

A Review of Machine Learning-based Techniques for Digital Image Dehazing

Somaye Rastgarifar¹, Yasser Elmi Sola²

¹ Department of Computer Engineering, Sab.C., Islamic Azad University, Sabzevar, Iran

² Department of Computer Engineering, Sab.C., Islamic Azad University, Sabzevar, Iran
Yasser.Elmi@iau.ac.ir

Received: 25 April 2025

Revised: 17 June 2025

Accepted: 19 June 2025

Abstract:

Fog is a natural phenomenon that occurs when water droplets or ice crystals become suspended in the air near the ground, reducing visibility. It can be thought of as a cloud at ground level. Fog is essentially a cloud that forms at the ground level. It forms when the air near the surface cools to its dew point, the temperature at which the air becomes saturated with moisture, leading to condensation. Fog formation usually requires three main ingredients: moisture, cooling air, and condensation nuclei.

Considering that adverse weather conditions such as fog, rain and snow expose outdoor images to a wide range of disturbances (such as noise, quality degradation, etc.), which will cause many problems in CCTV cameras, tracking, navigation and applications where reliability is important (such as outdoor surveillance, object recognition), dehazing is essential and is widely used in the fields of machine vision and computer graphics. This article reviews the techniques of dehazing in digital images and aims to analyze the different methods available in the scientific literature. Considering the importance of improving image quality in various fields such as photography, medicine and surveillance, this study analyzes and collects information from previous articles and research. The results show that deep learning-based methods are significantly more successful in improving the quality of foggy images than traditional methods. Finally, it is suggested that future research focus on optimizing algorithms and investigating real-world applications of these techniques.

Keywords: Image dehazing, deep learning, image restoration, convolutional neural network, image quality assessment

Corresponding Author: Yasser Elmi Sola

Corresponding Author Address: Department of Computer Engineering- Siadati Blvd. -Islamic Azad University- Sabzevar Branch, Iran

مروری بر تکنیک های مبتنی بر یادگیری ماشین جهت مه زدایی تصاویر دیجیتال

سمیه رستگاری فر^۱، کارشناسی ارشد، یاسر علمی سولا^۲، استادیار

۱- گروه مهندسی کامپیوتر، واحد سبزوار، دانشگاه آزاد اسلامی، سبزوار، ایران

۲- گروه مهندسی کامپیوتر، واحد سبزوار، دانشگاه آزاد اسلامی، سبزوار، ایران
Yasser.Elmi@iau.ac.ir

تاریخ پذیرش مقاله: ۱۴۰۴/۰۵/۲۵

تاریخ بازنگری مقاله: ۱۴۰۴/۰۴/۱۷

تاریخ ارسال مقاله: ۱۴۰۴/۰۳/۱۷

چکیده:

مه یک پدیده طبیعی است و زمانی رخ می دهد که قطرات آب یا کریستال های یخ در هوا در نزدیکی سطح زمین معلق می شوند و منجر به کاهش دید می شود. می توان آن را به صورت ابری در سطح زمین در نظر گرفت. مه در اصل ابری است که در سطح زمین تشکیل می شود. زمانی ایجاد می شود که هوای نزدیک سطح تا نقطه شبنم خنک می شود، دمای که در آن هوا از رطوبت اشباع می شود و منجر به تراکم می شود. تشکیل مه معمولاً به سه جزء اصلی نیاز دارد: رطوبت، هوای خنک کننده و هسته های تراکم. با توجه به آنکه شرایط نامساعد آب و هوایی مانند مه، باران و برف، تصاویر در فضای باز را، در معرض طیف گسترده ای از اختلالات (مانند نویز، کاهش کیفیت و غیره) قرار می دهد که باعث ایجاد بسیاری از مشکلات در دوربین های مدار بسته، ردیابی، ناوبری و برنامه های کاربردی که قابلیت اطمینان در آنها مهم اند (مانند نظارت در فضای باز، تشخیص شیء) خواهد شد حذف مه ضروری می باشد و به طور گسترده در حوزه بینایی ماشین و گرافیک کامپیوتری کاربرد دارد. این مقاله مروری به بررسی تکنیک های مه زدایی در تصاویر دیجیتال می پردازد و هدف آن تحلیل روش های مختلف موجود در ادبیات علمی است. با توجه به اهمیت بهبود کیفیت تصاویر در زمینه های مختلف مانند عکاسی، پزشکی و نظارت، این مطالعه به تحلیل و جمع آوری اطلاعات از مقالات و تحقیقات پیشین می پردازد. نتایج نشان می دهد که روش های مبتنی بر یادگیری عمیق به طور قابل توجهی در بهبود کیفیت تصاویر مه زده موفق تر از روش های سنتی هستند. در نهایت، پیشنهاد می شود که تحقیقات آینده بر روی بهینه سازی الگوریتم ها و بررسی کاربردهای واقعی این تکنیک ها تمرکز کنند.

کلمات کلیدی: مه زدایی تصاویر، یادگیری عمیق، بازیابی تصویر، شبکه عصبی کانولوشنی، ارزیابی کیفیت تصویر

نام نویسنده ی مسئول: یاسر علمی سولا

نشانی نویسنده ی مسئول: سبزوار - بلوار سیادت - دانشگاه آزاد اسلامی واحد سبزوار - گروه کامپیوتر

۱- مقدمه

هسته‌های متراکم ذرات کوچکی مانند گرد و غبار یا آلودگی هستند که سطحی را برای متراکم شدن بخار آب فراهم می‌کنند. رایج‌ترین انواع مه عبارتند از: مه تشعشع، مه فرارفتی، مه سربالایی، مه تبخیر، مه یخ. مه می‌تواند بر شرایط آب و هوایی تأثیر بگذارد و تحت تأثیر قرار گیرد. انتظار می‌رود تغییرات آب و هوایی بر فراوانی و توزیع مه در مناطق مختلف تأثیر بگذارد. دماهای گرم‌تر ممکن است منجر به تغییر در الگوهای تشکیل مه شود. به عنوان مثال، مه ساحلی، که برای برخی از اکوسیستم‌ها حیاتی است، ممکن است در مناطقی که دمای دریا در حال افزایش است، کاهش یابد. قطرات مه معمولاً بسیار کوچک هستند و قطر آنها بین ۱ تا ۲۰ میکرومتر است. علی‌رغم اندازه آنها، این قطرات می‌توانند به‌طور قابل توجهی نور را پراکنده کنند و منجر به کاهش دید شوند. چگالی مه بر حسب محتوای آب مایع (LWC) اندازه گیری می‌شود که می‌تواند از ۰.۰۵ تا ۰.۵ گرم بر متر مکعب متغیر باشد. شرایط نامساعد آب و هوایی مانند مه (haze, fog, mist)، باران سنگین و برف باعث کاهش کیفیت تصاویر در فضای باز می‌شوند. این وضعیت‌ها طیف گسترده‌ای از اختلالات مانند نویز و کاهش وضوح تصویر را به همراه دارند. در نتیجه، دوربین‌های مداربسته، سیستم‌های ردیابی و ناوبری و برنامه‌های کاربردی که قابلیت اطمینان در آنها حیاتی است (مانند نظارت در فضای باز و تشخیص شیء)، با مشکلات زیادی مواجه می‌شوند.

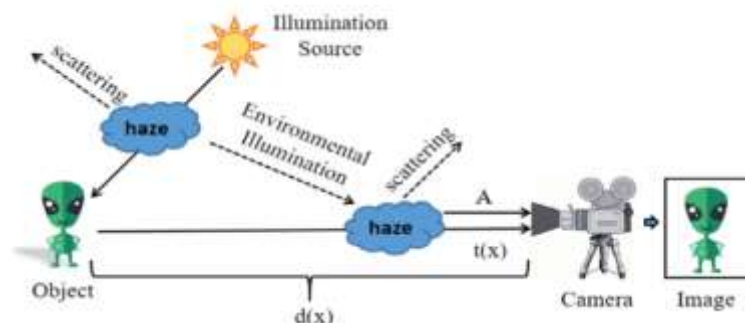
یکی از رایج‌ترین مشکلات در شرایط بد جوی که بر کیفیت تصویر در فضای باز تأثیر می‌گذارد، مه است. تصاویر مخدوش شده توسط مه، کیفیت سیستم‌های مبتنی بر دید را کاهش می‌دهند. این سیستم‌ها شامل سیستم‌های خودکار نظارتی، تشخیص در فضای باز، ردیابی، حمل و نقل هوشمند، تصاویر ماهواره‌ای و تصاویر زیر آب هستند. در این شرایط، دید اجسام از فاصله دور کاهش یافته و منجر به سوانح جاده‌ای، تصادم کشتی‌ها و مشکلات در حرکت هواپیماها می‌شود.

۱-۱- بیان مساله

با توجه به آنکه شرایط نامساعد آب و هوایی مانند مه، باران و برف، تصاویر در فضای باز را، در معرض طیف گسترده‌ای از اختلالات (مانند نویز، کاهش کیفیت و غیره) قرار می‌دهد که باعث ایجاد بسیاری از مشکلات در دوربین‌های مدار بسته، ردیابی، ناوبری و برنامه‌های کاربردی که قابلیت اطمینان در آنها مهم‌اند (مانند نظارت در فضای باز، تشخیص شیء) خواهد شد حذف مه ضروری می‌باشد و به‌طور گسترده در حوزه بینایی ماشین و گرافیک کامپیوتری کاربرد دارد. از سوانح معروف ایجاد شده ناشی از مه، می‌توان به برخورد هواپیما به ساختمان امپایر استیت و یا برخورد دو کشتی اقیانوس پیما اس‌اس اندریا دوریا و ام‌اس استکهلم را اشاره کرد. با توجه به پیشرفت‌های سیستم‌های دیجیتال، بهبود کیفیت و حذف نویز یکی از شاخه‌های جالب و مورد توجه در زمینه‌ی بهبود تصاویر است و کاربردهای گسترده‌ای در بسیاری از زمینه‌های علمی مانند پردازش تصویر و بینایی ماشین پیدا کرده است. مه زدایی تصاویر، فرایندی برای بازیابی یک تصویر تمیز از یک ورودی تخریب‌شده که در صحنه‌های مه گرفته شده‌است، بسیار ارزشمند است و در دو دهه گذشته به‌طور گسترده مورد مطالعه قرار گرفته است. روش‌های SID موجود را می‌توان به دو دسته طبقه‌بندی کرد، روش‌های سنتی پردازش تصویر و روش‌های مبتنی بر یادگیری عمیق. اخیراً، یادگیری عمیق در وظایف مختلف پردازش تصویر استفاده شده و عملکرد مناسبی نشان داده است. در مجموع با توجه به نتایج در حوزه SID روش‌های مبتنی بر یادگیری عمیق برتر از الگوریتم‌های سنتی هستند. با توجه به توضیحات فوق هدف این پژوهش، حذف مه از تصاویر با ایجاد حداقل تخریب در آنهاست که در این میان از ترکیب مجموعه‌ای از ویژگی‌های حوزه زمان و فرکانس و یادگیری عمیق استفاده خواهد شد.

مقاله مروری حاضر به بررسی جامعی از روش‌های حذف مه از تصاویر دیجیتال با رویکرد یادگیری ماشین می‌پردازد و در عین حال به روش‌های مدلسازی مسئله و تکنیک‌های غیر از یادگیری ماشین نیز اشاره می‌کند. این رویکرد چندجانبه به خوانندگان این امکان را می‌دهد که درک بهتری از چالش‌ها و راه‌حل‌های موجود در این حوزه داشته باشند. با بررسی تکنیک‌های سنتی و غیر یادگیری ماشین، مقاله به تحلیل نقاط قوت و ضعف هر یک پرداخته و روش‌های نوآورانه‌ای را که می‌توانند مکمل رویکردهای یادگیری ماشین باشند، معرفی می‌کند. این امر به غنای علمی مقاله افزوده و به پژوهشگران و مهندسان در انتخاب روش‌های مناسب‌تر کمک می‌کند.

علاوه بر این، مقاله مروری حاضر تمامی روش هایی را که در مقالات مروری قبلی به طور کامل پوشش داده نشده اند، به دقت بررسی می کند. این امر شامل تکنیک های جدید و نوآورانه ای است که در سال های اخیر معرفی شده اند و به طور قابل توجهی می توانند در بهبود کیفیت تصاویر در شرایط مه آلود موثر باشند. همچنین، مقاله به صورت انتقادی تمامی زیرمجموعه های یادگیری ماشین را بررسی و مقایسه می کند، به طوری که نقاط قوت و ضعف هر یک از این زیرمجموعه ها به طور مجزا تحلیل شده و در نهایت، یک نمای کلی از وضعیت فعلی و روندهای آینده در این حوزه ارائه می دهد. این تحلیل عمیق و جامع می تواند راهنمایی ارزشمند برای پژوهشگران و توسعه دهندگان در انتخاب و بهبود روش های حذف مه باشد.



شکل (۱) مدل پراکندگی جوی (ASM) [۱]

مه یک پدیده طبیعی است که می توان با استفاده از مدل ASM آن را توضیح داد. مک کارتنی [۲] اولین بار مدل ASM را برای توصیف اصول تشکیل مه پیشنهاد داد. این مدل در سال های اخیر توسعه یافته است به گونه ای که اکنون به طور گسترده ای استفاده می شود. این مدل یک پایه قابل اعتماد برای تحقیق در مورد حذف مه فراهم می کند. فرمول مربوط به آن به شکل زیر است:

$$I(x) = J(x)t(x) + A(1 - t(x)) \quad (۱)$$

که در آن، محل پیکسل با x نشان داده می شود و A به معنای نور جوی است. در روش های رفع مه گرفتنی بر اساس ASM، معمولاً A ناشناخته است. $I(x)$ نشان دهنده تصویر مه گرفته و $J(x)$ نشان دهنده تصویر صحنه واضح است. برای بیشتر مدل های رفع مه گرفتنی، $I(x)$ ورودی و $J(x)$ خروجی مطلوب است. $t(x)$ به معنای نقشه انتقال متوسط است که به صورت زیر تعریف می شود:

$$t(x) = e^{-\beta d(x)} \quad (۲)$$

در این مدل، β و $d(x)$ به ترتیب نشان دهنده پارامتر پراکندگی جو و عمق $I(x)$ هستند. بنابراین، $t(x)$ توسط $d(x)$ تعیین می شود که می تواند برای سنتز تصویر مه آلود استفاده شود. اگر $t(x)$ و A قابل تخمین باشند، تصویر بدون مه $J(x)$ را می توان با فرمول زیر به دست آورد:

$$J(x) = \frac{I(x) - A(1 - t(x))}{t(x)} \quad (۳)$$

اصل تصویرسازی ASM در شکل ۱ نشان داده شده است. این مدل نشان می دهد که چگونه نوری که از یک جسم به دوربین می رسد، تحت تأثیر ذرات موجود در هوا قرار می گیرد. برخی از تحقیقات از ASM برای توصیف فرآیند تشکیل مه استفاده می کنند و پارامترهای موجود در مدل پراکندگی جو را به صورت صریح یا ضمنی حل می کنند. همانطور که در جدول ۱ نشان داده شده است، ASM تأثیر مهمی بر تحقیق در زمینه رفع مه دارد. این تحقیقات شامل الگوریتم های نظارت شده، نیمه نظارت شده و بدون نظارت می شود. برای وظایف بینایی کامپیوتر مانند تشخیص اشیاء، بخش بندی تصویر و طبقه بندی تصویر، برچسب های دقیق واقعیت زمینی با نشانه گذاری دقیق به دست می آیند. این برچسب ها می توانند برای آموزش و ارزیابی الگوریتم های بینایی کامپیوتر استفاده شوند.

