

## استفاده از شبکه عصبی با انتشار هدایت یافته در خوشه‌بندی زیر

### کلمات متون چاپی فارسی<sup>۱</sup>

محمد جواد فدایی اسلام<sup>۲</sup>، ناصر مزینی<sup>۳</sup>

#### چکیده

در این مقاله، از شبکه عصبی با انتشار هدایت یافته برای تصحیح نتایج حاصل از جداسازی زیر کلمات متون چاپی فارسی استفاده شده است. زیرکلمات بدون نقطه، با استفاده از روشی به زیرحروف شکسته می‌شوند. هر زیرحرف با زیرحرف‌های به دست آمده از فونت لوتوس مقایسه شده و میزان شباهت آن با استفاده از معیار فاصله اقلیدسی و ویژگی گشتاور مرکزی نرمال شده سنجیده می‌شود. سپس این مقادیر برای شناخت دنباله زیر حروف به شبکه داده می‌شود. شبکه با انتشار هدایت یافته یک ساختار درختی دارد و به صورت بدون ناظر عمل می‌کند و برای هر کلمه ممکن است چندین پیشنهاد با درجه شباهت گوناگون ارائه کند. پایگاه داده، زیر کلمات متداول روزنامه‌های فارسی بودند که در معرض نوعی از اعوجاج قرار گرفتند. نتایج شناسایی پس از اعمال شبکه در مقایسه با معیار فاصله اقلیدسی بهبود چشمگیری یافت.

#### کلمات کلیدی

شبکه عصبی با انتشار هدایت یافته، بازشناسی متون چاپی فارسی، پردازش تصویر، خوشه‌بندی

## Application of Guided Propagation Network on Persian Subword Clustering

M. J. Fadaeieslam, N. Mozayani

#### Abstract

In this study, a Guided Propagation Network is used to improve the results of Persian subword recognition. The subwords which are without points are broken to subcharacters. Each subcharacter is compared to Lotus Font subcharacters and its similarity is determined using Euclidian distance criterion and normalized central moments feature. The results are given to the network for recognizing the sequence of subcharacters. The network has a tree structure. It operates unsupervisedly and might propose several solutions with different degrees of similarity. The data set is collected from the common subwords of Persian newspapers. We have also simulated a kind of distortion to the subwords. After applying the network, the results of recognition improved drastically, compared to the Euclidian distance criterion.

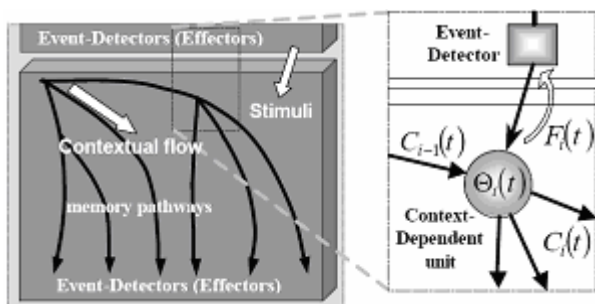
#### Keywords

Guided Propagation Network, GPN, OCR, Optical Character Recognition, Image Processing, Clustering, Text Recognition.

۱- این تحقیق با حمایت شورای عالی اطلاع رسانی انجام شده است.

۲- دانشجوی دوره دکترا، دانشکده کامپیوتر، دانشگاه علم و صنعت ایران، fadaei@iust.ac.ir

۳- عضو هیئت علمی، دانشکده کامپیوتر، دانشگاه علم و صنعت ایران، mozayani@iust.ac.ir



شکل ۱- معماری GPN [۴] (یک ماژول شبکه)

## ۱- مقدمه

شبکه عصبی با انتشار هدایت یافته، با الهام از مغز انسان به وجود آمده و سعی دارد از بعد زمان در معماری خود استفاده کند. در این شبکه اطلاعات مربوط به یک دنباله در همزمانی بین تحریک ورودی و جریان داخلی آن ذخیره می‌شود. از ویژگی‌های مهم این شبکه آن است که فرآیند پردازش اطلاعات و یادگیری به طور همزمان انجام می‌شود، بدین معنا که به محض ورود دنباله‌ای که قبلاً دیده نشده است شبکه می‌تواند آن را یاد بگیرد.

