J. of Marine Eng. Vol. 21, Iss. 47, 2025. (53-66)

Journal of Marine Engineering

Journal homepage: marine-eng.ir

# A Novel Intelligent Method for Underwater Image Classification Using Capsule Networks

**Hamid Hooshmand[1], Hamed Jabbari[2*]**

[1] *Faculty of Electrical and Robotics Engineering, Shahrood University of Technology, Semnan* Hamidhooshmand85@gmail.com

[2*] *Department of Electrical Engineering - Control, Faculty of Engineering, Imam Khomeini International University, Qazvin* H.Jabbari@edu.ikiu.ac.ir

## ARTICLE INFO

*Article History:*
Received: 8 Mar 2025
Last modification: 1 Oct 2025
Accepted: 6 Oct 2025
Available online: 8 Oct 2025

*Article type:*
Research paper

*Keywords:*
Underwater Images
Capsule Neural Network
Image Classification
Underwater Image Processing
Deep Learning

## ABSTRACT

Underwater images play a vital role in marine environmental research. Accurate analysis of these images is essential for managing marine ecosystems, preserving biodiversity, and monitoring environmental changes. However, various factors such as poor lighting, quality degradation due to light absorption and scattering, noise from suspended particles, and color distortions pose significant challenges in processing underwater images. These issues reduce the effectiveness of traditional image processing methods. Deep learning, as a powerful approach for automatically extracting complex features, can help mitigate or overcome these challenges. In this study, an intelligent deep learning method based on Capsule Neural Networks (CapsNets) is proposed for underwater image classification. By preserving spatial relationships among features and reducing reliance on pooling operations, CapsNets offer a better understanding of complex patterns. These characteristics enable effective handling of underwater image processing challenges. The proposed model, utilizing an advanced architecture, achieved higher classification accuracy compared to conventional methods and demonstrated robust performance under varying lighting conditions and image qualities. Experimental results showed that the proposed model outperformed existing approaches, achieving an overall accuracy of 96.75%. Moreover, the model exhibited stable performance across different underwater image classes, with accuracy ranging from 95.5% (lowest) to 98% (highest). Additionally, the average accuracy, sensitivity, and F1-score of the model were calculated as 96.75%, 96.75%, and 96.73%, respectively, indicating the model's robustness in underwater image classification. Based on these findings, the proposed model demonstrates strong potential for a wide range of applications, including marine habitat monitoring, underwater exploration, conservation of rare species, and tracking environmental changes.

ISSN: 2645-8136

**DOI:** http://dx.doi.org/10.66224/marineeng.21.47.5

**Copyright:** © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license [https://creativecommons.org/licenses/by/4.0/]

نشریه مهندسی دریا، جلد ۲۱، شماره ۴۷، ۱۴۰۴، (۶۶-۵۳)

نشریه مهندسی دریا

صفحه اصلی نشریه: marine-eng.ir

# یک روش هوشمند جدید برای طبقه‌بندی تصاویر زیر آب با استفاده از شبکه‌های کپسولی

**حمید هوشمند[۱]، حامد جباری[۲]***

[۱] استادیار، دانشکده مهندسی برق و رباتیک، دانشگاه صنعتی شاهرود، Hamidhooshmand85@gmail.com

[۲] استادیار، گروه مهندسی برق کنترل، دانشکده فنی و مهندسی، دانشگاه بین‌المللی امام خمینی (ره) قزوین، H.Jabbari@edu.ikiu.ac.ir

**اطلاعات مقاله**

*تاریخچه مقاله:*

تاریخ دریافت مقاله: ۱۴۰۳/۱۲/۱۸

تاریخ اصلاح مقاله: ۱۴۰۴/۰۷/۰۹

تاریخ پذیرش مقاله: ۱۴۰۴/۰۷/۱۴

تاریخ انتشار مقاله: ۱۴۰۴/۰۷/۱۶

*نوع مقاله:*

مقاله پژوهشی

*کلمات کلیدی:*

تصاویر زیر آب

شبکه‌های عصبی کپسولی

طبقه‌بندی تصاویر

پردازش تصاویر زیر آب

یادگیری عمیق

**چکیده**

تصاویر زیر آب نقش مهمی در پژوهش‌های زیست‌محیطی دریایی ایفا می‌کنند. تحلیل دقیق آن‌ها برای مدیریت اکوسیستم‌های دریایی، حفظ تنوع زیستی و نظارت بر تغییرات محیطی ضروری است. با این حال، عوامل مختلفی مانند نورپردازی نامناسب، افت کیفیت به‌دلیل جذب و پراکندگی نور، نویزهای ناشی از ذرات معلق و تغییرات رنگی، چالش‌های جدی در پردازش این تصاویر ایجاد می‌کنند. این مشکلات باعث کاهش کارایی روش‌های سنتی در پردازش تصاویر زیر آب می‌شوند. یادگیری عمیق به‌عنوان رویکردی قدرتمند در استخراج خودکار ویژگی‌های پیچیده، می‌تواند این چالش‌ها را کاهش دهد یا برطرف سازد. در این مقاله، یک روش یادگیری عمیق هوشمند مبتنی بر شبکه‌های عصبی کپسولی برای طبقه‌بندی تصاویر زیر آب ارائه شد. شبکه‌های کپسولی با حفظ روابط فضایی میان ویژگی‌ها و کاهش وابستگی به عملیات ادغام، درک بهتری از الگوهای پیچیده فراهم می‌کنند. این خصوصیات امکان مقابله مؤثر با چالش‌های پردازش تصاویر زیر آب را فراهم می‌سازند. مدل پیشنهادی با بهره‌گیری از ساختاری پیشرفته، نسبت به روش‌های متداول صحت طبقه‌بندی بالاتری ارائه داد و در برابر تغییرات شرایط نوری و کیفیت تصویر عملکرد پایداری داشت. نتایج آزمایش‌ها نشان داد که مدل پیشنهادی با صحت کلی ۹۶/۷۵ درصد عملکرد بهتری نسبت به سایر روش‌های موجود داشته است. هم‌چنین، مدل در طبقه‌بندی کلاس‌های مختلف تصاویر زیر آب، عملکردی پایدار با دامنه صحت بین ۹۵/۵٪ (کمترین) تا ۹۸٪ (بیشترین) از خود نشان داد. علاوه بر این، مقدار میانگین صحت، حساسیت و امتیاز F1 مدل به‌ترتیب ۹۶/۷۵٪، ۹۶/۷۵٪ و ۹۶/۷۳٪ محاسبه شد که نشان‌دهنده عملکرد مقاوم مدل در طبقه‌بندی تصاویر زیر آب است. بر اساس یافته‌های این پژوهش، مدل پیشنهادی از پتانسیل بالایی برای کاربردهای گسترده‌ای همچون نظارت بر زیستگاه‌های دریایی، اکتشافات زیرآبی، حفاظت از گونه‌های نادر و پایش تغییرات زیست‌محیطی برخوردار است.

ISSN: 2645-8136

**DOI:** http://dx.doi.org/10.66224/marineeng.21.47.5

**حق نشر:** © ۲۰۲۵ توسط نویسندگان. این اثر برای انتشار با دسترسی آزاد، تحت شرایط و ضوابط مجوز (CC BY)ارسال شده است.

در عصر حاضر، تصاویر زیر آب یکی از منابع اصلی اطلاعات برای پژوهش و مطالعه محیط‌های دریایی و اقیانوسی محسوب می‌شوند[۱]. این تصاویر از اهمیت بالایی در حوزه‌های علمی، زیست‌محیطی و صنعتی برخوردار است و برای تحلیل دقیق و کاربردی به روش‌های پردازش تصویر پیشرفته و دقیق نیاز دارند. شناسایی و طبقه‌بندی موجودات دریایی مانند ماهی‌ها، نقشی اساسی در مدیریت اکوسیستم‌های دریایی، حفظ تنوع زیستی، تحلیل تفاوت‌های گونه‌ای و حفاظت از گونه‌های در معرض خطر ایفا می‌کند[۲]. بررسی نحوه توزیع این موجودات، امکان تحلیل تأثیرات تغییرات اقلیمی جهانی و بهره‌برداری انسان از منابع دریایی را فراهم کرده است و زمینه‌ساز استفاده بهینه از این منابع می‌شود. با این حال، تحلیل تصاویر زیر آب با چالش‌های متعددی همراه است؛ از جمله نورپردازی ضعیف، تخریب لبه‌ها و جزئیات، تنوع رنگ و نویزهای مختلف. الگوریتم‌های طبقه‌بندی سنتی به دلیل ناتوانی در تشخیص ویژگی‌های کلیدی و استخراج اطلاعات معنادار، اثربخشی محدودی در این حوزه دارند و طبقه‌بندی تصاویر زیر آب را به چالشی قابل توجه تبدیل کرده‌اند[۳]. این مشکلات باعث می‌شوند بسیاری از روش‌های پردازش تصویر سنتی قادر به ارائه نتایج دقیق و قابل اعتماد نباشند. یکی از چالش‌های اصلی در طبقه‌بندی تصاویر زیر آب، استخراج ویژگی‌های مفید از تصاویری است که تحت تأثیر نویز قرار گرفته‌اند. روش‌های اولیه طبقه‌بندی تصاویر زیر آب عمدتاً بر اساس فناوری‌های پردازش تصویر و تشخیص الگو انجام می‌شدند و از تکنیک‌هایی مانند فیلترینگ و روش‌های پیش‌پردازش برای بهبود کیفیت تصویر و انجام قطعه‌بندی استفاده می‌کردند. در این زمینه، استفاده از روش‌های مدرن و پیشرفته برای طبقه‌بندی و تحلیل این تصاویر به یکی از نیازهای اساسی تبدیل شده است.

