



Semnan University

# Journal of Modeling in Engineering

Journal homepage: <https://modelling.semnan.ac.ir/>



## Research Article

# A New Capsule Generative Adversarial Network for Imbalanced Classification of Human Sperm Images

Hamed Jabbari<sup>1,\*</sup>, Nooshin Bigdeli<sup>2</sup>

1. Assistant Professor, Control Engineering Department, Faculty of Technical and Engineering, Imam Khomeini International University, Qazvin, Iran.

2. Professor, Control Engineering Department, Faculty of Technical and Engineering, Imam Khomeini International University, Qazvin, Iran.

\*Corresponding Author: [H\\_jabbari@edu.ikiu.ac.ir](mailto:H_jabbari@edu.ikiu.ac.ir)

## PAPER INFO

### Paper history:

Received: 13 September 2022

Revised: 04 January 2023

Accepted: 04 April 2023

### Keywords:

Imbalanced Classification,  
Human Sperm,  
Infertility,  
Generative Adversarial  
Network (GAN),  
Capsule Neural Network  
(CapsNet),  
Deep Learning.

## ABSTRACT

Male infertility as an effective factor can affect the lives of infertile couples. Sperm morphology is an important step in evaluating and examining semen in male infertility. The lack of samples of sperm head abnormalities compared to natural sperm samples can make the classification of sperm head images into an imbalanced classification problem. With the inability of common classification algorithms, capsule neural networks (CapsNet) provide a suitable platform for designing imbalanced classification models compared to other deep networks. Also, Generative Adversarial Networks (GANs) help improve the imbalanced classification of images by producing appropriate artificial samples. To this end, in this paper a new architecture is introduced based on CapsNet and GAN to evaluate the imbalanced classification of human sperm images. Reviewing and comparing the proposed model with other deep learning models in the balanced and imbalanced classification of human sperm images showed the superiority of the proposed model. Investigating the general methods of increasing data with the proposed model to increase data, it was concluded that the general methods have less resistance to reducing the number of data than the proposed model. Balanced classification of human sperm images was done by proposed model with 98.1 % accuracy. The proposed model also maintained a high sensitivity to the minority to the majority of 1:25, indicating its proper performance in the imbalanced classification of sperm images.

© 2023 Published by Semnan University Press.

DOI: <https://doi.org/10.22075/jme.2023.28349.2333>

How to cite this article:

Jabbari, H., & Bigdeli, N. (2023). A New Capsule Generative Adversarial Network for Imbalanced Classification of Human Sperm Images. *Journal of Modeling in Engineering*, 21(73), 279-294.

doi: 10.22075/jme.2023.28349.2333

## یک شبکه مولد متخاصم کپسولی جدید برای طبقه‌بندی نامتوازن تصاویر اسپرم انسان

حامد جباری<sup>۱\*</sup>، نوشین بیگدلی<sup>۲</sup>

اطلاعات مقاله	چکیده
<p>نوع مقاله: پژوهشی            دریافت مقاله: ۱۴۰۱/۰۶/۲۲            بازنگری مقاله: ۱۴۰۱/۱۰/۱۴            پذیرش مقاله: ۱۴۰۲/۰۱/۱۵</p>	<p>ناباروری مردان به عنوان یک عامل موثر، می‌تواند زندگی زوج‌های نابارور را تحت تاثیر قرار دهد. مورفولوژی سر اسپرم یک مرحله مهم در ارزیابی و بررسی مایع منی در ناباروری مردان است. کمبود نمونه‌های مربوط به ناهنجاری‌های سر اسپرم در مقایسه با نمونه‌های سر اسپرم طبیعی، می‌تواند طبقه‌بندی تصاویر سر اسپرم را به یک مسئله طبقه‌بندی نامتوازن تبدیل کند. با عدم توانایی الگوریتم‌های طبقه‌بندی رایج، شبکه‌های عصبی کپسول با در نظر گرفتن ارتباطات فضایی ویژگی‌ها، در مقایسه با سایر شبکه‌های عمیق بستر مناسبی را برای طراحی مدل‌های طبقه‌بندی نامتوازن فراهم می‌کنند. هم‌چنین شبکه‌های مولد متخاصم با تولید نمونه‌های مصنوعی مناسب کمک شایانی به بهبود طبقه‌بندی نامتوازن تصاویر می‌کنند. به‌همین منظور در این مقاله یک معماری جدید بر اساس شبکه‌های کپسولی و شبکه‌های مولد متخاصم برای ارزیابی طبقه‌بندی نامتوازن تصاویر سر اسپرم انسان معرفی می‌گردد. بررسی و مقایسه مدل پیشنهادی با سایر مدل‌های یادگیری عمیق در طبقه‌بندی متوازن و نامتوازن تصاویر سر اسپرم انسان، نشان از برتری مدل پیشنهادی داشت. در بررسی روش‌های عمومی افزایش داده با مدل پیشنهادی برای افزایش داده این نتیجه حاصل شد که روش‌های عمومی از مقاومت کمتری در کاهش تعداد داده‌ها نسبت به مدل پیشنهادی برخوردار است. این مدل، طبقه‌بندی متوازن تصاویر سر اسپرم را با دقت ۹۸/۱ درصد انجام داد. هم‌چنین مدل پیشنهادی تا عدم توازن دسته اقلیت به اکثریت ۱:۲۵، حساسیت بالای ۸۰ درصد را حفظ کرد که نشان از عملکرد مناسب آن در طبقه‌بندی نامتوازن تصاویر اسپرم دارد.</p>
<p><b>واژگان کلیدی:</b>            طبقه‌بندی نامتوازن،            اسپرم انسان،            ناباروری،            شبکه مولد متخاصم،            شبکه عصبی کپسول،            یادگیری عمیق.</p>	

### ۱- مقدمه

بررسی این کیفیت کمک بسزایی کند [۳]. عوامل مردانه در ناباروری را می‌توان با بررسی مشکلات در خروج مایع منی، ویسکوزیته، غلظت، تحرک و مورفولوژی اسپرم تعیین نمود. هم‌چنین طبق تحقیقات صورت گرفته ارتباط مستقیمی بین نتایج لقاح آزمایشگاهی در درمان زوجین نابارور و مورفولوژی طبیعی اسپرم وجود دارد و ناهنجاری‌های جدی در سر اسپرم باعث کاهش نرخ بارداری می‌شود [۴]. بر این اساس، کیفیت مورفولوژیکی اسپرم می‌تواند به عنوان یک

ناباروری یک اختلال رایج است که میلیون‌ها نفر از افراد در سنین باروری را در سراسر جهان تحت تاثیر قرار می‌دهد. این مسئله بر خانواده و جوامع آنها تاثیرگذار است و ممکن است منجر به افسردگی و مشکلات زوجین نیز شود [۱]. برآوردها نشان می‌دهد که عامل بیش از یک سوم موارد در بین زوجین نابارور، مردان هستند [۲]. کیفیت اسپرم نقش مهمی در لقاح تخمک دارد و آنالیز مایع منی می‌تواند در

\* پست الکترونیک نویسنده مسئول: H\_jabbari@edu.ikiu.ac.ir

۱. استادیار، گروه مهندسی برق-کنترل، دانشکده فنی و مهندسی، دانشگاه بین‌المللی امام خمینی (ع) قزوین  
 ۲. استاد، گروه مهندسی برق-کنترل، دانشکده فنی و مهندسی، دانشگاه بین‌المللی امام خمینی (ع) قزوین

کردن فضای ویژگی دارد [۱۶]. در استراتژی یادگیری حساس به هزینه، توسط ماتریس‌های هزینه از طبقه‌بندی غلط نمونه‌های کلاس اقلیت، جلوگیری می‌شود [۱۷]. استراتژی سوم نیز استفاده از شبکه‌های عمیق نظیر شبکه‌های مولد تخصصی<sup>۳</sup> (GAN) به منظور تولید داده است [۱۸]. این شبکه‌ها مدل‌های مولدی هستند که داده‌های جدید شبیه داده‌های آموزشی تولید می‌کنند. اخیراً تشخیص و طبقه‌بندی سر اسپرم انسان با استفاده از الگوریتم‌های یادگیری ماشین به خصوص یادگیری عمیق رونق گرفته است. جیسون رایوردن<sup>۴</sup> و همکارانش یک روش یادگیری عمیق برای طبقه‌بندی اسپرم معرفی کردند [۱۹]. در روش مذکور به منظور اصلاح نامتوازن بین کلاس‌ها، تصاویر کلاس اقلیت کپی شدند و در نهایت درصد صحتی معادل ۹۴/۱ درصد به دنبال داشت [۱۹]. یک مدل یادگیری عمیق بر اساس شبکه‌های کانولوشنی نیز برای تشخیص ناهنجاری مورفولوژی اسپرم انسان توسط سروش جوادی و همکارش پیشنهاد شد [۲۰]. در این الگوریتم، نمونه‌گیری مجدد بر مبنای چرخش، انتقال و تغییر روشنایی برای رفع نامتوازن داده‌ها استفاده شد. شبکه عصبی عمیق پیشنهادی آنها با دقت تقریبی ۸۴ درصد موفق به تشخیص ناهنجاری سر اسپرم شد. امران اقبال<sup>۵</sup> و همکارانش یک شبکه عصبی کانولوشنی برای طبقه‌بندی دقیق سر اسپرم انسان بر اساس تصاویر اسپرم طراحی کردند [۴]. آنها به منظور غلبه بر نامتوازن داده‌ها از سه تکنیک رایج چرخش، جابه‌جایی و قرینه‌سازی استفاده کردند و در نهایت به دقت ۹۵ درصدی رسیدند. عبدالکریم و همکارش الگوریتمی به منظور پیش‌بینی کیفیت اسپرم بر اساس شبکه‌های عصبی با الگوریتم جلبک مصنوعی ارائه دادند. آنها از روش SMOTE برای متعادل کردن کلاس‌ها استفاده کردند [۲۱]. یک روش یادگیری عمیق بر اساس معماری شبکه‌های عصبی کانولوشنی مبتنی بر منطقه<sup>۶</sup> (R-CNN) برای ارزیابی تحرک سر اسپرم در ویدیوهای اسپرم انسان توسط ویکتوریا و همکارانش ارائه شد [۲۲]. در این روش معیار غیرطبیعی بودن اسپرم‌ها، سرعت حرکت سر آنها بود که منجر به دقت ۹۱/۷۷ درصدی در تشخیص غیرطبیعی بودن سر اسپرم شد. در یک مطالعه دیگر، یک مدل هیبریدی در

شاخص مهم در ناباروری مردان باشد. در موارد سنتی، ارزیابی کیفیت مورفولوژیکی با تجزیه و تحلیل آزمایشگاهی و به صورت دستی توسط تکنسین انجام می‌شود که زمانبر است. محدودیت زمان باعث رشد الگوریتم‌های یادگیری ماشین در آنالیز مورفولوژی اسپرم مردان شده است [۵]. با توجه به پیشرفت‌ها و هم‌چنین خصوصیات الگوریتم‌های یادگیری ماشین به ویژه دقت بالا و قابل اعتماد بودن و توانایی در آشکارسازی ویژگی‌های پنهان، آنالیز و طبقه‌بندی خودکار تصاویر اسپرم امکان‌پذیر شده است.