حروف در متون فارسی و عربی به هم می‌پیوندند که گاهی این پیوستگی باعث تغییر در شکل حروف می‌شود. با توجه به این نکته دسته مهمی از روش‌های شناسایی متون فارسی ابتدا کلمه را به حروف آن می‌شکنند سپس بخش‌ها را مورد شناسایی قرار می‌دهند. در این روش که مبتنی بر جداسازی است دنباله‌ای از حروف و زیرحروف پدید می‌آید. از آنجا که شبکه با انتشار هدایت یافته دنباله تحریک‌ها را به خوبی تشخیص می‌دهد در این مقاله از آن برای بهبود نتایج حاصل از جداسازی زیر کلمات فارسی استفاده شده است. ساختار مقاله بدین صورت است که در بخش دوم معماری شبکه‌های عصبی با انتشار هدایت یافته معرفی می‌شود. در بخش سوم بخش بندی کلمه به حروف و زیرحروف و پیاده‌سازی شبکه بررسی شده است بخش آخر به جمع‌بندی و نتیجه‌گیری و ارائه پیشنهادها اختصاص دارد.

## ۲- معماری شبکه عصبی با انتشار هدایت یافته<sup>۱</sup>

شبکه‌های با انتشار هدایت یافته از یک یا چند ماژول تشکیل می‌شوند. هر ماژول به وسیله یک جریان داخلی اولیه که خود به خود در آن وجود دارد، و نام آن نیز جریان بافتی<sup>۲</sup> است فعال می‌شود. جریانی بافتی در طول زمان به سمت یک خروجی مشخص هدایت می‌شود. البته این هدایت با استفاده از سیگنال تحریک (سیگنال فعالیت)<sup>۳</sup>  $S(t)$  که از ماژول‌های جانبی (یا محیطی) می‌آید و سیگنال تسهیل کننده<sup>۴</sup> (سیگنال کاهش دهنده آستانه)  $F(t)$  که از ماژول‌های پایین‌تر یا تنظیم کننده مرکزی می‌آید انجام می‌شود. شکل (۱) معماری GPN را نشان می‌دهد. در سمت چپ یک ماژول GPN نمایش داده شده که دارای سه جریان منشعب از ریشه است. این جریان‌های بافتی به سمت آشکارسازهای رویداد در حرکت هستند، در سمت راست هم یک سلول بافتی نشان داده شده که از پدر خود سیگنال  $C_{i-1}(t)$  را دریافت می‌کند. این سلول قبل از فعال شدن به آشکارساز ماژول دیگر سیگنال تسهیل کننده  $F_i(t)$  می‌فرستد و با این سیگنال آستانه فعالیت سیگنال تحریک وابسته به خود را کاهش می‌دهد. وقتی مجموع  $C_{i-1}(t)$  و تحریک رسیده، از آستانه بیشتر شد  $C_i(t)$  تولید شده و به سلول بافتی بعدی فرستاده می‌شود.

هر واحد یک آستانه تحریک  $\theta(t)$  دارد که ممکن است به وسیله سیگنال تسهیل کننده  $F(t)$  کاهش یابد. به عبارتی سیگنال تسهیل کننده به معنای انتظار برای رسیدن ورودی تسهیل شده است. مقدار خروجی یک واحد پردازش اولیه وقتی ورودی از حد آستانه بیشتر باشد، خطی است و در پایین این حد واحد پردازش اولیه خروجی ندارد. ماکزیمم خروجی  $A_{max}$  زمانی حاصل می‌شود که ورودی مقدار ماکزیمم خود را به دست آورد [۶]. مقدار آستانه یک واحد بافتی از رابطه (۱) محاسبه می‌شود که در این رابطه،  $R_{i-1}$  برای تنظیم نسبت بین  $C_{i-1}(t)$  و  $S_i(t)$  است و بسیار آهسته با قانونی شبیه هب تغییر می‌کند،  $E_n$  تحریک-پذیری<sup>۵</sup> واحد بافتی را مشخص می‌کند.

$$\theta_i(t) = \frac{\max(S_i(t) + R_{i-1} \times C_{i-1}(t))}{E_n} \quad (1)$$

$$= \frac{(1 + R_{i-1}) \times A_{max}}{E_i}$$

وقتی ورودی از حد آستانه فراتر رود، خروجی واحد از رابطه زیر محاسبه می‌شود:

$$C_i(t) = \frac{S_i(t) + R_{i-1} \times C_{i-1}(t) - 1}{1 + R_{i-1}} \quad (2)$$

در رابطه (۱)،  $A_{max}$  بیشترین مقدار هر ورودی است و همیشه  $E_n > 1$  است. اگر بیشینه مقدار تحریک ورودی، بتواند به تنهایی بیشتر از مقدار آستانه شود (برقراری رابطه (۳))، واحد فقط با استفاده از حداکثر مقدار تحریک، می‌تواند فعال شود. به این حالت "انتشار اجباری" گویند، و این نوع واحد بافتی "هدایت کننده داده" نام دارد.