یکی از فناوری‌های نوظهوری که در سال‌های اخیر توجه زیادی را به خود جلب کرده است، هوش مصنوعی[1] (AI) و الگوریتم‌های یادگیری ماشین است[۴]. هوش مصنوعی بر توسعه سیستم‌هایی تمرکز دارد که قادر به انجام خودکار وظایفی هستند که نیازمند هوش انسانی‌اند، مانند تشخیص الگو، تصمیم‌گیری و یادگیری از تجربیات[۵]. الگوریتم‌های یادگیری ماشین به عنوان زیرمجموعه‌ای از هوش مصنوعی، این امکان را فراهم می‌سازند که سیستم‌ها بدون برنامه‌نویسی صریح از داده‌ها بیاموزند و عملکرد خود را بهبود بخشند. این الگوریتم‌ها به طور کلی به دو دسته تقسیم می‌شوند: یادگیری نظارت‌شده، که مدل‌ها را با داده‌های برچسب‌دار آموزش می‌دهد و یادگیری بدون نظارت، که به شناسایی الگوها و ساختارهای پنهان در داده‌های بدون برچسب می‌پردازد. در این میان، شبکه‌های یادگیری عمیق به عنوان یکی از پیشرفته‌ترین شاخه‌های یادگیری ماشین ظهور کرده‌اند. این شبکه‌ها با بهره‌گیری از ساختارهای پیچیده و چندلایه، قادر به استخراج ویژگی‌های سطح بالا و معنادار از داده‌های خام هستند[۶]. شبکه‌های عصبی کانولوشنی[2] (CNN) به ویژه در زمینه طبقه‌بندی و پردازش تصاویر عملکرد قابل توجهی دارند[۷]. این شبکه‌ها با استفاده از ساختارهای چندلایه و پیچیده خود می‌توانند ویژگی‌های مختلفی را از تصاویر استخراج کنند و در شناسایی و طبقه‌بندی اشیا به صحت بالایی دست یابند. هم‌چنین، در حوزه‌هایی مانند تشخیص تصویر، پردازش زبان طبیعی و پیش‌بینی رفتار، شبکه‌های یادگیری عمیق به ابزارهای قدرتمندی تبدیل شده‌اند که می‌توانند به طور قابل توجهی عملکرد و صحت سیستم‌های هوش مصنوعی را بهبود بخشند.

CNNها چندین مزیت کلیدی به‌ویژه در زمینه طبقه‌بندی تصاویر ارائه می‌دهند. یکی از مهم‌ترین مزایای CNNها توانایی آن‌ها در استخراج خودکار و کارآمد ویژگی‌ها از تصاویر بدون نیاز به مهندسی دستی ویژگی‌هاست[۸]. این قابلیت از طریق معماری سلسله‌مراتبی آن‌ها حاصل می‌شود، جایی که لایه‌های پی‌درپی از فیلترهای کانولوشنی الگوهای پیچیده‌تری را یاد می‌گیرند؛ از لبه‌ها و بافت‌های ساده در لایه‌های اولیه تا اجزای سطح بالای اشیاء و کل اشیاء در لایه‌های عمیق‌تر. مزیت مهم دیگر CNNها مقاومت آن‌ها در برابر تغییرات مقیاس، جهت و شرایط نوری تصویر است. لایه‌های کانولوشنی به‌گونه‌ای طراحی شده‌اند که از نظر مکانی نامتغیر باشند؛ به این معنی که می‌توانند الگوها را بدون توجه به موقعیت آن‌ها در تصویر تشخیص دهند. این ویژگی برای وظایف طبقه‌بندی تصاویر که اشیاء ممکن است در مکان‌ها یا جهت‌های مختلف ظاهر شوند، بسیار مهم است.

علاوه بر این، CNNها از لایه‌های ادغام بهره می‌برند؛ این لایه‌ها از یک سو ابعاد داده‌ها را کاهش می‌دهند و از سوی دیگر ویژگی‌های اساسی را حفظ می‌کنند. عملیات ادغام مانند ادغام ماکزیمم، بار محاسباتی را کاهش داده و با خلاصه‌سازی حضور ویژگی‌ها در نواحی محلی تصویر از بیش‌برازش جلوگیری می‌کند[۹]. هم‌چنین، CNNها در پردازش مجموعه داده‌های تصویری در مقیاس بزرگ بسیار کارآمد هستند، زیرا می‌توانند نمایش‌های سلسله‌مراتبی از ویژگی‌ها را یاد بگیرند. این کارایی آن‌ها را برای وظایفی مانند طبقه‌بندی تصاویر که نیاز به پردازش حجم عظیمی از داده‌ها دارند، بسیار مناسب می‌سازد. به‌طور کلی ترکیب استخراج خودکار ویژگی‌ها، نامتغیری فضایی، کاهش ابعاد و مقیاس‌پذیری، CNNها را به ابزاری قدرتمند و مؤثر برای طبقه‌بندی تصاویر تبدیل کرده است.

در دهه گذشته، مطالعات متعددی برای طبقه‌بندی تصاویر زیرآب انجام شده است. طبقه‌بندی تصاویر زیرآب به دلیل پراکندگی نور

[1] Artificial Intelligence

[2] Convolutional Neural Network

وابسته به طول موج، جذب و اعوجاج که منجر به کاهش کنتراست و کیفیت تصویر در محیط‌های عملیاتی دشوار می‌شود، چالشی اساسی به شمار می‌رود. آریدوس[1] و همکارانش مدل طبقه‌بندی تصویر زیرآب عمیق را با استفاده از CNN برای طبقه‌بندی خودکار تصاویر زیرآب ارائه کردند[۱۰]. مدل آنها با استفاده از تصاویر آموزشی و تکنیک‌های طبقه‌بندی، تصاویر کدر را براساس ویژگی‌های منتخب از مجموعه داده Benchmark Turbid Image طبقه‌بندی می‌کرد. نتایج آزمایش‌های آریدوس نشان داد که مدل آنها صحت مناسبی در طبقه‌بندی تصاویر کدر زیرآب دارد. هم‌چنین، مدل آنها با استفاده از تصاویر کدر با ویژگی‌های مختلف مورد اعتبارسنجی قرار گرفت تا توانایی تعمیم‌پذیری آن اثبات شود[۱۰]. محمود و همکارانش ویژگی‌های جدیدی به نام ResFeats را معرفی کردند که از لایه‌های کانولوشنی مختلف شبکه باقی‌مانده[2] (ResNet) که بر روی مجموعه داده ImageNet از پیش آموزش دیده بود، استخراج شدند[۱۱]. آنها هم‌چنین ResFeats استخراج‌شده از لایه‌های مختلف را ترکیب کردند تا ویژگی‌های عمیق فشرده و قدرتمندی به دست آورند. نتایج آنها نشان داد که ResFeats به طور مداوم عملکرد بهتری نسبت به ویژگی‌های مشابه CNN ارائه می‌دهد. روش آنها در مجموعه داده‌های Caltech-101، MLC، Benthoz15، EILAT و RSMAS عملکرد مؤثری داشت[۱۱].

ام-یانگ[3] و همکارانش شبکه توجه-آغازین[4] (IA-Net) را معرفی کردند که یک مدل مبتنی بر CNN برای طبقه‌بندی تصاویر زیرآب بود[۱۲]. در مدل آنها، ماژول توجه-آغازین[5] (I-A) برای شبیه‌سازی مکانیزم همبستگی بصری در طبقه‌بندی تصاویر محیط‌های خاص مانند مه، شب و زیرآب طراحی شده است. آزمایش‌های آنها بر روی مجموعه داده‌ای شامل ۴۰۰۰ تصویر زیر آب و ۵۰۰۰ تصویر غیر زیر آب به دقت ۹۹/۳ درصد در طبقه‌بندی تصاویر زیر آب دست یافت که به طور قابل توجهی بهتر از شبکه‌های کلاسیک طبقه‌بندی تصویر مانند AlexNet، InceptionV3 و ResNet بود[۱۲]. هم‌چنین جی-یانگ[6] و همکارانش الگوریتمی برای طبقه‌بندی اهداف زیرآب مبتنی بر الگوریتم جهت جریان[7] (FDA) بهبودیافته و استراتژی عامل جستجو پیشنهاد کردند که به طور همزمان می‌تولند پارامترهای وزن، بایاس و فرلپارامترهای ماشین یادگیری افراطی[8] (ELM) را بهینه‌سازی کند[۱۳]. در مرحله اول شبکه،DenseNet201 که با مجموعه داده ImageNet از پیش آموزش دیده بود، برای استخراج ویژگی‌ها و کاهش ابعاد تصاویر زیرآب استفاده شد. در مرحله دوم، طبقه‌بند ELM بهینه‌شده آموزش داده شد و مورد آزمایش قرار گرفت[۱۳]. برای رفع مشکلات مربوط به کاهش کیفیت تصاویر زیرآب، لی[9] و همکارانش شبکه توجه چندکاناله[10] (MCANet) را با یک رمزگذار فضای رنگی چندگانه پیشنهاد کردند که در طراحی روش‌های خاص بهبود تصویر زیرآب نقش مهمی داشت[۱۴]. آنها معماری MCANet را با تعبیه و پشته‌سازی ماژول‌های توجه چندکلنلله برای افزایش مداوم ادراک ویژگی‌های تصویر طراحی کردند. مدل آنها همچنین به صحت طبقه‌بندی ۹۸/۷۴ درصد دست یافت[۱۴].

طبقه‌بندی تصاویر زیرآب مبتنی بر بهبود تصویر و ارزیابی کیفیت اطلاعات توسط شائو[11] و همکارانش انجام شد[۱۵]. آنها عوامل مختلفی را که در فرآیند تصویربرداری زیرآب نویز ایجاد می‌کنند، تجزیه و تحلیل کردند و یک روش حذف نویز تصویر مبتنی بر فیلتر مختلط بلوکی براساس ویژگی‌های آماری نویز پیشنهاد دادند[۱۵]. هم‌چنین، یک مدل مبتنی بر CNN برای طبقه‌بندی و بهبود تصاویر زیرآب توسط جاماندلامودی[12] و همکارانش ارائه شد[۱۶]. آنها از معماری CNN که به طور خاص برای پردازش تصاویر زیرآب طراحی شده بود استفاده کردند تا به طور خودکار کیفیت بصری صحنه‌های زیرآب را بیاموزند و بهبود بخشند. شبکه پیشنهادی آنها بر روی یک مجموعه داده بزرگ از تصاویر زیرآب با تمرکز بر بهبود کنتراست و اصلاح رنگ آموزش دیده بود[۱۶].

