# ارائه یک روش جدید دیموزائیک با استفاده از شبکه عصبی

فروزان ذاکریان۱، دکتر علیمحمد لطیف ۲و دکتر علیمحمد اسمعیلی زینی۳
۱ دانشجو، دانشکده فنی مهندسی، دانشگاه علم و هنر، یزد، ایران – foroozan.zakerian@yahoo.com
۲ دانشیار، دانشکده فنی مهندسی ، دانشگاه یزد، یزد، ایران – alatif@yazd.ac.ir
۳ مربی، دانشکده فنی مهندسی ، دانشگاه علم و هنر، یزد، ایران – esmailizaini@gmail.com

## چکیده:

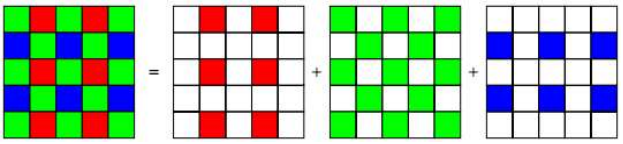
دوربین‌های دیجیتال تصاویر رنگی را با استفاده از یک سنسور پوشیده شده از آرایه فیلتر رنگی دریافت می‌کنند. آرایه فیلتر رنگی در هر پیکسل تنها یک مؤلفه از رنگ را دریافت کرده و دو مؤلفه باقی‌مانده با تکنیک دیموزائیک بدست می‌آیند. تکنیک‌های رایج دیموزائیک بر کیفیت تصویر مصنوعاتی مانند ***Zipper*** و ***False color*** ایجاد می‌نمایند. در این مقاله برای کاهش این مصنوعات پدید آمده یک تکنینک دیموزائیک جدید با استفاده از شبکه عصبی مطرح شده است. شبکه عصبی برای بازیابی مؤلفه‌های سبز ناموجود در کانال رنگی سبز به‌کار برده شده و بازیابی مؤلفه‌های قرمز و آبی ناموجود با استفاده از کانال رنگی سبز بدست آمده، صورت می‌گیرد. به این منظور از دو شبکه عصبی پرسپترون دو لایه و شبکه عصبی ***RBF*** استفاده شده است.ارزیابی و مقایسه نتایج بدست آمده با استفاده از معیار ***PSNR*** نشان می‌دهد که نتایج حاصل از روش ارائه شده نتایج قابل قبولی می‌باشند و این روش می‌تواند جایگزین خوبی برای روش‌های پایه‌ای دیموزائیک قرار بگیرد.

## کلید واژه:

الگوی بایر، درونیابی، دیموزائیک، موزائیک، شبکه عصبی

## ۱– مقدمه

تصاویر رنگی با ترکیبی از سه کانال رنگی (قرمز،سبز و آبی) ایجاد می‌شوند. بنابراین برای ایجاد یک تصویر رنگی، باید سه نمونه رنگ در هر پیکسل وجود داشته باشد. دوربین دیجیتال برای گرفتن تصویر رنگی به سه سنسور پوشیده از فیلتر رنگی نیاز دارد[۱]. اکثر دوربین‌های دیجیتالی در جهت کاهش هزینه و توان مصرفی، به جای استفاده از سه سنسور از یک سنسور که با آرایه فیلتر رنگ پوشیده شده است، استفاده می‌کنند. الگوی(چیدمان) رنگ‌های دریافتی توسط یک سنسور با آرایه فیلتر رنگی مطابق با الگوی بایر و به شکل ۱–۱ می‌باشد. الگوی بایر تصویر سبز را بر روی یک شبکه با آرایش پنج تایی و تصاویر آبی و قرمز را بر روی شبکه‌های مستطیل شکل نمایش می‌دهد. در این الگو نیمی از پیکسل‌های تصویر را، کانال رنگی سبز در بر می‌گیرد و یک چهارم پیکسل‌ها را مؤلفه‌های رنگی قرمز و آبی تشکیل می‌دهند.[۲]. درنتیجه یک سنسور پوشیده شده از آرایه فیلتر رنگی قادر است در هر پیکسل تنها یک رنگ از سه رنگ اصلی را دریافت نماید و مقادیر عددی دو رنگ دیگر، ناموجود می‌باشند. تصویر بدست آمده از این طریق یک تصویر موزائیک نامیده می‌شود و روند بازیابی مقادیر رنگی ناموجود دیموزائیک نام گرفته است.



