



## استفاده از قوانین گرادیان توام به منظور افزایش دقت شبکه عصبی در دسته بندی تصاویر

معصومه بورجندی، کیمیا رضائی کلانتری

عضو هیات علمی دانشگاه آزاد اسلامی واحد علی آباد کتول

گرگان، ایران

[m.boorjandi@yahoo.com](mailto:m.boorjandi@yahoo.com),

مدرس دانشگاه آزاد واحد ساری

ساری، ایران

[rezaeikalantari@gmail.com](mailto:rezaeikalantari@gmail.com)

### چکیده:

یکی از مسائل مهم در زمینه پردازش تصویر، طبقه بندی داده های یک تصویر است. روش های مختلفی جهت انجام این عمل ارائه شده است. تمامی این روش ها سعی در کاهش خطا در دسته بندی نموده اند. به توجه به اینکه عمل دسته بندی نوعی یادگیری با ناظر میباشد، لذا می توان از شبکه عصبی برای انجام آن استفاده نمود. شبکه عصبی **back-propagation** یکی از روش هایی است که اخیراً برای دسته بندی تصاویر استفاده شده است. اما این روش همیشه به نتایجی با دقت بالا همگرا نمی شوند. از آنجائیکه قوانین گرادیان توام، جستجویی در جهت گرادیان انجام می دهند بنابراین در این مقاله قصد داریم تا با افزودن آن به روش **back-propagation** در هر مرحله تکرار بتوانیم با انجام جستجو، بهترین مقدار را برای نرخ آموزش تعیین نماییم. نتایج آزمایشات حاکی از این است که روش بیان شده باعث بهبود قابل توجهی در دقت شبکه عصبی در دسته بندی تصاویر می گردد.

### کلمات کلیدی:

**back propagation**, قوانین گرادیان توام, Fletcher-Reeves, Polak-Ribieri, Powell-Beale, **Scaled conjugate gradient**

### ۱. مقدمه:

ها به محیط خارجی مرتبط هستند که به آنها گره های داخلی یا خارجی گفته می شود. هر گره، مجموعه ای از لبه های ورودی را دریافت کرده و مجموعه ای از لبه های خروجی نیز به دیگر گره ها

شبکه عصبی مصنوعی مشابه با ساختار شبکه عصبی انسان طراحی شده است. این شبکه از واحدهایی بنام گره و ارتباطاتی به نام اتصال (**link**) تشکیل یافته است. وزن ها عامل اصلی یادگیری در شبکه های عصبی هستند که در طول زمان بهنگام خواهند شد. برخی از گره

در بخش دوم نیز به بررسی نحوه کار شبکه back propagation پرداخته و در بخش سوم، روشها و تاثیرات قوانین گرادیان توام در شبکه بررسی شده و در ادامه نیز نتایج عملی از بکار بردن این قوانین و اثرات آنها روی مجموعه ای از تصاویر دیده می شود.

## ۲. روش backpropagation:

Backpropagation بر اساس قانون یادگیری widrow-hoff برای شبکه های چند لایه و توابع غیرخطی مشتق پذیر ایجاد می شود. بردارهای ورودی به همراه بردارهای هدف آن جهت آموزش شبکه برای تخمین یک تابع، تعیین میزان تشابه بردار ورودی به خروجی و طبقه بندی بکار می رود. شبکه هایی با داشتن بایاس، لایه ای با تابع سیگموئید و یک لایه خروجی با تابع خطی، قابلیت تخمین زدن هر تابعی را دارند. backpropagation استاندارد یک الگوریتم Gradient descent است، بدین ترتیب که وزن ها در جهت خلاف گرادیان تابع حرکت می کنند. یک شبکه backpropagation که به صورت مناسبی آموزش داده شده است، باید بتواند بهترین پاسخ را در مقابل ورودی که تاکنون ندیده است، ارائه نماید.

برای انجام فرآیند آموزش در شبکه چهار مرحله وجود دارد [۸]:

- ۱- جمع آوری داده های آموزشی
- ۲- ایجاد شبکه
- ۳- آموزش شبکه

۴- دادن ورودی جدید برای بدست آوردن پاسخ ها

هدف ما در این مقاله بررسی مرحله سوم می باشد. به منظور آموزش شبکه های backpropagation یک نکته حائز اهمیت این است که بتوان مشتق توابع انتقال مورد استفاده را محاسبه نمود. ساده ترین الگوریتم بدین صورت است که برورسانی وزن ها و بایاس ها در جهتی انجام می گیرد که تابع اجرای شبکه کاهش می یابد.