علی‌رغم موفقیت‌های CNN، پژوهش‌های پیشین نشان داده‌اند که این شبکه‌ها دارای محدودیت‌های ذاتی هستند که چالش‌هایی را به همراه دارند. یکی از معایب عمده آنها ناتوانی در ثبت مؤثر روابط فضایی بین ویژگی‌ها است. CNNها برای کاهش ابعاد فضایی به لایه‌های ادغام متکی هستند که این موضوع می‌تواند منجر به از دست رفتن اطلاعات مهم مربوط به موقعیت و سلسله‌مراتب فضایی شود. این مشکل به ویژه هنگام مواجهه با تغییرات زاویه دید یا جهت‌گیری اشیاء در تصاویر، اثرگذار و مشکل‌ساز است. علاوه بر این، CNNها معمولاً در برابر تغییرات هندسی[13] مقاومت کمی دارند و برای دستیابی به دقت بالا نیازمند حجم زیادی از داده‌های آموزشی هستند. با وجود این محدودیت‌ها و خلأ پژوهشی در حوزه تحلیل تصاویر زیرآب، شبکه‌های عصبی کپسولی[14] (CapsNets) به عنوان یک رویکرد نوین و نویدبخش معرفی شده‌اند که می‌توانند بسیاری از این چالش‌ها را برطرف کنند.

شبکه‌های کپسولی از کپسول‌هایی تشکیل شده‌اند که گروه‌هایی از نورون‌ها هستند و به طور مشترک برای نمایش یک ویژگی خاص و

[1] Aridoss
[2] Residual Network
[3] M. Yang
[4] Inception-Attention Network
[5] Inception-Attention
[6] J. Yang
[7] Flow Direction Algorithm
[8] Extreme Learning Machine
[9] Li
[10] Multi-Channel Attention Network
[11] Xiao
[12] Jamandlamudi
[13] Affine Transformations
[14] Capsule Neural Networks

خصوصیات آن مانند جهت‌گیری و موقعیت کار می‌کنند [۱۷]. هر کپسول نه تنها وجود یک ویژگی را رمزگذاری می‌کند، بلکه روابط فضایی آن با سایر ویژگی‌ها را نیز ثبت می‌نماید. این رویکرد باعث حفظ سلسله‌مراتب فضایی و روابط میان ویژگی‌ها می‌شود که برای تشخیص اشیاء در زوایا و جهت‌گیری‌های مختلف اهمیت بسیاری دارد [۱۸]. شبکه‌های کپسولی هم‌چنین با استفاده از الگوریتم‌های مسیریابی پویا، اتصالات بین کپسول‌ها را بر اساس ویژگی‌هایی که هر کپسول نمایندگی می‌کند، تنظیم می‌کنند. این فرآیند موجب حفظ اطلاعات مهمی می‌شود که ممکن است در فرآیند ادغام در CNNها از دست برود. در نتیجه، این روش به درک بهتر و دقیق‌تری از ساختار فضایی تصاویر منجر می‌شود. CapsNetها در وظایف طبقه‌بندی تصویر مزایای متعددی نسبت به CNNهای سنتی دارند؛ به‌خصوص در سناریوهایی که شامل آرایش‌های فضایی پیچیده و تغییرات در ظاهر اشیاء هستند، عملکرد بهتری نشان می‌دهند. توانایی آن‌ها در مدیریت تغییرات هندسی و مقاومت در برابر الگوهای متنوع با داده‌های آموزشی محدود، این مدل‌ها را به گزینه‌ای ایده‌آل برای کاربردهایی تبدیل می‌کند که در آن‌ها تشخیص دقیق ویژگی‌ها اهمیت حیاتی دارد.

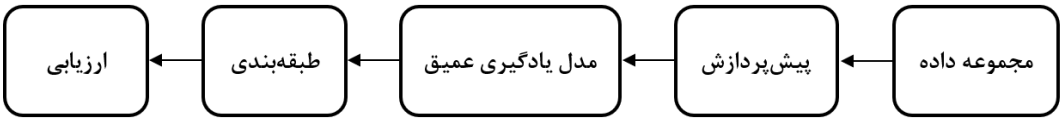
با توجه به پیچیدگی‌های خاص تصاویر زیرآب و نیاز به تحلیل دقیق‌تر، توسعه مدل‌های جدیدی که قادر به مقابله با این چالش‌ها باشند، می‌تواند پیشرفت‌های علمی و کاربردی چشمگیری ایجاد کند. به همین دلیل، در این مقاله طبقه‌بندی تصاویر زیرآب با استفاده از یک مدل هوشمند جدید مبتنی بر CapsNets ارائه شده است. معرفی مدل مبتنی بر CapsNets در این مقاله، گامی مؤثر و حیاتی در این مسیر است که هدف آن بهبود صحت و کارایی در تحلیل و طبقه‌بندی تصاویر زیر آب است. مدل پیشنهادی با بهره‌گیری از ساختارهای پیشرفته شبکه‌های کپسولی، توانست هم صحت و هم عملکرد را در طبقه‌بندی تصاویر زیر آب بهبود بخشد. هم‌چنین، نتایج این پژوهش می‌تواند به کاربردهای عملی گسترده‌ای در زمینه‌هایی مانند اکتشافات دریایی، حفاظت محیط زیست و تحلیل داده‌های محیطی منجر شود. نتایج تجربی نشان دادند که مدل پیشنهادی در مقایسه با روش‌های رایج، دقت بالاتری در طبقه‌بندی تصاویر زیرآب داشته و از پایداری بیشتری در طول آموزش برخوردار است. هم‌چنین این مدل با سرعت بیشتری به همگرایی می‌رسد. در مجموع، این ویژگی‌ها موجب می‌شوند تا مدل پیشنهادی گزینه‌ای کارآمد برای به‌کارگیری در سامانه‌های شناسایی اهداف زیرآبی باشد و به بهبود عملکرد و دقت این سامانه‌ها کمک کند.

ساختار این مقاله شامل بخش‌های مختلفی از جمله طراحی مدل پیشنهادی، ارزیابی عملکرد، و مقایسه با روش‌های پیشین است. در بخش دوم، روش و معماری مدل پیشنهادی به‌تفصیل معرفی و تشریح می‌شود؛ بخش سوم به ارائه و تحلیل نتایج تجربی اختصاص دارد؛ و در نهایت، بخش چهارم به جمع‌بندی یافته‌ها و ارائه نتیجه‌گیری کلی می‌پردازد.

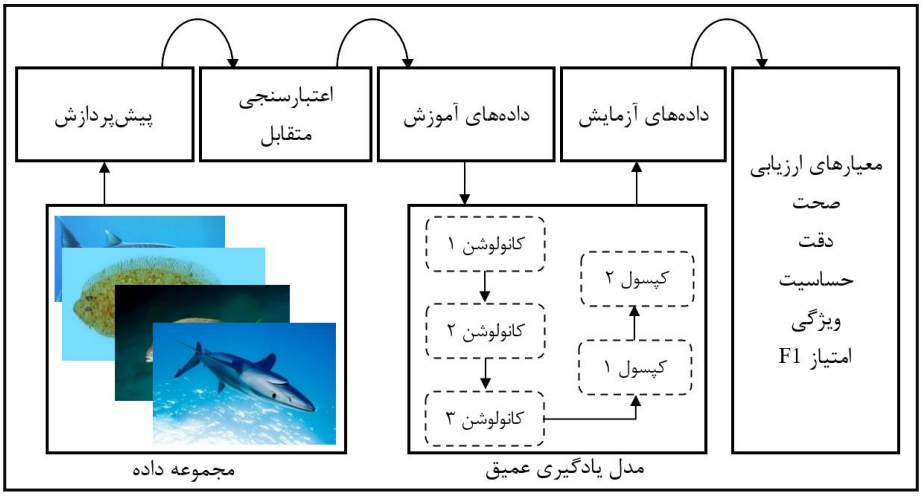
## ۲ – الگوریتم پیشنهادی

شکل ۱ فلوچارت الگوریتم پیشنهادی جهت طبقه‌بندی تصاویر زیرآب را نمایش می‌دهد. در این پژوهش، از یک مجموعه‌داده ترکیبی استفاده شده است که با ادغام چند پایگاه‌داده تخصصی در حوزه‌ی تصاویر زیرآب تهیه شده است. در مرحله‌ی پیش‌پردازش، تمامی تصاویر از نظر کیفیت کنتراست مورد بررسی قرار گرفتند. تصاویر دارای کنتراست پایین یا وضوح نامناسب با بهره‌گیری از روش‌های بهبود کنتراست اصلاح شدند تا کیفیت بصری مجموعه یکنواخت‌تر گردد. همچنین، برای کاهش اثرات سوگیری ناشی از تفاوت ابعاد تصاویر، تمامی نمونه‌ها به اندازه‌ای استاندارد مقیاس‌گذاری شدند. سپس، طبقه‌بندی تصاویر با استفاده از یک مدل یادگیری عمیق پیشنهادی انجام شد.

شکل ۲ معماری مدل و روند یادگیری آن را نشان می‌دهد. تصاویر مجموعه‌داده در چهار کلاس مختلف دسته‌بندی شده‌اند. برای ارزیابی عملکرد مدل و بررسی میزان تعمیم‌پذیری آن، از روش اعتبارسنجی متقابل ۴-پوشه‌ای استفاده گردید. بر این اساس، داده‌ها به‌صورت تصادفی به چهار زیرمجموعه تقسیم شدند. سپس در هر تکرار، یکی به عنوان مجموعه‌ی آزمون و سایر زیرمجموعه‌ها برای آموزش به‌کار گرفته شدند. این فرآیند چهار بار تکرار شد تا هر زیرمجموعه یک‌بار به‌عنوان داده‌ی آزمون مورد استفاده قرار گیرد. در پایان، میانگین معیارهای ارزیابی جهت کاهش سوگیری ناشی از تقسیم‌بندی تصادفی محاسبه گردید. در مرحله‌ی آموزش، پارامترهای مدل پیشنهادی که متشکل از پنج لایه اصلی است با استفاده از الگوریتم بهینه‌سازی مبتنی بر گرادیان و داده‌های آموزشی تنظیم شدند. سپس، عملکرد مدل آموزش‌دیده بر روی مجموعه‌ی آزمون ارزیابی گردید تا اثربخشی آن در طبقه‌بندی تصاویر زیرآب مورد سنجش قرار گیرد.