شکل(۱) الگوی بایر از آرایه فیلتر رنگی بکار برده شده در دوربین دیجیتال تک سنسوری

برای این هدف الگوریتم‌های بکار برده شده به دو دسته تقسیم شده‌اند. دسته اول الگوریتم‌های پایه‌ای دیموزائیک اعم از نزدیک‌ترین همسایه، درونیابی خطی ساده تا درونیابی‌های غیرخطی می‌باشند. این روش‌های درونیابی معمولا با مسائلی مانند پدیده **Zipper** و **False color** روبرو می‌باشند. . برای غلبه بر این مسائل الگوریتم های دسته دوم که از الگوریتم‌های یادگیری ماشین در کنار الگوریتم‌های پایه‌ای بهره می‌برند پدید آمدند.

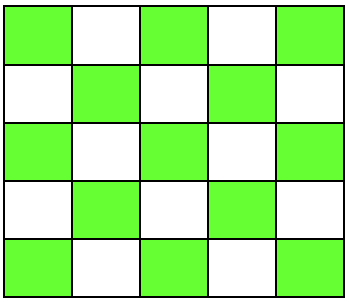
در [۳] ، یک روش خود-آموز جدید با روش **SVR** مطرح شده است. این روش یادگیری، داده آموزشی مورد نیاز خود را از طریق کوچک کردن اندازه تصویر از $m \times n$ به $\frac{m}{2} \times \frac{n}{2}$ و به $\frac{m}{4} \times \frac{n}{4}$ و الی آخر استخراج می‌کند و از درونیابی خطی برای دیموزائیک کردن تصویر و **SVR** برای بهبود تصویر دیموزائیک شده بهره می‌برد.

در [٤]، برای دیموزائیک کردن تصویر از الگوریتم **Edge-Adaptive Demosaicing(EAD)** همراه با روش خود آموز **SVR** استفاده شده است. هدف اصلی استفاده از **EAD** این است که با استفاده از اطلاعات فرکانس بالای کانال سبز، خطا و مصنوعات رنگی ایجاد شده از کانال قرمز و آبی را کاهش دهد. نحوه‌ی کار بدین صورت است که ابتدا برای درونیابی کانال‌های قرمز و آبی از روش خطی و برای درونیابی کانال سبز از یک روش غیرخطی استفاده می‌شود. بعد از درونیابی این صفحات، برای صفحات قرمز و آبی از یک فیلتر پایین گذر و برای صفحه سبز از یک فیلتر بالا گذر استفاده می‌شود. بعد از فیلتر کردن هر صفحه رنگی، با اضافه کردن مؤلفه‌های فرکانس بالا از کانال سبز به کانال‌های قرمز و آبی، این دو کانال مجددا بازسازی می‌شوند. این تصویر دیموزائیک حاصل از الگوریتم **EAD** برای تصحیح و بهبود به **SVR** داده می‌شود. در [۵]، از یک روش تکمیل ماتریس غیر محلی با مرتبه پایین با استفاده از لبه‌یابی و شبکه عصبی استفاده شده است. ابتدا تصویر با وضوح پایین را با استفاده از عملگر دیفرانسیل کسری به سه ناحیه هموار، لبه و ناحیه بافت تقسیم می‌کند. سپس برای نواحی هموار از درونیابی خطی، برای نواحی لبه از درونیابی **edge-directed** و برای نواحی بافت از درونیابی شبکه عصبی استفاده نموده است. نحوه‌ی استفاده از درونیابی شبکه عصبی بدین صورت است که تصاویر در دسترس با وضوح بالا را با استفاده از روش **down-sampling** به تصاویر با وضوح پایین تبدیل می‌کند، سپس بر روی تصویر بدست آمده با وضوح پایین از درونیابی **bicubic** استفاده می‌کند تا تصویر با وضوح بالا بدست آید. از مقایسه تصاویر با وضوح بالا و تصاویر با وضوح پایین، شبکه عصبی آموزش می‌بینند.

