

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی برق و ریاضیات

رشته برق گرایش الکترونیک

پایان نامه کارشناسی ارشد

بهبود کارآیی فشرده‌سازی تصاویر طبیعی رنگی مبتنی بر شبکه عصبی

نگارنده: علی کارگران

استاد راهنما

دکتر هادی گرایلو

بهمن ۱۳۹۶

شماره: ۱۵۳۹ / ۳۰۵
تاریخ: ۱۰ آذر ۹۶

باسمه تعالی



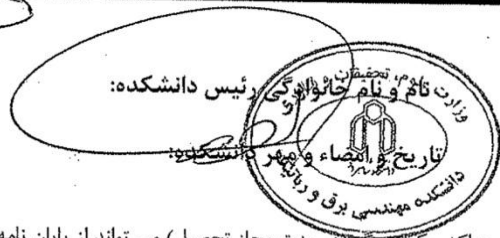
مدیریت تحصیلات تکمیلی

فرم شماره (۳) صورتجلسه نهایی دفاع از پایان نامه دوره کارشناسی ارشد

با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خانم / آقای علی کارگران با شماره دانشجویی ۹۳۱۳۸۰۴ رشته مهندسی الکترونیک گرایش سیستم تحت عنوان: بهبود کارآیی فشرده سازی تصاویر طبیعی رنگی مبتنی بر شبکه عصبی که در تاریخ ۱۳۹۶/۱۱/۱۰ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می گردد:

قبول (با امتیاز ۱۷/۴۵ درجه ۱) مردود
نوع تحقیق: نظری عملی

عضو هیأت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱- استاد راهنمای اول	حاجی کریمی	استاد	
۲- استاد راهنمای دوم	---	---	---
۳- استاد مشاور	---	---	---
۴- نماینده تحصیلات تکمیلی	ولایتی	استاد	
۵- استاد ممتحن اول	عزیز زاده	دانشیار	
۶- استاد ممتحن دوم	سید علی سلیمان زاده	دانشیار	



تبصره: در صورتی که کسی مردود شود حداکثر یکبار دیگر (در مدت مجاز تحصیل) می تواند از پایان نامه خود دفاع نماید (دفاع مجدد نباید زودتر از ۲ ماه برگزار شود).

شکر و قدردانی

به مصداق «من لم یشکر المخلوق لم یشکر الخالق» بر خود لازم می دانم از کلیه کسانی که بنده را در تدوین و نگارش این پایان نامه یاری نمودند صمیمانه تشکر و قدردانی نمایم. به خصوص از استاد فرزانه جناب آقای دکتر گرایلو (استاد راهنما) که در کلیه مراحل انجام این پژوهش با خوشروئی، یاری و راهنمایی ام نمودند و همچنین که وقت خود را بی شائبه در اختیار من گذاشته و با دقت نظر خاصی راهنمایی لازم در این خصوص ارائه نمودند صمیمانه تشکر و قدردانی نمایم.

تعهد نامه

اینجانب علی کارگران دانشجوی دوره کارشناسی ارشد برق گرایش الکترونیک دانشکده مهندسی برق و رباتیک دانشگاه صنعتی شاهرود نویسنده پایان نامه تحت عنوان: **بهبود کارآیی فشرده‌سازی تصاویر طبیعی رنگی مبتنی بر شبکه عصبی تحت راهنمایی دکتر هادی گرایلو متعهد می شوم.**

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است.

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزارها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

چکیده

رشد روز افزون داده‌های دیجیتالی در سال‌های اخیر منجر به افزایش توجهات به فشرده‌سازی شده است. فشرده‌سازی تصویر با شبکه عصبی به ما یک راه امیدبخش برای توسعه تکنیک‌های فشرده‌سازی موثر و جدید را معرفی می‌کند. هدف اصلی این حوزه کاهش افزونگی و افزایش میزان فشرده‌سازی داده است. تاکنون بهبود خوبی در تکنیک‌های فشرده‌سازی تصویر و ویدئو انجام شده است و این بهبودها به دلیل صرفه جویی در حافظه برای انتقال اطلاعات بصری انجام شده است. یکی از چالش‌های اصلی در فشرده‌سازی تصویر کیفیت تصویر بازسازی شده است. در این پایان‌نامه از شبکه عصبی عمیق (DNN) و فیلتر حذف نویز برای فشرده‌سازی تصاویر طبیعی رنگی استفاده شده است. نتایج شبیه‌سازی نشان می‌دهد که روش پیشنهادی ما در کاهش حجم تصویر و میزان فشرده‌سازی موثر است و همچنین کیفیت تصویر بازسازی شده روش پیشنهادی در مقایسه با روش‌های دیگر معرفی شده بهبود پیدا کرده است. بعنوان نمونه در روش پیشنهادی با استفاده از فیلتر حذف نویز در مقایسه با روش فشرده‌سازی تصویر مبتنی بر شبکه عصبی (بر اساس پیکسل‌های مشابه) در نسبت فشرده‌سازی برابر، مقدار PSNR در تصویر Lena از 22.66db به 24.3db و تصویر Mandrill از 19.73db به 22.10db افزایش پیدا کرده است که هم مقادیر را بهبود بخشیده و هم کیفیت تصویر بهبود یافته است.

کلمات کلیدی: فشرده‌سازی تصویر رنگی، شبکه عصبی عمیق، خودرمنزنگار، فیلتر حذف نویز

مقالات استخراج شده از پایان نامه:

کارگران ع، گرایلو ه، (۱۳۹۶)، "فشرده سازی تصاویر رنگی مبتنی بر شبکه عصبی عمیق"، سومین کنفرانس پردازش سیگنال و سیستم های هوشمند ایران، ص ۶۹، شاهرود.

فهرست مطالب

فصل اول:مقدمه

- ۱-۱ مقدمه ۲
- ۲-۱ ضرورت انجام تحقیق ۴
- ۳-۱ روش پیشنهادی ۴
- ۴-۱ ساختار پایان نامه ۵

فصل دوم:مروری بر کارهای پیشین

- ۱-۲ فشرده سازی و بازسازی تصویر با استفاده از رویکردی جدید توسط شبکه عصبی ۸
- ۲-۲ فشرده سازی تصویر با استفاده از شبکه عصبی و موجک هار ۱۰
- ۳-۲ رویکرد شبکه عصبی برای فشرده سازی تصویر رنگی در حوزه تبدیل ۱۴
- ۴-۲ فشرده سازی تصاویر ثابت با استفاده از شبکه های عصبی ۱۶
- ۵-۲ فشرده سازی تصویر با روش ترکیبی ۱۸
- ۶-۲ فشرده سازی تصویر با استفاده از شبکه عصبی پس انتشار ۲۲
- ۷-۲ فشرده سازی تصویر با استفاده از شبکه عصبی چند لایه ۲۳
- ۸-۲ شبکه عصبی برای انتخاب بهترین موجک در فشرده سازی تصویر رنگی ۲۵
- ۹-۲ یک روش فشرده سازی برای تصاویر رنگی با استفاده از شبکه های عصبی چند لایه ۲۷
- ۱۰-۲ فشرده سازی تصویر و اصلاح خطای نویز کانال با استفاده از شبکه عصبی عمیق ۲۹
- ۱۱-۲ رمزگذاری و فشرده سازی تصویر مبتنی بر یادگیری عمیق ۳۱

فصل سوم:مباحث تئوری

- ۱-۳ پردازش تصویر ۳۴
- ۲-۳ شبکه عصبی ۳۴
- ۳-۳ یادگیری عمیق ۳۵

۳۷	تعداد متغیرهای هر لایه در شبکه‌های عمیق
۳۸	یادگیری بدون نظارت برای شبکه‌های عمیق
۳۸	کاربردهای یادگیری بدون نظارت
۳۹	شبکه عصبی عمیق
۴۱	فیلترگذاری روی تصاویر و انواع فیلترهای خطی و غیرخطی
۴۱	فیلترهای خطی
۴۳	فیلترهای غیر خطی
۴۴	فیلترهای پایین گذر
۴۴	فیلترهای بالا گذر
۴۵	خودرمنگار
۴۷	آموزش خودرمنگار

فصل چهارم: روش پیشنهادی و نتایج شبیه‌سازی

۵۰	مقدمه
۵۵	پارامترهای ارزیابی
۵۶	پایگاه داده
۵۶	نتایج شبیه‌سازی
۶۵	نتایج آزمون MOS

فصل پنجم: نتیجه‌گیری و پیشنهادات

۶۸	نتیجه‌گیری
۶۸	پیشنهادات
۶۹	منابع

فهرست جداول

جدول (۱-۴): مقایسه عملکرد PSNR(db) روش پیشنهادی با روش قبل و بعد از اعمال فیلتر حذف نویز.....	۵۶
جدول (۲-۴): مقایسه عملکرد PSNR روش پیشنهادی با روش مشابه بعد و قبل از اعمال فیلتر حذف نویز.....	۶۰
جدول (۳-۴): مقدار زمان تست روش پیشنهادی با مقادیر نشان داده شده در جدول (۲-۴).....	۶۰
جدول (۴-۴): مقایسه بصری فضای رنگ YUV [۴۳] و روش پیشنهادی برای تصویر lena.....	۶۵
جدول (۵-۴): مقایسه بصری فشرده‌سازی بر اساس پیکسل‌های مشابه و روش پیشنهادی برای تصویر Mandrill.....	۶۵
جدول (۶-۴): مقایسه بصری فشرده‌سازی بر اساس پیکسل‌های مشابه و روش پیشنهادی برای تصویر lena.....	۶۶

فهرست اشکال

- شکل (۱-۱) بلوک دیاگرام فرآیند فشرده سازی و رمزگشایی ۳
- شکل (۱-۲) بلوک دیاگرام آموزش شبکه عصبی به منظور فشرده سازی تصویر ۹
- شکل (۲-۲) معماری شبکه عصبی ۹
- شکل (۳-۲) بلوک دیاگرام بازسازی تصویر ۱۰
- شکل (۴-۲) نسبت بین تصاویر و میزان فشرده سازی بهینه ۱۲
- شکل (۵-۲) الف: سیستم فشرده ساز تصاویر بهینه با استفاده از ANN32 ۱۳
- شکل (۵-۲) ب: سیستم فشرده ساز تصاویر بهینه با استفاده از ANN64 ۱۳
- شکل (۶-۲) مقایسه بین اندازه فایل در حوزه فضایی و تبدیل ۱۶
- شکل (۷-۲) تجزیه یک تصویر به تقریب و جزئیات با استفاده از موجک ۱۷
- شکل (۸-۲) عملکرد سیستم با افزایش کیفیت ۱۸
- شکل (۹-۲) ساختار شبکه عصبی سلسله مراتبی ۱۹
- شکل (۱۰-۲) ساختار شبکه عصبی تطبیقی ۲۰
- شکل (۱۱-۲) ساختار شبکه عصبی هیبریدی ۲۱
- شکل (۱۲-۲) خروجی آموزش شبکه ۲۲
- شکل (۱۳-۲) تصویر اصلی و بازسازی شده با ۱۱۰۰ epochs ۲۳
- شکل (۱۴-۲) شبکه رو به جلو چهار لایه برای فشرده سازی ۲۴
- شکل (۱۵-۲) بلوک دیاگرام کدگذاری تصویر مبتنی بر موجک ۲۶
- شکل (۱۶-۲) توپولوژی MLP با دو لایه مخفی ۲۷
- شکل (۱۷-۲) روش فشرده سازی تصویر پیشنهاد شده ۲۸
- شکل (۱۸-۲) توزیع بلوکها در روش پیشنهاد شده ۲۹

- شکل (۲-۱۹) بلوک دیاگرام سیستم فشرده سازی در ۳۰
- شکل (۲-۲۰) ساختار شبکه عصبی عمیق ۳۰
- شکل (۲-۲۱) معماری Stacked Auto-Encoder ۳۱
- شکل (۳-۱) یادگیری لایه لایه ۳۶
- شکل (۳-۲) شبکه عصبی عمیق ۳۸
- شکل (۳-۳) نمونه ای از خوشه بندی بدون نظارت ۳۹
- شکل (۳-۴) معماری شبکه عصبی عمیق [..... ۴۱
- شکل (۳-۵) نمونه ای از فیلتر خطی ۴۲
- شکل (۳-۶) نحوه کانولوشن و اعمال فیلتر بر روی پیکسل های یک تصویر (فیلتر خطی) ۴۳
- شکل (۳-۷) تصویر مات شده پس از اعمال فیلتر ۴۴
- شکل (۳-۸) فیلتر تیزکننده یک ماسک 3×3 ۴۵
- شکل (۳-۹) ساختار کلی اتوانکودر عمیق ۴۶
- شکل (۳-۱۰) نمونه ای از فشرده سازی عدد دستنویس ۴۷
- شکل (۴-۱) الف: بلوک دیاگرام روش پیشنهادی برای آموزش خودرمننگار در فشرده سازی تصویر .. ۵۳
- شکل (۴-۱) ب: بلوک دیاگرام روش پیشنهادی برای آموزش خودرمننگار در بازسازی تصویر ۵۴
- شکل (۴-۲) الف: تصویر lena اصلی ۵۷
- شکل (۴-۲) ب: تصویر lena فشرده شده در فضای YUV ۵۷
- شکل (۴-۲) ج: تصویر lena فشرده شده قبل از اعمال فیلتر در روش پیشنهادی ۵۷
- شکل (۴-۲) د: تصویر lena فشرده شده بعد از اعمال فیلتر در روش پیشنهادی ۵۸
- شکل (۴-۳) الف: نمودار همگرایی فرآیند آموزش مولفه ی رنگ قرمز در روش پیشنهادی ۵۸
- شکل (۴-۳) ب: نمودار همگرایی فرآیند آموزش مولفه ی رنگ سبز در روش پیشنهادی ۵۹
- شکل (۴-۳) ج: نمودار همگرایی فرآیند آموزش مولفه ی رنگ آبی در روش پیشنهادی ۵۹
- شکل (۴-۴) الف: تصویر Mandrill اصلی ۶۱
- شکل (۴-۴) ب: تصویر Mandrill فشرده شده در روش مشابه [۳۴] ۶۱
- شکل (۴-۴) ج: تصویر Mandrill فشرده شده قبل از اعمال فیلتر حذف نویز در روش پیشنهادی ۶۱

- شکل ۴-۴ (د: تصویر Mandrill فشرده شده بعد از اعمال فیلتر حذف نویز در روش پیشنهادی .. ۶۲
- شکل ۴-۵ (الف: تصویر lena اصلی ۶۲
- شکل ۴-۵ (ب: تصویر lena فشرده شده در روش مشابه ۶۲
- شکل ۴-۵ (ج: تصویر lena فشرده شده پس از اعمال فیلتر حذف نویز در روش پیشنهادی ۶۳
- شکل ۴-۶ (الف: نمودار همگرایی فرآیند آموزش مولفه‌ی رنگ قرمز در روش پیشنهادی ۶۳
- شکل ۴-۶ (ب: نمودار همگرایی فرآیند آموزش مولفه‌ی رنگ سبز در روش پیشنهادی ۶۴
- شکل ۴-۶ (ج: نمودار همگرایی فرآیند آموزش مولفه‌ی رنگ آبی در روش پیشنهادی ۶۴

فصل اول

مقدمه

۱-۱ مقدمه

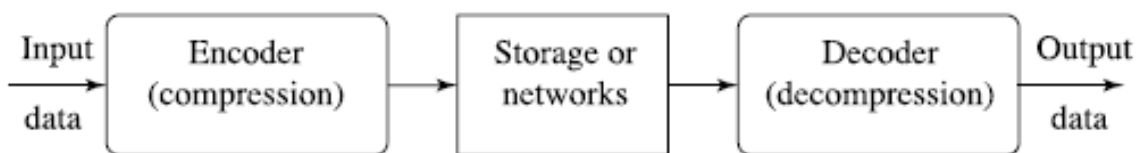
پردازش تصویر یکی از علوم است که بطور چشمگیری روز به روز در زندگی روزمره بکار می رود. بطور گسترده در نشر رومیزی، تصویربرداری پزشکی، تجزیه و تحلیل اهداف نظامی، کنترل اتوماسیون، بینایی ماشین، تصاویر جغرافیایی، گرافیک، هنر و تولیدات چند رسانه ای، نجوم و فضانوردی، استفاده می شود [۱]. امروزه کمتر کارخانه پیشرفته ای وجود دارد که بخشی از خط تولید آن توسط برنامه های هوشمند بینایی ماشین^۱ کنترل نشود. خطای بسیار کم، سرعت زیاد، هزینه نگهداری بسیار پایین، عدم نیاز به حضور اپراتور ۲۴ ساعته و خیلی مزایای دیگر باعث شده که صنایع و کارخانه ها به سرعت به سمت پردازش تصویر و بینایی ماشین روی بیاورند. از آنجایی که در علم هواشناسی تشخیص و پیش بینی آب و هوا اکثرا از طریق تصاویر هوایی و ماهواره ای انجام می گیرد، پردازش تصویر در این علم کاربرد زیادی دارد و دقت و سرعت پیش بینی آب و هوا و طوفان ها را بسیار بالا می برد. کاربرد دیگر پردازش تصویر می تواند در کنترل ترافیک باشد. با گرفتن عکس های هوایی از زمین ترافیک هر قسمت از شهر مشخص می شود. تصاویر ماهواره ای که از شهرها گرفته می شود، می تواند توسط فیلترهای مختلف پردازش تصویر فیلتر شود و اطلاعات مختلفی از آن استخراج شود [۲].

کاربرد فشرده سازی تصویر، فشرده سازی داده ها در تصاویر دیجیتال است. در واقع هدف از این کار کاهش افزونگی داده های تصویر به منظور توانایی ذخیره و یا انتقال داده در قالب کارآمد می باشد. داده های چندرسانه ای فشرده نشده مانند گرافیک، صوت و تصویر به فضای ذخیره سازی قابل توجه و انتقال پهنای باند نیاز دارند. یک ویژگی مشترک بسیاری از این تصاویر این است که پیکسل های همسایه در ارتباط هستند و در فشرده سازی باید این اطلاعات زائد حذف شوند. هدف از کاهش افزونگی برداشتن تکرار از منبع سیگنال (تصویر، ویدئو) است [۳].

¹ Machine vision

فشرده‌سازی می‌تواند به دو صورت با اتلاف و بدون اتلاف صورت گیرد. در فشرده‌سازی بدون اتلاف، کاهش بیت بر اساس شناسایی و حذف افزونگی آماری انجام می‌گیرد و هیچ اطلاعاتی از بین نمی‌رود. برای بعضی تصاویر مثل عکس‌های پزشکی و نقشه‌کشی از فشرده‌سازی بدون اتلاف استفاده می‌شود. در فشرده‌سازی با اتلاف، کاهش بیت بر اساس شناسایی و حذف اطلاعات کم ارزش‌تر صورت می‌گیرد و این اطلاعات برگشت‌پذیر نخواهد بود. این روش معمولاً برای تصاویر طبیعی مناسب است که از دست رفتن درستی برای دست یافتن به کاهش نرخ بیت قابل توجه است. بلوک دیاگرام فرآیند فشرده‌سازی و رمزگشایی در شکل (۱-۱) نشان داده شده است.

اخیراً، روش‌های بهبود کیفیت جهت افزایش کیفیت تصویر به یکی از حوزه‌های فعال تحقیقاتی تبدیل شده است. مزیت مهم روش‌های پردازش تصویر، هزینه پایین آن است و بنابراین کاربردهای فراوانی در سیستم‌های تصویر برداری هوشمند یافته است. الگوریتم‌های متعددی در این پایان‌نامه مورد بررسی قرار گرفته شده است که با فشرده‌سازی تصویر مبتنی بر شبکه عصبی ارتباط دارد، اما روش‌های کمی برای فشرده‌سازی تصاویر رنگی وجود دارد.



شکل (۱-۱) بلوک دیاگرام فرآیند فشرده‌سازی و رمزگشایی [۳۵]

فشرده‌سازی بر اساس معیاری تحت عنوان میزان فشرده‌سازی ارزیابی می‌شود که رابطه آن بصورت (۱-۱) می‌باشد:

$$\text{Compression ratio} = B_0/B_1 \quad (1-1)$$

که در آن B_0 تعداد بیت‌های قبل از فشرده‌سازی و B_1 تعداد بیت‌های بعد از فشرده‌سازی است.

۲-۱ ضرورت انجام تحقیق

ظهور حوزه فشرده‌سازی تصویر مبتنی بر شبکه عصبی به ما یک راه امیدبخش برای توسعه تکنیک‌های فشرده‌سازی موثر و جدید را معرفی می‌کند که هدف اصلی آن کاهش افزونگی و انجام فشرده‌سازی داده است. تاکنون بهبود خوبی در تکنیک‌های فشرده‌سازی تصویر و ویدئو انجام شده است و این بهبودها به دلیل صرفه جویی در حافظه برای انتقال اطلاعات بصری انجام شده است. گرچه تاکنون فشرده‌سازهای استاندارد ویدئو و تصویر مثل JPEG، JPEG2000، MPEG-4، AVC/H.264 و غیره معرفی شده‌اند. به هر حال این روش‌های فشرده‌سازی و تکنیک‌های کدگذاری ممکن است در برخی کاربردهای فشرده‌سازی مخصوص مانند کدگذاری تصویر مقاوم در برابر نویز و فشرده‌سازی تصاویر رمزگذاری شده موثر نباشند. بخش عمده پردازش در فرآیند فشرده‌سازی غیر ترتیبی است لذا پردازش موازی می‌تواند موجب افزایش سرعت آن گردد. در بعضی از تکنیک‌ها به قابلیت یادگیری نیاز می‌باشد، یعنی سیستم یاد بگیرد چگونه به ورودی‌ها پاسخ دهد و این یادگیری را در حافظه خود نگهداری نماید. در فشرده‌سازی زمان و آموزش سریع و همچنین حفظ کیفیت تصویر نکات قابل توجه‌ای است که باید مورد توجه محققین قرار بگیرد. در تکنیک‌های فشرده‌سازی بیان شده می‌توان پس از یادگیری و انطباق نتایج حاصل را به موارد مشابه تعمیم داد.

