

Attractor Analysis in Associative Neural Networks and its Application to Facial Image Analysis

Seyyede Zohreh Seyyedsalehi¹, Seyyed Ali Seyyedsalehi 1

¹ Bioelectric, Biomedical Engineering, Amirkabir University, Tehran, Iran

Abstract :

Autoassociative neural networks can be used for nonlinear processing and normalization of data. Because, firstly, they are able to learn and simulate complex nonlinear communications, and secondly, the communications can be learned through analyzing and distributing information on the neurons and weights and then combining the results of their processing. In this way, they actually make an interpolation between the input data and their communications. But these neural networks cannot model attractor dynamics that is obviously used in brain function. In this paper, the output of autoassociative neural network is connected to its input, and through recursive connections the ability of attractor behavior in these models is provided. This study showed that a recursive neuron with a logistic function forms two attractors, in its training point and its symmetry, but for it with a sigmoid nonlinear function can be formed an attractor in a certain range. In the experiments on face images, it was shown that the absorbance of the images to their attractors was improved from 52.67% to 87.27% by increasing the number of layers and the supervised layer-by-layer pre-training in order to adjust the attractors.

Keywords: Associative Neural Networks, Attractor Dynamics, Face Images, Fixed Point, Recursive Connection.

تحلیل جاذب‌ها در شبکه‌های عصبی خودانجمنی و کاربرد آن در آنالیز تصاویر چهره

سیده زهره سیدصالحی^۱، سید علی سیدصالحی^۲

۱- فارغ‌التحصیل دکتری، گروه بیوالکترونیک، دانشکده مهندسی پزشکی، دانشگاه صنعتی امیرکبیر، تهران، ایران

z.seyyedsalehi@aut.ac.ir

۲- دانشیار، گروه بیوالکترونیک، دانشکده مهندسی پزشکی، دانشگاه صنعتی امیرکبیر، تهران، ایران

ssalehi@aut.ac.ir

چکیده: شبکه‌های عصبی خودانجمنی بالقوه امکان به‌کارگیری برای پردازش و هنجارسازی غیرخطی داده‌ها را در خود دارند؛ زیرا نخست، به یادگیری و شبیه‌سازی ارتباطات غیرخطی پیچیده قادرند و دوم، این ارتباطات را با تجزیه و گسترده‌سازی اطلاعات روی نورون‌ها و وزن‌ها و سپس ترکیب نتایج پردازش آنها به انجام می‌رسانند و از این طریق روی اطلاعات ورودی و ارتباطات مابین آنها عملاً درونیابی انجام می‌دهند. ایراد این شبکه‌ها این است که نمی‌توانند رفتار جاذب‌گونه را توضیح دهند که یکی از واضح‌ترین عملکردهای مغز است. در این مقاله با برقراری اتصال بازگشتی در ساختار این شبکه‌ها، قابلیت رفتار جاذب‌گونه نیز به آنها افزوده شده است. تحلیل‌های انجام‌شده در این راستا نشان داد برای یک نورون با تابع غیرخطی تانژانت هیپربولیک، با برقراری این اتصال بازگشتی، همواره در محل نمونه‌تعلیم داده شده و قرینه آن دو جاذب ایجاد می‌شود؛ اما در صورتی که تابع غیرخطی، سیگموئید باشد برای محدوده خاصی جاذب تشکیل می‌شود. در آزمایشات روی تصاویر چهره نشان داده شد با افزایش تعداد لایه‌های شبکه خودانجمنی، قابلیت ذخیره تصاویر بیشتری فراهم می‌شود؛ به طوری که با پیش‌تعلیم لایه به لایه باسرپرست، به‌منظور جهت‌دهی به نحوه تشکیل جاذب‌ها، میزان جذب تصاویر به جاذب‌هایی با حالت مشابه از ۵۲/۶۷٪ به ۸۷/۲۷٪ بهبود یافت.

واژه‌های کلیدی: شبکه‌های عصبی خودانجمنی، دینامیک‌های جاذب، نقاط تعادل، اتصال بازگشتی، تصاویر چهره

۱- مقدمه

نمی‌گیرند، منیفلدی که تخمین زده می‌شود، نسبت به منیفلد ذاتی داده‌ها خطای زیادی دارد؛ بنابراین در بسیاری از موارد، زمانی که الگوهای ورودی جدید از نمونه‌های تعلیم داده شده فاصله می‌گیرند، قابلیت تعمیم‌دهی به‌شدت افت می‌کند. در این موارد سعی می‌شود با در اختیار گرفتن داده‌های بیشتر و میانگین‌گیری از آنها تا حدودی نویز داده‌ها کاهش داده شود [۲ و ۳]؛ اما باید در نظر داشت که در خیلی از مواقع امکان دسترسی به داده‌های بیشتر وجود ندارد. بنابراین راهکار دیگر در این زمینه، هدایت نمونه جدید با قابلیت جاذب‌ها به سمت یکی از الگوهای یادگیری‌شده است که به آن نزدیک‌تر است.

وجود نویز در داده‌ها از مشکلاتی است که روش‌های کاهش بعد و استخراج منیفلد با آن مواجه‌اند [۱]. در چنین مواردی با توجه به اینکه نمونه‌ها دقیقاً روی منیفلد قرار

^۱ تاریخ ارسال مقاله: ۱۳۹۳/۱۱/۰۶

تاریخ پذیرش مقاله: ۱۳۹۷/۰۴/۱۶

نام نویسنده مسئول: سیده زهره سیدصالحی

نشانی نویسنده مسئول: ایران - تهران - دانشگاه صنعتی امیرکبیر، دانشکده مهندسی پزشکی

این شبکه‌ها را برای مقابله با مسئله تنوعات کارآمد دانسته‌اند [۹]. ایراد این شبکه‌ها این است که رفتار جاذب‌گونه را توضیح نمی‌دهند که یکی از واضح‌ترین عملکردهای مغز است [۱۰]. هنگامی که به شبکه عصبی خودانجمنی الگویی ارائه می‌شود که برخی مؤلفه‌های آن شبیه یکی از الگوهای یادگیری شده است و برخی دیگر به الگوی دیگر مشابه است، شبکه الگویی با ترکیب این مؤلفه‌ها ارائه می‌دهد که لزوماً به هیچ دو دسته الگوی یادگیری شده متعلق نیست که این مطلوب نیست.