## ۲- روش پیشنهادی

### ۲-۱بازیابی مؤلفه‌های سبز

از تصویر موزائیک بدست آمده از دوربین تک سنسوره مطابق با شکل ۱ سه ماتریس با سه رنگ قرمز، سبز و آبی که دارای درایه‌هایی در جایگاه‌هایی مشخص و مطابق با الگوی بایر هستند، در دسترس است و درایه‌ها با مقادیر ناموجود باید بازیابی شوند. از آنجا که رنگ سبز در الگوی بایر بیشترین تعداد موجود پیکسل‌ها را دارا است عملیات بازیابی بر روی این ماتریس، شروع می‌شود. صفحه‌ی سبز به‌این‌صورت است (شکل ۲) :



شکل(٢) صفحه سبز حاصل از الگوی بایر

از شکل ٢ زیر ماتریس هایی ٣×٣ را مطابق شکل ٣ در نظر بگیرید، هدف از آموزش شبکه عصبی یافتن مقدار پیکسل مرکزی یعنی مقدار پیکسل ٥ می‌باشد. بعد از آموزش، شبکه عصبی باید یاد بگیرد که با گرفتن چهار مقدار از داده‌های پیکسل‌های موجود یعنی پیکسل‌های ٢، ٤، ٦ و ٨، مقدار عددی پیکسل وسط را تخمین بزند.

| | | |
|---|---|---|
| ١ | ٢ | ٣ |
| ٤ | ٥ | ٦ |
| ٧ | ٨ | ٩ |

شکل( ٣)زیر ماتریس٣×٣ انتخاب شده برای تعیین مقدار پیکسل سبز وسط

در این مقاله برای بازیابی مؤلفه‌های ناموجود از دو شبکه عصبی، شبکه عصبی پرسپترون چند لایه و شبکه عصبی RBF استفاده شده است.

## ٢-٢ورودی شبکه عصبی

برای آموزش شبکه عصبی، زیر ماتریس هایی ٣×٣ مطابق شکل ٤ به گونه‌ای انتخاب می‌شوند که مقدار پیکسل مرکزی در آن‌ها مشخص باشد یعنی پیکسل‌هایی که مقادیر آن‌ها مشخص است و به‌صورت آرایش شبکه‌ای ٥تایی از رنگ سبز وجود دارند به‌عنوان داده‌ی آموزش درنظر گرفته شده‌اند. مقادیر پیکسل های ١، ٣، ٧ و ٩ داده‌هایی هستند که از آن‌ها به‌عنوان بردار ورودی شبکه عصبی استفاده می‌شود و مقدار پیکسل ٥، نتیجه خروجی می‌باشد، که خروجی شبکه عصبی با مقدار پیکسل ٥ مقایسه می‌گردد و باید شبکه با رسیدن به این خروجی آموزش ببیند.

| | | |
|---|---|---|
| ١ | ٢ | ٣ |
| ٤ | ٥ | ٦ |
| ٧ | ٨ | ٩ |

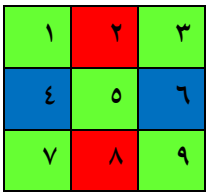
شکل (٤) زیرماتریس انتخاب شده برای آموزش شبکه عصبی

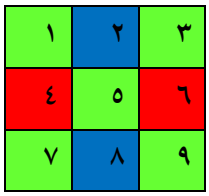
مقادیر عددی پیکسل‌های ١، ٣، ٧ و ٩ به‌صورت مستقیم به شبکه داده نمی‌شوند. از این چهار پیکسل، ویژگی‌هایی محاسبه شده و به‌عنوان ورودی شبکه عصبی مورد استفاده قرار می‌گیرد. شبکه عصبی با سه ورودی مختلف مورد آزمایش و تست قرار گرفته است. ورودی آزمون اول شش ویژگی از پیکسل‌های سبز بوده و به قرار زیر می‌باشد:

١. میانه چهار پیکسل
٢. میانگین چهار پیکسل
٣. میانگین اختلاف پیکسل‌های افقی
٤. میانگین اختلاف پیکسل‌های عمودی
٥. اختلاف پیکسل‌های قطر اصلی
٦. اختلاف پیکسل‌های قطر فرعی

