

# بهینه‌سازی ساختاری شبکه عصبی عمیق سلسله مراتبی خودتوجه و رمزگذاری پویای تک متغیره برای بزرگ‌نمایی تصاویر دیجیتال

علی دلشادی<sup>۱</sup>، وحید مهرداد<sup>۲\*</sup>، محمدباقر دولتشاهی<sup>۳</sup>

## چکیده

بزرگ‌نمایی تصاویر دیجیتالی یکی از روش‌های پردازش تصویر می‌باشد، که وضوح تصویر را در زمینه دید رایانه‌ای بهبود می‌بخشد. در اصل این کار برای بزرگ‌نمایی تصاویر ثابت و متحرک که از زمان تصویربرداری آن‌ها گذشته و دسترسی به دوربین یا صحنه‌ها برای زوم وجود ندارد، استفاده می‌شود. در این مقاله از شبکه سلسله مراتبی جهت استخراج ویژگی‌های سطح بالا برای حل چالش مرزبندی بین رنگ‌ها و از بلوک‌های توجه خودآموز برای کاهش عملیات کانولوشن پیشنهاد می‌شود. در ادامه برای بهینه‌سازی شبکه از روش جستجو تصادفی و تقسیم دودویی برای یافتن پاسخ بهینه پارامترها و فرآیندها استفاده می‌گردد. با استفاده از روش جستجوی ذکر شده علاوه بر جستجوی وزن‌ها و پارامترهای شبکه، می‌توان ساختار معماری را نیز جستجو کرد، این عمل باعث تنظیم خودکار فرآیندها و بهینه‌سازی ساختار شبکه خواهد شد. برای بررسی کارایی روش پیشنهادی، نتایج شبیه‌سازی بر روی پایگاه داده تصاویر در این حوزه تست شده که این نتایج برتری روش پیشنهادی نسبت به روش‌های دیگر را نمایش می‌دهد. با توجه به نتایج بدست آمده در بهینه‌سازی معماری با استفاده از روش ذکر شده در بزرگ‌نمایی چهار برابر با بلوک سلسله مراتبی چهار طبقه و استفاده از بلوک توجه در بخش بزرگ‌نمایی به عدد سیگنال به نویز ۳۲/۶۶ دست یافت.

## کلید واژه‌ها

بزرگ‌نمایی تصاویر دیجیتال، شبکه عصبی عمیق سلسله مراتبی، بهینه‌سازی ساختاری، بهینه‌سازی پویای تک متغیره، بلوک توجه

## ۱ - مقدمه

در دو دهه گذشته، مفهوم بزرگ‌نمایی تصاویر که با عناوین مختلفی در متون مختلف ذکر شده است. در اینجا از عبارت بزرگ‌نمایی تصاویر دیجیتال<sup>۱</sup> SR به عنوان روشی برای تولید تصاویر با وضوح بالا از یک تصویر با وضوح پایین<sup>۲</sup> LR تعریف شده است. کاربردهای بزرگ‌نمایی تصاویر دیجیتال شامل طیف وسیعی از کاربردهای دنیای واقعی مانند تصویربرداری پزشکی

این مقاله در دی‌ماه ۱۴۰۱ دریافت، در اسفندماه بازنگری و در فروردین ۱۴۰۲ پذیرفته شد.

۱ دانشجوی دکترای برق، دانشگاه لرستان

رایانامه: [delshadi.al@fe.lu.ac.ir](mailto:delshadi.al@fe.lu.ac.ir)

۲ استادیار گروه برق دانشگاه لرستان

رایانامه: [mehرداد.v@lu.ac.ir](mailto:mehرداد.v@lu.ac.ir)

۳ دانشیار گروه کامپیوتر دانشگاه لرستان

رایانامه: [dowlatshahi.mb@lu.ac.ir](mailto:dowlatshahi.mb@lu.ac.ir)

superresolution<sup>۱</sup> معنی لغوی آن وضوح فوق العاده است که در مقاله به عنوان بزرگ‌نمایی تصاویر دیجیتال یاد شده است.

<sup>۲</sup>low-resolution

نویسنده مسئول: وحید مهرداد

[doi.org/10.23831/197.1402.10.3.3.5](https://doi.org/10.23831/197.1402.10.3.3.5)

افزایش عمق شبکه برای بهبود عملکرد بازسازی تصویر نهایی دارند، که این موضوع باعث بزرگ‌شکبه و در پی آن کند شدن شبکه در یادگیری و تست می‌شود. حتی در مواقعی شبکه خیلی عمیق می‌تواند بیش برآش شود. در ضمن آنها به استخراج ویژگی-های سطح بالا و نادیده‌گرفتن ویژگی‌های سلسله‌مراسته‌پیداخته‌اند. برای حل مشکل استخراج ویژگی‌ها در همه سطوح مورد نیاز از روش استخراج ویژگی سلسله مراتبی استفاده شده است [۳۹]، که این روش از مشکل عمق و عدم بهینه ساختاری رنج می‌برد. در نهایت، با در نظر گرفتن این بررسی‌ها در بزرگ‌نمایی تصاویر دیجیتال، برای حل مشکل استخراج مرز بین رنگ‌ها از شبکه سلسله مراتبی با تعداد چهار لایه استفاده شده است.

## ۱-۲- بلوک توجه

در شبکه‌های عصبی عمیق از نظر محاسباتی، میدان‌های دریافتی محلی و اشتراک وزن به ترتیب در کاهش تعداد عملیات ممیز شناور FLOP و تعداد پارامترها در مقایسه با لایه‌های کاملاً متصل بسیار مفید واقع شده‌اند.

با اینحال، استفاده از وزنه‌های کانولوشنی یکسان در همه موقعیت‌های پیکسل ممکن است لزوماً انتخاب بهینه نباشد. با استفاده از موفقیت روش‌های بلوک توجه [۲۶] و [۲۷] در وظایف پردازش زبان طبیعی، چندین روش اخیر عمل کانولوشن را با مکانیزم خودتوجهی در [۲۸] و [۲۹] جایگزین کردند. درحالی‌که اندازه میدان جستجو لزوماً تغییر نمی‌کند، استفاده از توجه به خود اجازه می‌دهد تا وزنه‌های تجمع پیکسلی ضرایب توجه وابسته به محتوا باشد. این کار، در واقع، منجر به یک هسته تطبیقی می‌شود که وزن آن در مکان‌های مکانی بر اساس محتویات میدان جستجو متفاوت است.

در این مقاله، بر یکدیگر از عملیات اساسی در کاربردهای بینایی تمرکز شده که توجه نسبتاً کمی را به خود جلب کرده است، که کانولوشن انتقالی یا عکس کانولوشن گام‌به‌گام می‌باشد. بر خلاف عمل کانولوشن گام دار استاندارد که بعد فضایی ورودی را با ضریب گام کاهش می‌دهد، کانولوشن انتقالی گام‌به‌گام ابعاد فضایی ورودی را با ضریب گام افزایش می‌دهد. بنابراین، کانولوشن انتقالی گام دار یک عملیات طبیعی برای استفاده در مورد نمونه برداری تصویر است. در کل در این مقاله، هسته ثابت مورد استفاده در کانولوشن انتقالی گام‌دار را با مکانیزم خودتوجهی جایگزین کرده، که وزن هسته را در هر موقعیت پیکسل بر اساس محتویات میدان گیرنده محاسبه می‌کند. این کار باعث می‌شود که تعداد عملیات در پردازش شبکه کاهش یابد و یافتن ساختار معماری و بهینه‌سازی وزن‌ها سریعتر انجام شود.

[۱] و [۲]، نظارت و امنیت [۳] و [۴] و غیره می‌باشد. به غیر از بهبود کیفیت ادراکی تصویر، به بهبود سایر وظایف بینایی رایانه نیز کمک می‌کند [۵] و [۶]. همچنین این روش‌ها برای کاربردهای دیگری در پردازش تصویر امکان استفاده را دارد [۷].

به طور کلی، ساخت تصویر با کیفیت بالا از تصویر با کیفیت پایین مشکل بسیار چالش برانگیزی می‌باشد. انواع روش‌های کلاسیک SR تا به امروز پیشنهاد شده‌اند، از جمله روش‌های مبتنی بر پیش‌بینی [۸] و [۹]، روش‌های مبتنی بر لبه [۱۰] و [۱۱]، روش‌های آماری [۱۲] و [۱۳]، روش‌های مبتنی بر پیچ [۱۴] و [۱۵] و روش‌های نمایش پراکنده [۱۶] و [۱۷] را می‌توان نام برد. روش‌ها بسیار متنوع و زیاد می‌باشند، در ادامه به بررسی سه بخش اصلی جهت دستیابی به شبکه پیشنهادی در این مقاله پرداخته شده است.

## ۱-۱- استخراج ویژگی

اگر چه بزرگ‌نمایی تصاویر دیجیتالی برای چندین دهه توسط محققین حوزه‌های مختلف بررسی شده است، ولی همچنان یک کار چالش برانگیز در بینایی ماشین و دید رایانه‌ای می‌باشد. این مشکلات می‌تواند بصورت ذاتی از ابتدای تصویربرداری ایجاد شده باشد، زیرا هنگام تصویربرداری، تغییرات جزئی در زاویه دوربین، عمق رنگ، میزان روشنایی، پیچیدگی تصویر و سایر متغیرها برای هر تصویر کم کیفیت LR می‌تواند وجود داشته باشد، که باعث سردرگمی و خطای سامانه می‌شود.

در دسته بندی اصلی برای بزرگ‌نمایی تصاویر دیجیتال، می‌توان آن را به دو دسته، بزرگ‌نمایی تک تصویری<sup>۱</sup> SI و بزرگ‌نمایی چند تصویری<sup>۲</sup> طبقه‌بندی کرد [۱۸] و [۱۹]. در بزرگ‌نمایی تصاویر دیجیتال تک تصویر، یادگیری برای یک جفت تصویر با کیفیت بالا و پایین LR-HR برای یک تصویر انجام می‌شود، در حالی که در بزرگ‌نمایی چند تصویری، یادگیری برای تعداد زیادی جفت تصویر با کیفیت بالا و پایین، برای یک صحنه خاص انجام می‌شود، در نتیجه امکان تولید یک تصویر با کیفیت بالا از یک صحنه یا از چندین صحنه متوالی می‌باشد [۲۰] و [۲۱]. برای کار پیشنهادی از گروه بزرگ‌نمایی تک تصویر استفاده شده است.

می‌توان روش‌های بزرگ‌نمایی تک تصویر دیجیتال را به چهار دسته طبقه‌بندی کرد: که روش‌های کلاسیک، روش‌های مبتنی بر یادگیری نظارت شده، روش‌های مبتنی بر یادگیری بدون نظارت، و روش‌های حوزه خاص می‌باشند.

روش‌های زیادی مبتنی بر CNN برای دست‌یابی به بزرگ‌نمایی تصویر دیجیتال پیشنهاد شده‌اند، از جمله شبکه کانولوشن بسیار عمیق [۲۲] VDSR، شبکه باقیمانده [۲۳] SRResNet، شبکه بازگشتی عمیق [۲۴] DRRN و شبکه متراکم باقیمانده [۲۵] RDN را می‌توان نام برد. با اینحال، این شبکه‌ها به طور کلی تمایل به

<sup>۱</sup>single image

<sup>۲</sup>multiple images

## ۱-۳- بهینه سازی

محل استفاده شده است. در اینجا از روش بلوک سلسله مراتبی جهت استخراج ویژگی‌های سطوح مختلف و بزرگ‌نمایی تصویر نهایی از طریق بلوک توجه پرداخته شده است. در ضمن برای بهینه سازی ساختار شبکه و وزن‌ها از بهینه سازی ساختاری با روش رمزگذاری پویای تک متغیره استفاده شده است که باعث می‌شود شبکه نهایی کم حجم تر و چابکتر با عملکردی بهتر دست یابد.

در مقاله از سه بخش اصلی برای انجام کار استفاده شده که می‌توان به بخش استخراج ویژگی سلسله مراتبی، بخش بزرگ‌نمایی با کمک بلوک توجه و بخش بهینه سازی ساختاری با کمک روش جستجوی تصادفی و باینری تقسیم کرد. در بخش ۲ مروری بر کارهای انجام شده، در بخش ۳ شرح کامل طرح پیشنهادی در هر سه بخش کاری طبق جدول (۳)، در بخش ۴ تست و نتیجه گیری و در بخش ۵ چالش‌ها و آینده تحقیقاتی پرداخته شده است.

## ۲- کارهای مرتبط

کارهای مرتبط در سه بخش جداگانه بررسی شده است.