مسئله طبقه‌بندی تصاویر نامتوازن یکی از چالش‌های مهم در طبقه‌بندی تصاویر است [۶]. مجموعه نامتوازن تصاویر به مجموعه تصاویری اشاره دارد که در آن، یک یا چند کلاس شامل نمونه‌های بسیار بیشتری نسبت به سایر کلاس‌هاست. رایج‌ترین کلاس، کلاس اکثریت و کمیاب‌ترین کلاس، کلاس اقلیت نامیده می‌شود که در این میان، اکثر کلاس اقلیت مورد توجه است [۷]. به این دلیل که در بیشتر موارد ناهنجاری مهم مورد نظر در کلاس اقلیت جای می‌گیرد. به عنوان مثال تشخیص و طبقه‌بندی ناهنجاری‌های اسپرم [۸]، تشخیص تومورهای بدخیم موجود در بافت پستان [۹، ۱۰]، تشخیص تصاویر دارای چندشاخگی در ترک در سازه‌های بتنی [۱۱، ۱۲] و شناسایی تومورهای مغز [۱۳] مجموعه‌های نامتوازن هستند که تشخیص درست کلاس اقلیت منجر به تشخیص درست ناهنجاری مربوطه می‌شود. نتایج زیربهبوده<sup>۱</sup>، تمایل دقت پیش‌بینی نسبت به کلاس اکثریت، کمبود نمونه‌های کلاس اقلیت از جمله چالش‌های مدل‌های یادگیری سنتی در طبقه‌بندی نامتوازن است که باعث می‌شوند این مدل‌ها در تشخیص الگوهای نادر و اقلیت موفق نشوند. قابل ذکر است طبقه‌بندی اشتباه کلاس اقلیت خصوصاً در موارد پزشکی ممکن است منجر به هزینه‌های سنگینی شود.

در دهه گذشته، بسیاری از روش‌های یادگیری ماشینی برای مقابله با طبقه‌بندی نامتوازن تصاویر توسعه یافته‌اند که بیشتر آنها بر اساس تکنیک‌های پیش‌پردازش، یادگیری حساس به هزینه و یادگیری عمیق است [۱۴]. رویکردهای پیش‌پردازش شامل روش‌های نمونه‌برداری مجدد<sup>۲</sup> [۱۵] و روش‌های انتخاب و استخراج ویژگی است که سعی در بهینه

<sup>4</sup> Jason Riordon

<sup>5</sup> Imran Iqbal

<sup>6</sup> Region Based Convolutional Neural Networks

<sup>1</sup> Suboptimal

<sup>2</sup> Re-Sampling

<sup>3</sup> Generative Adversarial Network

با چند کلاس اقلیت است. به همین دلیل نیاز به روش‌هایی است که علاوه بر توانایی حل مسئله طبقه‌بندی نامتوازن، نیاز به تعداد داده کمتری در مرحله آموزش داشته باشد. یکی از این روش‌ها، استفاده از معکوس گرافیک کامپیوتری است. در گرافیک کامپیوتری برای رسیدن از یک داده سه بعدی به یک داده دو بعدی (تصویر) کافی است که یک تغییر خطی (ضرب ماتریسی) بر روی داده سه بعدی صورت بگیرد. حال اگر داده دو بعدی موجود باشد و معکوس ماتریس قبل در آن ضرب شود، ویژگی‌های سه بعدی آن تصویر در دسترس خواهد بود که این روند همان روند شناسایی اشیاست. استفاده از این ایده علاوه بر دقت مناسب، داده‌های لازم برای آموزش شبکه‌ها را بسیار کاهش خواهد داد. این راهکار تحت عنوان شبکه‌های عصبی کپسول ۲ (CapsNet) شناخته می‌شود [۳۰، ۳۱].

با توجه به موارد ذکر شده، برای شناسایی و طبقه‌بندی مؤثر سرهای غیرطبیعی اسپرم انسان به عنوان یکی از شاخص‌های مهم ناباروری که یک کلاس اقلیت به شمار می‌رود و به‌منظور تولید یک مدل طبقه‌بندی قوی نیاز به استخراج ویژگی‌های پیشرفته از تصاویر اسپرم انسان است. به همین منظور در این مطالعه، یک روش هوشمند و ترکیبی بر اساس شبکه‌های CapsNet و GAN برای طبقه‌بندی نامتوازن تصاویر اسپرم ارائه می‌شود. در واقع ابتدا از یک شبکه مولد متخاصم کپسولی به منظور تولید داده استفاده شد. سپس توسط یک شبکه کپسولی کانولوشنی، طبقه‌بندی نامتوازن تصاویر اسپرم انسان انجام شد. نوآوری‌ها و دستاوردهای این مقاله عبارتند از:

- ۱- طراحی یک مدل جدید مبتنی بر شبکه‌های CapsNet برای طبقه‌بندی متوازن تصاویر اسپرم انسان با دقت ۹۸/۱.
- ۲- طراحی یک مدل جدید براساس شبکه‌های مولد متخاصم برای تولید تصاویر مصنوعی اسپرم انسان.
- ۳- عملکرد مناسب مدل پیشنهادی در طبقه‌بندی نامتوازن تصاویر اسپرم انسان تا نسبت نامتوازنی ۱:۲۵ که حساسیت بالای ۸۰ درصد را حفظ می‌کند.
- ۴- مقاومت مناسب معماری پیشنهادی برای تولید تصاویر جدید در کاهش تعداد داده‌ها در مقایسه با سایر روش‌های معمول تولید داده‌های تصویری.

بستر گوشی هوشمند که آنالیز مورفولوژیک اسپرم را به طور کامل خودکار می‌کند، معرفی شد [۲۳]. در این مدل که توسط حمزا و همکارانش پیشنهاد شد، به منظور متعادل‌سازی داده‌ها از روش‌های مرسوم نمونه‌برداری مجدد نظیر مقیاس‌بندی، تغییر اندازه، جابجایی، برش استفاده شد. در [۲۴] شش مدل مختلف شبکه‌های عصبی کانولوشنی برای خودکارسازی طبقه‌بندی مورفولوژیکی تصاویر اسپرم معرفی شد. در این مطالعه از چهار عملیات اصلی بزرگنمایی، برش، جابجایی و قرینه‌سازی تصاویر برای نمونه‌برداری مجدد و متعادل‌سازی تصاویر استفاده شد و در نهایت دقت طبقه‌بندی ۹۰/۷۳ درصد را به دنبال داشت.

بررسی مطالعات پیشین در زمینه طبقه‌بندی سر اسپرم انسان نشان می‌دهد که الگوریتم‌های یادگیری عمیق نسبت به سایر الگوریتم‌ها به مراتب کارآمدتر و موثرتر عمل کرده‌اند. یادگیری عمیق زیرشاخه‌ای از یادگیری ماشین است که از مهم‌ترین مزایای آن میتوان به یادگیری خودکار و چند لایه ویژگی‌ها، دقت بالا، قدرت تعمیم بالا و پشتیبانی گسترده در حوزه‌های مختلف طبقه‌بندی تصاویر اشاره کرد [۲۵، ۲۶]. اکثر الگوریتم‌های یادگیری عمیق مانند شبکه‌های عصبی کانولوشنی<sup>۱</sup> (CNN)، به تعداد بسیار زیادی از تصاویر برای مدل‌سازی مناسب در مرحله آموزش نیاز دارند [۲۷-۲۹]. این ویژگی باعث می‌شود که این الگوریتم‌ها در طبقه‌بندی مجموعه نامتوازن تصاویر (مثلاً تشخیص اسپرم غیرطبیعی) عملکرد مناسبی نداشته باشند. همچنین وجود تعداد تصاویر بالا در مدل‌سازی الگوریتم‌های یادگیری عمیق، الزام وجود سیستم‌های سخت‌افزاری قوی را در پی دارد که همین نکته استفاده از آنها را با محدودیت‌هایی مواجه می‌کند. از طرفی، اگرچه می‌توان طبقه‌بندی سر اسپرم انسان را به دو دسته طبیعی (کلاس اکثریت) و غیرطبیعی (کلاس اقلیت) انجام داد ولی ناهنجاری‌های سر اسپرم انسان خود شامل چندین دسته از جمله سرهای کوچک، بی‌شکل، مخروطی و گلابی شکل می‌شوند. در حالت کلی به دلیل کمبود هریک از این نمونه‌ها در مقایسه با نمونه‌های سر اسپرم طبیعی، می‌توان هر کدام از آنها را در کلاس اقلیت جای داد. به بیان دیگر طبقه‌بندی سر اسپرم انسان یک مسئله طبقه‌بندی نامتوازن

<sup>2</sup> Capsule Neural Network (CapsNet)

<sup>1</sup> Convolutional Neural Network (CNN)

این مجموعه داده از نمونه منی ۱۵ بیمار در مرکز باروری و ناباروری اصفهان جمع‌آوری و رنگ‌آمیزی شده است [۳۲]. در این تصاویر، سرهای اسپرم توسط سه متخصص به ۴ دسته طبیعی<sup>۱</sup>، گلابی شکل<sup>۲</sup>، مخروطی<sup>۳</sup> و بی‌شکل<sup>۴</sup> طبقه‌بندی شده‌اند. این مجموعه داده از ۴ پوشه تصویر سر اسپرم با ابعاد ۱۳۱×۱۳۱ شامل ۵۴ سر اسپرم نرمال، ۵۳ سر گلابی شکل، ۵۷ سر مخروطی و ۵۲ سر بی‌شکل تشکیل شده است. شکل (۲) نمونه‌ای از تصاویر این مجموعه داده را نشان می‌دهد.



