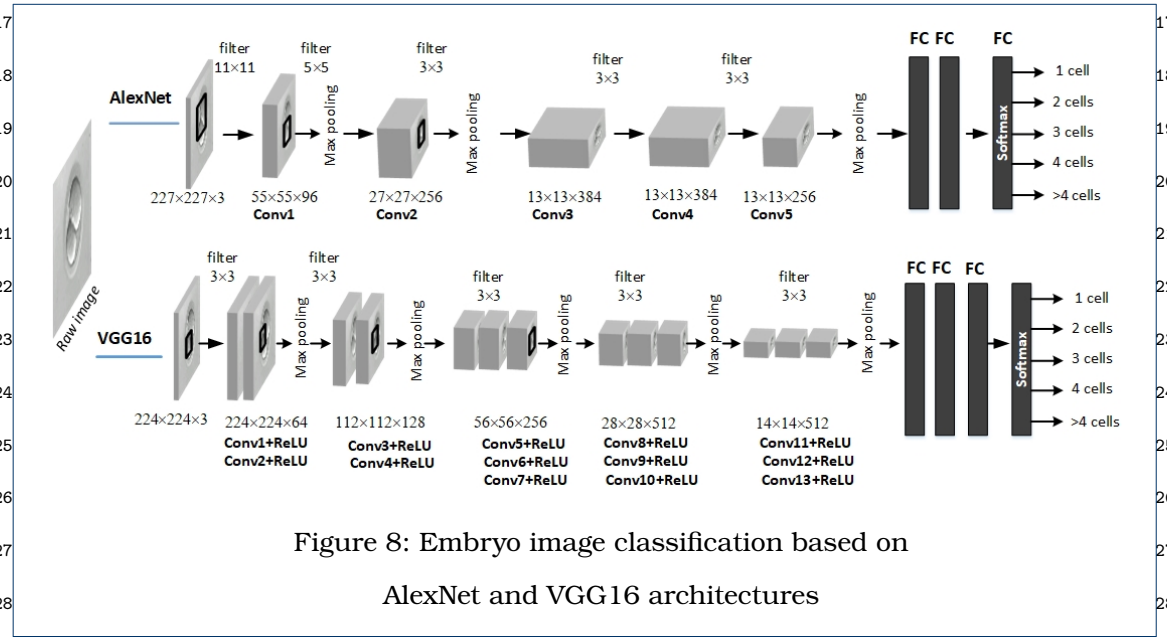


Figure 8: Embryo image classification based on AlexNet and VGG16 architectures

Comparison of these two architectures reveals that VGG16 has twice as many parameters (~527 MB of required memory) as AlexNet (~232 MB of required memory), which makes it more likely to observe VGG16 demonstrating ~15% higher classification accuracy over AlexNet [42]. However, the computational complexity of VGG16 is very high, being 10 times greater than that of AlexNet. Notably, AlexNet is one of a few CNNs models that capable of achieving super real-time performance with very small batch sizes, thus allowing it to reduce the consumption of system memory (e.g., a batch size of 1 requires less than 1 GB memory). In this research, both architectures are used to explore and estimate their possibilities of achieving high accuracy results (more than 90%) in identifying a total cell number in images of an embryo.

The classification model has been implemented using MatConvNet [43], an open-source implementation of CNNs in the MATLAB environment that can be easily extended in order to develop new CNNs architectures. Specific software and hardware requirements exist for deep learning model implementations, such as MATLAB 2016a (or later version), a C\C++ compiler, and a computer with a CUDA-enabled NVIDIA GPU supporting compute capability 2.0 or above.

In general, different types of measures are used to evaluate the performance of the selected classifiers. In the multi-class setting, the outcome is produced for many predefined classes $\{C_1, \ldots, C_i, \ldots, C_K\}$, where K is the class cardinality [19, 20]. Accordingly, for an individual class C_i, the main counts are defined as true positives TP_i, false positives FP_i, false negatives FN_i, and true negatives TN_i. These are the main entrances for the confusion matrix. A list of measures used to assess the performance of a multi-class predictor is richer compared to binary classification. The conventional performance measures are modified to consider the class distribution resulting in macro-averaging or micro-averaging computation. A macro-average defines the performance treating all classes equally, whereas a micro-average considers the contributions of all classes to compute the selected measure. Obviously, in a multi-class setting, a micro-average is preferable if the class imbalance is prominent.