## ۱-۲- استخراج ویژگی

شبکه‌های عصبی کانولوش قابلیت استخراج ویژگی‌های قوی و عملکرد خوبی را در زمینه SR نشان داده‌اند. بنابراین، روش‌های زیادی مبتنی بر CNN برای دستیابی به SR تصویر پیشنهاد شده‌اند. با اینحال، این شبکه‌ها به طور کلی تمایل به افزایش عمق شبکه برای بهبود عملکرد خود در SR دارند. آنها به استخراج ویژگی‌های سطح بالا و ویژگی‌های سلسله مراتبی اختصاص داده شده‌اند. با این وجود، استخراج ویژگی‌های سلسله مراتبی هنگام دستیابی به SR در مقیاس‌های مختلف با یک مدل بسیار مهم است زیرا مقیاس مختلف در تصویر نهایی به ویژگی‌های سطوح متفاوتی برای SR نیاز دارند.

برخی از روش‌های پرکاربرد بزرگ‌نمایی تصاویر دیجیتال بررسی شده که از شبکه‌های عصبی کانولوشنی عمیق بهره می‌برند، که در ادامه آورده شده‌اند: شبکه بزرگ‌نمایی عمیق پیشرفته [۳۵] EDSR، شبکه بزرگ‌نمایی کانولوشن [۳۶] SRCNN، شبکه باقی‌مانده چند مقیاسی [۳۷] MSRN، شبکه متراکم باقی‌مانده [۳۸] RDN، شبکه عمیق برگشتی [۳۹] DRCN، شبکه تکراری غیر محلی [۴۰] NLRN، شبکه سلسله مراتبی چند سطحی [۴۱] RSI-HFAS، ماسک پراکنده [۴۲] SMSR، شبکه تجمیع ویژگی‌های باقیمانده [۴۳] RFANet، شبکه توجه کل نگر [۴۴] HAN، وضوح تصویر فوق العاده با توجه پراکنده غیر محلی [۴۵] CNN و NLSN با وضوح فوق العاده کارآمد و سبک همراه با ترانسفورماتور [۴۶] ESRT را می‌توان نام برد.

## ۲-۲- بلوک توجه

روش‌های بلوک توجه که در ابتدا در [۴۷] معرفی شدند، عملکرد چشمگیری را در وظایف مدل‌سازی دنباله‌ای از خود نشان دادند، جایکه به طور قابل توجهی از روش‌های سنتی مبتنی بر شبکه‌های

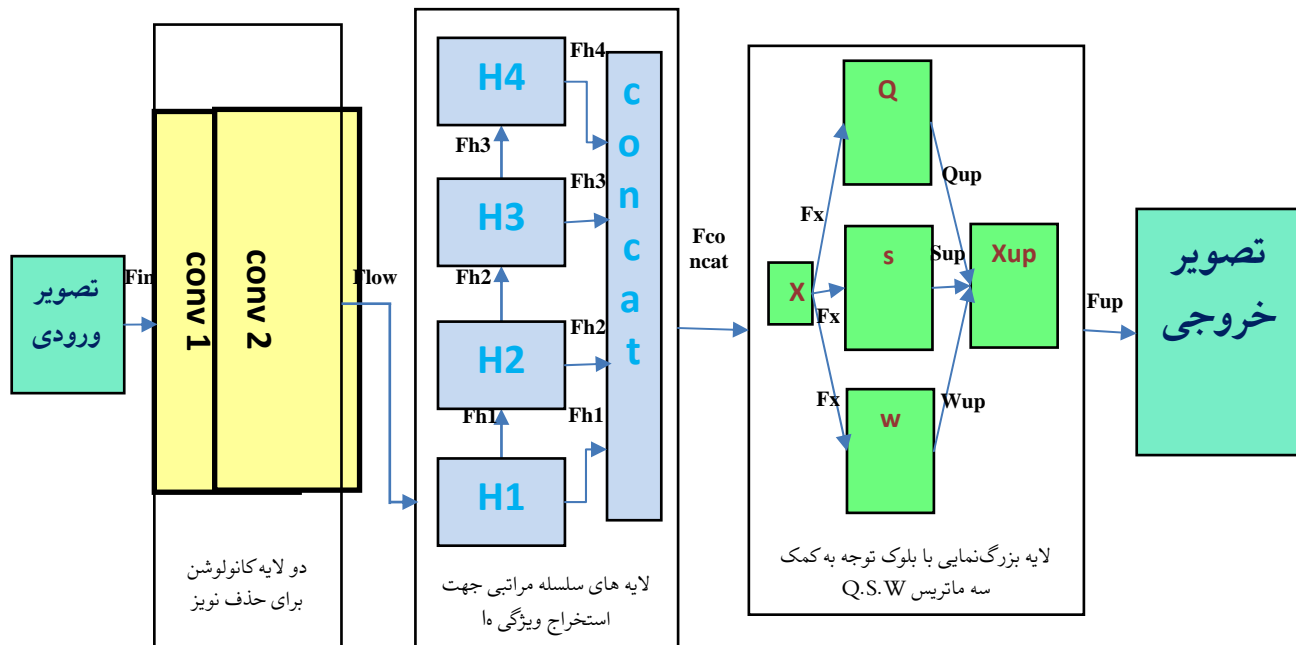
هنگامی که جستجوی شبکه به بهینه‌سازی فرا پارامترها اعمال می‌شود، تعداد نقاط شبکه قابل آموزش به طور قابل توجهی برای شبکه‌های عصبی افزایش می‌یابد. در این حالت مقایسه تعداد نقاط شبکه در همه موارد غیر ممکن است، بنابراین بهینه‌سازی فراپارامترها می‌تواند براساس تجربه یک متخصص انجام شود. در این حالت بدست آوردن پاسخ با روش محاسبه متعارف دشوار و زمانبر خواهد بود، زیرا بهینه‌سازی به یک مساله دشوار چند هدفه تبدیل می‌شود. در روش‌های اولیه برای بهینه‌سازی فراپارامترها برای یادگیری ماشین برای حل این مشکل از جستجوی گرادینان [۳۰]، جستجوی گرادینان تصادفی [۳۱] و الگوریتم‌های مسابقه [۳۲]، برای بهینه‌سازی یادگیری عمیق استفاده شدند. این روش‌ها موقعی کارساز است که تابع زیان مشتق پذیر باشد. حال مسایلی که به یافتن پاسخ مناسب برای ساختار شبکه می‌پردازد، بطور کل مشتق پذیر نمی‌باشد. همچنین شبکه‌های چند هدفه بطور همزمان می‌خواهد در چند هدف بهینه باشد، نیز مشتق پذیر نمی‌باشد و تنها راه حل استفاده از روشهایی غیر از روشهای کلاسیک است. به تازگی روش‌های بهینه‌سازی فراپارامترها معرفی شده‌اند که کارآمد بوده و با استفاده از جستجوهای سراسری و محلی به حل آنها پرداخته‌اند. روش‌های مختلف جستجوی سراسری و محلی معرفی شده‌اند مانند روش‌های تکاملی و خرد جمعی که خود آنها نیز به روشهای مختلف دیگری دسته بندی می‌شوند.

در کل بهینه‌سازی فراپارامترها برای مشکلات مهندسی می‌تواند به مشکلات بهینه‌سازی هزینه‌ها در یادگیری کمک کند. روش‌های بهینه‌سازی هزینه را می‌توان به دو رویکرد طبقه‌بندی کرد: روش‌های غیرمستقیم و مستقیم [۳۳]. در بین آنها، رویکرد با استفاده از الگوریتم رمزگذاری پویای تک متغیره برای جستجوها در چندین مساله بهینه‌سازی متعدد دو برای تنظیم فراپارامترها در مسایل مهندسی استفاده شده‌است. این را باید بدانیم که یک شبکه عصبی عمیق برای بزرگ‌نمایی تصاویر دیجیتال غیرخطی بوده و تابع هزینه پیچیده و زمانبری خواهد داشت. بنابراین بهینه‌سازی روش کلاسیک با استفاده از مشتق تابع هزینه برای چنین شبکه‌ای دشوار است. برای مقابله با این مسایل، بهینه‌سازی عاری از مشتق می‌تواند یک راه حل مناسب باشد، که از اطلاعات مشتق در مفهوم برای یافتن راه حل‌های بهینه استفاده نکند. در این مقاله برای بهینه‌سازی شبکه عصبی عمیق مورد نظر از الگوریتم جستجوی سراسری تصادفی و جستجوی محلی باینری استفاده شده است [۳۴]. این روش یک تکنیک بهینه‌سازی غیرخطی می‌باشد، که بهینه‌سازی شبکه را بدون استفاده از اطلاعات تفاضلی تابع هزینه می‌تواند جستجو را انجام می‌دهد.

در این مقاله برای حل مشکل یادگیری شبکه‌های عمیق و تنظیم فراپارامترها از روش جستجوی تصادفی در فضای جستجوی سراسری و روش جستجوی دودویی در بخش فضای جستجوی

نام برد. همچنین این روشها برای یافتن سریع نقطه بهینه سراسری استفاده شده‌اند، که می‌توان روشهایی مانند، روش تصادفی آلفی-پنتینی [۶۴]، الگوریتم تکامل یونگ [۶۵]، روش پیوند تک سطحی کان و تایمر [۶۶]، تکنیک حسابی بازه ای راتچک و راکنه [۶۷]، روش تونل زنی لوی و مونتالوو [۶۸] و طرح جستجوی تابوی سیویویچ و کلینوفسکی [۶۹] را نام برد.

عصبی بازگشتی [۴۸] بهتر عمل کردند. روش‌های توجه در نحوه محاسبه وزن‌های تجمع یا ضرایب توجه متفاوت هستند [۴۹] و [۵۰]. در این مقاله، عمدتاً از روش توجه محصول نقطه‌ای مقیاس‌شده استفاده شده، که توسعه ایازر و شمول محصول نقطه‌ای مقیاس ناپذیر [۵۱] است. استفاده از آنها در توالی وظایف، روش‌های توجه برای جایگزینی



شکل (۱): معماری شبکه عصبی عمیق در استفاده برای بزرگ‌نمایی تصویر

نسل‌های بعدی الگوریتم‌های بهینه‌سازی که منطبق با متا اکتشافی هستند، الگوریتم‌های دقیقی می‌باشند [۷۰]. الگوریتم‌های دقیق برای بهینه‌سازی چندین موضوع مهندسی مورد استفاده قرار گرفته‌اند [۷۱]. در [۷۲]، محققان بهینه‌سازی زنجیره تامین چند خریدار چند محصولی یکپارچه را تحت سیاست‌های جریمه و مدیریت شده با استفاده از تقریب بیرونی مورد مطالعه قرار دادند. به طور کلی، روش ذکر شده را می‌توان به جستجوی سراسری و جستجوی محلی تقسیم کرد. برای بهینه‌سازی سراسری، در این مقاله از روش انتخاب تصادفی که ساده‌ترین و سریعترین روش می‌باشد، استفاده شده است. بهینه‌سازی محلی توسط یک جستجوی دو بخشی و یک جستجوی یک طرفه انجام می‌شود که بر اساس اطلاعات به دست آمده از منطقه همسایه، بهترین نقطه محلی را می‌یابد [۷۳] و [۷۴].

### ۳- روش پیشنهادی

باید این موضوع را مدنظر قرار داد، که در مقایسه با تصاویر صحنه طبیعی، تصاویر خروجی از دوربین‌ها دارای پس‌زمینه پیچیده، وضوح پایین اجسام و بافت‌های ضعیف بوده، که شامل بسیاری از اجزای فرکانس پایین، متوسط و بالا است. بنابراین، استخراج کورکورانه ویژگی‌های سطح بالا برای SR در مقیاس بالا مناسب نخواهد بود و باعث تار شدن تصویر نهایی می‌شود. در این مقاله

یا تقویت شبکه‌های عصبی کانولوشن مورد استفاده قرار گرفته‌اند [۵۲] و [۵۳]. کانولوشن مبتنی بر توجه به جای استفاده از یک هسته ثابت برای جمع‌آوری ویژگی‌ها از همه پیکسل‌ها در یک میدان پذیرنده، از مکانیزم توجه برای محاسبه پویا، این وزن‌های تجمع بر اساس محتویات میدان پذیرنده استفاده می‌کند [۵۴].

### ۲-۳- بهینه‌سازی ساختاری

چندین محقق در زمینه روش بهینه‌سازی فرآپارامترها تحقیقاتی انجام داده‌اند. برای بهینه‌سازی فرآپارامترها، یکی از مشکلات اساسی در یک مدل شبکه عصبی روش جستجوی سراسری است [۵۵] و [۵۶] که در صورت انتخاب درست روش جستجوی سراسری مناسب شبکه سریعتر به جستجوی محلی رسیده و می‌تواند در زمان کوتاهتری پاسخ نزدیک یا درست را بیابد [۵۷] و [۵۸]، محققین از روشهایی مانند جستجوی تصادفی، جستجو بیزی و جستجو شبکه ای برای یافتن راه‌حل‌های بهینه استفاده کرده‌اند [۵۹] و [۶۰].