شکل ۲- نمونه‌ای از تصاویر مجموعه داده HuSHeM. (الف) طبیعی، (ب) گلابی شکل، (ج) مخروطی و (د) بی‌شکل

## ۲-۲- شبکه‌های مولد متخاصم

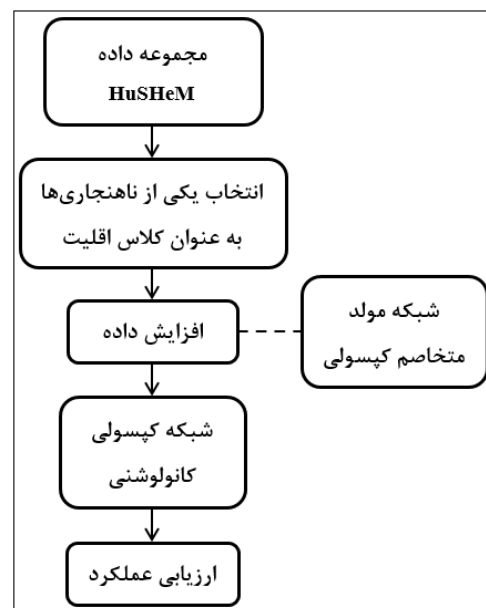
شبکه‌های مولد متخاصم<sup>۵</sup> (GAN) شامل دو شبکه مولد<sup>۶</sup> ( $G$ ) و متمایزگر<sup>۷</sup> ( $D$ ) است که در رقابت با یکدیگر هستند که نتیجه این رقابت تولید نمونه‌های مصنوعی از داده‌های واقعی است [۳۳]. ورودی شبکه  $G$  نویز تصادفی با توزیع گاوسی و خروجی آن داده مصنوعی است. در واقع این شبکه تبدیل  $G(z)$  از توزیع نویز  $(z \sim P_z)$  به توزیع داده  $(x \sim P_x)$  را بر عهده دارد. شبکه  $D$  نیز مشخص می‌کند که نمونه‌های ورودی به آن از توزیع داده‌های واقعی آموزشی  $(x \sim P_x)$  هستند یا از توزیع داده‌های مصنوعی تولید شده توسط شبکه  $G$   $(G(z) \sim P_G)$ . این دو مدل که معمولاً با شبکه‌های عصبی طراحی می‌شوند، یک بازی خصمانه مجموع صفر را با هدف ذکر شده در رابطه (۱)، انجام می‌دهند [۱۸]:

$$\min_G \max_D V(G,D) = E_{X \sim P_{data}(x)} [\log D(x)] + E_{Z \sim P_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

شبکه‌های GAN دو کاربرد عمده دارند: یکی تولید تصاویر جعلی جدید بر اساس تصاویر واقعی آموزش دیده و دیگری ترمیم تصویر است. شکل (۳) (الف) یک مدل نمونه از شبکه‌های GAN را نشان می‌دهد.

## ۲- روش پژوهش

الگوریتم پیشنهادی برای طبقه‌بندی نامتوازن سر اسپرم انسان در شکل (۱) ارائه شده است. در این مقاله از تصاویر موجود در مجموعه داده HuSHeM برای طراحی و ارزیابی مدل پیشنهادی استفاده شده است [۳۲]. این مجموعه داده شامل چهار کلاس است که سه تا از آنها، ناهنجاری‌های موجود در اسپرم انسان می‌باشد. به این ترتیب هر یک از این سه کلاس ناهنجاری می‌تواند به عنوان کلاس اقلیت و سایر کلاس‌ها به عنوان کلاس اکثریت انتخاب شوند. مطابق شکل (۱)، پس از انتخاب یک کلاس (از سه کلاس ناهنجاری موجود) به عنوان کلاس اقلیت، نوبت به افزایش داده می‌رسد. در این مقاله افزایش داده با استفاده از یک شبکه مولد متخاصم کیپسولی جدید انجام می‌شود. پس از افزایش داده، یک شبکه کیپسولی کانولوشنی جدید برای طبقه‌بندی نامتوازن تصاویر اسپرم انسان آموزش می‌بیند. در نهایت ارزیابی عملکرد مدل پیشنهادی توسط معیارهای مختلف صورت می‌پذیرد.



شکل ۱- الگوریتم پیشنهادی برای طبقه‌بندی نامتوازن تصاویر اسپرم انسان

## ۲-۱- مجموعه داده

الگوریتم پیشنهادی برای طبقه‌بندی سرهای اسپرم انسان روی مجموعه داده HuSHeM مورد آزمایش قرار گرفت.

<sup>5</sup> Generative Aversarial Network

<sup>6</sup> Generator

<sup>7</sup> Discriminator

<sup>1</sup> Normal

<sup>2</sup> Pyriform

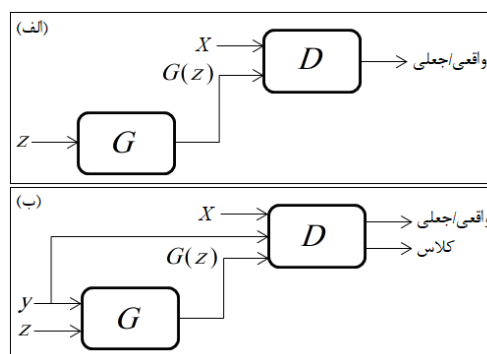
<sup>3</sup> Tapered

<sup>4</sup> Amorphous

## ۲-۳- شبکه‌های مولد متخاصم شرطی

چالش شبکه‌های GAN عدم کنترل روی فرآیند تولید تصاویر است. شبکه‌های مولد متخاصم شرطی<sup>۱</sup> (CGAN) نوعی از شبکه‌های GAN است که به مدل اجازه می‌دهد برای بهبود کیفیت نمونه‌های تولیدی از اطلاعات خارجی استفاده کند [۳۴]. در این شبکه‌ها بخش  $G$  و  $D$  بر اساس برچسب کلاس تصاویر مشروط می‌شوند. در واقع در شبکه‌های CGAN یادگیری نظارت نشده وجود دارد. شبکه‌های مولد متخاصم با طبقه‌بند کمکی<sup>۲</sup> (ACGAN) نوعی از شبکه‌های CGAN هستند که شبکه  $D$  را برای پیش‌بینی برچسب کلاس یک تصویر به جای دریافت آن به عنوان ورودی تغییر می‌دهد [۳۵]. این تغییر، فرآیند آموزش را تثبیت و امکان تولید تصاویر با کیفیت بالا را مستقل از برچسب کلاس فراهم می‌کند. در شکل (۳) (ب) شبکه ACGAN به عنوان نمونه‌ای از شبکه‌های CGAN نشان داده شده است. تابع هدف در شبکه‌های CGAN به صورت رابطه (۲) می‌باشد که  $y$  بردار اطلاعات خارجی مثل بردار کلاس تصاویر است [۳۴]:

$$\min_G \max_D V(G,D) = E_{X \sim P_{\text{data}}(x)} [\log D(x|y)] + E_{Z \sim P_z(z)} [\log(1 - D(G(z|y)))] \quad (2)$$



شکل ۳- (الف) شبکه GAN، (ب) شبکه ACGAN

## ۲-۴- شبکه عصبی کپسول (CapsNet)

شبکه CapsNet یک سیستم یادگیری ماشین است که می‌تواند برای مدل کردن بهتر روابط سلسله مراتبی مورد استفاده قرار گیرد [۳۰]. کار اصلی این شبکه‌ها، اضافه کردن ساختارهایی به نام کپسول به شبکه‌های CNN و استفاده مجدد خروجی آن برای کپسول‌های مرتبه بالاتر برای

رسیدن به فرم باثبات‌تر است [۳۶]. شبکه‌های CapsNet دارای دو بخش هستند: بخش رمزگذار و بخش رمزگشا. بخش رمزگذار معمولاً از سه لایه کانولوشن، کپسول‌های اولیه و کپسول‌های کلاس‌ها تشکیل می‌شود. بخش رمزگذار، تصاویر را به عنوان ورودی می‌گیرد و طی آموزش یاد می‌گیرد که آن‌ها را درون بردارهایی از پارامترها رمزگذاری کند. بخش رمزگشا که از لایه‌های تماماً متصل تشکیل شده است، یک بردار را از کپسول کلاس‌ها می‌گیرد و می‌آموزد که آن را به صورت یک تصویر رمزگشایی کند. در واقع بخش رمزگشا به عنوان تنظیم‌کننده مورد استفاده قرار می‌گیرد. به طوری که خروجی صحیح کپسول کلاس‌ها را به عنوان ورودی دریافت می‌کند و یاد می‌گیرد که یک تصویر را متناظر با آن بازسازی کند.

