



فشرده‌سازی تطبیقی تصاویر با استفاده از شبکه‌های عصبی

منصور جم زاد

هادی ویسی

گروه هوش مصنوعی، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شریف

jamzad@sharif.edu

veisi@ce.sharif.edu

۱. مقدمه

به رغم پیشرفتهای چشمگیری که در زمینه وسایل ذخیره سازی و سرعت خطوط انتقال داده وجود داشته است، مساله فشرده‌سازی داده‌ها هنوز به عنوان یکی از مسایل کلیدی و جدی باقی مانده است. در فشرده‌سازی تصاویر به عنوان یکی از داده‌های پرکاربرد، در کنار روشهای تجاری و مرسوم، استفاده از روش‌های جدید نیز برای نشان دادن توانایی‌های آنها و به عنوان مسیری فراروی آینده کد کردن مورد بررسی قرار گرفته است. در میان این روشها، شبکه‌های عصبی مصنوعی با توجه به موفقیت‌های خوبی که در برخی کاربردها داشته است مورد توجه قرار گرفته است. قابلیت‌های ساختار موازی، توانایی یادگیری، استخراج تبدیل‌ها، تقریب بهینه و مقاوم بودن به نویز باعث محبوبیت و استفاده زیاد آنها شده است. انواع مختلف شبکه‌های عصبی مثل پرسپترون چند لایه^۱ [۷-۱]، هب^۲ [۷]، هاپفیلد^۳ [۱۰]، LVQ^۴ و خود سازمان یافته کوهنن^۵ [۱۱، ۱۲] و شبکه‌های

چکیده: در این مقاله به استفاده از شبکه‌های عصبی به منظور فشرده‌سازی تصاویر به صورت تطبیقی بر اساس انتروپی آنها، پرداخته شده است. یک شبکه پرسپترون سه لایه برای فشرده سازی تصاویر بلوک بندی شده به عنوان شبکه مبنا مورد استفاده قرار گرفته است. برای بهبود این ساختار از شبکه‌های مختلف با نرخ فشرده سازی‌های متفاوت برای تصاویر با میزان پیچیدگی‌ها و جزئیات مختلف، می‌توان استفاده کرد. برای محاسبه میزان پیچیدگی تصویر از انتروپی به عنوان یک معیار مفهومی برای تخمین اطلاعات موجود در تصویر استفاده شده است. نتایج حاصل از فشرده‌سازی با این ساختار با همپوشانی بلوک‌های تصویر و بدون آن به همراه انتخاب نوع شبکه‌ها براساس معیار محاسبه انتروپی و براساس معیار بهترین نسبت سیگنال به نویز، بررسی شده‌اند. PSNR این نتایج با روش تجاری و کارای JPEG در نرخ فشرده‌سازی یکسان، مقایسه شده‌اند. نتایج حاصله بهبود خوبی را بر اساس این معیار و بعلاوه در کیفیت بصری تصاویر و افشرده شده نشان می‌دهند.

واژه‌های کلیدی: فشرده سازی تصاویر، شبکه‌های عصبی

پرسپترون چند لایه، انتروپی، فشرده‌سازی تطبیقی.

^۱ - Multi-Layer Perceptron (MLP)

^۲ - Hebb

^۳ - Hopfield

^۴ - Learning Vector Quantization

^۵ - Self-Organizing Map (SOM)

عصبی اجزای اصلی^۱ [۱۳] در زمینه فشرده سازی تصاویر مورد استفاده قرار گرفته‌اند. از میان آنها شبکه‌های پرسپترون چند لایه به عنوان یک ساختار ساده و کارا بیشتر مورد توجه بوده‌اند. یکی از کارهای اولیه در این زمینه شبکه ای سه لایه است که در [۱] ارائه شده است. در این روش مشابه بسیاری از روشهای فشرده سازی دیگر، تصویر ورودی ابتدا به بلوکهایی با اندازه کوچکتر تقسیم می‌شود و سپس هر بلوک به لایه ورودی داده می‌شود تا فشرده شده آن در لایه وسط و بلوک بازسازی شده تصویر در لایه خروجی قرار گیرد. پیاده سازی این روش که روی یک کامپیوتر موازی انجام گرفته است، نشان دهنده کارایی نه چندان خوب این روش حتی برای تصاویر آموزش دیده است. به طوری که تصویر بازسازی شده با این روش فشرده سازی برای نرخ فشرده سازی چهار برابر، حتی دارای نرخ سیگنال به نویز^۲ کمتری نسبت به حالتی است که به جای کل بلوک، میانگین آن فرستاده شود و در مقصد از آن به جای کل پیکسل‌های آن بلوک استفاده شود [۶]. از کارهایی که در ادامه برای بهبود این ساختار انجام شده است می‌توان به ساختار سلسله مراتبی [۸] اشاره کرد که در آن از دو لایه مخفی بیشتر استفاده شده است تا همبستگی بین بلوکها در دو لایه مخفی خارجی و همبستگی بین پیکسل‌ها در داخلی ترین لایه در نظر گرفته شود. اما این ساختار هم بهبود چندانانی را به دنبال نداشته است. گروه دیگری از این تلاشها روش‌های تطبیقی است که برای فشرده سازی بلوک‌های مختلف از شبکه‌های مختلف با نرخ فشرده سازی‌های مختلف با توجه به میزان جزئیات و پیچیدگی تصویر استفاده می‌کنند. نمونه‌هایی از این کارها نتایج خوبی را به همراه داشته است. در [۶] روشهای مختلفی برای آموزش چنین ساختاری پیشنهاد شده است. در یکی از این روشها [۲] بلوک‌های تصویر را بر اساس معیاری که به عنوان فعالیت^۳ بلوک تعریف می‌کنند، به چهار دسته تقسیم کرده و بر اساس آن چهار شبکه مختلف با نرخهای فشرده سازی متفاوت برپا می‌کنند و تصاویر را با آن شبکه‌ها فشرده