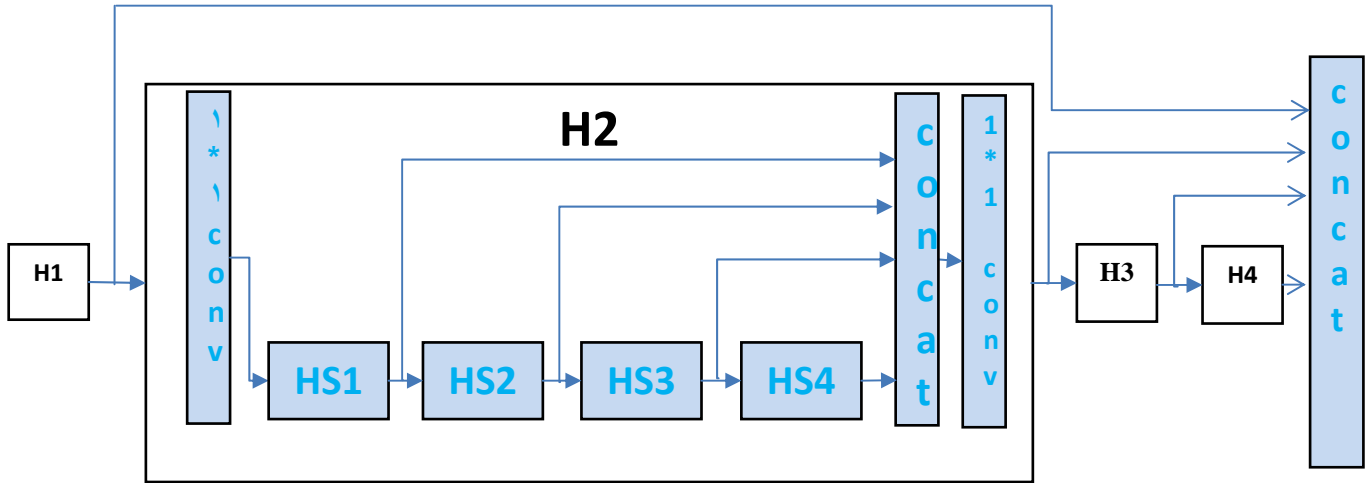
در مورد استفاده از روش‌های بهینه‌سازی فرآپارامترها می‌توان مثال‌های کاربردی که در بخش‌های مختلف صنعت هم مورد استقبال قرار گرفته‌اند را نام برد. بهینه‌سازی پارامترهای کنترلی مختلف، مواردی مانند پارامتر فازی کنترل‌کننده ربات [۶۱]، تخمین پارامتر موتور [۶۲] و تخمین دمای بیلت [۶۳] را می‌توان

ویژگی‌های سطوح مختلف و بخش سوم بلوک توجه برای بزرگ-نمایی برای ساخت تصویر نهایی می باشد.

### ۳-۱- استخراج ویژگی با روش سلسله مراتبی

در ابتدا با توجه به اهمیت ویژگی‌های سلسله مراتبی، یک بلوک

یک شبکه میانی برای دستیابی به همه ویژگی‌های مورد نیاز پیشنهاد شده است. ابتدا، با توجه به اهمیت ویژگی‌های سلسله مراتبی، یک بلوک سلسله مراتبی طراحی کرده که میتواند بسیاری از اطلاعات ویژگی‌های سلسله مراتبی، از جمله اطلاعات ویژگی‌های سطح پایین، متوسط و بالا را استخراج کند. با الهام از [۷۵] و [۷۶]



شکل (۲): بلوک سلسله مراتبی دارای چهار سطح H1 الی H4 است. بخش داخلی هر کدام از این سطوح شبیه بهم می باشند. در این شکل بخش داخلی سطح H2 دیده می شود.

جمع‌آوری ویژگی سلسله مراتبی طراحی شده، که می‌تواند اطلاعات ویژگی سلسله مراتبی زیادی از جمله اطلاعات ویژگی سطح پایین تا سطح بالا را در ارتباط با فرکانس‌های پایین بصورت پیوسته تا فرکانس بالا استخراج کند. دوم اینکه، مکانیزم بازخورد به طور گسترده در بسیاری از زمینه‌های بینایی کامپیوتر استفاده شده است. یک ساختار تکراری برای تقلید از مکانیزم بازخورد موجود در دید انسان اعمال کرده، که می‌تواند گام به گام اطلاعات ویژگی سلسله مراتبی را اصلاح کند. در نهایت با توجه به عوامل مقیاس مختلف، شبکه می‌تواند با استفاده از فیلترهای مناسب برای تکمیل عملیات به تنهایی انجام وظیفه کند و سپس برای دستیابی به بازسازی تصویر نهایی، انتخاب شود.

یک ماژول مخصوص استخراج ویژگی در شکل شماره (۱) دیده می‌شود، این ماژول متشکل از چهار سطح استخراج ویژگی سلسله مراتبی می‌باشد، که در یک روش سری به هم متصل شده‌اند تا در نهایت با یک لایه ترکیب سریالی خروجی هر سطح، ترکیب شوند. هر سطح می‌تواند ویژگی‌های سلسله مراتبی مختص خود را استخراج کند، که می‌تواند بافت غنی از تصویر را ایجاد کند و برای دستیابی به ویژگی‌های استراتژیک مقیاس مختلف با یک مدل منفرد، مناسب می‌باشد.

یک ساختار سطح بالا با یادگیری جدید در این بلوک‌ها طبق شکل شماره (۲) پیشنهاد می‌شود، بطوریکه یک شبکه غیرخطی بسیار قدرتمند برای برآزش یک تابع نقشه برداری شده از مرزها را شامل می‌شود. با توجه به عوامل مهم مقیاس‌های متفاوت، یادگیرنده

یک ساختار تکراری را برای تقلید از روش بازخوردی که در بینایی انسان وجود دارد، اعمال می‌شود، که می‌تواند اطلاعات ویژگی سلسله مراتبی را گام به گام اصلاح کند. در نهایت، با توجه به عوامل مقیاس مختلف، پیش‌بینی می‌شود شبکه معرفی شده بتواند از فیلترهای نمونه برداری مناسب برای تکمیل عملیات نمونه برداری به تنهایی استفاده کند و سپس در بلوک توجه به بازسازی تصویر نهایی یدست یابد.

این مقاله روشی را برای پیدا کردن تنظیم فرآپارامترها<sup>۱</sup> برای یک مدل شبکه عصبی با استفاده از uDEAS<sup>۲</sup> استفاده می‌کند. معادل فارسی uDEAS الگوریتم کدگذاری پویای تک متغیره برای جستجوها می‌باشد. روش پیشنهادی با استفاده از مدل شبکه عصبی کانولوشن سلسله مراتبی با ساختار شناسایی ویژگی‌های لبه تصویر و بزرگ‌نمایی توسط بلوک توجه می‌باشد. برای بهینه‌سازی فرآپارامترها، برای به حداقل رساندن هزینه توسط uDEAS، به عنوان میانگین اختلاف بین مقدار کشف شده با تصویر اصلی و معکوس دقت ارزیابی برای شبکه سلسله مراتبی انتخاب شده است.

همانطور که از شکل شماره (۱) پیداست، ساختار اصلی کار از چهار بخش اصلی تشکیل شده است، که بخش اول پیش پردازش برای حذف نویز فرکانس پایین و کوچک کردن فضای جستجو، بخش دوم مراحل بلوک‌های سلسله مراتبی برای آشکار کردن

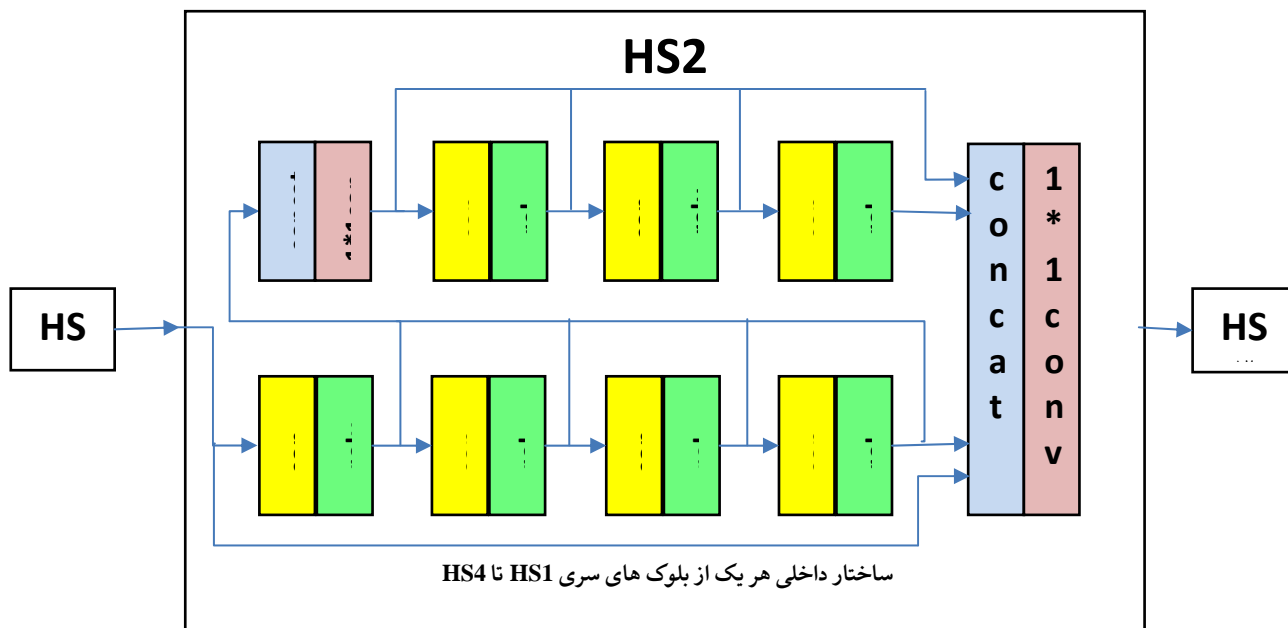
<sup>۱</sup>hyperparameter

<sup>۲</sup>univariate dynamic encoding algorithm for searches



مراتبی حاوی اطلاعات جامع ویژگی‌ها هستند، که برای بازسازی بافت مفید می‌باشند. همانطور که در شکل شماره (۲) نشان داده شده، با لایه های کانولوشن پیوسته و باقیمانده، مشخصه‌ها و ویژگی‌ها به سطوح بالاتر و بالاتری تبدیل خواهند شد. بنابراین، طبقات مختلف می‌توانند ویژگی‌ها را در سطوح متفاوتی ایجاد کنند. همانطور که قبلاً گفته شد، یک بلوک H دارای چهار سطح ویژگی مختلف است، که می‌تواند ترکیب ویژگی را با یک عملیات کانولوشن  $1 \times 1$  برای بدست آوردن ویژگی‌های سلسله مراتبی اجرا کند. از طرفی ساختار سلسله مراتبی بلوک های H1 الی H4 در داخل خود از چهار بلوک HS1 الی HS4 تشکیل شده است،

می‌تواند به طور پویا تعداد و وزن‌های فیلتر را برای رسیدن به هدف پیش‌بینی کند. چارچوب شبکه همانطور که در شکل شماره (۲) نشان داده شده است، شامل یک بلوک چهار سطحی H1 الی H4 است، که همه بخش‌ها به نوعی می‌خواهند ویژگی‌های سطح پایین تا بالا را استخراج کنند و در نهایت با هم ترکیب شوند. ماهیت اصلی بدین صورت است که اطلاعات کمی در فرکانس‌های پایین تصویر وجود دارد، پس کار پردازشی روی آن باید کمتر انجام شود تا حجم شبکه کمتر و سرعت کار بیشتر شود. از طرفی اطلاعات بیشتری در فرکانس بالاتر تصویر وجود دارد و بایستی کار بیشتر روی آن انجام



شکل (۳): لایه های داخلی بلوک‌های HS

که حاوی اطلاعات و ویژگی سلسله مراتبی مفیدی می‌باشد. این کار باعث می‌شود بافت SR به ظاهر پرت به نظر برسد. طبق شکل شماره (۲) می‌توان گفت که، در ابتدا از بخش ورودی، تصویر وارد بلوک شده که آنها را باهم از طریق یک لایه کانولوشن  $1 \times 1$  بصورت فشرده قرار می‌دهد. پس از عبور از چهار بلوک HS ویژگی‌های استخراج شده با هم ترکیب می‌شوند، که در نهایت ویژگی‌های سلسله مراتبی خروجی به دست می‌آید. در پایان از یک فیلتر  $1 \times 1$  عبور داده می‌شود. از هر بلوک یک خروجی به بلوک بعدی رفته و یک خروجی برای ترکیب شدن<sup>۱</sup> به خروجی نهایی می‌رود. سپس با یک لایه کانولوشن  $1 \times 1$  ابعاد را کاهش داده تا خروجی با ویژگی سطوح مختلف را بسازد. بطور خلاصه در کل بلوک سلسله مراتبی از ۴ ماژول سلسله مراتبی H استفاده شده و هر کدام از بلوک های H داخل خود از ۴ ماژول سریالی HS تشکیل شده، که در کل در این ساختار ۱۶ عدد از ماژول HS وجود دارد.