۳-۱ روش پیشنهادی

شبکه عصبی این امکان را ایجاد نموده است که بتوانیم سیستم فشرده‌سازی با قابلیت‌های اشاره شده در زیربخش قبلی را پیاده‌سازی کنیم که هم برای فشرده‌سازی تصاویر مناسب است و هم قابلیت پردازش موازی اطلاعات و هم آموزش سریع و همچنین امکان توانایی تعمیم دادن را دارا می‌باشد.

در این پایان‌نامه، بهبود کارایی فشرده‌سازی تصاویر طبیعی رنگی مبتنی بر شبکه عصبی معرفی می‌شود و در آن نوعی از معماری یادگیری عمیق^۱ به نام اتوانکودر عمیق^۲ که در ساختارش توانایی

¹ Deep learning

² Deep Autoencoder

رمزگذاری و هم رمزگشایی را دارد استفاده شده است. این روش توانسته عملکرد فشرده‌سازی را بهبود بخشد.

۴-۱ ساختار پایان‌نامه

ادامه پایان‌نامه بصورت زیر سازمان‌دهی شده است. در فصل دوم کارهای انجام شده در فشرده‌سازی تصویر با استفاده از شبکه عصبی آورده شده است و در فصل سوم به معرفی مفاهیم اولیه و مباحث تئوری روش پیشنهادی می‌پردازیم. همچنین روش پیشنهادی و نتایج شبیه‌سازی روش پیشنهاد شده در فصل چهارم آورده شده است. نتیجه‌گیری و همچنین پیشنهادات برای کارهای آتی در این زمینه، در فصل پنجم ارائه شده است.

فصل دوم

مروری بر کارهای پیشین

در سال‌های اخیر تلاش‌های نسبتاً زیادی در زمینه فشردن‌سازی تصویر با استفاده از شبکه عصبی صورت گرفته است که در این فصل تعدادی از این تکنیکها مورد بررسی قرار گرفته است.

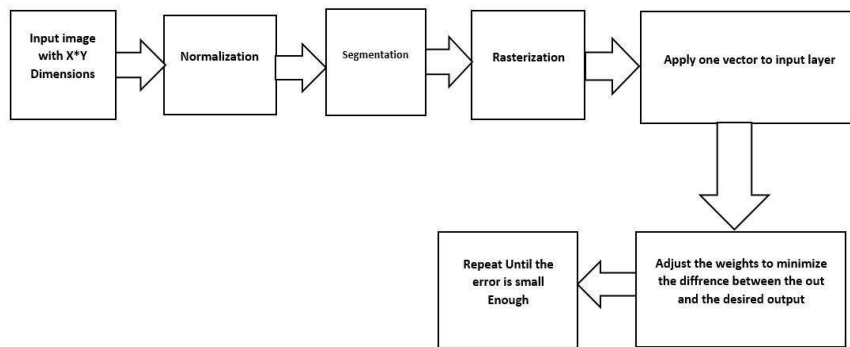
۱-۲ فشردن‌سازی و بازسازی تصویر با استفاده از رویکردی جدید توسط شبکه عصبی

در [۴] یک روش فشردن‌سازی تصویر مبتنی بر شبکه عصبی ارائه شده است. این شبکه کاربردی، شبکه پس انتشار خطا (bp)^۱ است که مقدار وسیعی از داده تصویر را برای ذخیره‌سازی یا انتقال و بازیابی در زمان مطلوب فراهم می‌کند. هم‌چنین یک رویکرد جدید برای کاهش زمان آموزش، توسط بردارهای بازسازی شده نیز ارائه شده است و عملکرد شبکه با استفاده از برخی تصاویر دنیای واقعی ارزیابی شده است. ساختار توسعه‌یافته و الگوریتم آموزشی، نرخ فشردن‌سازی بالا و نویز کم را ایجاد میکند. از یک شبکه رو به جلو شامل سه لایه، یک لایه ورودی با N نرون، یک لایه خروجی با N نرون، و یک لایه مخفی با Y نرون استفاده شده است که Y کوچکتر از N انتخاب می‌شود. مقادیر لایه مخفی همان اطلاعات فشردن شده هستند. در این تحقیق تصویر بطور خطی از محدوده (۰-۲۵۵) به محدوده (۰-۱) تبدیل شده است. لازم است هر یک از تصاویر ورودی به تعدادی بلوک غیر هم‌پوشان با اندازه برابر قطعه‌بندی شوند. این بلوک‌ها می‌توانند $۴*۴$ یا $۸*۸$ یا $۱۶*۱۶$ باشند. در این مقاله بلوک‌های $۸*۸$ انتخاب شده است. هر بلوک $۸*۸$ به یک بردار ۶۴ عنصری تبدیل شده و به ورودی شبکه عصبی دارای ۶۴ نرون ورودی اعمال می‌گردد. با برابر قرار دادن مقادیر نرون‌های لایه ورودی و خروجی شبکه عصبی آموزش می‌بیند. فرآیند آموزش و اصلاح وزن‌ها با دو الگوریتم لونیبرگ - مارکواد^۲ و لونیبرگ - مارکواد اصلاح شده^۳ انجام و مقایسه گردیده است. در شکل (۱-۲) بلوک دیاگرام آموزش شبکه عصبی نشان داده شده است.

^۱ Back Propagation

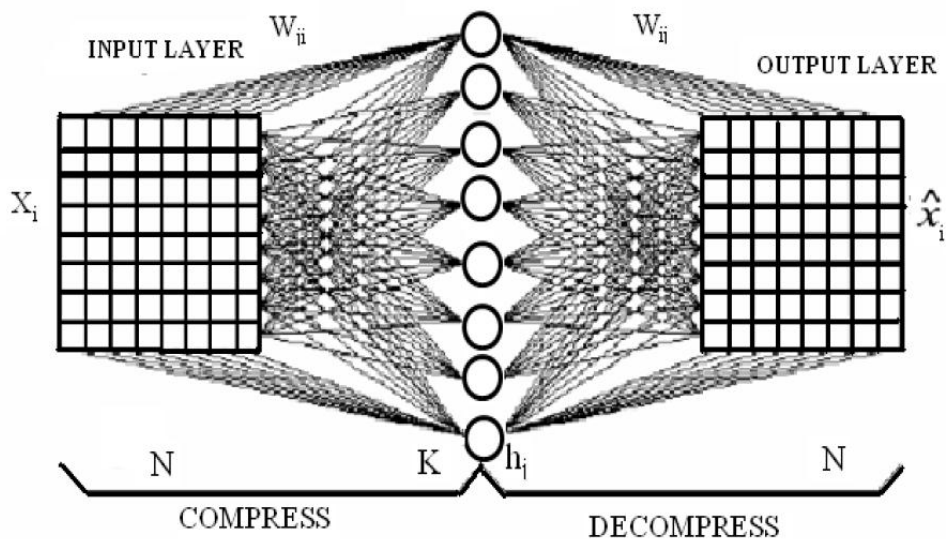
^۲ levenberg marquardt(LM)

^۳ Modified levenberg marquardt(MLM)



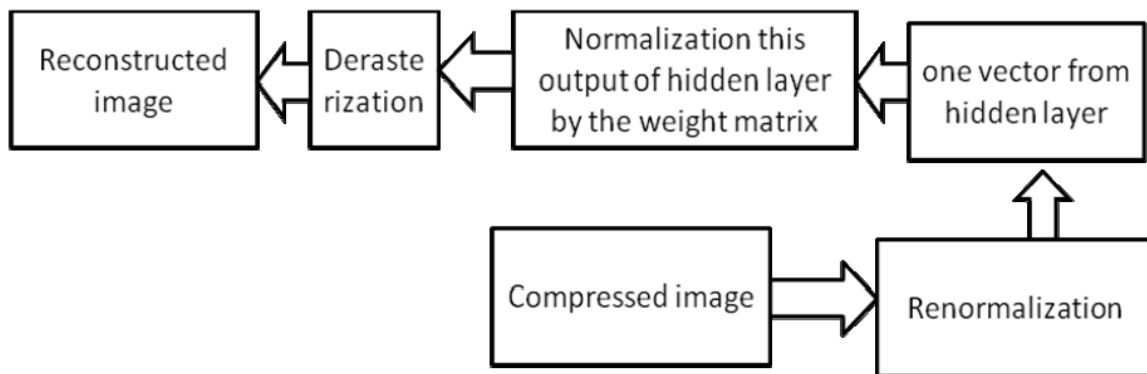
شکل ۲-۱) بلوک دیاگرام آموزش شبکه عصبی به منظور فشرده‌سازی تصویر [۴]

پس انتشار (bp) یکی از شبکه‌های عصبی است که بطور مستقیم برای فشرده‌سازی تصویر در روش ذکر شده اعمال شده است. این ساختار که از نوع رو به جلو می‌باشد لایه ورودی و خروجی بطور کامل به لایه مخفی وصل شده است. فشرده‌سازی بدست آمده تعداد نرون‌ها در لایه مخفی کمتر از نرون‌های لایه خروجی و ورودی است. بلوک دیاگرام معماری شبکه عصبی در شکل (۲-۲) نشان داده شده است.



شکل ۲-۲) معماری شبکه عصبی [۴]

همچنین در شکل (۳-۲) بلوک دیاگرام بازسازی تصویر نشان داده شده است.



شکل ۲-۳) بلوک دیاگرام بازسازی تصویر [۴]

همچنین استفاده از الگوریتم اصلاح شده^۱ LM موجب کاهش زمان آموزش استفاده شده است [۶۵]. در الگوریتم LM هر بار وزن‌ها در هر تکرار اصلاح شده و معکوس گیری انجام می‌شود. اما در LM اصلاح شده اندازه ماتریس در هر تکرار کوچکتر می‌شود و به همین خاطر زمان آموزش کم می‌شود. سپس نتایج آزمایش با معیارهای MSE^۲ و PSNR^۳ بررسی شدند. با نتایج بدست آمده می‌توان نتیجه گرفت این ساختار برای فشرده‌سازی تصویر نتایج مطلوبی را به همراه داشت. که می‌توان استنباط نمود که روش پیشنهاد شده خوب و نتایج فشرده‌سازی بالا است. علاوه بر نرخ فشرده‌سازی بالا کیفیت تصویر را حفظ نموده است.

۲-۲ فشرده‌سازی تصویر با استفاده از شبکه عصبی و موجک هار^۴

فشرده‌سازی تصویر مبتنی بر موجک بهبود قابل توجهی از کیفیت را در نرخهای فشرده‌سازی بالا فراهم می‌کند [۷]. فشرده سازی مبتنی بر تبدیل موجک هار یکی از روشهایی است که برای فشرده‌سازی تصویر می‌توان اعمال نمود. با فشرده‌سازی مبتنی بر تبدیل موجک، کیفیت تصاویر فشرده‌شده معمولاً بالاست و رسیدن به میزان فشرده‌سازی ایده آل مشکل است؛ چون بسته به محتوی عکس تغییر می‌کند [۷]. در [۷] پیشنهاد داده شده شبکه عصبی می‌تواند با ایجاد رابطه

^۱ Levenberg Marquardt

^۲ Mean Square Error

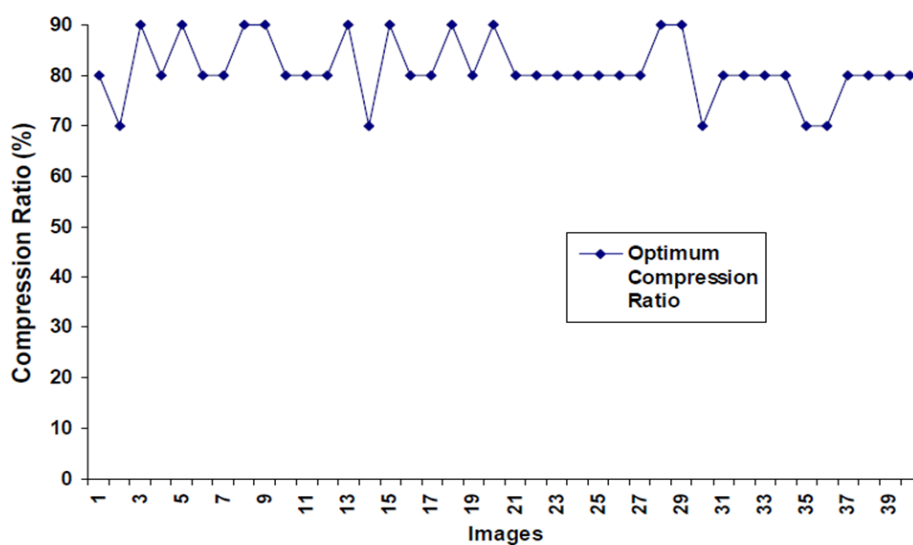
^۳ Peak Signal Noise Ratio

^۴ Haar

غیرخطی بین شدت (روشنایی) تصویر و نرخ‌های فشرده‌سازی در هر جستجو برای بهینه نمودن نرخ فشرده‌سازی آموزش دیده شود. هدف از کار ارائه شده در این تحقیق توسعه یک سیستم فشرده‌سازی تصویر بهینه با استفاده از تبدیل موجک هار و یک شبکه عصبی است. در روش مطرح شده پیشنهاد گردید که یک شبکه عصبی آموزش دیده می‌تواند رابطه‌ای غیرخطی بین شدت (مقادیر پیکسل) از یک تصویر و نرخ فشرده‌سازی بهینه آن یاد بگیرد. بر اساس این فرضیه می‌توان با ارائه یک تصویر به شبکه عصبی آموزش دیده نرخ فشرده‌سازی بهینه موجک هار را تشخیص داد. توسعه و پیاده‌سازی این سیستم از فشرده‌سازی تصویر با ۱۰۰ تصویر از موضوعات مختلف و با شدت روشنایی و کنتراست^۱ های متفاوت انجام شده است. ابتدا با استفاده از روش [۳۶] میزان بهینه فشرده‌سازی موجک هار برای ۷۰ تصویر بدست می‌آید. میزان فشرده‌سازی بهینه موجک هار برای ۳۰ تصویر دیگر نامعلوم است و این ۳۰ تصویر برای تست استفاده خواهد شد. از ۷۰ تصویر دارای میزان بهینه فشرده‌سازی موجک هار معین، ۴۰ تصویر برای آموزش شبکه عصبی و ۳۰ تصویر برای ارزیابی آموزش مورد استفاده قرار می‌گیرد.

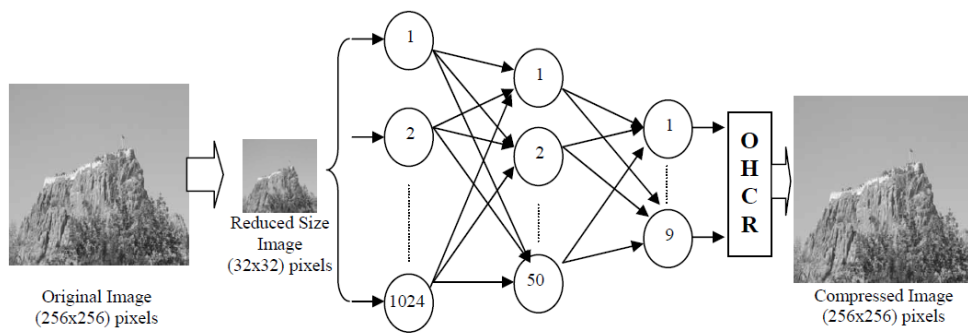
سیستم فشرده‌سازی بهینه تصویر با استفاده از یک شبکه عصبی با نظارت مبتنی بر الگوریتم یادگیری bp است. در شکل (۲-۴) مقادیر مختلف از میزان فشرده‌سازی بهینه برای مجموعه داده تصویر را مشاهده می‌کنید.

¹ Contrast

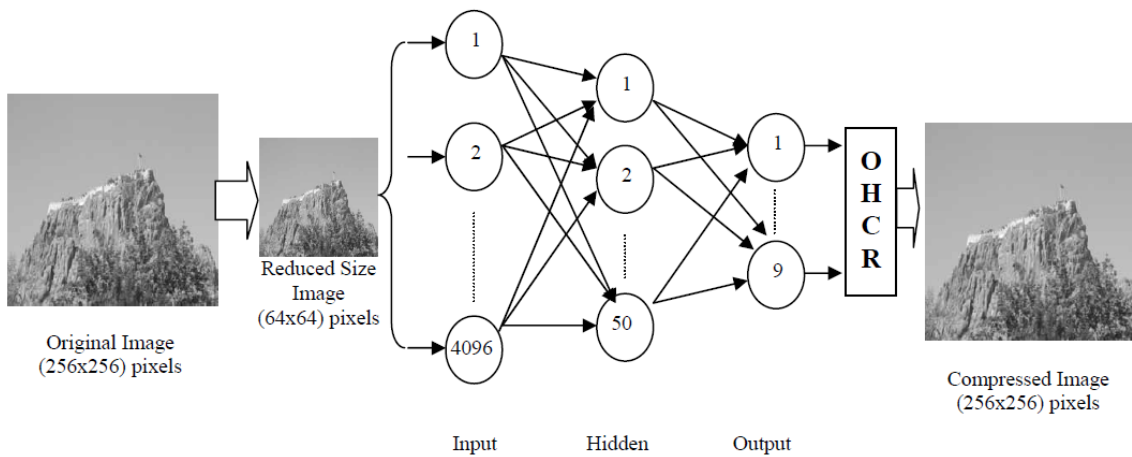


شکل ۲-۴) نسبت بین تصاویر و میزان فشرده‌سازی بهینه [۷]

در این تحقیق از دو شبکه عصبی با اندازه‌های متفاوت استفاده شده است که در شکل‌های (۲-۵ الف و ب) نشان داده شده است. شبکه عصبی نشان داده شده در شکل (۲-۵ ب) نسبت به شبکه عصبی نشان داده شده در شکل (۲-۵ الف) در تشخیص میزان فشرده‌سازی بهینه بهتر عمل کرده است. تست شبکه عصبی آموزش دیده ANN64 با استفاده از ۳۰ تصویر از مجموعه تست اول که قبلاً در شبکه ظاهر نشده بودند، نرخ تشخیص ۹۶/۶۷ درصد را نتیجه داد که در آن از ۳۰ تصویر با نرخ تراکم بهینه معلوم، ۲۹ تصویر به نرخ صحیح خود نسبت داده شدند. شبکه عصبی آموزش دیده ANN32 با استفاده از همان ۳۰ تصویر، نتیجه نرخ تشخیص ۹۳/۳ درصد را داشت، که در آن از ۳۰ تصویر با نرخ تراکم بهینه معلوم، ۲۸ تصویر به نرخ صحیح خود نسبت داده شده‌اند.



شکل ۲-۵ الف: سیستم فشرده ساز تصاویر بهینه با استفاده از ANN32 [۷]



شکل ۲-۵ ب: سیستم فشرده ساز تصاویر بهینه با استفاده از ANN64 [۷]

تصاویر ۶۴*۶۴ پیکسلی برای آموزش شبکه عصبی اول (ANN64) و تصاویر ۳۲*۳۲ پیکسلی برای آموزش شبکه عصبی دوم (ANN32) استفاده شده‌اند. تعداد نرون‌های مخفی در هر دو شبکه ۵۰، و تعداد نرون‌های خروجی ۹ عدد می‌باشد.

موجک‌ها عموماً به شکل ریاضی فشرده‌تری و با استفاده از توابعی ویژه به نام «موجک‌های هار»

ارائه می‌شود [۸].

۲-۳ رویکرد شبکه عصبی برای فشرده‌سازی تصویر رنگی در حوزه تبدیل

همچون بیشتر شبکه‌های عصبی مصنوعی، نقشه‌های خودسازمان دهنده^۱ در دو حالت کار می‌کنند: آموزش و نگاشت. “آموزش”، نگاشت را با استفاده از نمونه‌های ورودی ایجاد می‌کند در حالی که “نگاشت” یک بردار ورودی را به صورت خودکار دسته‌بندی می‌کند. نگاشتی که از یک فضای ورودی ابعاد بالا به فضای نگاشتی با ابعاد پایین را ایجاد می‌کند. اساس کار شبکه بر مبنای کم کردن فاصله خود از الگوهای ورودی است. در شبکه خود سازمانده از روش یادگیری رقابتی برای آموزش استفاده می‌شود. یادگیری رقابتی که در این قبیل شبکه‌ها بکار گرفته می‌شود بدین صورت است که در هر قدم یادگیری واحدها برای فعال شدن با یکدیگر به رقابت می‌پردازند، در پایان یک مرحله رقابت تنها یک واحد برنده می‌شود که وزن‌های آن نسبت به وزن‌های سایر واحدها به شکل متفاوتی تغییر داده می‌شود [۹].

