

ارزیابی کیفیت تصاویر زیر آب با استفاده از شبکه عصبی کانولوشن عمیق چندشاخه با بلوک توجه ECA

مریم کریمی^۱

چکیده

ارزیابی کیفیت تصاویر زیر آب به دلیل چالش‌هایی مانند نورپراکنی، جذب طول موج‌های رنگی، و کاهش تابین، نیازمند روش‌های تخصصی است. روش‌های سنتی ارزیابی کیفیت تصویر (IQA) به دلیل عدم توجه به ویژگی‌های منحصر به فرد محیط زیر آب، عملکرد محدودی دارند. در این مقاله، یک روش نوین مبتنی بر یادگیری عمیق برای ارزیابی کیفیت تصاویر زیر آب ارائه شده است. این روش از یک شبکه عصبی کانولوشنی (CNN) چندشاخه با بلوک توجه کانالی کارآمد (ECA) بهره می‌برد که به طور همزمان تصویر اصلی، نقشه تیزی، و نقشه ساختار را پردازش می‌کند. هر شاخه با استخراج خودکار ویژگی‌های مرتبط با کیفیت، اطلاعات مکمل را در اختیار مدل قرار می‌دهد. ترکیب این ویژگی‌ها با استفاده از بلوک‌های تجمیع و توجه، منجر به پیش‌بینی دقیق‌تر نمره کیفیت می‌شود. ارزیابی جامع بر روی مجموعه داده‌های متنوع SAUD2.0 و UID2021 نشان می‌دهد که مدل در هر دو مجموعه عملکرد قابل قبولی دارد و نسبت به روش‌های موجود دقت و همبستگی بالاتری را ارائه می‌دهد. همچنین، آزمایش ارزیابی متقابل بین این دو مجموعه داده تأیید کننده توانایی تعمیم و پایداری روش پیشنهادی در مواجهه با داده‌های متفاوت و ناشناخته است. این پژوهش گامی موثر در جهت بهبود ارزیابی کیفیت تصاویر زیر آب با رویکردی کارا و مقاوم ارائه می‌دهد.

کلید واژه‌ها

تصاویر زیر آب، ارزیابی کیفیت بدون مرجع، یادگیری عمیق، شبکه کانولوشنی، ارزیابی کیفیت تصاویر زیر آب، ارزیابی کیفیت تصاویر رنگی، شبکه‌های چند شاخه، سازوکار توجه.

۱- مقدمه

تغییرات رنگی پیچیده زیر آب کارایی لازم را ندارند. همچنین، پیشرفت‌های اخیر مانند روش‌های حذف مه و شبکه‌های عصبی عمیق علی‌رغم نتایج امیدوارکننده، همچنان نسبت به نویز، شرایط محیطی و شرایط تصویربرداری حساس هستند [۳-۶]. نمونه‌های پیشرفته شامل ADMNNet [۷] با ساختار چندشاخه‌ی پویا و مکانیزم توجه کانالی، و Hybrid-Net [۸] با ادغام مسیرهای پیش پردازش متفاوت است که بازسازی رنگ و کنتراست را بهبود داده‌اند. با وجود این پیشرفت‌ها، هیچ‌یک از روش‌های UIE قادر نیستند در تمام سناریوهای نوری و عمق‌های مختلف نتیجه‌ای یکنواخت ارائه دهند؛ افزون بر آن، خروجی این الگوریتم‌ها غالباً دستخوش تقویت بیش‌ازحد یا باقی‌ماندن اعوجاج پنهان است. بنابراین، معمولاً نمی‌توان یک روش بهبود تصویر را برای همه

تصویربرداری زیر آب در حوزه‌های مختلفی مانند حسگرهای زیر آب، زیست‌شناسی دریایی، اکتشافات زیر آب و پایش محیط زیست اهمیت زیادی دارد [۱]. با این حال، به دست آوردن تصاویر با کیفیت بالا در زیر آب بسیار چالش برانگیز است، چرا که عواملی مانند دید محدود، تابین پایین، کدورت آب، پراکندگی و تضعیف نور، جریان آب و اعوجاج رنگی بر کیفیت تصاویر تأثیر می‌گذارد [۲]. روش‌های معمول بهبود تصویر، مانند تنظیم تابین و همسان‌سازی هیستوگرام، اغلب در مواجهه با شرایط نوری و

این مقاله در اردیبهشت‌ماه ۱۴۰۴ دریافت شد؛ در مرداد بازنگری و در شهریور پذیرفته گردید.

مورد توجه قرار نگرفته‌اند. بنابراین، توسعه روش‌هایی که بتوانند به صورت خودکار و عمیق، ویژگی‌های پیچیده و متنوع تصاویر زیر آب را از داده‌ها استخراج و مدل‌سازی کنند، می‌تواند منجر به ارزیابی دقیق‌تر کیفیت این تصاویر در شرایط مختلف محیطی شود [۱۷].

در این مقاله، یک روش نوین و مبتنی بر یادگیری عمیق برای ارزیابی کیفیت تصاویر زیر آب ارائه شده است. ساختار مدل پیشنهادی به صورت چندشاخه طراحی شده که در آن سه نوع ورودی مکمل شامل تصویر اصلی، نقشه تیزی و نقشه ساختار به صورت مجزا پردازش می‌شوند. این سه ویژگی، بر اساس تحلیل جنبه‌های ادراکی مؤثر در کیفیت تصاویر زیر آب انتخاب شده‌اند، چراکه کاهش وضوح و تخریب ساختار از عوامل اصلی افت کیفیت در محیط‌های آبی به شمار می‌آیند. هر شاخه از مدل با استفاده از شبکه‌های کانولوشنی و بلوک توجه کانالی کارآمد (ECA) ۱۱ ویژگی‌های مرتبط را به‌طور خودکار استخراج کرده و سپس خروجی شاخه‌ها با هم ترکیب می‌شود تا نمره نهایی کیفیت پیش‌بینی شود. برای ارزیابی عملکرد مدل، آزمایش‌ها با استفاده از اعتبارسنجی متقابل روی دو مجموعه داده واقعی و دارای نمرات ادراکی انسانی (UID2021 و SAUD 2.0) انجام شده و نتایج به تفکیک زیر مجموعه‌های محیطی گزارش شده‌اند. بطور خلاصه نوآوری‌های اصلی روش پیشنهادی عبارتند از:

۱. طراحی معماری چندشاخه برای استخراج هدفمند ویژگی‌های ادراکی از سه نمایه مکمل تصویر؛
 ۲. استفاده از نقشه‌های تیزی و ساختار به‌عنوان ورودی‌های مستقل برای افزایش تفسیرپذیری و دقت ارزیابی؛
 ۳. به‌کارگیری بلوک توجه ECA در هر شاخه برای تقویت تطبیقی ویژگی‌ها؛
 ۴. تعریف تابع هزینه ترکیبی برای هم‌زمان‌سازی بهینه‌سازی عددی و ادراکی؛
- در ادامه این مقاله، بخش ۲ کارهای مرتبط را بررسی خواهد کرد. بخش ۳ به تفصیل روش پیشنهادی شرح داده خواهد شد. بخش ۴ به آزمایشات و نتایج خواهد پرداخت و در نهایت در بخش ۵ جمع‌بندی از کلیه قسمت‌ها ارائه خواهد شد.

۲- کارهای مرتبط

۲-۱- روش‌های عمومی ارزیابی کیفیت تصویر

در سال‌های اخیر، پیشرفت‌های چشمگیری در ارزیابی کیفیت تصویر (IQA) برای تصاویر ثبت شده در هوا حاصل شده است.

تصاویر زیر آب به‌طور هم‌زمان مناسب دانست [۹]. ارزیابی دقیق کیفیت تصاویر زیر آب نقش کلیدی در انتخاب، تنظیم و ارزیابی عملکرد الگوریتم‌های بهبود تصویر ایفا می‌کند. از آنجا که کیفیت تصاویر زیر آب تحت تأثیر شرایط متغیر محیطی به شدت نوسان دارد، وجود یک معیار قابل اعتماد، پایدار و قابل تکرار برای ارزیابی کیفیت، برای توسعه و اعتبارسنجی این الگوریتم‌ها ضروری است [۱۰].

ارزیابی کیفیت تصویر (IQA) ۱ به دو دسته کلی روش‌های انسانی ۲ و محاسباتی ۳ تقسیم می‌شود. از آنجا که انسان‌ها بینندگان نهایی تصاویر هستند، ارزیابی انسانی کیفیت تصویر به عنوان معتبرترین و دقیق‌ترین روش شناخته می‌شود. با این حال، این رویکرد در کاربردهای عملی قابل اجرا نیست. در نتیجه، ارزیابی انسانی کیفیت به عنوان معیار مرجع ۴ برای روش‌های محاسباتی مورد استفاده قرار می‌گیرد [۱۱، ۱۲]. بنابراین، روش‌های محاسباتی که بیشترین همبستگی را با ارزیابی انسانی دارند، ترجیح داده می‌شوند. روش‌های محاسباتی ارزیابی کیفیت تصویر از نظر استفاده از تصویر مرجع به سه دسته دارای مرجع کامل ۵ (FR)، مرجع کاهش‌یافته ۶ (RR) و بدون مرجع ۷ (NR) تقسیم می‌شوند. از آنجا که در سناریوهای زیر آب معمولاً تصویر مرجع در دسترس نیست، ارزیابی کیفیت تصاویر زیر آب عمدتاً با روش‌های بدون مرجع انجام می‌شود [۱۳].

در سال‌های اخیر، پیشرفت‌های قابل توجهی در حوزه ارزیابی کیفیت تصویر بدون مرجع (NR-IQA) حاصل شده که روی تصاویر طبیعی در محیط‌های جوی عملکرد بسیار خوبی دارند و به‌طور گسترده مورد استفاده قرار می‌گیرند؛ مانند BRISQUE [۱۴]، IL-NIQE [۱۵] و DIIVINE [۱۶] با این حال، به دلیل ویژگی‌های خاص تصویربرداری در زیر آب، این روش‌ها برای ارزیابی تصاویر زیر آب مناسب نیستند. علت اصلی این موضوع، پراکندگی و جذب نور در آب است که باعث کاهش تباين^۸ و دید، تغییر رنگ‌ها به سمت طیف سبز-آبی (به دلیل جذب بیشتر طول موج‌های بلند مانند قرمز و زرد) و همچنین ایجاد پدیده پراکندگی معکوس ۹ می‌شود که باعث ایجاد مه و نویز شده و جزئیات تصویر را کاهش می‌دهد و شفافیت تصویر را کاهش می‌دهد [۱۷].

بنابراین، توسعه روش‌های قابل اعتماد و عملی برای ارزیابی کیفیت تصاویر زیر آب (UIQA) که ویژگی‌های خاص زیر آب را در نظر بگیرند، ضروری است. برای ارزیابی دقیق کیفیت تصاویر زیر آب، برخی روش‌های UIQA ارائه شده‌اند، اما همبستگی بین نمرات پیش‌بینی شده توسط این روش‌ها و نمرات انسانی کافی نبوده و بسیاری از جنبه‌های پیچیده افت کیفیت در این تصاویر همچنان

No-Reference^v
Contrast^h
Backscatter^a
Underwater Image Quality Assessment^l
Efficient Channel Attention¹¹

Image Quality Assessment^l
Subjective^v
Objective^v
ground Truth^h
Full-Reference^a
Reduced-Reference^e

متوسط هر بلوک را محاسبه می‌کند [۲۴]. همان نویسندگان، روش UCIQE را در [۲۵] ارائه کردند که بر ترکیب خطی ویژگی‌های فضای رنگ، اشباع رنگ در فضای CIELab و تباین استوار است. با این حال، این دو روش عمدتاً بر حل مشکل تقویت بیش از حد تمرکز دارند و عوامل مهمی مانند حفظ گرادیان و یکپارچگی ساختاری را که برای ارزیابی دقیق انواع تخریب در تصاویر زیر آب حیاتی هستند، نادیده می‌گیرند.