به همین خاطر این فرکانس‌ها وارد سطوح بالایی استخراج ویژگی برای پردازش بیشتر می‌شوند. برای ازدست ندادن ماهیت کلی ویژگی‌ها، ورودی لایه بعدی از خروجی لایه قبلی بوده و در نهایت کل اطلاعات و ویژگی‌های استخراج شده با هم ترکیب می‌شوند. کل ماهیت بلوک سلسله مراتبی بدین صورت می‌باشد.

طبق شکل شماره (۲) داخل هر کدام از این سطوح از چهار بلوک HS استفاده شده است که محتوای داخلی هر یک از آنها در شکل شماره (۳) بررسی شده است. در داخل هر سطح در ابتدای لایه کانولوشن  $1 \times 1$  وجود دارد و پس از آن تصویر وارد بلوک‌های HS می‌شود و بصورت سریالی و باقیمانده در خروجی وارد لایه سریالی سازی خروجی هر یک از بلوک‌های HS شده و در نهایت وارد یک لایه کانولوشن  $1 \times 1$  می‌شود و وارد سطح بعدی H می‌شود.

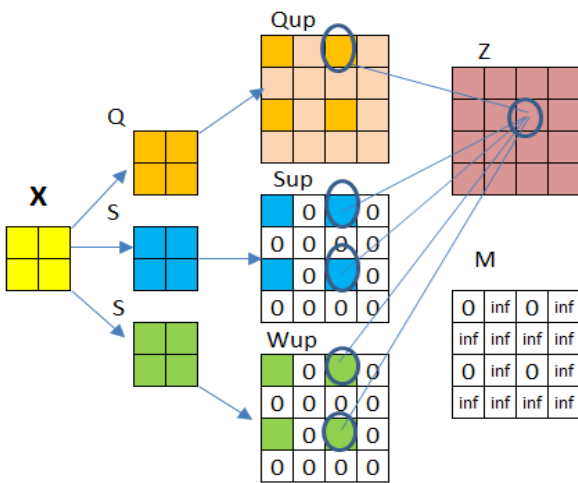
وزن هر بار یادگیری در هر تکرار به اشتراک گذاشته می‌شود. ویژگی ایجاد شده توسط بلوک سری در هر تکرار به عنوان ورودی برای تولید تصویر باقی‌مانده استفاده می‌شود و با تصویر قبلی برای تولید تصویر بعدی بصورت سری ترکیب می‌شود.

با توجه به تغییرپذیری توپوگرافی سطحی، ویژگی‌های بافتی غیرقابل پیش‌بینی هستند. از طرفی ذکر شد که ویژگی‌های سلسله

<sup>۱</sup>concat

توجه می‌دهد، زیرا لزوماً مجبور نیستیم این بلوکها را از همان مجموعه نقشه‌های ویژگی اولیه به دست آوریم. در واقع، به دست آوردن بخش‌های پرس و جو و بلوکهای کلیدی از این روش به طور طبیعی اجازه می‌دهد که بلوک مبتنی بر توجه همه آنها را در یک مرحله انجام دهد. عملیات کانولوشن انتقالی گام به گام معمولی این قابلیت را ندارد. بنابراین مدل نمونه برداری مشترک مبتنی بر توجه دقیق‌تر و کارآمدتر است، که در این مقاله استفاده شده است. مکانیزم توجه یکه در این مقاله استفاده شده است، مکانیسم توجه محصول نقطه‌ای مقیاس شده است [۸۵]. معماری بزرگمایی از معماری خود توجه برپایه مدل بزرگنمایی عمیق مفصلی شکل گرفته است که برای روش رمز گزاری از روش رمزگذاری موقعیتی نسبی الهام گرفته شده است تا به ماتریس تصویر برسیم [۸۶].

در این مقاله، از نوعی عملیات کانولوشن که کانولوشن انتقالی گام‌دار نام دارد استفاده شده است. به این عمل انحراف کسری یا معکوس کانولوشن [۸۷] نیز می‌گویند. پارامترهای معماری یک کانولوشن انتقالی گام‌دار، اندازه هسته، گام، تعداد نقشه‌های ویژگی ورودی و تعداد نقشه‌های ویژگی خروجی است. پارامترهای چیدن صفر دور تصویر در نظر گرفته نمی‌شوند، زیرا همیشه آنها طوری انتخاب می‌شوند که عملیات دقیقاً تمام ابعاد فضایی را با ضرب بزرگنمایی S مقیاس کند.



شکل (۴): ساختار بلوک توجه

طبق شکل شماره (۴) بلوک بزرگنمایی برپایه توجه<sup>۳</sup> که از تصویر ورودی طی سه بخش مجزا کانولوشن گرفته می‌شود و بلوکهای Sup, Wup, Qup ساخته می‌شوند. بلوکهای Sup, Wup بصورت بزرگنمایی صفر ابعادشان بزرگ می‌شود و بلوک Qup از طریق درون یابی بزرگنمایی می‌شود. در نهایت این سه بلوک در ماتریس M نگاشت می‌شوند. تا تصویر بزرگنمایی در ماتریس Z حاصل شود.

$$Z \in \mathbb{R}^{4 \times 4 \times 4 \times 4} \quad (1)$$

S ضرب بزرگنمایی می باشد.

طبق شکل شماره (۳) در داخل هر کدام از بلوک های HS از ۷ لایه کانولوشن، ۷ لایه استخر<sup>۱</sup>، ۱ لایه ترکیب<sup>۲</sup>، ۱ لایه کانولوشن ۱\*۱ در ورودی، ۱ لایه ترکیب و یک لایه کانولوشن ۱\*۱ در خروجی ساخته شده است. هر طبقه خودش بصورت مجرا داری ساختار شبکه باقیمانده است. در ساختار شبکه از روش بهینه سازی برای انتخاب ابعاد فیلترها و دیگر فرآیندها استفاده شده است. این کار باعث می‌شود فیلترهایی پویا طراحی شود که متناسب با میزان بزرگنمایی تنظیم شوند و تصویر نهایی از کیفیت بالاتری برخوردار شود.

### ۳-۲- بخش بزرگنمایی با کمک بلوک توجه

لایه های کانولوشن بخش جدایی ناپذیر بسیاری از راه حل‌های شبکه عصبی عمیق در بینایی کامپیوتر هستند. کارهای اخیر نشان می‌دهند که جایگزینی عملیات کانولوشن استاندارد با روش‌های مبتنی بر توجه به خود منجر به بهبود عملکرد در طبقه‌بندی تصویر و وظایف تشخیص اشیا می‌شود. در این کار، نشان داده شده که چگونه می‌توان از روش‌های توجه برای جایگزینی یک عملیات متعارف دیگر استفاده کرد که کانولوشن انتقالی گام‌به‌گام نام گذاری شده است.

عملیات جدید مبتنی بر توجه را نمونه برداری مبتنی بر توجه نام گذاری شده است، زیرا ابعاد فضایی نقشه‌های ویژگی را افزایش یا نمونه سازی می‌کند. نمونه برداری مبتنی بر توجه به طور مداوم در حالیکه از پارامترهای کمتری استفاده می‌کند، از روش‌های نمونه برداری سنتی مبتنی بر کانولوشن انتقالی گام‌دار یا براساس فیلترهای تطبیقی بهتر عمل می‌کند.

لایه های کانولوشن بخش جدایی ناپذیر از اکثر شبکه های عصبی عمیق هستند که در کاربردهای بینایی مانند تشخیص تصویر [۷۷] و [۷۸] تقسیم بندی تصویر [۷۹] و [۸۰] فیلترکردن تصویر مشترک [۸۱] و [۸۲] استفاده میشوند. بزرگنمایی تصاویر دیجیتال [۸۳] و [۸۴] نام برد. عملیات کانولوشن هنگامی که با ادغام همراه شود، ویژگی‌های مفیدی را در وظایف طبقه‌بندی تصویر معرفی می‌کند.

روش مبتنی بر توجه در این مقاله نیز از پارامترهای بسیار کمتری استفاده می‌کند. بنابراین، لایه نمونه برداری مبتنی بر توجه پیشنهادی می‌تواند به عنوان جایگزینی برای لایه کانولوشن انتقالی گام‌به‌گام استفاده شود، که عملکرد مدل را به دلیل ماهیت تطبیقی محتوا و هسته‌های تولیدشده با بلوک توجه، بهبود می‌بخشد. بر خلاف عملیات کانولوشن انتقالی استاندارد گام به گام، فرمول مبتنی بر توجه در این مقاله سه بلوک میانی تولید می‌کند، یعنی بلوکهای پرسوجو، کلید و مقدار می‌باشد. این موارد امکان انعطاف پذیری بیشتری را در تعریف روش نمونه برداری مبتنی بر

<sup>۱</sup>pooling

<sup>۲</sup>concat

<sup>۳</sup>attention-based upsampling

بهینه‌سازی فرایامترها عملاً یک بهینه‌سازی کلی تابع خطا با عنوان جعبه سیاه است که ارزیابی آن پرهزینه و زمانبر است، تابع  $f$  یک انتخاب فرایامتر از فضای جستجو  $D$  قابل پیکربندی را به خطای اعتبار سنجی یک الگوریتم یادگیری عمیق با پارامترهای شبکه می‌آموزد. تابع بهینه ساز  $f$  به صورت زیر روشی برای جستجوی خودکار برای یافتن فرایامترهای بهینه ارائه می‌دهد.

$$\min_{\theta \in D} f(x, Z_{val}) \quad (۳)$$

$$= \operatorname{argmin}_{\theta \in D} f(x, Z_{train}) \quad (۴)$$

$Z_{val}$  و  $Z_{train}$  نشان‌دهنده آموزش و اعتبار سنجی هستند،  $\theta$  از طریق به حداقل رساندن خطای آموزش و  $x$  از یک مجموعه محدود می‌باشد. حل مساله رابطه بالا به دلیل پیچیدگی بالای عملکرد چالش برانگیز است. از طرفی می‌دانیم شبکه عصبی عمیق بطور کامل غیر خطی است. بنابراین بهینه‌سازی روش کلاسیک با استفاده از مشتق تابع غیر ممکن می‌باشد. بهینه‌سازی بدون مشتق، که بهینه‌سازی جعبه سیاه نامیده می‌شود، می‌تواند یک راه‌حل خوبی باشد که از اطلاعات مشتق در مفهوم کلاسیک برای یافتن راه‌حل‌های بهینه استفاده نکند.

فرا پارامترهای یک شبکه عصبی از جمله آن‌هایی که ساختار خود شبکه را مشخص کرده و آن‌هایی که نحوه آموزش شبکه را مشخص می‌کنند را می‌توان فرا پارامتر یک شبکه در نظر گرفت. از آنجا که سرعت آموزش این شبکه‌ها کند است و چندین پارامتر برای پیکربندی وجود دارد، تنظیم فرا پارامترها برای این شبکه‌ها دشوار است. در هنگام آموزش چنین شبکه‌ای، نتیجه این مدل نه تنها بر روی ساختار انتخاب شده بلکه در روش آموزشی بکار گرفته شده برای تعیین پارامترهای شبکه بستگی خواهد داشت.

خود روش آموزشی دارای چندین فرایامتر مثل نرخ یادگیری، تابع زیان، اندازه دسته‌ای کوچک، تعداد تکرارهای آموزش و تکانه-هاست. علاوه بر این، ساختار شبکه عصبی شامل فرایامتر متعدد در طراحی آن، شامل اندازه و غیر خطی بودن هر لایه، از جمله تعداد واحدهای پنهان، وزن، مقداردهی اولیه لایه‌ها، مقداردهی اولیه وزن‌ها، مقادیر تصادفی و میانگین مدل و داده‌های ورودی پیش پردازش شده است.

روش جستجوی کلی در فضای جستجو بدین شرح است، که بهینه‌سازی فرایامتر یک مساله بسیار مهم برای مدل‌های یادگیری ماشین است که نسبت به مقادیر پارامتر حساس هستند. شناخته‌شده‌ترین شیوه‌ها در این زمینه جستجوی شبکه‌ای و جستجوی تصادفی می‌باشد. هنگامی که جستجوی شبکه برای مدل یادگیری ماشین مورد استفاده قرار می‌گیرد، می‌توان بهترین روش را با تنظیم محدوده و فاصله مناسب پیدا کرد؛ با این حال، زمان زیادی برای یادگیری مدل مورد نیاز است. به طور خاص، همانطور که تعداد فرا پارامترها افزایش می‌یابد، تعداد نقاط شبکه را می‌تواند به صورت نمایی افزایش یابد و ساخت فرا پارامتر بهینه را تقریباً غیر ممکن بسازد. بنابراین وقتی تعداد فرا پارامترها زیاد

$$M(L, J) = \begin{cases} 0 & \text{if } (L \bmod s = 0) \text{ and } (J \bmod s = 0) \\ \infty & \text{otherwise} \end{cases} \quad (۲)$$

بنابراین، ماسک  $M$  در تمام موقعیت‌های صفر شده در ماتریس‌های  $Q$ ،  $W$ ، نمونه برداری شده، دارای مقدار  $\infty$  است که از طریق آموزش بدست می‌آید. سپس از مکانیزم توجه پوشانده شده برای بدست آوردن نقشه‌های ویژگی خروجی استفاده می‌کند.