در [۱۰] یک روش جدید برای فشرده‌سازی تصویر رنگی در حوزه تبدیل ارائه شده است. برای فشرده‌سازی تصویر از تکنیک کوانتیزیشن بردار (VQ)^۲ (بدین صورتکه VQ یک برداری را در یک فضای برداری می‌گیرد و به مجموعه‌ای از اندیس‌ها تبدیل می‌کند، این بردار می‌تواند هر نوع داده‌ای در یک فضای برداری باشد) استفاده شده است. و لغت نامه در VQ با استفاده از SOFM طراحی شده است. این کار از ویژگی‌های خاص SOFM برای تولید لغت نامه کلی است، که منجر می‌شود فقط یکبار نیاز به ساختن لغت نامه باشد. جدا از کوانتیزه فضایی، رویکرد لغت نامه‌ها عمومی در حوزه‌های تبدیل مانند DCT و DWT اجرا شده است. در تصاویر رنگی ثابت دیجیتال، هر مولفه از رنگ کوانتیزه شده با ۸ بیت، بطوری که هر رنگ با ۲۴ بیت مشخص شده است. که در نتیجه 2^{24} رنگ ممکن برای تصویر وجود دارد. در [۱۰] یک شماتیک برای طراحی تبدیل VQ برای فشرده‌سازی تصویر رنگی با استفاده از SOFM پیشنهاد داده شده، و فشرده‌سازی تصاویر رنگی با تبدیل تصاویر رنگی از RGB به فضای رنگی HSV انجام شده است. تصاویر می‌توانند با پیدا کردن خروجی کدگذاری شوند. هر بردار

¹ Self Organizing Maps(SOM)

² Vector Quantization

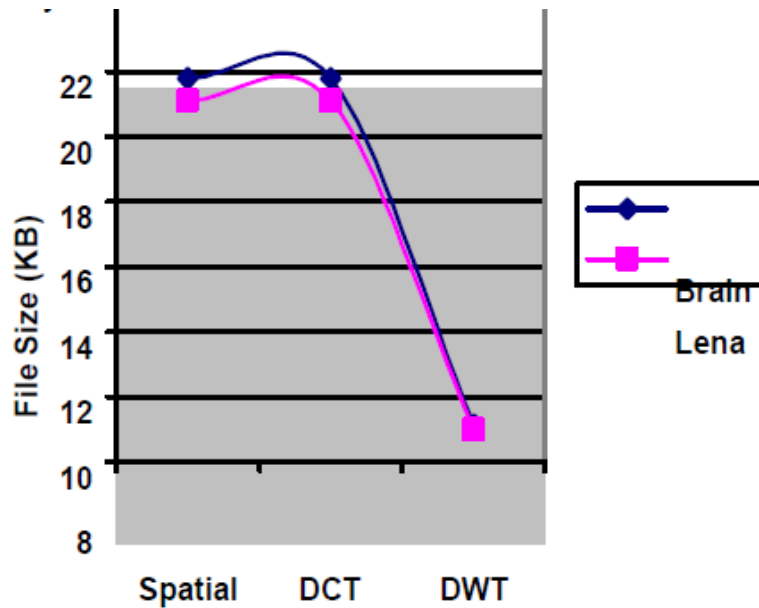
ویژگی تصویر با یافتن بردار کدی که کمترین فاصله اقلیدسی با آن را دارد کدگذاری می شود. البته تمام VQ ها موجب قطعه قطعه شدن تصویر بازسازی شده می گردند؛ یعنی مرز بین بلوکها قابل مشاهده خواهد بود. اگرچه تصویر حاصل PSNR خوبی دارد اما این اتفاق در کیفیت بصری اثر می گذارد. معمولاً برای مقابله با این اثر از کدگذاری تبدیل و / یا زیر باند استفاده می شود. هدف از فرآیند تبدیل ناهمبستگی پیکسلهای هر زیر تصویر، یا اطلاعات زیاد ممکن است به عدد کوچکتری از ضرایب منجر شود. از این رو تصویر خام باید قبل از کدگذاری با اسفاده از کوانتیزر به حوزه فرکانس تبدیل شود، همچنین در رمزگشایی باید عکس آن به حوزه فضایی (مکان) انجام شود. در [۱۰] از هر دو روش DCT و DWT برای آزمایش استفاده شده است. شبکه SOFM با رویکرد آموزشهای انتخابی که عملکرد بهتری در سطوح اشباع رنگ دارند آموزش دیده اند. ساخت لغت نامه عمومی میتواند با رمزگذاری برای هر تصویر با درستی قابل قبولی استفاده شود. برای ایجاد این لغت نامه تصاویری با ماهیتهای گوناگون از حیث جزئیات، کنتراست و بافت استفاده می شود.

سنجش کیفیت تصویر با استفاده از معیارهایی مانند PSNR، محتوی ساختمانی^۱، درستی تصویر^۲ و متوسط شاخص تشابه ساختاری تصاویر^۳ در فضای رنگی HSV انجام شده است. همچنین تجزیه و تحلیل کیفیت با استفاده از معیارهای کیفیت بصری مانند (MOS) انجام شده است. این بخاطر این است بعضی زمانها صرفنظر از کیفیت خوب تصویر، PSNR و معیارهای دیگر مقادیر نامطلوبی بدست می دهند. در نمودار (۲-۶) مقایسه بین اندازه فایل فشرده شده در حوزه فضایی و تبدیل نشان داده شده است.

¹ Structural content

² Image fidelity

³ Normalized Correlation Quality

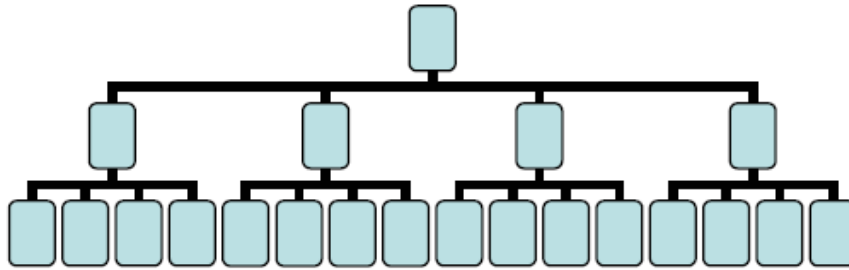


شکل ۲-۶) مقایسه بین اندازه فایل در حوزه فضایی و تبدیل [۱۰]

در شکل (۲-۶) فشرده‌سازی تصویر با DWT میزان فشرده‌سازی بهتری با کیفیت بصری قابل قبول را می‌دهد.

۲-۴ فشرده‌سازی تصاویر ثابت با استفاده از شبکه‌های عصبی

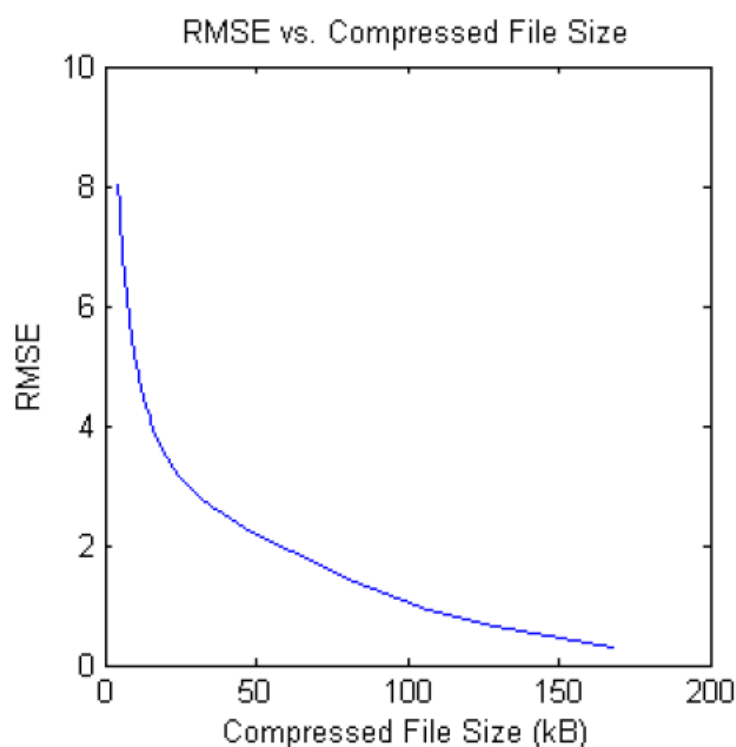
در [۱۱] از ترکیب تبدیل موجک گسسته (DWT) و شبکه‌های عصبی بعنوان روش پیشنهادی استفاده شده است. DWT یکی از فرآیندهای اساسی در الگوریتم فشرده‌سازی تصویر JPEG 2000 است. DWT تبدیلی است که میتواند بلوکی از داده را از حوزه مکان به حوزه فرکانس نگاشت نماید. DWT یک تصویر را در ۴ بلوک تجزیه می‌نماید: ضرایب تقریب و ۳ تا از ضرایب جزئیات. جزئیات شامل ضرایب افقی، عمودی و مورب می‌باشد. بخش فرکانس پایین از تصویر می‌تواند حفظ شود در حالیکه بخش فرکانس بالا (جزئیات) می‌تواند با اتلاف بیشتر و بدون اینکه افت کیفیت خیلی نمایان بشود تقریب زده شود. در شکل (۲-۷) یک نمونه از تجزیه موجک را نشان می‌دهد.



شکل ۷-۲) تجزیه یک تصویر به تقریب و جزئیات با استفاده از موجک [۱۱]

تبدیل DWT را میتوان یک بار برای تصویر و سپس برای ضرایب تبدیل DWT تولید شده اعمال نمود. می‌توان بعنوان یک ساختار درختی معکوس که تصویر اصلی در بالا قرار گرفته تصور نمود. تبدیل فوریه DWT سطح اول تصویر به چهار بخش یا شاخه تجزیه شده، که از هر ۴ قسمت جدید می‌توان تبدیل DWT اعمال شده به آن‌ها را به صورت جداگانه تقسیم بر بخش‌ها یا شاخه‌های مجزا کرد که این روش عموماً بعنوان تبدیل بسته موجک شناخته شده است.

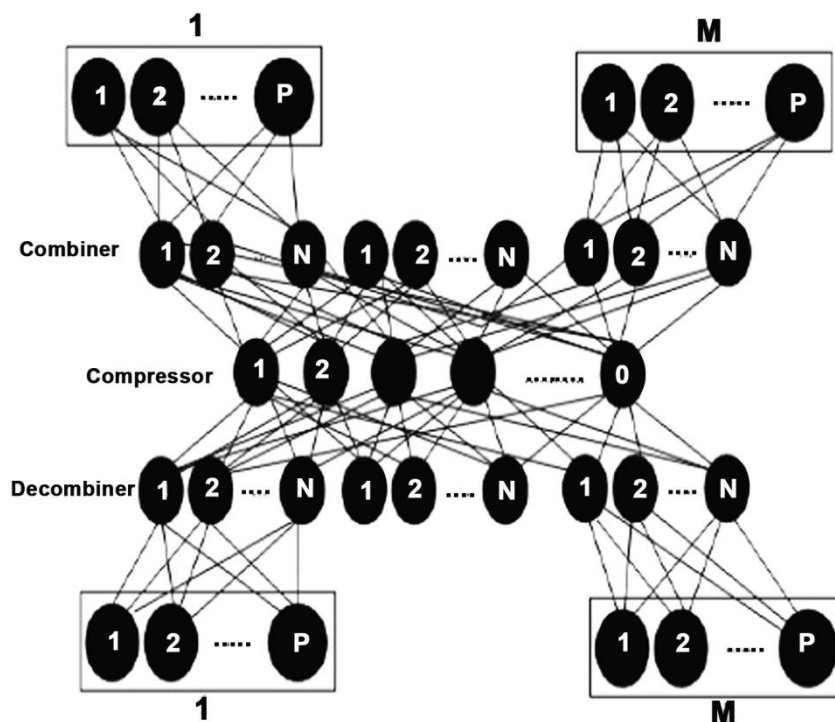
نوع شبکه عصبی استفاده شده یک شبکه دو لایه رو به جلو با یک تابع یادگیری پس‌انتشار استاندارد بوده است. لایه اول می‌تواند شامل یک تعداد از نرون‌های متغیر با روش فشرده‌سازی انتخاب شده باشد. این لایه ورودی از هر نرون یک تابع انتقال سیگموئید تانژانت هیپربولیک دارد. مقدار خروجی هر نرون باید بین -1 و 1 باشد. برای ذخیره کردن یک شبکه عصبی تنها دو چیز مورد نیاز است برای هر نرون، وزن‌ها و بایاس. در روش مطرح شده برای پویس الگو تنها یک ورودی وجود دارد بطوریکه تعداد وزن‌ها دو برابر تعداد نرون‌های لایه ورودی می‌باشند. دو برابر شدن وزن‌ها بدلیل وزن‌های بین لایه اول نرون‌ها و لایه خروجی است. با توجه به نتایجی که گرفته شد بجز از لحاظ مقدار زمان آموزش عملکرد سیستم مطلوب بوده است. همچنین نسبت فشرده‌سازی با در نظر گرفتن تعداد کل پیکسل در تصویر نسبت به اندازه فایل فشرده‌شده در بایت محاسبه شده است. در نمودار (۲-۸) با افزایش کیفیت، RMSE (نشان‌دهنده انحراف معیار از تفاضل بین مقادیر پیش‌بینی شده و مشاهده شده است) نیز کاهش پیدا می‌کند.



شکل ۲-۸) عملکرد سیستم با افزایش کیفیت [۱۱]

۲-۵ فشرده‌سازی تصویر با روش ترکیبی

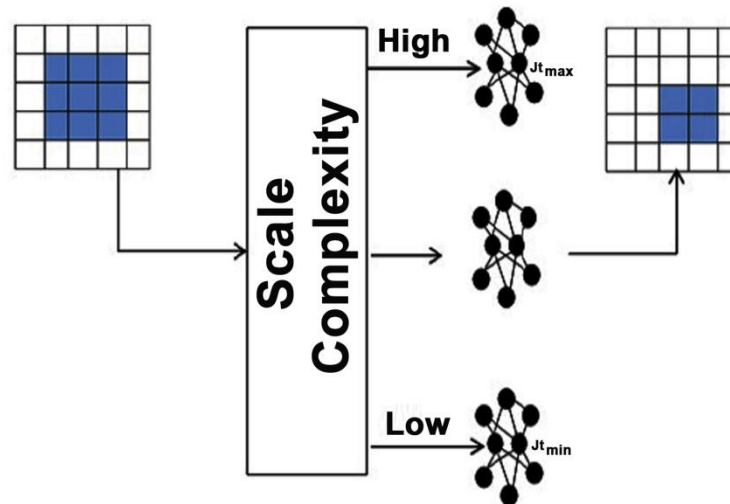
مرجع [۱۲] یک مطالعه تجربی از برخی روشهای فشرده سازی تصویر مبتنی بر شبکه های عصبی، بخصوص الگوریتم مبتنی بر پس انتشار نزولی خطا انجام شده است. همچنین یک روش ترکیبی جدید مبتنی بر پرسپترون چند لایه که ترکیبی از طرحهای سلسله مراتبی و تطبیقی می باشد ارائه شده است. ایده این است که الگوریتم پس انتشار بر روی یک شبکه عصبی تطبیقی یافته که از ریز شبکه های عصبی با ساختار سلسله مراتبی استفاده می کند و به صورت موازی انجام می شود. ساختار شبکه عصبی بر اساس پس انتشار را می توان برای ایجاد نمودن یک شبکه سلسله مراتبی با اضافه نمودن بیش از دو لایه مخفی تعمیم داد [۱۳ و ۱۴]. در شکل (۲-۹) ساختار شبکه سلسله مراتبی نشان داده شده است.



شکل ۲-۹) ساختار شبکه عصبی سلسله مراتبی [۱۲]

اساس این ایده تقسیم کردن تصویر ورودی به M زیرتصویر غیرمشترک و هر زیرتصویر نیز به T بلوک با اندازه‌های $P \times P$ تقسیم بندی شوند. بعد از ایجاد نمودن ساختار شبکه، الگوریتم پس انتشار اعمال می‌شود ولی بخاطر ساختار لایه مخفی زمان مرحله آموزش بیش از حد طولانی می‌شود. بدین منظور [۱۵] پیشنهاد می‌کند که از شبکه‌های عصبی متفاوتی برای فشرده‌سازی تصاویر با سطوح پیچیدگی متفاوت استفاده شود. با استفاده از یک پس انتشار تطبیقی برای آموزش شبکه عصبی که برای فشرده‌سازی تصویر شامل طبقه بندی در تصویر ورودی است می‌توان bp را به زیر مجموعه‌های مختلف اعمال نمود. زمان بازسازی با استفاده از این الگوریتم کاهش یافته است. فرق اساسی بین شبکه عصبی سلسله مراتبی و شبکه عصبی تطبیقی در این است که در شبکه عصبی سلسله مراتبی از سه لایه مخفی برای فشرده‌سازی استفاده می‌شود اما شبکه عصبی تطبیقی با محاسبه نویز بین

بلوک‌های نزدیک به هم از تصویر ورودی فشرده‌سازی را ممکن می‌کند. شکل (۲-۱۰) ساختار شبکه عصبی تطبیقی را نشان می‌دهد.

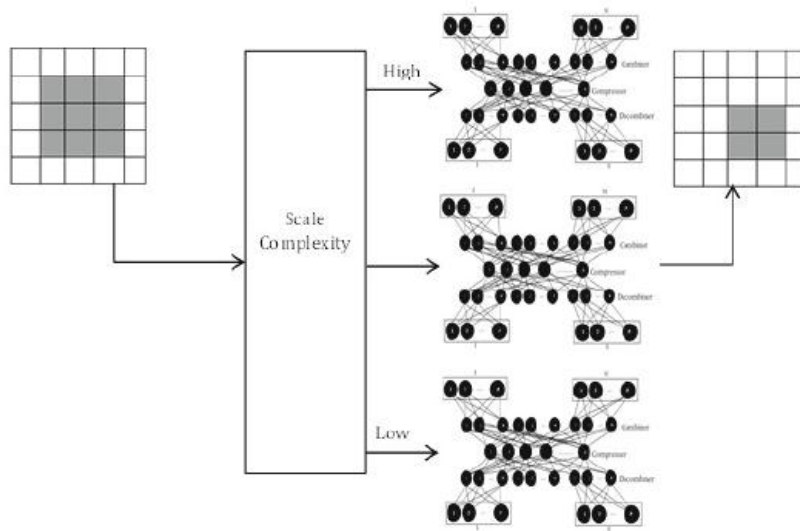


شکل (۲-۱۰) ساختار شبکه عصبی تطبیقی [۱۲]

ساختار شبکه عصبی تطبیقی بدین گونه که تعدادی از ترکیبات بر مبنای شبکه‌های عصبی متفاوت که برای فشرده‌سازی بلوک‌های تصویر با سطوح پیچیدگی مختلف استفاده می‌شود. سپس N زیرشبکه عصبی آموزش داده شده و تصویر ورودی به N زیرمجموعه که هر زیرمجموعه مطابق با یک سطح پیچیدگی است تقسیم می‌شود. شکل (۲-۱۰) شبکه‌های عصبی با تعداد فزاینده‌ای از نرون‌های مخفی طراحی شده است. استفاده از شبکه عصبی تطبیقی در واقع طبقه‌بندی روی تصاویر ورودی است.

با بهره گرفتن از مزایای هر دو شبکه، یک روشی بنام هیبرید (ترکیبی) معرفی شده است تا با استفاده از آن بهینه ترین نتیجه ممکن بدست آید. ایده این است که یک شبکه پرسپترون چند لایه ایجاد شود که ساختار یک شبکه عصبی تطبیقی را دارا باشد اما در عین حال از زیرشبکه‌هایی با ساختار سلسله مراتبی برای طبقه‌بندی تصاویر ورودی بر اساس سطح پیچیدگی‌شان بهره گیرد. بنابراین تصویر ورودی به N زیرمجموعه تقسیم خواهد شد که با N زیرشبکه متناظر هستند. هر

زیرمجموعه با زیرشبکه‌ی متناظرش که ساختار سلسله مراتبی دارد آموزش داده خواهد شد. در شکل (۱۱-۲) ساختار شبکه هیبریدی نشان داده شده است.



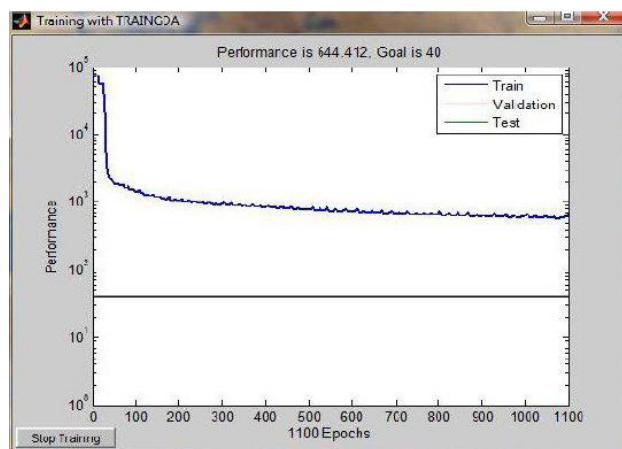
شکل (۱۱-۲) ساختار شبکه عصبی هیبریدی [۱۲]

قاعده کلی روش هیبرید شبیه همان bp تطبیقی است، اما زیر شبکه ها برای طبقه بندی بلوک های تصویر در ورودی با توجه به سطوح پیچیدگیشان استفاده شدند، که ساختاری از یک شبکه عصبی سلسله مراتبی است. پس از تکمیل نمودن طبقه بندی، هر شبکه عصبی از طریق الگوریتم bp آموزش می بیند. استفاده از موازی کاری برای محاسبه روش هیبرید مطرح شده منجر به کاهش تقریباً نیمی از زمان های محاسبات (آموزش، فشرده سازی و بازسازی) می شود. با بهره گیری از مزایای هر دو روش پس انتشار سلسله مراتبی و پس انتشار تطبیقی تکنیکهای فعلی مبتنی بر پس انتشار که خودشان به تنهایی هم نتایج قابل قبولی را ارائه می کنند را بیشتر بهبود می دهند. بلوک های تصویر آموزش برای فشرده سازی و بازسازی که در شبکه خطی بکار برده می شود بصورت رابطه (۱-۲) است.

$$\sum_{i=1}^N w_{ji} x_i \quad 1 \leq j \leq k \quad (1-2)$$

۲-۶ فشرده سازی تصویر با استفاده از شبکه عصبی پس انتشار

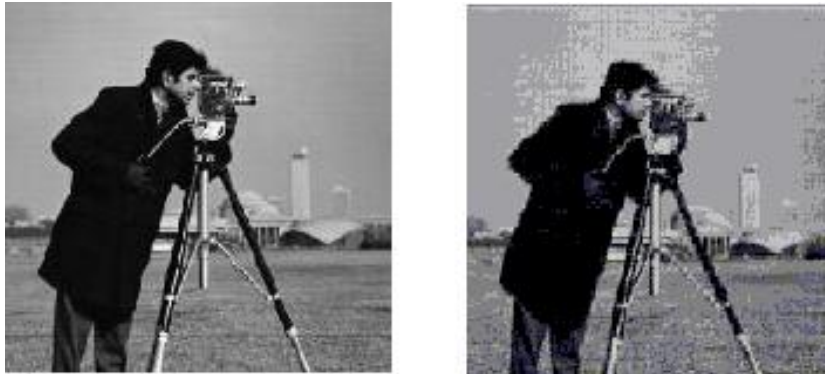
در [۱۶] از الگوریتم آموزش شبکه عصبی پس انتشار استفاده شده است. الگوریتم پس انتشار شبکه عصبی کارآیی سیستم را افزایش و زمان همگرایی آموزش شبکه را کاهش می‌دهد. تکنیک جدیدی در این مقاله اتخاذ شده که مقداردهی وزن‌های بین نرون‌های لایه ورودی و خروجی بجای وزن‌های تصادفی، و تصویر هم به بردار یک بعدی تبدیل و نرمالیزه (۱ و ۰) می‌شود که نرخ سریع همگرایی الگوریتم آموزش را برای تعدادی از تصاویر آزمایش شده نشان می‌دهد. کیفیت تصویر فشرده شده با تعدادی از پارامترهای معین سنجش، و با تکنیک‌های مختلف فشرده سازی مقایسه شده است. بطور کلی هرچه PSNR (یکی از پارامترهای ارزیابی می‌باشد) بزرگتر باشد کیفیت تصویر بهتر است. نمودار نشان داده شده در شکل (۲-۱۲) خروجی آموزش شبکه و ۱۱۰۰ EPOCHS گرفته شده از شبکه با استفاده از تابع آموزش را نشان می‌دهد.