با این حال، برچسب های دقیق، واضح و پیکسلی (یعنی تصاویر بدون مه همراه) برای تصاویر مه آلود در صحنه های طبیعی بسیار دشوار است. در حال حاضر، دو روش اصلی برای به دست آوردن تصاویر همراه مه آلود و بدون مه وجود دارد. روش اول، ساخت

داده‌های مصنوعی با استفاده از یک مدل ASM است (مانند D-HAZY, HazeRD و RESIDE). با انتخاب پارامترهای مختلف برای یک ASM، پژوهشگران می‌توانند به راحتی تصاویر مه‌آلود با چگالی‌های مختلف مه را به دست آورند. برای سنتز یک تصویر مه‌آلود، چهار جزء مورد نیاز است: یک تصویر واضح، یک نقشه عمق $d(x)$ مربوط به محتوای تصویر واضح، نور جو A و پارامتر پراکندگی جو β . بنابراین، می‌توانیم مجموعه داده مصنوعی را به دو مرحله تقسیم کنیم. در مرحله اول، تصاویر واضح و نقشه‌های عمق مربوطه باید به صورت جفت جمع‌آوری شوند.



شکل ۲) یک تصویر واضح و نقشه عمق مربوطه در مجموعه داده NYU-Depth



شکل ۳) تصاویر مه‌آلود سنتز شده با چگالی‌های مختلف برای پارامتر پراکندگی جو β بر اساس مجموعه داده NYU-Depth



شکل ۴) مثال‌هایی از NH-HAZE و Dense-Haze, O-HAZE, I-HAZE

برای اطمینان از اینکه مه سنتز شده تا حد ممکن به مه دنیای واقعی نزدیک باشد، اطلاعات عمق باید به اندازه کافی دقیق باشد. شکل ۲ تصویر واضح و نقشه عمق مربوطه را در مجموعه داده NYU-Depth نشان می‌دهد. در مرحله دوم، نور جو A و پارامتر پراکندگی جو β به عنوان مقادیر ثابت یا به صورت تصادفی انتخاب می‌شوند.

روش دوم برای ایجاد تصویر مه آلود با استفاده از یک مولد مه مانند O-HAZE, I-HAZE, Dense-Haze و NH-HAZE است. شکل ۴ چهار جفت نمونه تصویر مه آلود و بدون مه را نشان می دهد که در این مجموعه داده ها شبیه سازی شده اند. مزیت اصلی مه مصنوعی و تولید شده این است که دشواری جمع آوری داده ها را کاهش می دهد. با این حال، تصاویر مه آلودی که بر اساس ASM شبیه سازی شده اند یا توسط مولد مه تولید شده اند نمی توانند فرآیند تشکیل مه در دنیای واقعی را به طور کامل شبیه سازی کنند. بنابراین، تفاوت ذاتی بین داده های مصنوعی و داده های واقعی وجود دارد. چندین تحقیق مشکلات داده های مصنوعی را مشاهده کرده اند و سعی در ساخت مجموعه داده های واقعی مانند MRFID و BeDDE داشته اند. اما به دلیل هزینه ها و دشواری های بالا در جمع آوری داده ها، مجموعه داده های واقعی به اندازه مجموعه داده های مصنوعی مانند RESIDE تعداد نمونه های کافی ندارند. برای تسهیل مقایسه مجموعه داده های مختلف، ویژگی های گوناگون این مجموعه داده ها را در جدول ۱ خلاصه کرده ایم.

جدول (۱) مجموعه داده ها برای مه زدایی تصویر

Dataset	Type	Nums	I/O	P/NP
D-HAZY	Syn	1,400+	I	P
HazeRD	Syn	15	O	P
I-HAZE	HG	35	I	P
O-HAZE	HG	45	O	P
RESIDE	S&R	10,000+	I&O	P&NP
Dense-Haze	HG	33	O	P
NH-HAZE	HG	55	O	P
MRFID	Real	200	O	P
BeDDE	Real	200+	O	P
4KID	Syn	10,000	O	P

Syn به معنی تصاویر مه آلود مصنوعی است. HG نشان دهنده تصاویر مه آلود تولید شده توسط تولیدکننده مه است. Real به معنای صحنه های واقعی است. S&R به معنای مصنوعی و واقعی است. I و O به ترتیب به معنای داخلی و خارجی هستند. P و NP به ترتیب به معنای جفت و غیر جفت هستند.

شبکه های عصبی کانولوشنی (CNNs) به طور گسترده در شبکه های دیپ لرنینگ مبتنی بر حذف مه رایج هستند. ماژول های استفاده شده معمولاً شامل کانولوشن استاندارد، کانولوشن توسعه یافته، ادغام چند مقیاسی، پیرامید ویژگی، اتصالات متقاطع لایه ای و توجه هستند. معمولاً، چندین بلوک ابتدایی با یک شبکه حذف مه ترکیب می شوند. به منظور تسهیل درک اصول الگوریتم های مختلف حذف مه، بلوک های ابتدایی معمولاً در معماری های شبکه استفاده شده و به شرح زیر خلاصه شده اند.

- تبدیل استاندارد: نشان داده شده است که استفاده از تبدیل استاندارد به صورت متوالی در ساخت شبکه های عصبی مؤثر است. بنابراین، تبدیل های استاندارد معمولاً در مدل های رفع مه به همراه بلوک های دیگر استفاده می شوند.
- کانولوشن گسترش یافته: کانولوشن گسترش یافته می تواند میدان دید را افزایش دهد در حالی که اندازه هسته کانولوشن تغییر نمی کند. مطالعات نشان داده اند که کانولوشن گسترش یافته می تواند عملکرد استخراج ویژگی های global را بهبود بخشد. علاوه بر این، ترکیب لایه های کانولوشن با نرخ های گسترش مختلف می تواند ویژگی ها را از میدان های دید مختلف استخراج کند.
- تلفیق چند مقیاسی: شبکه های عصبی کانولوشن چند مقیاسی ابزار قدرتمندی هستند که در استخراج ویژگی ها از تصاویر برای حل مسائل مختلف تصویری مورد استفاده قرار می گیرند. این شبکه ها از هسته های کانولوشن با اندازه های مختلف استفاده می کنند تا ویژگی های مختلفی را از تصاویر استخراج کنند. سپس، این ویژگی های استخراج شده ترکیب می شوند. یکی از روش های متداول در فرآیند تلفیق ویژگی ها، افزودن یا اتصال مکانی ویژگی های خروجی است که توسط هسته های کانولوشن با اندازه های مختلف به دست آمده اند. این روش ها در مسائل مختلفی مانند حذف مه اثبات شده اند. شبکه تقویت کننده شفافیت مبتنی بر یادگیری عمیق از این استراتژی در لایه میانی شبکه استفاده می کند تا اطلاعات چندگانه فضا و کانال را استخراج کند [۹-۱۴].

- اتصال میان‌لایه‌ای: برای بهبود تبادل اطلاعات بین لایه‌های مختلف و افزایش توانایی استخراج ویژگی‌های شبکه، در شبکه‌های عصبی پیچشی (CNNs) از اتصالات میان‌لایه‌ای استفاده می‌شود. سه نوع اصلی از اتصالات میان‌لایه‌ای که در شبکه‌های زدودن مه استفاده می‌شوند عبارتند از: اتصال باقیمانده که توسط ResNet معرفی شده‌است، اتصال که توسط DenseNet طراحی شده‌است، و اتصال جهشی که از U-Net الهام گرفته شده‌است [۲۱-۱۶].
- توجه در حذف مه: مکانیزم توجه با موفقیت در تحقیقات پردازش زبان طبیعی به کار رفته‌است. بلوک‌های توجه معمولاً مورد استفاده در بینایی کامپیوتر شامل توجه کانال (channel attention) و توجه فضایی (spatial attention) هستند. برای فرایند استخراج و بازسازی ویژگی تصویر دوبعدی، توجه کانال می‌تواند بخش‌های مفید نقشه ویژگی را استخراج کند. این استراتژی پردازش ناهمگن نقشه ویژگی اجازه می‌دهد تا مدل بیشتر بر روی اطلاعات ویژگی مؤثر تمرکز کند. مکانیزم توجه فضایی (spatial attention) بر تفاوت‌ها در مناطق مکانی داخلی نقشه ویژگی تمرکز می‌کند، مانند توزیع (distribution of haze) بر روی کل نقشه. با جاسازی ماژول توجه در شبکه، چندین روش پاکسازی از مهبه عملکرد عالی در پاکسازی از مه دست‌یافته‌اند [۳۴-۲۸].
- در این بخش، توابع از دست رفتگی (Loss Function) معمولاً در مدل‌های تصویری کاهش مه، نظارت‌شده، نیمه‌نظارت‌شده و بی‌نظارت‌شده معرفی و استفاده می‌شود که می‌توان از آن‌ها برای تخمین نقشه انتقال، پیش‌بینی تصویر واضح، بازسازی تصویر مه‌دار، انطباق نور جوی زمینه و غیره استفاده کرد. چندین الگوریتم از چندین تابع از دست رفتگی به ترکیب استفاده می‌کنند تا عملکرد بهتری در کاهش مه داشته باشند. یک طبقه‌بندی دقیق و خلاصه از توابع از دست رفتگی استفاده شده توسط روش‌های مختلف کاهش مه در جدول ۲ ارائه شده‌است. در تابع از دست رفتگی که در زیر معرفی شده‌است، X و Y به ترتیب مقدار پیش‌بینی شده و مقدار واقعی را نشان می‌دهند.

جدول ۲) توابع از دست دادن (Loss Function) برای وظیفه حذف مه

Loss Function	Algorithms
L1	[۱, ۱]
L2	[۳۷, ۱]
SSIM	[۱, ۱]
MS-SSIM	[۱, ۲۳]
Perceptual	[۱۷, ۳۲]
TV Loss	[۴۲, ۱]
Gradient	[۴, ۲۷]

کاهش وفاداری. توابع از دست دادن وفاداری متداول برای تحقیقات در زمینه حذف مه شامل تابع از دست دادن L1 و تابع از دست دادن L2 هستند که به شرح زیر تعریف می‌شوند:

$$L1 = \|X - Y\|_1 \quad (۴)$$

$$L2 = \|X - Y\|_2 \quad (۵)$$

- از دست رفتگی ویژگی‌ها. پژوهش‌های انجام شده در زمینه افزایش وضوح تصویر و انتقال سبک نشان می‌دهد که ویژگی‌های سیستم بینایی انسان در فرآیند ارزیابی دریافتی به‌طور کامل توسط از دست دادن (loss) L1 یا L2 نمایان نمی‌شود. به همین دلیل، L2 می‌تواند منجر به خروجی‌های بیش از حد صاف شود. پژوهش‌های اخیر از شبکه‌های عصبی طبقه‌بندی پیش‌آموزش دیده برای محاسبه از دست رفتگی ادراکی (perceptual loss) در فضای ویژگی استفاده می‌کنند. مدل پیش‌آموزش داده شده معمول‌ترین مدل استفاده شده VGG است و برخی از لایه‌های آن برای محاسبه فاصله بین تصویر پیش‌بینی شده و تصویر مرجع در فضای ویژگی استفاده می‌شود، یعنی،

$$L_{per}(X, Y) = \sum_{i=1}^N \|\psi_i(X) - \psi_i(Y)\|_2, \quad (۶)$$

در اینجا ما تعدادی ویژگی را انتخاب می‌کنیم (که توسط N نشان داده شده) و برای هر ویژگی، یک شاخص داریم (که توسط i نشان داده شده).

ما از یک شبکه VGG که قبلاً آموزش داده شده استفاده می کنیم (که توسط ψ نشان داده شده). در حین محاسبات ما، پارامترهای این شبکه تغییر نمی کنند.

- در روش های رفع مه، "شباهت ساختاری" یا SSIM نیز به عنوان یک تابع از دست رفتن در فرآیند بهینه سازی استفاده می شود. تحقیقات نشان داده اند که از دست رفتن SSIM می تواند شباهت ساختاری را در هنگام بازیابی تصویر بهبود بخشد. MS-SSIM ارزیابی چند مقیاسی را در SSIM معرفی می کند، که همچنین به عنوان تابع از دست رفتن توسط الگوریتم های رفع مه استفاده می شود، یعنی،

$$L_{ssim} = 1 - SSIM(X, Y) \quad (7)$$

$$L_{mssim} = 1 - MSSIM(X, Y) \quad (8)$$

- از دست دادن گرادیان: از دست دادن گرادیان که به عنوان از دست دادن لبه نیز شناخته می شود، برای بازسازی بهتر اطلاعات کناره و لبه های تصویر بدون مه استفاده می شود. استخراج لبه می تواند با استفاده از عملگر لاپلاس، عملگر کنی و غیره انجام شود. برای مثال، SA-CGAN از لاپلاسیان گوسی با انحراف معیار σ استفاده می کند تا تفاضل مرتبه دوم روی تصویر دو بعدی F انجام دهد:

$$L(m, n) = \nabla^2 F(m, n) = \frac{\partial^2 F}{\partial m^2} + \frac{\partial^2 F}{\partial n^2} = -\frac{1}{\pi \sigma^4} \left[1 - \frac{m^2 + n^2}{2\sigma^2} \right] \exp \left(-\frac{m^2 + n^2}{2\sigma^2} \right) \quad (9)$$

در اینجا (m, n) به معنای مکان پیکسل است و $L(m, n)$ برای هر دو محور X و Y محاسبه می شود. سپس، از توابع هدف رگرسیون مانند L_1 و L_2 برای محاسبه از دست دادن گرادیان استفاده می شود.

- از دست دادن تغییرات کل. خسارت تغییرات کلی یا TV می تواند برای هموار کردن یک تصویر و حذف نویز استفاده شود. هدف آموزش، کمینه کردن تابع زیر است:

$$L_{TV} = \|\partial_m X\|_1 + \|\partial_n X\|_1 \quad (10)$$

در اینجا m و n به ترتیب نشان دهنده مختصات افقی و عمودی هستند. از این فرمول می توان فهمید که از دست دادن تغییرات کل (TV) را می توان به شبکه هایی که به صورت بدون نظارت آموزش داده می شوند اضافه کرد، بدون نیاز به استفاده از Y واقعی (ground-truth Y).