**شکل ۱ – فلوچارت الگوریتم پیشنهادی برای طبقه‌بندی تصاویر زیر آب**




**شکل ۲ – جزئیات روش و مدل یادگیری پیشنهادی**

### ۲-۱- مجموعه داده

در فرآیند توسعه‌ی مدل‌های طبقه‌بندی تصاویر زیرآب، انتخاب و بهره‌گیری از یک مجموعه‌داده‌ی آموزشی مناسب نقشی کلیدی ایفا می‌کند. به‌ویژه در حوزه‌ی یادگیری عمیق دسترسی به داده‌های بزرگ‌مقیاس و متنوع برای آموزش مؤثر چارچوب‌های پیچیده، امری حیاتی است. با این حال، به‌دست آوردن تصاویر یا ویدئوهای با کیفیت بالا در محیط‌های دریایی همواره با چالش‌های متعددی همراه است؛ عواملی چون شرایط نوری متغیر، تحرک موجودات، شفافیت آب و پس‌زمینه‌های پیچیده، گردآوری داده‌های دقیق و یکنواخت را با دشواری مواجه می‌سازند. در این پژوهش، به‌منظور فراهم‌سازی مجموعه‌داده‌ای با تنوع کافی و نماینده‌ی شرایط واقعی، از ادغام دو پایگاه‌داده‌ی معتبر در حوزه‌ی تصویربرداری زیر دریایی استفاده شده است: مجموعه‌داده‌ی Fish Species Image Data و مجموعه‌داده‌ی Shark Species.

مجموعه‌داده‌ی Fish Species شامل ۳۹۶۰ تصویر از ۴۶۸ گونه‌ی مختلف ماهی است که در سه سناریوی تصویربرداری متفاوت گردآوری شده‌اند. دسته‌ای از این تصاویر در شرایط کاملا کنترل‌شده و با نورپردازی استاندارد ثبت شده‌اند؛ در این تصاویر، نمونه‌های ماهی با پاله‌های گشوده در برابر پس‌زمینه‌ای یکنواخت قرار گرفته‌اند. دسته دوم تصاویر به ماهی‌هایی تعلق دارند که پس از خروج از آب و در شرایط نوری و پس زمینه متنوع عکاسی شده‌اند. دسته سوم نیز شامل تصاویر گرفته شده در زیستگاه‌های طبیعی ماهیان است که در آنها هیچ کنترلی بر نور یا پس‌زمینه وجود نداشته و شرایط ثبت کاملا واقع‌گرایانه است. چنین تنوعی در شرایط تصویربرداری موجب افزایش غنای بصری مجموعه داده و ارتقای کارایی مدل‌های آموزشی در مواجهه با تصاویر دنیای واقعی می‌گردد.

در کنار این مجموعه‌داده‌ی Fish Species، از مجموعه‌داده‌ی Shark Species نیز بهره گرفته شده است که شامل تصاویری از ۱۴ گونه‌ی مختلف کوسه بوده و برای هر گونه به‌طور میانگین حدود ۱۰۰ تصویر فراهم شده است. با توجه به تنوع ظاهری موجودات دریایی و برای حفظ تعادل و انسجام در داده‌های آموزشی، در این پژوهش در مجموع از ۴۰۰۰ تصویر منتخب در چهار کلاس استفاده شده است. این کلاس‌ها نه بر اساس برچسب‌های اولیه‌ی گونه‌ای، بلکه بر پایه‌ی شباهت‌های هندسی و ویژگی‌های بصری تصاویر انتخاب و سازمان‌دهی شده‌اند؛ به‌گونه‌ای که نمونه‌های هر کلاس از نظر ساختار بدنی و بافت‌های غالب، بیشترین همگنی درون‌کلاسی را دارا باشند. جدول ۱ نمونه‌هایی از این چهار کلاس را نمایش می‌دهد. همان‌طور که مشخص است کلاس‌ها از نظر ساختار بدنی و ابعاد ظاهری متفاوت‌اند.

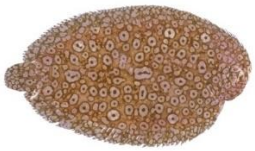
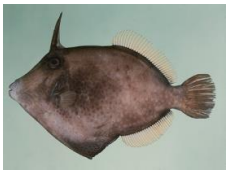
### ۲-۲- پیش‌پردازش

پیش‌پردازش داده‌ها گامی کلیدی در بهبود کیفیت تصاویر ورودی و افزایش دقت مدل‌های یادگیری عمیق است. در این پژوهش، تصاویر موجود در مجموعه داده ابتدا از نظر کیفیت بصری و کنتراست ارزیابی شدند. به منظور یکنواخت‌سازی کیفیت تصاویر و بهبود قابلیت یادگیری مدل، روی نمونه‌هایی که دارای کنتراست پایین یا وضوح نامناسب بودند (مانند تصاویر تیره یا محو) از روش متعادل‌سازی هیستوگرام[1] استفاده شد[۱۹]. علاوه بر این، به دلیل تنوع ابعاد تصاویر در دیتاست ترکیبی، تمامی تصاویر به ابعاد استاندارد ۵۲۶×۵۲۶ تغییر اندازه داده شدند تا ابعاد ورودی مدل یکنواخت باشد و از سوگیری ناشی از تفاوت اندازه جلوگیری شود. برای اثبات اثرگذاری پیش‌پردازش بر کیفیت داده‌ها، چند نمونه از تصاویر قبل و بعد از اعمال پیش‌پردازش در جدول ۲ ارائه شده است. تصاویر موجود در این شکل نشان‌دهنده افزایش وضوح و

[1] Histogram Equalization

کنتراست هستند که به بهبود کیفیت بصری و ثبات داده‌ها کمک می‌کند و در نهایت منجر به عملکرد بهتر مدل گردیده است.

**جدول ۱ – نمونه‌ای از تصاویر چهار کلاس انتخاب شده**

| کلاس ۱ | کلاس ۲ | کلاس ۳ | کلاس ۴ |
|---|---|---|---|
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**جدول ۲ – نمونه‌ای از تصاویر پیش‌پردازش شده**

| | | | |
|---|---|---|---|
| تصویر اصلی |  |  |  |
| تصویر پیش‌پردازش شده |  |  |  |

### ۳-۲- مدل عمیق پیشنهادی

بر اساس شکل ۲، پس از انتخاب ۴۰۰۰ تصویر برای ایجاد دیتاست، پیش‌پردازش اولیه‌ای روی تصاویر انجام شد. این پیش‌پردازش شامل بهبود کنتراست تصاویر با وضوح پایین و استانداردسازی ابعاد در بین کلاس‌های مختلف بود؛ به‌طوری که تمامی تصاویر به ابعاد ۵۲۶×۵۲۶ پیکسل تغییر اندازه داده شدند. مدل پیشنهادی این پژوهش یک مدل یادگیری ماشین بر اساس است که در آن الگوریتم از داده‌های آموزشی برچسب‌خورده یاد می‌گیرد و هدف آن پیش‌بینی یا تصمیم‌گیری بر اساس داده‌های جدید و نادیده‌گرفته‌شده است. این فرآیند شامل دو مرحله اصلی است: آموزش و آزمون. در مرحله آموزش، مدل عمیق با نمونه‌های برچسب‌خورده (جفت‌های ورودی-خروجی) ارائه می‌شود و یاد می‌گیرد که ویژگی‌های ورودی را به خروجی مطلوب متصل کند. این کار با حداقل کردن تابع از دست دادن انجام می‌شود که اختلاف بین خروجی پیش‌بینی‌شده و برچسب واقعی را اندازه‌گیری می‌کند. پس از آموزش مدل، آن را بر روی مجموعه‌ای جداگانه از نمونه‌های برچسب‌خورده (مجموعه آزمون) آزمایش می‌کنند تا عملکرد آن ارزیابی شود.

هدف یادگیری نظارت‌شده، ایجاد مدلی تعمیم‌پذیر است که بتواند الگوهای آموخته‌شده از داده‌های آموزشی را به داده‌های جدید و مشاهده نشده به‌درستی اعمال کند. مدل عمیق پیشنهادی این مقاله یک مدل یادگیری تحت نظارت بر اساس شبکه‌های CNN و CapsNet است. CNNها یک نوع از شبکه‌های عصبی عمیق هستند که بیشتر برای تجزیه و تحلیل تصاویر بصری کاربرد دارند. آن‌ها برای وظایف طبقه‌بندی تصاویر به‌ویژه مؤثر هستند زیرا توانایی یادگیری نمایش‌های سلسله‌مراتبی ویژگی‌های تصویر را دارند. CapsNetها با استفاده از کپسول‌ها برای تشکیل کپسول‌های سطح بالاتر برای نمایش پایدار، CNNها را تقویت می‌کنند. آن‌ها از یک رمزگذار و یک رمزگشا تشکیل شده‌اند. بخش رمزگذار یاد می‌گیرد که تصاویر را به پارامترهای برداری کدگذاری کند، در حالی که بخش رمزگشا بردارها را به عنوان تصاویر بازسازی می‌کند. CapsNets از مسیریابی پویا برای انجام تعاملات بین کپسول‌ها استفاده می‌کنند که جایگزین ادغام ماکزیمم شده است و ویژگی‌های سطح بالای تصویر را حفظ می‌کند [۲۰]. تابع هزینه نهایی که با Lc نمایش داده می‌شود و برای لایه آخر کپسول استفاده می‌شود، به شرح زیر است:

$$L_c = T_c.\max(0,m^+ - \|V_c\|)^2 + \lambda(1-T_c).\max(0,\|V_c\| - m^-)^2 \quad (۱)$$

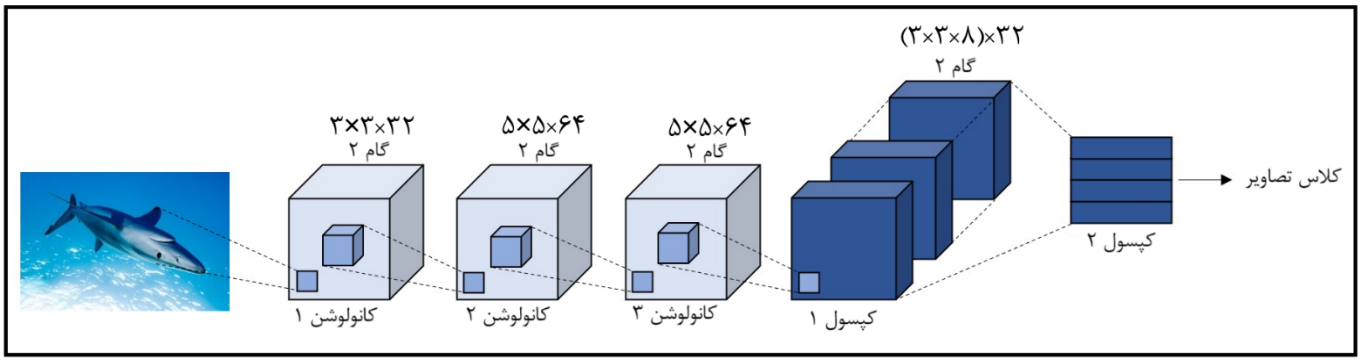
به طوری که $T_c = 1$ نشان‌دهنده برچسب صحیح کلاس $c$ است. $\lambda$ ، $m^+$ و $m^-$ پارامترهای ثابت هستند و $V_c$ خروجی لایه نهایی است.