می‌نمایند. این ساختار که نتایج خوبی را به همراه داشته با یک تقسیم بندی سطح دوم، بلوک‌های با فعالیت بالا را بر اساس جهت بلوک، به چهار دسته دیگر تقسیم می‌کنند تا کلاً از شش شبکه مختلف استفاده نمایند. این نحوه تقسیم بندی به نوبه خود باعث بهبود بصری بیشتری در تصویر شده است. یکی دیگر از بهترین نتایج بدست آمده از این شبکه‌ها در [۵] ارائه شده است که در آن از نه شبکه مختلف برای فشرده سازی با دسته بندی تصاویر با توجه به جهات بلوک‌ها و بر اساس کاهش میزان لبه آنها استفاده شده است. در این ساختار تصاویر بعد از کم کردن هر پیکسل تصویر از میانگین سطح خاکستری بلوکی که تصویر در آن قرار دارد، به شبکه‌ها داده می‌شود تا فشرده شود که این کار باعث مقداری سربراف اضافی برای این روش نسبت به موارد مشابه است.

در این مقاله ما از معیار انتروپی به عنوان تخمینی از میزان اطلاعات موجود در تصویر برای محاسبه سطح جزئیات آن استفاده می‌کنیم و بر اساس این معیار از شبکه‌های مختلفی جهت فشرده سازی تصویر ورودی استفاده می‌نماییم. در فاز استفاده برای فشرده سازی یک بلوک از تصویر می‌توان یکی از شبکه‌ها را بر اساس معیار تعریف شده انتخاب کرد و بلوک تصویر را با آن شبکه فشرده نمود. یک راه بهتر این است که برای هر بلوک از تصویر بهترین شبکه را بر اساس برخی معیارهای خطا، مثل نسبت سیگنال به نویز، انتخاب کنیم و با آن شبکه تصویر را فشرده نماییم. روش دیگر برای بهبود کیفیت تصویر بازسازی شده همپوشانی دادن بلوک‌های تصویر است که به همراه موارد فوق در اینجا مورد بررسی قرار می‌گیرد و نتایج حاصل با روش فشرده سازی استاندارد JPEG مقایسه می‌شوند.

این مقاله شامل سه بخش دیگر است، در بخش ۲ به استفاده از شبکه‌های عصبی پرسپترون چند لایه در فشرده سازی تصاویر به صورت ساده و تطبیقی پرداخته می‌شود. در بخش ۳ نتایج حاصل از پیاده سازی‌ها و مقایسه انجام شده برای تعیین کارایی الگوریتم جدید با برخی معیارهای معمول فشرده سازی

¹ - Principal Component Neural Networks

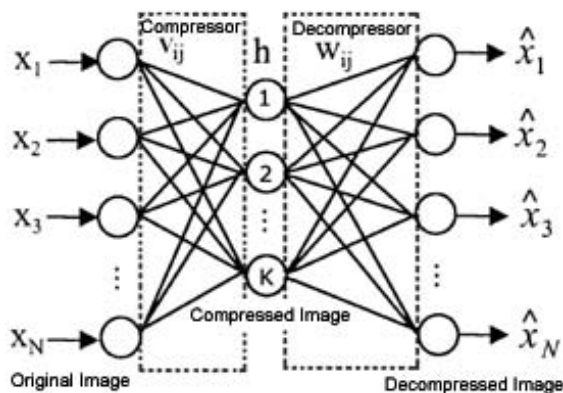
² - Signal to Noise Ratio (SNR)