List of abbreviations

CNNs – convolutional neural networks

FC – fully connected

IVF – in-vitro fertilization

PCA – principal component analysis

ReLU – rectified linear unit

ROI – rectangular region of interest

SVM – support vector machine

TL – time-lapse

Declarations

Acknowledgements

Not applicable.

Author contributions

Conceptualisation, VR and APT; methodology, VR and KS; software and data curation, VR and DJ; validation, APT and KS. All authors assisted in writing and improving the paper.

Funding

This research received no external funding.

Availability of data and material

The image data set used to support the findings of this study has not been made publicly available because the images are owned by a private IVF laboratory (ESCO MEDICAL Ltd., company code 303705851, Draugystes str. 19, 51230 Kaunas, Lithuania) and are available by request only.

Ethics approval and consent to participate

The images were captured by a private IVF laboratory (ESCO MEDICAL Ltd., company code 303705851, Draugystes str. 19, 51230 Kaunas, Lithuania) and then shared with the research group "Smart Automatic Control Systems" led by prof. Vidas Raudonis for research purposes under Data Use Agreement established on 3rd of September, 2018.

Consent for publication

The images were used as anonymised data set for research purpose only.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Automation, Kaunas University of Technology, 51367 Kaunas, Lithuania. [2]Department of Applied Informatics, Kaunas University of Technology, 51368 Kaunas, Lithuania. [3]Department of Mathematical Modelling, Kaunas University of Technology, 51368 Kaunas, Lithuania.

References

1. InkWood Research: Global in Vitro Fertilization Market Forecast 2018-2026. Accessed 01 March 2019 (2018). https://www.inkwoodresearch.com/reports/in-vitro-fertilization-market
2. Mordor Intelligence: Global In-Vitro Fertilization Market - Segmented By Instruments, Technology, End Users, and Geography - Growth, Trends, and Forecast 2019-2024. Accessed 22 March 2019 (2019). https://www.mordorintelligence.com/industry-reports/invitro-fertilization-market
3. FertilityIQ: IVF - In Vitro Fertilization. The most successful, invasive, and expensive fertility treatment. Accessed 04 January 2019 (2019). https://www.fertilityiq.com/ivf-in-vitro-fertilization/
4. Paulson, R.J.: Time-lapse imaging. Fertility and Sterility 109(4), 583–583 (2018). doi:10.1016/J.FERTNSTERT.2018.02.013
5. Wang, S.-S., Sun, H.-X.: Blastocyst transfer ameliorates live birth rate compared with cleavage-stage embryos transfer in fresh in vitro fertilization or intracytoplasmic sperm injection cycles: reviews and meta-analysis. Yonsei medical journal 55(3), 815–825 (2014). doi:10.3349/ymj.2014.55.3.815
6. Wong, C.C., Loewke, K.E., Bossert, N.L., Behr, B., De Jonge, C.J., Baer, T.M., Pera, R.A.R.: Non-invasive imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage. Nature Biotechnology 28, 1115–1121 (2010). doi:10.1038/nbt.1686
7. Lyssek-Boroń, A., Wylęgała, A., Polanowska, K., Krysik, K., Dobrowolski, D.: Longitudinal Changes in Retinal Nerve Fiber Layer Thickness Evaluated Using Avanti Rtvue-XR Optical Coherence Tomography after 23G Vitrectomy for Epiretinal Membrane in Patients with Open-Angle Glaucoma. Journal of Healthcare Engineering (2017). doi:10.1167/iovs.12-9782
8. Gao, X., Boccalini, S., Kitslaar, P.H., Budde, R.P.J., Tu, S., Lelieveldt, B.P.F., Dijkstra, J., Reiber, J.H.C.: A novel software tool for semi-automatic quantification of thoracic aorta dilatation on baseline and follow-up computed tomography angiography. The International Journal of Cardiovascular Imaging 35(4), 711–723 (2019). doi:10.1007/s10554-018-1488-9
9. Goceri, E., Goceri, N.: Deep learning in medical image analysis: Recent advances and future trends. In: 11th Int.l Conf. on Computer Graphics, Visualization, Computer Vision and Image Processing (CGVCVIP 2017), pp. 305–310 (2017)
10. Lundervold, A.S., Lundervold, A.: An overview of deep learning in medical imaging focusing on mri. Zeitschrift für Medizinische Physik 29(2), 102–127 (2019). doi:10.1016/j.zemedi.2018.11.002. Special Issue: Deep Learning in Medical Physics
11. Khosravi, P., Kazemi, E., Zhan, Q., Malmsten, J.E., Toschi, M., Zisimopoulos, P., Sigaras, A., Lavery, S., Cooper, L.A.D., Hickman, C., Meseguer, M., Rosenwaks, Z., Elemento, O., Zaninovic, N., Hajirasouliha, I.: Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization. npj Digital Medicine 2(1), 21 (2019). doi:10.1038/s41746-019-0096-y
12. Iwata, K., Sato, M., Matsumoto, I., Shimura, T., Yumoto, K., Negami, A., Mio, Y.: Deep learning based on images of human embryos obtained from high-resolusion time-lapse cinematography for predicting good-quality embryos. Fertility and Sterility 110(4), 213 (2018). doi:10.1016/j.fertnstert.2018.07.615
13. Jeanray, N., Marée, R., Pruvot, B., Stern, O., Geurts, P., Wehenkel, L., Muller, M.: Phenotype classification of zebrafish embryos by supervised learning. PLOS ONE 10(1), 1–20 (2015). doi:10.1371/journal.pone.0116989
14. Rocha, J.C., Passalia, F.J., Matos, F.D., Takahashi, M.B., Ciniciato, D.d.S., Maserati, M.P., Alves, M.F., de Almeida, T.G., Cardoso, B.L., Basso, A.C., Nogueira, M.F.G.: A method based on artificial intelligence to fully automatize the evaluation of bovine blastocyst images. Scientific Reports 7(1) (2017). doi:10.1038/s41598-017-08104-9
15. Viswanath, P.S., Weiser, T., Chintala, P., Mandal, S., Dutta, R.: Grading of mammalian cumulus oocyte complexes using machine learning for in vitro embryo culture. In: 2016 IEEE-EMBS International Conference on Biomedical and Health