هر جا که مقدار ماسک  $\infty$  باشد، ضرایب توجه را در این موقعیت‌ها به دلیل عملکرد softmax به طور موثر صفر می‌کند. در واقع، این باعث می‌شود که این موقعیت‌ها توسط مکانیزم توجه نادیده گرفته شود. مکانیزم نمونه برداری مبتنی بر توجه که در شکل (۴) نشان داده شده، ملاحظه می‌شود که با استفاده از درونیایی دوخطی، بردارهای پرسوجو را در موقعیت‌های فضایی جدید تولید می‌کند. سپس مکانیزم توجه از نقشه پرسوجو با وضوح بالا استفاده می‌کند تا اطلاعات را فقط از موقعیت‌های معتبر (غیر صفر) در نقشه مقادیر نمونه برداری صفر با استفاده از بردارهای کلیدی در این موقعیت‌های معتبر جمع کند. لایه نمونه برداری پیشنهادی همچنین می‌تواند با اجازه دادن به ادغام مبتنی بر توجه ویژگی‌ها از روش‌های مختلف، فراتر از لایه کانولوشن انتقالی گام به گام برود که مرز بین رنگ‌ها در سطوح مختلف حفظ شود.

### ۳-۳- روش بهینه‌سازی ساختاری

اکثر الگوریتم‌های یادگیری ماشین نیازمند تنظیمات پارامترهایی هستند که محققین یا مهندسين باید با سعی و خطا یا تجربه آنها را مشخص کنند. این تنظیمات فرا پارامتر<sup>۱</sup> نامیده می‌شوند و برای بهینه‌سازی عملکرد الگوریتم شبکه عصبی مورد نیاز می‌باشد [۸۸]. برای مثال، یک شبکه عصبی عمیق فرا پارامترهای مختلفی مثل نرخ یادگیری، تابع هزینه، پارامتر تنظیم‌کننده، تعداد لایه‌های پنهان و غیره دارد. سرعت یادگیری، سرعت حرکت در جهت شیب را تعیین می‌کند. اگر نرخ یادگیری بسیار کوچک باشد، سرعت یادگیری بسیار کند است و بالعکس اگر به طور قابل توجهی بزرگ باشد، استفاده از مجموعه آموزشی تصادفی بوده و نتیجه آن همگرا نخواهد بود. نرخ یادگیری نیز باید مطابق با مجموعه داده آموزشی و شبکه مورد استفاده تنظیم شود. به عنوان مثالی دیگر، تابع هزینه ممکن است از روش حداقل مربع خطا و یا از تابع آنتروپی بسته به نوع کاربری شبکه استفاده شود. تعداد لایه‌های پنهان نیز یکی از فرایامترها است. هرچه تعداد لایه‌های پنهان بزرگ‌تر باشد، شبکه می‌تواند برای داده‌های آموزشی خاصی بهینه شود. تنظیم این متغیرها برای عملکرد بهینه لازم می‌باشد، که هیچ راه ساده و سریعی برای تضمین بهترین عملکرد در برخی از مجموعه داده‌ها و شبکه‌ها وجود ندارد [۸۹] و [۹۰].

<sup>۱</sup>hyperparameters



در جستجو با uDEAS، فضای جستجو به طور پیوسته به وسیله تعداد محدودی از شبکه‌بندی با فواصل محدود تقسیم می‌شود، به این معنی که مسیرهای جستجو به طور خودکار با توجه به شکل توابع هزینه تعیین می‌شوند. بنابراین اگر uDEAS برای هر دو ماتریس یکسان اعمال شود، فرآیند جستجو دقیقاً همان نتایج جستجو را بدست خواهد آورد. علاوه بر این، uds را قادر می‌سازد که به صورت عمودی جستجو کند.

برای جلوگیری از محاسبات غیر ضروری در انتخاب‌های سراسری مشابه، باید در هر نمونه جدید از یک بررسی تشخیص داده استفاده شود، تا روند تکراری حذف شود. ساده‌ترین راه برای اجرای این مورد، حفظ ماتریس باینری فعلی در یک جدول جستجو به عنوان یک عدد حقیقی است و آن را به صورت پی‌درپی با داده‌های گذشته مقایسه می‌کند. برای این منظور، نمایندگان جستجو برای یک رشته از یک ماتریس ذخیره می‌شوند و در جدول جستجو مقایسه می‌شوند.

گام بعدی شبکه را با مقادیر اولیه داده‌شده آموزش می‌دهد. هنگامی که آموزش تمام شد، تابع هزینه به uDEAS به روزرسانی می‌شود. اکثر توابع هزینه را می‌توان با دقت مدل، انتخاب کرد. با این حال، تابع هزینه می‌تواند با تعریف جدید تغییر کند. به عنوان مثال، در مورد تابع  $f$ ، یک تابع هزینه می‌تواند به عنوان تفاوت بین مقدار کدگذاری و مقدار بازسازی شده توسط رمزگشایی تعریف شود. گام بعدی اجرای uDEAS با هزینه به روزرسانی شبکه است. هنگامی که جستجوی uDEAS اجرا می‌شود، فرآیندها برای مرحله بعدی به روز رسانی می‌شوند. پس از به روز رسانی پارامترها، مدل شبکه برای راه‌اندازی مجدد با استفاده از این فرآیندها تنظیم شده‌است و مدل آموزش داده می‌شود. تا زمانی که این uDEAS به پایان برسد، روش پیشنهادی می‌تواند مدلی را با فرآیندها بهینه پیدا کند که هزینه را به حداقل می‌رساند.

#### ۴- آزمایشات و بررسی نتایج

##### ۴-۱- جزئیات شبیه سازی

در شکل شماره (۱) معماری اصلی بکار رفته در این مقاله دیده می‌شود. در بخش اول دو لایه کانولوشن برای حذف نویز وجود دارد، که عمق هر یک ۶۴ می‌باشد و ابعاد فیلتر متناسب با بخش-بندی تصویر ورودی انتخاب می‌شود. در بخش دوم سطوح چهارگانه مراحل سلسله‌مراتبی قرار داده شده‌است. در ابتدا تصویر وارد H1 می‌شود و خروجی آن به سطح بعدی و یک خروجی به لایه نهایی سریال سازی ویژگی‌ها برای بخش بعدی می‌رود. هر یک از سطوح H1 الی H4 دارای بلوک داخلی طبق شکل شماره (۲) می‌باشد. در این قسمت ویژگی‌های بلوک سلسله‌مراتبی وارد یک لایه کانولوشن  $1*1$  شده سپس وارد ۴ بلوک H1 الی HS4 می‌شوند و خروجی آنها بصورت مجزا وارد واحد سریال سازی و در نهایت وارد یک لایه کانولوشن  $1*1$  شده و در نهایت وارد سطح

است، روش جستجوی شبکه پیچیدگی محاسباتی عظیمی دارد. بنابراین، یک جستجوی تصادفی می‌تواند یک راه‌حل مناسب و ساده در اجرا باشد چون نتیجه بهبود یافته بهینه‌سازی شده فرآیندها را ارائه می‌دهد.

بهینه‌سازی پیشنهاد شده برای شبکه‌های عصبی عمیق الگوریتم کدگذاری پویا برای جستجوها<sup>۱</sup> می‌باشد. الگوریتم رمزگذاری پویای تک متغیری برای جستجوها uDEAS یک تکنیک بهینه‌سازی غیرخطی است که به دنبال بهینه‌سازی سراسری تابع هزینه بدون استفاده از اطلاعات مشتق تابع هزینه است. بطور کلی، uDEAS را می‌توان به جستجوی سراسری و جستجوی محلی تقسیم کرد. علاوه بر این، بهینه‌سازی محلی توسط جستجوی BSS<sup>۲</sup> و یک جستجوی هدایتی<sup>۳</sup> uds انجام می‌شود که از منطقه همسایه بر پایه اطلاعات بدست‌آمده، کار جستجو را انجام می‌دهد. برای بهینه‌سازی در سطح سراسری، این مطالعه روش تصادفی را اتخاذ کرده که ساده‌ترین پیاده‌سازی در عین حال سریع‌ترین است.

استراتژی جستجوی محلی در uDEAS توسط BSS و یک uds اجرا می‌شود. BSS از اضافه کردن ۰ و ۱ در سمت راست یا چپ یک عدد باینری دلخواه بدست می‌آید، که منجر به کاهش یا افزایش در مقدار حقیقی معادل با مقدار حقیقی عدد باینری اولیه می‌شود.

یک رشته دودویی  $n$  بیتی را فرض کنیم. اگر ۰ به کم ارزش‌ترین بیت اضافه شود، رشته جدید به عنوان رشته‌ای به دست چپ خوانده می‌شود. در مقابل اگر ۱ اضافه شود، یک رشته دست‌راست نامیده می‌شود، این رشته‌ها به اعداد حقیقی نرمال شده در بین ۰ و ۱ توسط تابع رمزگشایی، رمز گشایی می‌شوند میتوان به سمت محل پاسخ حرکت کرد، با این حال، تابع کدگذاری منجر به تفاوت‌های حقیقی بین رشته‌ای والد و دو حلقه فرزند آن می‌شود.

BSS فرآیند قرار دادن یک رقم دودویی است و در مقایسه با پخش-کردن شاخه‌های فرزند می‌باشد، که نقاط جستجوی مجاور را بررسی می‌کنند، اصل اساس BSS مشابه روش جستجوی دودویی در علوم رایانه، جستجوی دو بخشی و روش نصف کردن فاصله در بهینه‌سازی غیرخطی است. بنابراین، BSS نرخ هم‌گرایی بالایی را حفظ می‌کند. با این حال، اگر چه یک جستجو در بالاترین گره ۰ یا ۱ در درخت شروع می‌شود، BSS به تنهایی نمی‌تواند از شاخه بالایی یا پایینی فرار کند، چون روال‌های جستجو در یک شاخه واحد محدود شده‌اند. به دلیل ویژگی‌های مکمل BSS و uds، ترکیبی از BSS و چندین uds واحد در DEAS به عنوان یک جلسه در هر افزایشی در طول رشته استفاده می‌شود.

<sup>۱</sup>Univariate dynamic encoding algorithm for searches

<sup>۲</sup>Bisectional search

<sup>۳</sup>uni-directional search

#### ۴-۲- انجام آزمایشات

در این مقاله برای آموزش و تست از مجموعه تصاویر جدول شماره (۴) استفاده شده است. در این جدول می‌توان نام برخی از پایگاه داده تصاویر که در این حوضه فعالیت دارند را با جزئیات بیشتری دید.

برای آموزش شبکه از ۳۰۰ تصویر و برای تست از ۳۰ تصویر بصورت تصادفی از پایگاه‌های داده انتخاب شده اند. معماری شبکه به عنوان داده اولیه با دانش قبلی برنامه ریزی شده و فضای جستجو با داده اولیه معماری شبکه و متغیرهای فرآیند در نظر گرفته شده‌اند. در طول آموزش، خطای آموزشی مبتنی بر میزان سیگنال به نویز برای بهینه سازی شبکه پیشنهادی انتخاب و برای مقایسه از رابطه میزان شباهت در نظر گرفته شد. تنظیمات فرآیند آموزش پس از دوره اول آموزش طبق روال گفته شده، مقادیر به سمت بهینه حرکت می‌کنند.

برای بدست آوردن میزان رضایت نهایی از تصویر خروجی، رابطه سیگنال به نویز استفاده شده است. در رابطه زیر مقدار سیگنال به نویز تصویر خروجی را می‌توان بدست آورد.

$$PSNR = 10 \log \left( \frac{L^2}{\frac{1}{N} \sum_{i=1}^N (I(i) - I_o(i))^2} \right) \quad (5)$$

L حداکثر مقدار پیکسل، N تعداد تصاویر، I(i) تصویر اصلی و I<sub>o</sub>(i) تصویر بازسازی شده می‌باشد.