³ - Activity

$$Hin_j = \sum_{i=1}^N V_{ij} X_i, h_j = f(Hin_j), 1 \leq j \leq K, \quad (1)$$

$$\hat{X} in_i = \sum_{j=1}^K W_{ij} h_j, \hat{X}_i = g(\hat{X} in_i), 1 \leq i \leq N, \quad (2)$$

همانطور که در شکل ۱ هم نشان شده است، X_i پیکسل‌های تصویر اصلی و \hat{X}_i پیکسل‌های تصویر بازسازی شده از پیکسل‌های تصویر فشرده شده h_j می‌باشد. توابع f و g نیز می‌توانند خطی یا غیر خطی باشند.



شکل ۱: شبکه عصبی پرسپترون سه لایه به عنوان شبکه مبنای فشرده سازی تصاویر

کاری که عملاً توسط این ساختار انجام می‌شود تخمینی است که شبکه از تبدیل تحلیل اجزای اصلی^۳ بردارهای تصویر می‌زند. این تبدیل که با اسامی دیگری چون هتلینگ^۴، کارهانن-لوه^۵ و بردارهای ویژه نیز شناخته می‌شود، تبدیلی بهینه در کارهای فشرده سازی و کاهش بعد است و خطای میانگین مربعات حاصل از عمل فشرده سازی را کمینه می‌کند. هر چند این تبدیل به علت وابسته بودن به داده ورودی، جای خود را در الگوریتم‌های کاربردی و عملی، به تبدیل گسسته کسینوسی^۶ داده است اما در صورت کافی بودن داده ورودی می‌تواند تبدیل بهینه را استخراج نماید.

^۳ - Principal Component Analysis (PCA)

^۴ - Hostelling Transform

^۵ - Karhunen-Loeve Transform (KLT)

^۶ - Discrete Cosine Transform (DCT)

مورد بررسی قرار می‌گیرد و در بخش ۴ جمع بندی و خلاصه کار ارایه می‌شود.

۲. استفاده از شبکه‌های عصبی پرسپترون در فشرده

سازی تصاویر

ساده ترین ساختار برای کار فشرده سازی تصویر با یک شبکه سه لایه پرسپترون به صورت نشان داده شده در شکل ۱ می‌باشد. برای استفاده از این ساختار تصاویر ورودی بایستی ابتدا به بلوک‌هایی با تعداد پیکسل‌های برابر تعداد نرون‌های لایه ورودی، N ، تقسیم شود. یعنی بلوک‌های تصویری بایستی در ابعاد $\sqrt{N} * \sqrt{N}$ باشند تا به صورت بردار N بعدی درآیند و به ورودی شبکه داده شوند. لایه مخفی این ساختار نشان دهنده تصویر فشرده شده بلوک ورودی است که از N پیکسل به K پیکسل ($K \leq N$) فشرده شده است. نهایتاً تصویر بازسازی شده از تصویر فشرده شده، در لایه خروجی با همان تعداد پیکسل/نرون، N ، بدست می‌آید. در این ساختار وزن‌های لایه ورودی به لایه مخفی $\{V_{ij}, 1 \leq i \leq N, 1 \leq j \leq K\}$ ماتریس تبدیلی است که کار فشرده سازی را انجام می‌دهد. به طور مشابه وزن‌های لایه مخفی به لایه خروجی $\{W_{ij}, 1 \leq i \leq K, 1 \leq j \leq N\}$ تبدیل و فشرده کننده^۱ تصویر فشرده شده است. آموزش این ساختار با الگوریتم پس انتشار خطا^۲ انجام می‌گیرد که برای این کار لازم است ابتدا بازه سطوح خاکستری پیکسل‌های تصویر به بازه $[0-1]$ مقیاس شود و برای همه الگوهای مجموعه آموزش بلوک‌های ورودی به عنوان هدف به لایه خروجی هم داده شود تا با توجه به آن خطای شبکه برای الگوی جاری بدست آید. به این ترتیب الگوریتم آموزش، وزن‌های لایه آخر را بر اساس این خطاها و وزن‌های لایه اول را با انتشار مقادیر این خطاها به این لایه تنظیم می‌کند. عملکرد این شبکه را می‌توان به صورت روابطی که در زیر آورده شده‌اند، بیان کرد:

1 - Decompress

2 - Error Back Propagation

نرخ فشرده سازی^۱ که به عنوان یک معیار در مسایل فشرده سازی مورد استفاده قرار می‌گیرد، طبق تعریف برابر با نسبت بیت‌های تصویر اولیه به بیت‌های تصویر فشرده شده می‌باشد. این معیار در ساختار شبکه عصبی مورد استفاده ما بصورت رابطه زیر تعریف می‌شود:

$$CR = \frac{N B_I}{K B_H} \quad (۳)$$

در این رابطه N و K همان تعداد نرونها/پیکسل‌های لایه‌های ورودی و مخفی هستند و B_I و B_H به ترتیب تعداد بیت‌های لازم برای کد کردن خروجی‌های لایه ورودی و لایه مخفی می‌باشند. اگر تعداد بیت‌های لازم برای کد کردن تصویر فشرده شده در لایه مخفی را برابر تعداد بیت‌های ورودی بگیریم نرخ فشرده سازی برابر نسبت تعداد نرونهای لایه ورودی به تعداد نرونهای لایه مخفی خواهد بود. مثلاً برای تصویر سطوح خاکستری که ۸ بیتی هستند، اگر تصویر فشرده شده را نیز با همین تعداد بیت کد کنیم برای یک بلوک ۸×۸ که به شبکه ای با ۱۶ نرون مخفی داده شود، نرخ فشرده سازی ۴:۱ خواهد بود. در صورتی که برای همان شبکه از اعداد ممیز شناور ۳۲ بیتی برای کد کردن تصویر فشرده شده استفاده شود این نرخ ۱:۱ خواهد بود که در واقع هیچ فشرده سازی اتفاق نیفتاده است. مساله تعداد بیت‌های لازم جهت کد کردن تصویر فشرده شده در لایه مخفی، بوسیله محققان دیگری مورد بررسی قرار گرفته است که از ۲ تا ۱۰ بیت برای این کار در نظر گرفته شده است. بعلاوه نشان داده شده است که استفاده از بیش از ۶ بیت برای این کار بهبود چندانی در کارایی بوجود نمی‌آورد [۱].

۱-۲. روش تطبیقی مبتنی بر انتروپی

نتایج حاصل از بکارگیری شبکه پایه فوق نشان می‌دهد که کارایی این ساختار ساده در عمل ضعیف تر از آنست که بتواند مورد استفاده قرار گیرد. روشی که در اینجا برای بهبود

کارایی این ساختار ارائه شده است بدین صورت عمل می‌کند که از شبکه‌های مختلفی برای فشرده سازی بلوک‌های مختلف تصویر استفاده شده می‌نماید. ابتدا بلوک‌های تصویر با توجه به میزان انتروپی آنها به چند گروه تقسیم می‌شوند و هر کدام از گروه‌ها به شبکه ای جداگانه آموزش داده می‌شود. تفاوت شبکه‌ها در نرخ فشرده سازی تصویر یا همان تعداد نرون‌های لایه مخفی است. این تعداد متناسب با میزان اطلاعات موجود بلوک‌های تصویری آن گروه انتخاب می‌شود به صورتی که تصاویر با میزان اطلاعات و سطح جزئیات بیشتر کمتر فشرده می‌شوند تا در بازسازی تصویر داده کمتری گم شود تا در نتیجه خطای فشرده سازی کاهش یابد. با در نظر گرفتن شبکه نشان داده شده در شکل ۱ به عنوان یک ساختار پایه، می‌توان رویکرد تطبیقی را به صورت شکل ۲ نشان داد. انتخاب شبکه‌ها که بر اساس معیار پیچیدگی هر بلوک تصویر انجام می‌شود، با محاسبه انتروپی پیکسل‌های آن بلوک به صورت رابطه ۴ می‌باشد:

$$Entropy = -\sum_{i=1}^L P(x_i) \log P(x_i) \quad (۴)$$

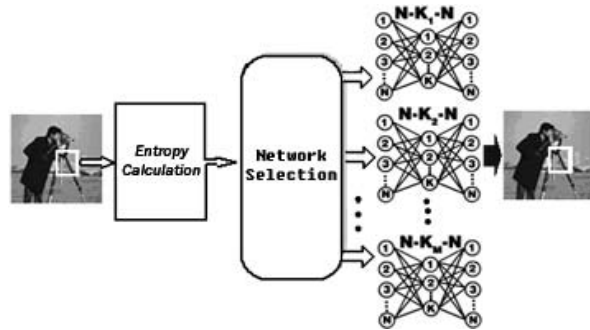
در رابطه فوق $P(x_i)$ احتمال رخداد سطح خاکستری x_i در بلوکی با L سطح خاکستری مختلف است. برای آموزش این ساختار ابتدا انتروپی هر کدام از بلوک‌های تصویر مجموعه آموزش محاسبه شده و بعد از انتخاب شبکه مربوطه به لایه ورودی آن شبکه داده می‌شود تا مشابه ساختار پایه شکل ۱ با الگوریتم پس انتشار خطا آن شبکه را آموزش دهد. این کار برای هر الگوی آموزشی انجام می‌شود. در فاز استفاده از شبکه‌های آموزش دیده بایستی برای هر بلوک تصویر ورودی نوع شبکه ای که قرار است با آن آموزش داده شود انتخاب شود و شماره آن شبکه به عنوان سرباری از ذخیره سازی، به همراه تصویر فشرده شده بلوک ذخیره شود تا در مرحله بازسازی تصویر اصلی از تصویر فشرده شده، مشخص باشد که از کدامیک از شبکه‌ها استفاده شود. در این حالت نرخ فشرده سازی با فرض وجود M شبکه مختلف با تعداد نرونهای مخفی K_1 تا K_M برای