در این مقاله در راستای فراهم کردن قابلیت رفتارهای جاذب‌گونه در شبکه‌های عصبی خودانجمنی راهکاری معرفی شده است و آن برقراری اتصالات بازگشتی در این ساختارهاست که با حفظ توانمندی آنها در استخراج مؤلفه‌های اساسی، قابلیت رفتار جاذب‌گونه را نیز به آنها می‌افزاید. انتظار می‌رود این جاذب‌های سراسری بتوانند قابلیت این شبکه‌ها را در حذف تنوعات الگوهای جدید تا حد چشمگیری بهبود بخشند و آنها را مقاوم کنند.

۱- شبکه خودانجمنی با اتصال بازگشتی تأخیردار

می‌توان نشان داد هنگامی که به شبکه خودانجمنی با اتصال بازگشتی تأخیردار، تعدادی الگو تعلیم داده می‌شود، در صورت داشتن ظرفیت کافی و تعلیم صحیح شبکه، با شرایطی به تعداد الگوها در ساختار شبکه، جاذب ساخته می‌شود [۱۱ و ۱۲]. این جاذب‌ها در راستای حذف تنوعات داده‌ها استفاده می‌شوند. به عبارتی هنگامی که به چنین ساختاری نمونه‌ای داده می‌شود که به دلیل نویز یا سایر عوامل از منیفلد ذاتی فاصله گرفته‌است، شبکه جاذب، آن را به سمت منیفلد هدایت می‌کند. در ادامه، تحلیل‌های ریاضی اولیه مبنی بر اثبات وجود این جاذب‌ها ارائه می‌شود.

۱-۱- تحلیل رفتار جاذب‌گونه نوروں با اتصال

بازگشتی تأخیردار

هرگاه به شبکه با نوروں غیرخطی، ورودی X_1 به صورت خودانجمنی تعلیم داده شود، می‌توان ثابت کرد

شواهد فیزیولوژیکی نیز بر وجود این جاذب‌ها در مغز دلالت دارند [۴]. برخلاف ساختار منحصراً جلوسوی یک شبکه عصبی جلوسوی چند لایه، بین قسمت‌های مختلف مغز اتصالات رفت و برگشتی بسیاری دیده می‌شود. مشخص شده است در ساختار محلی تک‌تک قسمت‌های مغز، بازخوردهای محلی وجود دارند که به رفتارهای جاذب‌گونه منجر می‌شوند؛ برای مثال می‌توان به پردازش در منطقه CA3 از هیپوکمپوس اشاره نمود. همچنین یکی از کاراترین قابلیت‌های مغز انسان، عملکرد آن به عنوان حافظه انجمنی است. این حافظه، پیونددهنده اطلاعات ورودی با سایر اطلاعات و نیز سایر اطلاعات با یکدیگر در مغز است. بازیابی اطلاعات در این حافظه از یک حالت دلخواه که می‌تواند یک بازنمایی گذشته از یک تصویر بینایی باشد، آغاز و در نهایت فعالیت مغز به یکی از حالت‌ها همگرا می‌شود. این حالت، وضعیت پایدار یا همان جاذب است که قبلاً مغز آن را به خاطر سپرده است [۵].

پویایی غیرخطی را با استفاده از اتصالات بازگشتی در شبکه‌های عصبی و حافظه‌های ذخیره‌شده در آنها می‌توان مدل‌سازی نمود [۶]. شبکه‌های عصبی بازگشتی با توجه به اینکه در ساختار خود بازخورد دارند، قادر به تولید رفتارهای جاذب‌گونه‌اند؛ به همین دلیل، این شبکه‌ها اهمیت زیادی در دانش عصبی محاسباتی دارند و نشانه‌های زیادی وجود دارد که از این سازوکارها برای پردازش اطلاعات در مغز بسیار استفاده می‌شود [۷]. این شبکه‌ها کاربردهایی شامل ذخیره حافظه انجمنی و بازسازی تصاویر نویزی دارند. در این کاربردها یک سری محاسبات توزیع‌شده با پویایی جاذب‌گونه انجام می‌شود و به یک نقطه ثابت پایدار منجر می‌شود [۸]؛ اما مشکل این مدل‌های بازگشتی این است که این شبکه‌ها را حتماً باید با نمونه‌های تمیز بدون نویز تعلیم داد؛ زیرا همان نمونه‌های تعلیمی‌اند که جاذب‌ها را تشکیل می‌دهند و دسته‌های الگوها را نمایندگی می‌کنند؛ درحالی‌که انسان قادر است با دیدن نمونه‌های نامحدود از اشیاء و مفاهیم، تعداد بسیار کوچک‌تری گروه مفهومی بسازد و هنگام برخورد با مفاهیم، آنها را در یکی از گروه‌های موجود در ذهن طبقه‌بندی کند. این خاصیتی است که در شبکه‌های عصبی خودانجمنی وجود دارد؛ به همین دلیل،