Informatics (BHI), pp. 172–175 (2016). doi:10.1109/BHI.2016.7455862

16. Filho, E.S., Noble, J.A., Poli, M., Griffiths, T., Emerson, G., Wells, D.: A method for semi-automatic grading of human blastocyst microscope images. Human Reproduction 27(9), 2641–2648 (2012)
17. Malmsten, J., Zaninovic, N., Zhan, Q., Toschi, M., Rosenwaks, Z., Shan, J.: Automatic prediction of embryo cell stages using artificial intelligence convolutional neural network. Fertility and Sterility 110(4), 360 (2018). doi:10.1016/j.fertnstert.2018.07.1005
18. Wang, Y., Moussavi, F., Lorenzen, P.: Automated embryo stage classification in time-lapse microscopy video of early human embryo development. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013, pp. 460–467. Springer, Berlin, Heidelberg (2013). doi:10.1007/978-3-642-40763-5-57
19. Döring, M.: Performance Measures for Multi-Class Problems. Accessed 05 April 2019 (2018). https://www.datascienceblog.net/post/machine-learning/performance-measures-multi-class-problems
20. Sokolova, M., Lapalme, G.: A systematic analysis of performance measures for classification tasks. Information Processing & Management 45(4), 427–437 (2009). doi:10.1016/j.ipm.2009.03.002
21. Montag, M.: How a Decision Support Tool Based on Known Implantation Data Can Enhance Embryo Selection. Accessed 06 September 2019 (2015). https://blog.vitrolife.com/togetherallthewayhow-a-decision-support-tool-based-on-known-implantation-data-can-enhance-embryo-selection
22. Lienhart, R., Maydt, J.: An extended set of haar-like features for rapid object detection. In: Proceedings. International Conference on Image Processing, vol. 1, pp. 900–903 (2002). doi:10.1109/ICIP.2002.1038171
23. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, vol. 1, pp. 511–518 (2001). doi:10.1109/CVPR.2001.990517
24. OpenCV: Cascade Classification. Haar Feature-based Cascade Classifier for Object Detection. Accessed 25 February 2019 (2019). https://docs.opencv.org/2.4/modules/objdetect/doc/cascade_classification.html
25. Gupta, S., Mazumda, S.G.: Sobel edge detection algorithm. International Journal of Computer Science and Management Research 2(2), 1578–1583 (2013)
26. Flanagan, C.: Bresenham Line-Drawing Algorithm. Accessed 20 February 2019 (1996). https://www.cs.helsinki.fi/group/goa/mallinnus/lines/bresenh.html
27. Goceri, E.: Diagnosis of alzheimer's disease with sobolev gradient based optimization and 3d convolutional neural network. International Journal for Numerical Methods in Biomedical Engineering 35, 3225 (2019). doi:10.1002/cnm.3225
28. Goceri, E.: Full-automated liver segmentation using sobolev gradient based level set evolution. International Journal for Numerical Methods in Biomedical Engineering 32 (2016). doi:10.1002/cnm.2765
29. Goceri, E., Martínez, E.D.: A level set method with sobolev gradient and haralick edge detection. Global Journal on Technology, 131–140 (2014)