شکل (۲-۱۲) خروجی آموزش شبکه [۱۶]

پیاده سازی الگوریتم پس انتشار شبکه عصبی در سیستم فشرده سازی تصویر با عملکرد خوبی نشان داده شده است. شبکه عصبی پس انتشار آموزش دیده و تست شده برای تجزیه و تحلیل تصاویر مختلف، که مشاهده شده زمان همگرایی برای آموزش شبکه عصبی پس انتشار خیلی سریع است. این روش مقداردهی وزن‌ها نرخ سریع همگرایی را به همراه داشته و از مجموع وزن آموزش دیده کیفیت

خوب تصاویر بازسازی شده در پایان در دسترس هستند. با افزایش تعداد تکرارهای^۱ فرآیند آموزش از ۱۱۰۰ به ۲۲۰۰ نرخ فشرده سازی از ۰.۹۹ به ۰.۹۵۵۶ کاهش یافت و PSNR از ۱۹.۳۱۸db به ۲۰.۷۲۲db افزایش یافت. دلیل آن زمان بیشتری است که شبکه برای تنظیم وزنهایش در اختیار داشته است. در شکل (۲-۱۳) تصویر اصلی و بازسازی شده cameraman آمده است.



شکل (۲-۱۳) تصویر اصلی و بازسازی شده با ۱۱۰۰ epochs [۱۶]

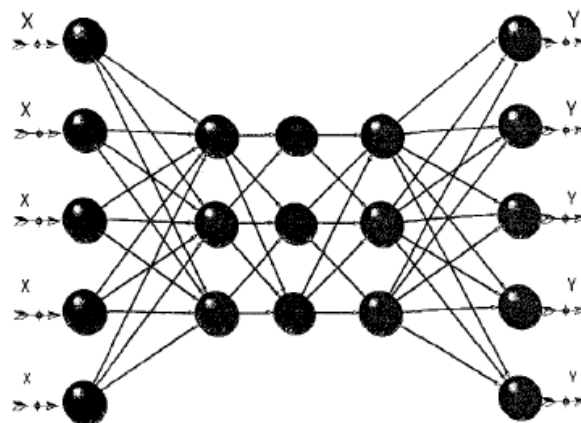
همانطور که در این کار نشان داده شده است، شبکه‌های عصبی پس انتشار می‌توانند با موفقیت برای پردازش تصویر اجرا شوند.

۷-۲ فشرده سازی تصویر با استفاده از شبکه عصبی چند لایه

در [۱۷] از یک شبکه عصبی جدید برای فشرده‌سازی داده استفاده شده است. شبکه‌های دو لایه به شبکه‌های چند لایه توسعه داده شده است. نتایج نشان از برتری شبکه‌های چند لایه نسبت به شبکه دو لایه، بخصوص در نرخ فشرده‌سازی بالا دارد. همچنین یک خطای فیدبک اصلاح شده برای کاهش زمان آموزش و افزایش کیفیت تصویر مطرح شده است. همچنین یک بازتوزیع مجدد از سطوح خاکستری در مرحله آموزش برای حداقل رساندن خطای میانگین مربع سیستم دید انسان پیشنهاد شده است. شبکه‌های عصبی به دلیل معماری موازی‌شان و انعطاف‌پذیری برای فشرده‌سازی تصویر خود را ثابت نمودند. یکی از روشهایی که موفقیت آمیز بوده فشرده‌سازی تصویر مبتنی بر تبدیل که

^۱ Epochs

رمزگذاری تبدیل Autoassociative نامیده می‌شود و فقط در شبکه‌های دو لایه محدود می‌شود. در [۱۷] این تکنیک برای شبکه‌های چند لایه تعمیم داده شده است. که نرخ فشرده‌سازی بالا با همان کیفیت را نشان می‌دهد. مراجع [۱۸ و ۱۹] یک مدل شبکه عصبی دو لایه با تعدادی از نرون‌ها در لایه اول و تعداد کمتر در لایه دوم پیشنهاد داده‌اند. این مدل با مجموعه بردار 0 بعنوان ورودی و خروجی در همان زمان با استفاده از تکنیکهای آموزش شبکه عصبی با نظارت استاندارد مانند روش پس‌انتشار خطا آموزش دیده است. مرجع [۲۰] از فشرده‌ساز کسکود که نرخ فشرده‌سازی بالا و کیفیت خوب گزارش شده است استفاده نموده است. مرجع [۲۱] الگوریتم آموزش سریع برای این مدل [۲۰] را توسعه دادند. در شکل (۲-۱۴) شبکه رو به جلو چهار لایه برای فشرده‌سازی نشان داده شده است.



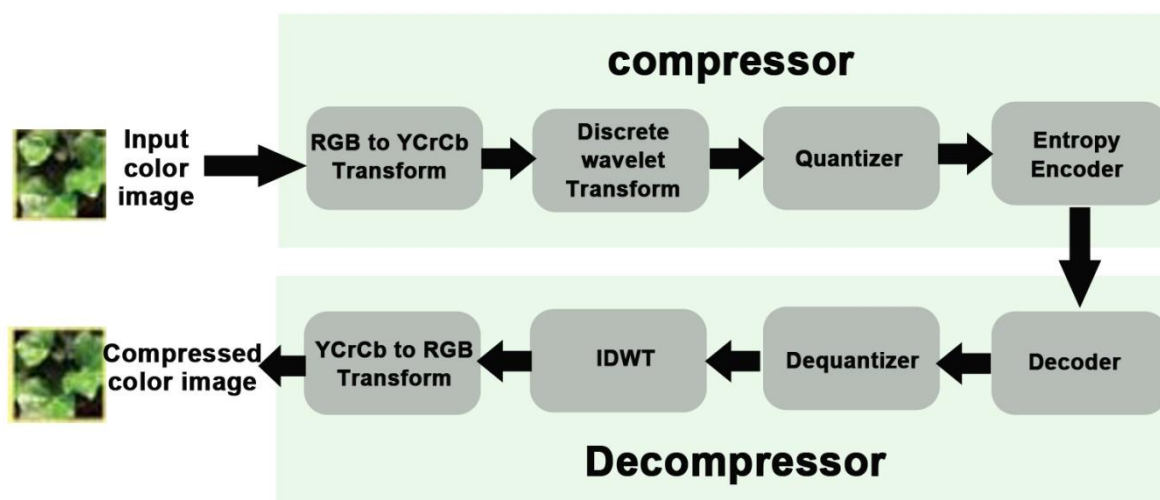
شکل (۲-۱۴) شبکه رو به جلو چهار لایه برای فشرده‌سازی [۱۷]

الگوریتم معرفی شده بطور همزمان در چهار لایه آموزش می‌بیند. آموزش شبکه عصبی چهار لایه پردازش سنگین و زمان زیادی را صرف میکند. در این الگوریتم خطای میانگین مربع بین خروجی مطلوب و واقعی با حفظ ورودی‌های غیر خطی به حداقل رسیده است. این در تضاد است با الگوریتم استاندارد که خطای میانگین مربع بین خروجی‌های مطلوب و واقعی با حفظ وزن‌ها به حداقل رسیده است. برای چیره شدن بر زمان طولانی آموزش برای شبکه‌های عصبی چند لایه اخیراً الگوریتم آموزش بکاربرده شده را توسعه دادند. با پیشرفت فعلی در اجرای شبکه عصبی (سخت افزار یا نرم

افزار)، افزایش محاسبات مشکل جدی نیست. این افزایش نتایج در پیچیدگی محاسبات هنوز کمتر از دیگر روش های مرسوم مانند JPEG است. مرجع [۲۲] الگوریتم جدیدی که سریعتر و پایداری بیشتری دارند را توسعه دادند.

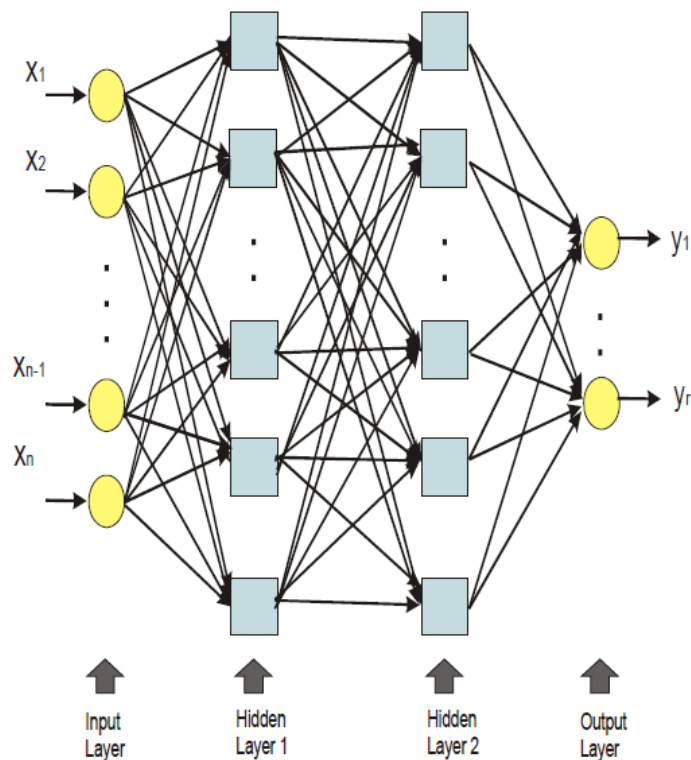
۸-۲ شبکه عصبی برای انتخاب بهترین موجک در فشرده سازی تصویر رنگی

بر خلاف روش فشرده سازی مبتنی بر تبدیل موجک مرسوم که انتخاب بصورت دستی انجام می شود، در [۲۳] از یک شبکه عصبی برای انتخاب موجک استفاده می شود. یک شبکه عصبی بعنوان یک روش برای حل این مسئله بجای انتخاب دستی تحت کدک مبتنی بر موجک مرسوم ارائه شده است. نتایج نشان می دهد شبکه عصبی مبتنی بر ویژگی های تصویر می تواند بعنوان یک راه حل برای حل مسئله استفاده شود. متغیرهای ورودی برای شبکه عصبی، دو ویژگی تصویر است، گرادیان تصویر و فرکانس فضایی (SF) از سه مولفه رنگ (قرمز، سبز، آبی) و خروجی شبکه عصبی نوع موجک است. اهمیت اصلی این مقاله فشرده سازی تصاویر رنگی است. این مقاله یک فشرده سازی تصویر رنگی مبتنی بر موجک است که ارائه شده و سیستم از سه قسمت اساسی تشکیل شده، تبدیل موجک گسسته (DWT) ضرایب تبدیل شده از کوانتیزه و آنتروپی ضرایب کوانتیزه شده را رمزگذاری می کند. DWT برای تبدیل در فضای رنگ (YCrCb) بکار برده می شود. شکل [۲-۱۵] یک نوع رمزگذاری تصویر رنگی مبتنی بر موجک است.



شکل ۲-۱۵) بلوک دیاگرام کدگذاری تصویر مبتنی بر موجک [۲۳]

هدف اصلی این مقاله انتخاب بهترین موجک با استفاده از شبکه عصبی بعنوان یک روش برای حل این مسئله بطور خودکار بجای انتخاب دستی است. در این مقاله برای حل مسئله از MLP استفاده شده است. بعنوان مثال شکل [۲-۱۶] MLP با دو لایه مخفی را نشان می‌دهد. دایره‌ها ورودی و خروجی نرونها را نشان می‌دهد و مربع نشان دهنده نرون‌ها در لایه مخفی است. تابع سیگموئید معمولاً در MLP استفاده می‌شود. با استفاده از شبکه عصبی و طی فرآیندی دو مرحله‌ای مناسب‌ترین موجک انتخاب می‌گردد. مرحله اول فرآیند آموزش و مرحله دوم فرآیند تست است. فرآیند آموزش یک MLP معمولاً با استفاده از الگوریتم BP انجام می‌شود.



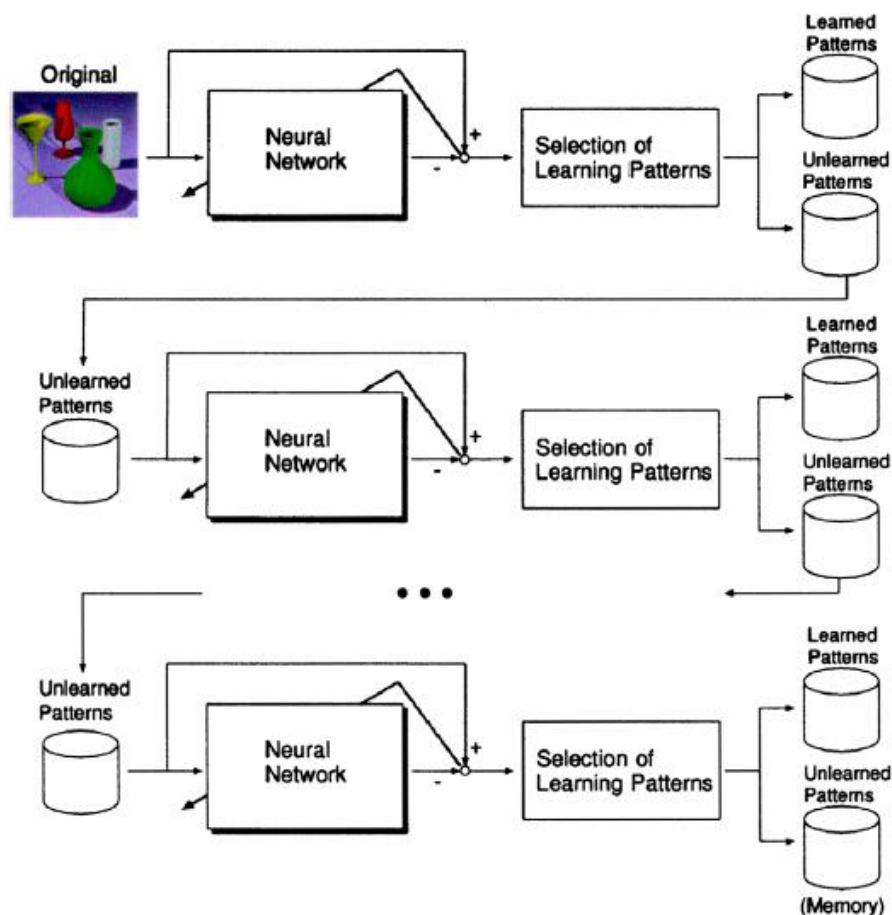
شکل ۲-۱۶) توپولوژی MLP با دو لایه مخفی [۲۳]

انتخاب بهترین موجک برای فشرده‌سازی تصویر بر اساس ویژگی‌ها و مقدار بالاترین PSNR بعنوان سنجش عملکرد MLP انجام شده است. توپولوژی‌های مختلف MLP برای حل مسئله بکار برده شده است. شبکه چهار لایه بعنوان بهترین توپولوژی از MLP برای این مسئله انتخاب شده است. در پایان نشان داده شده که استفاده از یک سیستم فشرده‌ساز مبتنی بر موجک تطبیقی برای انتخاب مناسب‌ترین موجک مبتنی بر ویژگی‌های تصویر امکان‌پذیر است [۲۳].

۹-۲ یک روش فشرده‌سازی برای تصاویر رنگی با استفاده از شبکه‌های عصبی چند لایه

در [۲۴] یک روش فشرده‌سازی تصاویر رنگی مبتنی بر یک گروه از شبکه‌های عصبی چند لایه که می‌تواند خطای بازسازی را قدم به قدم کاهش دهد پیشنهاد شده است. بسپازی از تکنیک‌های فشرده‌سازی تصویر به منظور کاهش نیازهای انتقال توسعه داده شده است [۲۵]. ساختار این مدل از شبکه عصبی شامل چندین شبکه عصبی با اندازه بلوک‌های مختلف برای تقسیم‌بندی هر بخش است.

با تقسیم‌بندی هر بخش، هر شبکه عصبی به یک بخش مانند لبه یا منطقه اختصاص داده می‌شود. از نتایج بدست آمده از این روش، عملکرد فشرده‌سازی بهتر از روش با استفاده از تک شبکه عصبی نشان داده شده است. در شکل (۲-۱۷) روش فشرده‌سازی تصویر پیشنهاد شده نشان داده شده است.

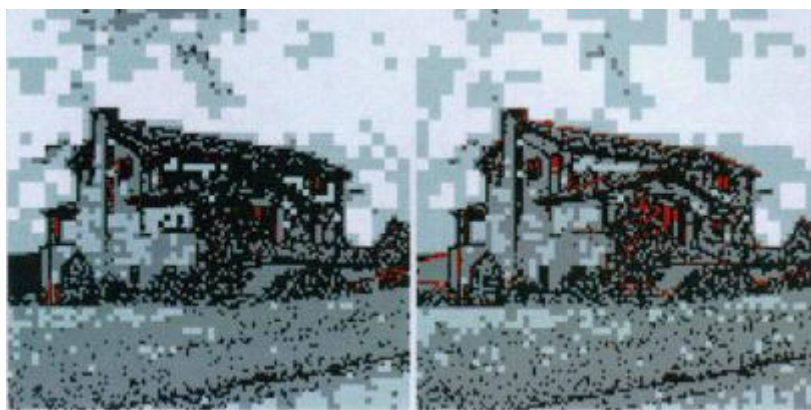


شکل (۲-۱۷) روش فشرده‌سازی تصویر پیشنهاد شده [۲۴]

شکل (۲-۱۷) روش فشرده‌سازی براساس یک شبکه عصبی با ساختار مدولار را معرفی می‌کند، در این روش از چندین شبکه عصبی با اندازه‌های بلوک مختلف برای تقسیم‌بندی ناحیه هدف استفاده شده است. بطور واضحی تعداد واحدهای ورودی، مخفی و خروجی در هر شبکه عصبی با یکدیگر متفاوت است. در این مقاله [۲۴] تجزیه‌سازی کار قطعه بندی ناحیه را حل می‌کند. فرض می‌کنیم تصویر هدف از مناطق تخت تشکیل شده است، سپس تصویر هدف با شبکه عصبی با بزرگترین اندازه

بلوک فشرده می‌شود. سپس بلوک با بیشترین خطای یادگیری از تصویر هدف حذف می‌شود. با تکرار فرآیندهای انتخاب برای هر بلوک، تقسیم‌بندی مناطق می‌تواند براساس توانایی یادگیری هر شبکه عصبی اجرا شود.

با در نظر گرفتن سرعت و دقت فشرده‌سازی از یک شبکه عصبی و روش پیشنهاد شده دیده می‌شود روش پیشنهاد شده می‌تواند سرعت بازسازی نرخ فشرده‌سازی را نسبت به یک شبکه عصبی بهبود ببخشد. توزیع بلوکها در روش فشرده‌سازی مطرح شده در شکل (۲-۱۸) نشان داده شده است. دیده می‌شود مناطق سفید در اندازه بلوک بزرگ فشرده شده و مناطق سیاه در اندازه بلوک کوچک فشرده شده‌اند.



شکل ۲-۱۸) توزیع بلوکها در روش پیشنهاد شده [۲۴]

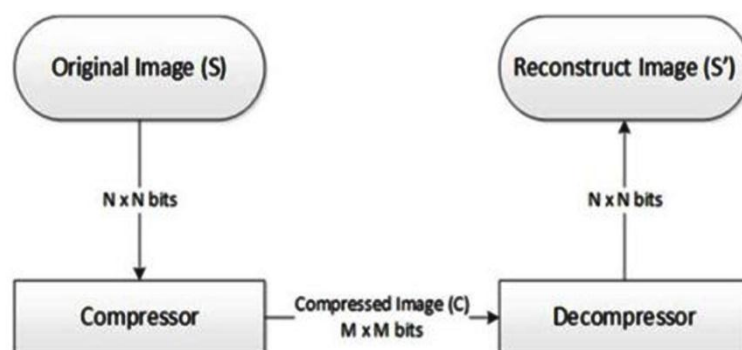
در روش فشرده‌سازی پیشنهاد شده مناطق تخت با اندازه بلوک بزرگتر و مناطق لبه با اندازه بلوک کوچکتر فشرده شوند.