جدول (۱): پایگاه داده تصاویر

نام پایگاه داده	تعداد تصویر	ابعاد تصویر	فرمت	دسته بندی تصاویر
BSDS300 [91]	300	(435; 367)	JPG	حیوان، ساختمان، غذا، منظره، مردم، گیاه و غیره
Set5,14 [92]	5+14	(492; 446)	PNG	انسان ها، حیوانات، حشرات، گل ها، سبزیجات، کمیک، اسلاید و غیره
Urban100 [93]	100	(984; 797)	PNG	معماری، شهر، سازه، شهری و غیره
UC Merced [94]	100	(1972; 1437)	PNG	سازه، شهری و غیره

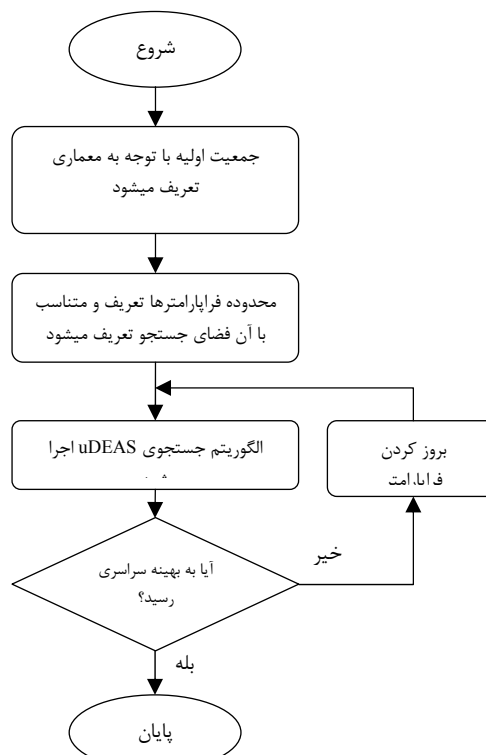
در بررسی ها از پارامتر دیگری برای سنجیدن تصویر خروجی بکار می‌رود که به آن میزان شباهت گفته میشود. SSIM شباهت ساختاری، سنجشی بین تصاویر از نظر روشنایی و کنتراست می‌باشد [95].

در شکل شماره (۶) بخشی از تصویر در مقادیر ۲ و ۴ بزرگ‌نمایی شده‌اند که برای یافتن پارامترهای شبکه از روش uDEAS و بدون استفاده از آن یعنی روش گرادیان کاهشی استفاده شده است. در

بعدی خواهد شد. طبق شکل شماره (۳) داخل بلوک‌های HS از چندین لایه کانولوشن، سریال سازی و فعال سازی تشکیل شده است. ابعاد لایه ها متناسب با مقدار ورودی تصویر تعیین می‌شود و عمق تصویر نهایی خروجی از بخش استخراج ویژگی ها ۶۴ می‌باشد.

در بخش سوم شکل شماره (۱) بخش بزرگ‌نمایی می‌باشد. با توجه به شکل شماره (۴) از بلوک بزرگ‌نمایی برپایه توجه که از تصویر ورودی طی سه بخش مجزا، کانولوشن به صورت کانولوشن انتقالی گام‌دار، گرفته می‌شود و بلوک‌های Qup, Sup, Wup ساخته می‌شوند. بلوک‌های Sup, Wup بصورت بزرگ‌نمایی صفر ابعادشان بزرگ می‌شود و بلوک Qup از طریق درون یابی بزرگ‌نمایی می‌شود. میزان بزرگ شدن این ماتریس‌ها به میزان بزرگ شدن تصویر خروجی بستگی دارد. در نهایت این سه بلوک در ماتریس M نگاشت می‌شوند، تا تصویر بزرگ‌نمایی در ماتریس Z حاصل شود.

مطابق با شکل شماره (۵) روند کار کاملاً مشخص می‌باشد. بعد از اینکه معماری پیاده سازی شد، جمعیت اولیه ایجاد می‌شود، سپس محدوده فرآیند آنها داده می‌شود تا فضای جستجو ایجاد شود. به عنوان مثال میزان ابعاد تصویر ورودی، میزان نرخ یادگیری، عمق لایه‌های کانولوشن، نوع لایه فعال سازی، نوع نگاشت ماتریس‌ها و مانند آنها را می‌توان در ایجاد فضای جستجو دخالت داد. در این بخش با استفاده از الگوریتم uDEAS به جستجوی فرآیند و پارامترهای بهینه معماری ساخته شده رفته و پس از یافتن بهینه سراسری کار به اتمام می‌رسد.



شکل (۵): فرآیند یافتن فرآیند آنها

جدول (۳): مقایسه ای در نوع بزرگ‌نمایی تصویر خروجی

میزان بزرگ‌نمایی	نوع بزرگ‌نمایی			
	بدون بلوک توجه		با بلوک توجه	
	PSNR	SSIM	PSNR	SSIM
x2	32.67	0.9311	<b>33.02</b>	<b>0.9381</b>
x4	26.36	0.7821	<b>26.74</b>	<b>0.8101</b>

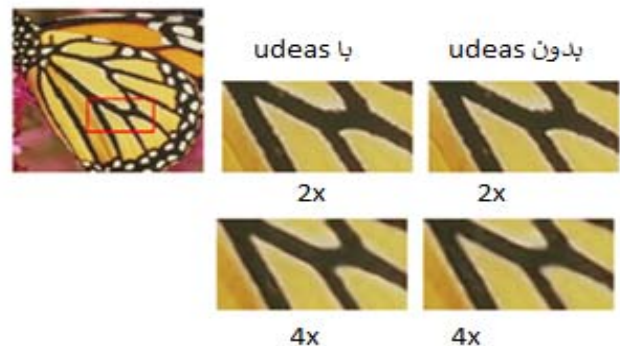
در ادامه به بررسی تفاوت در مقادیر سیگنال به نویز و شباهت دو شبکه بزرگ‌نمایی پرداخته شده است. در بخش بزرگ‌نمایی در کارهای گذشته از روش‌های مختلفی استفاده شده است که عمده مشکل آنها در نوع افزایش طول و عرض تصویر نهایی و میزان عملیات پردازشی در انجام این کار می‌باشد. در این مقاله از بلوک بزرگ‌نمایی برپایه توجه که از تصویر ورودی طی سه بخش مجزا، کانولوشن به صورت کانولوشن انتقالی گام‌دار، گرفته می‌شود و بلوک‌های Sup, Wup, Sup, Wup ساخته می‌شوند. بلوک‌های Sup, Wup از صورت بزرگ‌نمایی صفر ابعادشان بزرگ می‌شود و بلوک Qup از طریق درون‌یابی بزرگ‌نمایی می‌شود. در نهایت این سه بلوک در ماتریس M نگاشت می‌شوند، تا تصویر بزرگ‌نمایی در ماتریس Z حاصل شود. تمامی این عملیات باعث می‌شود تا تصویر نهایی از بافت رنگ و مرز بین رنگ قابل قبولی برخوردار باشد. همچنین میزان عملیات برای انجام کار کانولوشن بدلیل استفاده از روش کانولوشن گام‌دار، کاهش یافته است. برای نشان دادن میزان بهبود تصویر، آزمایشی انجام شد که در شبکه اول از بزرگ‌نمایی با توابع کانولوشن مرسوم و در شبکه دوم از بزرگ‌نمایی با بلوک توجه با کانولوشن گام‌دار استفاده شده است که نتایج این آزمایش را می‌توان در جدول شماره (۳) دید. در انجام شبیه‌سازی تعداد طبقات بلوک سلسله مراتبی چهار طبقه و نوع بهینه‌سازی uDEAS می‌باشد و فقط نوع بزرگ‌نمایی تغییر کرده است. ملاحظه می‌شود بزرگ‌نمایی با استفاده از روش بلوک خودتوجه در میزان بزرگ‌نمایی چهار برابره مقدار ۲۶/۷۴ در عدد سیگنال به نویز و همین عدد در بهینه‌سازی با روش گرادیان کاهشی به مقدار ۲۶/۳۶ رسیده است، که نشان دهنده عملکرد بهتر بلوک خودتوجه می‌باشد.

جدول (۴): بررسی تاثیر تعداد طبقات بلوک سلسله مراتبی استخراج ویژگی‌ها

میزان بزرگ‌نمایی	تعداد طبقات بلوک سلسله مراتبی			
	سه طبقه		چهار طبقه	
	PSNR	SSIM	PSNR	SSIM
x2	32.11	0.9301	<b>33.22</b>	<b>0.9381</b>
x4	26.25	0.7815	<b>26.54</b>	<b>0.8001</b>

همانطور که در بخش استخراج ویژگی صحبت شد، در این مقاله برای آشکارسازی ویژگی‌ها از بلوک سلسله مراتبی استفاده شده است. دلیل این کار، پردازش تصویر در سطوح مختلف می‌باشد.

روش uDEAS برای یافتن پارامترها، جستجو دو بخشی می‌باشد. در ابتدا از روش جستجوی تصادفی در فضای جستجوی ایجاد شده، بخشی بصورت تصادفی انتخاب می‌شود و در مرحله دوم از روش جستجوی دودویی بهینه محلی همان بخش یافت می‌شود، در نهایت پس از چندین بار تکرار بهترین بهینه محلی به عنوان بهینه سراسری در نظر گرفته شده و الگوریتم به سراغ بهینه سراسری بعدی می‌رود. حسن این روش سریع بودن در یافتن بهینه سراسری و فرار از گیرافتادن در بهینه محلی می‌باشد. در روش دوم برای یافتن بهینه پارامترها از روش گرادیان کاهشی استفاده شده است. ایراد این روش این است که امکان گیرکردن در بهینه محلی بسیار است و سرعت یافتن بهینه سراسری کند یا حتی ناممکن می‌باشد. با مقایسه بصورت بصری تفاوت‌های خروجی آنها مشهود می‌باشد. همچنین در جدول شماره (۲) مقادیر مقایسه ای را می‌توان مشاهده کرد، که مشخص می‌کند بهینه‌سازی با استفاده از uDEAS نتایج بهتری را رقم زده است. ملاحظه می‌شود بهینه‌سازی با استفاده از روش uDEAS در میزان بزرگ‌نمایی چهار برابره مقدار ۲۶/۸۴ در عدد سیگنال به نویز و همین عدد در بهینه‌سازی با روش گرادیان کاهشی به مقدار ۲۶/۶۳ رسیده است، که نشان دهنده عملکرد بهتر جستجوی uDEAS می‌باشد. در انجام شبیه‌سازی تعداد طبقات بلوک سلسله مراتبی چهار طبقه و نوع بزرگ‌نمایی با استفاده از بلوک توجه می‌باشد و فقط نوع بهینه‌سازی تغییر کرده است.



شکل (۶): بهینه‌سازی با دو روش استفاده از uDEAS و بدون استفاده از آن

جدول (۲): بررسی تفاوت تصویر خروجی با دو نوع بهینه‌سازی

میزان بزرگ‌نمایی	نوع بهینه‌سازی			
	گرادیان		Udeas	
	PSNR	SSIM	PSNR	SSIM
x2	32.82	0.9337	<b>33.12</b>	<b>0.9481</b>
x4	26.63	0.7836	<b>26.84</b>	<b>0.8121</b>

تفاوت در روش بهینه‌سازی شبکه عصبی، همانطور که از جدول شماره (۲) دیده می‌شود، می‌توان تفاوت بین دو روش بهینه‌سازی را ملاحظه کرد. در هر دو مقدار بزرگ‌نمایی، میزان سیگنال به نویز و شباهت بهبود پیدا کرده است.

پارامترها و فرایارامترها، از بین روش‌های مختلف از روش uDEAS استفاده شد، که جستجو را هدفمند می‌کند. این روش بغیر از پارامترهای شبکه عصبی، می‌تواند فرایارامترها را نیز پیدا کرده و تنظیم کند. استفاده از این روش برای نخستین بار است که در چنین کاری استفاده می‌شود و طبق آزمایشات و شبیه‌سازی نتیجه قابل قبولی حاصل شد که در جداول ۲، ۳ و ۴ قابل مشاهده می‌باشد.