به دلیل وجود مه، اشباع و کنتراست تصویر کاهش می یابد و رنگ تصویر به دلیل حالت نامشخص تغییر می کند. برای اندازه گیری تفاوت بین تصویر بدون مه و تصویر اصلی بدون مه، به معیارهای عینی نیاز داریم تا نتایج به دست آمده توسط الگوریتم های مختلف از بین بردن مه را ارزیابی کنیم. بیشتر مقالات از نسبت اوج سیگنال به نویز (PSNR) و SSIM برای ارزیابی کیفیت تصویر پس از از بین بردن مه استفاده می کنند. برای محاسبه PSNR باید از رابطه ۱۱ برای به دست آوردن میانگین مربعات خطا (MSE) استفاده کرد:

$$MSE = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (X(i, j) - Y(i, j))^2 \quad (11)$$

در اینجا X و Y به ترتیب نشان دهنده دو تصویری هستند که قرار است ارزیابی شوند. H و W ارتفاع و عرض آن ها هستند، به این معنی که ابعاد X و Y باید کاملاً یکسان باشد. شاخص مکان پیکسل در تصویر با i و j نشان داده می شود. سپس، PSNR با محاسبه لگاریتمی به صورت زیر به دست می آید:

$$PSNR = 10 \log_{10} \left[\frac{(2^N - 1)^2}{MSE} \right], \quad (12)$$

در اینجا N برای تصاویر ۸ بیتی برابر ۸ است. SSIM بر اساس همبستگی بین درک بصری انسان و اطلاعات ساختاری استوار است و فرمول آن به صورت زیر تعریف می شود:

$$SSIM(X, Y) = \frac{2\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad (13)$$

در اینجا، μ_x و μ_y میانگین های X و Y را نشان می دهند؛ σ_x و σ_y واریانس های X و Y هستند؛ σ_{xy} کواریانس بین دو متغیر است؛ و C_1 و C_2 ثابت هایی هستند که برای تضمین پایداری عددی استفاده می شوند. به دلیل اینکه مه می تواند باعث تغییر رنگ صحنه ها و اشیاء شود، برخی از تحقیقات از CIEDE2000 به منظور ارزیابی درجه تغییر رنگ استفاده می کنند. PSNR،

SSIM و CIEDE به معیارهای ارزیابی با مرجع کامل تعلق دارند، به این معنی که باید تصویر واضح متناظر با تصویر مه دار به عنوان یک مرجع استفاده شود. با این حال، جفت‌های واقعی تصویر مه‌آلود و بدون مه سخت است که دقیقاً یکسان باشند. به عنوان مثال، روشنایی و اشیا موجود در صحنه ممکن است قبل و بعد از ظهور مه تغییر کنند. بنابراین، برای حفظ دقت فرآیند ارزیابی، لازم است تصویر مه‌آلود متناظر را با یک تصویر واضح ترکیب کنیم.

علاوه بر معیارهای مرجع کامل PSNR، SSIM و CIEDE، کارهای اخیر از معیارهای بدون مرجع SSEQ و BLINDS-II برای ارزیابی تصاویر بدون مرجع استفاده می‌کنند. معیارهای بدون مرجع برای ارزیابی تصاویر در دنیای واقعی اهمیت زیادی دارند. با این حال، ارزیابی الگوریتم‌های فعلی از بین بردن مه معمولاً روی مجموعه داده‌هایی با جفت‌های تصاویر مه‌آلود و بدون مه انجام می‌شود. از آنجایی که معیارهای مرجع کامل برای مجموعه داده‌های جفت‌شده مناسب‌تر هستند، در این مرحله بیشتر از معیارهای بدون مرجع استفاده می‌شود.

مدل‌های رفع مه به صورت نظارت شده معمولاً به انواع مختلف سیگنال‌های نظارتی در فرآیند آموزش نیاز دارند، مانند نقشه انتقال، نور جوی، برچسب تصویر بدون مه و غیره. به طور مفهومی، روش‌های رفع مه با نظارت می‌توانند به دو دسته مبتنی بر ASM و غیر مبتنی بر ASM تقسیم شوند. با این حال، ممکن است همپوشانی‌هایی در این دسته‌بندی وجود داشته باشد، زیرا هر دو الگوریتم مبتنی بر ASM و غیر مبتنی بر ASM ممکن است با سایر عملیات‌های مبتنی بر بینایی کامپیوتر مانند بخش‌بندی، تشخیص و تخمین عمق ترکیب شوند. بنابراین، این بخش الگوریتم‌های نظارتی را بر اساس مشارکت اصلی آن‌ها دسته‌بندی می‌کند تا تکنیک‌هایی که برای تحقیقات رفع مه مفید هستند به وضوح مشاهده شوند.

۲- بررسی الگوریتم‌های مه زدایی تصاویر

همان‌طور که در بخش ۱ بیان شد، بر اساس مدل ASM، فرآیند از بین بردن مه به سه بخش تقسیم می‌شود: تخمین نقشه انتقال، پیش‌بینی نور جوی و بازیابی تصویر بدون مه. کلیه روش‌های انتخاب شده در این بخش مبتنی بر شبکه عصبی عمیق هستند و در یکی از این سه مرحله یا همه آن‌ها از یادگیری عمیق استفاده کرده‌اند. در انتهای بخش یک جدول مقایسه‌ای نیز قرار داده شده است.

MSCNN سه مرحله زیر را برای حل مدل ASM پیشنهاد می‌کند: (۱) استفاده از شبکه عصبی کانولوشنی (CNN) برای برآورد نقشه انتقال $t(x)$ ، (۲) استفاده از قوانین آماری برای پیش‌بینی نور جوی A ، (۳) حل $J(x)$ با استفاده از A و $t(x)$ به صورت مشترک.

MSCNN یک مدل کانولوشنی چند مقیاسی را برای برآورد نقشه انتقال اتخاذ می‌کند و آن را با خطای $L2$ بهینه می‌کند. علاوه بر این، A با انتخاب 0.1% تیره‌ترین پیکسل‌ها در $t(x)$ که با بالاترین شدت در $I(x)$ متناظر است، به دست می‌آید. بنابراین، تصویر واضح $J(x)$ می‌تواند به دست آید.

$$J(x) = \frac{I(x) - A}{\max\{0, 1 - t(x)\}} + A \quad (14)$$

مقالات مختلف ممکن است از پیش‌فرض‌های آماری متفاوتی برای تخمین A استفاده کنند، اما استراتژی‌های آن‌ها برای رفع مه مشابه MSCNN است. شبکه ABC-Net از عملیات max pooling برای به دست آوردن بیشترین مقدار از هر کانال تصویر $I(x)$ استفاده می‌کند. روش SID-JPM از یک فیلتر مینیمم برای هر کانال تصویر ورودی RGB استفاده می‌کند و سپس بیشترین مقدار هر کانال به عنوان A تخمین زده می‌شود. روش LAPTNet نیز فیلتر مینیمم و فیلتر ماکسیمم را برای پیش‌بینی A به کار می‌گیرد. این روش‌ها معمولاً به برچسب‌گذاری نور جوی نیاز ندارند، اما به تصاویر مه‌آلود و نقشه‌های انتقال مرتبط نیاز دارند.

۲-۱ یادگیری مشترک A و $t(x)$

به جای استفاده همزمان از شبکه‌های کانولوشنی و پیش‌فرض‌های آماری برای تخمین پارامترهای فیزیکی مدل ASM، برخی از تحقیقات، تخمین پارامترهای فیزیکی را به طور کامل از طریق شبکه‌های عصبی کانولوشن (CNN) انجام می‌دهند. روش DCPDN نقشه انتقال را از طریق یک شبکه رمزگشای هرمی متراکم تخمین می‌زند و از یک شبکه U-Net متقارن برای پیش‌بینی نور جوی A استفاده می‌کند. برای بهبود دقت لبه‌های نقشه انتقال، DCPDN یک تابع خطای ترکیبی با حفظ لبه

طراحی کرده که شامل خطای L2، خطای گرادیان دو جهته، و خطای لبه ویژگی است. روش DHD-Net الگوریتمی برای تخمین تراکم مه بر اساس تقسیم بندی طراحی کرده که می تواند مناطق مه غلیظ را تقسیم و مناطق کاندیدای نور جوی A را مشخص کند. روش HRGAN از یک شبکه کانولوشن با همجوشی چند مقیاسی برای پیش بینی $t(x)$ استفاده کرده و یک مدل کانولوشن تک لایه برای تخمین A به کار می برد. روش PMHLD از یک تولید کننده نقشه پچ (patch map generator) و شبکه پالایش شده (refines the network) برای تخمین نقشه انتقال استفاده کرده و از VGG-16 برای تخمین نور جوی استفاده می کند. شایان ذکر است که اگر از آموزش رگرسیون برای به دست آوردن A استفاده شود، معمولاً برچسب های واقعی نیاز است. در جدول ۳، چهار الگوریتم مه زدایی شامل DHD-Net، HRGAN و PMHLD با توجه به ویژگی های مختلف مقایسه شده اند:

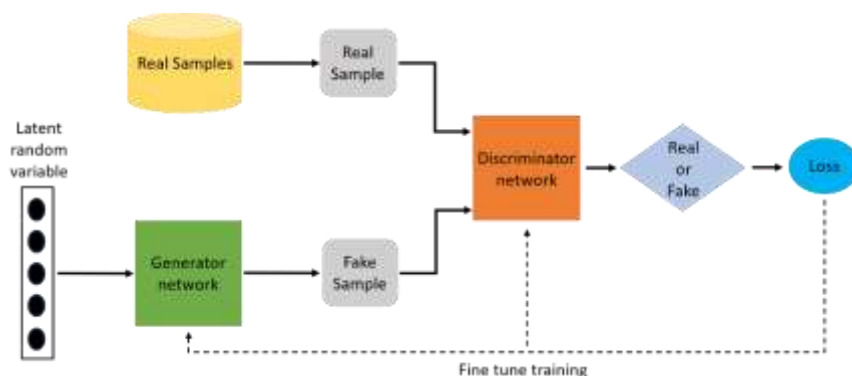
جدول ۳) مقایسه چهار الگوریتم مه زدایی شامل DHD-Net، HRGAN و PMHLD

ویژگی ها	DCPDF	DHD-Net	HRGAN	PMHLD
دقت	دقت بالا در شرایط مختلف	دقت بالا در تخمین عمق	دقت بالا در بازسازی جزئیات	دقت بالا در حذف مه
پیچیدگی محاسباتی	متوسط	بالا	بالا	متوسط
نیاز به داده های آموزشی	کم	زیاد	زیاد	متوسط
سرعت پردازش	متوسط	پایین	بالا	متوسط
قابلیت تعمیم	خوب	خوب	عالی	خوب
حساسیت به پارامترها	کم	متوسط	بالا	کم
استفاده از اطلاعات عمق	خیر	بله	خیر	بله

تحلیل جدول ۳ نشان می دهد که :

- DCPDF: مناسب برای شرایط مختلف با دقت بالا و پیچیدگی متوسط. نیاز به داده های آموزشی کمی دارد.
- DHD-Net: از دقت بالایی برخوردار است ولی پیچیدگی محاسباتی و نیاز به داده های آموزشی زیادی دارد. مناسب برای کاربردهایی که نیاز به تخمین عمق دارند.
- HRGAN: با استفاده از شبکه های GAN، دقت و سرعت بالایی در بازسازی جزئیات تصویر ارائه می دهد، اما حساسیت بالایی به پارامترها دارد.
- PMHLD: ترکیب یادگیری عمیق و اطلاعات عمق، دقت خوبی در حذف مه دارد و نیاز به داده های آموزشی متوسطی دارد.

۲-۲ استفاده از ASM بدون جاسازی صریح



شکل ۵) شبکه مقابله ای تولیدی برای حذف مه.

ASM می تواند به صورت بازفرمایی شده یا جاسازی شده در شبکه عصبی کانونی (CNN) قرار گیرد. مطالعه AOD-Net نشان داد که می توان از شبکه عصبی end-to-end برای حل مسئله ASM استفاده کرد بدون اینکه به صورت مستقیم از واقعیت زمینه $t(x)$ و A استفاده شود. طبق ASM اصلی، عبارت $J(x)$ به شکل زیر است:

$$J(x) = \frac{1}{t(x)} I(x) - A \frac{1}{t(x)} + A \quad (15)$$

AOD-Net پارامتری به نام $K(x)$ را پیشنهاد می‌کند که معنای فیزیکی خاصی ندارد، اما به عنوان یک پارامتر میانی برای توصیف $t(x)$ و A استفاده می‌شود. $K(x)$ به این صورت تعریف می‌شود که

$$K(x) = \frac{\frac{1}{t(x)}(I(x)-A)+(A-b)}{I(x)-1}, \quad (16)$$

در اینجا b برابر با ۱ است. بر اساس نظریه ASM، $J(x)$ را می‌توان با استفاده از $K(x)$ به طور منحصربه‌فرد تعیین کرد به این صورت که

$$J(x) = K(x)I(x) - K(x) + 1 \quad (17)$$

در شبکه AOD-Net نقشه انتقال غیر مشترک و فرآیند پیش‌بینی نور جوی ممکن است خطاهای تجمعی ایجاد کنند. بنابراین، برآورد مستقل $K(x)$ می‌تواند خطای سیستماتیک را کاهش دهد.