$$A_{max} \times (1 + R_{i-1}) \times A_{max} / E_i \quad (3)$$

یک واحد بافتی "هدایت کننده دانش" است؛ اگر بیشینه مقدار ورودی بافتی به تنهایی بتواند از حد آستانه رد شود. که به این حالت "انتشار گسترش یافته" می‌گویند (رابطه ۴). در این مورد بدون آمدن سیگنال تحریک جریان می‌تواند ادامه یابد.

$$A_{max} \times R_{i-1} \times (1 + R_{i-1}) \times A_{max} / E_i \quad (4)$$

جدول ۱- دقت بازشناسی بدون استفاده از شبکه

دقت تشخیص	ویژگی داده ورودی
۱۰۰٪	فونت لوتوس با بزرگی ۱۲ بدون اعوجاج و نویز
۵۷٫۴۹٪	فونت لوتوس با بزرگی ۱۲ (نویزدار)

این روش  $E_i$  برای کلیه گره‌ها ثابت در نظر گرفته شده است. با توجه به اینکه بیشینه مقدار صورت کسر مد نظر است. رابطه (۱) برای هر گره به صورت زیر می‌شود.

$$\theta_i = \frac{100 \times (1 + R)}{E} \quad (۵)$$

مقدار سیگنال تحریک از رابطه (۶) به دست می‌آید.

$$S_i(t) = 100 - dist \times 50 \quad (۶)$$

در رابطه (۶) مقدار  $dist$  فاصله اقلیدسی گشتاور مرکزی نرمال شده بخش ورودی با بخشی است که این سیگنال تحریک مربوط به آن است. این جریان تا گره‌های برگ ادامه می‌یابد و میزان این جریان در گره برگ مشخص‌کننده میزان شباهت دنباله ورودی با دنباله‌ای است که این گره معرف آن است.

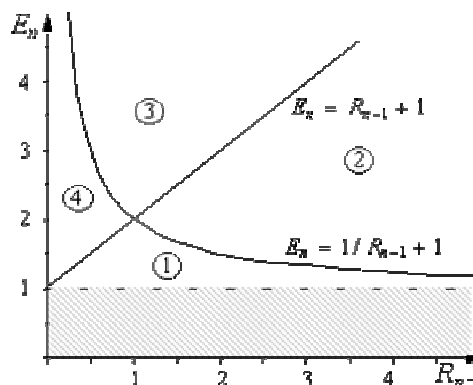
سیگنال تسهیل‌کننده به صورت صفر و یک صادر می‌شود بدین صورت که از مرحله جداسازی فاصله هر بخش با تک بخش‌های الگو محاسبه می‌شود. اما با استفاده از سیگنال تسهیل‌کننده تنها آنهایی مد نظر قرار می‌گیرند که جزو فرزندان آن گره باشند و در یک رابطه متقابل فرزندان مورد بررسی قرار می‌گیرند که از دسته بخش ورودی باشند.

هر چند شبکه با انتشار هدایت‌یافته، فرآیند خوشه‌بندی و یادگیری را باهم انجام می‌دهد. اما برای مقاوم بودن بیشتر در برابر نویز، این دو فرآیند را از هم جدا نمودیم. زیرا مطمئن هستیم که تمام حالات متداول را می‌توانیم با یک بار دادن داده‌های بدون نویز به شبکه، به آن آموزش دهیم. پارامتر  $R$  نسبت بین سیگنال تحریک و جریان بافتی را مشخص می‌کند. از دیدگاه دیگر جریان رسیده از پدر را بین فرزندان تقسیم می‌کند. پس چه بهتر است که بر اساس تعداد تکرار آن فرزند یا حرف بعد از پدر تنظیم شود.

۲-۳- نتایج پیاده‌سازی

مشخصات شبکه عصبی با انتشار هدایت‌یافته که برای پیاده‌سازی استفاده شده، در بخش قبل آمد. درخت مناسب برای پیمایش نیز به همراه ضرایب  $R$  ساخته شد. تنها پارامتر باقی‌مانده میزان تحریک هر سلول بافتی،  $E$  است. ابتدا این پارامتر را برای هر سلول برابر ۱٫۳ قرار دادیم (انتشار اجباری). در این حالت سیگنال تحریک به تنهایی می‌تواند یک سلول پیش‌فعال شده را فعال کند، اما هیچ سلولی در حالتی قرار نمی‌گیرد، که بدون سیگنال تحریک فعال شود [۹ و ۳]. نتایج بالا نشان می‌دهد در صورتیکه  $R$  کمتر از یک باشد، و به سمت صفر میل کند اهمیت سیگنال تحریک بیشتر می‌شود.