## ۵- نتیجه‌گیری

در روند مقاله مشاهده شد که بزرگ‌نمایی تصاویر دیجیتال یکی از کارهای مهم در پردازش تصویر می‌باشد. این کار در حوزه‌های مختلف پردازش تصویر قابل استفاده می‌باشد. در روش‌های گذشته از حوزه‌های مختلف، کارهای بسیاری ارایه شده است، که هر کدام دارای نقاط قوت و ضعفی بودند. در اینجا از روش سلسله مراتبی برای استخراج ویژگی ترکیب رنگ و مرز بین آن‌ها استفاده شد. نتیجه آزمایشات در خصوص شبکه با بلوک سلسله مراتبی سه طبقه و چهار طبقه نشان داد که تعداد طبقه بالاتر بافت مرزی بین رنگها را بهتر مشخص می‌کند، که تصویر نهایی از کیفیت بهتری برخوردار می‌شود. پس از آن در مرحله بزرگ‌نمایی از بلوک خودتوجه با روش نوینی در کانونوشن، برای رسیدن به کارایی بهتر و سریع‌تر استفاده شد که نتایج مورد قبولی را به همراه داشت. بهینه‌سازی پارامترها و فرایارامترها که این بخش یک حوزه بسیار وسیع در یادگیری ماشین و شبکه عصبی عمیق با نام جستجوی معماری عصبی معروف شده، می‌باشد. در این مقاله برای یافتن بهترین پارامتر و فرایارامتر از روش uDEAS استفاده شد. استفاده از این روش برای نخستین بار بود که در چنین کاربردی استفاده می‌شد و طبق آزمایشات و شبیه‌سازی نتیجه قابل قبولی حاصل شد. طبق جدول شماره (۵) می‌توان دید که در عمل معماری این مقاله توانست در پایگاه داده Urban100 مقدار سیگنال به نویز به عدد ۲۶/۹۴ و میزان شباهت به عدد ۰/۸۱۱۷، در پایگاه داده BSD300 مقدار سیگنال به نویز به عدد ۲۷/۸۵ و میزان شباهت به عدد ۰/۷۶۱۰، در پایگاه داده set14 مقدار سیگنال به نویز به عدد ۲۸/۹۸ و میزان شباهت به عدد ۰/۷۹۲۳، در پایگاه داده set5 مقدار سیگنال به نویز به عدد ۳۲/۶۶ و میزان شباهت به عدد ۰/۹۰۲۰ در چهار برابر کردن تصویر دست یابد. در بررسی بصری نیز می‌توان نتایج را از روی تصاویر بررسی کرد. در شکل شماره (۶) بخشی از بال پروانه بزرگ‌نمایی شده است و مشخص است بهینه‌سازی با کمک روش uDEAS از تفکیک رنگ‌بندی و مرز بین رنگ بهتری برخوردار است. در شکل شماره (۷) این موضوع با تعداد طبقات سلسله مراتبی بررسی شده است که ملاحظه می‌شود پردازش فرکانس بالای تصویر در چهار طبقه بلوک سلسله مراتبی نتیجه بهتری را رقم زده است. تمامی این آزمایشات نشان داد که استفاده همزمان از روش‌های گفته شده نتیجه خوبی به همراه دارد و توانست رقبای خود را شکست دهد. که بررسی مقادیر با رقا در جدول شماره (۵) آمده است.

در یک تصویر، فرکانس‌های بالاتر می‌تواند مرز بین رنگ‌ها را ایجاد کند که متناسب با میزان تغییر در سطوح رنگ‌ها میزان مرز مشخص می‌شود. حال با پردازش بصورت سلسله مراتبی و سطوح چهارگانه در این بلوک، هرچه تصویر به لایه‌های بیشتری وارد می‌شود میزان پردازش روی فرکانس‌های بالاتر بیشتر می‌شود و ویژگی‌های سطوح بالاتر آشکار می‌شود. پس با انجام این کار فرکانس‌های بالاتر که حاوی اطلاعات مهمی از مرز بین رنگ‌ها هستند، پردازش عمیقتری را خواهند داشت، تا مرز بین رنگ‌ها مشخص شود. در شکل شماره (۷) مقایسه‌ای از تعداد طبقات شبکه با بلوک سلسله مراتبی نتیجه آشکارسازی ویژگی‌ها را نشان می‌دهد، که مشخص می‌شود بلوک چهار طبقه دارای بافت رنگ بهتری می‌باشد. پس از انجام آزمایش، نتیجه این موضوع در جدول شماره (۴) از نظر پارامتر سیگنال به نویز و شباهت مورد بررسی قرار گرفته است.

در انجام شبیه‌سازی نوع بهینه‌سازی uDEAS و نوع بزرگ‌نمایی با استفاده از بلوک توجه می‌باشد و فقط تعداد طبقات بلوک سلسله مراتبی تغییر کرده است. با توجه به جدول شماره (۴) مقادیر مقایسه‌ای را می‌توان مشاهده کرد، که مشخص می‌کند بلوک سلسله مراتبی چهار لایه نتایج بهتری را رقم زده است. ملاحظه می‌شود بلوک چهار لایه در میزان بزرگ‌نمایی چهار برابر به مقدار ۲۶/۵۴ در عدد سیگنال به نویز و همین عدد در بهینه‌سازی با روش گرادیان کاهشی به مقدار ۲۶/۲۵ رسیده است، که نشان دهنده عملکرد بهتر بلوک چهار لایه می‌باشد.



شکل (۷): مقایسه بزرگ‌نمایی با بلوکهای سلسله مراتبی ۳ و ۴ طبقه

با توجه به استفاده از بخش‌های مختلفی که در قبل به بررسی آنها پرداخته شد، می‌توانیم در این مقاله نام شبکه را، بزرگ‌نمایی تصاویر دیجیتالی خودتوجه با بهینه‌سازی uDEAS به اختصار AUSR نامید. در شکل شماره (۸) می‌توان سه نوع بزرگ‌نمایی به نام های Bicubic، HFAS و AUSR را به میزان ۴ برابر بزرگ‌نمایی در دو بخش از تصویر را ملاحظه کرد. در این بخش ملاحظه شد که با استفاده از روش سلسله مراتبی برای استخراج ویژگی ترکیب رنگ و مرز بین آن‌ها با کمک بلوک سلسله مراتبی چهار طبقه، بافت مرزی بین رنگها را بهتر مشخص مس‌کند، سپس در مرحله بزرگ‌نمایی از بلوک خودتوجه با روش جدیدی به نام کانونوشن گام‌دار، برای رسیدن به کارایی بهتر و سریع‌تر استفاده شد که این کار میزان پارامترها را کاهش داد. در بهینه‌سازی



جدول (۵): بررسی شبکه نهایی و مقایسه آن با روش‌های دیگر در این حوزه

Method	Scale	Set5		Set14		BSD300		Urban100	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bicubic	2	33.66	0.9299	30.24	0.8688	29.56	0.8431	26.88	0.8403
SRCNN	2	36.66	0.9542	32.45	0.9067	31.36	0.8879	29.5	0.8946
VDSR	2	37.53	0.959	33.05	0.913	33.05	0.896	30.77	0.914
DRCN	2	37.63	0.9588	33.04	0.9118	31.85	0.8942	30.75	0.9133
LapSRN	2	37.52	0.9591	33.08	0.913	31.08	0.895	30.41	0.9101
MemNet	2	37.78	0.9597	33.28	0.9142	32.08	0.8978	31.31	0.9195
IDN	2	37.83	0.96	33.3	0.9148	32.08	0.8985	31.27	0.9196
CARN	2	37.76	0.959	33.52	0.9166	32.09	0.8978	31.92	0.9256
NLRN	2	38	0.9603	33.46	0.9159	32.19	0.8992	31.81	0.9249
SMSR	2	38.00	0.9601	33.64	0.9179	32.17	0.8990	32.19	0.9284
MSRN	2	38.08	0.9605	33.74	0.917	32.23	0.9013	32.22	0.9326
D-DBPN	2	38.09	0.96	33.85	0.919	32.27	0.9	32.55	0.9324
EDSR	2	38.11	0.9602	33.92	0.9195	32.32	0.9013	32.93	0.9351
HFAFS	2	38.25	0.9644	34.11	0.9262	32.36	0.9047	32.88	0.9383
RFANet	2	38.26	0.9615	34.16	0.9220	32.41	0.9026	33.33	0.9389
HAN	2	38.27	0.9614	34.16	0.9217	32.41	0.9027	33.35	0.9385
NLSN	2	38.34	0.9618	34.08	0.9231	32.43	0.9027	33.42	<b>0.9394</b>
AUSR	2	<b>38.35</b>	<b>0.9664</b>	<b>34.32</b>	<b>0.9282</b>	<b>32.72</b>	<b>0.908</b>	<b>33.421</b>	0.9345
Bicubic	4	28.42	0.8104	26	0.7027	25.96	0.6675	23.14	0.6577
SRCNN	4	30.48	0.8628	27.5	0.7513	26.9	0.7101	24.52	0.7221
VDSR	4	31.35	0.883	28.02	0.768	27.29	0.0726	25.18	0.754
DRCN	4	31.53	0.8854	28.02	0.767	27.23	0.7233	25.14	0.751
LapSRN	4	31.54	0.885	28.19	0.772	27.32	0.727	25.21	0.756
MemNet	4	31.74	0.8893	28.26	0.7723	27.4	0.7281	25.5	0.763
IDN	4	31.82	0.8903	28.25	0.773	27.41	0.7297	25.41	0.7632
NLRN	4	31.92	0.8916	28.36	0.7745	27.48	0.7306	25.79	0.7729
MSRN	4	32.07	0.8903	28.6	0.7751	27.52	0.7273	26.04	0.7896
SMSR	4	32.12	0.8932	28.55	0.7808	27.55	0.7351	26.11	0.7868
CARN	4	32.13	0.8937	28.6	0.7806	27.58	0.7349	26.07	0.7837
ESRT	4	32.19	0.8947	28.69	0.7833	27.69	0.7379	26.39	0.7962
EDSR	4	32.46	0.8968	28.8	0.7876	27.71	0.742	26.64	0.8033
D-DBPN	4	32.47	0.898	28.82	0.786	27.72	0.74	26.38	0.7946
HFAFS	4	32.48	0.899	28.87	0.7881	27.73	0.7429	26.61	0.8078
NLSN	4	32.59	0.9000	28.87	0.7891	27.78	0.7444	26.96	0.8109
HAN	4	32.64	0.9002	28.90	0.7890	27.80	0.7442	26.85	0.8094
RFANet	4	32.66	0.9004	28.88	0.7894	27.79	0.7442	26.92	0.8112
AUSR	4	<b>32.66</b>	<b>0.902</b>	<b>28.98</b>	<b>0.7923</b>	<b>27.85</b>	<b>0.761</b>	<b>26.94</b>	<b>0.8117</b>



شکل (۸): مقایسه بزرگ‌نمایی از دو منطقه با سه روش مختلف



## ۶- بحث و آینده تحقیقاتی

برای بزرگ‌نمایی تصاویر روش‌های مختلفی ابداع شده است که بصورت اجمالی بررسی شدند. از روش‌های موجود عملکرد شبکه‌های عصبی نمود پیدا کرد و توانستند قابلیت خود را به اثبات برسانند. طبق مطالعات انجام شده مشخص شد که برای یافتن ساختار شبکه و فرآیندها بهترین روش این است که آنها را به روش‌های جستجوی تکاملی و هوش جمعی بسپاریم. ما در این مقاله خود از روش uDEAS بهره بردیم و ملاحظه شد که شبکه نهایی چابکتر و سریعتر به عملکردی بهتر رسید. این روش بهینه‌سازی فعال برای شبکه‌های بزرگ‌نمایی فراگیر نشده و ساختارهای مختلف شبکه‌های بزرگ‌نمایی، بدین صورت بررسی نشده است و جای کار بسیاری دارد. از طرفی هنوز روی بزرگ‌نمایی تصاویر ویدیویی و تصاویر واقعی کار جدی صورت نگرفته است.

## مراجع

- [11] J. Sun, Z. Xu, and H.-Y. Shum, "Image super-resolution using gradient profile prior," in CVPR, 2008.
- [12] K. I. Kim and Y. Kwon, "Single-image super-resolution using sparse regression and natural image prior," TPAMI, vol. 32, 2010.
- [13] Z. Xiong, X. Sun, and F. Wu, "Robust web image/video super-resolution," IEEE Transactions on Image Processing, vol. 19, 2010.
- [14] Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in CVPR, 2004.
- [15] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in ICCV, 2009.
- [16] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," IEEE Transactions on Image Processing, vol. 19, 2010.
- [17] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in ECCV, 2014.
- [18] E. Smith, S. Fujimoto, and D. Meger, "Multi-view silhouette and depth decomposition for high resolution 3d object representation," in Advances in Neural Information Processing Systems, 2018, pp. 6478-6488.
- [19] O. Thawakar, P. W. Patil, A. Dudhane, S. Murala, and U. Kulkarni, "Image and video super resolution using recurrent generative adversarial network," in 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 2019, pp. 1-8.
- [20] X. Wang, K. C. Chan, K. Yu, C. Dong, and C. Change Loy, "Edvr: Video restoration with enhanced deformable convolutional networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2019, pp. 0-0.
- [21] Y. Jo, S. Wug Oh, J. Kang, and S. Joo Kim, "Deep video super-resolution network using dynamic upsampling filters without explicit motion compensation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 3224-3232.
- [22] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 1646-1654.
- [23] C. Ledig et al., "Photo-realistic single image super-resolution using a generative adversarial network," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jul. 2017, pp. 4681-4690.
- [1] J. S. Isaac and R. Kulkarni, "Super resolution techniques for medical image processing," in ICTSD, 2015.
- [2] Y. Huang, L. Shao, and A. F. Frangi, "Simultaneous super-resolution and cross-modality synthesis of 3d medical images using weakly-supervised joint convolutional sparse coding," in CVPR, 2017.
- [3] L. Zhang, H. Zhang, H. Shen, and P. Li, "A super-resolution reconstruction algorithm for surveillance images," Elsevier Signal Processing, vol. 90, 2010.
- [4] P. Rasti, T. Uiboupin, S. Escalera, and G. Anbarjafari, "Convolutional neural network super resolution for face recognition in surveillance monitoring," in AMDO, 2016.
- [5] Haris, G. Shakhnarovich, and N. Ukita, "Task-driven super-resolution: Object detection in low-resolution images," Arxiv:1803.11316, 2018.
- [6] S. Sajjadi, B. Schölkopf, and M. Hirsch, "Enhancenet: Single image super-resolution through automated texture synthesis," in ICCV, 2017.
- [7] مهدی جم‌پور، ملیحه جاویدی "یک معماری شبکه عصبی عمیق مشترک با ویژگی‌های صریح برای بازشناسی امضاء" مجله ماشین بینایی و پردازش تصویر سال هفتم، شماره دوم، زمستان ۱۳۹۹
- [8] C. E. Duchon, "Lanczos filtering in one and two dimensions," Journal of Applied Meteorology, vol. 18, 1979.
- [9] M. Irani and S. Peleg, "Improving resolution by image registration," CVGIP: Graphical Models and Image Processing, vol. 53, 1991.
- [10] G. Freedman and R. Fattal, "Image and video upscaling from local self-examples," TOG, vol. 30, 2011