معیار کیفیت تصویر زیر آب دیگری به نام UIQM سه ویژگی اصلی رنگی بودن، وضوح و تباین را در نظر می‌گیرد [۲۶]. هر یک از این معیارها جنبه خاصی از تخریب تصویر زیر آب را هدف قرار می‌دهند و از ویژگی‌های سیستم بینایی انسان الهام گرفته‌اند. با این حال، به دلیل استفاده از وزن‌دهی ثابت، این روش همبستگی کافی با معیارهای انسانی را به دست نیاورده است. وانگ و همکاران، معیار CCF را معرفی کردند که تباین، کرومینانس و تراکم مه را برای ارزیابی از دست رفتن رنگ به دلیل جذب، تاری ناشی از پراکندگی رو به جلو و مه‌آلودگی ناشی از پراکندگی معکوس ترکیب می‌کند و با استفاده از رگرسیون خطی چندگانه، سه ضریب وزن‌دهی را محاسبه می‌کند [۲۷]. با این حال، این روش بافت، ساختار و پراکندگی رو به جلو را نادیده می‌گیرد.

ویژگی‌های به کار رفته در روش FDUM شامل رنگی بودن، تباین و وضوح هستند که با در نظر گرفتن حساسیت متفاوت انسان به اطلاعات فرکانس بالا و پایین طراحی شده‌اند. همچنین، یک معیار تباین وزنی برای بهبود تباین تحت تأثیر پراکندگی معکوس پیشنهاد شده است و ترکیب این سه معیار به عنوان امتیاز نهایی کیفیت در نظر گرفته می‌شود [۲۸]. اخیراً، روشی به نام UICQA معرفی شده که بر استخراج ویژگی‌های آماری از نقشه انتقال ۴ برای تفکیک کیفیت تصویر تمرکز دارد [۲۹]. این روش همچنین یک ویژگی تباین وزنی مبتنی بر نگاشت انتقال و عمق صحنه ارائه می‌دهد و ویژگی‌های رنگ و بافت را بر اساس ابعاد فراکتالی پیشنهاد می‌کند. روش NUIQ تصاویر RGB زیر آب را به فضای رنگ مخالف ۵ تبدیل می‌کند تا ویژگی‌های روشنایی، رنگ و ساختار استخراج و ارزیابی محاسباتی کیفیت تسهیل شود [۳۰]. بی و همکاران در [۳۱] به طور جامع از آمار فضای رنگ، تباین و وضوح استفاده کردند. با توجه به اینکه تصاویر زیر آب توزیع رایلی دارند، یک ویژگی رنگی محدود در این مطالعه طراحی شده است. همچنین، روشی نوین مبتنی بر اختلاف آنتروپی و آمار گشتاور برای تخمین تباین پس از تقویت پیشنهاد شده است. روش MSAEQA یک نگاشت پیشینه رنگینگی برای شناسایی رنگ‌های مخدوش در آب ایجاد می‌کند و ویژگی‌های آماری چندمقیاسه را از این نگاشت، همراه با الگوهای محلی و جهت‌گیری گرادیان به عنوان ویژگی‌های کیفی برای بافت و ساختار زیر آب استخراج

چارچوب BIQI با استفاده از آمار صحنه‌های طبیعی (NSS)، رویکردی دومرحله‌ای برای تشخیص انواع اعوجاج و ارزیابی کیفیت ارائه می‌دهد [۱۸]. BRISQUE که توسط میتال و همکاران توسعه یافت، روشی سریع و کارآمد برای ارزیابی کیفیت با بهره‌گیری از NSS در حوزه مکانی است [۱۴]. متعاقباً، روش NIQE برای ایجاد مدل آماری چندمقیاسه از تصاویر با کیفیت بدون اتکا به نظرات انسانی معرفی شد که به عنوان مرجعی برای ارزیابی کیفیت تصاویر دیگر عمل می‌کند [۱۹]. نسخه پیشرفته‌تری به نام IL-NIQE، ویژگی‌های محلی تصویر را همراه با ویژگی‌های چندمقیاسه ادغام می‌کند [۱۵]. الگوریتم BLINDS-II توسط سعد و همکاران به عنوان روشی تک‌مرحله‌ای و کارآمد برای ارزیابی کیفیت تصویر بدون مرجع معرفی شد که NSS را از ضرایب تبدیل کسینوسی گسسته (DCT) استخراج می‌کند. این الگوریتم با استفاده از مدل استنتاج بیزی ساده، نمرات کیفیت تصویر را پیش‌بینی می‌کند [۲۰]. مدل DIIVINE مبتنی بر تحلیل ویژگی‌های آماری تصاویر طبیعی در حوزه موجک طراحی شده و با عملکردی جامع در دسته‌بندی‌های مختلف اعوجاج، کیفیت را ارزیابی می‌کند [۱۶].

روش BIQME با استخراج ۱۷ ویژگی شامل تباین، وضوح، و روشنایی، کیفیت بصری را از طریق مازول رگرسیون پیش‌بینی می‌کند [۲۱]. رویکرد GM-LOG از آمار ترکیبی دو ویژگی شناخته شده تباین محلی (گرادیان قدرتی و لاپلاسین گاوسی) برای پیش‌بینی کیفیت بصری استفاده می‌کند [۲۲]. مدل FRIQUEE نیز بر آمار طبیعی تصاویر دارای اعوجاج واقعی تمرکز دارد و با استفاده از «کیف نقشه‌های ویژگی» و آموزش رگرسیون بر پایه پایگاه داده بزرگی از تصاویر واقعی، کیفیت تصاویر با اعوجاج مرکب را پیش‌بینی می‌کند [۲۳].

اگرچه این روش‌های شناخته شده NR-IQA بر جنبه‌های مختلفی از جمله آمارگان تصاویر طبیعی، انواع اعوجاج ۱، وفاداری اطلاعاتی ۲ و ویژگی‌های ادراکی متمرکزند، اما در ارزیابی کیفیت تصویربرداری زیر آب به دلیل ویژگی‌های منحصر به فرد آن با محدودیت‌هایی روبه‌رو می‌شوند. برای غلبه بر این چالش‌ها، معیارهای NR-IQA خاصی برای تصویربرداری زیر آب توسعه یافته‌اند.

۲-۲- روش‌های مخصوص زیر آب مبتنی بر استخراج

ویژگی

نخستین ارزیاب کیفیت تصویر و ویدئوی زیر آب در سال ۲۰۱۴ معرفی شد که از طیف توان تباین لگاریتمی متناسب با محیط‌های زیر آب بهره می‌برد. این معیار مبتنی بر تقسیم‌بندی بلوکی عمل می‌کند و میانگین وزنی تباین خاکستری، روشنایی کلی و گرادیان

در ادامه این روند، مدل ATUIQP [۳۸] با بهره‌گیری از معماری دوشاخه و ترکیب سازوکارهای توجه فضایی و کانالی، کیفیت تصاویر زیر آب را به صورت بدون مرجع و با استفاده همزمان از ویژگی‌های محلی و سراسری پیش‌بینی می‌کند. مدل دیگری با عنوان CET [۳۹] از ساختار مدل ۳ تطبیق محتوا برای استخراج تطبیقی ویژگی‌ها استفاده کرده و با طراحی بلوک‌های مبدل و رگرسیون محتوایی، بهبود قابل توجهی در دقت ارزیابی نشان داده است. همچنین در مطالعه [۴۰] یک معماری عمیق چندشاخه مبتنی بر نقشه‌های روشنایی و بافت و ماژول‌های توجه با نام UIQAVSI پیشنهاد شده که توانایی آن در تفکیک نواحی دچار افت کیفیت و پیش‌بینی نمرات ادراکی به صورت خودکار مورد ارزیابی قرار گرفته است. در مقاله [۴۱] نیز یک معماری عمیق شامل بلوک‌های ادغام درک چندجهته و تشخیص انحراف ارائه شده که به صورت ویژه برای تشخیص نواحی دچار افزایش یا کاهش بیش از حد بهبود طراحی شده است. این روش با بهره‌گیری از مبدل بینایی و ساختار پردازش موازی، امتیاز کیفیت را بر پایه ویژگی‌های استخراج شده از چند مسیر محاسبه می‌نماید.

در مجموع، اگرچه روش‌های مبتنی بر استخراج ویژگی و یادگیری عمیق پیشرفت‌های قابل توجهی در ارزیابی کیفیت تصاویر زیر آب داشته‌اند، هر یک با چالش‌هایی مواجه‌اند. روش‌های سنتی مبتنی بر ویژگی عمدتاً بر معیارهایی نظیر تباین، رنگ و وضوح متمرکزند و غالباً از وزن‌دهی‌های دستی و مدل‌سازی ساده استفاده می‌کنند که دقت و تعمیم‌پذیری آن‌ها را در شرایط پیچیده محیطی محدود می‌سازد. از سوی دیگر، بسیاری از روش‌های عمیق نیازمند معماری‌های پیچیده و پرهزینه هستند و با وجود بهره‌گیری از یادگیری خودکار ویژگی‌ها، اغلب به یک نمای خاص از تصویر مانند تصویر اصلی یا یک ویژگی منفرد بسنده کرده و کمتر به ترکیب هدفمند چند نقشه ویژگی با معماری اختصاصی پرداخته‌اند. روش پیشنهادی این پژوهش با طراحی یک معماری چندشاخه‌ی سبک‌وزن، به‌طور هم‌زمان از تصویر اصلی و دو نقشه‌ی مکمل تیزی و ساختار به‌عنوان ورودی استفاده می‌کند. این انتخاب مبتنی بر تحلیل اهمیت ادراکی این ویژگی‌ها در کاهش کیفیت تصویر زیر آب بوده و برخلاف رویکردهای رایج، آن‌ها را به‌عنوان ورودی‌های صریح در شبکه مدل‌سازی می‌کند. این طراحی، در کنار بهره‌گیری از بلوک توجه و تابع هزینه ترکیبی امکان استخراج مؤثر ویژگی‌های چندبعدی را فراهم کرده و منجر به بهبود دقت ارزیابی کیفیت در شرایط متنوع زیر آب شده است. نتایج تجربی نشان می‌دهد که این مدل از نظر همبستگی با نمرات ادراکی انسانی و پایداری عملکرد، نسبت به بسیاری از روش‌های موجود، به‌ویژه در مجموعه داده‌های متنوع، برتری دارد.

چارچوب اصلی مدل در شکل ۱ نمایش داده شده است.

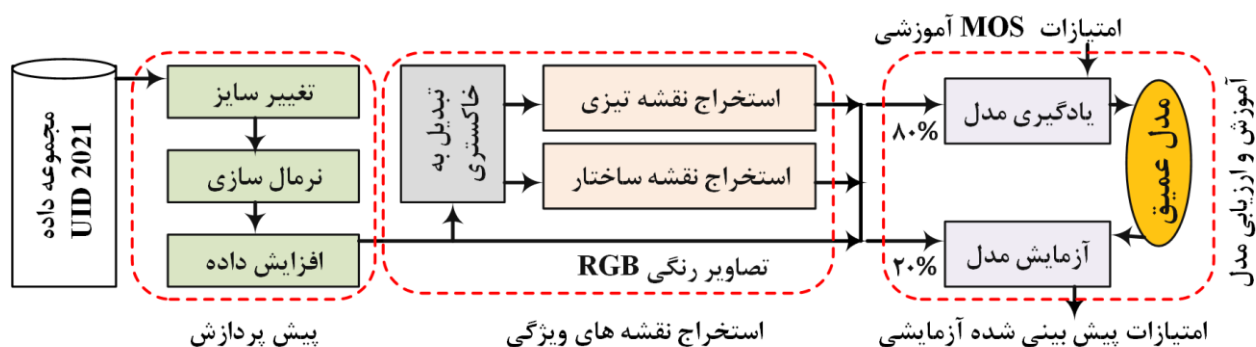
می‌کند. در نهایت، یک توزیع ویبول ۱ برای مدل‌سازی توزیع انرژی تصاویر در فضای رنگی CIELab توسعه یافته است [۱۷]. در مقاله [۳۲]، روشی کور معرفی شده که با ترکیب ویژگی‌های ساختاری کانال‌محور، نرخ پراکندگی رنگ، و اشباع و فام کلی تصویر، امتیاز نهایی کیفیت را با استفاده از رگرسیون خطی پیش‌بینی می‌کند. در مقاله [۳۳] نیز یک مطالعه تجربی در زمینه تصویربرداری قطبی شده زیر آب ارائه شده است که کیفیت تصاویر را در زمینه کاربردهای خاص نوری ارزیابی می‌کند.