^۱ - Compression Ratio (CR)

تصویری با n بلوک N پیکسلی، با توجه به رابطه ۳، به صورت زیر خواهد بود:

$$CR_a = \frac{NB_l}{\left(\frac{B_H}{n} \sum_{i=1}^n K_{ij}\right) + q}, \quad 1 \leq j \leq M \quad (5)$$

در این رابطه K_{ij} بیانگر یکی از M تا K_j برای بلوک i ام است و q تعداد بیت‌هایی است که برای کد کردن شماره شبکه لازم می‌باشد. بعبارت دیگر q برابر کوچکترین عدد صحیح مثبتی است که: $2^q \geq M$.



شکل ۲: ساختار تطبیقی شبکه عصبی پرسپترون برای فشرده سازی تصاویر

۳. نتایج پیاده سازی

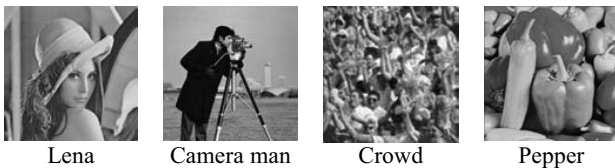
در این بخش نتایج حاصل از پیاده سازی روشهای فشرده سازی با ساختار مبنای شکل ۱ و ساختار تطبیقی شکل ۲ مورد بررسی قرار می‌گیرند و نتایج حاصل با الگوریتم فشرده سازی JPEG مقایسه می‌شوند. معیارهایی که برای سنجش کارایی این الگوریتم‌ها به کار برده می‌شوند، نرخ فشرده سازی و قله نسبت سیگنال به نویز^۱ است. محاسبه نرخ فشرده سازی برای دو ساختار مورد بررسی در روابط ۳ و ۵ آمده‌اند، قله نسبت سیگنال به نویز هم برای تصویری با $Rows$ سطر و $Cols$ ستون به صورت زیر تعریف می‌شود:

$$PSNR = 10 \log_{10} \left(\frac{255^2}{\frac{1}{Rows \times Cols} \sum_{i=1}^{Rows} \sum_{j=1}^{Cols} (X_{ij} - \hat{X}_{ij})^2} \right) \quad (6)$$

^۱ - Peak Signal to Noise Ratio (PSNR)

در این رابطه ۲۵۵ بیانگر ماکزیمم سطح خاکستری موجود در تصویر ۸ بیتی ورودی است. این معیار اگرچه کیفیت تصویر بازسازی شده را تا حدی جدا از آنچه از نظر بینایی قابل تشخیص است، معین می‌کند اما به علت سادگی آن، در اغلب موارد مورد استفاده قرار می‌گیرد.

در هر دو ساختار مورد نظر از بلوک‌های تصویر 8×8 استفاده شده است و در نتیجه لایه‌های ورودی و خروجی همه شبکه‌ها ۶۴ نرونی هستند. در ساختار پایه شکل ۱ از تعداد ۱۶ نرون در لایه مخفی استفاده شده است که با توجه به ۸ بیتی بودن تصاویر ورودی و در نظر گرفتن همین تعداد بیت برای کد کردن تصویر فشرده شده در لایه مخفی، نسبت ثابت فشرده سازی ۴:۱ را نتیجه می‌دهد. در این ساختار برای آموزش شبکه، از تصویر 256×256 Lena که در شکل ۳ نشان داده شده است، استفاده گردیده که بدون همپوشانی دادن بلوک‌ها تشکیل ۱۰۲۴ الگوی آموزشی می‌دهند. نتایج آزمایش تصاویر 256×256 شکل ۳ با این شبکه با نرخ فشرده سازی ثابت ۴:۱ در جدول ۱ نشان داده شده است. آنچه از این نتایج مشهود است و قابل پیش بینی هم می‌باشد، این است که این ساختار برای تصویر آموزش دیده به نسبت سایر موارد نتیجه خوبی می‌دهد. همانطور که در مقدمه هم اشاره شد نتایج حاصل از این ساختار، مخصوصاً برای تصاویر غیر آموزش دیده، دارای کیفیت قابل قبولی نیست.