مقادیر خروجی در شبکه‌های CapsNet به صورت یک بردار ذخیره می‌شود [۳۱]؛ ویژگی بردار خواص مهمی از جمله اندازه و جهت دارد. به طوری که وقتی ویژگی‌ها در هر جای تصویر تغییر می‌کنند، اندازه بردار آنها حفظ می‌شود ولی جهت آنها تغییر می‌کند. در حالت کلی یک کپسول از چند بخش مهم تشکیل می‌گردد: ضرب ماتریسی بردارهای ورودی با ضرایب وزنی، وزن‌دهی عددی به بردارهای ایجادشده در مرحله قبل، جمع وزن‌دار بردارها و غیرخطی‌سازی برداری. تعامل بین کپسول‌های لایه‌های پایین و بالا توسط الگوریتم مسیریابی پویا انجام می‌شود که یک روش جدید برای تعیین محل خروجی هر کپسول است [۳۱]. این مکانیسم جایگزین ادغام ماکزیمم در شبکه‌های CNN است که باعث حفظ ویژگی‌های سطح بالای تصویر می‌شود. سپس یک تابع فعال‌ساز غیرخطی که یک بردار را دریافت و با حفظ جهت و ماکزیمم طول آن را لهیده می‌کند، استفاده می‌گردد. تابع هزینه نهایی  $L_c$  که برای لایه کپسولی آخر مورد استفاده قرار می‌گیرد، عبارت است از [۳۱]:

$$L_c = T_c \cdot \max(0, m^+ - \|V_c\|)^2 + \lambda(1 - T_c) \cdot \max(0, \|V_c\| - m^-)^2 \quad (3)$$

$T_c = 1$  نشان‌دهنده صحیح بودن برچسب کلاس  $c$  است.  $\lambda$ ،  $m^+$  و  $m^-$  پارامترهای ثابت و  $V_c$  خروجی لایه نهایی است.

<sup>2</sup> Auxiliary Classifier GAN<sup>1</sup> Conditional GAN

### ۳- مدل پیشنهادی

مدل پیشنهادی این مقاله برای تولید داده در شکل (۴) مشاهده می‌گردد. این مدل که بر اساس شبکه‌های مولد متخاصم کپسولی شکل گرفته است، از دو بخش مولد (شبکه  $G$ ) و متمایزگر (شبکه  $D$ ) تشکیل می‌شود. تصاویر واقعی ( $x$ ) به همراه کلاس‌های متناظر آنها ( $y$ ) به هر دو شبکه مولد و متمایزگر وارد می‌شود. هم‌چنین برای تولید تصاویر جعلی توسط شبکه مولد، لازم است بردار نویز تصادفی ( $z$ ) نیز به آن داده شود. در شبکه مولد لازم است ابتدا برای تصاویر واقعی و کلاس‌های متناظر با آنها، عملیات گنجاندن ۱ یا تعبیه‌سازی انجام شود. معمولاً هر تعبیه، نگاشتی از یک بردار در فضای با ابعاد بالا به فضایی با ابعاد کمتر است. حاصل عملیات گنجاندن تصاویر واقعی و کلاس‌های متناظر باید با بردار تصادفی  $z$ ، الحاق ۲ یا ترکیب شود. سپس خروجی این الحاق، توسط عملیات دکانولوشن طی ۴ مرحله به ۳ ماتریس با ابعاد  $۱۳۱ \times ۱۳۱$  تبدیل می‌شوند که همان تصاویر اسپرم رنگی و جعلی تولیدی است.

تصاویر جعلی تولیدی از بخش مولد ( $G(z|y)$ ) به همراه تصاویر واقعی ( $x$ ) و کلاس‌های متناظرشان ( $y$ )، ورودی‌های بخش متمایزگر مدل پیشنهادی هستند. مطابق شکل (۴) بخش متمایزگر از دو لایه کانولوشن به ترتیب با فیلترهای با ابعاد  $۱۶ \times ۱۶$  و  $۳۲ \times ۳۲$  تشکیل شده است. پس از این دو لایه کانولوشنی لایه کپسول‌های اولیه که شامل ۳۲ کپسول با ۸ فیلتر  $۳ \times ۳$  است، وجود دارد. این لایه در طبقه‌بند اولیه و ثانویه مشترک است. طبقه‌بند اولیه تصویر را از لحاظ واقعی/جعلی بودن طبقه‌بندی و طبقه‌بند ثانویه محتمل‌ترین کلاس تصویر را پیش‌بینی می‌کند. از آنجا که مدل پیشنهادی مبتنی بر شبکه‌های مولد متخاصم و شبکه‌های کپسول است، تابع هدف ترکیبی از روابط (۲) و (۳) و به صورت رابطه (۴) خواهد بود:

$$\min_G \max_D V(G,D) = -E_{X \sim P_{data}(x)} [L_C(D(x|y))] - E_{Z \sim P_z(z)} [L_C(D(G(z|y)))] \quad (4)$$

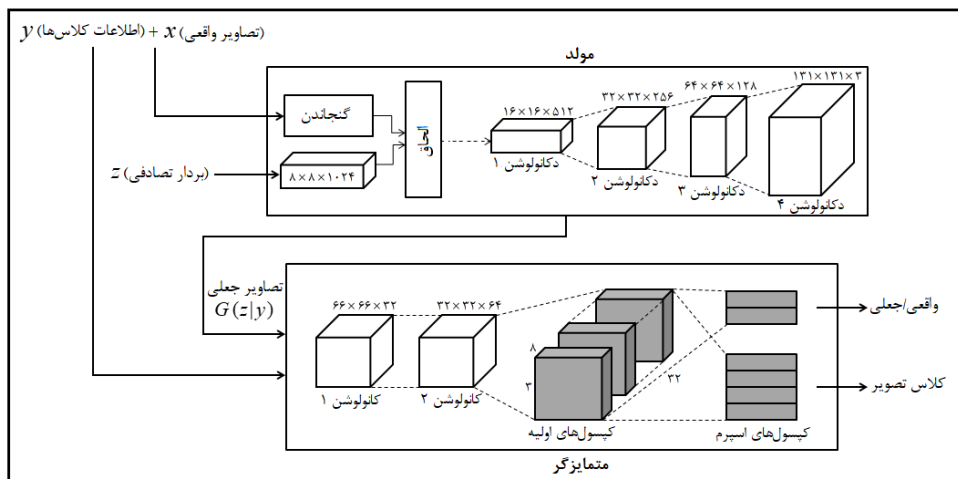
در این رابطه  $x$  تصاویر واقعی،  $y$  کلاس متناظر با تصاویر واقعی و  $z$  بردار نویز تصادفی است.  $x$  می‌تواند هر یک از تصاویر واقعی اسپرم طبیعی، گلابی‌شکل، مخروطی و یا

بی‌شکل باشد. شکل (۵) مدل پیشنهادی این مقاله برای طبقه‌بندی نامتوازن تصاویر اسپرم انسان است. این مدل بر اساس شبکه‌های عصبی کپسول و کانولوشن شکل گرفته است. این مدل از دو قسمت رمزگذار و رمزگشا تشکیل شده است. قسمت رمزگذار از سه لایه کانولوشن، لایه کپسول‌های اولیه و لایه کپسول‌های اسپرم تشکیل شده است. سه لایه تماماً متصل که از خروجی لایه کپسول‌های اسپرم تغذیه می‌شوند، قسمت رمزگشا را تشکیل می‌دهند. تصاویر اسپرم انسان با ابعاد  $۱۳۱ \times ۱۳۱$  به عنوان ورودی وارد بخش رمزگذار می‌شوند و این قسمت طی آموزش یاد می‌گیرد که آن‌ها را درون بردارهای ۱۶ بعدی از پارامترها رمزگذاری کند. اینجا جایی است که کپسول‌ها کار خود را انجام می‌دهند. خروجی شبکه در طول پیش‌بینی، بردارهای ۱۶ بعدی از طول خروجی‌های کپسول‌های لایه آخر (لایه کپسول‌های اسپرم) است.

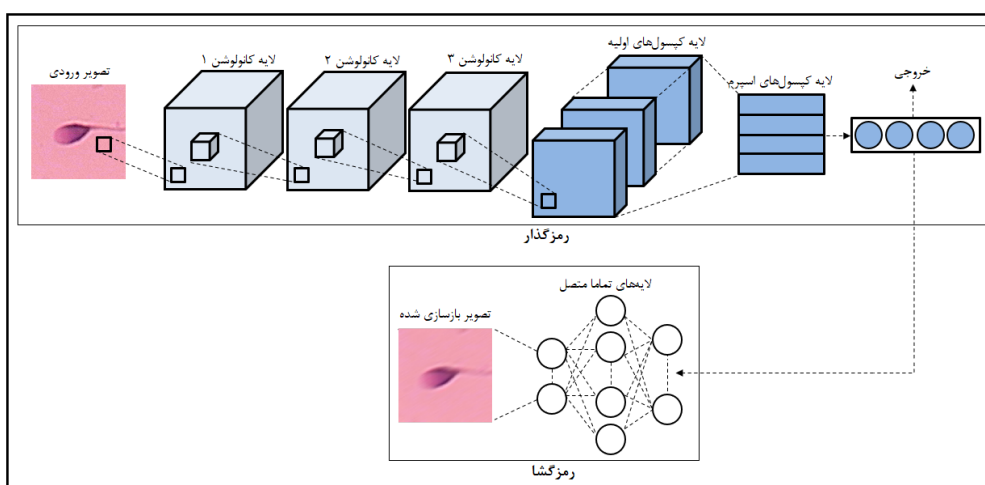
در جدول ۱ جزئیات لایه‌های بخش رمزگذار و رمزگشای مدل پیشنهادی برای طبقه‌بندی نامتوازن ثبت شده است. بر اساس این جدول، لایه‌های کانولوشن اول، دوم و سوم به ترتیب از ۳۲ فیلتر  $۵ \times ۵$  با گام ۲، ۶۴ فیلتر  $۳ \times ۳$  با گام ۲ و ۶۴ فیلتر  $۳ \times ۳$  با گام ۱ (همگی با تابع فعال‌ساز  $ReLU$ ) تشکیل شده است. هم‌چنین لایه کپسول‌های اولیه دارای ۳۲ کپسول است که وظیفه آن دریافت ویژگی‌های اساسی‌ست که توسط لایه‌های کانولوشن قبلی کشف شده و ایجاد ترکیبی از این ویژگی‌هاست. هر کپسول در این لایه از ۸ فیلتر با ابعاد  $۳ \times ۳$  و گام ۲ تشکیل می‌شود. از آنجا که ورودی این لایه نگاشت‌های ویژگی با ابعاد  $۳۲ \times ۳۲ \times ۶۴$  است، نگاشت‌های ویژگی با ابعاد  $۱۶ \times ۱۶ \times ۸$  خروجی هر کپسول را تشکیل می‌دهند. لایه آخر (لایه کپسول‌های اسپرم) دارای ۴ عدد کپسول است که برای هر دسته از تصاویر اسپرم (طبیعی، گلابی‌شکل، مخروطی و بی‌شکل) یک کپسول در نظر گرفته شده است. هر کپسول به عنوان ورودی یک نگاشت ویژگی با ابعاد  $۱۶ \times ۱۶ \times ۸$  را در نظر می‌گیرد. در واقع ورودی هر کپسول در این لایه تعداد  $۱۶ \times ۱۶ \times ۳۲ = ۸۱۹۲$  بردار ۸ بعدی است. طبق اساس کار داخلی کپسول، هر یک از این بردارهای ورودی، ماتریس وزنی با ابعاد  $۸ \times ۱۶$  خود را دارند که فضای ورودی ۸ بعدی را به فضای خروجی کپسول ۱۶ بعدی نگاشت می‌کند.