ازآنجاکه در اطراف این پیکسل‌های سبز، پیکسل‌های قرمز و آبی نیز وجود دارند (شکل ٥)، علاوه‌براین شش ویژگی دو ویژگی دیگر مربوط به رنگ‌های قرمز و آبی در اطراف این پیکسل‌های سبز نیز در نظر گرفته می‌شوند. ورودی آزمون دوم علاوه بر شش ورودی ذکر شده در آزمایش اول، دو ویژگی زیر را نیز در نظر می‌گیرد:

١. میانگین پیسکل‌های قرمز
٢. میانگین پیسکل‌های آبی




شکل (٥) زیرماتریس انتخاب شده برای آموزش شبکه عصبی

در آزمون سوم چهار ویژگی زیر به‌عنوان ورودی در نظر گرفته شده است.

١. میانه چهار پیکسل
٢. میانگین چهار پیکسل
٣. میانگین پیسکل‌های قرمز
٤. میانگین پیسکل‌های آبی

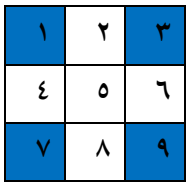
### ٢-٣ بازیابی مؤلفه های قرمز و آبی

برای درونیابی صفحات قرمز و آبی از صفحه سبز که کامل شده است و مؤلفه‌های ناموجود آن از طریق شبکه عصبی بدست آمده استفاده می‌شود. چراکه فرضیات خاصی را می‌توان با توجه به ارتباط بین صفحات رنگی در نظر گرفت. یکی از این فرضیات به خوبی شناخته شده، این است که صفحات رنگی در یک همسایگی به اندازه کافی کوچک، کاملا مرتبط هستند. در نتیجه در یک همسایگی به اندازه کافی کوچک، معادلات برای صفحات آبی و سبز و همچنین صفحات قرمز و سبز با ثابت‌های J و K به این صورت صحیح می‌باشند :

$$G = B + K$$ رابطه١

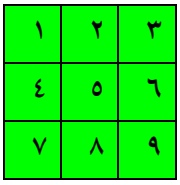
$$G = R + J$$ رابطه٢

نحوه‌ی این ارتباط در صفحه آبی و سبز بدین صورت است که زیر ماتریس ٣×٣ از صفحه آبی مطابق با شکل ٦ را در نظر بگیرید. مؤلفه‌های ١، ٣، ٧ و ٩ دارای مقدار مشخص می‌باشند و مؤلفه‌های ٢، ٤، ٥، ٦ و ٨ با استفاده از مؤلفه‌های متناظرشان در صفحه سبز و مؤلفه‌های آبی مشخص در اطرافشان بازیابی می‌شوند.




شکل (٦) زیرماتریس ٣×٣ از صفحه آبی

زیر ماتریس صفحه سبز نظیر با زیر ماتریس شکل٦ ، شکل ٧ می‌باشد.




شکل (٧) زیرماتریس ٣×٣ از صفحه سبز متناظر با زیر ماتریس انتخابی از صفحه آبی

مقدار B ناموجود به‌صورت زیر تخمین زده می‌شود:

$$B_5 = \frac{B_1+B_3+B_7+B_9}{4} + \frac{4\times G_5 - G_1 - G_3 - G_7 - G_9}{4}$$ رابطه٣

$$B_2 = \frac{B_1+B_3}{2} + \frac{2\times G_2 - G_1 - G_3}{2}$$ رابطه٤

$$B_4 = \frac{B_1+B_7}{2} + \frac{2\times G_4 - G_1 - G_7}{2}$$ رابطه٥

مقادیر $B_6$ و $B_8$ متناظر با مقادیر $B_2$ و $B_4$ محاسبه می‌شوند. مقادیر ناموجود از رنگ قرمز نیز به همین صورت با استفاده از صفحه سبز بدست می‌آید.