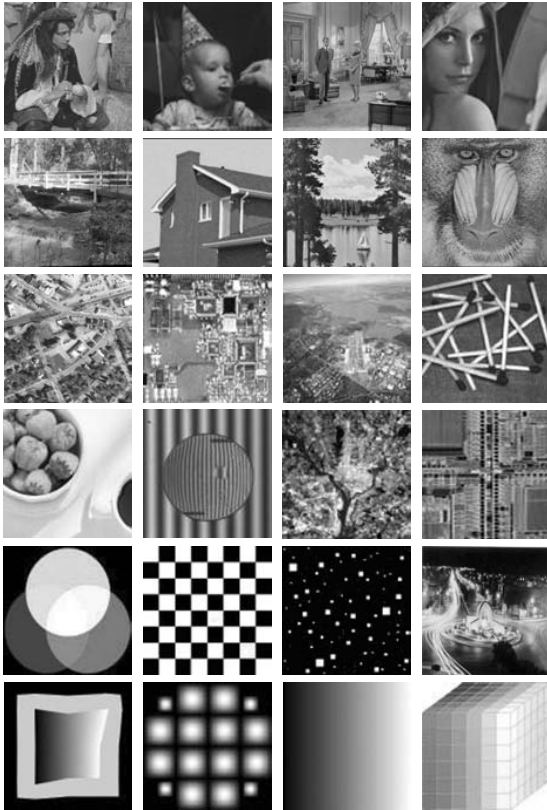


شکل ۳: تصاویر مورد استفاده برای تست الگوریتم‌ها

جدول ۱: نتایج فشرده سازی تصاویر شکل ۳ با ساختار شبکه عصبی

سه لایه پرسپترون

Test image ▼	PSNR (dB)
Lena	34.92
Camera man	26.67
Crowd	23.41
Pepper	22.20



شکل ۴: نمونه هایی از تصاویر مجموعه آموزشی برای آموزش شبکه های با ساختار تطبیقی

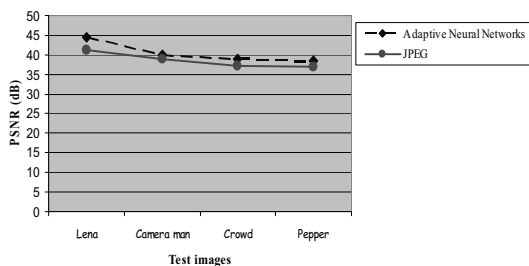
همانطور که نتایج جدول ۲ نیز نشان می دهد این کار علاوه بر بهبود کیفیت بصری باعث افزایش PSNR تا حدود ۱۵dB شده است. یکی دیگر از راه هایی که برای انتخاب بهینه شبکه ها وجود دارد این است که برای هر بلوک تصویر از میان شش شبکه، بدون توجه به معیار انتروپی، شبکه ای را انتخاب کنیم که معیار خطای خاصی را کمینه نماید. معیاری که در اینجا برای انتخاب بکار برده شده است ماکزیمم کردن نسبت سیگنال به نویز است. این کار اگرچه زمان فشرده سازی را بالا می برد اما منجر به بهینه کردن نسبت سیگنال به نویز می شود که از این نظر خطای فشرده سازی را کمینه می کند. همانطور که در جدول ۲ دیده می شود، نتایج انجام این کار بهبود قابل ملاحظه ای در معیار PSNR بوجود آورده است، این بهبود برای کیفیت بصری تصاویر

برای بهبود ساختار پایه با استفاده از معیار محاسبه اطلاعات انتروپی، کل بلوک های مجموعه آموزشی به شش شبکه مختلف با نرخ های فشرده سازی متفاوت داده می شود. با توجه به میزان انتروپی بلوک های تصویر، بلوک های با انتروپی بیشتر به شبکه های با نرخ فشرده سازی کمتر و بلوک های با انتروپی کمتر به شبکه های با نرخ فشرده سازی بیشتر داده می شوند. برای آموزش این ساختار از تعداد زیادی تصویر (۱۰۵) تصویر 256×256 استفاده شده است. اگرچه می توان این شبکه ها را با تعداد تصاویر کمتر از این هم آموزش داد اما برای تنظیم خوب وزنها با تغییرات زیاد این تعداد انتخاب شده است. هر چند این کار خود منجر به بالا بردن زمان آموزش شبکه ها گردید. تصاویر مورد استفاده به عنوان مجموعه آموزشی در این حالت از دادگان استاندارد انتخاب نشده است بلکه با توجه به ماهیت شبکه ها انتخاب شده اند به طوری که تصاویر انتخابی این خاصیت را دارند که تعداد نمونه های آموزشی برای همه شبکه های مورد استفاده تقریباً برابر باشند و شبکه ها بصورت مشابه آموزش داده شوند. نمونه هایی از تصاویر بکار گرفته شده برای آموزش دادن شبکه ها در این حالت در شکل ۴ آورده شده است.