در نقطه ثابت x_1 ، خروجی شبکه بالا یعنی y_1 ، دقیقاً با همان x_1 ورودی برابر خواهد بود؛ ولی برای هر نقطه x_2 دلخواه غیر از x_1 که به ورودی شبکه داده شود، y_2 برابر می‌شود با:

$$y_1 = f(x_1 w_1) \quad (5)$$

$$y_2 = f(x_2 w_1) = f(x_1 w_1 + \Delta x w_1) \quad \Delta x = x_2 - x_1 \quad (6)$$

$$\Delta y = y_2 - y_1 = f(x_1 w_1 + \Delta x w_1) - f(x_1 w_1) \quad (7)$$

اگر $f(x_2 w_1)$ حول x_1 بسط داده شود:

$$\begin{aligned} f(x_2 w_1) &= f(x_1 w_1 + \Delta x w_1) \\ &= f(x_1 w_1) + \frac{df(x_1 w_1)}{d(x_1 w_1)} \cdot \frac{d(x_1 w_1)}{d(x_1)} \Delta x \\ &\quad + \frac{1}{2} \frac{d^2 f(x_1 w_1)}{d^2(x_1^2)} \cdot \Delta x^2 + \dots \quad (8) \\ &\cong f(x_1 w_1) + \frac{df(x_1 w_1)}{d(x_1 w_1)} \cdot w_1 \cdot \Delta x \end{aligned}$$

اگر Δx به اندازه کافی کوچک باشد، از جمله دوم به بعد صرف نظر می‌شود؛ در نتیجه Δy برابر است با:

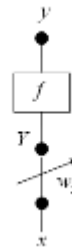
$$\Delta y = y_2 - y_1 = f(x_2 w_1) - f(x_1 w_1) = \frac{df(x_1 w_1)}{d(x_1 w_1)} \cdot w_1 \cdot \Delta x \quad (9)$$

$$\frac{\Delta y}{\Delta x} = \frac{df(x_1 w_1)}{d(x_1 w_1)} \cdot w_1 = \frac{df(x_1 w_1)}{d(x_1)} \quad (10)$$

به ازای $\Delta x > \Delta y$ تغییرات خروجی نورون حول y_1 ناشی از Δx ، همواره از آن کوچک‌تر خواهد بود؛ بنابراین با دادن $x_2 = x_1 + \Delta x$ به ورودی نورون، پس از یک دورزدن، مقدار ورودی نورون به x_1 نزدیک‌تر می‌شود و با ادامه دورزدن در نهایت به سمت x_1 جذب خواهد شد. باید توجه داشت $\Delta x > \Delta y$ شرط نقطه ثابت پایدار مجانبی بودن x_1 را برای تابع $f(x_1 w_1)$ تضمین می‌کند که این خود تأییدکننده جاذب بودن نقطه تعلیم داده شده است. این شرط به صورت رابطه (۱۱) بیان می‌شود که با شرط پایداری مجانبی نقطه ثابت در [۱۳] نیز مطابقت دارد.

$$\left| \frac{\Delta y}{\Delta x} \right| = \left| \frac{df(x_1 w_1)}{d(x_1)} \right| < 1 \quad (11)$$

اگر یک اتصال بازگشتی تأخیردار از خروجی به ورودی برقرار شود، نقطه تعلیم داده شده در شرایطی به‌عنوان یک نقطه ثابت پایدار در شبکه عمل می‌کند؛ یعنی به‌ازای ورودی x_1 ، خروجی همواره x_1 خواهد بود یا به‌عبارتی، خروجی و ورودی شبکه در مقدار x_1 پایدار باقی می‌مانند. همچنین، برای هر مقدار x دیگر که به‌عنوان نقطه آغازین به ورودی این شبکه داده می‌شود، پس از مدتی تغییرات، شبکه به نقطه x_1 همگرا خواهد شد. به‌عبارت دیگر، تمامی محور x ورودی، بستر جذب این نقطه ثابت پایدار خواهند بود و نقطه x_1 در فضای ورودی همانند قعر یک دره عمل می‌کند که سایر نقاط را به خود جذب می‌کند.



شکل (۱): ساختار یک شبکه تک‌لایه با یک نورون غیرخطی

به منظور ساده‌سازی مسئله، تحلیل‌های اولیه روی یک شبکه تک‌لایه با ساختار شکل (۱) صورت می‌گیرد. این شبکه یک شبکه تک‌لایه با یک نورون با تابع غیرخطی f است که به‌صورت خودانجمنی تعلیم داده می‌شود. مقدار w_1 با کمینه‌کردن خطای تعلیم تعیین می‌شود که با رابطه ۱ تعریف می‌شود و در آن، d خروجی مطلوب شبکه است.

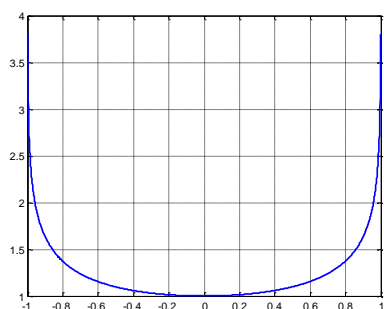
$$E = 1/2(d - y)^2 \quad (1)$$

با توجه به اینکه یک نمونه به‌صورت خودانجمنی تعلیم داده می‌شود، خطای تعلیم با رابطه (۲) برابر خواهد بود که با کمینه‌کردن این خطا w_1 به دست می‌آید.