۲-۱۰ فشرده‌سازی تصویر و اصلاح خطای نویز کانال با استفاده از شبکه عصبی

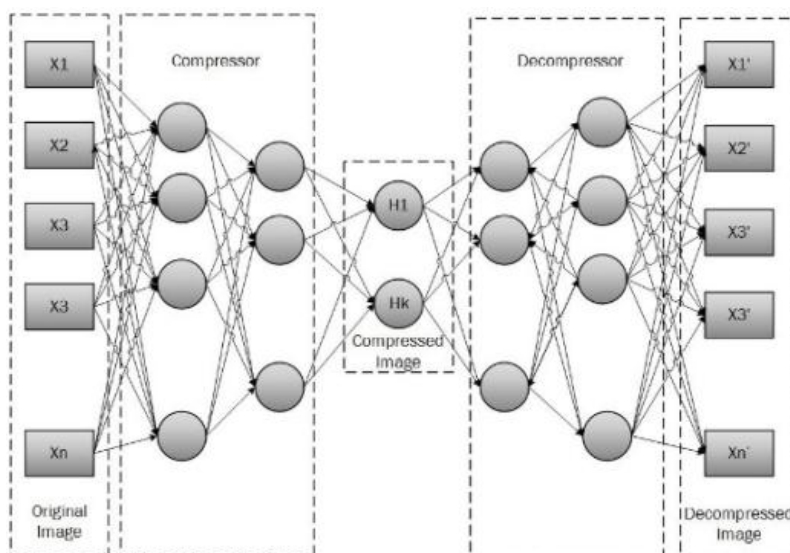
عمیق

مرجع [۲۶] که بعنوان مرجع این پایان‌نامه استفاده شده است از یک شبکه عصبی عمیق برای فشرده‌سازی تصاویر خاکستری و همچنین اصلاح خطای نویز کانال استفاده نموده است. جزئیات شبکه عصبی عمیق در فصل سوم این پایان‌نامه تشریح می‌گردد. نتایج تجربی تصاویر بازسازی شده

کیفیت بهتری را نشان می‌دهد. در شکل (۲-۱۹) بلوک دیاگرام سیستم فشرده‌سازی تصویر روش مذکور نشان داده شده است. که این سیستم فشرده‌ساز شامل واحد فشرده‌سازی و بازسازی می‌باشد. همچنین ساختار شبکه عصبی عمیق طراحی شده برای حل مسئله فشرده‌سازی تصویر در شکل (۲-۲۰) نمایش داده شده است. در نتایج بدست آمده با توجه به استفاده تعداد لایه‌های بیشتر در معماری روش ارائه شده، نشان از برتری این روش نسبت به روش‌های مقایسه شده را دارد.



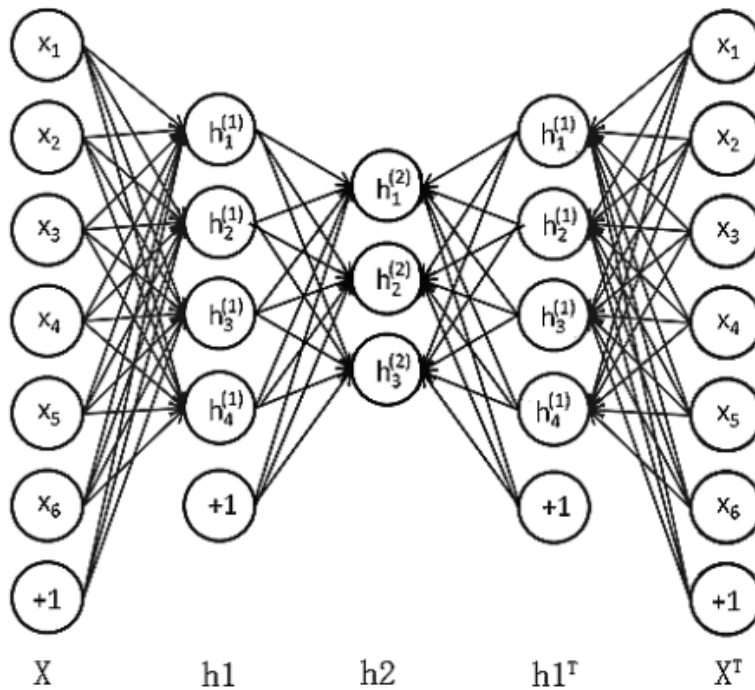
شکل (۲-۱۹) بلوک دیاگرام سیستم فشرده سازی در [۲۶]



شکل (۲-۲۰) ساختار شبکه عصبی عمیق [۲۶]

۱۱-۲ رمزگذاری و فشرده‌سازی تصویر مبتنی بر یادگیری عمیق

مرجع [۲۷] از خودرمزگذار انباشته^۱ که نوعی از الگوریتم یادگیری عمیق بدون نظارت است برای فشرده‌سازی تصویر استفاده شده است. در این تحقیق نتایج بدست آمده از کاربرد فشرده‌سازی و رمزگذاری تصویر با استفاده از SAE را به خوبی نشان می‌دهد که عملی و موثر می‌باشد که بطور همزمان برای انتقال و حفاظت تصویر در اینترنت استفاده می‌شود. در شکل (۲-۲۱) معماری روش مطرح شده نشان داده شده است. این روش جدید با چند مجموعه استاندارد تصاویر ارزیابی شده است. با توجه به نتایج بدست آمده یکی از معایب روش مقدار بالای MSE می‌باشد که ممکن است منجر به کاهش PSNR شود.



شکل (۲-۲۱) معماری [۲۷] Stacked Auto-Encoder

¹ Stacked Auto-Encoder

هدف این کار که فشرده‌سازی و بازسازی تصاویر می باشد توانسته است با SAE (خودرمزنگار انباشته) این عمل را انجام دهد. یادگیری عمیق این امکان را بوجود آورده است هر چه شبکه عمیقتر باشد قادر است اطلاعات بیشتری از داده‌های آموزشی را در خود ذخیره کند.

فصل سوم

مباحث تئوری

۳-۱ پردازش تصویر

پردازش تصویر روشی برای تبدیل یک تصویر به صورت دیجیتال و انجام برخی از عملیات بر روی آن، به منظور دریافت یک تصویر بهبود یافته و یا برای استخراج برخی از اطلاعات مفید از آن است. پردازش تصویر امروزه در حوزه‌های مختلفی کاربرد داشته و عضوی جدایی‌ناپذیر از تحقیق و پژوهش در این حوزه‌ها به شمار می‌آید. اما با این حال می‌توان زمینه‌های تحقیق در پردازش تصویر را با توجه به نوع نگاه به تصویر دیجیتال، به دودسته کلی تقسیم‌بندی کرد: بهبود تصاویر و بینایی ماشین. اگرچه در بسیاری از کاربردهای مختلف از دودسته فوق، عملیات تقریباً مشابهی بر روی تصویر دیجیتال ورودی صورت می‌پذیرد، اما علت دسته‌بندی را می‌توان در هدف متفاوت کاربردهای هر یک از این دودسته عنوان کرد. در بهبود تصاویر، هدف تغییر در شدت روشنایی پیکسل‌های تصویر است، به‌گونه‌ای که تصویر بهبودیافته (خروجی) از نقطه‌نظر خاصی کیفیت بهتری نسبت به تصویر ورودی داشته باشد. اما در کاربردهای مرتبط با بینایی ماشین، عموماً هدف اصلی درک هوشمند محتوای تصویر و استخراج اطلاعات موردنیاز از تصویر است. اگرچه بهبود تصاویر می‌تواند به‌عنوان یک پیش‌پردازش برای کاربردهای بینایی ماشین در نظر گرفته شود.

۳-۲ شبکه عصبی

عبارت 'شبکه عصبی' در حالت کلی به دو مفهوم مختلف اشاره دارد:

الف) شبکه عصبی زیستی

ب) شبکه عصبی مصنوعی

یک شبکه عصبی زیستی از مجموعه یا مجموعه‌هایی از نورون‌های به هم متصل یا از لحاظ عملکردی به هم وابسته تشکیل شده‌است. هر نورون می‌تواند به تعداد بسیار زیادی از نورون‌ها وصل باشد و تعداد کل نورون‌ها و اتصالات بین آن‌ها می‌تواند بسیار زیاد باشد. شبکه عصبی مصنوعی: در این شبکه‌ها به کمک دانش برنامه نویسی، ساختار داده‌ای طراحی می‌شود که می‌تواند همانند نورون

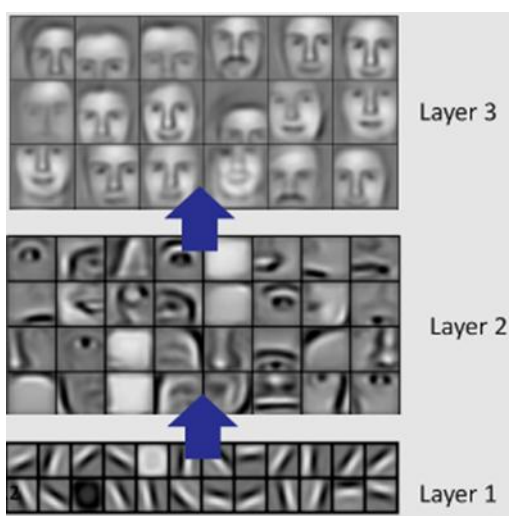
عمل کند. به این ساختار داده گره گفته می‌شود. سپس با ایجاد شبکه‌ای بین این گره‌ها و اعمال یک الگوریتم آموزشی به آن، شبکه را آموزش می‌دهند. در این حافظه یا شبکه عصبی گره‌ها دارای دو حالت فعال (روشن یا ۱) و غیرفعال (خاموش یا ۰) هستند و هر یال (سیناپس یا ارتباط بین گره‌ها) دارای یک وزن می‌باشد. یال‌های با وزن مثبت، موجب تحریک یا فعال کردن گره غیر فعال بعدی می‌شوند و یال‌های با وزن منفی، گره متصل بعدی را غیر فعال یا مهار (در صورتی که فعال بوده باشد) می‌کنند. شبکه‌های عصبی مصنوعی (ANN) در واقع از ساختار درهم و توده‌ای مغز پستانداران الهام گرفته شده است، که در آن میلیون‌ها سلول عصبی (نورون) از طریق ارتباطاتی که با یکدیگر دارند (سیناپس‌ها)، به حل مسائل یا ذخیره‌سازی اطلاعات می‌پردازند.

۳-۳ یادگیری عمیق

یادگیری عمیق شاخه‌ای از بحث یادگیری ماشینی و مجموعه‌ای از الگوریتم‌هایی است که تلاش می‌کنند مفاهیم انتزاعی سطح بالا را با استفاده از یادگیری در سطوح و لایه‌های مختلف مدل کنند. یادگیری عمیق در واقع نگرشی جدید به ایده شبکه‌های عصبی می‌باشد که سالیان زیادی است وجود داشته و هر چند سال یکبار در قالبی جدید خود را نشان می‌دهد. روشن کردن این نکته ضروری به نظر می‌رسد که شبکه‌های عصبی یک لایه مخفی درونی دارند و شبکه‌ای که چندین لایه مخفی درونی داشته باشد شبکه عمیق نامیده می‌شود. دو مزیت این شیوه یادگیری در زیر آمده است.

بازنمایی یادگیری: نیاز اصلی هر الگوریتم یادگیری ویژگی‌هایی است که از ورودی‌ها استخراج می‌شود. ممکن است این ویژگی‌ها از پیش به صورت دستی تهیه شده و به الگوریتم خوراندن شود که این روش در الگوریتم‌های با ناظر به کار می‌رود. در مقابل روش‌های بدون نظارت خواهد بود که خود اقدام به استخراج ویژگی‌ها از ورودی خواهد نمود. استخراج دستی ویژگی‌ها علاوه بر اینکه زمان بر است معمولاً هم ناقص و در عین حال بیش از حد نیاز ذکر شده می‌باشد. یادگیری عمقی برای ما یک راه استخراج خودکار ویژگی‌ها پدید می‌آورد.

یادگیری چندین لایه بازنمایی‌ها: یادگیری عمیق برای ما این امکان را به وجود می‌آورد که بتوانیم مفاهیم با سطح انتزاع بالا را با استفاده از یادگیری چند لایه از پایین به بالا بسازیم. شکل (۳-۱) مفهوم لایه لایه بودن مراحل یادگیری را در بازشناسی چهره انسان به خوبی نمایش می‌دهد [۴۵]. در شکل (۳-۱) هر چه تعداد لایه‌ها بیشتر می‌شود اطلاعات بیشتری از چهره انسان استخراج شده و به نمایش گذاشته شده است. یا در واقع هر چه به لایه‌های جلوتر می‌رویم می‌توانیم وابستگی‌های سطح بالا را در حد دسته مربوط به داده‌ها مدل کنیم.



شکل (۳-۱) یادگیری لایه لایه [۳۷]

یادگیری عمیق سعی می‌کند با دقت بالا عملیات وزن‌دهی را برای تعداد صفحات بالا انجام دهد. از آنجا که روش‌های آموزش برای شبکه عصبی در سطوح بالا عملاً ناکام ماند تا قبل از روش آموزش عمیق مدل‌های سطحی شبکه عصبی که دارای تعداد صفحات کمی هستند برای دهه‌های زیادی مورد استفاده قرار گرفتند. روش‌های پسانتشار جز معروف‌ترین روش‌های گرادیان نزولی برای مسائل با ناظر در دهه‌های ۶۰ و ۷۰ میلادی بوده است (در این روش اصلاح وزن‌ها برای در جهت حداقل شدن خطای کل برای نمونه‌های آموزشی معرفی شده انجام می‌شود که یک الگوریتم تکراری است) [۲۸].

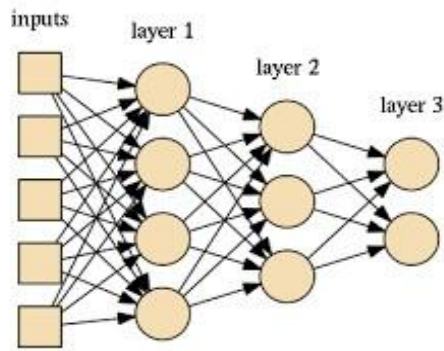
در دهه های ۸۰ و ۹۰ سعی شد از پس انتشار برای شبکه‌های عصبی عمیق با تعداد صفحات بالا استفاده شود، اما بسیار سخت و در عمل برای ۴ لایه به بالا نشدنی بود. یادگیری عمیق به صورت بدون ناظر توانست در بعضی زمینه‌ها عملی شود. در حدود سال های ۲۰۰۰ یادگیری عمیق توانست در مسایل باناظر بهبود یابد و در نهایت شبکه‌های عمیق عصبی که اعتبار خود را نسبتا از دست داده بودند توانستند به خاطر عملکرد خوب نسبت به بقیه‌ی روش‌های یادگیری ماشین توجه افراد زیادی را به خود جلب کنند. از جمله این دست آورده‌ها ماشین‌های کرنل بودند که روی تعداد بسیاری از برنامه های مهم و کاربردی با دقت بالا جواب می‌دادند. از سال ۲۰۰۹ به بعد شبکه‌های عمیق با ناظر در بسیاری از مسابقات جهانی شناسایی الگو برنده شدند. بعدها شبکه‌های عمیق، با مبحث کلی‌تری به اسم یادگیری تقویت شده^۱ مرتبط شدند که در آنها ناظری وجود ندارد. شبکه‌های یادگیری عمیق در واقع در لایه‌های پایینی آموزش را براساس انتقال ویژگی‌های سطح پایین به مجموعه‌ای از ویژگی‌های سطح بالا انجام می‌دهد (توسط ماشین بولتزمان تغییر یافته RBM). به این ترتیب وزنها در لایه‌های پایینی به صورت حدودی تعیین می‌شوند و نیازی به اصلاح چندانی ندارند.

یادگیری عمیق در حوزه‌های مختلفی مانند، دسته‌بندی تصاویر (تشخیص تصاویر)، دسته‌بندی متون، تشخیص صدا، فشرده‌سازی تصویر و... کاربردهای فراوانی دارد [۳۸].

۳-۱-۳ تعداد متغیرهای هر لایه در شبکه‌های عمیق

معمولا در همه کارهای شبکه عمیق مرسوم است که هرچه که به لایه‌های جلوتر می‌رویم تعداد متغیرها کاهش می‌یابد. کم شدن تعداد متغیرهای لایه‌های بعدی را می‌توان از دو جنبه دید: ۱. اصلا آیا اطلاعات موردنیاز از داده‌ها را می‌توان با تعداد کمتری متغیر نشان داد؟ ۲. چرا کم شدن تعداد متغیرها باعث بهبود دقت می‌شود؟ در مورد پرسش اول می‌توان گفت که این کار در لایه‌های اول غیرممکن است ولی هرچه به لایه‌های بعد می‌رویم می‌توانیم وابستگی‌های سطح بالا در حد دسته مربوط به داده‌ها را مدل کنیم در مورد پرسش دوم، نیز پاسخ مشابهی وجود دارد.

¹ Reinforcement



شکل ۳-۲) شبکه عصبی عمیق

۳-۳-۲ یادگیری بدون نظارت برای شبکه‌های عمیق

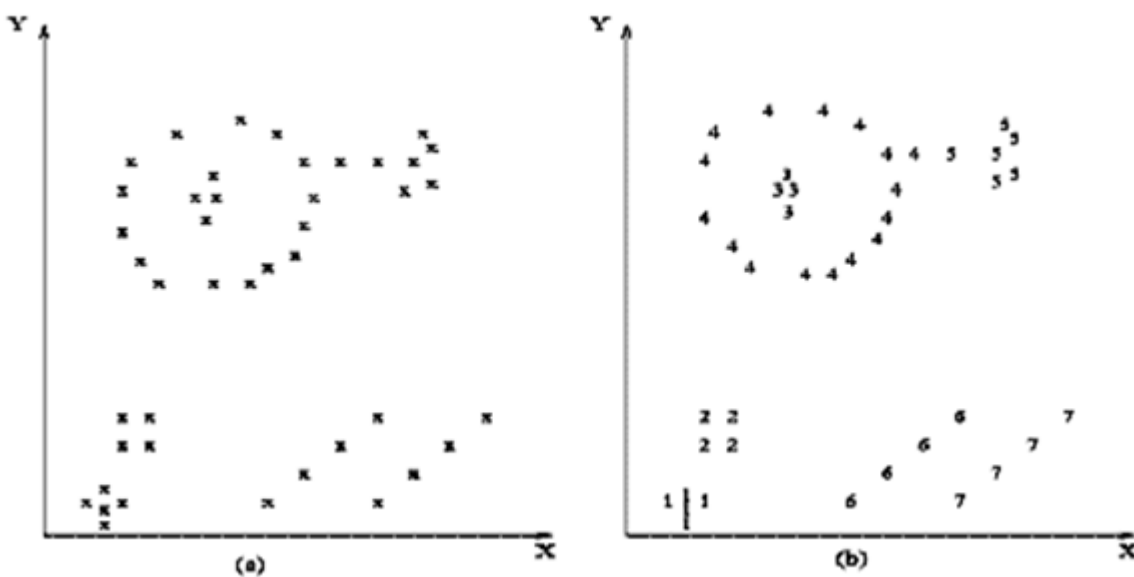
یادگیری بدون نظارت اشاره به استفاده نکردن از اطلاعات نظارتی خاص (مانند برچسب‌های کلاس هدف) در فرآیند یادگیری می‌باشد. یکی از مزیت‌های یادگیری بدون نظارت حجم بالای داده‌های بدون برچسب در مقابل برچسب‌دار و یادگیری ویژگی‌های مناسب برای توصیف داده‌ها می‌باشد. شبکه‌های عمیق برای یادگیری بدون ناظر یا (تولیدی) که برای استخراج همبستگی‌های سطح بالا از داده‌های قابل مشاهده برای تحلیل یا ترکیب اهداف زمانی که هیچ اطلاعاتی در مورد برچسب‌های کلاس هدف موجود نیست در نظر گرفته شده‌اند. همچنین در مدل تولیدی کلاس داده‌ها و مشاهدات بصورت مشترک آموزش داده می‌شوند [۲۹].

۳-۳-۱ کاربردهای یادگیری بدون نظارت

- دسته‌بندی خودکار داده‌های حجیم (متون وبسایت، رفتار مشتری‌ها برای بازاریابی هدفمند)

- داده کاوی و استخراج دانش از داده‌ها
- کاهش ابعاد ویژگی‌ها
- قطعه‌بندی تصاویر (segmentation)
- خوشه‌بندی (clustering) [۳۹].

در شکل (۳-۳) یک مثالی از خوشه بندی آمده است که نمونه‌های ورودی در قسمت a و خوشه‌های مطلوب در b نمایش داده شده است. شماره‌هایی که به هر کدام از نمونه‌ها داده شده است، برچسب آن خوشه‌بند تلقی میشود. وجود تکنیکهای متعدد برای نمایش داده، اندازه گیری مجاور(تشابه) بین عناصر داده و گروه‌بندی آنها باعث شده است که روش‌های خوشه‌بندی به انواع مختلف در طبقه‌بندیهای مختلف معرفی شوند.



شکل (۳-۳) نمونه ای از خوشه بندی بدون نظارت

همچنین در روش پیشنهادی که در فصل چهار آورده شده است نوعی از الگوریتم یادگیری عمیق بدون نظارت برای فشردسازی تصاویر استفاده شده است.