با وجود پیشرفت‌های قابل توجه، روش‌های ارزیابی کیفیت تصویر زیر آب همچنان با چالش‌های مهمی روبه‌رو هستند. رویکردهای سنتی مبتنی بر ویژگی اغلب به معیارهایی مانند رنگ، تباین و وضوح متکی‌اند که ممکن است نتوانند عوامل پیچیده تخریب در محیط زیر آب را به درستی شناسایی کنند. این روش‌ها معمولاً از وزن‌دهی دستی برای ترکیب ویژگی‌ها استفاده می‌کنند که منجر به تعمیم ضعیف و ناسازگاری عملکرد، به‌ویژه در مجموعه داده‌های متنوع می‌شود. همچنین، بسیاری از جنبه‌های پیچیده تخریب تصویر زیر آب را نادیده می‌گیرند.

۳-۲- روش‌های مخصوص زیر آب مبتنی بر یادگیری

عمیق

با گسترش تکنیک‌های یادگیری عمیق در پردازش تصویر، مدل‌های مبتنی بر یادگیری عمیق به عنوان گزینه‌ای مناسب برای ارزیابی کیفیت تصویر زیر آب توسعه یافته‌اند. مدل HPUIQA از رویکرد سلسله‌مراتبی برای جداسازی اطلاعات روشنایی و جزئیات بهره می‌برد و با استفاده از MobileNetV3 و ASPP ویژگی‌های بافت را استخراج می‌کند. همچنین، روش جدیدی برای محاسبه تابع هزینه جهت هدایت آموزش مدل معرفی می‌کند [۳۴]. مدل B-FEN از شبکه تعبیه ویژگی ۲ دوطرفه برای ارزیابی کیفیت حذف باران استفاده می‌کند و با سرکوب اطلاعات غیرمرتبط و تعبیه ویژگی‌های حساس به کیفیت، ویژگی‌های پیشرو و پسرو را با یک ماژول ادغام ترکیبی برای درک سراسری و تفاوت‌های محلی ترکیب می‌کند [۳۵]. چارچوب MCOLE از معماری چندجریانی همکاری‌کننده با استفاده از برجسب‌گذاری چندبعدی برای آموزش ویژگی‌های خاص اعوجاج بهره می‌برد. این ویژگی‌ها در یک شبکه پیش‌بینی ادغام شده و دقت و قابلیت تفسیر را افزایش می‌دهد [۳۶]. مدل JLSAU از یک ستون فقرات هرمی با ماژول‌های روشنایی و برجستگی برای استخراج ویژگی‌های چندمقیاسه استفاده می‌کند که با یک ماژول توجه برای ارزیابی بهتر کیفیت ادغام می‌شوند [۳۷].



شکل (۱): چارچوب پیشنهادی برای ارزیابی کیفیت تصاویر زیر آب

در جدول ۱ خلاصه شده‌اند. این روش‌ها به گونه‌ای انتخاب شدند که ویژگی‌های اصلی تصاویر حفظ شده و از ایجاد تغییرات غیرواقعی یا مخرب که ممکن است کیفیت تصویر را به صورت مصنوعی تغییر دهند، مانند تغییر روشنایی یا افزودن نویز، استفاده نشود. استفاده از این روش‌های افزایش داده باعث بهبود عملکرد مدل در داده‌های اعتبارسنجی و کاهش احتمال بیش‌برازش می‌گردد.

جدول (۱): انواع و میزان افزایش داده در روش پیشنهادی

نوع افزایش داده	میزان اعمال شده	توضیح
چرخش تصادفی	$\pm 20\%$	چرخش تصویر تا 20% در جهت تصادفی
برگردان تصادفی	افقی و عمودی	قرینه سازی تصویر در هر دو جهت
بزرگنمایی تصادفی	تا 10%	بزرگنمایی یا کوچک نمایی تا 10%

۳-۱- استخراج نقشه‌های ویژگی

در این بخش، فرآیند استخراج نقشه‌های ویژگی مورد استفاده به عنوان ورودی سه شاخه شبکه عصبی عمیق شرح داده می‌شود. هدف اصلی از استخراج این نقشه‌ها، فراهم آوردن اطلاعات تکمیلی و مکمل برای شبکه است تا علاوه بر تصویر اصلی که شامل اطلاعات رنگ نیز هست، ویژگی‌های مهمی مانند تیزی و ساختار ۳ تصویر نیز به صورت جداگانه به مدل ارائه شود. ارائه سه ورودی مجزا (تصویر اصلی، نقشه تیزی، نقشه ساختار) به سه شاخه مجزا در شبکه، باعث می‌شود هر شاخه بتواند به صورت تخصصی بر روی ویژگی‌های مربوط به خود تمرکز کند. این ساختار چندشاخه‌ای به مدل کمک می‌کند تا اطلاعات بصری و ساختاری تصویر را به طور جامع‌تری تحلیل کند و در نهایت دقت پیش‌بینی کیفیت تصویر را افزایش دهد.

۳- روش پیشنهادی

در این بخش، روش پیشنهادی برای ارزیابی کیفیت تصاویر زیر آب به طور مفصل شرح داده می‌شود. این روش با استفاده از یک شبکه عصبی کانولوشنی (CNN) ۱ چندشاخه و سازوکار توجه انتخابی، ویژگی‌های مهم تصاویر را استخراج کرده و کیفیت آن‌ها را ارزیابی می‌کند. گام‌های چارچوب پیشنهادی شامل: پیش پردازش داده‌ها، استخراج نقشه‌های ویژگی، اعمال سازوکار توجه انتخابی و در نهایت آموزش و تست شبکه CNN پیشنهادی است.

پیش پردازش داده‌ها

مرحله پیش‌پردازش نقش مهمی در بهبود عملکرد مدل دارد. در این مرحله، تصاویر و نقشه‌های ویژگی برای آموزش شبکه عصبی آماده می‌شوند.

- **تغییر اندازه تصاویر:** ابتدا، تصاویر با استفاده از روش درونیایی به اندازه ثابت 255×255 پیکسل تغییر اندازه داده می‌شوند. کاهش ابعاد تصاویر باعث کاهش قابل توجه تعداد پارامترهای ورودی مدل و حجم حافظه مورد نیاز می‌شود و در نتیجه سرعت آموزش و پیش‌بینی مدل افزایش می‌یابد. همچنین، این اندازه جدید به گونه‌ای انتخاب شد که ضمن حفظ اطلاعات ساختاری و محتوایی کافی برای ارزیابی کیفیت تصویر، امکان استفاده از اندازه دسته مناسب و پردازش موازی روی GPU را نیز فراهم کند. بنابراین، این تغییر اندازه یک توازن مناسب میان حفظ ویژگی‌های مهم تصویر و کارایی محاسباتی شبکه ایجاد می‌کند.
- **نرمال‌سازی تصاویر:** در مرحله بعد تصاویر به بازه $[0, 1]$ نرمال‌سازی می‌شوند. نرمال‌سازی تصاویر به بهبود آموزش شبکه عصبی و کاهش حساسیت به تغییرات مقیاس کمک می‌کند. برای نرمال‌سازی تصاویر، مقادیر پیکسل‌ها بر 255 تقسیم می‌شوند.
- **افزایش داده:** برای افزایش تنوع داده‌های آموزشی و تعمیم‌پذیری مدل، از تکنیک‌های افزایش داده استفاده شد که

A. تصویر رنگی اصلی

در نهایت، نقشه ساختار با محاسبه اندازه گرادیان نرمالسازی شده M_{norm} برای استفاده در شاخه سوم شبکه بدست می‌آید:

$$M_{norm}(x, y) = \frac{\sqrt{G_x^2(x, y) + G_y^2(x, y)}}{255} \quad (5)$$

۲-۳- معماری شبکه پیشنهادی

در این پژوهش، مطابق با شکل ۲، یک معماری چندشاخه‌ای مبتنی بر شبکه عصبی کانولوشنی (CNN) و بلوک‌های توجه ECA برای ارزیابی کیفیت تصاویر زیر آب ارائه شده است. مدل پیشنهادی به منظور استخراج ویژگی‌های مکمل، از سه شاخه موازی تشکیل شده که هر یک به ترتیب تصویر اصلی، نقشه تیزی و نقشه ساختار را به عنوان ورودی دریافت می‌کنند. هدف از این طراحی، پردازش جداگانه و هدفمند سه نمایه‌ی مکمل تصویر شامل تصویر اصلی، نقشه تیزی، و نقشه ساختار در مسیرهای مستقل است تا ویژگی‌های ادراکی متفاوت هر کدام به صورت اختصاصی استخراج شود. پردازش جداگانه این سه ورودی سبب می‌شود هر شاخه بتواند الگوهای اختصاصی همان ویژگی را بدون تضعیف متقابل یاد بگیرد؛ در حالی که در شبکه‌های تک‌مسیره یا روش‌های ادغام اولیه، ترکیب زود هنگام کانال‌ها احتمال محو شدن اطلاعات ظریف ادراکی را افزایش می‌دهد. هر شاخه شامل دو لایه کانولوشن متوالی با تعداد فیلترهای ۶۴ و ۱۲۸ و کرنل ۳×۳ است که پس از هر کدام، نرمال‌سازی دسته‌ای ۲ و فعال‌ساز ReLU به کار رفته و به دنبال آن بلوک ECA قرار دارد که با هدف تقویت تطبیقی ویژگی‌های کانالی طراحی شده است. لایه تجمیع بیشینه ۳، پس از اولین کانولوشن و لایه تجمیع سراسری ۴ در انتهای هر شاخه برای فشرده‌سازی ویژگی‌های مکانی استفاده شده است.

پس از استخراج بردار ویژگی از هر شاخه، خروجی‌ها با یکدیگر الحاق شده و یک بردار ویژگی با ابعاد ۳۸۴ تشکیل می‌دهند. این بردار، وارد دو لایه کاملاً متصل ۶ با ۵۱۲ و ۲۵۶ نورون می‌شود که هر یک با نرمال‌سازی دسته‌ای، فعال‌ساز ReLU و لایه حذف تصادفی ۷ همراه هستند تا از بیش‌برازش جلوگیری شود. در نهایت، یک لایه متراکم تک‌خروجی برای پیش‌بینی مقدار کیفیت تصویر قرار گرفته است. جدول ۲ ساختار دقیق هر شاخه و بخش ادغام مدل را نمایش می‌دهد. در این معماری، لایه‌های کانولوشن برای استخراج ویژگی‌های مکانی و کانالی، بلوک‌های ECA برای تقویت توجه به کانال‌های مهم، و لایه‌های تجمیع سراسری برای کاهش ابعاد و فشرده‌سازی اطلاعات به کار رفته‌اند. پس از ادغام سه شاخه، لایه‌های متراکم ۸ و حذف، یادگیری ترکیبی و کاهش بیش‌برازش را تضمین می‌کنند. این ساختار مطابق با

تصویر رنگی در فضای RGB با ابعاد ۲۵۵×۲۵۵×۳ بعنوان ورودی اول و اصلی شبکه مورد استفاده قرار می‌گیرد. اطلاعات رنگی و طیفی، تفاوت رنگ‌ها و ترکیب طیف‌های مختلف نور در تصویر که برای درک بصری و کیفیت رنگ‌ها اهمیت دارد، ویژگی‌های کلی شامل الگوهای بافتی، سایه‌ها، نورپردازی و جزئیات کلی تصویر که در فضای رنگی قابل تشخیص هستند، همچنین اطلاعات سطح بالا مانند اشکال، تباین و ترکیب‌بندی کلی تصویر از طریق تصاویر رنگی در اختیار شاخه اول شبکه قرار می‌گیرند. این شاخه به عنوان منبع اصلی داده، پایه تحلیل کیفیت تصویر است و سایر شاخه‌ها به عنوان مکمل‌های تخصصی برای استخراج ویژگی‌های خاص‌تر عمل می‌کنند.