<sup>2</sup> Concatenation

<sup>1</sup> Embedding



شکل ۴- مدل شبکه مولد متخاصم کپسولی برای تولید داده



شکل ۵- مدل شبکه کپسولی کانولوشنی برای طبقه‌بندی نامتوازن

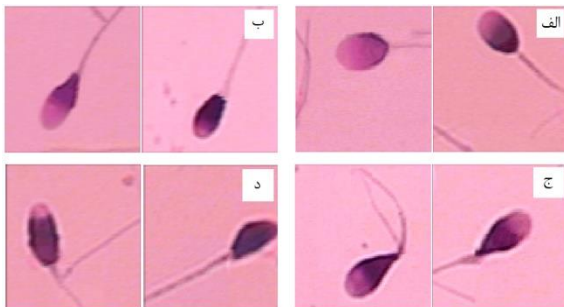
جدول ۱- جزئیات لایه‌های قسمت رمزگذار و رمزگشا

بخش	نام لایه	ورودی لایه	تعداد (فیلتر/کپسول/نورون)	خروجی لایه
رمز	کانولوشن ۱	تصویر ۱۳۱×۱۳۱	۳۲ فیلتر ۵×۵ با گام ۲	نگاشت‌های ویژگی ۶۵×۶۵×۳۲
	کانولوشن ۲	نگاشت‌های ویژگی ۶۵×۶۵×۳۲	۶۴ فیلتر ۳×۳ با گام ۲	نگاشت‌های ویژگی ۳۳×۳۳×۶۴
	کانولوشن ۳	نگاشت‌های ویژگی ۳۳×۳۳×۶۴	۶۴ فیلتر ۳×۳ با گام ۱	نگاشت‌های ویژگی ۳۲×۳۲×۶۴
گذار	کپسول‌های اولیه	نگاشت‌های ویژگی ۳۲×۳۲×۶۴	۳۲ کپسول دارای ۸ فیلتر ۳×۳ با گام ۲	نگاشت‌های ویژگی ۱۶×۱۶×۸×۳۲
	کپسول‌های اسپرم	نگاشت‌های ویژگی ۱۶×۱۶×۸×۳۲	۴ کپسول ۱۶ بعدی	ماتریس‌های ۴×۱۶
رمزگشا	تماما متصل ۱	ماتریس ۴×۱۶	۵۱۲ نورون	۵۱۲ نورون
	تماما متصل ۲	۵۱۲ نورون	۱۰۲۴ نورون	۱۰۲۴ نورون
	تماما متصل ۳	۱۰۲۴ نورون	۱۷۱۶۱ نورون (۱۳۱×۱۳۱)	۱۷۱۶۱ نورون



#### ۴-۱- تولید تصاویر جدید و دسته‌بندی آنها

بررسی مدل پیشنهادی به منظور طبقه‌بندی تصاویر اسپرم انسان به دو صورت متوازن و نامتوازن انجام شد. با توجه به تعداد تصاویر کم در مجموعه داده HuSHeM نیاز بود ابتدا تصاویر جعلی جدیدی به مجموعه داده اضافه شود. تصاویر جعلی توسط مدل مولد متخاصم کپسولی معرفی شده در شکل (۴) تولید شدند. تعداد این تصاویر تولیدی به گونه‌ای بود که تعداد تصاویر در هر کلاس به ۳۰۰۰ تصویر رسید. شکل (۶) نمونه‌ای از این تصاویر تولیدی را نشان می‌دهد.



شکل ۶- نمونه‌ای از تصاویر تولیدی توسط مدل مولد متخاصم کپسولی، (الف) اسپرم نرمال، (ب) اسپرم گلابی شکل، (ج) اسپرم مخروطی و (د) اسپرم بی‌شکل

جدول ۲ جزئیات تعداد تصاویر موجود در هر مرحله در طبقه‌بندی متوازن و نامتوازن تصاویر اسپرم انسان را نشان می‌دهد. طبق این جدول در حالت متوازن (نسبت نامتوازنی کلاس اقلیت به اکثریت ۱ به ۱)، در هر کلاس ۳۰۰۰ تصویر اسپرم موجود است. اما در حالت طبقه‌بندی نامتوازن تغییرات نسبت نامتوازنی از ۱:۵ تا ۱:۳۰ (طی شش گام)، تعداد تصاویر موجود در کلاس اقلیت تغییر می‌کند. تا جایی که تعداد تصاویر موجود در کلاس اقلیت به ۱۰۰ تصویر (در نسبت نامتوازنی ۱:۳۰) می‌رسد.

#### ۴-۲- طبقه‌بندی متوازن تصاویر اسپرم انسان

پس از تولید تصاویر جعلی و به منظور ارزیابی حالت طبقه‌بندی متوازن تصاویر اسپرم، ۸۰ درصد تصاویر برای آموزش و ۲۰ درصد برای آزمایش به صورت تصادفی انتخاب شدند. البته تا جای ممکن سعی شد تصاویر آزمایش از بین تصاویر واقعی انتخاب شوند. اعمال تصاویر آزمایش به مدل کپسولی کانولوشنی پیشنهادی آموزش دیده صحت ۹۸/۱ درصدی را به دنبال داشت. همان‌طور که در جدول ۳ مشاهده می‌شود مدل پیشنهادی این مقاله در طبقه‌بندی متوازن تصاویر اسپرم موجود در مجموعه داده HuSHeM

خروجی لایه کپسول‌های اسپرم، ۴ بردار ۱۶ بعدی است. در طول پروسه آموزش برای هر نمونه آموزشی یک مقدار هزینه برای هر یک از این ۴ بردار طبق رابطه (۳) محاسبه می‌شود. در رابطه مذکور  $V_c (c=1,2,3,4)$  هر یک از بردارهای خروجی است. برای محاسبه هزینه نهایی هر نمونه، ۴ مقدار محاسبه شده با هم جمع می‌شوند. از آنجا که در اینجا یادگیری نظارت شده وجود دارد، هر نمونه در مجموعه داده آموزشی دارای برچسب صحیح است که در این حالت یک بردار چهار بعدی با سه عنصر ۰ و یک عنصر با مقدار ۱ (موقعیت صحیح) وجود دارد. این شیوه به کدبندی One-Hot معروف است [۳۷]. در فرمول تابع هزینه، برچسب صحیح، مقدار  $T_c$  را مشخص می‌کند: این مقدار چنانچه برچسب صحیح به کپسول اسپرم خاص مربوط باشد برابر ۱ خواهد بود و در غیر آن صورت برابر صفر است. در لایه آخر (لایه کپسول‌های اسپرم) کپسول‌های اول، دوم، سوم و چهارم به ترتیب مسئول رمزگذاری تصاویر اسپرم طبیعی، تصاویر اسپرم گلابی شکل، تصاویر اسپرم مخروطی و تصاویر اسپرم بی‌شکل هستند. اگر فرض شود که برچسب صحیح برابر کلاس تصاویر اسپرم طبیعی باشد، این نشان می‌دهد که کپسول نخست در لایه آخر مسئول رمزگذاری آن خواهد بود. برای این کپسول اسپرم، مقدار  $T_c$  برابر یک بوده و برای سه کپسول باقی مانده در این لایه،  $T_c$  برابر صفر خواهد بود. وقتی  $T_c$  برابر یک است، بخش اول تابع هزینه محاسبه می‌شود و بخش دوم برابر صفر خواهد بود. در شبکه پیشنهادی این مقاله  $m^+ = 0/9$ ،  $m^- = 0/1$  و  $\lambda = 0/5$  است. قسمت رمزگشا نیز مطابق شکل (۵) یک بردار ۱۶ بعدی را از کپسول اسپرم صحیح در لایه آخر بخش رمزگذار می‌گیرد و می‌آموزد که آن را به صورت یک تصویر اسپرم رمزگشایی کند. طبق جدول ۱ خروجی این لایه پس از تغییر شکل، تصویر اسپرم رمزگشایی شده با ابعاد  $131 \times 131$  را به خروجی تحویل می‌دهد.

