International Journal of Applied Operational Research
Vol. 12, No. 3, pp. 13-20, Summer 2024

Journal homepage: ijorlu.liau.ac.ir

An efficient framework for sperm cell analysis using deep learning methods to improve infertility treatment

P. Afzali*

Received: 12 March 2024 ; **Accepted:** 10 June 2024

Abstract Human infertility is a significant global health concern that impacts the lives of numerous couples attempting to conceive. A substantial proportion of infertility cases are attributed to male-related factors. A pivotal component in the diagnosis of male infertility is semen analysis, which heavily depends on the evaluation of sperm head morphology, particularly the shape and size of the sperm head. In many instances, this evaluation is conducted manually, with its accuracy relying on the expertise and experience of the physician. This manual process is not only time-intensive and laborious but also subject to considerable variations depending on different laboratories or physicians within the same laboratory. Such variation poses challenges in ensuring consistent and reliable diagnostic outcomes. This paper presents a hybrid deep learning-based framework designed to enable automated and accurate classification of human sperm heads based on sperm images. Experimental results obtained from two datasets, HuSHeM and SCIAN, demonstrate the proposed framework's superior effectiveness in analyzing and classifying sperm head morphology compared to existing methodologies.

Keyword: Sperm Morphology Analysis, Infertility, Image Classification, Deep Learning.

[DOI: 10.71885/ijorlu-2024-1-678]

1 Introduction

Infertility constitutes a considerable problem on a global scale, affecting a significant proportion of the population of reproductive age [1]. Beyond the evident negative impact on the quality of life of affected couples, it can also give rise to a range of social and psychological consequences [2]. Male factors are responsible for approximately half of all infertility cases [3]. Semen analysis, a common initial evaluation for male infertility, is performed by a clinician and encompasses various aspects, including sperm morphology, count, and motility [4]. Specifically, sperm morphology analysis can be performed on fixed and stained sperm as a preliminary screen for infertility diagnosis or during the selection of viable sperm for intracytoplasmic sperm injection (ICSI) which is a popular assisted reproductive technology [5].

Manual sperm assessment is contingent upon the operator's expertise and is susceptible to variability and inaccuracy [6]. This can result in inconsistent estimates [7], underscoring the necessity for the development of automated methodologies for precise evaluation.

---

* **Corresponding Author.** (✉)
E-mail: Parvaneh.afzali@iau.ac.ir; P.afali@gmail.com

P. Afzali
Department of Computer Engineering, Astaneh Ashrafieh Branch, Islamic Azad University, Astaneh Ashrafieh, Iran

Contemporary machine learning algorithms have facilitated the development of diverse automated sperm analysis techniques, thereby expediting sperm classification [8]. Given the significance and complexity of detecting abnormalities in the sperm head region, the automatic identification and segmentation of this region is considered a crucial step in the analysis and classification of sperm cell images [9].

In medical diagnoses and several branches of machine learning, the issue of imbalanced data represents a prominent challenge [10-11]. Imbalanced datasets are pervasive in numerous practical classification problems, including text classification, fraud detection, face recognition, social media analysis, and defect detection in medicine [12-13]. The term imbalanced dataset refers to a dataset in which the number of samples in one or more classes (minority class) is significantly less than the other classes (majority class) [14-15]. Conventional classification algorithms, including the Bayesian classifier, decision tree, k-nearest neighbor, neural network, and support vector machine, operate under the assumption of a balanced training dataset. Consequently, an imbalanced distribution in the data can lead to a bias toward the majority class training samples, resulting in an increased error in identifying minority samples and reduced prediction efficiency of these algorithms [16]. Indeed, the primary objective of these approaches is to enhance the recognition rate of all samples; consequently, the recognition rate of minority class samples can be readily compromised [10]. However, in numerous instances, the identification of minority class samples is of greater significance than the identification of majority class samples [13]. A notable example is in the diagnosis of medical defects and diseases, wherein the repercussions of the misdiagnosing a sick individual as a healthy individual prove to be more substantial than those resulting from misdiagnosing a healthy individual as a sick individual. This is due to the fact that a misdiagnosis can lead to a missing out on the patient's optimal course of treatment and facing deleterious consequences. Consequently, the pursuit of enhanced precision in the classification of minority class samples, along with elevated overall accuracy in imbalanced datasets, has emerged as a compelling research domain within the realm of machine learning [17]. In the context of ICSI, the repercussions of erroneously categorizing an abnormal sperm as a normal sperm are considerably more deleterious than the consequences of misclassifying a healthy sperm as an abnormal sperm [5].