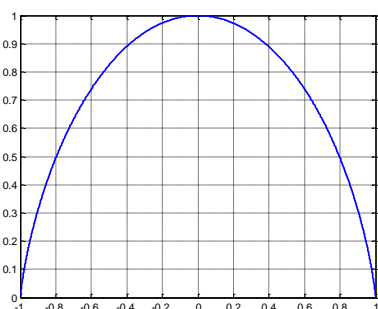
$$E = 1/2[x_1 - f(x_1 w_1)]^2 \quad (2)$$

$$\begin{aligned} \frac{\partial E}{\partial w_1} &= 0 \xrightarrow{w_1=?} [x_1 - f(x_1 w_1)] \frac{df(x_1 w_1)}{d(x_1 w_1)} x_1 \\ &= 0 \end{aligned} \quad (3)$$

$$x_1 = f(x_1 w_1) \xrightarrow{w_1=?} w_1 = \frac{1}{x_1} f^{-1}(x_1) \quad (4)$$



الف



ب

شکل (۳): (الف) نمودار مقادیر w_1 برای \tanh مطابق با

رابطه ۱۲ برای $-1 < x_1 < 1$ و (ب)، نمودار $\frac{d \tanh(x_1 w_1)}{d(x_1)}$

برای مقادیر مختلف $-1 < x_1 < 1$

همان‌طور که مشاهده می‌شود، w_1 در این بازه، مثبت و بزرگ‌تر از یک است؛ بنابراین با توجه به شکل (۲)، تابع فوق خط $y = x$ را در سه نقطه قطع می‌کند؛ یکی نقطه صفر است که نقطه ناپایدار یا دافع است و دو نقطه دیگر، دو نقطه قرینه هستند که هر دو پایدار و جاذبند. این مسئله با استفاده از شرط پایداری نیز نشان داده می‌شود. برای جاذب‌بودن، نقطه تعلیم داده شده، لازم است قدرمطلق تابع رابطه (۱۳) کمتر از یک باشد.

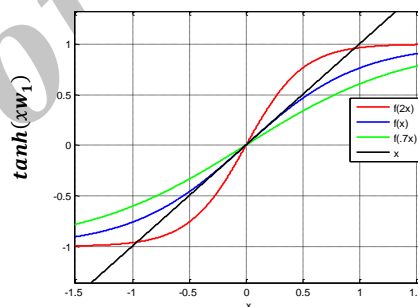
$$\begin{aligned} \frac{df(x_1 w_1)}{d(x_1)} &= (1 - \tanh^2(x_1 w_1)) \cdot w_1 \\ &= \frac{1}{2x_1} (1 - x_1^2) \ln\left(\frac{1+x_1}{1-x_1}\right) \end{aligned} \quad (13)$$

برای مقادیر مختلف $-1 < x_1 < 1$ در نمودار (۳-ب) رسم شده است. همان‌طور که مشاهده می‌شود، همواره مقداری بین صفر و یک دارد که با توجه به شرط پایداری مجانبی نقطه ثابت در [۱۳]، برای هر نقطه

همان‌طور که مشاهده می‌شود تابع f برای جاذب‌بودن نمونه x_1 نقش تعیین‌کننده‌ای دارد. با توجه به اینکه معمولاً دو تابع تانزانست هیپرپولیک و سیگموئید در ساختارها استفاده می‌شوند، در ادامه این دو تابع بررسی خواهند شد.

۱-۱-۱- تابع غیرخطی دوقطبی تانزانست هیپرپولیک

در شکل (۲)، تابع $\tanh(x_1 w_1)$ برای مقادیر مختلف w_1 نشان داده شده است. همان‌طور که مشاهده می‌شود، تابع فوق برای $w_1 < 1$ خط $y = x$ را در سه نقطه قطع می‌کند؛ یکی نقطه صفر است که می‌توان نشان داد نقطه ناپایدار یا دافع برای تابع است و دو نقطه دیگر، دو نقطه قرینه هستند که هر دو پایدار و جاذب‌اند؛ اما برای مقادیر $w_1 > 1$ یک تقاطع با خط $y = x$ در صفر است که جاذب است.

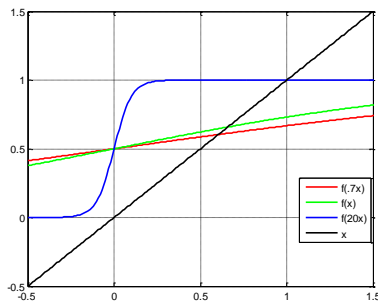


شکل (۲): تابع $\tanh(xw_1)$ برای مقادیر مختلف w_1

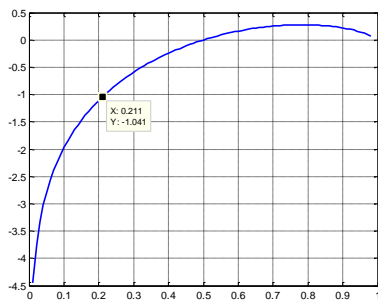
در شبکه تک‌نورونی خودانجمنی شکل (۱)، با فرض تعلیم کامل، مقدار وزن برای تابع تانزانست هیپرپولیک براساس رابطه ۴ برابر است با:

$$\begin{aligned} x_1 = f(x_1 w_1) &= \tanh(x_1 w_1) \xrightarrow{w_1=?} w_1 \\ &= \frac{1}{x_1} \tanh^{-1}(x_1) \\ &= \frac{1}{2x_1} \ln\left(\frac{1+x_1}{1-x_1}\right) \end{aligned} \quad (12)$$

از آنجا که x_1 مقداری بین ۱ و -۱ (خروجی تابع \tanh بین این دو مقدار است) دارد، مقدار w_1 که در رابطه (۱۲) محاسبه شده است برای مقادیر مختلف $-1 < x_1 < 1$ در نمودار (۳-الف) رسم شده است.



الف



ب

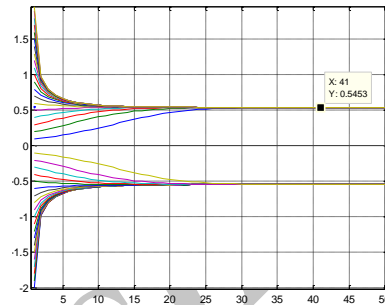
شکل (۵): (الف)، تابع $\text{logsig}(x)$ و نقاط آن با خط $y = x$

و (ب)، نمودار $\frac{d\text{logsig}(x_1)}{d(x_1)}$ برای مقادیر مختلف $0 < x_1 < 1$

مشتق تابع سیگموئید برای مقادیر مختلف $0 < x_1 < 1$ (خروجی تابع سیگموئید در این بازه است) در نمودار (۵-ب) نشان داده شده است. همان‌طور که مشاهده می‌شود، این تابع برای $x_1 < 0.211$ شرط جاذب نقطه‌ای بودن را دارد؛ اما برای نمونه‌های تعلیمی کمتر از این مقدار، x_1 یک نقطه ثابت دافع خواهد شد که بررسی‌ها نشان می‌دهد این نقطه در مرز بستر جاذب دو جاذب نوسانی قرار می‌گیرد؛ به طوری که خروجی برای $x_1 < x_0$ جاذب یکی از این جاذب‌ها و برای $x_1 > x_0$ جاذب نوسانی دیگر می‌شود.