## ٣- نتایج پیاده‌سازی

در این مقاله از دو نوع شبکه عصبی، شبکه عصبی پرسپترون چند لایه و شبکه عصبی RBF برای دیموزائیک کردن تصویر استفاده شده است. برای اعمال شبکه عصبی، زیرماتریس‌های ٣×٣ از صفحه سبز انتخاب شده است و هدف از اجرای شبکه عصبی تخمین مؤلفه سبز ناموجود مرکزی در این زیرماتریس می‌باشد. در شبکه عصبی پرسپترون، پس از بررسی لایه‌ها و نرون‌های مختلف و مقایسه نتایج حاصل از آن‌ها، یک شبکه عصبی پرسپترون دو لایه با تابع انتقال سیگموئید در لایه پنهان و تابع انتقال خطی Purelin در لایه خروجی انتخاب شد و تعداد بهینه نرون در لایه پنهان ٦ در نظر گرفته شد. شبکه عصبی RBF دارای دو لایه، لایه پنهان و لایه خروجی می‌باشد. این شبکه در لایه میانی خود تنها دارای یک لایه پنهان است. تابع محرک در لایه میانی تابع گوسین و در لایه خروجی تابع خطی است. پارامتر انتخاب مرکز در هر نرون از لایه میانی با الگوریتم K-means توسط خود شبکه تعیین می‌شود. مقدار بهینه پارامترهای تعداد نرون‌ها در لایه پنهان و انحراف معیار(σ) تابع گوسین با روش سعی و خطا به‌ترتیب ٢٥ و ٠,٥ در نظر گرفته شده است.

برای ارزیابی تصاویر دیموزائیک بدست آمده از هر دو شبکه عصبی از معیار PSNR که به‌ترتیب در رابطه ٦ تعریف شده، استفاده می‌شود.

رابطه٦

$$PSNR = 10\,log_{10}\left(\frac{255^2}{\frac{1}{m \times n}\sum_{x=1}^{m}\sum_{y=1}^{n}(I_1(x,y) - I_2(x,y))^2}\right)$$

نتایج بدست آمده از سه آزمون با سه ورودی مختلف با دو شبکه عصبی پرسپترون دولایه و شبکه عصبی RBF در جدول ١ قابل مشاهده است.

جدول (١) PSNR کانال سبز از تصویر Lenna از هر سه آزمون

| | شبکه عصبی | تعداد ویژگی ورودی | PSNR سبز |
|---|---|---|---|
| آزمون اول | MLP | ٦ | ٣٥/٧٨١٣ |
| آزمون دوم | MLP | ٨ | ٣٤/٥٤٧٥ |
| آزمون سوم | MLP | ٤ | ٣٦/٠٢٨٩ |
| آزمون اول | RBF | ٦ | ٣٦/٠٦٥٨ |
| آزمون دوم | RBF | ٨ | ٣٦/٢٤٦٨ |
| آزمون سوم | RBF | ٤ | ٣٦/٠٣٥١ |

نتایج بدست آمده از این سه آزمایش بسیار به هم نزدیک می‌باشند. نتایج نشان می‌دهد که شبکه عصبی پرسپترون در سه آزمایش انجام شده در آزمایش سوم یعنی با ٤ ورودی بهترین نتیجه را ارائه داده است و شبکه عصبی RBF در هر سه آزمایش دارای نتایج بسیار نزدیک و قابل قبولی می‌باشد. در نتیجه مؤثرترین ورودی‌ها، چهار ویژگی میانه پیکسل‌هاس سبز، میانگین پیکسل‌های سبز، میانگین پیکسل‌های قرمز و میانگین پیکسل‌های آبی می‌باشد چرا که هر دو شبکه عصبی نتایج قابل قبولی را ارائه داده اند. و سپس برای بازسازی تصویر و کاهش پدیده‌های مصنوعی ایجاد شده بر روی تصویر از شبکه عصبی بهره برده‌اند.