شبکه FAMED-Net این مفهوم را در یک چارچوب چند مقیاسی گسترش می‌دهد و از کانولوشن‌های کاملاً نقطه‌ای استفاده می‌کند تا عملکرد ضد مه‌زدایی سریع و دقیقی به دست آورد. DehazeGAN ایده برنامه‌نویسی قابل مشتق‌گیری را در فرآیند برآورد A و $t(x)$ به کار می‌گیرد. همراه با ASM بازفرموله‌شده، DehazeGAN همچنین یک سیستم ضد مه‌زدایی سرتاسری پیاده‌سازی می‌کند. PFDN، ASM را در طراحی شبکه جای می‌دهد و یک واحد ضد مه‌زدایی ویژگی‌ها را پیشنهاد می‌کند که مه را در یک فضای ویژگی‌های خوب طراحی شده به جای فضای تصویر خام حذف می‌کند. این نکته قابل توجه است که ASM هنوز می‌تواند به وظیفه ضد مه‌زدایی برای برآورد پارامترهای غیرمستقیم کمک کند. Generative adversarial networks تأثیر مهمی بر تحقیقات مربوط به مه‌زدایی دارند. به طور کلی، شبکه‌های مه‌زدایی نظارت‌شده که بر paired data تکیه می‌کنند، می‌توانند از adversarial loss به عنوان یک auxiliary supervisory signal استفاده کنند. Adversarial loss را می‌توان به عنوان دو بخش دید: هدف آموزشی مولد این است که تصاویری تولید کند تمییزدهنده آنها را به عنوان واقعی در نظر بگیرد. هدف بهینه‌سازی تمییزدهنده این است که تصویر تولید شده را از تصویر واقعی موجود در مجموعه داده به حد امکان تشخیص دهد. برای وظیفه مه‌زدایی، تأثیر adversarial loss این است که تصویر تولیدی را به تصویر واقعی نزدیک‌تر کند که این امر برای بهینه‌سازی $J(x)$ بدون مه و نقشه انتقال $t(x)$ مفید است.

patchGAN که می‌تواند اطلاعات فرکانس بالا را بهتر حفظ کند، DH-GAN، RI-GAN و DehazingGAN از $N \times N$ پیچ‌ها به جای یک مقدار تک برای خروجی discriminator استفاده می‌کنند. چندین تحقیق، مکانیسم‌های آموزش مشترک چند discriminator مانند EPDN و PGC-UNet را بررسی کردند. Discriminator D_1 برای هدایت مولد در مقیاس دقیق استفاده می‌شود، در حالی که discriminator D_2 به مولد کمک می‌کند تا خروجی global realistic را در coarse scale تولید کند. برای تحقق آموزش مشترک دو discriminators، EPDN تصویر ورودی D_1 را به نسبت ۲ به پایین نمونه‌برداری می‌کند و به عنوان ورودی D_2 استفاده می‌کند. adversarial loss به صورت زیر محاسبه می‌شود:

$$L_{adv} = \min_G \left[\max_{D_1, D_2} \sum_{k=1,2} \ell_A(G, D_k) \right] \quad (18)$$

در اینجا، شکل تابع ℓ_A adversarial loss مشابه با GAN ساده (single) است. استفاده از GAN ابزارهای قابل استفاده برای الگوریتم‌های نظارت‌شده فراهم می‌کند. از آنجا که آموزش تفکیک‌کننده (discriminator) به صورت بدون نظارت انجام می‌شود، کیفیت تصاویر مه‌زدایی شده بدون نیاز به برجسب‌های اضافی بهبود می‌یابد. با این حال، آموزش GAN گاهی با ناپایداری و عدم همگرایی مواجه می‌شود که ممکن است مشکلات اضافی برای آموزش شبکه مه‌زدایی ایجاد کند. در جدول ۴، الگوریتم‌های مه‌زدایی شامل AOD-Net، FAMED-Net، DehazeGAN، PFDN، ASM، Adversarial Net، PatchGAN، EPDN و PGC-UNet با توجه به ویژگی‌های مختلف مقایسه شده‌اند.

تحلیل جدول ۴ نشان می‌دهد که:

- AOD-Net: دقت خوبی دارد و پیچیدگی محاسباتی متوسطی را ارائه می‌دهد. نیاز به داده‌های آموزشی زیادی دارد.
- FAMED-Net: دقت عالی و قابلیت تعمیم خوبی دارد، اما پیچیدگی محاسباتی بالایی دارد.
- DehazeGAN: با استفاده از شبکه‌های GAN، دقت و کیفیت بالایی در حذف مه دارد، اما سرعت پردازش پایین است.

- ASM PFDN: دقت خوبی دارد و پیچیدگی محاسباتی متوسطی را ارائه می دهد. مناسب برای کاربردهای عمومی است.
- Adversarial Net: دقت خوبی دارد، اما پیچیدگی محاسباتی و نیاز به داده های آموزشی زیادی دارد.
- PatchGAN: دقت خوبی دارد و پیچیدگی محاسباتی متوسطی را ارائه می دهد. مناسب برای کاربردهای مختلف است.
- EPDN: دقت عالی و پیچیدگی محاسباتی بالایی دارد. نیاز به داده های آموزشی زیادی دارد.
- PGC-UNet: دقت و کیفیت بالایی در حذف مه دارد، اما پیچیدگی و نیاز به داده های آموزشی زیادی دارد.

جدول (۴) مقایسه الگوریتم های مه زدایی شامل AOD-Net, FAMED-Net, DehazeGAN, ASM PFDN, Adversarial Net, PatchGAN, EPDN و PGC-UNet

ویژگی ها	AOD-Net	FAMED-Net	DehazeGAN	ASM PFDN	Adversarial Net	PatchGAN	EPDN	PGC-UNet
دقت	خوب	عالی	عالی	خوب	خوب	خوب	عالی	عالی
پیچیدگی محاسباتی	متوسط	بالا	بالا	متوسط	بالا	متوسط	بالا	بالا
نیاز به داده های آموزشی	زیاد	زیاد	زیاد	متوسط	زیاد	زیاد	زیاد	زیاد
سرعت پردازش	متوسط	پایین	پایین	متوسط	پایین	متوسط	پایین	پایین
قابلیت تعمیم	خوب	خوب	عالی	خوب	خوب	خوب	خوب	خوب
حساسیت به پارامترها	متوسط	بالا	بالا	کم	بالا	متوسط	بالا	بالا
استفاده از اطلاعات عمق	خیر	خیر	خیر	خیر	خیر	خیر	خیر	خیر

۲-۳ روش های سطح-آگاه (Level-Aware)

بر اساس نظریه پراکندگی، هر چه صحنه دورتر از دوربین باشد، ذرات معلق بیشتری از آن عبور می کنند. این به این معناست که مناطقی در یک تصویر مه آلود که دورتر از دوربین هستند، چگالی مه بیشتری دارند. بنابراین، شبکه LAP-Net پیشنهاد می کند که الگوریتم باید تفاوت تراکم مه در داخل تصویر را در نظر بگیرد. با استفاده از آموزش چندمرحله ای مشترک، LAP-Net مدلی از easy-to-hard را پیاده سازی می کند که بر روی سطح مشخصی از چگالی مه تمرکز می کند، و این کار را از طریق یک stagewise loss انجام می دهد.

$$\hat{t}^s(x) = \begin{cases} \mathcal{F}(I(x), \theta^s), & \text{if } s = 1, \\ \mathcal{F}(I(x), \theta^s, \hat{t}^{s-1}(x)), & \text{if } s > 1 \end{cases} \quad (19)$$

در اینجا $\mathcal{F}$ نشان دهنده شبکه پیش بینی نقشه انتقال با پارامتر θ^s در مرحله s است. در مرحله اول، transmission map prediction network مسئول تخمین حالت با مه کم است. در مرحله دوم و مراحل بعدی، نتیجه پیش بینی مرحله قبلی و تصویر مه آلود به عنوان ورودی مشترک برای پردازش تراکم مه بیشتر استفاده می شوند. تراکم مه گرد ممکن است به شرایطی مانند دما، باد، ارتفاع و رطوبت وابسته باشد. بنابراین، تشکیل مه باید بر اساس فضا متغیر و غیر یکنواخت باشد. با توجه به این مشاهده، HardGAN ادعا می کند که estimating the transmission map برای حذف مه ممکن است دقت نداشته باشد. HardGAN با رمزگذاری روشنایی جوی به صورت یک ماتریس $1 \times 1 \times 2$ با عناصر γ_i^G و β_i^G و همچنین در نظر گرفتن اطلاعات فضایی هر پیکسل به عنوان یک ماتریس $H \times W \times 2$ با عناصر γ_i^L و β_i^L برای کانال نام از ورودی x تابع کنترل روشنایی جوی G_i و اطلاعات فضایی L_i را به صورت زیر طراحی و در نظر می گیرد:

$$\begin{aligned} G_i &= \gamma_i^G \frac{x - \mu}{\sigma} + \beta_i^G \\ L_i &= \gamma_i^L \frac{x - \mu}{\sigma} + \beta_i^L \end{aligned} \quad (20)$$

در اینجا، μ و σ به ترتیب میانگین و انحراف معیار برای x را نشان می دهند. پس از به دست آوردن G_i و L_i ، HardGAN از یک مدل خطی برای ترکیب آن ها استفاده می کند تا تصویر بدون مه را بازبایی کند.

$$J_{\text{pred}_i}(x) = (1 - HA_i) * G_i + HA_i * L_i, \quad (21)$$

در اینجا HA با استفاده از intermediate feature map of HardGAN با استفاده از نرمال سازی نمونه ای و سپس لایه سیگموئید محاسبه می شود، و علامت * ضرب عنصر در عنصر را نشان می دهد. در جدول ۵، دو الگوریتم مه زدایی شامل LAP-Net و HardGAN با توجه به ویژگی های مختلف مقایسه شده اند:

جدول ۵) مقایسه دو الگوریتم مه‌زدایی شامل LAP-Net و HardGAN

ویژگی‌ها	LAP-Net	HardGAN
دقت	خوب	عالی
پیچیدگی محاسباتی	متوسط	بالا
نیاز به داده‌های آموزشی	زیاد	زیاد
سرعت پردازش	متوسط	پایین
قابلیت تعمیم	خوب	عالی
حساسیت به پارامترها	متوسط	بالا
استفاده از اطلاعات عمق	خیر	خیر

تحلیل جدول ۵ نشان می‌دهد که :

- LAP-Net: دقت خوبی دارد و پیچیدگی محاسباتی متوسطی را ارائه می‌دهد. نیاز به داده‌های آموزشی زیادی دارد و برای کاربردهای عمومی مناسب است.
- HardGAN: دقت عالی و قابلیت تعمیم خوبی دارد، اما پیچیدگی محاسباتی بالایی دارد و سرعت پردازش آن پایین است.

۲-۴ ادغام چند عملکردی

DMMFD یک مدل جداسازی و تلفیق لایه‌ای برای بهبود توانایی یادگیری طراحی کرده است. این مدل شامل ASM بازفرمول شده، ضرب، جمع، توان و تجزیه لگاریتمی است:

$$\begin{aligned}
 J_0(x) &= \frac{I(x) - A_0 \times (1 - t_0(x))}{t_0(x)}, \\
 J_1(x) &= I(x) \times R_1(x), \\
 J_2(x) &= I(x) + R_2(x), \\
 J_3(x) &= (I(x))^{R_3(x)}, \\
 J_4(x) &= \log(1 + I(x) \times R_4(x)),
 \end{aligned} \tag{۲۲}$$

در اینجا $R_i(x)$ نمایانگر لایه‌های شبکه است؛ A_0 و $t_0(x)$ به ترتیب نور جوی و نقشه انتقال هستند که توسط شبکه استخراج ویژگی تخمین زده می‌شوند. $J_1(x)$ ، $J_2(x)$ ، $J_3(x)$ و $J_4(x)$ می‌توانند به عنوان چهار مدل مستقل برای جداسازی لایه‌های مه استفاده شوند. این مدل بر این فرض استوار است که تصویر مه‌دار ورودی $I(x)$ می‌تواند به یک لایه بدون مه $J(x)$ و یک لایه دیگر $H(x)$ تقسیم شود که به صورت $I(x) = \phi(J(x), H(x))$ نشان داده می‌شود. نتیجه نهایی حذف مه با ترکیب وزنی خروجی‌های میانی $J_0(x)$ ، $J_1(x)$ ، $J_2(x)$ ، $J_3(x)$ و $J_4(x)$ به وسیله پنج attention maps آموزش یافته W_0 ، W_1 ، W_2 ، W_3 و W_4 به دست می‌آید.

$$J_{\text{pred}}(x) = W_0 \times J_0(x) + W_1 \times J_1(x) + W_2 \times J_2(x) + W_3 \times J_3(x) + W_4 \times J_4(x) \tag{۲۳}$$

مطالعات ablation نشان داده است که مدل DMMFD با ترکیب چندین لایه می‌تواند کیفیت فرآیند بازسازی صحنه را بهبود بخشد.

۲-۵ تبدیل و تجزیه ورودی

GFN [۲۶] دو مشاهده در مورد تأثیر مه ارائه می‌دهد. اول، تحت تأثیر نور جوی، رنگ تصویر مه‌آلود ممکن است تا حدی تغییر کند. دوم، به دلیل پدیده‌های پراکندگی و کاهش نور، وضوح اشیای دورتر از دوربین در صحنه کاهش می‌یابد. بنابراین، GFN از سه استراتژی بهبود برای پردازش تصویر اصلی مه‌آلود استفاده می‌کند و آن‌ها را به عنوان ورودی‌های شبکه حذف مه به کار می‌برد. ورودی با تراز سفیدی $I_{wb}(x)$ از فرضیه دنیای خاکستری به دست می‌آید. ورودی با افزایش کنتراست $I_{ce}(x)$ از مقدار میانگین روشنایی $\tilde{I}(x)$ و عامل کنترل μ تشکیل شده، به صورت زیر:

$$I_{ce}(x) = \mu(I(x) - \tilde{I}(x)) \tag{۲۴}$$

در اینجا، $\mu = 2 \cdot (0.5 + \tilde{I}(x))$ است. با استفاده از اصلاح گاما غیرخطی، می‌توان از $I_{gc}(x)$ برای افزایش قابلیت رؤیت $I(x)$ به دست آورد:

$$I_{gc}(x) = \alpha I(x)^\gamma, \quad (25)$$

در اینجا α برابر با ۱ و γ برابر با ۲.۵ است. تصویر نهایی حذف مه $J(x)$ توسط ترکیب سه ورودی تعیین می شود، جایی که C_{wb} ، C_{ce} و C_{gc} در واقع confidence maps برای فرآیند تلفیق هستند:

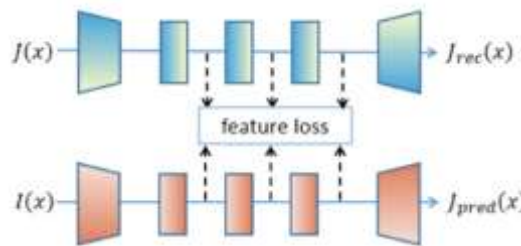
$$J(x) = C_{wb} \circ I_{wb}(x) + C_{ce} \circ I_{ce}(x) + C_{gc} \circ I_{gc}(x). \quad (26)$$

MSRL-DehazeNet تصویر مه آلود را به دو بخش فرکانس پایین و فرکانس بالا به عنوان مؤلفه پایه $I_{base}(x)$ و مؤلفه جزئیات $I_{detail}(x)$ تجزیه می کند. این مؤلفه پایه را می توان به عنوان محتوای اصلی تصویر در نظر گرفت، در حالی که مؤلفه فرکانس بالا حاوی لبه ها و بافت تصویر است. بنابراین، تصویر بدون مه می تواند با استفاده از تابع حذف مه $D(\cdot)$ از مؤلفه اصلی و تابع افزایش کیفیت $E(\cdot)$ از مؤلفه جزئی به دست آید. کل فرآیند می تواند توسط یک مدل خطی به شکل زیر نمایش داده شود:

$$I(x) = I_{base}(x) + I_{detail}(x) \quad (27)$$

$$J(x) = D(I_{base}(x)) + E(I_{detail}(x))$$

DIDH هم تصویر مه آلود ورودی $I(x)$ و هم تصویر پیش بینی شده بدون مه $J(x)$ را به صورت $(J(x), HF(J(x)))$ و $(I(x), LF(I(x)))$ تجزیه می کند، به صورتی که $LF(\cdot)$ و $HF(\cdot)$ به ترتیب فیلترهای گوسی و لاپلاسی را نشان می دهند. با تلفیق داده های تجزیه شده با داده های پیش تجزیه به عنوان ورودی دیسکریمیناتور، DIDH می تواند کیفیت تصویر تولید شده توسط adversarial training process را بهبود بخشد. با کمک دیسکریمیناتورهای DLF و DHF، شبکه حذف مه به شیوه adversarial بهینه می شود.