A substantial amount of research has been carried out in the domain of human sperm cell identification, segmentation, and classification. In their work, Chang et al. [18] presented a comprehensive framework for the identification and segmentation of the head, acrosome, and nucleus of human sperm. This framework comprises two distinct stages. Initially, the regions of interest for the heads are identified through the application of k-means clustering, and subsequently, the candidate heads are refined and enhanced through the utilization of morphological techniques and histogram statistical analysis. Following the completion of head segmentation, the second stage involves the division of the identified head into two regions, the acrosome and the nucleus. This is achieved through the application of the Otsu thresholding method and geometric constraints. A notable strength of this study lies in the provision of a reliable and standardized dataset comprising images of fixed and stained sperm. This dataset serves as a valuable resource, paving the way for the validation of novel methods for sperm identification and segmentation. Chang et al. [19] proposed a novel methodology for segmenting sperm heads, employing the method outlined in [18]. This approach involves a two-stage classification scheme that categorizes sperm heads into five distinct classes: normal sperm and four classes of abnormal sperm heads. Shaker et al. [20] sought to classify the sperm head region into four classes following automatic identification and segmentation. To this end, they examined the discriminative power of a number of suitable features extracted

[DOI: 10.71885/ijorlu-2024-1-678]

from microscopic images, called elliptical features, and used the LDA approach for automatic classification.

In recent years, deep learning algorithms, particularly convolutional neural networks (CNNs), have witnessed a surge in popularity within the domain of medical image analysis [21] and have rapidly superseded conventional machine learning algorithms that rely on human-extracted features for tasks such as object recognition and image classification [22]. Li et al. [23] have developed a framework, designated as CLU-CNNs, for the purpose of object recognition in medical images. This framework operates on a complete image as an input. This input is processed through three convolutional blocks to generate feature maps. Akbari Movaheed et al. [24] sought to identify the outer parts of sperm, including the head and axial filament, and the inner parts, including the acrosome and nucleus inside the head, and the tail and middle segment inside the axial filament. In this study, two CNN models are employed to segment the outer parts, i.e., the head and axial filament. To achieve the acrosome and nucleus segments, the k-means clustering approach is used on the head segments, and finally, the support vector machine classifier is used to classify each pixel of the axial filament segments in order to extract the tail and middle segment regions from the obtained segments. Javadi et al. [5] designed and trained a deep neural network from scratch to classify a sperm image dataset that they collected and publicly presented. The method presented in this article is capable of functioning effectively with non-color and low-resolution images and can analyze each sperm image in real time in less than 25 milliseconds. Consequently, the proposed method can assist embryologists in making rapid decisions regarding the selection of sperm.

Transfer learning constitutes a salient component of deep learning, serving as a reliable and efficient approach for training deep neural networks, particularly within the domain of medical image analysis, where labeled data sets are often scarce. Riordan et al. [5] proposed a deep learning methodology for the classification of sperm in two publicly accessible datasets, namely SCIAN and HuSHeM. Rather than training a deep convolutional neural network model from the beginning, the authors employed a transfer learning approach, retraining the VGG16 model with distinct configurations for the two datasets.

The aforementioned works collectively illustrate the substantial potential of diverse machine learning approaches in sperm identification, analysis, and classification. These approaches can facilitate rapid differentiation of subtle sperm shapes, which were previously limited to specialist physicians.

Given the significant financial burden imposed by assisted reproductive techniques and the potential adverse effects of hormonal medications employed in these procedures, this study endeavored to develop an automated and efficient approach for the precise classification of human sperm head images by introducing a two-way deep framework. The remainder of this paper is structured as follows: The basic concepts are presented in Section2. Section 3 details the proposed methodology and the dataset utilized in this study. Section 4 encompasses the experimental procedures conducted to assess the efficacy of the proposed approach and a thorough examination of the outcomes. Conclusions and suggestions are also presented in Section 5.

2 Preliminaries and definitions

This section presents the fundamental concepts necessitated and the preliminary information that we used in our study.

[DOI: 10.71885/ijorlu-2024-1-678]

2.1 Convolutional neural network

Convolutional neural networks (CNNs) represent a particular type of deep neural network that has been developed for the purpose of processing image data and machine vision. These networks are designed to extract visual features in a gradual, hierarchical manner by leveraging convolutional layers. The lower layers of the network are responsible for detecting simple visual details such as edges and color patterns, while the higher layers are tasked with identifying more complex structures such as objects and shapes. This unique structural design renders convolutional neural networks highly effective in a variety of applications, including image recognition, classification, and various machine vision tasks.