#### ۴- نتایج شبیه‌سازی

در این قسمت نتایج و معیارهای ارزیابی مدل پیشنهادی این مقاله ارائه می‌گردد. برای این منظور ابتدا تصاویر جدید تولیدی توسط مدل مولد متخاصم کپسولی ارائه و سپس طبقه‌بندی تصاویر اسپرم انسان به صورت متوازن و نامتوازن به همراه معیارهای ارزیابی مختلف، ارائه و بررسی می‌شود.

کلاس مخروطی و در حالت سوم کلاس بی‌شکل به عنوان کلاس اقلیت در نظر گرفته شدند. همان‌طور که از جدول ۳ پیداست تا نسبت نامتوازنی ۱:۳۰ صحت کل کلاس‌ها بالای ۸۰ درصد حفظ شده است. اما صحت کلاس‌های اقلیت تا نسبت نامتوازنی ۱:۲۵ بالای این مقدار حفظ شده است.

برای نشان دادن عملکرد و ارزیابی مدل پیشنهادی در طبقه‌بندی نامتوازن تصاویر اسپرم علاوه بر معیار صحت، از معیارهای دقت<sup>۲</sup> و حساسیت<sup>۳</sup> نیز استفاده شد. در حالت کلی صحت به این معناست که مدل آموزش دیده تا چه اندازه، خروجی را درست پیش‌بینی می‌کند. دقت نشان‌دهنده میزان اعتماد به خروجی طبقه‌بند و حساسیت، معیاری است که مشخص می‌کند طبقه‌بند، به چه اندازه در تشخیص ناهنجاری یا همان کلاس اقلیت موفق بوده است. این معیارها به ترتیب از روابط (۵) الی (۷) محاسبه می‌شوند:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$Pr = \frac{TP}{TP + FP} \quad (6)$$

$$Sen = \frac{TP}{TP + FN} \quad (7)$$

به‌طوری‌که  $TP$  بیان‌گر مثبت واقعی،  $TN$  نشان‌دهنده منفی واقعی،  $FP$  و  $FN$  به‌ترتیب نشان‌دهنده مثبت کاذب و منفی کاذب است.  $Acc$ ،  $Pr$  و  $Sen$  نیز به ترتیب صحت، دقت و حساسیت را نشان می‌دهند. جدول ۵ میانگین معیارهای صحت، دقت و حساسیت را برای طبقه‌بندی نامتوازن تصاویر اسپرم تا نسبت نامتوازنی ۱:۲۵ نشان می‌دهد. همان‌طور که از این جدول دریافت می‌شود طی تغییرات نسبت نامتوازنی از ۱:۵ تا ۱:۲۵ همه معیارهای عملکرد بالای ۸۵ درصد حفظ شده است که نشان از عملکرد مناسب شبکه کپسولی کانولوشنی پیشنهادی این مقاله دارد. از طرفی معیارهای صحت و دقت بالای ۸۸ درصد حفظ شده است. هم‌چنین جدول ۶ نمونه تصاویری را با پیش‌بینی درست و غلط توسط شبکه کپسولی پیشنهادی نشان می‌دهد.

درصد صحت بالاتری در مقایسه با سایر الگوریتم‌های یادگیری عمیق در چند سال اخیر داشته است که از توانایی‌های مدل پیشنهادی به شمار می‌رود.

جدول ۲- تعداد تصاویر موجود در هر مرحله در طبقه‌بندی نامتوازن تصاویر اسپرم انسان

کلاس	اقلیت (گلابی شکل، یا مخروطی و یا بی‌شکل)	اکثریت
نسبت نامتوازنی		
۱:۱ (حالت متوازن)	۳۰۰۰	۳۰۰۰
۱:۵	۶۰۰	۳۰۰۰
۱:۱۰	۳۰۰	۳۰۰۰
۱:۱۵	۲۰۰	۳۰۰۰
۱:۲۰	۱۵۰	۳۰۰۰
۱:۲۵	۱۲۰	۳۰۰۰
۱:۳۰	۱۰۰	۳۰۰۰

جدول ۳- مقایسه روش‌های طبقه‌بندی متوازن اسپرم انسان در

مجموعه داده HuSHeM

روش	سال	صحت متوسط
VGG16	۲۰۱۹	۹۴
[۲۰]CNN	۲۰۱۹	۹۲/۲
[۴]CNN	۲۰۲۰	۹۵
[۲۴]CNN	۲۰۲۱	۸۵/۱۸
مدل پیشنهادی	۲۰۲۲	۹۸/۱

#### ۴-۳- طبقه‌بندی نامتوازن تصاویر اسپرم انسان

با توجه به اینکه شناسایی و طبقه‌بندی ناهنجاری‌های موجود در تصاویر سر اسپرم انسان (سرهای گلابی شکل، مخروطی و بی‌شکل) از اهمیت بیشتری نسبت به تصاویر سر اسپرم نرمال برخوردار است و معمولاً در بررسی ناباروری مردان یک ناهنجاری و اتفاق نادر به شمار می‌رود، با نامتوازن کردن تصادفی هر یک از این کلاس‌ها در سه مرحله کارایی مدل پیشنهادی بررسی شد. این نامتوازن کردن کلاس‌ها توسط جدول ۲ صورت گرفت و نتایج درصد صحت ۱ در هر نسبت نامتوازنی در جدول ۴ ثبت شد. مطابق این جدول، طبقه‌بندی نامتوازن کلاس اقلیت در سه حالت بررسی شد. در حالت اول کلاس گلابی شکل، در حالت دوم

<sup>3</sup> Sensitivity

<sup>1</sup> Accuracy

<sup>2</sup> Precision


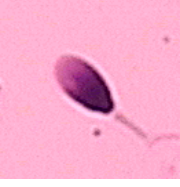
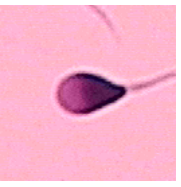
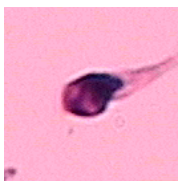


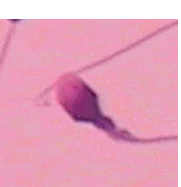

جدول ۴- تغییرات درصد صحت در نسبت‌های نامتوازن مختلف با کلاس‌های اقلیت متفاوت

نسبت نامتوازنی		حالت				
۱:۳۰	۱:۲۵	۱:۲۰	۱:۱۵	۱:۱۰	۱:۵	
۷۵/۶۸	۸۰/۶۵	۸۴/۷۲	۸۷/۳۹	۹۱/۶۳	۹۴/۴۵	کلاس اقلیت: گلابی شکل
۸۰/۱۱	۸۳/۹۰	۸۷/۱۹	۹۱/۰۴	۹۳/۷۹	۹۶/۸۷	کل کلاس‌ها
۷۵/۵۳	۸۰/۱۳	۸۴/۰۱	۸۶/۷۴	۹۰/۹۷	۹۳/۸۹	کلاس اقلیت: مخروطی
۸۰/۰۲	۸۳/۸۶	۸۷/۰۳	۹۰/۷۸	۹۳/۶۱	۹۶/۳۹	کل کلاس‌ها
۷۶/۲۲	۸۱/۲۵	۸۵/۲۵	۸۷/۸۱	۹۲/۰۹	۹۴/۶۷	کلاس اقلیت: بی‌شکل
۸۰/۷۲	۸۴/۳۴	۸۷/۷۶	۹۱/۳۷	۹۴/۴۶	۹۷/۱۴	کل کلاس‌ها

جدول ۵- میانگین معیارهای عملکرد در طبقه‌بندی نامتوازن تصاویر اسپرم انسان

معیار		حالت	
صحت	دقت	حساسیت	
۸۸/۸۶	۸۹/۹۲	۸۵/۷۵	کلاس اقلیت: گلابی شکل
۸۸/۲۱	۸۹/۶۵	۸۵/۲۱	کلاس اقلیت: مخروطی
۸۹/۰۷	۹۰/۳۷	۸۶/۲۲	کلاس اقلیت: بی‌شکل

جدول ۶- نمونه‌ای از تصاویر با پیش‌بینی درست و غلط توسط شبکه کپسولی ترکیبی

کلاس واقعی		پیش‌بینی		
اسپرم نرمال	اسپرم گلابی شکل	اسپرم مخروطی	اسپرم بی‌شکل	
				درست
(پیش‌بینی شده: نرمال)	(پیش‌بینی شده: گلابی شکل)	(پیش‌بینی شده: مخروطی)	(پیش‌بینی شده: بی‌شکل)	
				غلط
(پیش‌بینی شده: مخروطی)	(پیش‌بینی شده: بی‌شکل)	(پیش‌بینی شده: گلابی شکل)	(پیش‌بینی شده: گلابی شکل)	

## ۵- بحث

در این مقاله دو مدل جدید، یکی برای تولید تصاویر جدید و دیگری برای طبقه‌بندی نامتوازن تصاویر اسپرم انسان ارائه شد. بررسی درصد صحت در طبقه‌بندی متوازن و هم‌چنین





















معیارهای کمی در طبقه‌بندی نامتوازن نشان داد که مدل کپسولی کانولوشنی پیشنهادی می‌تواند در حالت نامتوازن تا نسبت نامتوازنی ۱:۲۵ صحت بالای ۸۸ درصد و حساسیت بالای ۸۰ درصد را حفظ کند. لازم به ذکر است

با تولید تصاویر مصنوعی جدید از تصاویر اصلی کمک می‌کند تا تصاویر جدیدی برای آموزش وجود داشته باشد [۳۸]. از جمله روش‌های افزایش داده‌های تصویری می‌توان به قرینه‌سازی<sup>۱</sup>، چرخش<sup>۲</sup>، مقیاس<sup>۳</sup>، برش<sup>۴</sup>، جابه‌جایی<sup>۵</sup> و نویز گوسی<sup>۶</sup> اشاره کرد.