شکل ۶ استراتژی کلی knowledge distillation در KDDN [۲۵]، KTDN و SRKTDN [۴۷]

نسبت به روش های سنتی افزایش داده ها، مانند چرخش، تقارن افقی یا عمودی، و برش تصادفی، تبدیل و تجزیه ورودی استراتژی موثرتری برای استفاده از تصاویر مه آلود است. در جدول ۶، سه الگوریتم مه زدایی شامل GFN، MSRL-DehazeNet و DIDH با توجه به ویژگی های مختلف مقایسه شده اند:

جدول ۶ مقایسه سه الگوریتم مه زدایی شامل GFN، MSRL-DehazeNet و DIDH

ویژگی ها	GFN	MSRL-DehazeNet	DIDH
دقت	عالی	عالی	خوب
پیچیدگی محاسباتی	بالا	متوسط	متوسط
نیاز به داده های آموزشی	زیاد	زیاد	متوسط
سرعت پردازش	پایین	متوسط	متوسط
قابلیت تعمیم	عالی	خوب	خوب
حساسیت به پارامترها	بالا	متوسط	کم
استفاده از اطلاعات عمق	خیر	خیر	بله

تحلیل جدول ۶ نشان می دهد که :

- GFN: دقت بسیار بالا و قابلیت تعمیم عالی دارد، اما پیچیدگی محاسباتی آن بالا است و سرعت پردازش پایین است.
- MSRL-DehazeNet: دقت عالی و پیچیدگی محاسباتی متوسطی دارد. مناسب برای کاربردهای عمومی با نیاز به داده های آموزشی زیاد.
- DIDH: دقت خوبی دارد و از اطلاعات عمق استفاده می کند. پیچیدگی محاسباتی و نیاز به داده های آموزشی متوسطی دارد.

۲-۶ روش‌های مبتنی بر دانش انتقالی (Knowledge Distillation)

دانش انتقالی (Knowledge Distillation) یک استراتژی برای انتقال دانش یادگرفته‌شده توسط teacher network به student network ارائه می‌دهد. این روش در high-level computer vision tasks مانند تشخیص اشیا و طبقه‌بندی تصاویر استفاده شده‌است. تحقیقات اخیر سه چالش برای به‌کارگیری knowledge distillation در dehazing task بیان می‌کند. اول، چه نوع teacher task می‌تواند به وظیفه dehazing کمک کند. دوم، چگونه teacher network در طول آموزش به dehazing network کمک می‌کند. سوم، کدام معیار شباهت بین teacher task و student task باید انتخاب شود. روش‌های مختلف ممکن است از تعداد و مکان‌های خروجی ویژگی‌های مختلفی برای محاسبه feature loss استفاده کنند. KDDN یک مکانیزم یادگیری process-oriented طراحی می‌کند که در آن teacher network T یک auto-encoder برای بازسازی تصاویر بدون مه با کیفیت بالا است. هنگام آموزش شبکه مه‌زدایی، teacher network در یادگیری ویژگی‌ها کمک می‌کند و $L_T = \|J(x) - T(J(x))\|_1$ را بهینه می‌کند. بنابراین، teacher task و student task پیشنهاد شده توسط KDDN دو different tasks هستند. برای استفاده کامل از feature information که توسط teacher network یاد گرفته شده و کمک به آموزش شبکه مه‌زدایی، KDDN از تابع تبدیل خطی g برای استفاده از feature matching loss و haze density aware loss استفاده می‌کند که به ترتیب توسط روابط ۲۸ و ۲۹ نشان داده شده‌اند.

$$L_{rm} = \sum_{(m,n) \in \mathcal{C}} |T^m(J(x)) - g(S^n(I(x)))| \quad (28)$$

$$L_{wrm} = \sum_{(m,n) \in \mathcal{C}} \psi \times |T^m(J(x)) - g(S^n(I(x)))| \quad (29)$$

در اینجا، T^m نشان‌دهنده لایه mام teacher network است و متناظر آن، S^n نشان‌دهنده لایه nام شبکه دانش‌آموز است؛ ψ از عملیات نرمال‌سازی به‌دست می‌آید. می‌توان با استفاده از تفاضل بین تصاویر مه‌دار و بی‌مه به‌جای real transmission map، KDDN را آموزش داد. KTDN به کمک teacher network و dehazing network با ساختار یکسان مشترک آموزش داده می‌شود. از طریق feature level loss، prior knowledge possessed by the teacher network به شبکه مه‌زدایی انتقال داده می‌شود. SRKTDN از ResNet18 که بر روی ImageNet و به‌صورت teacher network پیش آموزش داده شده (با حذف لایه‌های طبقه‌بندی) و statistical experiences بسیاری را به رمزگشای Res2Net101 برای مه‌زدایی انتقال می‌دهد. DALF آموزش dual adversarial را در training process of knowledge distillation ادغام می‌کند تا imitation ability of the student network to the teacher network را بهبود بخشد. استفاده از knowledge distillation به شبکه‌های مه‌زدایی یک روش جدید و کارآمد برای معرفی external prior knowledge ارائه می‌دهد. در جدول ۷، چهار الگوریتم مه‌زدایی شامل KDDN، KTDN، SRKTDN و DALF با توجه به ویژگی‌های مختلف مقایسه شده‌اند.

جدول ۷، مقایسه چهار الگوریتم مه‌زدایی شامل KDDN، KTDN، SRKTDN و DALF

ویژگی‌ها	KDDN	KTDN	SRKTDN	DALF
دقت	خوب	عالی	عالی	خوب
پیچیدگی محاسباتی	متوسط	بالا	بالا	متوسط
نیاز به داده‌های آموزشی	متوسط	زیاد	زیاد	کم
سرعت پردازش	متوسط	پایین	پایین	متوسط
قابلیت تعمیم	خوب	عالی	عالی	خوب
حساسیت به پارامترها	کم	بالا	بالا	متوسط
استفاده از اطلاعات عمق	خیر	خیر	بله	خیر

تحلیل جدول ۷ نشان می‌دهد که:

- KDDN: دقت خوبی دارد و پیچیدگی محاسباتی متوسطی را ارائه می‌دهد. نیاز به داده‌های آموزشی متوسطی دارد و برای کاربردهای عمومی مناسب است.
- KTDN: دقت عالی و قابلیت تعمیم بالایی دارد، اما پیچیدگی محاسباتی و نیاز به داده‌های آموزشی زیادی دارد.
- SRKTDN: دقت و قابلیت تعمیم عالی دارد، اما پیچیدگی و نیاز به داده‌های آموزشی زیادی دارد. همچنین از اطلاعات عمق استفاده می‌کند.

- DALF: دقت خوبی دارد و پیچیدگی محاسباتی متوسطی را ارائه می دهد. نیاز به داده های آموزشی کمتری دارد و برای کاربردهای عمومی مناسب است.

۲-۷ روش های مبتنی بر تبدیل فضای رنگی

ورودی شبکه ی حذف مه، معمولاً تصاویر رنگی سه کاناله در حالت RGB هستند. بیانکو و همکارانش ۴۸ با محاسبه ی MSE تصاویر با مه و بدون مه در فضای RGB و YCrCb روی مجموعه داده ی Dense-Haze دریافتند که مه در این دو فضا تفاوت های عددی آشکاری نشان می دهد. مقدار خطا در کانال های قرمز، سبز و آبی در فضای RGB بسیار نزدیک به هم هستند. اما در فضای YCrCb، مقدار خطا در کانال روشنایی به طور قابل توجهی بیشتر از مولفه های رنگ آبی و قرمز است. AIP-Net نیز مقایسه ی مشابهی از تبدیل فضای رنگی روی synthetic datasets انجام داده و به همان نتیجه رسیده است. نتایج کمی که از آموزش شبکه ی حذف مه در فضای رنگی YCrCb به دست آمده، نشان می دهد که فضای RGB تنها حالت مؤثر برای روش های حذف مه مبتنی بر یادگیری عمیق نیست. علاوه بر این، TheiaNet عملکرد training the dehazing model در فضاهای رنگی RGB، YCrCb، و HSV را به طور جامع تحلیل کرده است. آزمایش ها نشان می دهند که تبدیل تصاویر از فضای RGB به دیگر فضاهای رنگی برای آموزش مدل، روشی مؤثر است.

۲-۸ روش های مبتنی بر یادگیری متضاد

در فرآیند آموزش یک شبکه ی حذف مه نظارتی که بر پایه ی ASM نیست، معمولاً تصویر دارای مه به عنوان ورودی شبکه استفاده می شود و انتظار می رود که تصویر واضحی به دست آید. در این فرآیند، تصویر واضح به عنوان یک positive example برای guide optimization of network استفاده می شود. AECD-Net با طراحی pairs of positive and negative examples دیدگاه جدیدی برای بررسی تصاویر مه آلود و بدون مه ارائه می دهد. به طور خاص، تصویر واضح $J(x)$ و تصویر بدون مه $J_{pred}(x)$ به عنوان یک positive sample pair در نظر گرفته می شوند، در حالی که تصویر دارای مه $I(x)$ و تصویر بدون مه $J_{pred}(x)$ به عنوان یک negative sample pair در نظر گرفته می شوند. بر اساس یادگیری تقابلی، $J_{pred}(x)$ ، $I(x)$ و $I(x)$ به ترتیب به عنوان نقطه مرجع، نمونه مثبت و نمونه منفی در نظر گرفته می شوند. برای مدل از پیش آموزش دیده $G(\cdot)$ ، خطای حذف مه به عنوان مجموع reconstruction loss و regularization term در نظر گرفته می شود، به شرح زیر:

$$\min \|J(x) - \psi(I(x))\|_1 + \lambda \sum_{i=1}^N \omega_i \cdot \frac{\|G_i(J(x)), G_i(\phi(I(x)))\|_1}{\|G_i(I(x)), G_i(\phi(I(x)))\|_1} \quad (30)$$

در اینجا G_i خروجی ویژگی لایه ی i -ام از مدل از پیش آموزش دیده را نشان می دهد؛ λ نسبت وزن بین خطای بازسازی تصویر و regularization term است؛ w_i عامل وزنی برای خروجی ویژگی ها است؛ و ϕ شبکه ی حذف مه است. AECD-Net استراتژی universal contrastive regularization strategy برای روش های موجود فراهم می کند بدون اینکه پارامتر اضافی به آنها اضافه کند.

۲-۹ روش های مبتنی بر خروجی غیرقطعی

روش های حذف مه مبتنی بر یادگیری عمیق معمولاً هدف خود را به دست آوردن یک تصویر بدون مه مشخص می گذارند. اما روش pWAE با معرفی تنسورهای نهان دو بعدی، می تواند تصاویر بدون مه با سبک های مختلف تولید کند که باعث گسترش هدف کلی آموزش می شود. روش pWAE برای حذف مه از فضای نهان «dehazing» (z_h) و فضای نهان «style» (z_s) استفاده می کند و از توابع میانگین (μ) و انحراف استاندارد (σ) برای transformation of the space به صورت زیر استفاده می شود:

$$z_{h \rightarrow s} = \sigma(z_s) \left(\frac{z_h - \mu(z_h)}{\sigma(z_h)} \right) + \mu(z_s) \quad (31)$$

یک سؤال طبیعی این است که چگونه می توان شدت نقشه برداری فضایی را تنظیم کرد تا degree of style transformation کنترل شود. pWAE از مازول خطی $z_h^s = \alpha z_{h \rightarrow s} + (1 - \alpha) z_h$ برای تنظیم وزن بین dehazed image و style information استفاده می کند. با تنظیم متغیر α ، می توان تصاویر dehazed با درجه های مختلف تبدیل شده به stylized متناظر با style

image را به دست آورد. تصور کنید عکسی دارید که مه گرفته است (مثل مه صبحگاهی). DehazeFlow یک روش جالب برای رفع این مه است. روش‌های قدیمی سعی می‌کردند با یک محاسبه‌ی قطعی، مه را به طور کامل از تصویر حذف کنند. اما DehazeFlow می‌گوید این کار سخت و غیرقابل اطمینان است. به همین دلیل، از یک شبکه‌ی عصبی جدید به نام "conditional normalizing flow network" استفاده می‌کند. این شبکه می‌تواند حدس بزند که تصویر اصلی (بدون مه) چه شکلی بوده است. نکته‌ی مهم این است که DehazeFlow فقط یک حدس نمی‌زند، بلکه چندین حدس احتمالی می‌زند. یعنی به جای اینکه بگوید تصویر اصلی حتماً این شکلی بوده، می‌گوید با توجه به این تصویر مه گرفته، چند حالت مختلف برای تصویر اصلی وجود دارد. این چند حالت مختلف به ما انعطاف‌پذیری می‌دهد. چون سلیقه‌ی آدم‌ها برای میزان و نوع رفع مه می‌تواند متفاوت باشد. پس DehazeFlow با حدس‌های مختلفش، به ما این امکان را می‌دهد که بهترین حالت را برای عکس خود انتخاب کنیم. در جدول ۸، دو الگوریتم مه‌زدایی شامل pEAE و DehazeFlow با توجه به ویژگی‌های مختلف مقایسه شده‌اند:

جدول ۸) مقایسه دو الگوریتم مه‌زدایی شامل pEAE و DehazeFlow

ویژگی‌ها	pEAE	DehazeFlow
دقت	عالی	عالی
پیچیدگی محاسباتی	متوسط	بالا
نیاز به داده‌های آموزشی	زیاد	زیاد
سرعت پردازش	متوسط	پایین
قابلیت تعمیم	خوب	عالی
حساسیت به پارامترها	متوسط	بالا
استفاده از اطلاعات عمق	خیر	خیر

تحلیل جدول ۸ نشان می‌دهد که :

- pEAE: دقت عالی و پیچیدگی محاسباتی متوسطی دارد. نیاز به داده‌های آموزشی زیادی دارد و برای کاربردهای عمومی مناسب است.
- DehazeFlow: دقت بسیار بالا و قابلیت تعمیم عالی دارد، اما پیچیدگی محاسباتی آن بالا است و سرعت پردازش پایین است.