شکل ۳ جزئیات مدل یادگیری عمیق پیشنهادی را نشان می‌دهد. این مدل شامل سه لایه کانولوشن با تابع فعال‌سازی ReLU و دو لایه کپسولی است. تصاویر زیرآب با ابعاد ۵۲۶×۵۲۶ به شبکه وارد می‌شوند که در حین آموزش، این تصاویر را به بردارهای پارامتری ۳۲ بعدی (۳۲D) رمزگذاری می‌کند. جدول ۳ جزئیات لایه‌های مدل طبقه‌بندی نظارت‌شده پیشنهادی را نشان می‌دهد. طبق این جدول، لایه‌های اول تا سوم کانولوشن از ۳۲، ۶۴ و ۶۴ فیلتر تشکیل شده‌اند که به ترتیب با گام ۲ و ابعاد خاصی تنظیم شده‌اند. تابع فعال‌سازی برای این سه لایه کانولوشن، تابع ReLU است. ورودی لایه کپسول اولیه، نگاشت‌های ویژگی با ابعاد ۶۴×۶۴ است. این لایه شامل ۳۲ کپسول است که هر کپسول از ۸ فیلتر با ابعاد ۳×۳ و گام ۲ تشکیل شده است. در این حالت ابعاد نگاشت‌های ویژگی خروجی به ۳۲×۳۲ کاهش می‌یابد. بنابراین خروجی هر کپسول به صورت یک تنسور سه بعدی با ابعاد (۳۲×۳۲)×۸ است که نشان‌دهنده بردارهای ویژگی ۸ بعدی در هر موقعیت مکانی است. با توجه به وجود ۳۲ کپسول، خروجی کلی این لایه به صورت یک تنسور چهار بعدی با ابعاد (۳۲×۳۲)×۸×۳۲ خواهد بود. بدین معنی که در هر یک از موقعیت مکانی، هر یک از ۳۲ کپسول، برداری ۸ بعدی تولید می‌کند.

لایه کپسول ماهی که لایه کپسول‌های کلاس است، دارای ۴ کپسول است. هر کپسول نمایانگر یک کلاس از تصاویر کشتی‌های دریایی است. در این لایه هر کپسول یک تنسور با ابعاد (۳۲×۳۲)×۸×۳۲ (یعنی ۳۲۷۶۸ بردار ۸ بعدی) را به عنوان ورودی می‌گیرد. برای نگاشت فضای ۸ بعدی به فضای خروجی ۳۲ بعدی، هر بردار ورودی در هر کپسول با یک ماتریس وزنی با ابعاد ۸×۳۲ ضرب می‌شود. بنابراین، برای هر کپسول کلاس ۳۲۷۶۸ بردار ۳۲ بعدی به‌دست می‌آید. سپس این بردارها از طریق مکانیزم مسیریابی پویا ترکیب و تجمیع می‌شوند تا یک بردار نهایی ۳۲ بعدی به عنوان نماینده کلاس مربوطه تولید گردد. در نتیجه، خروجی نهایی این لایه شامل ۴ بردار ۳۲ بعدی است که هر بردار نشان‌دهنده احتمال حضور تصویر ورودی در یکی از ۴ کلاس تعریف‌شده می‌باشد. در طول فرآیند آموزش، برای هر نمونه آموزشی یک مقدار هزینه برای هر یک از این چهار بردار طبق معادله (۱) محاسبه می‌شود. برای محاسبه هزینه نهایی هر نمونه، مقادیر محاسبه شده چهار بردار با هم جمع می‌شوند.

طبق مدل یادگیری عمیق پیشنهادی هر نمونه در مجموعه داده آموزشی یک برچسب صحیح دارد. در این حالت یک بردار چهار بعدی حاوی سه عنصر ۰ و یک عنصر ۱ تولید می‌شود که در آن مقدار ۱ نشان‌دهنده موقعیت صحیح (کلاس هدف) و مقدار ۰ بیانگر موقعیت‌های نادرست (کلاس‌های غیر هدف) است. این به عنوان کدگذاری One-Hot شناخته می‌شود [۲۱]. در معادله (۱)،

برچسب صحیح کلاس مقدار $Tc$ را تعیین می‌کند: این مقدار برابر با ۱ خواهد بود اگر برچسب صحیح با کپسول کلاس مطابقت داشته باشد و در غیر این صورت برابر با ۰ خواهد بود. در لایه آخر (لایه کپسول کلاس)، اولین، دومین، سومین و چهارمین کپسول به ترتیب مسئول رمزگذاری کلاس‌های ۱ تا ۴ تصاویر زیر آب هستند. اگر فرض کنیم که برچسب صحیح مربوط به کلاس اول است، این نشان می‌دهد که کپسول اول در لایه آخر مسئول رمزگذاری آن خواهد بود. برای این کپسول کلاس، مقدار $Tc = 1$ است و برای سه کپسول باقی‌مانده در این لایه، مقدار $Tc = 0$ خواهد بود. وقتی $Tc = 1$ باشد، بخش اول تابع هزینه محاسبه می‌شود و بخش دوم آن صفر است. در مدل پیشنهادی، مقدار $m+ = 0.9$ و $m- = 0.1$ تنظیم شده است.




**شکل ۳ – جزئیات مدل یادگیری عمیق پیشنهادی**

**جدول ۳ – جزئیات لایه‌ها در مدل پیشنهادی**

| لایه | ورودی لایه | تعداد فیلتر/کپسول | خروجی لایه |
|---|---|---|---|
| کانولوشن ۱ | تصویر با ابعاد ۲۵۶×۲۵۶ | ۳۲ فیلتر با ابعاد ۳×۳ با گام ۲ و تابع فعالساز ReLu | نگاشت ویژگی با ابعاد ۲۶۳×۲۶۳ |
| کانولوشن ۲ | نگاشت ویژگی با ابعاد ۲۶۳×۲۶۳ | ۶۴ فیلتر با ابعاد ۵×۵ با گام ۲ و تابع فعالساز ReLu | نگاشت ویژگی با ابعاد ۱۳۱×۱۳۱ |
| کانولوشن ۳ | نگاشت ویژگی با ابعاد ۱۳۱×۱۳۱ | ۶۴ فیلتر با ابعاد ۵×۵ با گام ۲ و تابع فعالساز ReLu | نگاشت ویژگی با ابعاد ۶۵×۶۵ |
| کپسول ۱ | نگاشت ویژگی با ابعاد ۶۴×۶۴ | ۳۲ کپسول حاوی ۸ فیلتر با ابعاد ۳×۳ با گام ۲ | تنسور با ابعاد (۳۲×۳۲)×۸×۳۲ |
| کپسول ۲ | تنسور با ابعاد (۳۲×۳۲)×۸×۳۲ | ۴ کپسول ۳۲ بعدی (۳۲D) | ماتریس ۳۲×۴ |

## ۳ – نتایج

در این بخش، نتایج حاصل از شبیه‌سازی مدل پیشنهادی مورد تحلیل قرار می‌گیرد. ابتدا ماتریس درهمی به‌منظور ارزیابی دقت طبقه‌بندی بررسی می‌شود و نمونه‌هایی از تصاویر با پیش‌بینی‌های درست و غلط ارائه می‌گردد. در ادامه، معیارهای ارزیابی کمی به‌کار رفته و تغییرات آن‌ها برای مدل پیشنهادی تحلیل می‌شود. در نهایت، عملکرد مدل پیشنهادی با برخی از مدل‌های شناخته شده در مطالعات پیشین مقایسه می‌گردد. تمام آزمایش‌ها به زبان برنامه‌نویسی پایتون بر روی سیستمی با پردازنده Intel® Core i7نو کارت گرافیک NVIDIA GeForce RTX 3050 با ۸ گیگابایت رم و سیستم عامل ویندوز ۱۱ انجام شده است.