در شکل (۶)، نحوه همگراشدن نمونه‌ها برای نورون خودانجمنی با تابع غیرخطی سیگموئید برای ورودی‌های تعلیم داده شده مختلف نشان داده شده است. همان‌طور که مشاهده می‌شود برای دو نمونه $x_1 = 0.22$ و $x_1 = 0.9$ که بزرگتر از 0.211 هستند در نقطه تعلیم داده شده جاذب نقطه‌ای ایجاد شده است و تمام فضای ورودی به آن جذب شده‌اند؛ درحالی‌که برای دو نمونه دیگر جاذب‌های نوسانی مشاهده می‌شود.

تعلیمی در بازه $[-1, 1]$ ، دو جاذب برای آن نقطه و قرینه آن ایجاد می‌شود. شکل (۴) نیز نحوه جذب شدن نمونه‌ها به نقطه تعلیم داده شده (0.54) و قرینه آن و دفع شدن از آن را در شبکه تک‌لایه شکل (۱) با نورون غیرخطی با تابع \tanh نشان می‌دهد.



شکل (۴): نحوه جذب شدن نمونه‌ها با مقادیر اولیه مختلف به نقطه تعلیم داده شده (0.54) و قرینه آن و دفع شدن از آن

۱-۲- تابع غیرخطی تک‌قطبی سیگموئید

تابع دیگری که بیشتر استفاده می‌شود، تابع سیگموئید است که در نمودار (۵-الف) نشان داده شده است. همان‌طور که مشاهده می‌شود، تابع فوق خط $y = x$ را در یک نقطه قطع می‌کند؛ اما جاذب یا دافع بودن این نقطه به بررسی بیشتری نیاز دارد. در شبکه تک‌نورونی خودانجمنی شکل (۱)، با توجه به رابطه (۴)، مقدار وزن برای تابع سیگموئید برابر است با:

$$\begin{aligned} x_1 &= f(x_1 w_1) = \text{logsig}(x_1 w_1) \\ &= \frac{1}{1 + e^{-x_1 w_1}} \xrightarrow{w_1=?} w_1 \quad (14) \\ &= -\frac{1}{x_1} \ln\left(\frac{1}{x_1} - 1\right) \end{aligned}$$

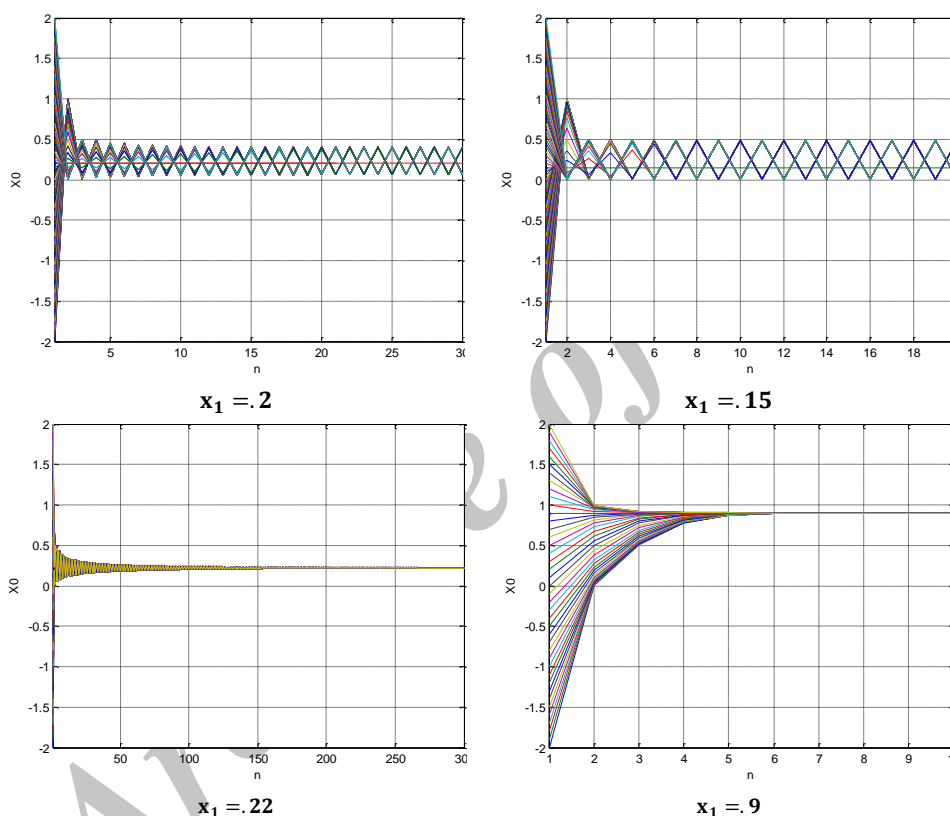
که در نتیجه

$$\begin{aligned} \frac{df(x_1 w_1)}{d(x_1)} &= x_1(1 - x_1)w_1 \quad (15) \\ &= -(1 - x_1) \ln\left(\frac{1}{x_1} - 1\right) \end{aligned}$$

برای جاذب بودن، نقطه تعلیم داده شده، لازم است قدرمطلق این تابع کمتر از یک باشد.