۲-۱۰ مدل رتینکس (Retinex)

تحقیقاتی که بر پایه یادگیری عمیق نیستند ارتباط duality بین مدل Retinex و image dehazing را نشان می‌دهند. علاوه بر روش ASM که قبلاً به طور گسترده مورد استفاده قرار می‌گرفت، کار اخیر [۵۲] ترکیب مدل Retinex و شبکه عصبی پیچیده (CNN) را برای حذف مه بررسی می‌کند. تئوری Retinex این را پیشنهاد می‌دهد که تصویر می‌تواند به عنوان حاصل ضرب نقطه‌ای تصویر بازتاب R و نقشه روشنایی L در نظر گرفته شود. با فرض اینکه بازتاب R ثابت نور است، ارتباط بین تصویر مه‌آلود و تصویر بدون مه را می‌توان با استفاده از یک مدل تجزیه بر پایه retinex مدل کرد، مانند:

$$I(x) = J(x) * \frac{L_I}{L_J} = J(x) * L_r, \quad (32)$$

علامت * برای ضرب کردن مولفه به مولفه استفاده می‌شود. L_r را می‌توان به عنوان نور جذب شده و پراکنده شده در اثر مه در نظر گرفت که توسط نسبت نقشه روشنایی تصویر مه‌آلود L_I و نقشه روشنایی طبیعی L_J تعیین می‌شود. به عبارتی ساده‌تر، L_r نشان‌دهنده‌ی میزان تأثیر مه روی روشنایی تصویر است. این مقدار با ضرب کردن مقدار نور موجود در تصویر (L_I) در نسبتی از میزان نور طبیعی محیط (L_J) که تحت تأثیر مه قرار گرفته، به دست می‌آید. هرچه این مقدار بیشتر باشد، نشان می‌دهد مه نور بیشتری را جذب یا پراکنده کرده است. تحقیقات نشان داده که روش Retinex نسبت به روش رایج ASM از نظر فیزیکی مفهوم ساده‌تری و پارامترهای کمتری برای تخمین نیاز دارد اما هنوز به طور گسترده با شبکه‌های عصبی پیچیده (CNN) برای رفع مه ترکیب نشده است.

۲-۱۱ روش‌های مبتنی بر یادگیری باقیمانده (Residual Learning)

فرض کنید می خواهیم مه را از یک تصویر برداریم. به طور معمول، شبکه های عصبی باید یاد بگیرند که تصویر با مه را به تصویر بدون مه تبدیل کنند. اما روش های جدیدی به نام «یادگیری باقیمانده» وجود دارند که ادعا می کنند یادگیری شبکه را ساده تر می کنند. در یادگیری باقیمانده، به جای اینکه مستقیماً تصویر بدون مه را از تصویر با مه یاد بگیریم، شبکه یاد می گیرد که چه تغییراتی روی تصویر با مه انجام دهد تا به تصویر بدون مه برسیم. این تغییر، همان «باقیمانده» است. چندین روش حذف مه بر اساس یادگیری باقیمانده وجود دارد که در کل تصویر عمل می کنند. برای مثال، شبکه ای به نام GCANet [7] تفاوت (باقیمانده) $\{J(x) - I(x)\}$ بین تصویر بدون مه و تصویر با مه را به عنوان هدف بهینه سازی در نظر می گیرد. روش های دیگر مانند DRL، SID-HL و POGAN معتقدند که یادگیری باقیمانده با روشی به نام (ASM) ارتباط دارد و می توان این ارتباط را با بازتعریف ASM به دست آورد:

$$I(x) = J(x) + (A - J(x))(1 - t(x)) = J(x) + r(x), \quad (33)$$

در اینجا فرض بر این است که خطای موجود در تصویر مه گرفته $r(x)$ را می توان به صورت $r(x) = (A - J(x))(1 - t(x))$ نمایش داد. در این رابطه، A شدت مه است، $J(x)$ تصویر بدون مه و $t(x)$ هم یک سیگنال وابسته به خود تصویر است. به عبارت ساده تر، به جای اینکه مستقیماً تصویر بدون مه $J(x)$ را حدس بزنیم، شبکه سعی می کند خطای موجود در تصویر مه گرفته $r(x)$ را یاد بگیرد. با کم کردن این خطا از تصویر مه گرفته، می توانیم به تصویر بدون مه $J(x) = I(x) - r(x)$ برسیم. در جدول ۹، چهار الگوریتم مه زدایی شامل GCANet، DRL، SID-HL و POGAN با توجه به ویژگی های مختلف مقایسه شده اند:

جدول ۹. مقایسه چهار الگوریتم مه زدایی شامل GCANet، DRL، SID-HL و POGAN

ویژگی ها	GCANet	DRL	SID-HL	POGAN
دقت	عالی	خوب	عالی	عالی
پیچیدگی محاسباتی	بالا	متوسط	بالا	بالا
نیاز به داده های آموزشی	زیاد	کم	زیاد	زیاد
سرعت پردازش	پایین	متوسط	پایین	پایین
قابلیت تعمیم	عالی	خوب	خوب	عالی
حساسیت به پارامترها	بالا	متوسط	بالا	بالا
استفاده از اطلاعات عمق	خیر	خیر	بله	خیر

تحلیل جدول ۹ نشان می دهد که :

- GCANet: دقت عالی و قابلیت تعمیم بالایی دارد، اما پیچیدگی محاسباتی آن بالا است و نیاز به داده های آموزشی زیادی دارد.
- DRL: رویکرد یادگیری تقویتی دارد و نیاز به داده های آموزشی کمتری دارد. دقت آن خوب است و پیچیدگی محاسباتی متوسطی دارد.
- SID-HL: دقت بسیار بالا و پیچیدگی محاسباتی بالایی دارد. همچنین از اطلاعات عمق استفاده می کند و نیاز به داده های آموزشی زیادی دارد.
- POGAN: دقت عالی و قابلیت تعمیم خوبی دارد، اما پیچیدگی محاسباتی آن بالا است و سرعت پردازش پایین است.

۲-۱۲ روش های حوزه فرکانس (Frequency Domain)

روش های حذف مه با استفاده از شبکه های عصبی پیچشی (CNN) معمولاً از کم کردن و زیاد کردن ابعاد تصویر برای استخراج ویژگی و بازسازی تصویر بدون مه استفاده می کنند. اما این فرایند کمتر به اطلاعات فرکانسی موجود در تصویر توجه می کند. در حال حاضر دو رویکرد برای ترکیب تحلیل فرکانسی و شبکه های حذف مه وجود دارد: اولین روش: قرار دادن تابع تجزیه فرکانسی (مثل تبدیل موجک) درون شبکه ی عصبی پیچشی است. دومین روش: استفاده از تجزیه فرکانسی به عنوان یک محدودیت برای محاسبه ی خطا است. به عنوان مثال، با استفاده از تبدیل موجک می توان یک تصویر را به یک جزء فرکانس پایین و سه جزء فرکانس بالا تجزیه کرد. شبکه ی "Wavelet U-net" از تبدیل موجک گسسته (DWT) و تبدیل معکوس موجک گسسته (IDWT) استفاده می کند تا بتواند جزئیات مهم تصویر بدون مه، خصوصاً لبه ها، را با کیفیت بالا بازایی کند. از

طریق تابع مقیاس‌دهی یک‌بعدی $\phi(\cdot)$ و تابع موجک $\psi(\cdot)$ ، موجک‌های دو بعدی با سطوح پایین-پایین (LL)، پایین-بالا (LH)، بالا-پایین (HL)، و بالا-بالا (HH) به صورت زیر محاسبه می‌شوند:

$$\begin{aligned}\Phi_{LL}(m, n) &= \phi(m)\phi(n) \\ \Psi_{LH}(m, n) &= \phi(m)\psi(n) \\ \Psi_{HL}(m, n) &= \psi(m)\phi(n) \\ \Psi_{HH}(m, n) &= \psi(m)\psi(n),\end{aligned}\quad (34)$$

در اینجا m و n به ترتیب مختصات افقی و عمودی را نشان می‌دهند. شبکه‌ی MsGWN از تجزیه‌ی گابور موجک برای استخراج ویژگی و بازسازی تصویر استفاده می‌کند. این روش با تنظیم مقادیر زاویه می‌تواند اطلاعات فرکانسی را از جهات مختلف تصویر استخراج کند. شبکه‌ی EMRA-Net به جای کم کردن اندازه‌ی تصویر با روش‌های ساده، از تجزیه‌ی هار موجک (Haar Wavelet Decomposition) استفاده می‌کند. این روش به حفظ جزئیات بافت تصویر کمک می‌کند. تحقیقات اخیر نشان داده‌اند که با قرار دادن تحلیل موجک درون شبکه‌ی عصبی پیچشی یا تابع خطا، می‌توان از تبدیل موجک دوبعدی برای بهبود بازیابی اطلاعات فرکانس بالا در حوزه موجک استفاده کرد. این کارها با موفقیت نظریه‌ی موجک و شبکه‌های عصبی را برای حذف مه ترکیب می‌کنند. علاوه بر این، شبکه‌ی TDN از تابع خطای تبدیل سریع فوری به عنوان یک محدودیت در حوزه فرکانس استفاده می‌کند. این روش با آموزش داده‌شده بر روی دامنه و فاز تصویر (در مرحله‌ی آموزش با نظارت) کیفیت بصری تصاویر را بدون نیاز به محاسبات اضافی در مرحله‌ی خروجی بهبود می‌بخشد. در جدول ۱۰، چهار الگوریتم مه‌زدایی شامل Wavlet-U-net، MsGWN، EMRA-Net و TDN با توجه به ویژگی‌های مختلف مقایسه شده‌اند:

جدول ۱۰) مقایسه چهار الگوریتم مه‌زدایی شامل Wavlet-U-net، MsGWN، EMRA-Net و TDN

ویژگی‌ها	Wavlet-U-net	MsGWN	EMRA-Net	TDN
دقت	عالی	خوب	عالی	خوب
پیچیدگی محاسباتی	بالا	متوسط	بالا	متوسط
نیاز به داده‌های آموزشی	متوسط	زیاد	زیاد	متوسط
سرعت پردازش	متوسط	پایین	پایین	متوسط
قابلیت تعمیم	خوب	خوب	عالی	خوب
حساسیت به پارامترها	متوسط	بالا	بالا	متوسط
استفاده از اطلاعات عمق	خیر	خیر	بله	خیر

تحلیل جدول ۱۰ نشان می‌دهد که:

- Wavlet-U-net: دقت عالی و پیچیدگی محاسباتی بالایی دارد. نیاز به داده‌های آموزشی متوسطی دارد و برای کاربردهای عمومی مناسب است.
- MsGWN: دقت خوبی دارد و نیاز به داده‌های آموزشی زیادی دارد. پیچیدگی محاسباتی آن متوسط است و سرعت پردازش پایین است.
- EMRA-Net: دقت بسیار بالا و پیچیدگی محاسباتی بالایی دارد. همچنین از اطلاعات عمق استفاده می‌کند و نیاز به داده‌های آموزشی زیادی دارد.
- TDN: دقت خوبی دارد و پیچیدگی محاسباتی متوسطی را ارائه می‌دهد. نیاز به داده‌های آموزشی کمتری دارد و برای کاربردهای عمومی مناسب است.

۲-۱۳ حذف مه و تخمین عمق به صورت مشترک (Joint Dehazing and Depth Estimation)

ذرات معلق در محیط مه‌آلود می‌توانند بر دقت اطلاعات عمقی که توسط دستگاه لیدار (LiDAR) جمع‌آوری می‌شوند، تأثیر بگذارند. S2DNet نشان داده است که الگوریتم‌های تخمین عمق با کیفیت بالا می‌توانند به حذف مه کمک کنند. طبق مدل ASM، می‌دانیم که بین (نقشه انتقال نور/transmission map) $t(x)$ و (نقشه عمق/depth map) $d(x)$ یک رابطه‌ی نمایی معکوس وجود دارد: $t(x) = e^{-\beta d(x)}$ بر اساس این وابستگی فیزیکی، SDDE از چهار رمزگشا (decoder) استفاده می‌کند تا تخمین‌های نور هوا، تصویر بدون مه، نقشه انتقال نور و نقشه عمق را به صورت یکپارچه آموزش دهد. به طور خاص، SDDE یک

تابع خطای سازگاری عمق-انتقال (depth-transmission) را بر اساس این مشاهده پیشنهاد می کند که مقدار انحراف معیار (std) برای جفت نقشه انتقال نور (transmission map) و نقشه عمق (depth map) باید به سمت صفر میل کند:

$$L = \| \text{std}(\ln(t_{\text{pred}}(x)/d_{\text{pred}}(x))) \|_2, \quad (35)$$

در اینجا $t_{\text{pred}}(x)$ نشان دهنده ی نقشه انتقال نور پیش بینی شده (predicted transmission map) و $d_{\text{pred}}(x)$ نشان دهنده ی نقشه عمق پیش بینی شده است. شبکه ی TSDCN-Net برای بهبود عملکرد حذف مه، یک شبکه ی آبشاری (cascaded network) را برای روش های دو مرحله ای با پیش بینی اطلاعات عمق طراحی کرده است. نتایج آزمایشگاهی کمی نشان می دهند که این روش تخمین مشترک می تواند دقت کار حذف مه را بهبود بخشد. در جدول ۱۱، سه الگوریتم مه زدایی شامل S2DNet، SDDE و TSDCN-Net با توجه به ویژگی های مختلف مقایسه شده اند:

جدول ۱۱) مقایسه سه الگوریتم مه زدایی شامل S2DNet، SDDE و TSDCN-Net

ویژگی ها	S2DNet	SDDE	TSDCN-Net
دقت	عالی	خوب	عالی
پیچیدگی محاسباتی	بالا	متوسط	بالا
نیاز به داده های آموزشی	زیاد	متوسط	زیاد
سرعت پردازش	پایین	متوسط	پایین
قابلیت تعمیم	عالی	خوب	عالی
حساسیت به پارامترها	بالا	متوسط	بالا
استفاده از اطلاعات عمق	بله	خیر	بله

تحلیل جدول ۱۱ نشان می دهد که :

- S2DNet: دقت عالی و پیچیدگی محاسباتی بالایی دارد. نیاز به داده های آموزشی زیادی دارد و برای کاربردهای پیشرفته مناسب است.
- SDDE: دقت خوبی دارد و پیچیدگی محاسباتی متوسطی را ارائه می دهد. نیاز به داده های آموزشی کمتری دارد و برای کاربردهای عمومی مناسب است.
- TSDCN-Net: دقت عالی و پیچیدگی محاسباتی بالایی دارد. همچنین از اطلاعات عمق استفاده می کند و نیاز به داده های آموزشی زیادی دارد.

۳- نتیجه گیری

در نهایت برای رسیدن به یک جمع بندی کلی و برای نشان دادن یک تصویر کلی از نقاط ضعف روش های مختلف پیشنهاد شده تاکنون، یک کار مقایسه ای در جدول ۱۲ انجام شده است. روش های مبتنی بر خطا مانند [3,4,6,28,44] به دلیل نبود تصویر مرجع همواره تقریبی هستند. روش های مبتنی بر ویژگی و الگوی آماری [11,25] پویا و تطبیقی نیستند و ثابت حساب می شوند، بنابراین در شرایط مختلف افت کارایی دارند. [15] در مواقعی که تصویر ثابت و مقدار ویژگی صفر است دچار مشکل عدم تعریف لگاریتم است. شبکه های پیش آموزش داده شده مانند [17,36,47,49] عمومی طراحی شده اند و معمولاً کارایی نسبی مناسبی در حذف مه دارند. روش های مبتنی بر ارتباط بین بخش های مختلف تصویر دارای مه مانند [7,50-52] همه آماری هستند و در شرایط مختلف منجر به بروز خطاهای عدم قطعیت می شوند. تبدیل موجک [29,53] الگوی زمان فرکانس قوی تری نسبت به تبدیل فوریه دارد اما مشکلاتی در مدل کردن منحنی ها و جهت ها دارد. روش های مانند [54,55] نیز به داده های اضافه عمق و رادار یا سنسور نیاز دارند که همیشه در دسترس نیست. با توجه به این ایرادات هنوز نیاز به روش های مبتنی بر هوش مصنوعی و تطبیقی با هدف الگو کردن بهتر مه در شرایط مختلف و حذف آن احساس می شود.