۴-۳ شبکه عصبی عمیق

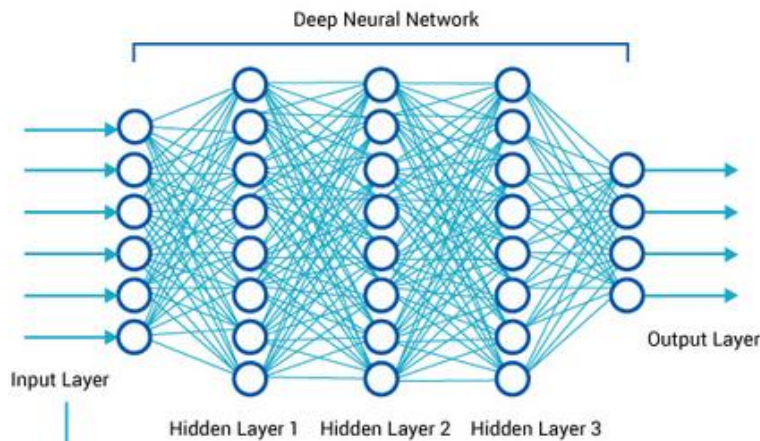
مدل‌های شبکه عمیق توسعه یافته مدل‌های شبکه عصبی برای یادگیری تبدیل غیرخطی روی داده‌ها هستند. این مدل‌ها در سال‌های اخیر در همه زمینه‌ها از جمله پردازش تصاویر ماهواری و پزشکی، پردازش متن و ... مورد توجه فراوان قرار گرفته‌اند. در این مدل‌ها علاوه بر کدگذار در هر لایه

یک کدگشا نیز وجود دارد. برای مثال در مدل کدگذار خودکار پشته‌ای (SAE)¹ در هر لایه یک تبدیل خطی همراه با یک تابع فعالیت روی آن برای بردن به نمایش جدید و یک تبدیل با همین مشخصات برای بازسازی داده داریم. مدل‌های مبتنی بر RBM در هر لایه یک مدل احتمالی مولد دارند، که به صورت خودکار نقش کدگشا را بازی می‌کند.

به عبارت دیگر تفاوت اساسی مدل‌های شبکه عمیق با شبکه عصبی این است که در هر لایه سعی می‌کنند قابلیت بازسازی داده را حفظ کنند یا شبیه مدل‌های RBM با مدل کردن توزیع داده‌ها در هر لایه، اطلاعات فضای ویژگی را حفظ کنند. این نکته اساسی در مدل‌های شبکه عمیق باعث می‌شود که جلوی بیش‌برازش مدل در شبکه عصبی گرفته شود و باعث می‌شود مدل قابلیت تعمیم بیشتری روی داده‌های آزمون داشته باشد. در واقع مدل‌های شبکه عصبی سنتی می‌توانند در تعداد لایه‌های زیاد، به راحتی وزن‌ها را به گونه‌ای تنظیم کنند که اطلاعات نظارتی توجیه شود درحالی‌که روی داده‌های آزمون نتیجه بسیار ناامید کننده به دست آید. در حالیکه در مدل‌های شبکه عمیق این تضمین وجود دارد که فقط نمایش داده‌ها به منظور رسیدن به برچسب یا توجیه اطلاعات نظارتی تغییر کرده است و در واقع فقط این اطلاعات حفظ شود در حالی که قابلیت تعمیم چندانی وجود نداشته باشد. به مدل‌های شبکه عمیق مبتنی بر RBM معمولاً نام شبکه باور عمیق² و به مدل‌های مبتنی بر انکدر-دیکدر شبکه عصبی عمیق یا مدل کدگذار خودکار پشته‌ای می‌گویند. مدل‌های شبکه عمیق در دو نسخه غیرنظارتی و نظارتی (مثل برچسب دسته‌ها) وجود دارند. مدل‌های غیرنظارتی به منظور انتقال فضای ویژگی و رسیدن به یک فضای بدون همبستگی در داده‌های با ابعاد بالا کاربرد دارند. مدل‌های نظارتی معمولاً به منظور دسته‌بندی یا رگرسیون همانند شبکه عصبی به صورت نظارتی آموزش می‌بینند. در شکل (۳-۴) معماری شبکه عصبی عمیق نشان داده شده است.

¹ Stacked-Auto-Encoder

² Deep Belief Network



شکل ۳-۴) معماری شبکه عصبی عمیق [۴۰]

مدل‌های شبکه عمیق توسعه‌یافته مدل‌های شبکه عصبی می‌باشند که امکان بالابردن تعداد لایه‌ها برای یادگیری ویژگی‌های سطح بالای جدید را دارند. مدل RBM و خودرمنزنگار دو مدل معروف به عنوان لایه‌های شبکه عمیق می‌باشند.

طبیعی است که در سیستم‌های فشرده‌سازی تصویر درجه کیفیت تصویر تاثیر زیادی در نتیجه کار دارد. حذف نویز باعث بهبود کیفیت تصویر و در نتیجه بهبود نتیجه می‌شود.

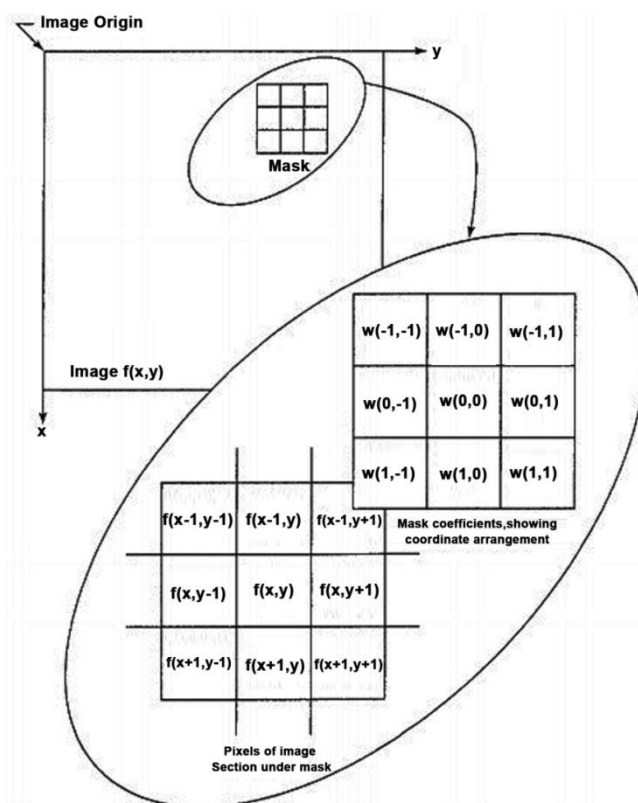
۳-۵ فیلترگذاری روی تصاویر و انواع فیلترهای خطی و غیرخطی

فیلترهای مکانی طوری روی یک تصویر اعمال می‌شوند که مقدار پیکسل را با توجه به مقادیر پیکسل‌های مجاور آن تغییر می‌دهند [۳۰]. مثلاً در فیلتر میانه، مقادیر هشت همسایه مرتب می‌شود و مقدار پیکسل میانی به عنوان مقدار پیکسل وسط انتخاب می‌شود. فیلترها به دو نوع خطی و غیرخطی تقسیم می‌شوند:

۳-۵-۱ فیلترهای خطی

فیلترهای نوع خطی را کانولوشن مکانی هم می‌نامند. در استفاده از فیلترهای خطی، همسایگی‌های یک پیکسل با وزن‌های مشخص شده با هم جمع می‌شوند و نرمالسازی شده و به عنوان مقدار پیکسل قرار داده می‌شوند. برای مشخص کردن وزن‌ها از یک ماتریس استفاده می‌کنیم به ماتریسی که در این

بخش استفاده میکنیم ماتریس وزن‌ها، کرنل، ماسک فیلتر و ... گفته می‌شود. بطور مثال اگر بخواهیم مقدار یک پیکسل مرکزی را متناسب با وزن‌های ویژه ای از هشت مجاورت آن بدست آوریم، کافی است یک ماتریس سه در سه شامل مقادیر وزن‌ها تعریف کرده و نظیر به نظیر در همسایگی‌ها ضرب کرده و جمع آن‌ها را به عنوان مقدار پیکسل مرکزی قرار دهیم [۳۰]. کلا از جمع وزن‌دار مجاورت های یک پیکسل، مقدار جدیدی برای آن پیکسل بدست می‌آید، بدین ترتیب این مدل از فیلترها را خطی می‌نامیم. در شکل (۵-۳) نمونه‌ای از فیلتر خطی نشان داده شده است.

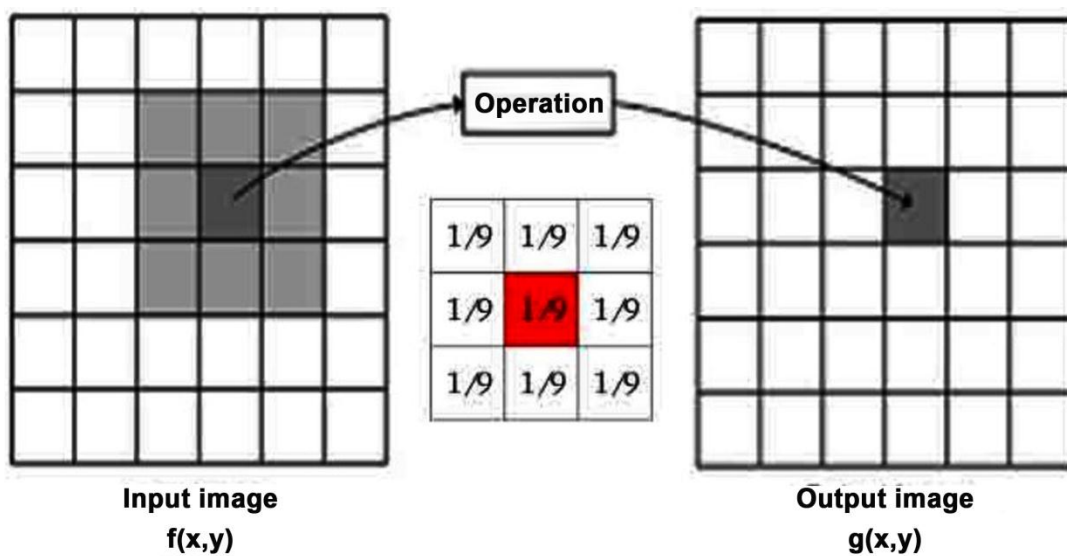


شکل (۵-۳) نمونه ای از فیلتر خطی [۴۱]

این کار باید برای تک تک پیکسل‌های تصویر صورت بگیرد و همچنین ضرب ماسک در تصویر به دو روش همبستگی و کانولوشن صورت می‌گیرد که تنها تفاوت این دو قرینه بودن خروجی آن‌ها نسبت به یکدیگر است.

۲-۵-۳ فیلترهای غیر خطی

فیلترهایی همانند Median غیرخطی هستند، چونکه مقدار حاصله در هر مرحله از عملیات کانولوشن بستگی به ترتیب داده‌ها دارد و مابقی فیلترها که کرنل‌های مشخص را با تصویر کانوال می‌کنند فیلترهای خطی هستند مانند فیلتر گوسین. در کاربرد فیلترهای خطی، تمامی پیکسل‌های تصویر تحت تاثیر ضرایب قرار می‌گیرند به همین دلیل برای اجتناب از این مسئله از فیلترهای غیر خطی می‌توان استفاده کرد. در این مدل، مقدار یک پیکسل مجاور براساس ویژگی خاص بعنوان مقدار پیکسل مرکزی در نظر گرفته می‌شود. مانند اینکه مقدار ماکزیمم مجاور را به عنوان مقدار مرکزی در نظر گرفت.



شکل ۳-۶) نحوه کانولوشن و اعمال فیلتر بر روی پیکسل‌های یک تصویر (فیلتر خطی) [۴۲]

معمولا در هر تصویر دو گروه پیکسل با فرکانس بالا و پایین وجود دارند. پیکسل فرکانس پایین پیکسلی است که اختلاف شدت روشنایی آن پیکسل با پیکسل‌های مجاورش اندک باشد (نقاط آرام تصویر). در مقابل پیکسل فرکانس بالا پیکسلی است که اختلاف شدت روشنایی آن پیکسل با پیکسل‌های مجاورش زیاد باشد (لبه‌ها و نویزها). برای استفاده تبدیلات در هر دو حوزه مکانی و فرکانس از فیلترها استفاده می‌کنیم [۳۰]. فیلترها به دو دسته تقسیم می‌شوند.

۳-۵-۳ فیلترهای پایین گذر

این مدل از فیلتر با گذر از پیکسل‌های با فرکانس کم، بر روی پیکسل‌های با فرکانس بالا اثرگذار خواهند بود. اثر اعمال فیلترهای پایین گذر تصویری معتدل خواهد بود (لبه‌ها و نویزها تا حدودی آرام می‌شوند). فیلترهای میانگین، مد و میانه از جمله فیلترهای پایین گذر هستند که برای مات کردن تصویر و یا حذف نویز استفاده می‌شوند. نمونه‌ای از تصویر مات شده پس از اعمال فیلتر در شکل زیر نشان داده شده است.



شکل ۳-۷) تصویر مات شده پس از اعمال فیلتر [۴۲]

۳-۵-۴ فیلترهای بالا گذر

این مدل از فیلتر با گذر از پیکسل‌های با فرکانس بالا، بر روی پیکسل‌های با فرکانس کم اثرگذار خواهند بود. اعمال فیلتر بالا گذر برای نمایان کردن بیشتر جزئیات تصویر به کار می‌روند. به این فیلترها آشکارکننده لبه‌ها نیز می‌گویند. ساده ترین فیلتر تیزکننده یک ماسک سه در سه است که مقادیر آن بصورت شکل (۳-۸) تعیین می‌گردند. فیلتر لاپلاسیان گووسی (LOG)^۱ برای آشکارسازی لبه استفاده می‌شود. ابتدا تصویر با فیلتر گووسی هموار می‌شود سپس با عبور از فیلتر لاپلاسیان لبه‌ها در تصویر آشکار می‌شود، در فیلتر LOG با افزایش سیکما لبه‌های ضعیف حذف می‌شوند و جزئیات بیشتر حذف می‌شود و کلیات در تصویر باقی می‌ماند و در نتیجه آشکارسازی بهتری از لبه ایجاد می‌شود. با افزایش کرنل نیز لبه‌های مهم و اصلی به صورت باریک آشکار می‌شود که بهینه است.

^۱ Laplacian Of Gaussian (LOG)

$-a$ $(a+1)$	$a-1$ $(a+1)$	$-a$ $(a+1)$
$a-1$ $(a+1)$	$a+5$ $(a+1)$	$a-1$ $(a+1)$
$-a$ $(a+1)$	$a-1$ $(a+1)$	$-a$ $(a+1)$

شکل ۳-۸) فیلتر تیزکننده یک ماسک 3×3 [۴۲]

در پاسخ به اینکه فیلترهای پایین و بالاگذر در پیکسل‌ها چه اثری می‌گذارند می‌توان گفت در تصویر به لبه‌ها فرکانس بالا می‌گویند، فیلتر پایین‌گذر قرار داده می‌شود تا لبه‌ها کم شود تا تصویر صاف شود، برای لبه‌یابی هم از فیلتر بالاگذر استفاده می‌شود.

۳-۶ خودرمزنگار^۱

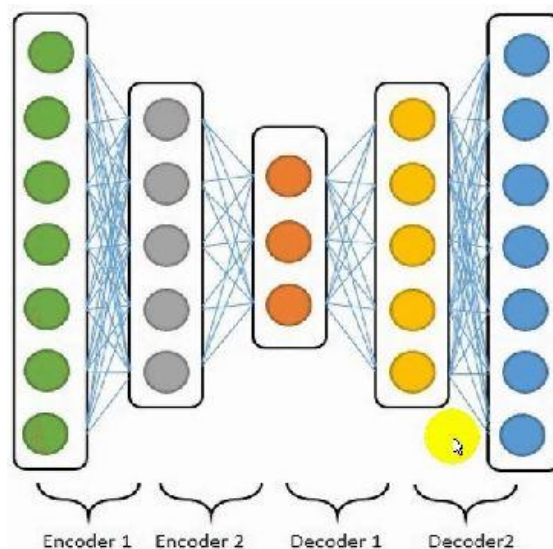
خودرمزنگار عمیق نوع خاصی از شبکه عصبی عمیق (البته بدون برچسب‌های کلاس) می‌باشد که ابعاد بردارهای خروجی برابر بردارهای ورودی می‌باشد. اغلب مقادیر لایه مخفی در خودرمزنگار، نماینده یا مقادیر رمزگذاری شده‌ی داده‌های اصلی که در لایه‌ی ورودی به آن وارد شده‌اند می‌باشند. به این نکته باید توجه داشت که خودرمزنگار یک روش استخراج ویژگی غیر خطی بدون استفاده از برچسب‌های کلاس می‌باشد.

هدف استخراج ویژگی نمایش بهتر و حفظ اطلاعات بجای انجام وظایف طبقه‌بندی می‌باشد. گرچه گاهی اوقات این دو هدف در ارتباط هستند. یک خودرمزنگار بطور طبیعی از یک لایه ورودی که نشان‌دهنده داده اصلی یا بردارهای ویژگی ورودی (مثل پیکسل‌های در تصویر و یا طیف در صوت) و چند لایه پنهان که نشان‌دهنده ویژگی‌های تبدیل شده و یک لایه خروجی که منطبق بر لایه ورودی برای بازسازی است تشکیل شده است [۳۱].

¹ Autoencoder

زمانیکه تعداد لایه‌های پنهان بیش از یک باشد خودرمزنگار عمیق در نظر گرفته می‌شود. زمانیکه هدف فشرده سازی است ابعاد لایه‌های پنهان می‌تواند نسبت به لایه‌های ورودی کوچکتر باشد و زمانیکه هدف نگاشت ویژگی به یک فضای چند بعدی است بزرگتر باشد [۳۲]. اتوانکودر اغلب با استفاده از روش‌های متفاوت پس انتشار و بطور معمول از روش گرادیان نزولی تصادفی آموزش می‌بیند. زمانیکه از پس انتشار برای آموزش شبکه با تعدادی لایه‌های پنهان استفاده می‌شود مشکلات اساسی وجود دارد.

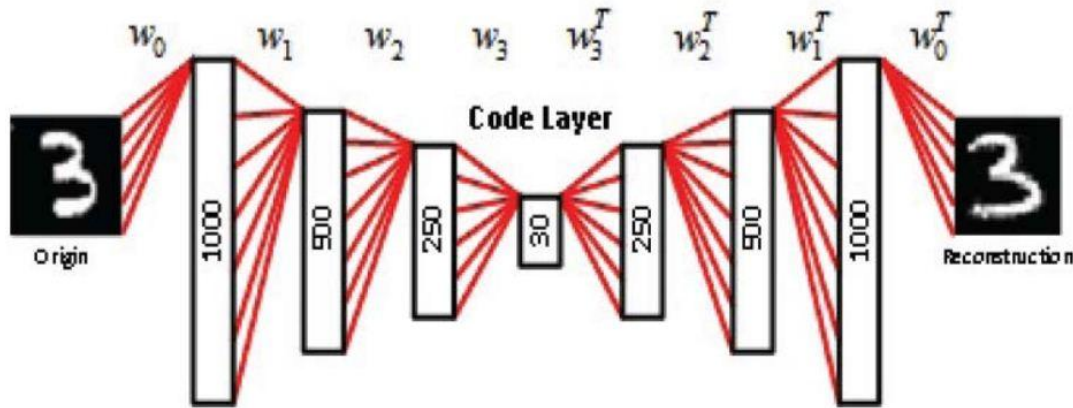
مدل‌های خودرمزنگار جز اولین مدل‌هایی در حوزه یادگیری عمیق بودند که از یادگیری بدون نظارت استفاده کردند. همان‌طور که ملاحظه می‌شود در شکل زیر ساختار کلی یک خودرمزنگار عمیق نشان داده شده است. همان‌طور که در این ساختار نشان داده شده در خودرمزنگار عمیق چند لایه رمزگذار و چند لایه رمزگشا وجود خواهد داشت.



شکل ۳-۹) ساختار کلی اتوانکودر عمیق

مثلا در شکل زیر یک عدد دست نویس ورودی به یک شبکه رمزگذاری داده شده که به یک سری کد (ویژگی) تبدیل می‌شود و سپس همان ویژگی‌ها به یک شبکه رمزگشایی داده شده و همان عدد در خروجی بازسازی می‌شود. یکی از مزایای خودرمزنگار ساختار ساده و منظم آن می‌باشد، همان‌طور

که در شکل نشان داده شده است برای بازسازی تصویر عدد دستنویس سه لایه رمزگذاری در قسمت رمزگشایی کنار هم قرار گرفته شده و با توجه به مقدار ویژگی، تصویر بازسازی می‌شود.



شکل ۳-۱۰) نمونه ای از فشرده‌سازی عدد دستنویس

فشرده‌سازی عدد دستنویس بدین صورت است که یک عدد دستنویس در ورودی داریم، آن را به یک بردار تبدیل نموده و به یک خودرمزنگار اعمال می‌نماییم. خودرمزنگار که شامل دو قسمت رمزگذار و رمزگشا می‌باشد، که بردار را در هر لایه کاهش بعد داده و سپس رمزگشا ویژگی‌ها را به همان حالت اولیه بازسازی می‌کند.

۳-۷ آموزش خودرمزنگار

بر اساس تعریف ارائه شده در [۳۳] اتوانکودر تابعی است که عمدتاً از دو بخش تشکیل می‌گردد:

الف) رمزگذار: که یک تابع استخراج ویژگی بوده و بردار ویژگی را از روی ورودی‌ها محاسبه می‌نماید. بنابراین اگر بردار ویژگی را با h ، رمزگذار را با f_θ و مجموعه داده‌ها را با $x^{(t)}$ نمایش دهیم رابطه زیر برقرار خواهد بود:

$$h^{(t)} = f_\theta(x^{(t)}), \quad x^{(t)} = \{x^{(1)}, \dots, x^{(T)}\} \quad (۱-۳)$$

که h بردار ویژگی یا نمایش یا کد محاسبه شده از x می‌باشد.

ب) رمزگشا: تابعی است که ان را با g_{θ} نمایش داده و با استفاده از رابطه زیر نگاشتی از فضای ویژگی به فضای ورودی انجام می دهد.

$$r = g_{\theta}(h) \quad (۲-۳)$$

در واقع از یک تابع احتمال مشخص، مدل های احتمالی تعریف شده و برای حداکثر نمودن (اغلب بصورت تقریبی) تشابه داده ها، آموزش داده می شوند.