B. نقشه تیزی تصویر

میزان تیزی تصویر معیاری از وضوح و تیزی لبه‌ها و جزئیات تصویر است که نقش مهمی در ادراک کیفیت تصویر دارد. برای استخراج نقشه تیزی S ، پس از تبدیل تصویر رنگی به خاکستری، از فیلتر لاپلاس ۱ (مشتمل بر مرتبه دوم) برای برجسته‌سازی لبه‌ها و محاسبه تیزی استفاده شد:

$$S(x, y) = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} * I_{gray}(x, y) \quad (1)$$

که در آن '*' عملگر کانولوشن است. مقدار مطلق نقشه S پس از نرمال‌سازی به بازه $[0, 1]$ با نام S_{norm} به عنوان نقشه تیزی نهایی در ورودی شاخه دوم شبکه مورد استفاده قرار می‌گیرد.

$$S_{norm}(x, y) = \frac{|S(x, y)|}{255} \quad (2)$$

C. نقشه ساختار تصویر

نقشه ساختار بیانگر جزئیات ساختاری تصویر مانند لبه‌ها و بافت‌ها است که برای ارزیابی کیفیت تصویر اهمیت دارد. برای استخراج این نقشه، گرادیان‌های افقی و عمودی تصویر توسط اعمال کانال‌های فیلترهای Sobel با تصویر خاکستری محاسبه می‌شوند:

$$G_x(x, y) = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * I_{gray}(x, y) \quad (3)$$

$$G_y(x, y) = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} * I_{gray}(x, y) \quad (4)$$

Concatenate^Δ
Fully Connected^Γ
Dropout^Λ
Dense^Α

Laplacian^Δ
Batch Normalization^Γ
Max Pooling^Λ
Global Average Pooling^Α

تجمیع سراسری به یک بردار تک بعدی تبدیل می‌شوند که نمایانگر اهمیت کلی هر کانال است. سپس به جای استفاده از لایه‌های کاملاً متصل سنگین، یک لایه کانولوشن یک بعدی با کرنل کوچک و متغیر اعمال می‌شود تا وابستگی‌های محلی میان کانال‌ها به صورت دقیق و کم هزینه مدل سازی شود. این کرنل به طور خودکار بر اساس تعداد کانال‌ها تعیین شده و معمولاً کوچک و فرد است تا تمرکز روی کانال‌های همسایه حفظ شود. خروجی این کانولوشن پس از عبور از تابع فعال سازی سیگموید، به عنوان وزن توجه کانالی برای هر کانال در نظر گرفته شده و با ضرب در ویژگی‌های ورودی، کانال‌های مهم تقویت و کانال‌های کم اهمیت تضعیف می‌شوند. نحوه تعیین اندازه کرنل کانولوشن یک بعدی در بلوک ECA بر اساس عبارت (۶) تعیین شده است:

$$k = \left\lfloor \frac{\log_2(C) + b}{\gamma} \right\rfloor_{odd} \quad (6)$$

که در آن C تعداد کانال‌های ورودی و b و γ پارامترهای تنظیمی هستند. $\lfloor \cdot \rfloor_{odd}$ نشان دهنده گرد کردن به نزدیک‌ترین عدد فرد است تا کرنل کانولوشن اندازه‌ی مناسب و متقارن داشته باشد. عملکرد کلی بلوک توجه ECA را می‌توان بصورت زیر بیان کرد:

$$y_c = \sigma \left(\text{Conv1D}(\text{GAP}(x)) \right) \cdot x_c \quad (7)$$

که در آن x ورودی بلوک با ابعاد (H, W, C) ، GAP عملیات تجمیع سراسری، Conv1D لایه کانولوشن یک بعدی با کرنل k ، و σ تابع فعال‌ساز سیگموید است. γ خروجی بلوک توجه با وزن‌دهی کانال‌ها است. در پیاده سازی مدل، ابتدا ویژگی‌های مکانی هر کانال با لایه تجمیع سراسری خلاصه می‌شوند. سپس داده‌ها تغییر شکل یافته و به لایه کانولوشن یک بعدی با کرنل تعیین شده وارد می‌شوند تا وابستگی‌های محلی بین کانال‌ها استخراج گردد. خروجی این لایه پس از تابع سیگموید به عنوان وزن توجه کانالی به ویژگی‌های ورودی اعمال می‌شود. این طراحی موجب می‌شود که بلوک ECA بتواند بصورت کارآمد و با هزینه محاسباتی پایین، کانال‌های مهم را تقویت کند.

استفاده از این بلوک در معماری چندشاخه‌ای ما اهمیت ویژه‌ای دارد، زیرا هر شاخه، ورودی متفاوتی (تصویر اصلی، نقشه تیزی و نقشه ساختار) دارد و ویژگی‌های استخراج شده ماهیت متفاوتی دارند. بلوک ECA با تقویت پویا و دقیق کانال‌های کلیدی در هر شاخه، امکان بهره‌گیری بهتر از اطلاعات مکمل را فراهم می‌کند. علاوه بر این، به دلیل ساختار ساده و سبک این بلوک، افزایش قابل توجهی در پارامترها و هزینه محاسباتی مدل ایجاد نمی‌کند که این موضوع در معماری چندشاخه‌ای با سه مسیر پردازشی موازی، بسیار حیاتی است. بنابراین، بهره‌گیری از بلوک توجه کانالی کارآمد، ضمن حفظ سادگی و کارایی، باعث تقویت

رویکردهای نوین ارزیابی کیفیت تصویر، امکان بهره‌گیری از اطلاعات مکمل چندمنظوره را فراهم می‌سازد.

جدول (۲): ساختار دقیق هر شاخه و بخش ادغام مدل

نام لایه	نوع لایه	خروجی لایه	پارامترها/توضیحات
Input	ورودی	(255, 255, 3/1)	(255, 255, 3) یا (255, 255, 1)
Conv2D	کانولوشن دو بعدی	(255, 255, 64)	فیلتر=۶۴، کرنل=۳×۳، padding='same'
Batch Normalization	نرمال سازی دسته‌ای	(255, 255, 64)	-
ReLU	فعال سازی	(255, 255, 64)	-
ECA Block	بلوک توجه	(255, 255, 64)	k متناسب با تعداد کانال
Max Pooling2D	تجمیع بیشینه	(127, 127, 64)	pool_size=2×2
Conv2D	کانولوشن دو بعدی	(127, 127, 128)	فیلتر=۱۲۸، کرنل=۳×۳، padding='same'
Batch Normalization	نرمال سازی دسته‌ای	(127, 127, 128)	-
ReLU	فعال سازی	(127, 127, 128)	-
ECA Block	بلوک توجه	(127, 127, 128)	k متناسب با تعداد کانال
Global Average Pooling2D	تجمیع سراسری	(128)	-
Concatenate	الحاق ویژگی‌ها	(384)	-
Dense	لایه کاملاً متصل	(512)	واحد=۵۱۲، بدون باپاس
Batch Normalization	نرمال سازی دسته‌ای	(512)	-
ReLU	فعال سازی	(512)	-
Dropout	حذف تصادفی	(512)	نرخ=۰/۵
Dense	لایه کاملاً متصل	(256)	واحد=۲۵۶، بدون باپاس
Batch Normalization	نرمال سازی دسته‌ای	(256)	-
ReLU	فعال سازی	(256)	-
Dropout	حذف تصادفی	(256)	نرخ=۰/۵
Dense	لایه خروجی	(1)	واحد=۱

۳-۳- بلوک توجه کانالی کارآمد

در معماری پیشنهادی، برای بهبود استخراج ویژگی‌های مهم و افزایش دقت پیش‌بینی کیفیت تصاویر زیر آب، از بلوک توجه ECA استفاده شده است [۴۲]. این بلوک توجه، با هدف مدل سازی وابستگی‌های میان کانال‌های ویژگی به روشی سبک و مؤثر طراحی شده است. ابتدا ویژگی‌های مکانی هر کانال با استفاده از لایه

در برخی موارد موجب افزایش زمان آموزش یا نوسان خطا می‌شود. بدین ترتیب ترکیب نهایی جدول ۴ با هدف دستیابی به تعادل میان دقت، پایداری و کارایی محاسباتی تثبیت شد و در تمام آزمایش‌های گزارش شده مورد استفاده قرار گرفت.

ویژگی‌های کلیدی و افزایش دقت پیش‌بینی می‌شود. در جدول ۳ ساختار لایه‌های تشکیل دهنده بلوک ECA به همراه ابعاد ورودی و خروجی هر لایه آورده شده است.

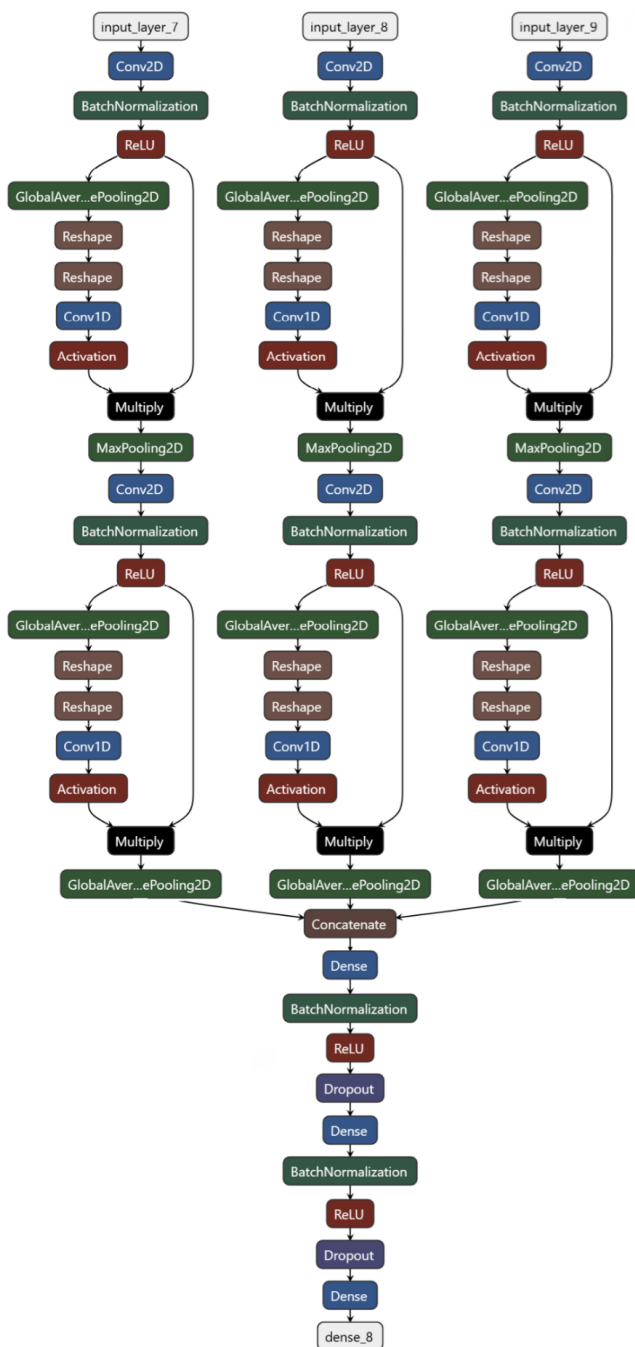
جدول (۳): ساختار لایه‌های تشکیل دهنده بلوک توجه کانالی ECA-Net

لایه	نوع لایه	پارامترها/توضیحات	ابعاد خروجی
ورودی	تسور ویژگی	(H, W, C)	(H, W, C)
Global Average Pooling	تجمع سراسری	-	$(1, C)$
Reshape	تغییر شکل	از $(C, 1)$ به $(1, C)$	$(1, C)$
Conv1D	کانولوشن یک‌بعدی	کرنل k (متغیر بر مبنای C)، padding=same	$(1, C)$
Sigmoid	تابع فعال‌سازی	-	$(1, C)$
ضرب کانالی	وزن‌دهی کانال‌ها	ضرب وزن توجه در ورودی	$(1, C)$

۳-۴- تنظیم پارامترها و هایپرپارامترها

تنظیم دقیق این پارامترها نقشی کلیدی در پایداری آموزش، جلوگیری از بیش‌برازش و بهبود عملکرد نهایی مدل ایفا می‌کند. در این پژوهش، برای تنظیم پارامترها و هایپرپارامترهای مدل، دو رویکرد مکمل در پیش گرفته شد. ابتدا محدوده‌ای از مقادیر متداول بر مبنای مطالعات اخیر در ارزیابی کیفیت تصویر زیر آب انتخاب شد؛ سپس برای سه پارامتر حساس (اندازه دسته ۱، نرخ یادگیری و حذف تصادفی) چند مقدار نماینده آزموده شد تا در چارچوب منابع محاسباتی محدود، ترکیبی پایدار و کارآمد به دست آید. مقایسه‌ی این پیکربندی‌های محدود بر روی مجموعه‌ی اعتبارسنجی پنج‌تایی نشان داد که اندازه دسته ۳۲ در برابر اندازه‌های کوچک‌تر گرادیان پایدارتر و در برابر اندازه‌های بزرگ‌تر حافظه‌ی پردازنده گرافیکی را بهینه می‌کند؛ نرخ یادگیری اولیه ۱۰-۴ سریع‌ترین همگرایی بدون نوسان شدید را فراهم کرد و مقدار ۰/۵ برای حذف تصادفی بهترین تعادل میان کاهش بیش‌برازش و حفظ دقت را ایجاد نمود. تفاوت میانگین PLCC در این آزمون‌های محدود به‌طور متوسط ۲ درصد و SRCC حدود ۱/۵ درصد به نفع ترکیب نهایی بود؛ بنابراین مقادیر جدول ۴ به عنوان تنظیم نهایی انتخاب شد.