مستقیم عمل می‌کند. همچنین، اگر نمونه دیگر $x_1 \neq x_2$ به ورودی شبکه داده شود، پس از محاسبه خروجی و سپس دورزدن در مسیر بازگشتی، x_2' به دست می‌آید که تخمین غیرخطی شبکه از x_2 است و به جاذب x_1 نزدیک‌تر خواهد بود. با دورزدن در این ساختار در نهایت x_2 به x_1 جذب خواهد شد.

این بررسی‌ها نشان داد برای یک نورون با تابع غیرخطی تانژانت هیپرپولیک برای هر نقطه تعلیم داده شده و با تابع غیرخطی سیگموئید برای $x_1 < 0.211$ در نقطه تعلیم داده شده جاذب تشکیل می‌شود؛ بنابراین، اگر ورودی تعلیم داده شده یعنی x_1 به ورودی این شبکه داده شود، خروجی شبکه دقیقاً برابر با همان x_1 خواهد بود؛ یعنی در نقطه x_1 نگاشت معکوس دقیقاً برابر با معکوس نگاشت



شکل (۶): نمودارهای همگرایی نمونه‌ها با مقادیر اولیه مختلف برای نقاط تعلیم داده شده به نورون غیرخطی با تابع $\text{logsig}(x)$

$$w_1 = \frac{1}{x_1} (\tanh^{-1}(x_1) - b) \quad (16)$$

$$= \frac{1}{x_1} \left(\ln \sqrt{\frac{1+x_1}{1-x_1}} - b \right)$$

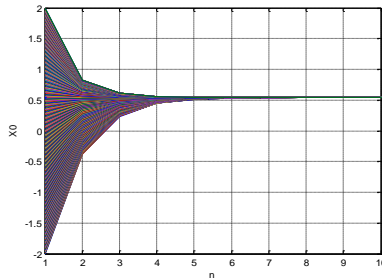
$$\frac{df(x_1 w_1 + b)}{d(x_1)} = \frac{1}{x_1} (1 - x_1^2) \left(\ln \sqrt{\frac{1+x_1}{1-x_1}} - b \right) \quad (17)$$

$$w_1 = -\frac{1}{x_1} \left(\ln \left(\frac{1}{x_1} - 1 \right) + b \right) \quad (18)$$

۱-۳- تأثیر سطح آستانه

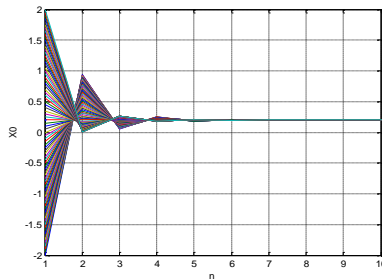
اگر در ورودی شبکه تک‌نورونی خودانجمنی شکل (۱)، مقدار سطح آستانه نیز اضافه شود، وجود آن هم در مقدار وزن و هم در مشتق تابع غیرخطی نورون مطابق روابط (۱۶) و (۱۷) برای تابع تانژانت هیپرپولیک و روابط (۱۸) و (۱۹) برای تابع سیگموئید تأثیرگذار است؛ بنابراین، رفتار نورون را تغییر می‌دهد.

درحالی که در نمودار (۸-الف) مشاهده می شود در صورت وجود سطح آستانه فقط یک جاذب در x_1 تشکیل می شود.



$x_1 = .54$

الف



$x_1 = .2$

ب

شکل (۸): نمودارهای همگرایی نمونه‌ها با مقادیر اولیه مختلف برای نقاط تعلیم داده شده به نورون غیرخطی به همراه سطح آستانه برای تابع تانژانت هیپرپولیک $x_1 = .54$ (ب) و تابع

سیگموئید $x_1 = .2$

همچنین در شکل (۷-ب) برخلاف شکل (۶) نشان داده شده است برای نمونه تعلیم داده شده $x_1 = .2$ به نورون غیرخطی با تابع سیگموئید به همراه سطح آستانه، جاذب به دست می آید. این بررسی‌ها نشان می دهد سطح آستانه به طور مستقیم در رفتار نورون تأثیرگذار است و در تشکیل جاذب‌ها نقش به سزایی دارد.

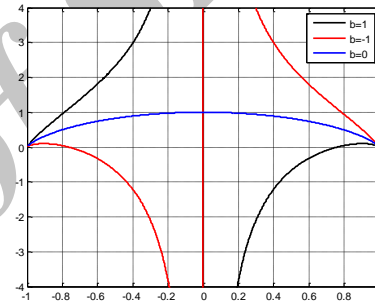
۲-۱- تحلیل رفتار جاذب گونه در شبکه

جلوسوی تک لایه

شبکه شکل (۹) را با n نورون در ورودی و خروجی در نظر بگیرید. فرض کنید به این شبکه تعدادی نمونه n بعدی به صورت (x_1, x_2, \dots, x_n) به طور خودانجمنی تعلیم داده شود. اگر خروجی شبکه به صورت

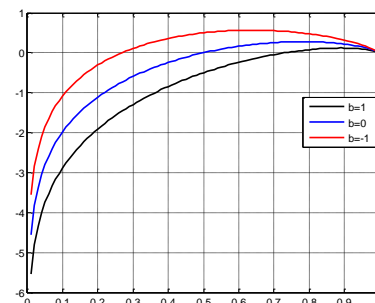
$$\frac{df(x_1 w_1 + b)}{d(x_1)} = -(1 - x_1) \left(\ln \left(\frac{1}{x_1} - 1 \right) + b \right) \quad (19)$$

شکل (۷)، تأثیر سطح آستانه را در تغییر رفتار نورون نشان می دهد. شکل (۷-الف) نشان می دهد برای تابع تانژانت هیپرپولیک برای $b \neq 0$ برای برخی مقادیر x_1 ، مقداری بزرگتر از یک خواهد داشت؛ در نتیجه در این نقاط و قرینه آنها جاذب تشکیل نخواهد شد. همچنین، با توجه به نمودار (۷-ب)، با تغییر b می توان آستانه مقدار x_1 را برای ایجاد جاذب تنظیم کرد. با کاهش b این آستانه کاهش می یابد.