### ۳-۱- معیارهای کمی ارزیابی

برای نشان دادن عملکرد مدل عمیق پیشنهادی در طبقه‌بندی تصاویر کشتی‌های دریایی، از معیارهای کمی مختلفی مانند صحت، دقت، حساسیت، ویژگی و امتیاز F1 استفاده شده است. صحت نشان می‌دهد مدل تا چه اندازه خروجی را به درستی پیش‌بینی کرده است. دقت سطح اطمینان مدل را در خروجی پیش‌بینی شده نمایش می‌دهد. حساسیت، عملکرد طبقه‌بند را در تشخیص صحیح نمونه‌های کلاس هدف نشان می‌دهد. ویژگی، توانایی طبقه‌بند در تشخیص صحیح نمونه‌های متعلق به سایر کلاس‌ها (غیر از کلاس هدف) را اندازه‌گیری می‌کند. امتیاز F1 میانگین همساز است که ترکیبی از دقت و حساسیت را توصیف می‌کند. بر اساس ماتریس درهمی، معیارهای کمی ارزیابی انتخاب شده عبارت است از [۲۲]:

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \quad (۲)$$

$$\text{Precision}=\frac{TP}{TP+FP} \quad (۳)$$

$$\text{Sensitivity}=\frac{TP}{TP+FN} \quad (۴)$$

$$\text{Specificity}=\frac{TN}{TN+FP} \quad (۵)$$

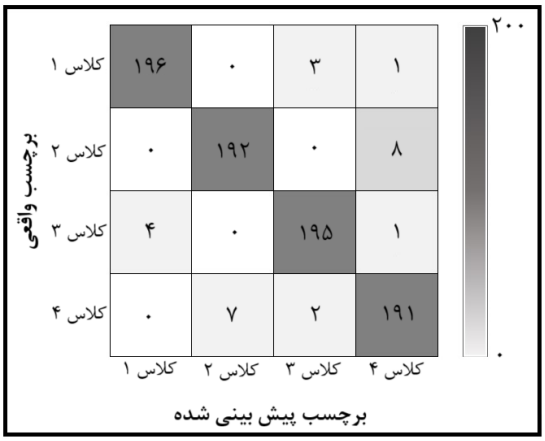
$$\text{F1-score}=\frac{2TP}{2TP+FP+FN} \quad (۶)$$

که به ترتیب نمایانگر صحت، دقت، حساسیت، ویژگی و امتیاز F1 هستند و در آنها، TP تعداد کشتی‌هایی است که به‌درستی طبقه‌بندی شده‌اند، TN تعداد نمونه‌هایی که به درستی به کلاس‌های دیگر (غیر از کلاس هدف) تخصیص یافته‌اند، FP تعداد کشتی‌های غیر هدفی که به اشتباه در کلاس هدف قرار گرفته‌اند و FN تعداد کشتی‌های هدف است که به اشتباه به کلاس‌های دیگر نسبت داده شده‌اند. عملکرد مدل تحت نظارت پیشنهادی بر اساس معیارهای ارزیابی ذکر شده از نظر پایداری، اثربخشی و توانایی تعمیم‌دهی اعتبارسنجی می‌شود.

### ۲-۳- طبقه‌بندی تصاویر زیرآب

همانطور که پیش‌تر ذکر شد، هر کلاس در مجموعه‌داده شامل ۱۰۰۰ تصویر زیرآب است. از این تعداد، ۸۰۰ تصویر (۸۰ درصد) به‌طور تصادفی برای آموزش و ۲۰۰ تصویر (۲۰ درصد) برای آزمایش در هر کلاس انتخاب شدند. مدل پیشنهادی همچنین با استفاده از اعتبارسنجی متقابل ۴-پوشه‌ای آموزش دید. ماتریس درهمی برای مدل پیشنهادی در این پژوهش در شکل ۴ نشان داده شده است. همانطور که از این شکل استنباط می‌شود، الگوریتم پیشنهادی با صحت ۹۶٪/۷۵ تصاویر زیرآب را به‌درستی طبقه‌بندی کرده است. این مقدار نشان می‌دهد که مدل توانسته است ۹۶/۷۵٪ از کل نمونه‌ها را به‌درستی طبقه‌بندی کند. بر اساس این شکل از ۲۰۰ تصویر موجود در کلاس ۱، ۱۹۶ تصویر به‌درستی طبقه‌بندی شدند. با این حال، از چهار تصویر باقی‌مانده، سه تصویر به اشتباه در کلاس ۳ و یک تصویر به‌اشتباه در کلاس ۴ طبقه‌بندی شدند. علاوه بر این، در کلاس ۲، هشت تصویر به اشتباه در کلاس ۴ طبقه‌بندی شدند. بیشترین میزان صحت طبقه‌بندی مربوط به کلاس ۱ بود که مدل توانست با صحت ٪۹۸ تعداد ۱۹۶ نمونه را به‌درستی شناسایی کند. در مقابل، پایین‌ترین عملکرد مدل در کلاس ۴ مشاهده شد؛ به‌طوری‌که از مجموع تصاویر این کلاس، تنها ۱۹۱ نمونه به‌درستی طبقه‌بندی شدند که معادل صحت ۹۵/۵٪ است.

دقت یکی از معیارهای کلیدی در ارزیابی عملکرد مدل‌های طبقه‌بندی است و نشان می‌دهد که از میان تمام نمونه‌هایی که مدل به عنوان مثبت پیش‌بینی کرده است، چه درصدی واقعاً مثبت بوده‌اند. به بیان دیگر، دقت مشخص می‌کند که پیش‌بینی‌های مثبت مدل چقدر قابل اعتماد هستند. دقت برای کلاس‌های ۱ تا ۴ به ترتیب ۹۸، ۹۶/۵، ۹۷/۵ و ۹۵ درصد بدست آمد. حساسیت که نشان‌دهنده نسبت نمونه‌های درست پیش‌بینی‌شده به کل نمونه‌های واقعی آن کلاس است، برای کلاس‌های ۱ تا ۴ به ترتیب برابر با ۹۸، ۹۶، ۹۷/۵ و ۹۵/۵ درصد بود. هم‌چنین ویژگی که میزان توانایی مدل در شناسایی نمونه‌های منفی را نشان می‌دهد، برای کلاس‌های ۱ تا ۴ به ترتیب ۹۹/۳، ۹۸/۸، ۹۹/۲ و ۹۸/۳ درصد محاسبه شد. این مقدار نشان می‌دهد که مدل در جلوگیری از تخصیص اشتباه نمونه‌های یک کلاس به سایر کلاس‌ها نیز عملکرد مناسبی داشته است. در نهایت، امتیاز F1 که میانگین هارمونیکی دقت و حساسیت است، برای کلاس‌های ۱ تا ۴ برابر با ۹۸، ۹۶/۲، ۹۷/۵ و ۹۵/۲ درصد محاسبه شد. جدول ۴ معیارهای ارزیابی کمی مدل پیشنهادی را نشان می‌دهد. مطابق این جدول میانگین معیارهای دقت، حساسیت، ویژگی و امتیاز F1 به ترتیب برابر با ۹۶٪/۷۵، ۹۶٪/۷۵، ۹۸٪/۹۰ و ۹۶/۷۳٪ است که نشان از عملکرد مناسب و پایدار مدل دارد.




**شکل ۴ – ماتریس درهمی برای مدل یادگیری پیشنهادی**

**جدول ۴ – معیارهای ارزیابی کمی مدل پیشنهادی**

| معیار / کلاس | دقت | حساسیت | ویژگی | امتیاز F1 |
|---|---|---|---|---|
| کلاس ۱ | ۹۸ | ۹۸ | ۹۹/۳ | ۹۸ |
| کلاس ۲ | ۹۶/۵ | ۹۶ | ۹۸/۸ | ۹۶/۲ |
| کلاس ۳ | ۹۷/۵ | ۹۷/۵ | ۹۹/۲ | ۹۷/۵ |
| کلاس ۴ | ۹۵ | ۹۵/۵ | ۹۸/۳ | ۹۵/۲ |

نتایج حاصل از ارزیابی مدل طبقه‌بندی با استفاده از معیارهای دقت، حساسیت، ویژگی و امتیاز F1 نشان می‌دهد که مدل در شناسایی کلاس‌ها عملکرد نسبتاً باثبات و مطلوبی دارد. مقدار دقت برای همه کلاس‌ها بالا بوده و در بازه ۹۵٪ تا ۹۸٪ قرار دارد که نشان‌دهنده این است که اکثر پیش‌بینی‌های مثبت مدل، واقعاً صحیح بوده‌اند. بالاترین مقدار دقت مربوط به کلاس ۱ با مقدار ۹۸٪ است، که بیانگر توانایی بالای مدل در پیش‌بینی صحیح این کلاس بدون اشتباه گرفتن نمونه‌های دیگر با آن است. از سوی دیگر، حساسیت که بیانگر توانایی مدل در شناسایی صحیح تمام نمونه‌های واقعی هر کلاس است، نیز در سطح مناسبی قرار دارد. کلاس‌های ۱ و ۳ با مقدار حساسیت ۹۸٪ و ۹۷/۵٪ بهترین عملکرد را داشته‌اند، در حالی که پایین‌ترین مقدار مربوط به کلاس ۲ با مقدار ۹۶٪ است. این تفاوت اندک ممکن است ناشی از تداخل بیشتر کلاس ۲ با کلاس ۴ باشد که در ماتریس درهمی نیز مشهود بود. معیار ویژگی برای تمامی کلاس‌ها بسیار بالا و در بازه‌ی ۹۸/۳٪

تا ۹۹/۳٪ قرار دارد. این موضوع نشان می‌دهد که مدل در تشخیص نمونه‌های منفی نیز عملکرد بسیار خوبی داشته و به ندرت دچار خطای مثبت کاذب شده است. ویژگی بالا به‌ویژه در مسائل حساس، مثل تشخیص بیماری، از اهمیت بالایی برخوردار است. در نهایت، مقادیر امتیاز F1 که ترکیبی متوازن از دقت و حساسیت است، تأیید می‌کند که مدل از نظر کلی عملکرد خوبی در شناسایی تمام کلاس‌ها داشته است. مقدار امتیاز F1 برای تمام کلاس‌ها بین ۹۵/۲٪ تا ۹۸٪ بوده است و کمترین مقدار آن نیز همچنان در سطح قابل قبول قرار دارد. این موضوع نشان می‌دهد که مدل نه تنها پیش‌بینی‌های دقیقی انجام می‌دهد، بلکه درصد قابل توجهی از نمونه‌های واقعی هر کلاس را نیز شناسایی می‌کند.

جدول ۵ نمونه‌هایی از تصاویر زیرآب با پیش‌بینی‌های درست و غلط توسط مدل پیشنهادی را نشان می‌دهد. همان‌گونه که در این جدول مشاهده می‌شود، برخی از موارد پیش‌بینی غلط توسط مدل پیشنهادی را می‌توان تا حدی قابل توجیه دانست. به‌عنوان مثال، نمونه‌ای که به اشتباه در کلاس ۴ طبقه‌بندی شده است، از نظر ویژگی‌های ظاهری شباهت قابل توجهی با تصاویر این کلاس دارد. در واقع، این مورد خاص به‌دلیل نزدیکی ظاهری بین کلاس‌ها، طبقه‌بندی دشواری را ایجاد کرده است؛ به‌گونه‌ای که حتی تشخیص صحیح آن توسط انسان نیز می‌تواند چالش‌برانگیز باشد. این موضوع برای نمونه‌ای از کلاس ۳ که به اشتباه در کلاس ۱ طبقه‌بندی شده است نیز صادق است.