سایر پارامترها از جمله تعداد تکرار آموزش (۲۰۰ دوره)، توقف زودهنگام ۳ با آستانه صبر ۲۰ دوره، بهینه‌ساز AdamW [۴۳] با وزن‌زدایی ۱۰-۴ و راهبرد کاهش پله‌ای ۴ نرخ یادگیری (ضریب ۰/۵ پس از ۱۰ دوره بدون بهبود) بر اساس توصیه‌های مقالات مرجع و آزمون یک یا دو مقدار جایگزین انتخاب شدند؛ نتایج نشان دادند تغییر این مقادیر به مقادیر بزرگ‌تر یا کوچک‌تر، بهبود معناداری در معیارهای PLCC و SRCC ایجاد نمی‌کند و



شکل (۲): معماری شبکه عمیق ارزیابی کیفیت پیشنهادی برای تصاویر زیر آب

که در آن ضرایب λ_1 و λ_2 جهت توازن بین دو جزء تابع هزینه به ترتیب برابر با $0/5$ تنظیم شدند. این ترکیب تابع هزینه باعث می‌شود مدل ضمن کاهش خطای عددی (RMSE)، همبستگی پیش‌بینی‌ها با نمرات انسانی را نیز بهینه کند که در نهایت منجر به بهبود عملکرد کلی مدل می‌شود.

$$L_{Huber}(y, \hat{y}) = \begin{cases} \frac{1}{2}(y, \hat{y})^2 & \text{if } |y, \hat{y}| \leq \delta \\ \delta \left(|y, \hat{y}| - \frac{1}{2}\delta \right) & \text{otherwise} \end{cases} \quad (9)$$

برای تحلیل روند آموزش، نمودار کاهش مقدار تابع هزینه در داده‌های آموزش و اعتبارسنجی رسم شده است (شکل ۳). این نمودار نشان‌دهنده همگرایی مناسب مدل و عدم وجود بیش‌برازش است. همچنین، نمودار پراکندگی ۶ پیش‌بینی‌های مدل در برابر نمرات واقعی کیفیت در شکل ۴ ارائه شده که همبستگی بالای پیش‌بینی‌ها با نمرات مرجع را به وضوح نشان می‌دهد.

۴- آزمایشات و نتایج

۴-۱- مجموعه داده‌ها

UID2021: مجموعه داده معیار بهبود تصاویر زیر آب (UID2021) یک مجموعه داده با دقت بالا و هدفمند است که برای ارزیابی کیفیت تصاویر زیر آب طراحی شده است [۴۷]. این مجموعه شامل ۶ زیرمجموعه از صحنه‌های زیر آب مختلف (شامل تصاویر با رنگ‌های مایل به سبز، مایل به آبی، آبی-سبز، کدر، کم‌نور و مه‌آلود) است که در مجموع ۶۰ تصویر مرجع از منابع مختلف دنیای واقعی انتخاب و به اندازه 384×512 برش خورده‌اند تا وضوح مناسبی روی نمایشگر داشته باشند. از این ۶۰ تصویر مرجع، تعداد ۹۶۰ تصویر بهبود یافته توسط ۱۵ الگوریتم پیشرفته شامل روش‌های رفع تاری، اصلاح رنگ، حذف مه، افزایش تباين و غیره تولید شده‌اند. نمونه‌هایی از تصاویر این مجموعه داده در شکل ۵ نمایش داده شده‌اند [۴۷]. برای ارائه چارچوب ارزیابی قوی، هر تصویر در این مجموعه با امتیاز میانگین نظرات ادراکی کاربران (MOS) ۸ برچسب‌گذاری شده است. این معیار کیفیت انسانی، میانگین نمرات نرمال شده‌ای است که توسط گروهی متشکل از ۵۲ ناظر انسانی با استفاده از روش مقایسه زوجی ۹ به تصاویر اختصاص یافته است. MOS کیفیت ادراک شده تصاویر را با در نظر گرفتن عوامل کلیدی مانند وضوح، دقت رنگ، تباين، بافت و قابلیت مشاهده منعکس می‌کند. با ارائه این امتیازهای استاندارد، مجموعه UID2021 یک معیار

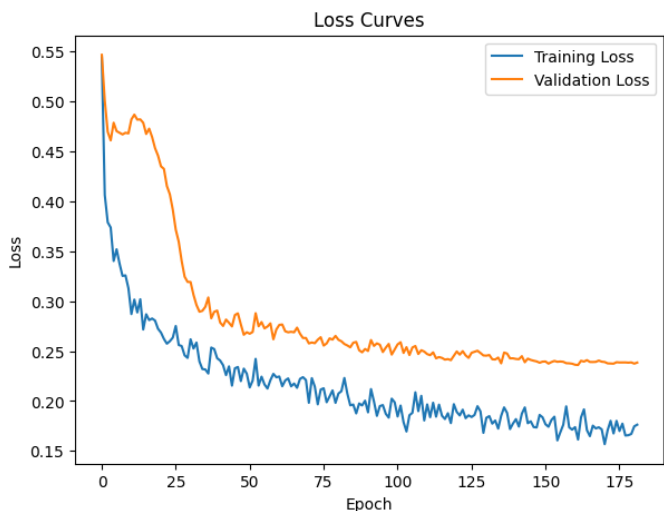
۵-۳- آموزش و ارزیابی مدل

آموزش و ارزیابی مدل پیشنهادی با هدف پیش‌بینی دقیق کیفیت تصاویر زیر آب بر اساس داده‌های مرجع انجام شده است. برای اطمینان از صحت عملکرد و افزایش قابلیت تعمیم مدل، فرآیند آموزش با استفاده از اعتبارسنجی متقابل ۵-تایی ۱ پیاده‌سازی شده است. در این روش، مجموعه داده به صورت تصادفی و با مقدار ثابت برای بذر تصادف ۲ به پنج زیرمجموعه مساوی تقسیم می‌شود. در هر تکرار، یکی از این زیرمجموعه‌ها به عنوان مجموعه آزمون و چهار بخش دیگر به عنوان مجموعه آموزش مورد استفاده قرار می‌گیرند. این تقسیم‌بندی تصادفی اما تکرارپذیر، به مدل امکان می‌دهد که با داده‌های متنوع‌تری آموزش دیده و در برابر داده‌های ناشناخته ارزیابی شود. در هر فولد، مدل با داده‌های آموزش تمرین داده شده و عملکرد آن بر روی داده‌های آزمون ارزیابی می‌شود. بهترین نسخه مدل بر اساس کمترین مقدار تابع هزینه روی داده اعتبارسنجی ذخیره می‌گردد. در پایان، نتایج نهایی به صورت میانگین بر حسب هر معیار از پنج تکرار گزارش می‌شود تا تأثیر نوسانات ناشی از تقسیم‌بندی کاهش یافته و ارزیابی دقیق‌تر و پایدارتری از توانایی تعمیم مدل ارائه گردد. نحوه تنظیم سایر پارامترها و هایپرپارامترهای آموزش، از جمله تعداد تکرار، اندازه دسته، راهبرد توقف زود هنگام و تنظیم نرخ یادگیری، در بخش قبل به تفصیل توضیح داده شده است.

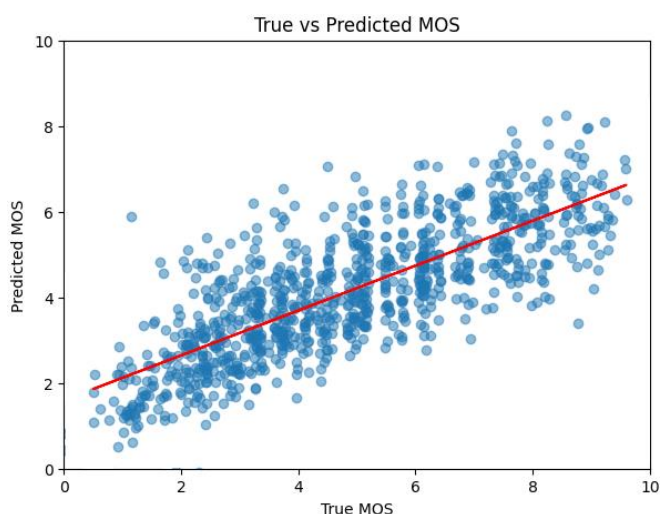
۱-۵-۳- تابع هزینه ترکیبی پیشنهادی

استفاده از توابع هزینه ترکیبی، در پژوهش‌های اخیر بسیار مورد توجه قرار گرفته است [۴۵، ۴۴]. در مدل پیشنهادی، تابع هزینه ۳ مدل، ترکیبی از تابع هابر ۴ [۴۶] و ضریب همبستگی پیرسون ۵ (PLCC) است تا همزمان دو هدف کاهش خطای عددی و افزایش همبستگی پیش‌بینی‌ها با نمرات انسانی را دنبال کند. این تابع به صورت میانگین روی هر دسته کوچک محاسبه می‌شود، به این صورت که ابتدا ضریب همبستگی خطی پیرسون (PLCC) بین خروجی مدل و نمرات مرجع MOS برای تمام نمونه‌های موجود در یک دسته محاسبه شده و سپس مقدار PLCC-1 به عنوان بخش دوم تابع هزینه لحاظ می‌گردد. این ترکیب باعث می‌شود مدل نه تنها خطاهای عددی را کاهش دهد، بلکه به طور مستقیم به افزایش همبستگی با نمرات انسانی نیز هدایت شود. این تابع به صورت زیر تعریف می‌شود:

$$L = \lambda_1 \times (1 - PLCC(y, \hat{y})) + \lambda_2 \times L_{Huber}(y, \hat{y}) \quad (8)$$



شکل (۳): نمودار تغییرات تابع هزینه برای مجموعه آموزش و اعتبار سنجی در حین روند آموزش



شکل (۴): نمودار پراکندگی پیش‌بینی‌های مدل در برابر نمرات واقعی کیفیت

۲-۴- معیارهای ارزیابی

برای ارزیابی اثربخشی روش‌های IQA، از چند معیار عملکرد پرکاربرد استفاده می‌شود که شامل ضریب همبستگی خطی پیرسون (PLCC) ۴، ضریب همبستگی رتبه‌ای اسپیرمن (SROCC) ۵، ضریب همبستگی رتبه‌ای کندال (KROCC) ۶ و جذر میانگین مربعات خطا (RMSE) ۷ است. این معیارها با مقایسه خروجی مدل با نمرات مرجع (MOS) محاسبه شده‌اند. لازم به ذکر است که این نمرات ادراکی به صورت مستقل توسط پژوهشگران این مقاله جمع‌آوری نشده‌اند، بلکه از مقادیر موجود در مجموعه‌های داده UID2021 و SAUD 2.0 استفاده شده است.