2.2 Capsule neural networks

Capsule neural networks are a type of neural network architecture designed to improve the ability to understand and identify more complex features from image data. Unlike convolutional neural networks, which may not properly account for spatial information and relationships between features, capsule neural networks use structures called capsules. These capsules are groups of neurons that are capable of maintaining and modeling spatial relationships between components and features. This architecture facilitates enhanced comprehension of information regarding the position, orientation, and scale of objects, thereby enhancing the accuracy of pattern recognition and mitigating errors stemming from rotation or variations in inputs.

2.3 Transfer learning

Transfer learning is a machine learning method that involves leveraging a model trained on a specific task or domain to enhance performance or reduce the reliance on extensive data sets for other tasks or domains. Rather than initializing a model from the beginning, transfer learning utilizes the knowledge and features previously acquired in the original model and generalizes them to a new task. This approach is particularly advantageous in scenarios where training data is scarce, as it enables tasks such as image recognition and classification to be executed more expeditiously and efficiently.

3 Proposed framework

This section delineates the utilized dataset of sperm head images, and the proposed hybrid architecture.

3.1 Data sets used

In this study, the training and evaluation of the proposed model was conducted using labeled sperm head images from two freely available datasets: HuSHeM [25], [26] (which contains four classes of normal, conical, pear-shaped, and amorphous sperm images) and SCIAN [27] (which contains five classes of normal, conical, pear-shaped, amorphous, and small sperm images). Representative sample images from these two datasets are presented in Fig. 1.

[DOI: 10.71885/ijorlu-2024-1-678]

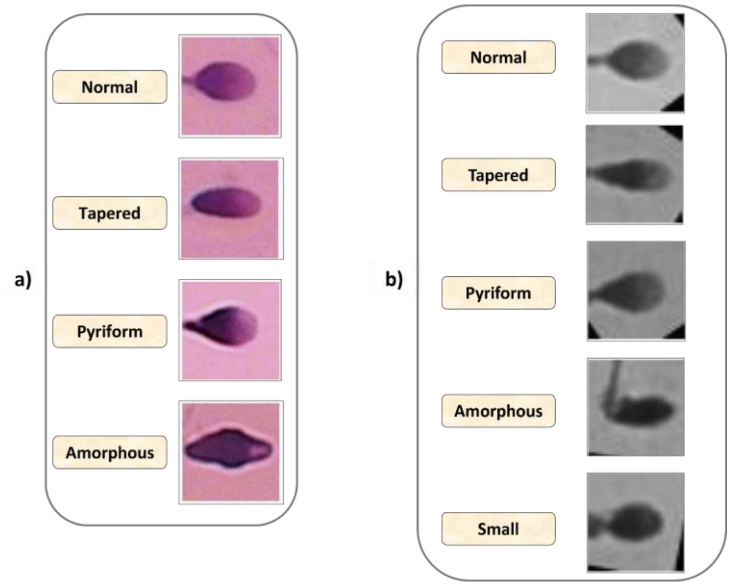

Fig. 1 Illustrative examples of sperm head images from: a) the HuSHeM, b) SCIAN datasets

In order to address the imbalance among classes and mitigate the classifier's bias toward a particular class, images within classes with fewer samples were augmented using data augmentation methods. These methods included rotation, brightness adjustment, and image enlargement and reduction, with the objective of achieving an equivalent number of images across all classes.

3.2 Proposed model architecture

The proposed deep dual-path architecture for human sperm head image classification is illustrated in Figure 2. This hybrid architecture incorporates a convolutional neural network model augmented with a capsule neural network in the first path to enhance the model's learning performance by considering the spatial relationships between features. In the second path of this architecture, a ResNet-50 model is embedded as a customized pre-trained CNN based on transfer learning to enhance the discrimination of the learned features. Finally, the features extracted from these two paths are merged, and the input sperm head image is classified.