به منظور بررسی و مقایسه روش‌های معمول افزایش داده‌های تصویری با مدل مولد متخاصم کپسولی این مقاله، ابتدا برای ۲۱۶ تصویر اسپرم موجود در مجموعه داده HuSheM، عملیات قرینه‌سازی، برش، جابجایی و چرخش انجام شد. در عملیات چرخش از زوایای مختلفی از جمله ۹۰، ۱۸۰ و ۲۷۰ درجه استفاده شد. جدول ۷ نمونه‌ای از تصاویر ایجاد شده توسط این روش‌ها را نشان می‌دهد. تعداد تصاویر تولیدی به گونه‌ای انجام شد که در هر کلاس ۳۰۰۰ تصویر موجود باشد. سپس طبقه‌بندی متوازن و نامتوازن چهار کلاس اسپرم طبیعی، گلابی شکل، مخروطی و بی‌شکل بررسی شد. هم‌چنین تاثیر کاهش تعداد داده در این دو نوع طبقه‌بندی نیز بررسی شد.

که اغلب الگوریتم‌های یادگیری عمیق برای آموزش، نیاز به حجم بالایی از داده‌های برچسب خورده دارند و فقط نیاز به زمان دارند که بتوانند بهترین ویژگی‌ها را با کمترین خطا بیاموزند. اما شبکه‌های عصبی کپسول با در نظر گرفتن بردار ویژگی (به جای مقدار ویژگی در مقایسه با شبکه‌های کانولوشنی) در عمل باعث می‌شوند به تعداد داده‌های کمتری برای آموزش نیاز باشد. با این وجود گاهی مجموعه داده مورد استفاده به اندازه‌ای کوچک است که امکان استفاده از آن تنها در صورت تولید تصاویر جدید امکان‌پذیر است. به عنوان مثال تصاویر مجموعه داده اسپرم انسان HuSheM در مجموع شامل ۲۱۶ تصویر در چهار کلاس مختلف است که برای آموزش یک شبکه عمیق بسیار ناکافی است و به همین دلیل این مجموعه داده نیاز به افزایش تعداد داده دارد. در این مقاله توسط مدل مولد متخاصم کپسولی تولید تصاویر اسپرم جدید صورت پذیرفت. علاوه بر تنوع روش‌های افزایش داده، کیفیت داده‌های تولید شده نیز مهم است. روش‌های افزایش داده<sup>۱</sup>

جدول ۷- نمونه‌ای از عملیات‌های مختلف به منظور افزایش تعداد تصاویر موجود در مجموعه داده HuSheM

عملیات	اسپرم نرمال	اسپرم گلابی شکل	اسپرم مخروطی	اسپرم بی‌شکل
تصویر اصلی				
چرخش ۹۰ درجه				
قرینه‌سازی				
برش				
جابه‌جایی				

<sup>5</sup> Cropping

<sup>6</sup> Translation

<sup>7</sup> Gaussian Noise

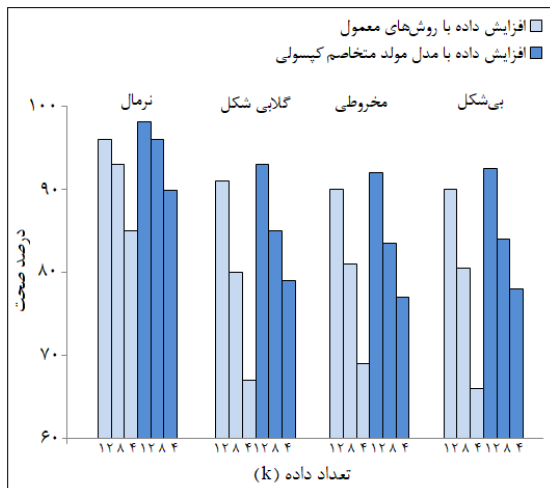
<sup>1</sup> Data augmentation

<sup>2</sup> Flipping

<sup>3</sup> Rotation

<sup>4</sup> Scale

پیشنهادی از جنس تصاویر اصلی با ویژگی‌های جدید است. اگر چه این ویژگی‌ها حاصل برهم‌کنش ویژگی‌های سایر تصاویر موجود است ولی در مقایسه با روش‌های معمول نظیر چرخش، برش و یا جابه‌جایی، تصاویر متفاوت‌تری تولید می‌کند. البته ذکر این نکته ضروری است که شبکه‌های مولد متخاصم نسبت به روش‌های معمول افزایش داده‌های تصویری، دارای محاسبات پیچیده‌تری است.

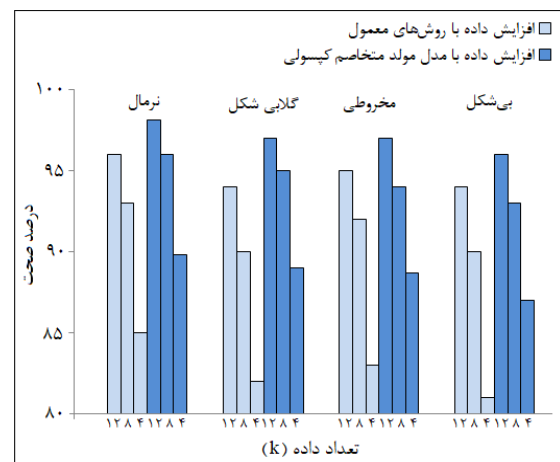


شکل ۸- بررسی اثر کاهش داده روی درصد صحت چهار کلاس برای طبقه‌بندی نامتوازن همراه با دو روش افزایش داده متفاوت

### ۶- نتیجه‌گیری

ناباروری یکی از مسائل مهم پزشکی است که عامل مردانه حدوداً یک سوم دلایل ناباروری در زوجین می‌باشد. بررسی نمونه‌های آزمایشگاهی اسپرم، اطلاعات اساسی و مفیدی در خصوص وضعیت بالینی بیمار برای درمان در اختیار پزشکان قرار می‌دهد و با تصمیم‌گیری درست پزشک، درمان می‌تواند بصورت موثرتر و با صرف هزینه‌های کمتر به باروری در زوجها منجر شود. ناهنجاری‌های اسپرم یکی از دلایل مهم ناباروری مردان است و اختلالات مورفولوژیکی اسپرم منشا اصلی این ناهنجاری‌ها شناخته شده است. بررسی تصاویر اسپرم بصورت دستی، علاوه بر بالا بردن احتمال خطا، باعث پایین آمدن سرعت بررسی نیز می‌شود. هم‌چنین وجود نمونه‌های کمتر مربوط به ناهنجاری‌های سر اسپرم در مقایسه با نمونه‌های اسپرم طبیعی، یکی از چالش‌های مهم بررسی مورفولوژی اسپرم به شمار می‌رود. استفاده از روش‌های کامپیوتری و مبتنی بر یادگیری عمیق می‌تواند با کاهش خطای تشخیص و طبقه‌بندی ناهنجاری‌ها، سرعت تصمیم‌گیری‌ها را نیز

شکل (۷) مقایسه درصد صحت برای طبقه‌بندی متوازن در چهار کلاس تصاویر اسپرم را نشان می‌دهد. همان‌گونه که در این شکل مشاهده می‌شود، افزایش داده با مدل مولد متخاصم کپسولی پیشنهادی در مقایسه با افزایش داده با روش‌های معمول، درصد صحت بالاتری را به دنبال داشت. هم‌چنین این شکل تاثیرات کاهش داده روی این دو روش در طبقه‌بندی متوازن را نشان می‌دهد. بدین منظور داده‌ها یکبار به ۸۰۰۰ تصویر و بار دیگر به ۴۰۰۰ تصویر کاهش پیدا کردند. طی این مراحل و مطابق شکل (۷) مشاهده می‌شود که داده‌های تولید شده با مدل پیشنهادی مقاومت بالاتری را از خود نشان می‌دهند.



شکل ۷- مقایسه اثر کاهش داده روی درصد صحت چهار کلاس برای طبقه‌بندی متوازن همراه با دو روش افزایش داده متفاوت

شکل (۸) مقایسه درصد صحت برای طبقه‌بندی نامتوازن در چهار کلاس تصاویر اسپرم را نشان می‌دهد. در این شکل برای بررسی هر کلاس (به جز کلاس نرمال) فرض شده است که آن کلاس، کلاس اقلیت است. مطابق با این شکل نیز درصد صحت در کلاس‌های نرمال و اقلیت گلابی‌شکل، مخروطی و بی‌شکل با افزایش داده مدل کپسولی کانولوشنی پیشنهادی به مراتب بالاتر از درصد صحت با افزایش داده توسط روش‌های معمول بود. اثر کاهش مجموعه داده در طبقه‌بندی نامتوازن نیز نشان از مقاومت مناسب داده‌های تولید شده توسط مدل پیشنهادی داشت. این موضوع تاییدی بر اهمیت ساختار شبکه‌های مولد متخاصم دارد که باعث می‌شود تصاویر تولیدی آن دارای تفاوت‌های بیشتر و بهتری در مقایسه با روش‌های معمول افزایش داده باشد. به بیان دیگر تصاویر تولیدی توسط مدل

مدل پیشنهادی تا نسبت نامتوازنی کلاس اقلیت به اکثریت ۱:۲۵ حساسیت بالای ۸۰ درصد را حفظ کرد. در بررسی کاهش داده‌ها و عدم توازن بین کلاس‌ها، عمل کاهش داده و نامتوازن‌سازی کلاس‌ها به صورت تصادفی صورت گرفت. در صورتی که با توجه به ساختارهای شبکه‌های کپسولی، میتوان با حذف هدفمند تصاویر، مقاومت این شبکه‌ها را نسبت به کاهش یا نامتوازنی داده‌ها بالا برد. هم‌چنین می‌توان در آینده با تغییراتی نظیر اضافه کردن توابع بهینه‌سازی، افزایش داده در دسته اقلیت، کاهش داده در دسته‌های اکثریت و یا ترکیبی از این موارد به مفاهیم شبکه‌های کپسولی در مسائل طبقه‌بندی نامتوازن تصاویر مطالعات بیشتری انجام داد.