- [38] X. Hu, H. Mu, X. Zhang, Z. Wang, T. Tan, and J. Sun, "Meta-sr: Amagnification-arbitrary network for super-resolution," in CVPR, 2019.
- [39] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Deeply recursive convolutional network for image super-resolution. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, pages 1637–1645, 2016
- [40] Ding Liu, Bihan Wen, Yuchen Fan, Chen Change Loy, and Thomas S Huang. Non-local recurrent network for image restoration. In *Proceedings of the Advances in Neural Information Processing Systems*, pages 1673–1682, 2018
- [41] Ning Ni, Hanlin Wu, Graduate Student Member, IEEE, and Libao Zhang, Member, IEEE Hierarchical Feature Aggregation and Self-Learning Network for Remote Sensing Image Continuous-Scale Super-Resolution IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 19, 2022
- [42] Wang, L.; Dong, X.; Wang, Y.; Ying, X.; Lin, Z.; An, W.; Guo, Y. Exploring sparsity in image super-resolution for efficient inference. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2022; pp. 4917–4926.
- [43] Liu, J.; Zhang, W.; Tang, Y.; Tang, J.; Wu, G. Residual feature aggregation network for image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 2359–2368.
- [44] Niu, B.; Wen, W.; Ren, W.; Zhang, X.; Yang, L.; Wang, S.; Zhang, K.; Cao, X.; Shen, H. Single image super-resolution via a holistic attention network. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 191–207.
- [45] Mei, Y.; Fan, Y.; Zhou, Y. Image super-resolution with non-local sparse attention. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 3517–3526.
- [46] Lu, Z.; Li, J.; Liu, H.; Huang, C.; Zhang, L.; Zeng, T. Transformer for single image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–24 June 2022; pp. 457–466.
- [47] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," arXiv preprint arXiv:1409.0473, 2014.
- [48] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. Gomez, Ł. Kaiser, and I. Polosukhin, [24] Y. Tai, J. Yang, and X. Liu, "Image super-resolution via deep recursive residual network," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Jul. 2017, pp. 2790–2798.
- [25] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual densenet for image super-resolution," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2018, pp. 2472–2481.
- [26] J. K. Chorowski, D. Bahdanau, D. Serdyuk, K. Cho, and Y. Bengio, "Attention-based models for speech recognition," in Advances in neural information processing systems, 2015, pp. 577–585.
- [27] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in Advances in neural information processing systems, 2017, pp. 5998–6008.
- [28] P. Ramachandran, N. Parmar, A. Vaswani, I. Bello, A. Levskaya, and J. Shlens, "Stand-alone self-attention in vision models," arXiv preprint arXiv:1906.05909, 2019.
- [29] I. Bello, B. Zoph, A. Vaswani, J. Shlens, and Q. V. Le, "Attention augmented convolutional networks," in Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 3286–3295.
- [30] Y. Bengio, Gradient-based optimization of hyperparameters, Neural Comput. 12(8) (2000) 1889–1900.
- [31] X. Guo, J. Yang, C. Wu, C. Wang, Y. Liang, A novel LSSVMs hyper-parameter selection based on particle swarm optimization, Neurocomput. 71 (16) (2008) 3211–3215.
- [32] Oded Maron, Andrew W. Moore, Hoeffding races: accelerating model selection search for classification and function approximation, Adv. Neural Inf. Process. Syst. 6 (1993) 59–66, NIPS 1993.
- [33] S.S. Rao, Engineering Optimization, John Wiley & Sons, 1996.
- [34] YoungJun Yoo "Hyperparameter optimization of deep neural network using univariate dynamic encoding algorithm for searches" Knowledge-Based Systems 178 (2019) 74–83
- [35] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep cnndenoiser prior for image restoration," in CVPR, 2017.
- [36] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Learning a deep convolutional network for image super-resolution. In Proceedings of the European Conference on Computer Vision, pages 184–199. Springer, 2014
- [37] J. Li, F. Fang, K. Mei, and G. Zhang, "Multi-scale residual network for image super-resolution," in ECCV, 2018.

- algorithms, *Adv. Neural Inf. Process. Syst.* 25 (2012) NIPS 2012.
- [61] J.W. Kim, T.G. Kim, Y.S. Park, S.W. Kim, On-load motor parameter identification using univariate dynamic encoding algorithm for searches, *IEEE Trans. Energy Convers.* 23 (3) (2008) 804–813.
- [62] Y.J. Yoo, D.S. Jung, Y.J. Jang, S.C. Won, Fuzzy weighted subtask controller for redundant manipulator, *Robotica* 33 (2) (2015) 295–313.
- [63] Y.J. Jang, S.W. Kim, An estimation of a billet temperature during reheating furnace operation, *Int. J. Control Autom. Syst.* 5 (1) (2007) 43–50.
- [64] F. Aluffi-Pentini, V. Paris, F. Zirilli, Global optimization and stochastic differential equations, *J. Optim. Theory Appl.* 47 (1985) 1–15.
- [65] L. Yong, K. Lishan, D.J. Evans, The annealing evolution algorithm as function optimizer, *Parallel Comput.* 21 (3) (1995) 389–400.
- [66] A.H.G.R. Kan, G.T. Timmer, A stochastic approach to global optimization, in: P.T. Boggs, R.H. Byrd, R.B. Schnabel (Eds.), *Numerical Optimization*, SIAM, Philadelphia, Pennsylvania, 1985, pp. 245–262.
- [67] A.V. Levy, A. Montalvo, The tunneling algorithm for the global minimization of functions, *SIAM J. Sci. Stat. Comput.* 6 (1985) 15–29.
- [68] H. Ratschek, J. Rokne, *New Computer Methods for Global Optimization*, Ellis Horwood, Chichester, UK, 1988.
- [69] D. Cvijovic, J. Klinowski, Taboo search: An approach to the multiple minimization problem, *Science* 267 (1995) 664–666.
- [70] C. Duan, C. Deng, A. Gharaei, J. Wu, B. Wang, Selective maintenance scheduling under stochastic maintenance quality with multiple maintenance actions, *Int. J. Prod. Res.* 56 (23) (2018) 1–19.
- [71] Seyed Ashkan Hoseini Shekarab, Abolfazl Gharaei, Mostafa Karimi, Modelling and optimal lot-sizing of integrated multi-level multi-wholesaler supply chains under the shortage and limited warehouse space: generalized outer approximation, *Int. J. Syst. Sci. Oper. Logist.* (2018) 21, <http://dx.doi.org/10.1080/23302674.2018.1435835>.
- [72] Abolfazl Gharaei, Mostafa Karimi, Seyed Ashkan Hoseini Shekarabi, An integrated multi-product, multi-buyer supply chain under penalty, green, and quality control policies and a vendor managed inventory with consignment stock agreement: The outer approximation with equality relaxation and augmented penalty algorithm, *Appl. Math. Model.* 69 (2019) 223–254, <http://dx.doi.org/10.1016/j.apm.2018.11.035>.
- “Attention is all you need,” in *Advances in neural information processing systems*, 2017, pp. 5998–6008.
- [49] M.-T. Luong, H. Pham, and C. Manning, “Effective approaches to attention-based neural machine translation,” arXiv preprint arXiv:1508.04025, 2015.
- [50] M.-T. Luong, H. Pham, and C. Manning, “Effective approaches to attention-based neural machine translation,” arXiv preprint arXiv:1508.04025, 2015.
- [51] A. Graves, G. Wayne, and I. Danihelka, “Neural Turing machines,” arXiv preprint arXiv:1410.5401, 2014.
- [52] I. Bello, B. Zoph, A. Vaswani, J. Shlens, and Q. Le, “Attention augmented convolutional networks,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 3286–3295.
- [53] H. Zhao, J. Jia, and V. Koltun, “Exploring self-attention for image recognition,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 10 076–10 085.
- [54] Souvik Kundu, Hesham Mostafa, Sharath Nittur Sridhar, Sairam Sundaresan” Attention-based Image Upsampling ” arXiv:2012.09904v1 [cs.CV] 17 Dec 2020
- [55] Ilija Ilijevski, Taimoor Akhtar, Jiashi Feng, Christine Annette Shoemaker, Efficient hyperparameter optimization of deep learning algorithms using deterministic RBF surrogates, in: *Thirty-First AAAI Conference on Artificial Intelligence*, AAAI Publications, 2016.
- [56] J. Bergstra, R. Bardenet, Y. Bengio, B. Kegl, Algorithms for hyper-parameter optimization, *Adv. Neural Inf. Process. Syst.* 24 (2011) NIPS 2011.
- [57] J. Bergstra, D. Yamins, D.D. Cox, Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures, Presented at the 30th International Conference on Machine Learning, ICML 2013, in: *JMLR Workshop and Conference Proceedings Vol. 28* (1), 2013, pp. 115–123.
- [58] K. Eggensperger, M. Feurer, F. Hutter, J. Bergstra, J. Snoek, H. Hoos, K. Leyton-Brown, Towards an empirical foundation for assessing Bayesian optimization of hyperparameters, in: *NIPS workshop on Bayesian Optimization in Theory and Practice 2013*, 2013.
- [59] K. Eggensperger, F. Hutter, H.H. Hoos, K. Leyton-Brown, Efficient benchmarking of hyperparameter optimizers via surrogates, in: *AAAI*, 2015, 2015, pp. 1114–1120.
- [60] J. Snoek, H. Larochelle, R.P. Adams, Practical Bayesian optimization of machine learning

- [85] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in neural information processing systems*, 2017, pp. 5998–6008.
- [86] P. Shaw, J. Uszkoreit, and A. Vaswani, "Self-attention with relative position representations," *arXiv preprint arXiv:1803.02155*, 2018.
- [87] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," *arXiv preprint arXiv:1603.07285*, 2016.
- [88] James Bergstra, Yoshua Bengio, Random search for hyper-parameter optimization, *J. Mach. Learn. Res.* 13 (2012) 281–305.
- [89] J. Snoek, H. Larochelle, R.P. Adams, Practical Bayesian optimization of machine learning algorithms, *Adv. Neural Inf. Process. Syst.* 25 (2012) 2951–2959.
- [90] T. Hinz, N. Navarro-Guerrero, S. Magg, S. Wermter, Speeding up the hyperparameter optimization of deep convolutional neural networks, *Int. J. Comput. Intell. Appl.* 17 (02) (2018) 1850008.
- [91] .D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *ICCV*, 2001.
- [92] R. Zeyde, M. Elad, and M. Protter, "On single image scale up using sparse-representations," in *International Conference on Curves and Surfaces*, 2010.
- [93] J.-B. Huang, A. Singh, and N. Ahuja, "Single image super-resolution from transformed self-exemplars," in *CVPR*, 2015.
- [94] Y. Yang and S. Newsam, "Bag-of-visual-words and spatial extensions for land-use classification," in *Proc. 18th SIGSPATIAL Int. Conf. Adv. Geograph. Inf. Syst.*, 2010, pp. 270–279.
- [95] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [73] Jong-Wook Kim, Sang Woo Kim, A fast computational optimization method: Univariate dynamic encoding algorithm for searches (uDEAS), *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.* E90-A (8) (2007) 1679–1689.
- [74] Jong-Wook Kim, Taegyu Kim, J.-Y. Choi, Sang Woo Kim, On the global convergence of univariate dynamic encoding algorithm for searches (uDEAS), *Int. J. Control Autom. Syst.* 6 (4) (2008) 571–582.
- [75] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1646–1654.
- [76] Z. Li, J. Yang, Z. Liu, X. Yang, G. Jeon, and W. Wu, "Feedback network for image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 3862–3871.
- [77] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [78] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9.
- [79] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. Le-Cun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," *arXiv preprint arXiv:1312.6229*, 2013.
- [80] J. Redmon and A. Farhadi, "Yolo9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271.
- [81] Y. Li, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Joint image filtering with deep convolutional networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 8, pp. 1909–1923, 2019.
- [82] B. Kim, J. Ponce, and B. Ham, "Deformable kernel networks for joint image filtering," *International Journal of Computer Vision*, pp. 1–22, 2020.
- [83] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, 2010.
- [84] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, vol. 2, 2001, pp. 416–423 vol. 2.