مرجع برای سنجش اثربخشی الگوریتم‌های مختلف بهبود تصاویر زیر آب بر اساس ادراک انسانی فراهم می‌آورد. SAUD 2.0: پایگاه داده SAUD 2.0 یک مجموعه مرجع جامع و هدفمند برای ارزیابی کیفیت تصاویر زیر آب پس از بهبود است [۳۶]. این پایگاه داده شامل ۲۰۰ تصویر خام زیر آب است که از مجموعه‌های شناخته‌شده UIEB و RUIE استخراج شده‌اند. هر یک از این تصاویر خام با استفاده از ۱۲ الگوریتم مطرح در حوزه بهبود تصاویر زیر آب پردازش شده‌اند و در مجموع ۲۴۰۰ تصویر بهبود یافته تولید شده است. این مجموعه تنوع بالایی از اعوجاج‌های رایج در تصاویر زیر آب و نتایج متنوع حاصل از الگوریتم‌های بهبود را پوشش می‌دهد. یکی از ویژگی‌های برجسته SAUD 2.0، وجود برجسب‌گذاری ادراکی چندبعدی ۲ است که نه تنها امتیاز ادراکی کلی (MOS) را شامل می‌شود، بلکه به صورت مجزا امتیازهایی برای وفاداری رنگ و وضوح ساختاری نیز ارائه می‌دهد. این امتیازها بر اساس استاندارد ITU-R BT.500-14 و با استفاده از مقیاس پنج‌درجه‌ای ACR ۳ توسط ۲۲ ارزیاب متخصص تعیین شده‌اند. چنین ساختار برجسب‌گذاری غنی، امکان ارزیابی دقیق‌تری از جنبه‌های مختلف کیفیت ادراکی مانند تصحیح رنگ، وضوح و شفافیت ساختاری را فراهم می‌سازد. توزیع متوازن امتیازهای MOS در سه بُعد یادشده، موجب تنوع و استحکام آماری مجموعه SAUD 2.0 شده و آن را به یکی از معتبرترین معیارهای ارزیابی در پژوهش‌های مربوط به ارزیابی کیفیت تصاویر زیر آب تبدیل کرده است. جدول ۵، مقایسه‌ای اجمالی میان دو مجموعه SAUD 2.0 و UID2021 ارائه می‌دهد.

جدول (۵): مقایسه‌ای مختصر بین SAUD 2.0 و UID2021

UID2021	SAUD 2.0	ویژگی
۶۰	۲۰۰	تعداد تصاویر اصلی
۹۶۰ (توسط ۱۵ روش UIE)	۲۴۰۰ (توسط ۱۲ روش UIE)	نسخه‌های بهبود یافته
MOS کلی (یک عدد برای هر تصویر)	سه نوع MOS (کلی، رنگ، وضوح)	نوع MOS
۵۲	۲۲ (مطابق استاندارد ITU-R BT.500-14)	تعداد ارزیاب‌ها
مقایسه زوجی و مرتب‌سازی	امتیازدهی مطلق (ACR) با مقیاس ۵ امتیازی	روش امتیازدهی
۶ دسته (سبز، آبی، آبی-سبز، کدر، کم‌نور، مه‌آلود)	صحنه‌های متنوع از دیپاست‌های UIEB و RUIE	نوع صحنه‌ها
ارزیابی کلی بهبود تصاویر زیر آب	ارزیابی چندبعدی بهبود زیر آب از نظر رنگ و وضوح	هدف

^۵Spearman Rank Order Correlation Coefficient

^۶Kendall Tau Rank Correlation Coefficient

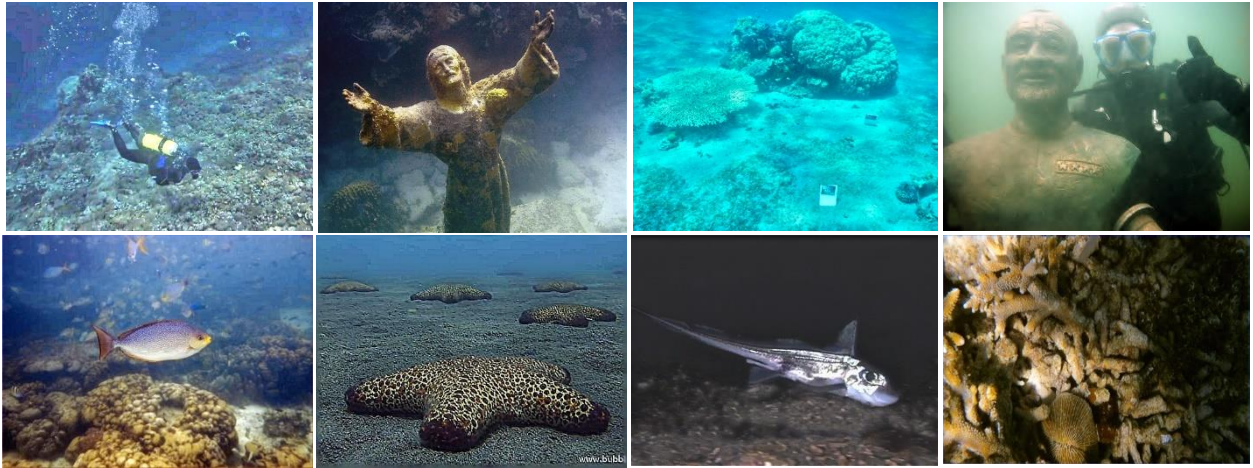
^۷Root Mean Squared Error

^۱Subjectively Annotated Underwater Dataset

^۲Multi-Dimensional Subjective Annotation

^۳Absolute Category Rating

^۴Pearson Linear Correlation Coefficient



شکل (۵): نمونه‌هایی از تصاویر موجود در مجموعه داده‌های مورد استفاده

تست و چهار بخش دیگر به عنوان مجموعه آموزش مورد استفاده قرار گرفته‌اند. در نهایت، میانگین نتایج به‌دست‌آمده از مجموعه تست هر پنج فولد بر حسب PLCC, SRCC, KRCC, RMSE به عنوان معیار نهایی عملکرد مدل گزارش شده است. این رویکرد موجب می‌شود ارزیابی مدل از پایداری و قابلیت تعمیم بیشتری برخوردار باشد و نتایج ارائه‌شده نمایانگر عملکرد واقعی مدل بر روی داده‌های ناشناخته باشند.

عملکرد مدل پیشنهادی بر اساس هر یک از این معیارها با هفت روش عمومی ارزیابی کیفیت تصویر شامل IL-NIQE, GM-, BIQME, BRISQUE, DIIVINE, BLINDS-II, FRIQUEE, LOG و همچنین سیزده روش ویژه تصاویر زیر آب شامل NUIQ, UICQA, FDUM, CCF, UIQM, UCIQE, JLSAU, MCOLE, B-FEN, HPUIQA, MSAEQA, Yi در جدول ۶ مقایسه شده است. برای برجسته‌سازی بهترین عملکردها، دو نتیجه برتر هر ستون مشخص شده‌اند؛ به طوری که بهترین نتیجه به صورت پررنگ و دومین نتیجه برتر به صورت ایتالیک نمایش داده شده است. سمت راست جدول مربوط به نتایج روی مجموعه داده UID2021 و سمت چپ جدول مربوط به نتایج روی SAUD2.0 است.

در سمت راست جدول ۶، همان‌طور که مشاهده می‌شود، در میان روش‌های عمومی، روش FRIQUEE با PLCC برابر با ۰/۷۱۳۴ و RMSE برابر با ۱/۴۹۸۴ بالاترین عملکرد را روی مجموعه UID2021 دارد، اما همچنان نسبت به روش پیشنهادی ضعیف‌تر عمل می‌کند. سایر روش‌های قابل توجه مانند BRISQUE و BIQME نیز عملکرد مناسبی از خود نشان می‌دهند، اما نتایج آن‌ها از روش پیشنهادی پایین‌تر است. در بررسی روش‌های ویژه تصاویر زیر آب، به وضوح مشخص است که این روش‌ها به طور کلی عملکرد بهتری نسبت به روش‌های عمومی دارند و این موضوع نشان‌دهنده تناسب بالاتر آن‌ها برای ارزیابی کیفیت تصاویر در شرایط زیر آب است. در میان این روش‌ها، مدل MCOLE عملکرد بسیار دقیقی در ارزیابی کیفیت تصاویر

این مجموعه داده‌ها از پیش شامل نمرات ادراکی انسانی هستند که توسط ارزیابان خبره و بر اساس استانداردهای معتبر بین‌المللی تهیه شده‌اند. در ادامه تعریف ریاضی این چهار معیار ارائه شده است.

$$PLCC = \frac{\sum_{i=1}^N (p_i - \bar{p})(m_i - \bar{m})}{\sqrt{\sum_{i=1}^N (p_i - \bar{p})^2} \cdot \sqrt{\sum_{i=1}^N (m_i - \bar{m})^2}} \quad (10)$$

$$SRCC = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)} \quad (11)$$

$$KRCC = \frac{(T - C)^2}{N(N - 1)} \quad (12)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (m_i - p_i)^2} \quad (13)$$

که در آن N تعداد تصاویر ارزیابی شده، p_i امتیاز پیش‌بینی شده توسط مدل برای تصویر i ام، m_i امتیاز MOS برای همان تصویر و \bar{p} و \bar{m} میانگین‌های مربوطه هستند. همچنین d_i اختلاف رتبه‌ی پیش‌بینی و اصلی برای تصویر i ، و T و C به ترتیب تعداد جفت‌های هم‌رتبه و ناهماهنگ در رتبه‌بندی هستند. مقادیر بالای PLCC و SROCC نشان‌دهنده همبستگی قوی و سازگاری بین نمرات پیش‌بینی شده توسط مدل و امتیازات MOS است. همچنین، مقدار بالاتر KROCC بیانگر توافق بهتر در رتبه‌بندی تصاویر است، در حالی که مقدار کمتر RMSE نشان می‌دهد که پیش‌بینی‌های روش‌های IQA دقت بیشتری دارند. این معیارها به طور کلی دید جامعی از میزان تطابق روش‌های ارزیابی کیفیت تصویر با قضاوت‌های ادراکی انسان ارائه می‌دهند و در نهایت راهنمایی برای بهبود الگوریتم‌ها محسوب می‌شوند.

۳-۴- ارزیابی کارایی

در این پژوهش، برای ارزیابی عملکرد مدل پیشنهادی از اعتبارسنجی متقابل پنج‌تایی (5-fold) استفاده شده است. بدین صورت که داده‌های موجود در هر مجموعه داده، به پنج بخش تقسیم شده و در هر تکرار، یکی از بخش‌ها به عنوان مجموعه

پیشنهادی بهبود چشمگیری نسبت به تکنیک‌های موجود ارزیابی کیفیت تصویر ارائه می‌دهد.

ارزیابی عملکرد مشابهی بر روی مجموعه داده SAUD 2.0 انجام شد و نتایج در سمت چپ جدول ۶ گزارش شده است. روش‌های IQA عمومی عموماً عملکرد ضعیف‌تری را در این مجموعه داده نشان دادند، به استثنای FRIQUEE که به PLCC برابر با ۰/۷۸۸۲ دست یافت اما همچنان نسبت به مدل‌های خاص زیر آب پایین‌تر بود. در میان روش‌های خاص زیر آب، MCOLE با PLCC برابر با ۰/۸۸۳۸ بالاترین عملکرد را به دست آورد و پس از آن UIQAVSI با PLCC برابر با ۰/۸۷۰۹ قرار گرفت که هر دو با اختلاف قابل توجهی از مدل‌های عمومی بهتر عمل کردند. روش پیشنهادی ما با PLCC برابر با ۰/۸۷۵۶، SRCC برابر با ۰/۸۶۳۱، KRCC برابر با ۰/۶۷۵۸ و RMSE برابر با ۱۰/۳۰۶۶ به نتایج رقابتی دست یافت که بسیار نزدیک به این دو روش اما کمی پایین‌تر در PLCC است، در حالی که عملکرد کلی بسیار رقابتی خود را حفظ می‌کند.

این مشاهده، توانایی تعمیم مدل پیشنهادی را در مجموعه داده‌های مختلف تأیید می‌کند. نکته قابل توجه این است که با وجود اینکه SAUD 2.0 حاوی تصاویر متنوع‌تر و حاشیه‌نویسی‌های چندبعدی MOS است، روش پیشنهادی، قابلیت پیش‌بینی قوی خود را حفظ می‌کند و اثربخشی آن را فراتر از مجموعه داده‌های UID2021 بیشتر تأیید می‌کند.