به عبارت دیگر در جهت عکس گرادیان، یک روش برای این منظور بصورت زیر می باشد:

$$x_{k+1} = x_k - \alpha_k g_k \quad (1)$$

که  $x_{k+1}$  بردار سس وزن ها و بایاس ها،  $g_k$  گرادیان فعلی و  $\alpha_k$  نرخ آموزش می باشد [۹].

## ۳. استفاده از قوانین گرادیان توام:

در الگوریتم پایه backpropagation وزن ها در جهت عکس گرادیان، یعنی جهتی که تابع اجرای شبکه به سرعت کاهش می یابد، تنظیم می شود. با اینکه تابع به سرعت در جهت خلاف گرادیان

دارد و یک سطح فعالیت دارد که معین می کند میزان تحریک ورودی در آن چه قدر است.

برای ساخت شبکه عصبی، ابتدا باید تعدادگره ها و نحوه اتصال آنها به یکدیگر را معین نمود، که این فرآیند به معماری شبکه عصبی مشهور است. سپس باید لایه های شبکه مقدا دهی اولیه شوند و از طریق یک الگوریتم یادگیری، وزن لایه ها بهنگام شود که این فرآیند معمولاً به کمک مثال های آموزشی صورت می گیرد [۱].

یکی از ساده ترین شبکه ها، شبکه عصبی پرسپترون می باشد. روش کار این شبکه بدین قرار است که در ابتدا سیگنال هایی توسط لایه های ورودی دریافت می شود که این اثر از تأثیر دو عامل تحریک پذیری، سلول ماقبل ( $a_j$ ) و وزن متعلق به لایه بین گره ماقبل ( $a_j$ ) و جاری ( $a_i$ ) است، که مقدار این وزن توسط  $w_{ji}$  نمایش داده می شود. فرمول محاسبه تحریک ورودی آن به قرار زیر است.

$$f(x) = b + \sum_{i=1}^n w_i \cdot a_i \quad (1)$$

یافتن این تابع کار بسیار ساده ای است. اما پرسپترون، تنها قادر به حل مسائل ابتدائی می باشد. در واقع مسائلی که توسط پرسپترون حل می شوند باید تابع جداپذیر خطی داشته باشند و یافتن این تابع کار بسیار ساده ای است [۲]. سال ها بعد مکانیزمی تحت عنوان انتشار به عقب (back-propagation) مطرح شد که با افزودن لایه های بینابینی به شبکه پرسپترون و به کمک این فرآیند قادر به حل مسائل جداپذیر غیرخطی شد [۳].

در شبکه back-propagation، برخلاف پرسپترون اولیه، لایه های بینابینی تحت عنوان لایه های مخفی به شبکه افزوده شدند بطوریکه حاصل کار، شبکه پرسپترون چندلایه بود [۶]. یادگیری در این شبکه بر مبنای فرایند انتشار به عقب صورت می گیرد. برای بهنگام سازی وزن لایه های بین لایه ورودی و مخفی، نیاز به تعریف مقدار خطا مشابه با مقدار خطا در خروجی هستیم. ایده چنان است که هرگره مخفی مسئول بخشی از خطای تولید شده در هر گره خروجی است، که به آن متصل است. پس به همین نحو، خطا در لایه های بین لایه مخفی و ورودی نیز توزیع می گردند [۴] [۵].

Wonil Kim, Han-Ku Lee در سال ۲۰۰۵ سعی نمودند با استفاده از توصیفگرهای MPEG7 و شبکه backpropagation باعث بالا رفتن دقت در دسته بندی تصاویر شوند [۷].

نوآوری این مقاله در بکار بردن قوانین گرادیان توام در شبکه back-propagation می باشد بطوریکه با استفاده از آن، یک مرحله جستجو در جهت گرادیان به شبکه اضافه می گردد که این جستجو در هر دوره تکرار باعث تعیین بهترین مقدار برای نرخ آموزش می گردد و بدین صورت دقت شبکه نیز به مقدار قابل توجهی افزایش می نماید. به منظور افزایش دقت شبکه عصبی در دسته بندی تصاویر می باشد.

### ۳.۳- روش Powell-Beale

در همه الگوریتم‌های گرادیان توأم جهت جستجو به صورت دوره‌ای مجدداً در عکس گرادیان قرار می‌گیرد. زمان تنظیم مجدد استاندارد هنگامی است که تعداد تکرارها برابر تعداد پارامترهای شبکه (وزن‌ها و بایاس‌ها) باشد. اما روش‌های دیگری مانند روش Powell باعث بهبود در آموزش شبکه می‌شود. در این روش زمان تنظیم مجدد زمانی است که تعامد کمی بین گرادیان فعلی و گرادیان قبلی وجود داشته باشد. این مقدار بصورت زیر محاسبه می‌شود:

$$\left\{ g_{k-1}^T g_k \geq 0.2 \|g_k\|^2 \right\} \quad (۹)$$

اگر شرط فوق برقرار شود، جهت جستجو مجدداً در جهت عکس گرادیان قرار خواهد گرفت.