در این مرحله با در اختیار داشتن شبکه های آموزش دیده، می توان برای بلوک تصویر ورودی با استفاده از محاسبه انتروپی آن یکی از شبکه ها را برای فشرده کردن بلوک در نظر گرفت. نتایج این کار برای تصاویر تست شکل ۳ در جدول ۲ با اسم انتخاب نوع شبکه بر اساس معیار انتروپی آورده شده اند. نرخ فشرده سازی CR_q از روی رابطه ۶ با $M = 6$ محاسبه شده است و در این جدول به صورت $1/CR_q$ آورده شده است. در اینجا نیز تصویر Lena جزو داده های آموزشی بوده است. برای کم کردن اثر جعبه ای^۱ تصویر بازسازی شده و بالاتر بردن کیفیت بصری آن، می توان بلوک های تصویر را با هم همپوشانی داد. این مساله در اینجا با همپوشانی های ۵۰٪ برای سطرها و ستون ها مورد آزمایش قرار گرفته است بطوریکه در نواحی مشترک از میانگین سطوح خاکستری بلوک ها استفاده شده است.

^۱ - Blocking effect

است. به علاوه کیفیت بصری تصاویر بازسازی شده نیز در این رویکرد بهتر از تصاویر JPEG می‌باشد. مقایسه نتایج بصری این رویکرد و روش JPEG برای دو تصویر *Pepper* و *Camera man* از مجموعه تست در شکل ۶ آورده شده است. هرچند در این مقایسه کیفی با توجه به بالا نسبی بودن PSNR تصاویر در هر دو حالت، تفاوت دو روش به مشابه مقایسه کمی آنها واضح نیست.



شکل ۵: مقایسه نتایج PSNR تصاویر تست فشرده شده با استفاده از شبکه عصبی تطبیقی و JPEG



ب

الف



ب

الف

شکل ۶: نتایج فشرده/وافشرده سازی تصاویر *Pepper* و *Camera man* از مجموعه تست با رویکرد تطبیقی با بهترین SNR و همپوشانی (الف) روش JPEG (ب)

وافشرده شده حاصل نیز قابل ملاحظه است. با انجام این کار نرخ فشرده سازی مقداری کاهش یافته است که از نظر عملی هم قابل توجه است. همپوشانی دادن بلوک‌های تصویری در این حالت نیز منجر به بالا بردن معیار PSNR و افزایش کیفیت تصویر بازسازی شده گردیده است.

جدول ۲: نتایج حاصل از فشرده سازی تصاویر با ساختار تطبیقی

شکل ۲ برای تصاویر تست مختلف و انتخاب نوع شبکه‌ها با معیارهای انترپی و بهترین SNR با همپوشانی و بدون همپوشانی بلوک‌ها.

انتخاب شبکه	Test image	PSNR (dB)	1/CR _a
انترپی بدون همپوشانی	Lena	35.50	0.36
	Camera man	35.11	0.35
	Crowd	37.92	0.58
	Pepper	35.62	0.45
انترپی با همپوشانی	Lena	36.70	0.36
	Camera man	36.57	0.35
	Crowd	38.68	0.59
	Pepper	36.25	0.45
بهترین SNR بدون همپوشانی	Lena	43.10	0.56
	Camera man	39.24	0.49
	Crowd	38.11	0.59
	Pepper	38.06	0.56
بهترین SNR با همپوشانی	Lena	44.55	0.56
	Camera man	40.04	0.49
	Crowd	38.97	0.59
	Pepper	38.48	0.57

برای اینکه برآورد بهتری از عملکرد الگوریتم بدست آید، نتایج حاصل از رویکرد فوق در بهترین حالت با ترکیب بهترین SNR و با وجود همپوشانی بین بلوک‌ها، با نمونه‌های فشرده شده همان تصاویر در همان نرخ فشرده سازی بوسیله الگوریتم JPEG در شکل ۵ با هم مقایسه شده‌اند. برای همه تصاویر مجموعه تست، این رویکرد منجر به بهبود PSNR شده است. با توجه به اینکه تصویر *Lena* در مجموعه آموزشی بوده است، نتایج بهتری برای آن به نسبت سایر تصاویر بدست آمده است. هرچند بعلت زیاد بودن داده‌های آموزشی و عمومیت آموزش شبکه‌ها، وابستگی روش به داده‌های آموزش کمتر شده است ولی برای تصاویر غیر آموزش دیده هم نتایج نسبتاً مشابهی حاصل شده