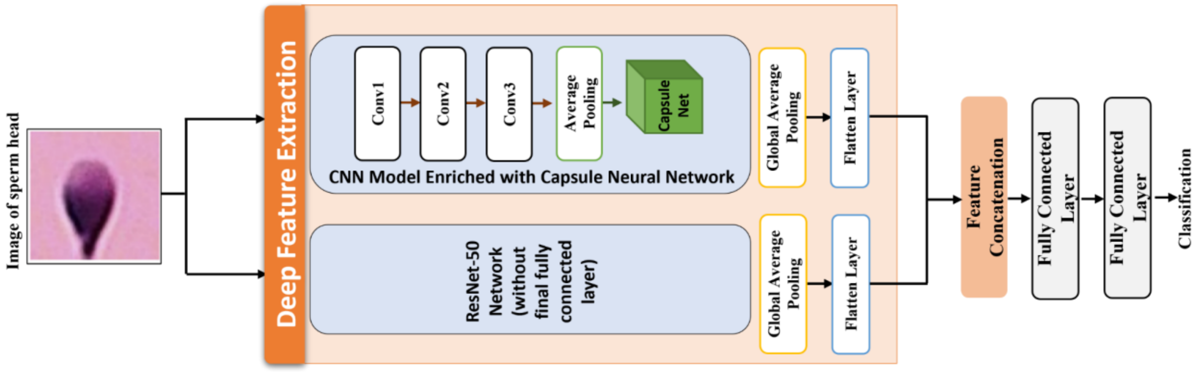

Fig. 2 Dual-path architecture of the proposed model for classifying sperm head images.

4 Experiments and discussion

To evaluate the proposed model, three metrics were used: accuracy, precision, and recall or true positive rate.

Accuracy represents the ratio of total correct predictions (both positive and negative) to total samples. In other words, accuracy indicates how well a model has classified samples. The accuracy formula is as follows:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \tag{1}$$

Where:
TP: Number of true positives that are correctly detected (True Positives)
TN: Number of true negatives that are correctly detected (True Negatives)
FP: Number of true negatives that are falsely detected (False Positives)
FN: Number of true positives that are falsely detected (False Negatives)

Precision represents the accuracy of the model in detecting positive samples and indicates how many of the samples that the model detects as positive are actually positive. The precision formula is as follows:

$$Precision = TP/(TP + FP) \tag{2}$$

Recall indicates how successful the model was in identifying true positives. In other words, what percentage of all positives were correctly identified. This metric is also referred to as the true positive rate and its formula is as follows:

$$Recall = TP/(TP + FN) \tag{3}$$

The proposed model was trained and evaluated on the HuSHeM dataset, with the minority class considered in three distinct cases: the Tapered class in the first case, the Pyriform class in the second case, and the Amorphous class in the third case. The performance of the proposed model on this image set was evaluated based on three criteria: accuracy, precision, and true positive rate, as illustrated in Table 1.

Table 1 Performance of the proposed model in classifying sperm head images from the HuSHeM dataset.

Minority Class	Accuracy	Precision	True Positive Rate
Tapered	92.82	93.80	92.51
Pyriform	95.15	96.32	94.92
Amorphous	93.72	94.84	93.41

In addition, the proposed model was trained and evaluated on the SCIAN dataset, with the minority class being considered in four distinct cases: the Tapered class in the first case, the Pyriform class in the second case, the Amorphous class in the third case, and the Small class in the fourth case. The performance of the proposed model, based on three criteria of accuracy, precision, and true positive rate on this dataset, is shown in Table 2.

[DOI: 10.71885/ijorlu-2024-1-678]

Table 2 Performance of the proposed model in classifying sperm head images from the SCIAN dataset.

Minority Class	Accuracy	Precision	True Positive Rate
Tapered	89.72	88.41	91.40
Pyriform	94.11	93.12	94.91
Amorphous	87.50	86.84	93.42
Small	92.24	91.50	92.33

As evidenced by the findings derived from both datasets, the model demonstrated a high degree of efficacy in the classification of sperm head images, as measured by the established evaluation criteria. Additionally, the true positive rate for the Pyriform class exhibited a notable level of effectiveness, suggesting that sperm head images classified as Pyriform are more readily identifiable than those of other types.

5 Conclusion

Infertility affects approximately 15% of couples, and sperm morphology analysis is a key element in diagnosing male infertility. Performing this examination by a specialist physician is time-consuming and prone to errors; therefore, the use of artificial intelligence technologies to automate and standardize this analysis is beneficial. In this study, we presented a deep framework that utilizes convolutional neural networks, capsule neural networks, and pre-trained networks for the classification of sperm head images. Our proposed model was trained and evaluated on sperm head images from two datasets, HuSHeM and SCIAN. The results indicate the model's high accuracy and good performance in image classification. Therefore, this study emphasizes the potential of using deep learning approaches for the automatic and reliable classification of sperm head images. Future research directions include the exploration of more advanced methods to address data imbalance among sperm head image classes, as well as the application of optimization methods.