30. Goceri, E.: Fully automated and adaptive intensity normalization using statistical features for brain mr images. Celal Bayar Üniversitesi Fen Bilimleri Dergisi, 125–134 (2018). doi:10.18466/cbayarfbe.384729
31. Goceri, E.: Intensity normalization in brain mr images using spatially varying distribution matching. In: 11th Int. Conf. on Computer Graphics, Visualization, Computer Vision and Image Processing (CGVCVIP 2017), pp. 300–304 (2017)
32. Khmag, A., Ramli, A.R., Al-haddad, S.A.R., Kamarudin, N.: Natural image noise level estimation based on local statistics for blind noise reduction. The Visual Computer 34(4), 575–587 (2018). doi:10.1007/s00371-017-1362-0
33. Goceri, E., Goksel, B., B Elder, J., Puduvalli, V., Otero, J., N Gurcan, M.: Quantitative validation of anti-ptbp1 antibody for diagnostic neuropathology use: Image analysis approach. International journal for numerical methods in biomedical engineering 33 (2016). doi:10.1002/cnm.2862
34. Alhindi, T.J., Kalra, S., Ng, K.H., Afrin, A., Tizhoosh, H.R.: Comparing lbp, hog and deep features for classification of histopathology images. 2018 International Joint Conference on Neural Networks (IJCNN), 1–7 (2018)

35. Chen, J., Zheng, H., He, N., Luo, Z., Zhu, R.: Object detection by learning oriented gradients. In: Zhao, Y., Kong, X., Taubman, D. (eds.) Image and Graphics, pp. 86–96. Springer, Cham (2017)
36. Ozturk, S., Akdemir, B.: Comparison of hog, mser, sift, fast, lbp and canny features for cell detection in histopathological images. HELIX 8, 3321–3325 (2018). doi:10.29042/2018-3321-3325
37. Huang, d., Shan, C., Ardabilian, M., Chen, L.: Local binary patterns and its application to facial image analysis: A survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C 41, 765–781 (2011). doi:10.1109/TSMCC.2011.2118750
38. Rawat, W., Wang, Z.: Deep convolutional neural networks for image classification: A comprehensive review. Neural Computation 29(9), 2352–2449 (2017). doi:10.1162/NECO_a_00990
39. Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., Stefanovic, D.: Deep neural networks based recognition of plant diseases by leaf image classification. Computational Intelligence and Neuroscience 2016, 1–11 (2016). doi:10.1155/2016/3289801
40. Krizhevsky, A., Sutskever, I., E. Hinton, G.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, vol. 1, pp. 1097–1105 (2012). doi:10.1061/(ASCE)GT.1943-5606.0001284
41. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: The 3rd International Conference on Learning Representations (ICLR2015), pp. 1–14 (2015). https://arxiv.org/pdf/1409.1556.pdf
42. Bianco, S., Cadene, R., Celona, L., Napoletano, P.: Benchmark analysis of representative deep neural network architectures. IEEE Access 4, 64270–64277 (2018). doi:10.1109/ACCESS.2018.2877890
43. Vedaldi, A., Lenc, K.: Matconvnet: Convolutional neural networks for matlab. In: Proceedings of the 23rd ACM International Conference on Multimedia. MM '15, pp. 689–692. ACM, New York, USA (2015). doi:10.1145/2733373.2807412

Proofreading Certificate

Response Letter