۴. جمع بندی و خلاصه

مساله استفاده از شبکه‌های عصبی پرسپترون برای فشرده سازی تصاویر مورد بررسی قرار گرفت. با توجه به کارایی نه چندان مطلوب استفاده از تنها یک شبکه برای فشرده سازی تصاویر، رویکردی تطبیقی بر اساس معیار محاسبه اطلاعات موجود در تصویر پیشنهاد شد. برای محاسبه اطلاعات و سطح جزئیات تصویر از معیار انتروپی استفاده شد. نتایج حاصل از بکارگیری این رویکرد در بهترین حالت که شامل همپوشانی بلوک های تصویر و استفاده از بهترین نرخ سیگنال به نویز برای انتخاب شبکه فشرده کننده آن بلوک است، با الگوریتم فشرده سازی JPEG مقایسه شده است که بهبود خوبی را با معیار PSNR نشان می‌دهد. در این مقاله از شش شبکه با نرخهای فشرده سازی متفاوت استفاده شده است که انتظار می‌رود در صورت استفاده از تعداد بیشتر شبکه و انتخاب بهینه نرخ‌های فشرده سازی شبکه‌ها، نتایج مطلوب تری نیز حاصل شود. جدای از راههای حریصانه^۱ و یا آزمون و خطا، می‌توان از شبکه‌های عصبی خاصی مثل همبستگی آبخاری^۲ که تعداد نرونهای لایه مخفی و به تبع آن نرخ فشرده سازی هر شبکه را به صورت بهینه انتخاب می‌کنند، بهره گرفت. استفاده از معیارهای محاسبه پیچیدگی بهینه تر و یا ترکیب چند معیار مختلف با هم، می‌تواند منجر به افزایش کارایی این ساختار شود. نظر به اینکه معیار انتروپی وابستگی اطلاعات کل بلوک تصویر بدون توجه به محل پیکسل‌ها در بلوک را اندازه گیری می‌کند، بنابراین برای متمایز نمودن بلوکی که نصف آن سفید و نصف دیگرش سیاه است با بلوکی که شطرنجی سیاه و سفید است، درست عمل نمی‌کند و انتروپی یکسانی را برای هر دو محاسبه می‌نماید در حالیکه برای این ساختار متمایز کردن این نوع تصاویر از همدیگر برای بهبود کارایی، مناسب خواهد بود.

منابع

- [1] N.Sonehara, M.Kawato, S.Miyake, K.Nakane, "Image compression using a neural network model", International Joint Conference on Neural Networks, Washington DC, 1989.
- [2] S. Marsi, G. Ramponi, G. L. Sicuranza, "Improved neural structure for image compression", Proceeding of International Conference on Acoustic Speech and Signal Processing, Toronto, 1991.
- [3] S. Carrato, G. Ramponi, "Improved structures based on neural networks for image compression", IEEE Workshop on Neural Networks for Signal Processing, New Jersey, September 1991.
- [4] S. Carrato, S. Marsi, "Compression of subband-filtered images via neural networks", IEEE Workshop on Neural Networks for Signal Processing, August 1992.
- [5] G. Qiu, M. Varley, T. Terrel, "Image compression by edge pattern learning using multilayer perceptron", Electronic letters, Vol 29, No 7, April 1993.
- [6] C. Cramer, "Neural networks for image and video compression: A review", European Journal of Operational Research, Vol. 108, July 1998.
- [7] J. Jiang, "Image compression with neural networks - A survey", Image Communication, ELSEVIER, Vol. 14, No. 9, 1999.
- [8] A.Namphol, S.Chin, M. Arozullah, "Image compression with a hierarchical neural network", IEEE Trans. Aerospace Electronic Systems Vol. 32 No.1, January 1996.
- [9] J. S. Lin, S.H. Liu, "A competitive continuous Hopfield neural network for vector quantization in image compression", Engineering Applications of Artificial Intelligence, Vol. 12, 1999.
- [10] G. Pavlidis, A. Tsompanopoulos, A. Atsalakis, N. Papamarkos, C. Chamzas, "A Vector Quantization - Entropy Coder Image Compression System", IX Spanish Symposium on Pattern Recognition and Image Processing, 2001.
- [11] C. Amerijckx, J. D. Legaty, M. Verleysenz, "Image Compression Using Self-Organizing Maps", Systems Analysis Modeling Simulation Vol. 43, No. 11, November 2003.
- [12] S. Costa, S. Fiori, "Image compression using principal component neural networks", Image and vision computing, Vol. 19, 2001.

1 - Heuristic

2 - Cascade correlation