۳-۱ اهمیت حذف مه برای تشخیص و دسته بندی اشیا

تحقیقات انجام شده [۳] نشان می دهند که وجود مه در تصویر می تواند مشکلات مختلفی را برای الگوریتم های تشخیص اشیا ایجاد کند، مانند جا افتادن اشیا (Missing Targets)، مکان یابی نادرست (Inaccurate Localizations) و پیش بینی غیرقابل اطمینان نوع شیء (Unconfident Category Prediction). همچنین، کارهای اخیر بیان کرده اند که مه باعث ایجاد مشکل

در درک معنایی صحنه‌ی تصویر می‌شود. به‌طور معمول، فرایند حذف مه به‌عنوان یک مرحله‌ی پیش‌پردازش برای کارهای سطح بالای بینایی کامپیوتر در نظر گرفته می‌شود و از آن‌ها جداگانه انجام می‌شود. لی و همکاران [۳] الگوریتم شناسایی اشیا و AOD-Net را به‌طور مشترک بهینه‌سازی کرده‌اند و نتایج نشان دادند که الگوریتم حذف مه می‌تواند به بهبود عملکرد تشخیص اشیا کمک کند. LEAAL [۵] پیشنهاد می‌کند که تنظیم دقیق پارامترهای آشکارساز اشیا در حین آموزش مشترک ممکن است منجر به یک آشکارساز متمایل به تصاویر بدون مه تولید شده توسط شبکه مه‌زدایی از پیش آموزش‌دیده شود. برخلاف عملیات تنظیم دقیق لی و همکاران، LEAAL از شناسایی اشیا به‌عنوان یک وظیفه کمکی برای مه‌زدایی استفاده می‌کند و پارامترهای آشکارساز اشیا به‌طور کامل در طول فرآیند آموزش به‌روزرسانی نمی‌شوند.

جدول ۱۲) مقایسه روش‌های مختلف حذف مه و نقاط قوت و ضعف آنها

دسته بندی	دقت بالا	انعطاف پذیری	داده های آموزشی	تعمیم پذیری	عدم نیاز به HW	حساسیت به پارامتر	احتمال بیش‌برازش	پیچیدگی زمانی
یادگیری مشترک $A_{t(x)}$	✓	×	×	✓	×	×	×	×
استفاده از ASM بدون جاسازی	×	✓	✓	×	✓	×	✓	✓
روش های سطح-آگاه	✓	×	×	×	✓	×	✓	×
ادغام چند عملکردی	✓	✓	×	×	✓	×	✓	×
تبدیل و تجزیه ورودی	×	×	×	×	✓	✓	✓	×
روش های مبتنی بر دانش انتقالی	×	✓	✓	×	✓	✓	✓	×
تبدیل فضای رنگی	×	×	×	×	✓	✓	✓	×
یادگیری متضاد	✓	×	×	✓	✓	✓	✓	×
خروجی غیرقطعی	✓	✓	×	×	✓	×	✓	×
مدل رتینکس (Retinex)	×	×	×	×	✓	×	✓	×
یادگیری باقیمانده	✓	×	×	✓	✓	×	×	×
حوزه فرکانس	×	×	×	×	✓	×	✓	×
حذف مه و تخمین عمق مشترک	✓	×	×	×	×	✓	✓	✓

UDnD [۲۰] از مزایای هر دو روش (حذف مه و تشخیص شیء) با آموزش مشترک یک شبکه‌ی حذف مه و یک تشخیص‌دهنده‌ی چندمنطقه‌ای اشیا با قابلیت تشخیص تراکم (dense-aware) استفاده می‌کند. شبکه‌ی تشخیص شیء با استفاده از اصطلاحات طبقه‌بندی و مکان‌یابی که در شبکه‌ی پیشنهاد ناحیه (Region Proposal Network) و ناحیه‌ی موردعلاقه (Region of Interest) استفاده می‌شود، آموزش داده می‌شود. رویکرد آموزش چندوظیفه‌ای که توسط UDnD به کار گرفته می‌شود، می‌تواند شکاف‌های کاهش‌یافته بین حوزه‌های مختلف و شکاف‌های باقیمانده درون یک حوزه برای سطوح تراکم متفاوت را در نظر بگیرد.

تحقیقات اخیر نشان داده‌اند که می‌توان حذف مه و کارهای سطح بالای بینایی کامپیوتر را بدون استفاده از مدل (ASM) به‌صورت همزمان انجام داد. SDNet [۲۴] تقسیم‌بندی معنایی و مه‌زدایی را در یک چارچوب یکپارچه ترکیب می‌کند تا از اطلاعات معنایی به‌عنوان یک محدودیت برای فرآیند بهینه‌سازی استفاده کند. با تعبیه‌ی پیش‌بینی‌های نقشه تفکیک در شبکه‌ی حذف مه، SDNet به‌طور همزمان یک بهینه‌سازی مشترک از تابع خطای طبقه‌بندی برای تک‌تک پیکسل‌ها و تابع خطای رگرسیون انجام می‌دهد. زبان طبقه‌بندی (Classification loss) عبارت است از:

$$L_{\text{sem}}(s, s^*) = -\frac{1}{p} \sum_i s_i^* \log(s_i), \quad (36)$$

در اینجا P تعداد کل پیکسل‌ها؛ s_i پیش‌بینی طبقه‌بندی برای موقعیت (پیکسل) i ؛ نشان‌دهنده‌ی برچسب‌گذاری دستی و دقیق برای آن موقعیت (ground-truth semantic annotation) است. تحقیقات اخیر [۲۰، ۲۴، ۴۵] نشان داده‌اند که با ترکیب کردن کارهای تفکیک و تشخیص اشیا با شبکه‌های حذف مه، می‌توان این کارها را به‌صورت همزمان انجام داد. این روش مشترک حذف مه و کارهای سطح بالای بینایی کامپیوتر می‌تواند تا حدودی بار محاسباتی را کاهش دهد، زیرا ویژگی‌های یاد گرفته شده در هر دو بخش قابل اشتراک هستند. همچنین، این رویکرد می‌تواند اهداف و کاربردهای پژوهش در زمینه‌ی حذف مه را گسترش دهد.

۲-۳ شبکه های عصبی پیچشی سر تاسر (End-to-End CNN) برای حذف مه

اصطلاح "End-to-End" برای الگوریتم های نظارت شده مبتنی بر عدم استفاده از ASM استفاده می شود که معمولاً از شبکه های عصبی با طراحی مناسب تشکیل شده اند که یک تصویر مه آلود را به عنوان ورودی می گیرند و یک تصویر بدون مه را به عنوان خروجی تولید می کنند. این شبکه ها بر اساس ایده های مختلفی ساخته شده اند که در ادامه به مهم ترین آن ها اشاره می کنیم:

- مکانیزم توجه: FFA-Net [۳۰]، GridDehazeNet [۳۲]، SAN [۱۸]، HFF [۵۶].
- رمز گشایی-رمز گشایی: CAE [۲۲].
- مبتنی بر بلوک متراکم: 123-CEDH [۲۳].
- ساختار U - شکل: MSBDN [۳۷]، DSEU [۸].
- شبکه سلسله مراتبی: DMHN [۵۶].
- ادغام با یادگیری شبکه بی درنگ: 4kDehazing [۵۷].

شبکه های عصبی کانولوشنال End-to-End تأثیر مهمی بر کل حوزه ی حذف مه داشته اند و نشان داده اند که مدل های یادگیری عمیق متعددی می توانند برای حذف مه مفید باشند.

جدول ۱۳) چالش های مساله مه زدایی تصاویر دیجیتال

چالش	علت اصلی	تأثیر بر تصویر	راهکارهای احتمالی
ابهام در مدل پراکندگی نور	عدم قطعیت در تخمین نور جو و نقشه انتقال	کاهش دقت در مناطق با مه غلیظ یا نورپردازی پیچیده	ترکیب مدل های فیزیکی با یادگیری عمیق (مثل شبکه های مبتنی بر توجه)
اختلال رنگ ها	تخمین نادرست نور جو یا اصلاح بیش از حد طیف رنگ	رنگ های غیرطبیعی در آسمان یا سطوح روشن	استفاده از محدودیت های رنگی در تابع هزینه شبکه عصبی
از بین رفتن جزئیات	حذف ناخواسته بافت ها هنگام حذف مه یا تقویت نویز	محو شدن لبه ها یا افزایش نویز در مناطق کم مغز (مثل برگ درختان)	بهینه سازی نقشه انتقال با فیلترهای حفظ جزئیات (مثل Guided Filter)
پردازش زمان واقعی	محاسبات سنگین در مدل های پیچیده (GANs، شبکه های عمیق)	تأخیر در پردازش تصاویر با رزولوشن بالا یا ویدیوها	کاهش پارامترهای مدل با معماری های سبک (معل) یا استفاده از سخت افزارهای خاص
عدم تعمیم پذیری	تفاوت بین داده های مصنوعی (آموزش) و تصاویر واقعی (تست)	عملکرد ضعیف در شرایط نوری متنوع یا مه های نامتعارف	افزایش داده های آموزشی با تصاویر واقعی و Augmentation

با بررسی سوابق موجود می توان در نهایت به چالش های اصلی در حوزه پژوهش های مساله مه زدایی تصاویر دیجیتال رسید. چالش های اصلی در زمینه مه زدایی تصاویر عبارتند از:

۱. ابهام در مدل های فیزیکی پراکندگی نور: رابطه بین مه و نور محیط (مثل نور خورشید) اغلب غیرخطی و وابسته به عمق صحنه است. مثلاً در مناطق با مه های غلیظ، تخمین نور جو (Atmospheric Light) و نقشه انتقال (Transmission Map) دقت کمتری دارد.
۲. اختلال در رنگ های طبیعی تصویر: علت: روش های سنتی ممکن است با اصلاح بیش از حد، رنگ های آسمان یا سطوح سفید را به صورت غیرواقعی نمایش دهند.
۳. از بین رفتن جزئیات ریز: چالش: الگوریتم های قوی ممکن است نویز را تقویت کنند یا بافت های ظریف (مثل برگ درختان) را محو کنند.
۴. پردازش در زمان واقعی: مدل های پیچیده (مثل GANs) برای تصاویر با رزولوشن بالا یا ویدیوها به منابع محاسباتی سنگین نیاز دارند.
۵. عدم تطابق داده های آموزشی با دنیای واقعی: داده های مصنوعی (مثل RESIDE) ممکن است شرایط نوری یا غلظت مه در تصاویر واقعی را پوشش ندهند.

چالش های اصلی مه زدایی تصاویر در جدول ۱۳ مقایسه ای ارائه می شود تا امکان تحلیل ساده تر فراهم شود. چالش های بیان شده بعضاً وابستگی متقابل دارند مثلاً بهبود تعمیم پذیری ممکن است پردازش زمان واقعی را دشوارتر کند. از سوی دیگر اولویت

یندی راهکارها در بررسی و پاسخ به این چالش‌ها می‌بایست در نظر گرفته شود به عنوان مثال در کاربردهای پزشکی، حفظ جزئیات از سرعت پردازش مهم‌تر است.

مسئله حذف مه از تصاویر دیجیتال به عنوان یک چالش مهم در زمینه بینایی کامپیوتری مطرح است. وجود مه در تصاویر می‌تواند مشکلاتی جدی برای الگوریتم‌های تشخیص اشیا ایجاد کند، از جمله جا افتادن اشیا، مکان‌یابی نادرست و پیش‌بینی غیرقابل اطمینان نوع شیء. به همین دلیل، حذف مه به عنوان یک مرحله پیش‌پردازش ضروری برای بهبود عملکرد سیستم‌های بینایی کامپیوتری در کاربردهای مختلف، از جمله خودروهای خودران، نظارت تصویری و تشخیص اشیا، اهمیت دارد. با افزایش استفاده از فناوری‌های هوش مصنوعی و یادگیری ماشین، نیاز به روش‌های مؤثر و کارآمد برای حذف مه و بهبود کیفیت تصاویر در شرایط مختلف بیش از پیش احساس می‌شود.

مدل‌سازی مسئله حذف مه شامل شناسایی و تجزیه و تحلیل عوامل مختلفی است که بر کیفیت تصویر تأثیر می‌گذارد. این عوامل شامل نور محیط، عمق صحنه و ویژگی‌های فیزیکی مه هستند. برای مدل‌سازی این مسئله، معمولاً از تکنیک‌های ریاضی و فیزیکی مانند مدل‌های پراکندگی نور و نقشه‌های انتقال استفاده می‌شود. این مدل‌ها به ما کمک می‌کنند تا رابطه بین مه و نور محیط را بهتر درک کنیم و به تبع آن، روش‌های مؤثری برای حذف مه از تصاویر توسعه دهیم. همچنین، استفاده از شبکه‌های عصبی عمیق و تکنیک‌های یادگیری ماشین به ما این امکان را می‌دهد که از داده‌های آموزشی برای بهبود دقت و کارایی الگوریتم‌ها بهره‌برداری کنیم.

روش‌های حذف مه به طور کلی به چند دسته تقسیم می‌شوند: روش‌های مبتنی بر خطا، روش‌های مبتنی بر ویژگی، روش‌های پیش‌آموزش داده‌شده و روش‌های آماری. روش‌های مبتنی بر خطا، مانند [۳، ۴، ۶، ۲۸، ۴۴]، به دلیل عدم وجود تصویر مرجع، معمولاً نتایج تقریبی ارائه می‌دهند. در مقابل، روش‌های مبتنی بر ویژگی و الگوهای آماری [۱۱، ۲۵] به دلیل عدم تطبیق با شرایط متغیر، در برخی موارد کارایی کمتری دارند. شبکه‌های پیش‌آموزش داده‌شده [۱۷، ۳۶، ۴۷، ۴۹] معمولاً کارایی مناسبی در حذف مه دارند، اما ممکن است در شرایط خاص دچار محدودیت شوند. همچنین، روش‌های آماری که به تحلیل روابط بین بخش‌های مختلف تصویر می‌پردازند [۷، ۵۰-۵۲] می‌توانند منجر به بروز خطاهای عدم قطعیت شوند. از سوی دیگر، تبدیل موجک [۲۹، ۵۳] نمایندگی زمان-فرکانس قوی‌تری نسبت به تبدیل فوریه دارد، اما در مدل‌سازی منحنی‌ها و جهت‌ها مشکلاتی دارد. به طور کلی، این مقایسه نشان می‌دهد که هر روش دارای مزایا و معایب خاص خود است و هنوز نیاز به توسعه روش‌های جدید و ترکیبی احساس می‌شود.