خودرمزنگارها با استفاده از اصول آموزشی متفاوتی آموزش داده می شوند. مجموعه پارامترهای θ رمزگذار و رمزگشا به طور مشابه برای بازسازی ورودی اصلی، یاد گرفته می شوند. یعنی تلاش می شود تا خطای بازسازی $L(x,r)$ تا حد امکان حداقل شود. این کار با بازسازی r از روی x با بیش آموزش و اندازه گیری اختلاف میان x و r انجام می شود. بطور خلاصه آموزش خودرمزنگار شامل پیدا نمودن بردار پارامتر θ برای حداقل نمودن خطای بازسازی می باشد.

$$J_{AE}(\theta) = \sum_t L(x^{(t)}, g_{\theta}(f_{\theta}(x^{(t)}))) \quad (۳-۳)$$

به حداقل رساندن این مقدار نیز معمولا به روش گرادیان نزولی تصادفی مشابه روش آموزش پرسپترون چند لایه انجام می شود. عمومی سازی خوب در اینجا به معنی خطای بازسازی کوچک برای نمونه های آزمایش است در حالی که این خطا برای سایر پیکربندی های دیگر x حداکثر باشد.

فصل چهارم

روش پیشنهادی و نتایج شبیه‌سازی

۴-۱ مقدمه

کاربرد روزافزون شبکه‌های عصبی در حل مسائل مهندسی و سیستم‌های هوشمند، سبب گرایش محققین به سمت ابداع روش‌های آموزش و طراحی معماری کارآمدتر آن‌ها شده است. معماری شبکه عصبی عمیق شامل تعدادی لایه مخفی، تعدادی نرون‌ها در لایه‌های مخفی و یک نوع تابع تحریک بوده و هر یک از این پارامترها بر روی عملکرد شبکه عصبی تاثیر مستقیم و بسزایی دارد. از سوی دیگر، عمیق‌تر بودن شبکه‌های عصبی منجر به استخراج بیشتر اطلاعات و یا ویژگی‌ها خواهد شد که این مطلب به نوبه خود اهمیت فرآیند عمیق بودن شبکه عصبی را آشکار می‌کند. تعیین دستی پارامترهای شبکه عصبی موجب سختی طراحی آن و همچنین عدم اطمینان از بهینه بودن عملکرد آن می‌شود؛ زیرا شبکه عصبی با ابعاد خیلی کوچک به طور کامل قادر به یادگیری تمام داده‌های آموزشی نیست. همچنین در شبکه عصبی با ابعاد خیلی بزرگ جنبه‌های نامربوط داده‌های آموزشی یاد گرفته می‌شوند و پدیده بیش برآزش رخ می‌دهد. در این پایان‌نامه مدلی از الگوریتم یادگیری عمیق بدون نظارت به نام خودرمنگار عمیق برای فشرده‌سازی تصویر استفاده شده است. همچنین بمنظور بهبود کارایی فشرده‌سازی و حفظ کیفیت تصویر فیلتر حذف نویز بکار رفته است.

۴-۲ روش پیشنهادی

در این پژوهش به منظور فشرده‌سازی تصاویر ورودی از طریق استخراج ویژگی‌های موردنظر از تصاویر، از یک شبکه عصبی عمیق با الگوریتم آموزشی پس انتشار خطا استفاده می‌شود. الگوریتم پس انتشار خطا الگوریتم محبوبی است که یادگیری را در یادگیری عمیق بسیاری از مدل‌ها پیاده می‌کند.

در این پژوهش به دلیل عملکرد مطلوب شبکه‌های عصبی عمیق در فرآیند فشرده‌سازی، از یک خورمنگار عمیق چهار لایه برای فشرده‌سازی استفاده شده است. معیار ارزیابی شبکه نیز PSNR¹ و

¹ Peak Signal to Noise Ratio

CR¹ در نظر گرفته شده است. از مشتق تابع خطا برای تنظیم وزن‌های شبکه استفاده می‌شود. در ادامه تعیین ساختار شبکه، خروجی‌های هر لایه از شبکه قبل از این که وارد لایه بعدی شوند ابتدا به یک تابع انتقال وارد شده و نرمال می‌گردند و بازه آن‌ها به طور صحیح مشخص می‌شوند.

برای هر مشاهده ثبت شده، ابتدا تمامی داده‌ها در بازه ۰ تا ۱+ نرمال می‌گردند. تصویر ورودی برای فشرده‌سازی به یک خودرمنگار عمیق اعمال می‌شود، برای هر رنگ یک خودرمنگار آموزش داده شده و تصویر ورودی بعد از نرمال‌سازی به سه تا تصویر قرمز، آبی و سبز تقسیم گردید. هر کدام از این تصاویر تک رنگ به بلوک‌های ۱۶*۱۶ تقسیم شده و سپس این ۲۵۶ پیکسل مربوط به هر بلوک تبدیل به یک بردار ۲۵۶ تایی شده و به ورودی خودرمنگار اعمال شده است. تعداد نرون‌های لایه مخفی خودرمنگار را با توجه به میزان فشرده‌سازی و ورودی انتخاب می‌نماییم. یک بار با صد تصویر (پایگاه داده) شبکه خودرمنگار را آموزش داده و بعد از به دست آمدن وزن‌ها تصویر آزمایش به آن اعمال می‌شود. با توجه به کمتر بودن نرون‌های لایه مخفی این لایه در بردارنده‌ی تصویر فشرده‌شده است. همچنین خروجی خودرمنگار تصویر بازسازی‌شده را بر اساس تصویر فشرده‌شده‌ی موجود در لایه‌ی مخفی ارائه می‌نماید. مشکلی که وجود دارد این است که چون خودرمنگار تصویر را بلوک به بلوک دریافت می‌نماید مرز بلوک‌ها در تصویر بازسازی شده نمایان می‌شود که در این پایان نامه برای رفع آن و همچنین بهبود کیفیت تصویر بکار برده شده این است که از فیلتر حذف نویز برای حذف لبه‌ها و نقاط تیز تصویر استفاده شده است. همچنین این ایده منجر به بهبود معیارهای ارزیابی شده است.

در فاز آموزش شبکه عصبی ابتدا مشاهدات را به صورت ترتیبی، یکی یکی به شبکه می‌دهیم و سپس با محاسبه خروجی شبکه میزان خطای MSE را برای مشاهده محاسبه می‌کنیم و از اطلاعات

¹ Compression Ratio

مشتق این خطا در جهت صفر کردن (کمتر کردن) مقدار خطا استفاده می‌کنیم. این کار را برای همه مشاهدات و در تعداد دوره‌های^۱ مشخصی تکرار می‌کنیم تا به یک حد خطای قابل قبول برسیم.

در فاز آزمایش نیز با اعمال مشاهدات آزمایشی به شبکه عصبی آموزش دیده خروجی را بدست آورده و پس از مقایسه با خروجی مطلوب (کلاس مشاهده) میزان خطا را محاسبه می‌کنیم و این کار را برای تمامی داده‌های آزمایشی تکرار می‌کنیم و در نهایت با مجموع همه خطاهای بدست آمده به خطای کلی برای داده آزمایشی می‌رسیم. مسلم است هرچه خطای داده آزمایشی کمتر باشد بدین مفهوم است که شبکه عصبی بهتر آموزش دیده است. همانگونه که در شکل (۴-۱ الف و ب) مشاهده می‌کنید بلوک دیاگرام روش پیشنهادی برای آموزش خودرمن‌نگار در فشرده‌سازی تصویر و همچنین بازسازی تصویر فشرده‌شده نشان داده شده است.

¹ Epoch



شکل ۴-۱) الف: بلوک دیاگرام روش پیشنهادی برای آموزش خودرمنگار در فشرده سازی تصویر

فرآیند آموزش فشرده‌سازی: در فرآیند آموزش فشرده‌سازی تصاویر رنگی را نرمال نمودیم تا مقادیر روشنایی‌ها بین صفر و یک قرار گیرد، سپس تصاویر رنگی را به سه تا تصویر خاکستری که هر کدام یکی از سه مولفه‌ی رنگ را نمایندگی می‌کنند تقسیم نمودیم به دلیل بزرگ بودن تصاویر ورودی تمام تصاویر را به بلوک‌های 16×16 بلوک بندی نموده و بلوک‌ها را به صورت بردارهای 256

تایی ایجاد نمودیم. یعنی هر بلوک به ۲۵۶ تا پیکسل تبدیل شده است و سپس بردارهای ۲۵۶ تایی به ورودی خودرمننگار اعمال شده است) به این دلیل بلوک را به بردار تبدیل نمودیم که ورودی شبکه عصبی بردار را قبول می کند). همچنین با روش گرادیان نزولی خودرمننگار آموزش داده شده و وزن‌ها اصلاح می‌شود. وقتی خطا از حد قابل قبول کمتر شد دیگر آموزش ادامه داده نخواهد شد و وزن‌های نهایی ذخیره می‌شوند.



شکل ۴-۱) ب: بلوک دیاگرام روش پیشنهادی برای آموزش خودرمننگار در بازسازی تصویر

فرآیند بازسازی: به ازای هر مولفه‌ی رنگ، مقادیر لایه مخفی مربوط به تصویر فشرده‌شده را به نرون‌های آخرین لایه مخفی اعمال کردیم. وزن‌های بین نرون‌های لایه مخفی و لایه خروجی که در فرآیند آموزش بدست آمده بود به آن اعمال می‌گردد. بردار ۲۵۶ تایی شامل ۲۵۶ پیکسل که در لایه خروجی بدست می‌آید به یک بلوک ۱۶*۱۶ تبدیل شده است. بعد از اینکه سه مرحله بالا را برای هر سه مولفه رنگ انجام دادیم، بلوک‌ها را در هر مولفه رنگ بهم می‌چسبانیم و همچنین سه مولفه‌ی رنگ را در هم ادغام می‌کنیم که دوباره تصویر رنگی حاصل شود. تصویر حاصل شده دارای نویز می‌باشد که به همین دلیل از فیلتر گووسی برای حذف نویز استفاده شده است.

۳-۴ پارامترهای ارزیابی

برای ارزیابی روش پیشنهادی از معیار CR و PSNR برای فشرده‌سازی تصاویر استفاده شده است. هدف نهایی جستجوی بهترین وزن‌های شبکه عصبی است که به کوچکترین میانگین مربعات خطا (MSE) منتهی شود، MSE مطابق رابطه (۱-۴) تعریف می‌شود.

$$MSE = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M (u_{ij} - u_{ij}^*)^2 \quad (1-4)$$

در رابطه فوق N تعداد سطرهای تصویر، M تعداد ستون‌های تصویر، u_{ij} مقدار پیکسل i و j در تصویر اصلی و u_{ij}^* مقدار پیکسل i و j در تصویر بازسازی شده می‌باشند.

یکی دیگر از پارامترهای ارزیابی PSNR می‌باشد. PSNR را می‌توان متداول‌ترین روش برای اندازه‌گیری کیفیت تصویر دانست.

$$PSNR = 10 \cdot \text{LOG}_{10} \frac{MAX^2}{MSE} \quad (2-4)$$

در رابطه (۲-۴) MAX برابر با بیشترین مقدار ممکن روشنایی می‌باشد. هرچه میزان PSNR بیشتر باشد نشان دهنده‌ی این است که بازسازی تصویر بهتر صورت گرفته است.

۴-۴ پایگاه داده

در این پژوهش به دلیل استفاده از تصاویر با اندازه‌های بزرگ در مقایسه کار انجام شده با کارهای پیشین از ۱۰۰ تصویر طبیعی رنگی در دسته‌های مختلف و اندازه‌ی یکسان ۵۱۲ در ۵۱۲ برای آموزش شبکه استفاده شده است [۴۴].

۴-۵ نتایج شبیه‌سازی

برای ارزیابی عملکرد روش پیشنهادی، آزمایش‌هایی بمنظور مقایسه روش پیشنهادی با چندین روش موجود انجام شده و نتایج در این فصل آورده شده است. مقایسه عملکرد PSNR و bpp قبل از اعمال فیلتر حذف نویز برای روش پیشنهادی با روش‌های مشابه دیگر در جدول (۴-۱) نشان داده شده است.

جدول (۴-۱): مقایسه عملکرد $PSNR(db)$ روش پیشنهادی با روش [۴۳] قبل و بعد از اعمال فیلتر حذف نویز

بیت در هر پیکسل (bpp)	۰/۴ (bit/pixel) تصویر lena	۰/۴۳ (bit/pixel) تصویر peppers	۰/۷۸ (bit/pixel) تصویر Baboon
فضای رنگ YUV	۲۳	۲۲	۱۶
فضای رنگ YIQ	۲۳/۶	۲۲	۱۶
فضای رنگ CIEXYZ	۲۷/۸	۲۴	۱۹/۳
روش پیشنهادی قبل از اعمال فیلتر	۲۲/۲	۲۰/۴	۱۵/۳
روش پیشنهادی بعد از اعمال فیلتر	۲۷/۳	۲۴/۲	۱۹/۸

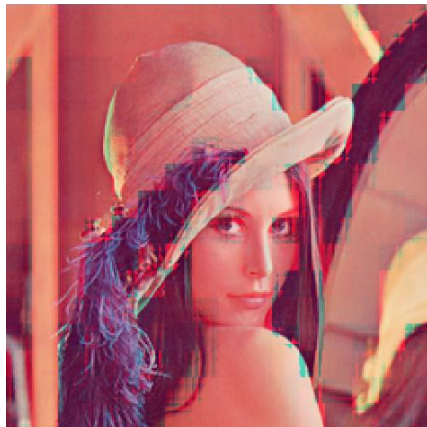
همان‌طور که در جدول (۴-۱) نشان داده شده است عملکرد روش پیشنهادی قبل از اعمال فیلتر نتوانسته نتایج را بهبود بخشد. حال فیلتر حذف نویز به نتایج فوق اعمال شده که در نتیجه هم مقادیر را بهبود بخشیده هم کیفیت تصویر بهبود یافته است.



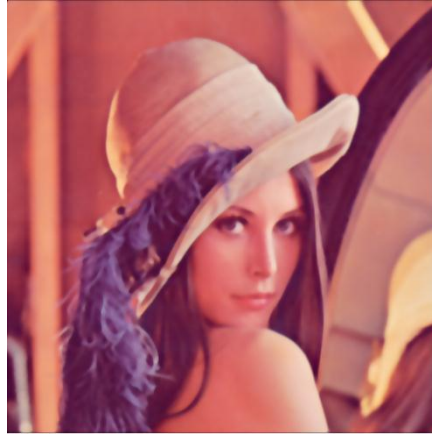
شکل ۴-۲) الف: تصویر lena اصلی



شکل ۴-۲) ب: تصویر lena فشرده شده در فضای YUV

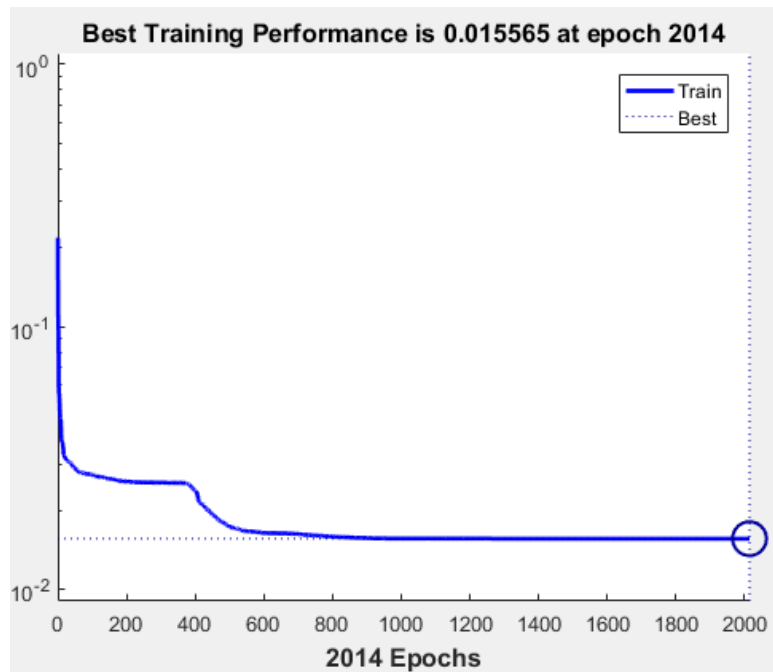


شکل ۴-۲) ج: تصویر lena فشرده شده قبل از اعمال فیلتر در روش پیشنهادی

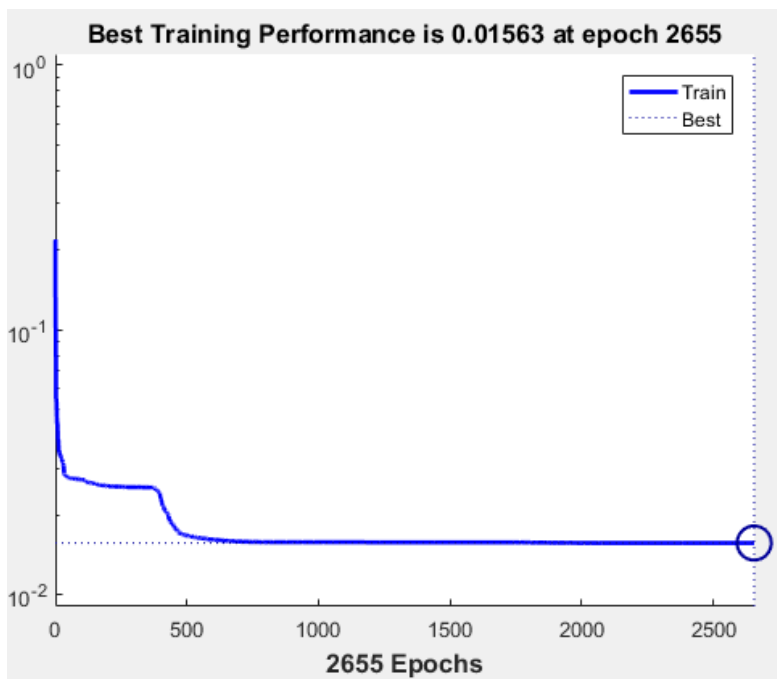


شکل (۲-۴) د: تصویر lena فشرده شده بعد از اعمال فیلتر در روش پیشنهادی

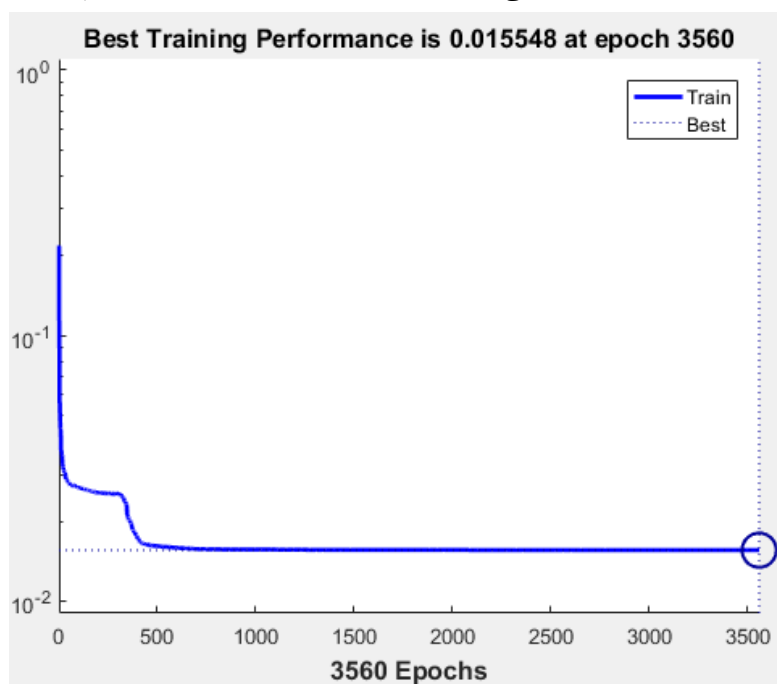
همچنین در شکل (۳-۴ الف، ب و ج) تعداد تکرار آموزش لازم برای همگرایی روش پیشنهادی در سه مولفه‌ی رنگ RGB در تصویر lena به ترتیب نشان داده شده است.



شکل (۳-۴) الف: نمودار همگرایی فرآیند آموزش مولفه‌ی رنگ قرمز در روش پیشنهادی



شکل (۳-۴) ب: نمودار همگرایی فرآیند آموزش مولفه‌ی رنگ سبز در روش پیشنهادی



شکل (۳-۴) ج: نمودار همگرایی فرآیند آموزش مولفه‌ی رنگ آبی در روش پیشنهادی

همچنین در جدول (۲-۴) عملکرد روش پیشنهادی با یکی دیگر از روشهای مشابه آورده شده

است.

جدول (۲-۴): مقایسه عملکرد $PSNR$ روش پیشنهادی با روش [۳۴] مشابه بعد و قبل از اعمال فیلتر حذف نویز

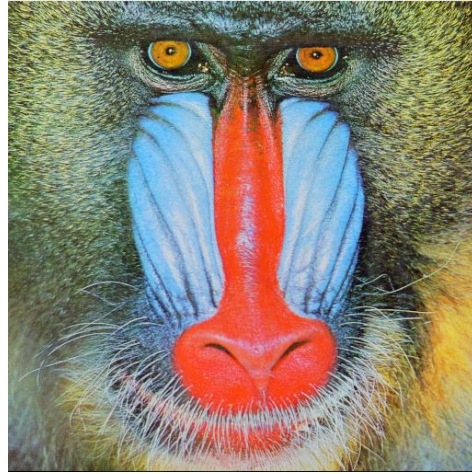
روش پیشنهادی بعد از اعمال فیلتر حذف نویز	روش پیشنهادی قبل از اعمال فیلتر حذف نویز	روش فشرده سازی تصویر بر اساس پیکسلهای مشابه [۳۴]	نرخ فشرده سازی
CR=4.68:1			
۲۴/۳	۲۱/۴۷	۲۲/۶۶	تصویر lena
CR=4.14:1			
۲۲/۱۰	۱۸/۸۹	۱۹/۷۳	تصویر Mandrill
CR=4.2:1			
۲۴/۵	۲۱/۸۸	۲۲/۰۰۶۹	تصویر Squal

همچنین در جدول (۳-۴) زمان تست روش پیشنهادی نشان داده شده است. زمان آموزش در روش پیشنهادی با مشخصات نشان داده شده در جدول (۲-۴) ۷ ساعت و ۴۳ دقیقه و ۵۲ ثانیه می باشد.