**جدول ۵ – نمونه‌هایی از تصاویر با پیش‌بینی درست و غلط توسط مدل یادگیری پیشنهادی**

| کلاس واقعی / نوع پیش‌بینی | کلاس ۱ | کلاس ۲ | کلاس ۳ | کلاس ۴ |
|---|---|---|---|---|
| **درست** |  |  |  |  |
| **غلط** |  پیش‌بینی: کلاس ۳ |  پیش‌بینی: کلاس ۴ |  پیش‌بینی: کلاس ۱ |  پیش‌بینی: کلاس ۳ |

جدول ۶ مقایسه مدل پیشنهادی را با برخی از مدل‌های پیشین نشان می‌دهد. با آنکه در این جدول مقایسه‌ای میان عملکرد روش پیشنهادی و برخی مدل‌های پیشین ارائه شده است، باید تأکید شود که این مقایسه به صورت مستقیم چندان قابل اتکا نیست. دلیل اصلی این امر، تفاوت بنیادین در نوع و ساختار مجموعه داده‌های مورد استفاده است. اکثر مطالعات پیشین از مجموعه داده‌های عمومی و متنوع تصاویر زیر آب بهره گرفته‌اند که هدف آن‌ها طبقه‌بندی کلی موجودات دریایی یا مناظر زیر آب بوده است. در مقابل، در پژوهش حاضر تمرکز بر شناسایی دقیق گونه‌های ماهی بر اساس ویژگی‌های ظاهری نظیر شکل بدن و ساختار اندام‌ها بوده و مجموعه داده‌ای خاص، متناسب با این هدف انتخاب شده است. بنابراین، تفاوت در اهداف، دامنه، و ساختار داده‌ها سبب می‌شود نتایج جدول ۶ صرفاً به‌عنوان یک ارزیابی تطبیقی تقریبی تلقی شوند و نه به عنوان مقایسه‌ای مطلق. با وجود این تفاوت‌ها، نتایج روش پیشنهادی از نظر صحت طبقه‌بندی قابل توجه است. این روش با دستیابی به دقت ۹۸٪ عملکردی بالاتر نسبت به اغلب روش‌های پیشین ارائه داده است. با توجه به چالش‌های موجود در مجموعه داده طراحی شده، به‌ویژه در زمینه تمایز بین گونه‌های ماهی با ویژگی‌های ظاهری نزدیک به هم، این میزان دقت نشان‌دهنده توانمندی مدل پیشنهادی در استخراج ویژگی‌های معنادار و تفکیک‌پذیر و به‌کارگیری آنها در فرآیند یادگیری و طبقه‌بندی است.

**جدول ۶ – مقایسه مدل پیشنهادی با مدل‌های قبلی**

| مدل | سال | مجموعه داده | تعداد مجموعه داده | تعداد کلاس | بالاترین صحت (٪) |
|---|---|---|---|---|---|
| [۱۱]ResFeats | ۲۰۲۰ | Caltech-101 | ~۹k | ۱۰۲ | ۹۴/۷ |
| [۱۳]FCMFDA | ۲۰۲۲ | URPC 2018 | ~۴k | ۴ | ۹۷/۵ |
| [۱۶]CNN based | ۲۰۲۴ | جمع‌آوری شده | ~۱۰k | ۶ | ۹۲ |
| [۱۵]Enhanced ResNet-101 | ۲۰۲۴ | جمع‌آوری شده | ۲/۵k | ۵ | ۸۰/۸ |
| **روش پیشنهادی** | **۲۰۲۵** | **Fish Species Shark Species** | **۴k** | **۴** | **۹۸** |

با وجود تفاوت‌های بنیادین در اهداف و ساختار مجموعه‌داده‌ها میان پژوهش حاضر و مطالعات پیشین، می‌توان عملکرد مدل پیشنهادی را با برخی از مدل‌های یادگیری عمیق شناخته‌شده نیز مورد بررسی قرار داد. جدول ۷ مقایسه‌ای میان مدل پیشنهادی و مدل‌هایی همچون GoogleNet، ResNet-50، Inception-V3 و VGG16 ارائه می‌دهد. بر اساس نتایج این جدول، مدل ResNet-50 با صحت ۹۵٪/۶۱ بالاترین صحت طبقه‌بندی را در میان مدل‌های پایه نشان داده است. با این حال، زمان اجرای آن نسبتاً زیاد (۳/۲ ساعت) ثبت شده است. در مقابل، مدل پیشنهادی با دستیابی به صحت کلی ۹۶٪/۷۵ ضمن ارائه عملکرد بهتر در زمینه صحت طبقه‌بندی، از زمان اجرای بسیار کمتری (۱/۱ ساعت) نیز برخوردار بوده است. این ویژگی، مدل پیشنهادی را به گزینه‌ای مناسب برای کاربردهای زمان-واقعی و برخط بدل می‌سازد. لازم به ذکر است که تمرکز این مطالعه بر طبقه‌بندی چهار گونه خاص از ماهیان زیر آب بوده که شباهت‌های ساختاری با برخی از تجهیزات نظامی زیرسطحی مانند زیردریایی‌ها دارند. بنابراین، صحت بالای مدل پیشنهادی در شناسایی دقیق این گونه‌ها می‌تواند گامی مؤثر در توسعه سامانه‌های هوشمند شناسایی زیرآبی به‌ویژه در حوزه‌های نظامی و شبیه‌سازی تجهیزات زیرسطحی محسوب شود.

**جدول ۷ – مقایسه مدل پیشنهادی با برخی مدل‌های شناخته شده**

| مدل | زمان اجرا | صحت کلی |
|---|---|---|
| GoogleNet | ۳/۱ ساعت | ۹۴/۷۰ |
| ResNet-50 | ۳/۲ ساعت | ۹۵/۶۱ |
| Inception-V3 | ۲/۵ ساعت | ۹۰/۱۰ |
| VGG16 | ۴ ساعت | ۸۸/۴۰ |
| **مدل پیشنهادی** | **۱/۱ ساعت** | **۹۶/۷۵** |

### ۳-۳- بحث

عملکرد برتر مدل پیشنهادی مبتنی بر CapsNets در مقایسه با روش‌های کلاسیک یادگیری عمیق، ریشه در ویژگی‌های ساختاری منحصربه‌فرد این شبکه‌ها دارد. برخلاف شبکه‌های کانولوشنی سنتی مانند GoogleNet، ResNet و VGGNet که از عملیات ادغام ماکزیمم[1] برای کاهش ابعاد استفاده می‌کنند و در نتیجه بخشی از اطلاعات مکانی تصویر را از دست می‌دهند[۲۳]، شبکه‌های کپسولی با بهره‌گیری از کپسول‌ها قادرند روابط فضایی بین ویژگی‌ها را به صورت برداری حفظ کنند. این ویژگی باعث می‌شود مدل در شناسایی اشیای زیرآب که شکل، مقیاس و موقعیت آن‌ها در تصویر می‌تواند متفاوت باشد، درک دقیق‌تری از ساختار فضایی و هندسی داشته باشد. از سوی دیگر، استفاده از الگوریتم مسیریابی پویا در CapsNets، باعث می‌شود که ویژگی‌های مهم‌تر به کپسول‌های سطح بالاتر منتقل شوند و مسیر انتقال اطلاعات به‌صورت پویا تنظیم شود. این مکانیسم، برخلاف مسیرهای ثابت در CNNها، مدل را قادر می‌سازد تا در برابر نویز و تنوع ظاهری تصاویر مقاومت بیشتری نشان دهد. به‌ویژه در شرایط نوری ضعیف یا وجود چندین شیء در تصویر، شبکه‌های کپسولی به‌دلیل این ویژگی‌ها عملکرد پایدارتری دارند. برتری مدل در سرعت همگرایی نیز قابل توجه است. معماری سبک‌تر نسبت به برخی شبکه‌های عمیق و حذف لایه‌های غیرضروری باعث کاهش زمان آموزش بدون افت دقت شده است. این موضوع نشان می‌دهد که شبکه کپسولی پیشنهادی نه تنها از نظر دقت، بلکه از نظر بهره‌وری محاسباتی نیز کارآمد است.

یکی از چالش‌های کلیدی در حوزه طبقه‌بندی تصاویر وجود توزیع نامتوازن در کلاس‌های مختلف است[۲۶-۲۴]. در بسیاری از مجموعه داده‌های واقعی، برخی کلاس‌ها (مانند گونه‌های خاص جانداران دریایی یا اشیای کم‌یاب) دارای نمونه‌های بسیار محدودتری نسبت به سایر کلاس‌ها هستند[۲۷]. این عدم توازن می‌تواند منجر به سوگیری مدل به سمت کلاس‌های پرتکرار و کاهش دقت در شناسایی نمونه‌های نادر شود. در مطالعات اخیر، راهکارهایی نظیر بازنمونه‌گیری هوشمند، طراحی توابع هزینه وزن‌دار، و به‌کارگیری معماری‌های مقاوم در برابر عدم توازن مانند شبکه‌های کپسولی ترکیب‌شده با تکنیک‌های یادگیری از داده‌های نامتوازن، جهت مقابله با این مشکل پیشنهاد شده‌اند[۲۸]. به‌کارگیری این رویکردها در کنار معماری پیشنهادی، می‌تواند به افزایش دقت در کلاس‌های کم‌نمونه و بهبود پایداری مدل در محیط‌های واقعی زیرآب منجر شود. بررسی این راهکارها در تحقیقات آتی می‌تواند گام مؤثری در ارتقای عملکرد مدل‌های طبقه‌بندی زیرآب باشد.

## ۴ – نتیجه‌گیری

با توجه به نقش حیاتی تصاویر زیرآب در تحقیقات زیست‌محیطی دریایی و اقیانوسی، تجزیه و تحلیل دقیق این تصاویر برای مدیریت مؤثر اکوسیستم و حفظ تنوع زیستی ضروری است. با این حال، انجام این تجزیه و تحلیل‌ها با چالش‌هایی مواجه است، مانند نوردهی ناکافی، کاهش کیفیت و انواع مختلف نویز که موانع قابل توجهی برای الگوریتم‌های طبقه‌بندی سنتی ایجاد می‌کند. این مطالعه به بررسی استفاده از مدل‌های پیشرفته شبکه‌های کپسولی برای بهبود طبقه‌بندی تصاویر زیرآب پرداخته است. شبکه‌های کپسولی در حفظ روابط فضایی بین ویژگی‌ها و کاهش از دست دادن اطلاعات ناشی از ادغام برجسته هستند و به این ترتیب قابلیت‌های برتری برای شناسایی و طبقه‌بندی اشیاء در تصاویر زیرآب ارائه می‌دهند.