اما در حالت پایه یا "انتشار محدود" واحد بافتی برای گذر از مقدار آستانه و فعال شدن واحد به هر دو ورودی بافت و تحریک نیازمند است. شکل (۲) این حالت‌ها را در قالب یک نمودار نشان می‌دهد [۳].



شکل ۲- حالت‌های مختلف فعالیت یک سلول بافتی. ناحیه ۱: انتشار محدود، ناحیه ۲: انتشار گسترش‌یافته، ناحیه ۳: انتشار اجباری، ناحیه ۴: انتشار اجباری، ناحیه ۳: انتشار اجباری و گسترش‌یافته [۴]

۳- بخش بندی زیر کلمه به حروف و زیر حروف به کارگیری شبکه برای بهبود آن

برای استخراج زیر کلمات فارسی از روش [۸ و ۶] استفاده شده است. به این ترتیب که نسخه الکترونیکی چند سال روزنامه‌ها مورد پردازش قرار گرفت و زیر کلمات آنها استخراج شد. با استفاده از این روش و بدون در نظر گرفتن نقطه تعداد زیر کلمات فارسی متداول به ۳۷۵۸ عدد رسید. به نحوی که در مرجع [۹] توضیح داده شده است به آن نویز اضافه شد.

برای بخش‌بندی از پروفایل بالای زیر کلمات استفاده شد که به جهت اختصار به آن نمی‌پردازیم (در مراجع [۹ و ۷] به طور کامل بحث شده است).

برای موتور تشخیص معیار فاصله اقلیدسی در نظر گرفته شد که با استفاده از ویژگی گشتاور مرکزی نرمال شده، داده‌های ورودی را با بخش‌هایی که از فونت لوتوس با بزرگی ۱۲ به دست آمده‌اند، مقایسه می‌کند. دقت میزان بازشناسی برای داده‌های ورودی مختلف، در جدول (۱) آمده است. دنباله‌هایی که در قسمت بخش‌بندی به دست آمده‌اند، در بخش بعد برای بهبود تشخیص، به شبکه می‌دهیم.

۳-۱- استفاده از شبکه با انتشار هدایت یافته برای

بهبود نتایج بخش بندی

شبکه پیاده‌سازی شده دارای یک ماژول است. البته در صورتی که اطلاعات ساختار زبانی بیشتری از فضای بازشناسی متون وجود داشته باشد، می‌توان ماژول‌های دیگری هم طراحی نمود.

جریان بافتی اولیه، از ریشه به فرزندان در ابتدا همیشه مقدار ماکزیمم خود را دارد که در اینجا ۱۰۰ در نظر گرفته شده است. در

و برای هر داده ورودی، ممکن است چند پیشنهاد ارایه دهد و میزان شباهت با داده ورودی را هم مشخص می‌کند. شبکه این قابلیت را دارد که از ماژول‌های متعدد تشکیل شود، با استفاده از این قابلیت می‌توان اطلاعات دستور زبانی مربوط به یک زبان را هم در شبکه استفاده کرد و یا چند ماژول با روش‌های مختلف بازشناسی ایجاد کرد که توسط سیگنال‌های تحریک و تسهیل کننده با یکدیگر ارتباط دارند و نتیجه در انتها، توسط یک ماژول که از اطلاعات به دست آمده از ماژول‌های دیگر استفاده می‌کند تولید شود.

## مراجع

- [1] P. Escande, D. Béroule, and P. Blanchet, "Speech recognition experiments with guided propagation", Proc. Int. Joint Conf. Neural Networks, Singapore, 1991.
- [2] M. Roques and D. Beroule, "Strategies of Unsupervised Learning for A Parallel Parsing Architecture," Proceeds. of the International Joint Conference on Neural Networks, Seattle.
- [3] C. Toffano-Nioche, D. Béroule, and J. P. Tassin, "A functional model of some Parkinson's disease symptoms using a guided propagation network," Artif. Intell. Medicine, no. 14, pp. 237-258, 1998.
- [4] D. Beroule, "An Instance of Coincidence Detection Architecture Relying on Temporal Coding," IEEE Transactions on Neural Networks, vol. 15, No. 5, September 2004.
- [5] C. Toffano-Nioche, H. Ruellan, C. Ménigault, J. C. Martin, A. Lainé, and D. Béroule, "The neverending learning II", LIMSI Internal Rep. N96-05, 1996.

[۶] ابراهیمی، ا. و کبیر، ا. "استفاده از یک روش دو مرحله‌ای برای طبقه‌بندی زیر-کلمات چاپی فارسی"، ششمین کنفرانس سیستم‌های هوشمند، کرمان، ۱۳۸۳.