x_1

الف



x_1

ب

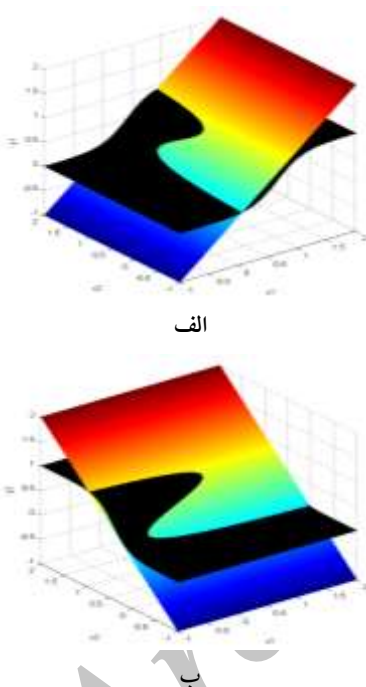
شکل (۷): نمودار $\frac{df(x_1 w_1 + b)}{d(x_1)}$ برای مقادیر مختلف x_1 ، (الف) برای تابع تانژانت هیپرپولیک و (ب) برای تابع سیگموئید

شکل (۸) نیز این تغییر رفتار نورون تأثیر گرفته از سطح آستانه را نشان می دهد. برای تابع تانژانت هیپرپولیک در شکل (۴) نشان داده شد که برای نمونه تعلیم داده شده $x_1 = .54$ ، دو جاذب در x_1 و $-x_1$ تشکیل می شود؛

نقاط تعادل کل شبکه نقاطی هستند که برای تمام نوروهای خروجی شبکه، نقطه تعادل باشند؛ بنابراین، این نقاط به صورت رابطه (۲۲) تعریف می‌شوند.

$$X_{f,p} = \{T_1: T_1 = T_2 = \dots = T_n\} \quad (22)$$

به عبارت دیگر، نقاط تقاطع ابرویه‌های T_1, T_2, \dots, T_n نقاط تعادل شبکه را تشکیل می‌دهند. برای روشن‌تر شدن مسئله، رویه‌های تصمیم برای شبکه شکل (۹) در نظر گرفته می‌شود که با ۳ نمونه دو بعدی ($n = 2$) تعلیم داده شده است.



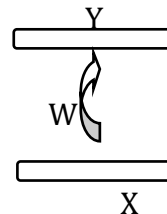
الف

ب

شکل (۱۰): نمودارهای رویه‌های y_1 و y_2 که به ترتیب با صفحات $x_1 = y_1$ و $x_2 = y_2$ قطع داده شده‌اند.

این رویه‌ها در شکل (۱۰) نشان داده شده‌اند. ابتدا باید نقاط تعادل در خروجی هر نورون تعیین شود. نقاط تعادل در خروجی y_1 نقاطی خواهند بود که برای آنها همواره $x_1 = y_1$ باشد؛ بنابراین، رویه y_1 با صفحه $x_1 = y_1$ مطابق نمودار (الف-۱۰) قطع می‌شود. نقاط روی منحنی تقاطع این دو، نقاط تعادل برای خروجی y_1 شبکه هستند. برای خروجی y_2 در (ب-۱۰) نیز نقاط تعادل، نقاط روی محل تقاطع رویه y_2 با صفحه $x_2 = y_2$ هستند. شکل (۱۱) نمودارهای شکل (۱۰) را در صفحه (x_1, x_2) نشان می‌دهد.

تعریف شود برای هر نقطه ثابت باید معادلات رابطه (۲۰) برقرار باشد.

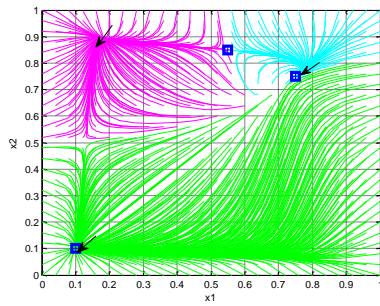


شکل (۹): شبکه جلوسوی تک لایه

$$\begin{aligned} x_1 = y_1 &= f(x_1 w_{11} + x_2 w_{21} + \dots + x_n w_{n1} + b_1) \\ x_2 = y_2 &= f(x_1 w_{12} + x_2 w_{22} + \dots + x_n w_{n2} + b_2) \\ &\vdots \\ x_n = y_n &= f(x_1 w_{1n} + x_2 w_{2n} + \dots + x_n w_{nn} + b_n) \end{aligned} \quad (20)$$

با دقت در روابط فوق مشاهده می‌شود که اطلاعات سایر نوروها برای تابع غیرخطی هر نورون خروجی مانند سطح آستانه عمل می‌کند با این تفاوت که مقدار آنها متغیر است. همچنین در بخش قبل نشان داده شد سطح آستانه بر رفتار یک نورون غیرخطی تأثیر مستقیم دارد؛ به گونه‌ای که رفتار آن را از حالت سه نقطه ثابت به شرایط ناپایدار تغییر می‌دهد. با این تفسیر در اینجا از خروجی هر یک از نوروها به‌ازای ورودی ثابت در همان بعد می‌توان رفتارهای مختلفی را انتظار داشت که بررسی آنها برای به‌دست آوردن رفتار کلی شبکه الزامیست. ابتدا باید نقاط تعادل در خروجی هر نورون تعیین شود. نقاط تعادل در خروجی y_i که در آن $i = 1, \dots, n$ ، نقاطی خواهند بود که برای آنها همواره $x_i = y_i$ باشد؛ بنابراین با تعیین نقاط تقاطع ابرویه y_i و ابرصفحه $x_i = y_i$ می‌توان این نقاط تعادل را به دست آورد. این نقاط تقاطع که در رابطه (۲۱) نیز تعریف شده‌اند یک ابرویه^۱ با بعد $n - 1$ تشکیل می‌دهند. به عبارت دیگر، T_i مکان هندسی نقاطی از ابرویه y_i است که برای آنها $y_i = x_i$ است.