[1] Max-Pooling

با توجه به هدف اصلی پژوهش که طراحی مدلی دقیق و کارا برای طبقه‌بندی گونه‌های خاصی از ماهیان در تصاویر زیرآب بود، مدل پیشنهادی مبتنی بر CapsNets توسعه یافت. این مدل با تمرکز بر استخراج ویژگی‌های ظاهری دقیق از جمله شکل بدن و ساختار اندام‌ها، موفق شد نتایجی فراتر از مدل‌های رایج ارائه دهد. مدل پیشنهادی در این مطالعه به دنبال بهبود دقت و کارایی تجزیه و تحلیل تصاویر زیرآب از طریق استفاده از معماری‌های پیشرفته شبکه‌های کپسولی است. نتایج تجربی نشان می‌دهند که این مدل به طور قابل توجهی در مقایسه با روش‌های موجود از نظر صحت طبقه‌بندی، پایداری عملکرد در میان کلاس‌ها، و زمان اجرای پایین (در مقایسه با مدل‌های سنگین‌وزنی مانند ResNet و Inception) برتر است. این بهبود عملکرد به‌ویژه در شرایطی که گونه‌های مورد نظر دارای شباهت‌های ظاهری بالا هستند، اهمیت بیشتری می‌یابد. علاوه بر این، این مدل برای کاربردهای گسترده عملی در زمینه‌هایی مانند اکتشافات دریایی، حفاظت از محیط زیست و تجزیه و تحلیل داده‌های زیست‌محیطی وعده‌های زیادی دارد. از جمله ظرفیت‌های بالقوه مدل پیشنهادی می‌توان به کاربرد آن در طبقه‌بندی خودکار تجهیزات نظامی زیرآبی که از نظر ساختار ظاهری مشابه با گونه‌های زیستی انتخاب‌شده در این پژوهش هستند، اشاره کرد. این کاربرد نشان می‌دهد که مدل حاضر علاوه بر توان علمی، قابلیت تعمیم به حوزه‌های عملیاتی را نیز داراست.

## ۵ - مراجع


1- S. Mittal, S. Srivastava, and J. P. Jayanth, "A Survey of Deep Learning Techniques for Underwater Image Classification," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–15, 2022. https://doi.org/10.1109/tnnls.2022.3143887.

2- T. T. Chungath, A. M. Nambiar, and A .Mittal, "Transfer learning and few-shot learning based deep neural network models for underwater sonar image classification with a few samples," *IEEE Journal of Oceanic Engineering*, vol. 49, no. 1, pp. 294-310, 2023.https://doi.org/10.1109/JOE.2022.3221127.

3- M. S. Mohammed, H. A. Khater, Y .F. Hassan, and A. Elsayed, "Proposed approach for automatic underwater object classification," *ICIC International*, vol. 12, no. 12, pp. 1205-1212, 2018. https://doi.org/10.24507/icicel.12.12.1205.

4- M. Goyal, T. Knackstedt, S. Yan, and S. Hassanpour, "Artificial intelligence-based image classification methods for diagnosis of skin cancer: Challenges and opportunities," *Computers in biology and medicine*, vol. 127, p. 104065, 2020. https://doi.org/10.1016/j.compbiomed.2020.104065.

5- X. Du-Harpur, F. Watt, N. Luscombe, and M. Lynch, "What is AI? Applications of artificial intelligence to dermatology," *British Journal of Dermatology*, vol. 183, no. 3, pp. 423-430, 2020. https://doi.org/10.1111/bjd.18880.

6- A. Saleh, M. Sheaves, and M. Rahimi Azghadi, "Computer vision and deep learning for fish classification in underwater habitats: A survey," *Fish and Fisheries*, vol. 23, no. 4, pp. 977-999, 2022. https://doi.org/10.1111/faf.12666.

7- F. Han, J. Yao, H. Zhu, and C. Wang, "Underwater image processing and object detection based on deep CNN method," *Journal of Sensors*, vol. 2020, no. 1, p. 6707328, 2020. https://doi.org/10.1155/2020/6707328.

8- J. Bharadiya, "Convolutional neural networks for image classification," *International Journal of Innovative Science and Research Technology*, vol. 8, no. 5, pp. 673-677, 2023.

9- M. Tripathi, "Analysis of convolutional neural network based image classification techniques," *Journal of Innovative Image Processing (JIIP)*, vol. 3, no. 02, pp. 100-117, 2021. https://doi.org/10.36548/jiip.2021.2.003.

10- M. Aridoss, C. Dhasarathan, A. Dumka, and J. Loganathan, "DUICM deep underwater image classification mobdel using convolutional neural networks," *International Journal of Grid and High Performance Computing (IJGHPC)*, vol. 12, no. 3, pp. 88-100, 2020.

11- A. Mahmood, M. Bennamoun, S. An, F. Sohel, and F. Boussaid, "ResFeats: Residual network based features for underwater image classification," *Image and Vision Computing*, vol. 93, p. 103811, 2020. https://doi.org/10.1016/j.imavis.2019.09.002.

12- M. Yang ,H. Wang, K. Hu, G. Yin, and Z. Wei, "IA-Net: An inception–attention-module-based network for classifying underwater images from others," *IEEE Journal of Oceanic Engineering*, vol. 47, no. 3, pp. 704-717, 2022. https://doi.org/10.1109/JOE.2021.3126090.

13- J. Yang, M. Cai, X. Yang, and Z. Zhou, "Underwater image classification algorithm based on convolutional neural network and optimized extreme learning machine," *Journal of Marine Science and Engineering*, vol. 10, no. 12, p. 1841, 2022. https://doi.org/10.3390/jmse10121841.

14- G. Li *et al.*, "MCANet: Multi-channel attention network with multi-color space encoder for underwater image classification," *Computers and Electrical Engineering*, vol. 108, p. 108724, 2023. https://doi.org/10.1016/j.compeleceng.2023.108724.

15- S. Xiao, X. Shen, Z. Zhang, J. Wen, M. Xi, and J. Yang, "Underwater image classification based on image enhancement and information quality evaluation," *Displays*, vol. 82, p. 102635, 2024. https://doi.org/10.1016/j.displa.2023.102635.

16- S. Jamandlamudi, D. P. Isravel, and J. P. M. Dhas, "CNN Based Model for Underwater Image Classification and Enhancement," in *2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science*

structures based on deep neural networks, "*Amirkabir Journal of Civil Engineering*, vol. 53, no. 8, pp. 3201-3220. Oct. 2021. https://doi.org/10.22060/ceej.2020.17738.6660.
24- H. Jabbari and Nooshin Bigdeli, "New conditional generative adversarial capsule network for imbalanced classification of human sperm head images," *Neural Computing and Applications*, vol. 35, no. 27, pp. 19919–19934, Jul. 2023. https://doi.org/10.1007/s00521-023-08742-3.
25- H. Jabbari and N. Bigdeli, "A new hierarchical algorithm based on CapsGAN for imbalanced image classification," *IET Image Processing*, vol. 18, no. 1, pp. 194–210, Oct. 2023. https://doi.org/10.1049/ipr2.12942.
26- H. Jabbari and N. Bigdeli, "Design and Evaluation of a New Capsule Neural Network (CapsNet) for Imbalanced Images Classification," *Journal of Machine Vision and Image Processing*, vol. 9, no. 1, pp. 1–15, Apr. 2022.
27- H. Jabbari and N. Bigdeli, "A New Capsule Generative Adversarial Network for Imbalanced Classification of Human Sperm Images," *Journal of Modeling in Engineering*, vol. 21, no. 73, pp. 279-294. June. 2023. https://doi.org/10.22075/JME.2023.28349.2333.
28- H. Jabbari, H. Hooshmand, and N. Bigdeli, "A Novel Intelligent Method Based on Capsule Networks for Maritime Ship," Classification. *Journal of Machine Vision and Image Processing*, Accepted Manuscript, Available Online from 11 August 2025.
*(AMATHE)*, 2024: IEEE, pp. 1-8 . https://doi.org/10.1109/AMATHE61652.2024.10582240.
17- M. Khodadadzadeh, X. Ding, P. Chaurasia, and D. Coyle, "A hybrid capsule network for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 11824-11839, 2021. https://doi.org/10.1109/JSTARS.2021.3126427.
18- Z. Sun, G. Zhao, R. Scherer, W. Wei, and M. Woźniak, "Overview of capsule neural networks," *Journal of Internet Technology*, vol. 23, no. 1, pp. 33-44, 2022.
19- M. Abdullah-Al-Wadud, Md. Kabir, M. Akber Dewan, and O. Chae, "A Dynamic Histogram Equalization for Image Contrast Enhancement," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 593–600, 2007. https://doi.org/10.1109/tce.2007.381734.
20- V. T, "COMPARATIVE STUDY OF CAPSULE NEURAL NETWORK IN VARIOUS APPLICATIONS," *Journal of Artificial Intelligence and Capsule Networks*, vol. 01, no. 01, pp. 19–27, Sep. 2019. https://doi.org/10.36548/jaicn.2019.1.003.
21- P. Rodríguez, M. A. Bautista, J. Gonzàlez, and S. Escalera, "Beyond one-hot encoding: Lower dimensional target embedding," *Image and Vision Computing*, vol. 75, pp. 21–31, Jul. 2018. https://doi.org/10.1016/j.imavis.2018.04.004.
22- A. Ali-Gombe and E. Elyan, "MFC-GAN: Class-imbalanced dataset classification using multiple fake class generative adversarial network," *Neurocomputing*, vol. 361, pp. 212-221, 2019. https://doi.org/10.1016/j.neucom.2019.06.043.
23- N. Bigdeli, H. Jabbari, and M. Shojaei, "An intelligent method for crack classification in concrete