[۷] ر. عزمی، "بازشناسی متون چاپی فارسی"، پایان نامه دکتری، دانشگاه تربیت مدرس، ۱۳۷۸

[۸] ابراهیمی، ا. و کبیر، ا. "یک روش دو مرحله‌ای برای طبقه‌بندی زیر-کلمات چاپی" نشریه مهندسی برق و مهندسی کامپیوتر ایران، سال ۲، شماره ۲، ۱۳۸۳.

[۹] فدایی اسلام، "استفاده از شبکه عصبی با انتشار هدایت یافته در خوشه‌بندی زیر کلمات متون چاپی فارسی"، پایان نامه کارشناسی ارشد، دانشگاه علم و صنعت ایران، ۱۳۸۴.

## زیر نویس‌ها

1- Guided Propagation Network

2- Contextual Flow

3- Stimuli

4- Facilitation

5- Excitability

شبکه دیگری مورد آزمون قرار گرفت. بدین صورت که  $R' = 1 + R$  به جای  $R$  استفاده شد. در این حالت تمام گره‌ها در حالت انتشار محدود قرار می‌گیرند (جدول (۲)). طبقه بندی درست، در انتخاب اول بهتر شده است، برای آنکه در این شبکه به جریان‌های داخلی بیشتر اهمیت داده شده است. اما میزان پیشنهادها کم و میزان واژدگی زیاد شده است. دلیل آن این است که سلول‌های پیش فعالی که می‌توانستند تنها با سیگنال تحریک فعال شوند، اکنون برای فعال شدن به جریان زیاد بافتی علاوه بر سیگنال تحریک نیازمندند [۴، ۳ و ۹].

تغییر دیگری در شبکه می‌دهیم، مقدار  $E$  را از ۱.۳ به ۱.۵ افزایش می‌دهیم. این افزایش حالت کاری سلول‌ها را تغییر نمی‌دهد و کماکان آنها را در انتشار محدود نگه می‌دارد. اما به علت کاهش میزان آستانه، باعث می‌شود جریان‌های و سلول‌های بیشتری در شبکه فعال باشند. (جدول (۲)). همانطور که دیده می‌شود نتایج، بهبود چشمگیری یافته است. توجه به این نکته مهم است که هرچند میزان بالاتر  $E$  باشد، میزان درستی در حداکثر ده انتخاب اول را افزایش می‌دهد، اما زیاد کردن بی رویه آن موجب زیاد شدن تعداد سلول‌های فعال می‌شود، یا عبارتی برای یافتن جواب تمام شبکه را می‌گردد. پس آن را باید در حد متعادلی تنظیم نمود.

جدول ۲- نتایج شبکه بر روی فونت لوتوس ۱۲ نوبت‌دار

E=1.5	E=1.3	E=1.3	
$R'$ و	$R'$ و	$R'$ و	
٪۷.۴۸	٪۱۸.۷۱	٪۱۳.۶۳	زیر کلمات تشخیص داده نشده (واژدگی)
٪۲.۱۳	٪۶.۴۱	٪۱۲.۸۳	زیر کلمات اشتباه تشخیص داده شده
٪۷۴.۰۶	٪۶۶.۸۴	٪۶۱.۷۶	طبقه بندی درست در انتخاب اول
٪۹۰.۳۷	٪۷۴.۵	٪۷۳.۵۲	میزان درستی در حداکثر ده انتخاب اول
۳.۷۱	۲.۱۵	۲.۶۱	متوسط کل پیشنهادها برای هر زیر کلمه

## ۴- نتیجه‌گیری و ارایه پیشنهادها

شبکه عصبی با انتشار هدایت یافته، یک ساختار درختی بدون ناظر است. این شبکه از شبکه‌های طبیعی الهام گرفته شده، و قابلیت یادگرفتن یک دنباله، با استفاده از همزمانی بین سیگنال‌های تحریک و جریان داخلی را دارد. فرایند یادگیری در این شبکه همزمان با پردازش و استفاده از آن است. این شبکه مانند سایر شبکه‌های عصبی که فیدبک ندارند، در مقایسه با برخی مدل‌های آماری مثل مدل مخفی مارکوف بسیار سریع است. در ابتدا، در این شبکه یک جریان اولیه وجود دارد که با استفاده از سیگنال‌های تحریک و سیگنال‌های تسهیل کننده به سمت خروجی هدایت می‌شود، میزان جریانی که به گره انتهایی می‌رسد نشان دهنده میزان شباهت دنباله ورودی به آن است. شبکه با انتشار هدایت یافته در خود مکانیزمی برای واژدگی دارد