افزایش دهد. در این مقاله، ما یک مدل یادگیری عمیق مبتنی بر شبکه عصبی کپسول ایجاد کردیم که طبقه‌بندی نامتوازن تصاویر اسپرم انسان را به صورت خودکار انجام می‌دهد. به این منظور و به دلیل کمبود مجموعه داده، ما مدلی بر اساس شبکه‌های مولد متخاصم ارائه دادیم تا تصاویر مصنوعی اسپرم تولید کند. بررسی‌ها نشان داد که این تصاویر تولیدی از جنس مجموعه داده با ویژگی‌های متمایز است. چرا که مقاومت بالاتری در مقایسه با سایر روش‌های معمول افزایش داده تصویری در مورد کاهش داده‌ها داشت. عملکرد طبقه‌بندی متوازن مدل پیشنهادی برای مجموعه داده HuSHeM، با صحت ۹۸/۱ درصد به دست آمد. هم‌چنین

## ۷- مراجع

[1] F. E. Okonofua et al., "Causes and risk factors for Male infertility: A scoping review of published studies," *International Journal of General Medicine*, 2022, pp. 5985-5997.

[2] S. Cilio et al., "Beneficial Effects of Antioxidants in Male Infertility Management: A Narrative Review," *Oxygen*, vol. 2, no. 1, 2022, pp. 1-11.

[3] X. Qi, F. Han, L. He, Y. Zhang, and G. Zhang, "Evaluation of microenvironment cleanliness for computer assisted sperm analysis system based on fusion of neutrosophic features," *Computer Methods and Programs in Biomedicine*, vol. 218, 2022, p. 106717.

[4] I. Iqbal, G. Mustafa, and J. J. D. Ma, "Deep learning-based morphological classification of human sperm heads," vol. 10, no. 5, 2020, p. 325.

[5] J. B. You, C. McCallum, Y. Wang, J. Riordon, R. Nosrati, and D. J. N. R. U. Sinton, "Machine learning for sperm selection," 2021, pp. 1-17.

[6] M. Imani and H. J. I. f. Ghassemian, "An overview on spectral and spatial information fusion for hyperspectral image classification: Current trends and challenges," vol. 59, 2020, pp. 59-83.

[7] I. Córdón, S. García, A. Fernández, and F. J. K.-B. S. Herrera, "Imbalance: Oversampling algorithms for imbalanced classification in R," vol. 161, 2018, pp. 329-341, 2018.

[8] N. V. Revollo, G. Sarmiento, C. Delrieux, M. Herrera, and R. González-José, "Supervised Machine Learning Classification of Human Sperm Head Based on Morphological Features," in *Trends and Advancements of Image Processing and Its Applications: Springer*, 2022, pp. 177-191.

[۹] حامد جباری، نوشین بیگدلی و علی خادم، "یک روش هیبریدی جدید جهت قطعه بندی و تشخیص تومورهای موجود در تصاویر ماموگرافی بافت پستان"، فصلنامه بیماری‌های پستان، دوره ۹، شماره ۳، دی ۱۳۹۵، صفحه ۲۴-۱۴.

[۱۰] نوشین بیگدلی، حامد جباری و نگار ملکی، "یک روش ترکیبی هوشمند برای تشخیص، مرزبندی و طبقه‌بندی توده‌های پستان مبتنی بر ویژگی‌های بافت جدید مستخرج از دو نمای تصاویر ماموگرافی"، مجله ماشین بینایی و پردازش تصویر، دوره ۵، شماره ۲، آذر ۱۳۹۷، صفحه ۸۳-۶۹.

[۱۱] نوشین بیگدلی، حامد جباری و مهدی شجاعی، "یک روش هوشمند برای طبقه‌بندی ترک در سازه های بتنی بر اساس شبکه‌های عصبی عمیق"، مهندسی عمران امیرکبیر، دوره ۵۳، شماره ۸، آبان ۱۴۰۰، صفحه ۳۲۲۰-۳۲۰۱.

[۱۲] جباری. حامد و بیگدلی. نوشین، "طراحی و ارزیابی یک شبکه عصبی کپسولی جدید برای طبقه‌بندی نامتوازن تصاویر"، مجله ماشین بینایی و پردازش تصویر، دوره ۹، شماره ۱، فروردین ۱۴۰۱، صفحه ۱۵-۱.

- [13] K. Muhammad, S. Khan, J. Del Ser, V. H. C. J. I. T. o. N. N. de Albuquerque, and L. Systems, "Deep learning for multigrade brain tumor classification in smart healthcare systems: A prospective survey," vol. 32, no. 2, 2020, pp. 507-522.
- [14] M. Vannucci and V. Colla, "Genetic algorithms based resampling for the classification of unbalanced datasets," in *International Conference on Intelligent Decision Technologies*, 2017: Springer, pp. 23-32 .
- [15] K. Polat, "A hybrid approach to Parkinson disease classification using speech signal: The combination of SMOTE and random forests," in *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, 2019: IEEE, pp. 1-3 .
- [16] G. Hu, B. Du, X. Wang, and G. Wei, "An enhanced black widow optimization algorithm for feature selection," *Knowledge-Based Systems*, vol. 235, 2022, p. 107638.
- [17] J. Zhang and Q. Dai, "A cost-sensitive active learning algorithm: toward imbalanced time series forecasting," *Neural Computing and Applications*, vol. 34, no. 9, 2022, pp. 6953-6972.
- [18] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. J. I. S. P. M. Bharath, "Generative adversarial networks: An overview," vol. 35, no. 1, 2018, pp. 53-65.
- [19] J. Riordon, C. McCallum, D. J. C. i. b. Sinton, and medicine, "Deep learning for the classification of human sperm," vol. 111, 2019, p. 103342.
- [20] S. Javadi, S. A. J. C. i. b. Mirroshandel, and medicine, "A novel deep learning method for automatic assessment of human sperm images," vol. 109, 2019, pp. 182-194.
- [21] A. M. Yibre, B. J. E. S. Koçer, and a. I. J. Technology, "Semen quality predictive model using Feed Forwarded Neural Network trained by Learning-Based Artificial Algae Algorithm," vol. 24, no. 2, 2021, pp. 310-318.
- [22] V. Valiuškaitė, V. Raudonis, R. Maskeliūnas, R. Damaševičius, and T. J. S. Krilavičius, "Deep learning based evaluation of spermatozoid motility for artificial insemination," vol. 21, no. 1, 2021, p. 72.
- [23] H. O. Ilhan, I. O. Sigirci, G. Serbes, N. J. M. Aydin, b. engineering, and computing, "A fully automated hybrid human sperm detection and classification system based on mobile-net and the performance comparison with conventional methods," vol. 58, no. 5, 2020, pp. 1047-1068.
- [24] M. Yüzkat, H. O. Ilhan, N. J. C. i. B. Aydin, and Medicine, "Multi-model CNN fusion for sperm morphology analysis," vol. 137, 2021, p. 104790.
- [۲۵] راضیه راستگو و کورش کیانی، "شناسایی چهره با استفاده از تنظیم دقیق شبکه‌های کانولوشنی عمیق و رویکرد یادگیری انتقالی"، *مجله مدل‌سازی در مهندسی*، دوره ۱۷، شماره ۵۸، مهر ۱۳۹۸، صفحه ۱۱۱-۱۰۳.
- [۲۶] مرضیه زارع نظری، محسن سرداری زارچی، سیما عمادی و هادی پورمحمدی، "چارچوبی برای استخراج آناتومی و طبقه‌بندی تصاویر پشه با رویکرد یادگیری عمیق"، *مجله مدل‌سازی در مهندسی*، دوره ۲۰، شماره ۷۰، مهر ۱۴۰۱، صفحه ۱۲۰-۱۰۷.
- [27] M. Sarıgül, B. M. Ozyildirim, and M. J. N. N. Avcı, "Differential convolutional neural network," vol. 116, 2019, pp. 279-287.
- [28] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, H. J. C. i. b. Adeli, and medicine, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," vol. 100, 2018, pp. 270-278.
- [۲۹] مسلم سردشتی بیرجندی، حسین رحمانی و سعید فراهت، "کاربرد شبکه‌های عصبی عمیق در طبقه‌بندی تصاویر آسیب‌های شبکه فاضلاب و مشخص کردن مسیرهای بحرانی آنها"، *مجله مدل‌سازی در مهندسی*، دوره ۲۰، شماره ۷۰، صفحه ۱۳۲-۱۲۱.
- [30] M. K. Patrick, A. F. Adekoya, A. A. Mighty, B. Y. J. J. o. K. S. U.-c. Edward, and i. sciences, "Capsule networks—a survey ", 2019.
- [31] S. Sabour, N. Frosst, and G. E. J. a. p. a. Hinton, "Dynamic routing between capsules," 2017.
- [32] Human Sperm Head Morphology dataset (HuSHeM) [Online] Available: <https://data.mendeley.com/datasets/tt3yj2pf38/3>
- [33] K. Wang, C. Gou, Y. Duan, Y. Lin, X. Zheng, and F.-Y. J. I. C. J. o. A. S. Wang, "Generative adversarial networks: introduction and outlook," vol. 4, no. 4, 2017, pp. 588-598.
- [34] M. Mirza and S. J. a. p. a. Osindero, "Conditional generative adversarial nets," 2014.

- [35] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier gans," in *International conference on machine learning*, 2017: PMLR, pp. 2642-2651 .
- [36] T. J. J. o. A. I. Vijayakumar, "Comparative study of capsule neural network in various applications," vol. 1, no. 01, 2019, pp. 19-27.
- [37] L. Yu, R. Zhou, R. Chen, and K. K. Lai, "Missing data preprocessing in credit classification: One-hot encoding or imputation?," *Emerging Markets Finance and Trade*, vol. 58, no. 2, 2022, pp. 472-482.
- [38] C. Shorten and T. M. J. J. o. B. D. Khoshgoftaar, "A survey on image data augmentation for deep learning," vol. 6, no. 1, 2019, pp. 1-48.