نحوه استفاده از شبکه عصبی در این مقاله با سایر پژوهش‌های انجام شده که از شبکه عصبی برای دیموزائیک کردن تصویر استفاده نموده‌اند کاملا متفاوت است. چرا که سایر مقالات که از روش شبکه عصبی استفاده نموده‌اند، ابتدا از یکی از روش‌های پایه، روش خطی یا cubic برای دیموزائیک کردن تصویر استفاده نموده و سپس برای بازسازی تصویر و کاهش پدیده‌های مصنوعی ایجاد شده بر روی تصویر از شبکه عصبی بهره

برده‌اند. نتایج حاصل ازمقایسه مقاله [۵] با روش استفاده شده در این پژوهش در جدول ۲ آورده شده است. نتایج بدست آمده از شبکه عصبی حاصل از پژوهش انجام شده، نتایج قابل قبولی می‌باشند و این روش می‌تواند جایگزین خوبی برای روش‌های پایه‌ای دیموزائیک قرار بگیرد.

جدول (۲) مقایسه PSNR از تصویر Lena بر روی تصویرروشنایی از مقاله [۵] و روش پیشنهادی

| PSNR / تصویر | PSNR روشنایی روش مطرح شده در مقاله [۵] | PSNR تصویر سبز روش پیشنهادی |
|---|---|---|
| Lena | ۳۳/۸٦ | ۳٦/۲٤ |

## ۴– نتیجه‌گیری

در این مقاله از دو نوع شبکه عصبی، شبکه عصبی پرسپترون چند لایه و شبکه عصبی *RBF* برای دیموزائیک کردن تصویر استفاده شده است. برای اعمال شبکه عصبی، زیرماتریس‌های ۳×۳ از صفحه سبز انتخاب شده است و هدف از اجرای شبکه عصبی تخمین مؤلفه سبز ناموجود مرکزی در این زیرماتریس می‌باشد. از معیار PSNR جهت ارزیابی و مقایسه تصاویر دیموزائیک بدست آمده از شبکه عصبی پرسپترون دو لایه و شبکه عصبی RBF استفاده شده است. نتایج بدست آمده از این سه آزمون بسیار به هم نزدیک می‌باشند. نتایج نشان می‌دهد که شبکه عصبی پرسپترون در سه آزمون انجام شده در آزمون سوم یعنی با ٤ ورودی بهترین نتیجه را ارائه داده است و شبکه عصبی RBF در هر سه آزمون دارای نتایج بسیار نزدیک و قابل قبولی می‌باشد. در نتیجه مؤثرترین ورودی‌ها، چهار ویژگی میانه پیکسل‌هاس سبز، میانگین پیکسل‌های سبز، میانگین پیکسل‌های قرمز و میانگین پیکسل‌های آبی می‌باشد چرا که هر دو شبکه عصبی نتایج قابل قبولی را ارائه داده اند. نتایج بدست آمده از شبکه عصبی حاصل از پژوهش انجام شده در مقایسه با روش ارائه شده در مقاله [۲۰]، دارای نتایج بهتری می‌باشد و این روش می‌تواند جایگزین خوبی برای روش‌های پایه‌ای دیموزائیک قرار بگیرد.

## مراجع

[۱] Shiny,G. Nagarajan , " A Self-Learning Approach to Edge-Adaptive Demosaicking Algorithm for CFA Images ", International Journal of Engineering Research and Applications,Vol.۹,No.۲,pp.۲۱-۲۸،۲۰۱٤.

[۲] Li, X., Gunturk, B., & Zhang, L. Image demosaicing: A systematic survey. In proc. SPIE (Vol. ٦٨۲۲, pp. ۷۱۰-۷۲۵،۲۰۰۸.

[۳] He, F. L., Wang, Y. C. F., & Hua, K. L. . Self-learning approach to color demosaicking via support vector regression. In Image Processing (ICIP), ۱۹th IEEE International Conference on pp. ۲۷٦۵-۲۷٦۸,۲۰۱۲.

[٤] Shiny,G. Nagarajan ," A Self-Learning Approach to Edge-Adaptive Demosaicking Algorithm for CFA Images ", International Journal of Engineering Research and Applications,Vol.۹,No.۲,pp.۲۱-۲۸،۲۰۱٤.

[۵] Chen, W., Tian, Q., Liu, J., & Wang, Q. Nonlocal low-rank matrix completion for image interpolation using edge detection and neural network. Signal, Image and Video Processing, ۸(٤), ٦۵۷-٦٦۳,۲۰۱٤.