چالش‌های اصلی در حوزه مه‌زدایی تصاویر دیجیتال شامل ابهام در مدل‌های فیزیکی پراکندگی نور، اختلال در رنگ‌های طبیعی، از بین رفتن جزئیات ریز، نیاز به پردازش در زمان واقعی و عدم تطابق داده‌های آموزشی با شرایط واقعی است. ابهام در مدل‌های پراکندگی نور به دلیل غیرخطی بودن رابطه بین مه و نور محیط، به ویژه در شرایط غلیظ، دقت تخمین نور جو و نقشه انتقال را کاهش می‌دهد. اختلال در رنگ‌ها نیز ممکن است منجر به نمایش غیرواقعی رنگ‌های آسمان یا سطوح روشن شود. همچنین، الگوریتم‌های قوی ممکن است نویز را تقویت کنند یا بافت‌های ظریف را محو کنند. پردازش در زمان واقعی نیز به منابع محاسباتی سنگین نیاز دارد که ممکن است در کاربردهای عملی مشکل‌ساز باشد. در نهایت، عدم تطابق داده‌های آموزشی با شرایط واقعی می‌تواند منجر به عملکرد ضعیف الگوریتم‌ها در شرایط نوری متنوع شود.

در جمع‌بندی، حذف مه از تصاویر دیجیتال یک چالش مهم در زمینه بینایی کامپیوتری است که نیاز به روش‌های مؤثر و کارآمد دارد. با توجه به پیشرفت‌های اخیر در تکنیک‌های یادگیری عمیق و هوش مصنوعی، امیدواری‌هایی برای بهبود عملکرد و دقت الگوریتم‌های حذف مه وجود دارد. با این حال، چالش‌های متعددی در این زمینه وجود دارد که نیاز به تحقیقات بیشتر و توسعه روش‌های نوآورانه را ضروری می‌سازد. به طور کلی، این حوزه همچنان به عنوان یک موضوع جذاب و فعال در تحقیقات علمی و صنعتی باقی مانده است.

References

مراجع

- [1] B. Cai, X. Xu, K. Jia, C. Qing, D. Tao, "Dehazenet: An end-to-end system for single image haze removal", *IEEE Trans. Image Process.*, vol. 25, no. 11, pp. 5187-5198, 2016.

- [2] E. J. McCartney, *Optics of the Atmosphere: Scattering by Molecules and Particles*, Physics Bulletin, 1976, pp. 1-421.
- [3] B. Li, X. Peng, Z. Wang, J. Xu, D. Feng, "AOD-Net: All-in-one dehazing network", in *Proc. Int. Conf. Computer Vision*, pp. 4780-4788, 2017.
- [4] H. Zhang and V. M. Patel, "Densely connected pyramid dehazing network", in *Proc. Conf. Computer Vision Pattern Recognit.*, pp. 3194-3203, 2018.
- [5] C. O. Ancuti, C. Ancuti, R. Timofte, C. De Vleeschouwer, "O-HAZE: A dehazing benchmark with real hazy and haze-free outdoor images", in *Proc. Conf. Computer Vision Pattern Recognit. Workshops*, pp. 754-762, 2018.
- [6] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks", in *Proc. Eur. Conf. Computer Vision*, pp. 154-169, 2016.
- [7] D. Chen, M. He, Q. Fan, J. Liao, L. Zhang, D. Hou, L. Yuan, G. Hua, "Gated context aggregation network for image dehazing and deraining", in *Proc. Winter Conf. Applications Computer Vision*, pp. 1375-1383, 2019.
- [8] Y.-W. Lee, L.-K. Wong, J. See, "Image dehazing with contextualized attentive U-NET", in *Proc. Int. Conf. Image Process.*, pp. 1068-1072, 2020.
- [9] A. Dudhane, H. S. Aulakh, S. Murala, "RI-GAN: An end-to-end network for single image haze removal", in *Proc. Conf. Computer Vision Pattern Recognit. Workshops*, pp. 2014-2023, 2019.
- [10] C. Wang, Y. Zou, Z. Chen, "ABC-NET: Avoiding blocking effect and color shift network for single image dehazing via restraining transmission bias", in *Proc. Int. Conf. Image Process.*, pp. 1053-1057, 2020.
- [11] A. Wang, W. Wang, J. Liu, N. Gu, "AIPNet: Image-to-image single image dehazing with atmospheric illumination prior", *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 381-393, 2018.
- [12] G. Tang, L. Zhao, R. Jiang, X. Zhang, "Single image dehazing via lightweight multi-scale networks", in *Proc. Int. Conf. Big Data*, pp. 5062-5069, 2019.
- [13] S. Chen, Y. Chen, Y. Qu, J. Huang, M. Hong, "Multi-scale adaptive dehazing network", in *Proc. Conf. Computer Vision Pattern Recognit. Workshops*, pp. 2051-2059, 2019.
- [14] A. Singh, A. Bhavne, D. K. Prasad, "Single image dehazing for a variety of haze scenarios using back projected pyramid network", in *Proc. Eur. Conf. Computer Vision*, pp. 166-181, 2020.
- [15] Z. Deng, L. Zhu, X. Hu, C.-W. Fu, X. Xu, Q. Zhang, J. Qin, P.-A. Heng, "Deep multi-model fusion for single-image dehazing", in *Proc. Int. Conf. Computer Vision*, pp. 2453-2462, 2019.
- [16] X. Chen, H. Lu, K. Cheng, Y. Ma, Q. Zhou, Y. Zhao, "Sequentially refined spatial and channel-wise feature aggregation in encoder-decoder network for single image dehazing", in *Proc. Int. Conf. Image Process.*, pp. 2776-2780, 2019.
- [17] M. Hong, Y. Xie, C. Li, Y. Qu, "Distilling image dehazing with heterogeneous task imitation", in *Proc. Conf. Computer Vision Pattern Recognit.*, pp. 3462-3471, 2020.
- [18] X. Liang, R. Li, J. Tang, "Selective attention network for image dehazing and deraining", in *Proc. ACM Multimedia Asia*, pp. 1-6, 2019.
- [19] Y. Qu, Y. Chen, J. Huang, Y. Xie, "Enhanced Pix2pix dehazing network", in *Proc. Conf. Computer Vision Pattern Recognit.*, pp. 8152-8160, 2019.
- [20] Z. Zhang, L. Zhao, Y. Liu, S. Zhang, J. Yang, "Unified density-aware image dehazing and object detection in real-world hazy scenes", in *Proc. Asian Conf. Computer Vision*, pp. 119-135, 2020.
- [21] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition", in *Proc. Conf. Computer Vision Pattern Recognit.*, pp. 770-778, 2016.
- [22] R. Chen and E. M.-K. Lai, "Convolutional autoencoder for single image dehazing", in *Proc. Int. Conf. Image Process.*, pp. 4464-4470, 2019.
- [23] T. Guo, V. Cherukuri, V. Monga, "Dense '123' color enhancement dehazing network", in *Proc. Conf. Computer Vision Pattern Recognit. Workshops*, pp. 2131-2139, 2019.
- [24] S. Zhang, W. Ren, X. Tan, Z.-J. Wang, Y. Liu, J. Zhang, X. Zhang, X. Cao, "Semantic-aware dehazing network with adaptive feature fusion", *IEEE Trans. Cybern.*, vol. 53, no. 1, pp. 454-467, 2023.
- [25] J. Dong and J. Pan, "Physics-based feature dehazing networks", in *Proc. Eur. Conf. Computer Vision*, pp. 188-204, 2020.
- [26] W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, M.-H. Yang, "Gated fusion network for single image dehazing", in *Proc. Conf. Computer Vision Pattern Recognit.*, pp. 3253-3261, 2018.

- [27] S. Yin, X. Yang, Y. Wang, Y.-H. Yang, "Visual attention dehazing network with multilevel features refinement and fusion", *Pattern Recognit.*, vol. 118, p. 108021, 2021.
- [28] L. Yan, W. Zheng, C. Gou, F.-Y. Wang, "Feature aggregation attention network for single image dehazing", in *Proc. Int. Conf. Image Process.*, pp. 923-927, 2020.
- [29] J. Wang, C. Li, S. Xu, "An ensemble multi-scale residual attention network (EMRA-net) for image dehazing", *Multimedia Tools Appl.*, vol. 80, pp. 29299-29319, 2021.
- [30] X. Qin, Z. Wang, Y. Bai, X. Xie, H. Jia, "FFA-Net: Feature fusion attention network for single image dehazing", in *Proc. AAAI Conf. Artificial Intelligence*, pp. 11908-11915, 2020.
- [31] K. Metwaly, X. Li, T. Guo, V. Monga, "Nonlocal channel attention for nonhomogeneous image dehazing", in *Proc. Conf. Computer Vision Pattern Recognit. Workshops*, pp. 452-453, 2020.
- [32] X. Liu, Y. Ma, Z. Shi, J. Chen, "GridDehazeNet: Attention-based multi-scale network for image dehazing", in *Proc. Int. Conf. Computer Vision*, pp. 7313-7322, 2019.
- [33] L. Liu, B. Liu, H. Huang, A. C. Bovik, "No-reference image quality assessment based on spatial and spectral entropies", *Signal Process.: Image Commun.*, vol. 29, no. 8, pp. 856-863, 2014.
- [34] H. Dong, X. Zhang, Y. Guo, F. Wang, "Deep multi-scale Gabor wavelet network for image restoration", in *Proc. Int. Conf. Acoustics, Speech Signal Processing*, pp. 2028-2032, 2020.
- [35] W.-T. Chen, J.-J. Ding, S.-Y. Kuo, "PMS-Net: Robust haze removal based on patch map for single images", in *Proc. Conf. Computer Vision Pattern Recognit.*, pp. 11673-11681, 2019.
- [36] W.-T. Chen, H.-Y. Fang, J.-J. Ding, S.-Y. Kuo, "PMHLD: Patch map-based hybrid learning DehazeNet for single image haze removal", *IEEE Trans. Image Process.*, vol. 29, pp. 6773-6788, 2020.
- [37] H. Dong, J. Pan, L. Xiang, Z. Hu, X. Zhang, F. Wang, M.-H. Yang, "Multi-scale boosted dehazing network with dense feature fusion", in *Proc. Conf. Computer Vision Pattern Recognit.*, pp. 2157-2167, 2020.
- [38] T. Guo, X. Li, V. Cherukuri, V. Monga, "Dense scene information estimation network for dehazing", in *Proc. Conf. Computer Vision Pattern Recognit. Workshops*, pp. 2122-2130, 2019.
- [39] Y. Dong, Y. Liu, H. Zhang, S. Chen, Y. Qiao, "FD-GAN: Generative adversarial networks with fusion-discriminator for single image dehazing", in *Proc. AAAI Conf. Artificial Intelligence*, pp. 10729-10736, 2020.
- [40] E. Jo and J.-Y. Sim, "Multi-scale selective residual learning for non-homogeneous dehazing", in *Proc. Conf. Computer Vision Pattern Recognit. Workshop*, pp. 507-515, 2021.
- [41] X. Cong, J. Gui, K.-C. Miao, J. Zhang, B. Wang, P. Chen, "Discrete haze level dehazing network", in *Proc. ACM Int. Conf. Multimedia*, pp. 1828-1836, 2020.
- [42] L. He, J. Bai, M. Yang, "Feature aggregation convolution network for haze removal", in *Proc. Int. Conf. Image Process.*, pp. 2806-2810, 2019.
- [43] Y. Shao, L. Li, W. Ren, C. Gao, N. Sang, "Domain adaptation for image dehazing", in *Proc. Conf. Computer Vision Pattern Recognit.*, pp. 2805-2814, 2020.
- [44] Y. Li, Q. Miao, W. Ouyang, Z. Ma, H. Fang, C. Dong, Y. Quan, "LAP-Net: Level-aware progressive network for image dehazing", in *Proc. Int. Conf. Computer Vision*, pp. 3276-3285, 2019.
- [45] Y. Li, Y. Liu, Q. Yan, K. Zhang, "Deep dehazing network with latent ensembling", in *Proc. Conf. Computer Vision Pattern Recognit. Workshops*, 2020.
- [46] C.-H. Yeh, C.-H. Huang, L.-W. Kang, "Multi-scale deep residual learning-based single image haze removal via image decomposition", *IEEE Trans. Image Process.*, vol. 29, pp. 3153-3167, 2019.
- [47] S. Bianco, L. Celona, F. Piccoli, R. Schettini, "High-resolution single image dehazing using encoder-decoder architecture", in *Proc. Conf. Computer Vision Pattern Recognit. Workshops*, pp. 1927-1935, 2019.
- [48] P. Khosla et al., "Supervised contrastive learning", in *Neural Inf. Process. Syst.*, pp. 18661-18673, 2020.
- [49] G. Kim, S. W. Park, J. Kwon, "Pixel-wise Wasserstein autoencoder for highly generative dehazing", *IEEE Trans. Image Process.*, vol. 30, pp. 5452-5462, 2021.
- [50] H. Li, J. Li, D. Zhao, L. Xu, "DehazeFlow: Multi-scale conditional flow network for single image dehazing", in *Proc. ACM Int. Conf. Multimedia*, pp. 257-261, 2021.
- [51] P. Li, J. Tian, Y. Tang, G. Wang, C. Wu, "Deep retinex network for single image dehazing", *IEEE Trans. Image Process.*, vol. 30, no. 2, 2021.
- [52] H.-H. Yang and Y. Fu, "Wavelet U-Net and the chromatic adaptation transform for single image dehazing", in *Proc. Int. Conf. Image Process.*, pp. 2736-2740, 2019.

- [53] P. Hambarde and S. Murala, "S2DNET: Depth estimation from single image and sparse samples", *IEEE Trans. Comput. Imaging*, vol. 6, pp. 806-817, 2020.
- [54] L. Cheng and L. Zhao, "Two-stage image dehazing with depth information and cross-scale non-local attention", in *Proc. Int. Conf. Big Data*, pp. 3155-3162, 2021.
- [55] X. Zhang, J. Wang, T. Wang, R. Jiang, "Hierarchical feature fusion with mixed convolution attention for single image dehazing", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 2, pp. 510-522, 2022.
- [56] S. D. Das and S. Dutta, "Fast deep multi-patch hierarchical network for nonhomogeneous image dehazing", in *Proc. Conf. Computer Vision Pattern Recognit. Workshops*, pp. 482-483, 2020.
- [57] Z. Zheng, W. Ren, X. Cao, X. Hu, T. Wang, F. Song, X. Jia, "Ultra-high-definition image dehazing via multi-guided bilateral learning", in *Proc. Conf. Computer Vision Pattern Recognit.*, pp. 16180-16189, 2021.