زیر آب نشان داده است. با این حال، مدل پیشنهادی در مقایسه با آن، به‌ویژه از نظر معیارهای همبستگی ادراکی، نتایج بهتری ارائه می‌دهد. این برتری را می‌توان به چند عامل کلیدی نسبت داد: نخست، استفاده همزمان از سه ورودی مکمل شامل تصویر اصلی، نقشه ساختار و نقشه تیزی، اطلاعات ادراکی غنی‌تری را در اختیار مدل قرار می‌دهد. دوم، ساختار چندشاخه‌ای طراحی شده امکان پردازش هدفمند و مستقل هر یک از این ویژگی‌ها را فراهم کرده و با بهره‌گیری از بلوک‌های توجه ECA، توجه شبکه به ویژگی‌های مؤثر در هر کانال بهبود یافته است. در مقابل، روش MCOLE علی‌رغم طراحی چندجریانی، از ورودی‌های ساده‌تری استفاده کرده و ترکیب ویژگی‌های آن عمدتاً به‌صورت متراکم در انتهای شبکه صورت می‌گیرد که می‌تواند منجر به تداخل یا تضعیف ویژگی‌های خاص شود. علاوه بر این، مدل پیشنهادی ساختار سبک‌تری دارد که منجر به کارایی بهتر در عین حفظ دقت شده است. در مجموع، روش پیشنهادی عملکرد برتری نسبت به اکثر مدل‌ها از خود نشان داده و مقادیر PLCC و SROCC بالایی را در میان روش‌های ارزیابی شده به دست آورده است که بیانگر همبستگی قوی با کیفیت ادراکی تصاویر است. همچنین با دستیابی به RMSE پایین، دقت بالایی در پیش‌بینی کیفیت تصویر ارائه می‌دهد و این موضوع نشان می‌دهد که مدل پیشنهادی برای کاربردهای عملی قابل اعتمادتر است. نتایج به وضوح بیانگر آن است که روش

جدول (۶): ارزیابی کیفیت روش‌های مختلف ارزیابی کیفیت تصویر روی مجموعه داده SAUD2.0 و UID2021

SAUD2.0				UID2021				روش	نوع روش
RMSE	KRCC	SROCC	PLCC	RMSE	KRCC	SROCC	PLCC		
-	-	-	-	1.9121	0.2419	0.4630	0.4644	IL-NIQE [۱۵]	روش‌های عمومی
19.4931	0.3060	0.4449	0.4816	1.7934	0.3688	0.5216	0.5451	BLIINDS-II [۲۰]	
17.9813	0.3523	0.5016	0.5151	1.6716	0.4363	0.6112	0.6264	DIIVINE [۱۶]	
17.9813	0.3533	0.5018	0.5276	1.6407	0.4623	0.6343	0.6439	BRISQUE [۱۴]	
17.6922	0.3168	0.4553	0.4938	1.6330	0.4566	0.6358	0.6491	BIQME [۲۱]	
20.0512	0.2644	0.3828	0.4084	1.7699	0.3779	0.5384	0.5624	GM-LOG [۲۲]	
13.1197	0.5843	0.7738	0.7882	1.4984	0.5035	0.7026	0.7134	FRIQUEE [۲۳]	
19.9196	0.2597	0.3719	0.3674	1.6969	0.4353	0.6030	0.6254	UCIQE [۲۵]	روش‌های تخصصی
19.9120	0.1993	0.2902	0.3498	1.7691	0.3807	0.5404	0.5816	UIQM [۲۶]	
20.9302	0.1088	0.1624	0.1664	1.9159	0.3097	0.4358	0.4713	CCF [۲۷]	
20.1235	0.1785	0.2613	0.2815	1.6756	0.4515	0.6269	0.6374	FDUM [۲۸]	
-	-	-	-	1.4358	0.5223	0.7363	0.7448	UICQA [۲۹]	
14.1152	0.5471	0.7480	0.7413	1.4762	0.5293	0.7168	0.7266	NUIQ [۳۰]	
12.3297	0.6326	0.8164	0.8174	1.4404	0.5420	0.7347	0.7423	Yi [۳۱]	
-	-	-	-	1.3683	0.5631	0.7604	0.7711	MSAEQA [۱۷]	روش‌های زیر آب
-	-	-	-	1.4718	0.5445	0.7497	0.7362	HPUIQA [۳۴]	
10.8642	0.6595	0.8441	0.8512	1.3696	0.5732	0.7674	0.7713	B-FEN [۳۵]	
9.7132	0.6887	0.8748	0.8838	1.2964	0.6024	0.7915	0.7977	MCOLE [۳۶]	
-	-	-	-	1.4512	0.5509	0.7467	0.7353	JLSAU [۳۷]	
-	0.6902	0.8703	0.8709	-	-	-	-	UIQAVSI [۴۰]	
10.4927	0.6671	0.8505	0.8632	1.3175	0.6035	0.7857	0.8060	پیشنهادی	

۴-۴- ارزیابی متقابل

برای ارزیابی قابلیت تعمیم مدل پیشنهادی، یک آزمایش ارزیابی متقابل ۱ انجام شد. بدین منظور به طور مشخص، مدل را روی مجموعه SAUD2.0 آموزش دادیم و سپس روی مجموعه تصاویر UID2021 تست کردیم. پیش از انجام آموزش، بازه امتیازات SAUD2.0 به بازه امتیازات مجموعه UID2021 با مقیاس کردن، تطبیق داده شد تا سازگاری بهتری در مقایسه نتایج حاصل شود. نتایج به دست آمده در جدول ۷ با نتایج اعلام شده توسط دیگر روش‌ها مقایسه شده است. چنانکه مشاهده می‌شود ضرایب همبستگی PLCC برابر با ۰/۵۲۳۱، SRCC برابر با ۰/۵۲۱۰، KRCC برابر با ۰/۳۵۴۸ و مقدار RMSE برابر با ۱/۸۱۲۷ بدست آمده است. این نتایج بیانگر عملکرد قابل قبول مدل تحت شرایط ارزیابی متقابل و نشان‌دهنده توانایی مدل در تعمیم به داده‌های مختلف و جدید است. بنابراین، آزمایش انجام شده، قوت و پایداری رویکرد پیشنهادی ما را در مواجهه با داده‌های خارج از مجموعه آموزش تایید می‌کند.

جدول (۷): نتایج ارزیابی متقابل آموزش روی SAUD2.0 و آزمایش

روی UID2021

RMSE	KRCC	SRCC	PLCC		روش‌های عمومی
1.9485	0.2484	0.3652	0.4274	BIQME [۲۱]	
2.1262	0.1145	0.1701	0.1635	BLINDS-II [۲۰]	
2.0384	0.2088	0.3097	0.3248	BRISQUE [۱۴]	
2.1393	0.0798	0.1199	0.1214	DIIVINE [۱۶]	
2.0455	0.2053	0.3042	0.3150	FRIQUEE [۲۳]	
2.0860	0.1488	0.2260	0.2513	GM-LOG [۲۲]	
2.0416	0.1164	0.1759	0.3204	CCF [۲۷]	روش‌های زیر آب
2.0397	0.2307	0.3364	0.3229	FDUM [۲۸]	
2.0391	0.2080	0.3080	0.3238	UCIQE [۲۵]	
2.0879	0.0926	0.1395	0.2480	UIQM [۲۶]	
2.0460	0.2078	0.3134	0.3143	NUIQ [۳۰]	
1.8564	0.3261	0.4766	0.4841	B-FEN [۳۵]	
1.8380	0.3478	0.4804	0.5223	Yi [۳۱]	
1.7690	0.37779	0.5340	0.5520	MCOLE [۳۶]	
1.8127	0.3548	0.5210	0.5231	پیشنهادی	

۴-۵- ارزیابی کیفیت روی هر زیرمجموعه تصاویر

علاوه بر مقایسه عملکرد کلی، روش پیشنهادی ما برای ارزیابی کیفیت تصاویر زیر آب (IQA) بر روی شش زیرمجموعه مختلف از مجموعه داده UID2021 نیز مورد ارزیابی قرار گرفت و نتایج به دست آمده با روش‌های پیشرفته ارزیابی کیفیت تصاویر زیر آب UCIQE، UIQM، CCF و FDUM که نتایج جداگانه برای هر زیرمجموعه ارائه کرده‌اند، در جدول ۸ مقایسه شده است.

در زیرمجموعه مایل به آبی، روش پیشنهادی ما به ترتیب مقادیر PLCC، SRCC و KRCC برابر با ۰/۷۸۸۶، ۰/۷۷۴۶ و

۰/۵۰۰۳ را به دست آورد که نسبت به سایر روش‌های رقیب عملکرد بهتری داشت. این برتری عملکرد در سایر زیرمجموعه‌ها از جمله آبی-سبز، مایل به سبز، مه‌آلود، کم نور و کدر نیز مشاهده شد. علاوه بر این، روش پیشنهادی ما در اکثر زیرمجموعه‌ها کمترین مقدار RMSE را نیز به دست آورد که این موضوع مؤید اثربخشی بالای آن در ارزیابی دقیق کیفیت تصاویر زیر آب است. این نتایج جامع به دست آمده بر روی زیرمجموعه‌های مختلف، علاوه بر کل مجموعه داده، مزایای روش پیشنهادی ارزیابی کیفیت تصاویر زیر آب را به خوبی نشان می‌دهد.

جدول (۸): ارزیابی کیفیت روش‌های مختلف ارزیابی کیفیت تصویر

روی هر زیرمجموعه داده UID2021

RMSE	KRCC	SRCC	PLCC	روش	زیرمجموعه
1.6149	0.4575	0.6182	0.6409	UCIQE [۲۵]	کدر
1.6972	0.3753	0.5393	0.5908	UIQM [۲۶]	
1.7685	0.3281	0.4570	0.5416	CCF [۲۷]	
1.5302	0.5047	0.6767	0.6862	FDUM [۲۸]	
1.4725	-	0.6686	0.6974	Farhadi [۳۲]	
1.4328	0.6003	0.7746	0.7886	پیشنهادی	
1.8012	0.3820	0.5434	0.5767	UCIQE [۲۵]	کم نور
1.7973	0.3531	0.5042	0.5792	UIQM [۲۶]	
2.1229	0.1510	0.2198	0.2699	CCF [۲۷]	
1.7880	0.3889	0.5513	0.5851	FDUM [۲۸]	
1.7263	-	0.5590	0.5911	Farhadi [۳۲]	
1.4856	0.6135	0.7923	0.8153	پیشنهادی	
1.6290	0.4892	0.6668	0.6905	UCIQE [۲۵]	مایل به سبز
1.7634	0.4164	0.5864	0.6220	UIQM [۲۶]	
2.1414	0.2945	0.4043	0.3097	CCF [۲۷]	
1.6686	0.4783	0.6588	0.6716	FDUM [۲۸]	
1.3831	-	0.7512	0.7736	Farhadi [۳۲]	
1.3874	0.6841	0.8456	0.8485	پیشنهادی	
1.5817	0.4558	0.6237	0.6688	UCIQE [۲۵]	مه‌آلود
1.7390	0.3103	0.4291	0.5761	UIQM [۲۶]	
1.7528	0.3496	0.4899	0.5668	CCF [۲۷]	
1.7339	0.3790	0.5261	0.5795	FDUM [۲۸]	
1.4270	-	0.6625	0.7241	Farhadi [۳۲]	
1.5034	0.6889	0.7532	0.7603	پیشنهادی	
1.4597	0.4755	0.6650	0.6727	UCIQE [۲۵]	سبز
1.6734	0.3356	0.4928	0.5297	UIQM [۲۶]	
1.7609	0.2978	0.4340	0.4510	CCF [۲۷]	
1.5383	0.4390	0.6200	0.6261	FDUM [۲۸]	
1.4453	-	0.6253	0.6618	Farhadi [۳۲]	
1.2748	0.6461	0.8120	0.8325	پیشنهادی	
1.7414	0.4060	0.5599	0.5844	UCIQE [۲۵]	آبی
1.5015	0.5357	0.7156	0.7145	UIQM [۲۶]	
1.5620	0.4508	0.6089	0.6383	CCF [۲۷]	
1.5571	0.4987	0.6725	0.6882	FDUM [۲۸]	
1.4728	-	0.6877	0.7118	Farhadi [۳۲]	
1.3295	0.6561	0.8389	0.8575	پیشنهادی	

۵- جمع بندی

در این مقاله، یک چارچوب مبتنی بر یادگیری عمیق برای ارزیابی کیفیت تصاویر زیر آب معرفی شد که با ترکیب نقشه‌های تیزی، ساختار، و تصویر اصلی در یک معماری چندشاخه‌ای، ویژگی‌های پیچیده تخریب را به صورت خودکار استخراج می‌کند. استفاده از بلوک‌های توجه ECA در هر شاخه، تمرکز مدل بر کانال‌های اطلاعاتی مؤثر را افزایش داده و دقت پیش‌بینی را بهبود بخشید. نتایج تجربی گسترده روی دو مجموعه داده معتبر SAUD2.0 و UID2021 نشان داد که مدل پیشنهادی، علاوه بر دقت بالاتر در پیش‌بینی کیفیت تصاویر نسبت به روش‌های پیشین، از قابلیت تعمیم و پایداری مناسبی برخوردار است. نتایج آزمایش ارزیابی متقابل که شامل آموزش روی SAUD2.0 و تست روی UID2021 بود، تأیید کننده توانایی مدل در تعمیم به داده‌هایی با منشأ و خصوصیات متفاوت بود. بنابراین، رویکرد پیشنهادی نه تنها به عنوان یک روش دقیق و کارا برای ارزیابی کیفیت تصاویر زیر آب قابل استفاده است، بلکه امکان توسعه و به‌کارگیری آن در کاربردهای واقعی و متنوع را نیز فراهم می‌آورد جهت توسعه آینده، می‌توان تأثیر افزودن نقشه‌های ویژگی جدید (مانند عمق صحنه) یا آموزش مدل روی داده‌های ترکیبی از محیط‌های متنوع را بررسی کرد. این روش به‌عنوان یک ابزار قابل اعتماد، پتانسیل بالایی برای استفاده در سیستم‌های پایش خودکار زیر آب دارد.