جدول (۳-۴): مقدار زمان تست روش پیشنهادی با مقادیر نشان داده شده در جدول (۲-۴)

روش پیشنهادی بعد اعمال فیلتر حذف نویز	روش پیشنهادی قبل اعمال فیلتر حذف نویز	نوع تصویر
01:19 sec	01:12 sec	تصویر lena
01:16 sec	01:09 sec	تصویر Mandrill
01:23 sec	01:32 sec	تصویر Squal

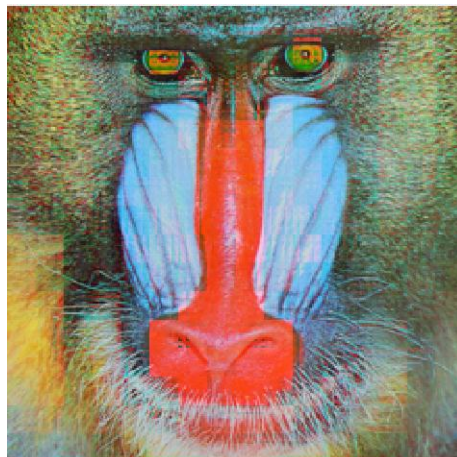
همانطور که در جدول بالا مشاهده می کنید بهبود عملکرد روش پیشنهادی بعد از اعمال فیلتر حذف نویز نشان داده شده است.



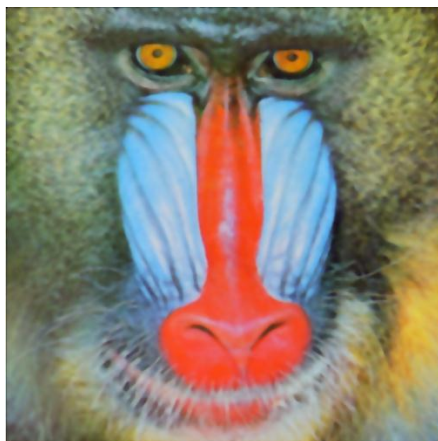
شکل ۴-۴ الف: تصویر Mandrill اصلی



شکل ۴-۴ ب: تصویر Mandrill فشرده شده در روش مشابه [۳۴]



شکل ۴-۴ ج: تصویر Mandrill فشرده شده قبل از اعمال فیلتر حذف نویز در روش پیشنهادی



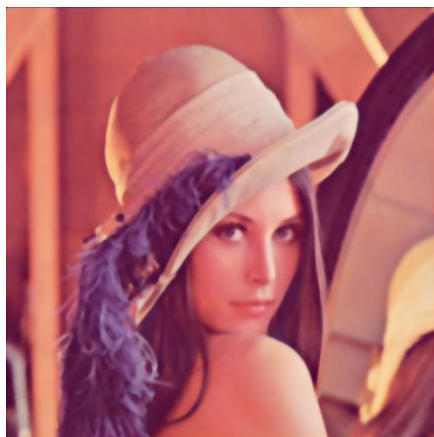
شکل ۴-۴) د: تصویر Mandrill فشرده شده بعد از اعمال فیلتر حذف نویز در روش پیشنهادی



شکل ۴-۵) الف: تصویر lna اصلی

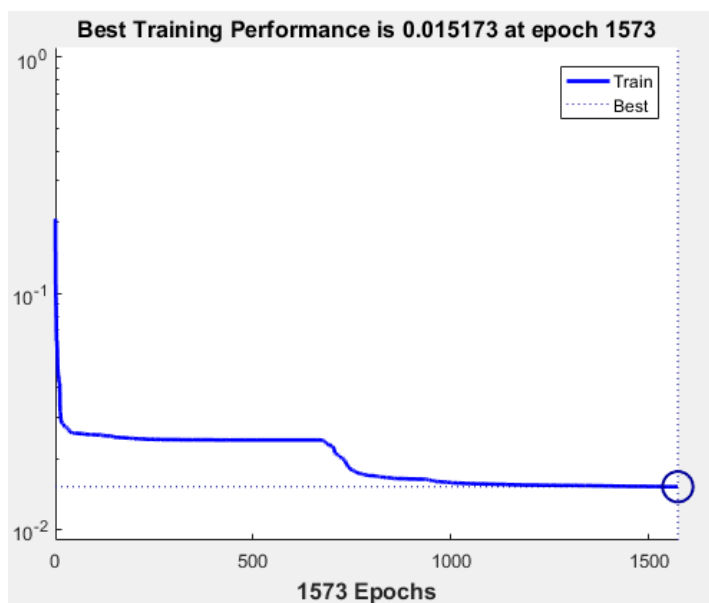


شکل ۴-۵) ب: تصویر lna فشرده شده در روش مشابه [۳۴]

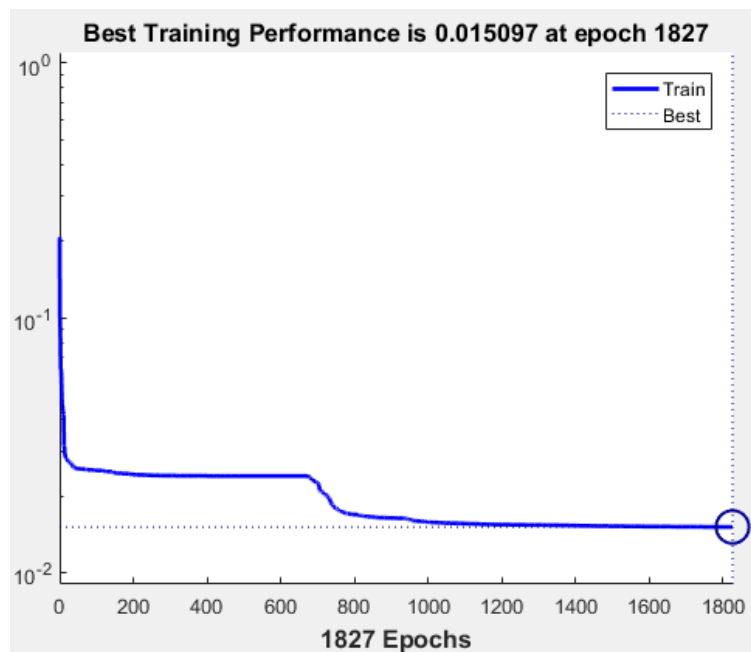


شکل ۴-۵) ج: تصویر lena فشرده شده پس از اعمال فیلتر حذف نویز در روش پیشنهادی

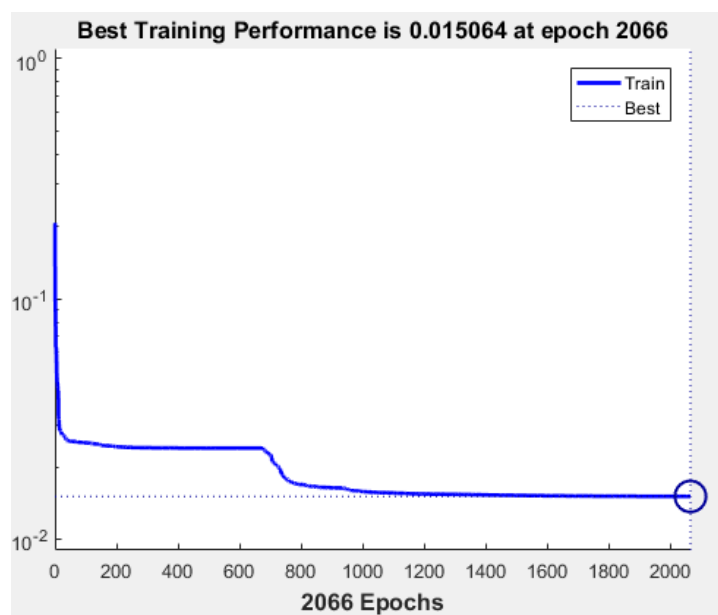
همچنین در شکل (۴-۶ الف، ب و ج) تعداد تکرار آموزش لازم برای همگرایی روش پیشنهادی در سه مولفه‌ی رنگ RGB در تصویر lena به ترتیب نشان داده شده است.



شکل ۴-۶) الف: نمودار همگرایی فرآیند آموزش مولفه‌ی رنگ قرمز در روش پیشنهادی



شکل ۴-۶) ب: نمودار همگرایی فرآیند آموزش مولفه‌ی رنگ سبز در روش پیشنهادی



شکل ۴-۶) ج: نمودار همگرایی فرآیند آموزش مولفه‌ی رنگ آبی در روش پیشنهادی

۴-۶ نتایج آزمون MOS^۱

در این پایان نامه برای ارزیابی بصری کیفیت تصاویر شبیه سازی شده از آزمون MOS استفاده شده است. همچنین تعداد افراد جامعه آماری استفاده شده ۵۱ نفر می باشد. حداقل سن افراد در این آزمون ۱۹ سال و حداکثر سن ۲۷ سال می باشد که از میان دانشجویان دانشگاه علوم پزشکی شاهرود این آزمون انجام شده است.

جدول ۴-۴): مقایسه بصری فضای رنگ YUV [۴۳] و روش پیشنهادی برای تصویر lena

بد	ضعیف	متوسط	خوب	عالی	کیفیت
۰	۳ نفر	۱۴ نفر	۱۹ نفر	۱۵ نفر	فضای رنگ YUV روش [۴۳]
۰	۲ نفر	۶ نفر	۲۶ نفر	۱۷ نفر	روش پیشنهادی

همانطور که در جدول (۴-۴) نشان داده شده می توان نسبتاً از برتری ارزیابی بصری روش پیشنهادی نسبت به روش مشابه نتیجه گرفت. همچنین در جداول (۴-۵ و ۴-۶) مقایسه بصری فشرده سازی تصویر براساس پیکسل های مشابه [۳۴] و روش پیشنهادی برای تصاویر lena و Mandrill نشان داده شده است.

جدول ۴-۵): مقایسه بصری فشرده سازی بر اساس پیکسل های مشابه [۳۴] و روش پیشنهادی برای تصویر Mandrill

بد	ضعیف	متوسط	خوب	عالی	کیفیت
۰	۹ نفر	۱۹ نفر	۱۸ نفر	۵ نفر	فشرده سازی بر اساس پیکسل های مشابه [۳۴]
۰	۰	۲ نفر	۱۶ نفر	۳۳ نفر	روش پیشنهادی

¹ Mean Opinion Score (MOS)

جدول ۴-۶): مقایسه بصری فشرده‌سازی بر اساس پیکسل‌های مشابه [۳۴] و روش پیشنهادی برای تصویر *lena*

کیفیت	عالی	خوب	متوسط	ضعیف	بد
فشرده‌سازی بر اساس پیکسل‌های مشابه [۳۴]	۴ نفر	۱۷ نفر	۱۴ نفر	۱۶ نفر	۰
روش پیشنهادی	۳۵ نفر	۱۳ نفر	۳ نفر	۰	۰

با توجه به نتایج آزمون فوق می‌توان نتیجه گرفت که روش پیشنهادی توانسته است عملکرد و کارایی بهتری نسبت به روش‌های مشابه داشته باشد.

فصل پنجم: نتیجه گیری

و پیشنهادات

۱-۵ نتیجه گیری

طرح فشرده سازی تصاویر طبیعی رنگی با استفاده از شبکه عصبی در این پایان نامه بررسی شد. از روش یادگیری عمیق، مدلی به نام اتوانکودر عمیق برای تقویت عملکرد سیستم بهره برداری شد. در واقع این روش برای فشرده سازی تصاویر طبیعی مناسب می باشد و دیدیم (جدول ۴-۱ و ۴-۲) که کیفیت تصویر بازسازی شده بهتر شده است.

در مقایسه معیارهای روش پیشنهادی با روش های مشابه در می یابیم که روش پیشنهادی معیارهای ارزیابی را بهبود داده است و همچنین با مشاهده نتایج شبیه سازی در می یابیم که روش پیشنهادی کیفیت تصاویر بازسازی را بهبود بخشیده است.

۲-۵ پیشنهادات

در این پایان نامه توانستیم یکی از چالش های مهم در فشرده سازی که حفظ کیفیت تصویر اصلی می باشد را در بازسازی تصویر حفظ نماییم، با توجه به اینکه از یادگیری عمیق (چندین لایه) برای فشرده سازی استفاده شده است نتوانسته زمان آموزش را بهبود آنچنانی بدهد. با توجه به حوزه گسترده یادگیری عمیق و مدل های مختلف آن این پیشنهاد می شود که از مدل های دیگر برای فشرده سازی تصاویر استفاده شود تا با حفظ کیفیت تصویر، زمان آموزش را بتوان بهبود داد.

- [1] Wang, Wei, M. N. S. Swamy, and M. Omair Ahmad. "RNS application for digital image processing" *IEEE System-on-Chip for Real-Time Applications*, vol. 1, no. 1, pp. 77-80, 2004.
- [2] Ramyachitra, D., and P. Manikandan. "Imbalanced dataset classification and solutions: a review" *International Journal of Computing and Business Research*, vol.5, no.4, pp. 1031-1059, 2014.
- [3] Dhawan, Sachin. "A review of image compression and comparison of its algorithms" *International Journal of Electronics & Communication Technology*, vol. 2, no. 1, pp. 22-26, 2011.
- [4] Reddy, K. Siva Nagi, et al. "Image Compression and reconstruction using a new approach by artificial neural network" *International Journal of Image Processing*, vol. 6, no. 2, pp. 68-85, 2012.
- [5] Wilamowski, Bogdan M., Yixin Chen, and Aleksander Malinowski. "Efficient algorithm for training neural networks with one hidden layer" *IEEE International Joint Conference*, vol. 3, no 1, 1999.
- [6] Chen, Tai-cong, et al. "Acceleration of Levenberg-Marquardt training of neural networks with variable decay rate" *Proceedings of the International Joint*, Vol. 3, no. 2, pp. 1725-1728, 2003.
- [7] Khashman, Adnan, and Kamil Dimililer. "Image compression using neural networks and haar wavelet" *WSEAS Transactions on Signal Processing*, vol. 4, no. 5, pp. 330-339, 2008.
- [8] Mulcahy, Colm. "Image compression using the Haar wavelet transform" *Spelman Science and Mathematics Journal*, vol. 1, no. 1, pp. 22-31, 1997.
- [9] Godfrey, Keith RL, and Yiannis Attikiouzel. "Applying neural networks to colour image data compression" *Technology Enabling Tomorrow: Computers, Communications and Automation* , vol. 1, no. 1, pp. 545-549, 1992.

- [10] Pandian, S. Immanuel Alex, and J. Anitha. "A neural network approach for color image compression in transform domain" *International journal of recent trends in engineering*, vol. 2, no. 2, pp. 152-154, 2009.
- [11] Gillespie, Wilford. "Still image compression using neural networks" *Utah University. Logan, UT*, 2005.
- [12] Kouamo, Stephane, and Claude Tangha. "Image Compression with Artificial Neural Networks" *International Joint Conference*, vol. 1, no. 2, pp. 240-245, 2013.
- [13] Jiang, J. "Image compression with neural networks—a survey" *Signal Processing: Image Communication*, vol. 14, no. 9, pp. 737-760, 1999.
- [14] Namphol, Aran, Steven H. Chin, and Mohammed Arozullah "Image compression with a hierarchical neural network." *IEEE transactions on Aerospace and Electronic Systems*, vol. 32, no. 1, pp. 326-338, 1996.
- [15] Carrato, Sergio. "Neural networks for image compression" *In Neural Networks, Adv*, vol. 2, no. 1, pp. 177–198, 1992.
- [16] Panda, S. S., et al. "Image compression using back propagation neural network" *International journal of engineering science and advance technology*, vol. 2, no. 1, pp. 74-78, 2012.
- [17] Abdel-Wahhab, Osama, and Moustafa M. Fahmy. "Image compression using multilayer neural networks" *IEE Proceedings-Vision, Image and Signal Processing*, vol. 144, no. 5, pp. 307-312, 1997.
- [18] Cottrell, Garrison W., and Paul Munro. "Principal components analysis of images via back propagation" *Visual Communications and Image Processing*, vol. 1, no. 3, pp. 343-348, 1988.
- [19] Sonehara, N., et al. "Image data compression using a neural network model" *Proc. IJCNN*, Vol. 2, 1989.
- [20] Namphol, Aran, Steven H. Chin, and Mohammed Arozullah. "Image compression with a hierarchical neural network" *IEEE transactions on Aerospace and Electronic Systems*, vol. 32, no. 1, pp. 326-338, 1996.

- [21] Russo, L. E., and E. C. Real. "Image compression using an outer product neural network" *Acoustics, Speech, and Signal Processing*, vol. 2, no. 1, pp. 377-380, 1992.
- [22] Abdel-Wahhab, Osama, and Maher A. Sid-Ahmed. "A new scheme for training feed-forward neural networks" *IEEE Pattern recognition*, vol. 30, no. 3, pp. 519-524, 1997.
- [23] Irijanti, E., V. V. Yap, and M. Y. Nayan. "Neural network for the best wavelet selection on colour image compression" *IEEE Information Technology*, vol. 3, no. 1, pp. 1-8, 2008.
- [24] Watanabe, Eiji, and Katsumi Mori. "A compression method for color images using multilayered neural networks" *Systems and Computers in Japan*, vol. 34, no. 2, pp. 13-22, 2003.
- [25] Harashima, H. "Image data compression" (1991).
- [26] YZ Watkins, MR Sayeh. " Image Data Compression and Noisy Channel Error Correction Using Deep Neural Network" *Procedia Computer Science*, vol. 95, no. 1, pp. 145-152, 2016.
- [27] F Hu, C Pu, H Gao, M Tang. " An image compression and encryption scheme based on deep learning", *CA*, vol. 1, pp. 1-6, 2016.
- [28] Bengio, Yoshua. " Learning deep architectures for AI" *Foundations and trends in Machine Learning*, vol. 2, no. 1, pp. 1-127, 2009.
- [29] Bengio, Yoshua. " Deep learning of representations: Looking forward" *International Conference on Statistical Language and Speech Processing*, vol. 1, no. 1, pp. 1-37, 2013.
- [30] Gonzalez, Rafael C., and Richard E. Woods. "Digital image processing" 1992.
- [31] Bengio, Yoshua, Pascal Lamblin, Dan Popovici and Hugo Larochelle. " In Advances in neural information processing systems, vol. 1, no. 1, pp. 153-160, 2007.
- [32] Sutskever, Ilya, et al. " On the importance of initialization and momentum in deep learning" *International conference on machine learning*, vol. 1, no. 1, pp. 1139-1147, 2013.

- [33] Bengio, Yoshua, Aaron Courville, and Pascal Vincent. " Representation learning: A review and new perspectives" *IEEE transactions on pattern analysis and machine intelligence*, vol. 35. No. 8, pp. 1798-1828, 2013.
- [34] Rahman, A., and Chowdhury Mofizur Rahman. " A new approach for compressing color images using neural network" *Processing of CIMCA*, vol. 12, no. 14, pp. 2003.
- [35] www.researchgate.net/figure/284113213_Block-diagram-of-an-image-compression-system.
- [36] A. Khashman, K. Dimililer, "Comparison Criteria for Optimum Image Compression" *IEEE*, vol. 2, pp. 935-938, 2005.
- [37] www.slideshare.net/mravendi/intro-deep-learning.
- [38] Krizhevsky A, Sutskever I, Hinton GE, editors "Imagenet classification with deep convolutional neural networks" *Advances in neural information processing systems*, pp.1097-1105, 2012.
- [39] Vaithyanathan, Shivakumar, and Byron Dom "Model selection in unsupervised learning with applications to document clustering" *ICML*, pp. 433-443, 1999.
- [40] www.deeplearning.ir .
- [41] www.anarcestes69.rssing.com/chan-67243744/all_p7.html .
- [42] www.gisman.ir/image-processing-filter-2/ .
- [43] Eltaweel, S. G. "Color image compression", *IEEE*, vol.1, pp. 83-86, 2011.
- [44] <http://www.noavard.co/> عکس-پروفایل-دانلود-عکس-پروفایل-جالب
- [45] <https://www.tanoco.ir/datamining/> یادگیری-عمیق

Abstract

The increasing growth of digital data in recent years has led to increased attention to compression. The emergence of the phenomenon of compression of image with the neural network introduces us a promising way of developing new effective compression techniques. The main purpose of this phenomenon is to reduce data redundancy. So far, good improvements have been made to image and video compression techniques, and improvements have been made to conserve memory for the transmission of visual information. In this thesis, the Deep Neural Network (DNN) and Noise Reduction Filter for compression of natural color images are proposed. Simulation results show that our proposed method is effective in reducing image size and compression rate, as well as improving the image quality compared other methods. For Instance, in the prposed method which uses noise-cancelling filter, the PSNR value has increased from 22.66db to 24.3db for lena image and from 19.73db to 22.10db for mandrill image compared to image compression method based on neural network(based on similar pixels) at an equal compression ratio. The prposed method has enhanced both image quality and the values.

Keywords: color image compression, deep neural network, autoencoder, noise-cancelling filter



Shahrood University of Technology

Faculty of Electrical and Robotics Engineering

MSc Thesis in Electronics

**An improvement in natural color image compression based
on neural networks**

By: Ali Kargaran

Supervisor:

Dr. Hadi Grailu

January 2018