References

1. Cherouveim, P., Velmahos, C., McCallum Bormann, C., (2023). Artificial intelligence for sperm selection-a systematic review. Fertility and Sterility; 120(1), 24-31.
2. Onat, G., Kizilkaya Beji, N., (2012). Marital Relationship and Quality of Life Among Couples with Infertility. Sexuality and Disability; 30(1), 39-52.
3. Maduro, M., Lamb, D., (2002). Understanding New Genetics of Male Infertility. The Journal of Urology; 168(5), 2197-2205.
4. Riordon, J., McCallum, C., Sinton, D., (2019). Deep learning for the classification of human sperm. Computers in Biology and Medicine; 111, 103342.
5. Javadi, S., Mirroshandel, S.A., (2019). A novel deep learning method for automatic assessment of human sperm and images. Computers in Biology and Medicine; 109, 182–194.
6. Brazil C., (2010). Practical semen analysis: from A to Z. Asian Journal of Andrology, 12(1), 14-20.
7. Maggavi, R.R., Pujari, S.A., Vijaykumar, C.N., (2019). Motility Analysis with Morphology: Study Related to Human Sperm. Procedia Computer Science;152, 179–185.
8. Iqbal, I., Mustafa, G., Ma, J., (2020). Deep Learning-Based Morphological Classification of Human Sperm Heads. Diagnostics; 10(5), 325.
9. Shaker, F., Monadjemi, S.A., Naghsh-Nilchi, A.R., (2016). Automatic detection and segmentation of sperm head, acrosome and nucleus in microscopic images of human semen smears. Computer Methods and Programs in Biomedicine; 132, 11-20.

10. Yamnin, S., Wong, A., S. Kamel, M., (2009). Classification of Imbalanced Data: a review. International Journal of Pattern Recognition and Artificial Intelligence, 23(4), 687–719.
11. Krawczyk, B., (2016). Learning from imbalanced data: open challenges and future directions. Prog Artif Intell; 5, 221–232.
12. García, S., Zhang, Z., Altalhi, A., (2018). Dynamic ensemble selection for multi-class imbalanced datasets. Information Sciences; 445–446, 22-37.
13. Yan, Y.T., (2018). A three-way decision ensemble method for imbalanced data oversampling. International Journal of Approximate Reasoning; 107, 1-16.
14. Haixiang, G., (2017). Learning from class-imbalanced data: Review of methods and applications. Expert Systems With Applications; 73, 220–239.
15. He, H., Garcia, E.A., (2009). Learning from Imbalanced Data. IEEE Transactions on Knowledge and Data Engineering; 21(9), 1263–1284.
16. Zhang, C., (2017). An Up-to-Date Comparison of State-of-the-Art Classification Algorithms. Expert Systems with Applications; 82, 128–150.
17. Yang, Q., Wu, X., (2006). 10 Challenging Problems in Data Mining Research. International Journal of Information Technology and Decision Making; 5(4), 597-604.
18. Chang, V., Saavedra, J., Castañeda, V., (2014). Gold-standard and improved framework for sperm head segmentation. Computer Methods and Programs in Biomedicine; 117(2), 225-237.
19. Chang, V., Heutte, L., Petitjean Härtel, S., (2017). Automatic classification of human sperm head morphology. Computers in Biology and Medicine; 84, 205-216.
20. Shaker, F., Monadjemi, S.A., Alirezaie, J., (2017). Classification of human sperm heads using elliptic features and LDA. s.l. : IEEE, 2017. 3rd International Conference on Pattern Recognition and Image Analysis (IPRIA).
21. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A., (2017). A survey on deep learning in medical image analysis. Medical Image Analysis; 42, 60–88.
22. Xie, Y., Richmond, D., (2019). Pre-training on Grayscale ImageNet Improves Medical Image. s.l. : Springer, Cham. Computer Vision – ECCV 2018 Workshops. Lecture Notes in Computer Science; 11134, 476-484.
23. Li, Z., Dong, M., Wen, S., Hu, X., Zhou, P., Zeng, Z., (2019). CLU-CNNs: Object detection for medical images. Neurocomputing; 350, 53-59.
24. Akbari Movahed, R., Mohammadi, E., (2019). Orooji M. Automatic segmentation of Sperm's parts in microscopic images of human semen smears using concatenated learning approaches. Computers in Biology and Medicine;109, 242-253.
25. Shaker F., (2017). Human Sperm Head Morphology Dataset, HuSHeM, p.1. https://doi.org/10.17632/tt3yj2pf38.1.
26. Shaker, F., Monadjemi, S.A., Alirezaie, J., Naghsh-Nilchi, A.R., (2017). A dictionary learning approach for human sperm heads classification. Computers in Biology and Medicine; 91, 181-190.
27. Chang, V., Garcia, A., Hitschfeld, N., Härtel, S., (2017). Gold-standard for computer-assisted morphological sperm analysis. Computers in Biology and Medicine; 83, 143–150.