مراجع

- [7] Yan X, Qin W, Wang Y, Wang G, Fu X. Attention-guided dynamic multi-branch neural network for underwater image enhancement. Knowledge-Based Systems. 2022 Dec 22;258:110041.
- [8] عباسی، امیررضا و هاشم زاده، مهدی و قویدل نیچران، جلیل، ۱۴۰۰، و Hybrid-Net: یک شبکه عصبی عمیق برای افزایش کیفیت تصاویر زیر آب، چهارمین کنفرانس بین‌المللی محاسبات نرم، <https://civilica.com/doc/1418617>
- [9] Liu R, Fan X, Zhu M, Hou M, Luo Z. Real-world underwater enhancement: Challenges, benchmarks, and solutions under natural light. IEEE transactions on circuits and systems for video technology. 2020 Jan 3;30(12):4861-75.
- [10] Zhang S, Li Y, Tan L, Yang H, Hou G. A no-reference underwater image quality evaluator via quality-aware features. Journal of Visual Communication and Image Representation. 2023 Dec 1;97:103979.
- [11] Karimi, M. and Nejati, M., Unsupervised Feature Learning for Blind Quality Assessment of Super-Resolved Images, Journal of Machine Vision and Image Processing (JMVIP), 8(3): 57-67, 2021.
- [12] Karimi M, Nejati M. HNQA: histogram-based descriptors for fast night-time image quality assessment. Multimedia Systems. 2024 Oct;30(5):259.
- [13] Varga D. No-reference image quality assessment with global statistical features. Journal of Imaging. 2021 Feb 5;7(2):29.
- [14] Mittal A, Moorthy AK, Bovik AC. No-reference image quality assessment in the spatial domain. IEEE Transactions on image processing. 2012 Aug 17;21(12):4695-708.
- [15] Zhang L, Zhang L, Bovik AC. A feature-enriched completely blind image quality evaluator. IEEE Transactions on Image Processing. 2015 Apr 24;24(8):2579-91.
- [16] Moorthy AK, Bovik AC. Blind image quality assessment: From natural scene statistics to perceptual quality. IEEE transactions on Image Processing. 2011 Apr 25;20(12):3350-64.
- [17] Li, X., Xu, H., Jiang, G., Yu, M., Chen, Y., Luo, T. and Ying, H., 2023. Underwater image quality assessment based on multiscale and antagonistic energy. IEEE Transactions on Instrumentation and Measurement.
- [18] Sun, S., Wang, H., Zhang, H., Li, M., Xiang, M., Luo, C. and Ren, P., 2022. Underwater image enhancement with reinforcement learning. IEEE Journal of Oceanic Engineering, 49(1), pp.249-261.
- [19] Moorthy AK, Bovik AC. A two-step framework for constructing blind image quality indices. IEEE Signal processing letters. 2010 Mar 18;17(5):513-6.
- [1] Naveen P. Advancements in underwater imaging through machine learning: Techniques, challenges, and applications. Multimedia Tools and Applications. 2024 Aug 24:1-20.
- [2] Raveendran S, Patil MD, Birajdar GK. Underwater image enhancement: a comprehensive review, recent trends, challenges and applications. Artificial Intelligence Review. 2021 Oct;54:5413-67.
- [3] Li C, Guo C, Ren W, Cong R, Hou J, Kwong S, Tao D. An underwater image enhancement benchmark dataset and beyond. IEEE transactions on image processing. 2019 Nov 28;29:4376-89.
- [4] Moghimi MK, Mohanna F. Real-time underwater image enhancement: a systematic review. Journal of Real-Time Image Processing. 2021 Oct;18(5):1509-25.
- [5] Fu Z, Wang W, Huang Y, Ding X, Ma KK. Uncertainty inspired underwater image enhancement. In European conference on computer vision 2022 Oct 23 (pp. 465-482). Cham: Springer Nature Switzerland.
- [6] Zhang, W., Zhuang, P., Sun, H.H., Li, G., Kwong, S. and Li, C., 2022. Underwater image enhancement via minimal color loss and locally adaptive contrast enhancement. IEEE Transactions on Image Processing, 31, pp.3997-4010.

- [33] Zhang Y, Liu C, Han Z, Jing J. Experimental Research of Underwater Polarization Imaging Based on Image Quality Assessment. In 2024 2nd International Conference on Signal Processing and Intelligent Computing (SPIC) 2024 Sep 20 (pp. 1045–1051). IEEE.
- [34] Chen, C., Li, Z., Zhong, Z., Wang, X. and Shao, F., 2024. Hierarchical prior-guided quality assessment method for underwater images. *Displays*, 83, p.102729.
- [35] Wu, Q., Wang, L., Ngan, K.N., Li, H., Meng, F. and Xu, L., 2020. Subjective and objective de-raining quality assessment towards authentic rain image. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(11), pp.3883–3897.
- [36] Jiang, Q., Yi, X., Ouyang, L., Zhou, J. and Wang, Z., 2024. Towards dimension-enriched underwater image quality assessment. *IEEE Transactions on Circuits and Systems for Video Technology*.
- [37] Lin, Z., He, Z., Jin, C., Luo, T. and Chen, Y., 2024. Joint Luminance-Saliency Prior and Attention for Underwater Image Quality Assessment. *Remote Sensing*, 16(16), p.3021.
- [38] Liu Y, Zhang B, Hu R, Gu K, Zhai G, Dong J. Underwater image quality assessment: Benchmark database and objective method. *IEEE Transactions on Multimedia*. 2024 Feb 28;26:7734–47.
- [39] [39] Zhu P, Ma H, Ma K, Liu Y, Liu S. Content-Adaptive Efficient Transformer for No-Reference Underwater Image Quality Assessment. In *OCEANS 2024-Singapore 2024 Apr 15* (pp. 1–9). IEEE.
- [40] [40] Chen T, Yang X, Wang T, Li N, Zhu S, Ji G. Underwater image quality assessment method via the fusion of visual and structural information. *Signal Processing: Image Communication*. 2025 Aug 1;136:117285.
- [41] [41] Tolie HF, Ren J, Cai J, Chen R, Zhao H. Blind Quality Assessment Using Channel-Based Structural, Dispersion Rate Scores, and Overall Saturation and Hue for Underwater Images. *IEEE Journal of Oceanic Engineering*. 2025 May 15.
- [42] [42] Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W. and Hu, Q., 2020. ECA-Net: Efficient channel attention for deep convolutional neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 11534–11542).
- [43] [43] Loshchilov, I. and Hutter, F., 2017. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*.
- [44] [44] Li, D., Jiang, T. and Jiang, M., 2020, October. Norm-in-norm loss with faster convergence and better performance for image quality assessment. In *Proceedings of the 28th ACM International conference on multimedia* (pp. 789–797).
- [45] [45] Chen, Z., Wang, J., Li, B., Yuan, C., Hu, W., Liu, J., Li, P., Wang, Y., Zhang, Y. and Zhang, C., 2024.
- [20] Mittal A, Soundararajan R, Bovik AC. Making a “completely blind” image quality analyzer. *IEEE Signal processing letters*. 2012 Nov 15;20(3):209–12.
- [21] Saad MA, Bovik AC, Charrier C. Blind image quality assessment: A natural scene statistics approach in the DCT domain. *IEEE Transactions on Image Processing*. 2012 Mar 21;21(8):3339–52.
- [22] Gu K, Tao D, Qiao JF, Lin W. Learning a no-reference quality assessment model of enhanced images with big data. *IEEE Transactions on neural networks and learning systems*. 2017 Mar 6;29(4):1301–13.
- [23] Xue W, Mou X, Zhang L, Bovik AC, Feng X. Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features. *IEEE Transactions on Image Processing*. 2014 Sep 8;23(11):4850–62.
- [24] Ghadiyaram D, Bovik AC. Perceptual quality prediction on authentically distorted images using a bag of features approach. *Journal of vision*. 2017 Jan 1;17(1):32–.
- [25] Yang M, Sowmya A. An underwater color image quality evaluation metric. *IEEE Transactions on Image Processing*. 2015 Oct 19;24(12):6062–71.
- [26] Panetta K, Gao C, Agaian S. Human-visual-system-inspired underwater image quality measures. *IEEE Journal of Oceanic Engineering*. 2015 Oct 26;41(3):541–51.
- [27] Wang, Y., Li, N., Li, Z., Gu, Z., Zheng, H., Zheng, B. and Sun, M., 2018. An imaging-inspired no-reference underwater color image quality assessment metric. *Computers & Electrical Engineering*, 70, pp.904–913.
- [28] Yang N, Zhong Q, Li K, Cong R, Zhao Y, Kwong S. A reference-free underwater image quality assessment metric in frequency domain. *Signal Processing: Image Communication*. 2021 May 1;94:116218.
- [29] Li X, Xu H, Jiang G, Yu M, Luo T, Zhang X, Ying H. Underwater image quality assessment from synthetic to real-world: Dataset and objective method. *ACM Transactions on Multimedia Computing, Communications and Applications*. 2023 Oct 23;20(3):1–23.
- [30] Jiang Q, Gu Y, Li C, Cong R, Shao F. Underwater image enhancement quality evaluation: Benchmark dataset and objective metric. *IEEE Transactions on Circuits and Systems for Video Technology*. 2022 Apr 5;32(9):5959–74.
- [31] Yi X, Jiang Q, Zhou W. No-reference quality assessment of underwater image enhancement. *Displays*. 2024 Jan 1;81:102586.
- [32] Liu C, Zhou J, Long B, Li T, Zhang D, Zhang W, Vivone G. Ocean Remote Sensing Image Quality Assessment Via Multi-Directional Perception Fusion and Deviation-Aware Processing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2025 May 12.

Gmc-iqa: Exploiting global-correlation and mean-opinion consistency for no-reference image quality assessment. arXiv preprint arXiv:2401.10511.

- [46] [46] Huber, P.J., 1992. Robust estimation of a location parameter. In Breakthroughs in statistics: Methodology and distribution (pp. 492–518). New York, NY: Springer New York
- [47] [47] Hou G, Li Y, Yang H, Li K, Pan Z. UID2021: An underwater image dataset for evaluation of no-reference quality assessment metrics. ACM Transactions on Multimedia Computing, Communications and Applications. 2023 Mar 1;19(4):1-24.



مریم کریمی استادیار گروه علوم کامپیوتر دانشگاه شهرکرد است. ایشان دانش آموخته دکتری مهندسی کامپیوتر از دانشگاه صنعتی اصفهان در سال ۱۳۹۶ است. همچنین مدارک کارشناسی و کارشناسی ارشد خود را به ترتیب در سالهای ۱۳۸۵ از دانشگاه صنعتی امیرکبیر و ۱۳۸۸ از دانشگاه صنعتی شریف اخذ نموده است. علاقمندی‌های تحقیقاتی ایشان شامل پردازش تصویر و ویدئو، ارزیابی کیفیت ادراکی تصویر و ویدئو و بینایی ماشین، یادگیری ماشین و یادگیری عمیق است.