### ۳.۴- روش Scaled conjugate gradient

در تمامی الگوریتم‌های گرادیان که تا اینجا بررسی شد نیاز به یک جستجوی خطی در هر دوره می‌باشد که این جستجو بسیار گران است. اما در روشی دیگر بنام گرادیان توأم درجه بندی شده، که ایده اصلی آن ترکیب دوروش Levenberg-Marquardt و گرادیان توأم می‌باشد دیگر نیازی به جستجوی خطی در هر دوره نمی‌باشد. با توجه به مطالب ذکر شده می‌توان به این نتیجه رسید که استفاده از این قانون هرچند به تعداد تکرار بیشتری نسبت به سایر قوانین گرادیان توأم نیاز دارد اما تعداد محاسبات در هر تکرار کاهش چشمگیری دارد.

### ۴- پیاده سازی

به منظور ارائه تاثیر بکارگیری قوانین گرادیان توأم بیان شده، ما از شبکه back propagation، که دارای ۶نود در لایه مخفی و ۴نود در لایه خروجی می‌باشد استفاده نموده ایم. پایگاه داده ای مورد استفاده حاوی ۱۲۶۰ تصویر از بانده مشخص مختلف که از سایت <http://archive.ics.uci.edu/ml> دانلود شده است می‌باشد.

بطوریکه از ۱۲۰۰ تصویر، به عنوان تصاویر آموزشی و ۶۰ تصویر، به عنوان تست استفاده نموده ایم. حاصل نتایج بدست آمده در شکل ۱ نشان داده شده است.

کاهش می‌یابد، اما این مسئله تضمین کننده همگرایی شبکه نخواهد بود. که این یکی از مشکلات در این شبکه می‌باشد. برای حل این مشکل و باتوجه به اینکه، نرخ آموزش در الگوریتم پایه جهت تشخیص اندازه قدم‌ها یعنی میزان بروزرسانی وزن بکار می‌رود، می‌توان با استفاده از قوانین گرادیان توأم در این شبکه، اندازه قدم‌ها در هر تکرار را تنظیم نمود. بطوریکه در هر مرحله، جستجویی در جهت گرادیان برای تشخیص اندازه بهترین قدم که تابع اجرا را مینیمم نماید، انجام می‌گیرد. بنابراین بدین صورت دقت شبکه بالا رفته و میزان خطا در هنگام تست کاهش می‌یابد. از جمله توابع گرادیان توأم به منظور بهنگام سازی وزن ها می توان به روشهای Fletcher-Reeves, Polak-Ribieri, Powell-Beale و گرادیان توأم درجه بندی شده اشاره کرد.

### ۳.۱- روش Fletcher-Reeves

تمام الگوریتم‌های گرادیان توأم با جستجویی در جهت عکس گرادیان در تکرار اول طبق فرمول ۱ آغاز می‌شوند.

$$p_0 = -g_0 \quad (۳)$$

سپس در این روش یک خط جستجو برای تشخیص فاصله بهینه برای حرکت در طول جهت جستجوی فعلی ایجاد می‌شود.

$$x_{k+1} = x_k + \alpha_k p_k \quad (۴)$$

سپس جهت بعدی جستجو بطوریکه توأم با جهت‌های جستجوی قبلی باشد، مشخص می‌شود. فرایند عمومی برای تشخیص جهت جدید بصورت زیر می‌باشد:

$$p_k = -g_k + \beta_k p_{k-1} \quad (۵)$$

میزان  $\beta_k$  باعث ایجاد ورژن‌های مختلفی در الگوریتم گرادیان توأم می‌شود. در این روش مقدار این پارامتر بصورت زیر تعیین می‌شود:

$$\beta_k = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}} \quad (۶)$$

### ۳.۲- روش Polak-Ribieri

مانند روش قبلی جهت جستجو در هر تکرار به صورت زیر تعیین می‌شود:

$$p_k = -g_k + \beta_k p_{k-1} \quad (۷)$$

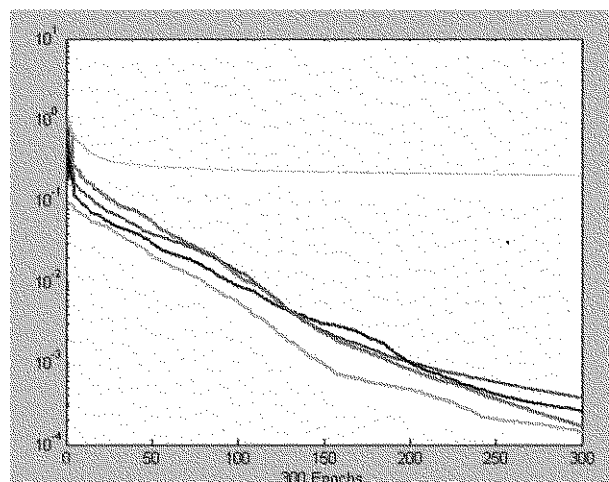
میزان  $\beta_k$  به صورت زیر تعیین می‌شود:

$$\beta_k = \frac{\Delta g_k^T g_k}{g_{k-1}^T g_{k-1}} \quad (۸)$$

تست این شبکه ها بر روی ۶۰ تصویر از پایگاه داده به نتایجی که در جدول ۱ ارائه شده است رسیده ایم. با توجه به جدول و نمودار بالا اگر چه دقت بکارگیری قوانین مختلف گرادیان توأم نزدیک به یکدیگر می باشد اما بکارگیری قانون powell-beale باعث افزایش دقت و نتایج بهتری در هنگام تست می شود.

## ۵- نتیجه گیری

در این مقاله الگوریتم backpropagation با قوانین مختلف گرادیان توأم برای مسئله تشخیص جهت چهره مورد بررسی قرار گرفت. همان طور که نتایج پیاده سازی ما نشان می دهد بکارگیری روش powell-beale نسبت به سایر روشها باعث افزایش دقت شبکه عصبی در دسته بندی تصاویر می گردد. نشان داده شد که الگوریتمهایی با قانون گرادیان توأم بسیار قوی تر از الگوریتمهای معمول (back propagation) می باشد، زیرا در این الگوریتمها، مراحل جستجویی قرار داده شده است که در هر مرحله با استفاده از آنها جستجویی جهت گرادیان صورت می گیرد تا بهترین مقدار برای نرخ آموزش به منظور مینیمم نمودن تابع اجرایی صورت گیرد که با توجه به نتایج بدست آمده از آزمایشات دیده شد که این امر خود باعث افزایش دقت در آموزش شبکه می شود.



شکل ۱: (نمودار درصد خطا در هنگام آموزش شبکه)

	Conjugate gradient fletcher-reeves
	Conjugate gradient polak-ribiere
	Scaled conjugate gradient
	Back propagation
	Conjugate gradient Powell-beale

آزمایشات فوق در ۳۰۰ اپک انجام شده است.

در نمودارهای فوق محور افقی نشان دهنده تعداد اپکها و محور عمودی نشان دهنده میزان مجموع مربعات خطا برای شبکه است. با

نوع شبکه عصبی	میزان خطای آموزشی	درصد پاسخ درست بهنگام تست
Back propagation	۰,۱۴۱۷۲۱	۸۰ درصد
قانون Fletcher-Reeves	۰,۰۰۰۲۵۴۱۴۴	۹۵ درصد
قانون Polak-Ribieri	۰,۰۰۰۲۹۶۱۵۹	۹۵ درصد
قانون Powell-Beale	۰,۰۰۰۱۵۸۵۴۲	۹۸ درصد
قانون Scaled conjugate gradient	۰,۰۰۰۲۳۷۸۱	۹۶ درصد

جدول ۱: (میزان خطا در هنگام آموزش و تست)

## مراجع :

- [1] B. Lerner, H. Guterman, M. Aladjem, and I. "Dinstein, Feature Extraction by Neural Network Nonlinear Mapping for Pattern Classification", 1996.
- [2] D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning Representations by Back Propagating Errors", Nature (London), Vol. 323, pp 533-536, 1986.
- [3] Lynne E. Parker, "Notes on Multilayer, Feedforward Neural Networks", Springer, 2006.
- [4] Holger Arnold, "A Recurrent Neural Network Model for Pattern Recognition", 2003.
- [5] Hugues Berry, Mathias Qouy, "structure and dynamics of random recurrent neural network", 2006.
- [6] G.M. FOODY, "Supervised image classification by MLP and RBF neural network with and without an exhaustively defined set of classes", ISSN 0143-1161 print, 2004.
- [7] Wonil Kim, Han-Ku Lee, Seong Joon Yoo, Sung Wook Baik, "Neural Network Based Adult Image Classification", Springer, 2005.
- [8] Shih-Wei Lin, Shou-Yan Chou, Shih-Chieh Chen, "Irregular shapes classification by back-propagation neural networks", Springer, 2007.
- [9] F. Fogelman Soulie, "Neural Network architectures and algorithms: A perspective", Elsevier, 1992.