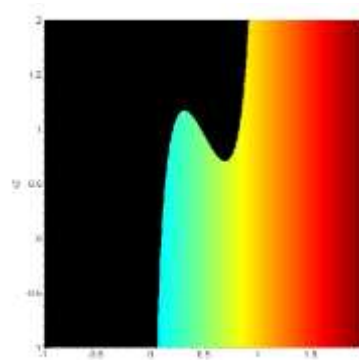
$$T_i = \{y_i: y_i = x_i\} \quad i = 1, \dots, n \quad (21)$$



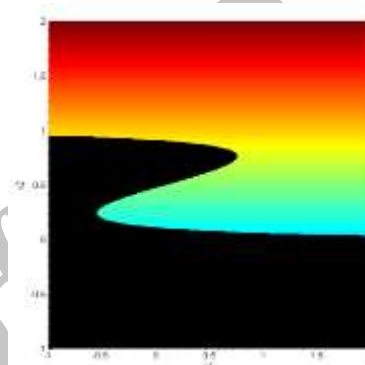
ب

شکل (۱۲): (الف)، نمودار تقاطع منحنی‌های T_1 و T_2 و (ب)، بستر جذب‌ها در فضای ورودی و نقاط تعلیم، مربع‌ها نقاط تعلیم و + خروجی‌های شبکه را به‌ازای این نقاط تعلیم نشان می‌دهد.

برای بررسی نوع تعادل این نقاط، یک‌سری نقاط با توزیع یکنواخت در بازه $[0,1]$ تعریف می‌شوند و با برقراری اتصال بازگشتی از خروجی به ورودی در طی نگاشت‌های تکراری خروجی محاسبه می‌شود. نمودار (۱۲-ب) نحوه جذب شدن نمونه‌های فضای ورودی به جاذب‌های شبکه را نشان می‌دهد. همان‌طور که مشاهده می‌شود این جاذب‌های پایدار همان نقاط تقاطعی هستند که در نمودار (۱۲-الف) با فلش نشان داده شده‌اند؛ اما نکته درخور توجه این است که جاذب‌ها لزوماً در نقاط تعلیم تشکیل نشده‌اند. با وجود تعلیم کامل شبکه به‌گونه‌ای که خروجی به‌ازای هر ورودی دقیقاً خود ورودی شود، مشاهده می‌شود در محل دو مورد از نقاط تعلیم دو نقطه ثابت با تعادل زینی شکل گرفته‌است. برخلاف شبکه تک‌نورونی که با برقراری شرایط حتماً در محل نمونه تعلیم داده شده یک جاذب ایجاد می‌شود، برای شبکه تک‌لایه چندبعدی آرایش نمونه‌های تعلیم در نوع تعادلشان نقش دارد؛ به‌طوری‌که در اینجا ۵ نقطه تعادل ایجاد شده است که از دو مورد جاذب، یک جاذب در محل نمونه‌های تعلیم تشکیل شده است.



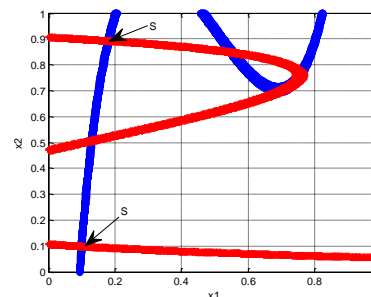
الف



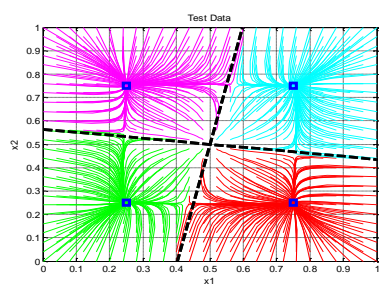
ب

شکل (۱۱): نمودارهای رویه‌های y_1 و y_2 در صفحه (x_1, x_2) است که به ترتیب با صفحات $x_1 = y_1$ و $x_2 = y_2$ قطع داده شده‌اند.

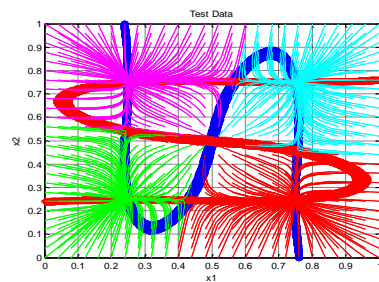
در شکل (۱۲)، تقاطع منحنی‌های T_1 و T_2 نشان داده شده است. همان‌طور که مشاهده می‌شود این دو منحنی در ۵ نقطه همدیگر را قطع می‌کنند که نقاط تعادل شبکه هستند؛ البته نوع تعادل این نقاط به بررسی نیاز دارد.



الف



الف



ب

شکل (۱۴): (الف)، نمونه‌های تعلیم، مرز نواحی تصمیم و جاذب‌ها در فضای ورودی و (ب)، منحنی‌های T_1 و T_2 نمایش جاذب‌ها در محل‌های تقاطع آنها

نمودار (ب) شکل (۱۴) نیز T_1 و T_2 را نشان می‌دهد. همان‌طور که مشاهده می‌شود جاذب‌ها دقیقاً در محل‌های تقاطع آنها تشکیل شده‌اند. T_1 و T_2 در ۵ نقطه دیگر نیز تقاطع دارند که ۴ مورد آنها تعادل زینی دارند و نقطه وسطی، تعادل ناپایدار دارد.

۱-۴- خوشه‌بندی در شبکه خودانجمنی با

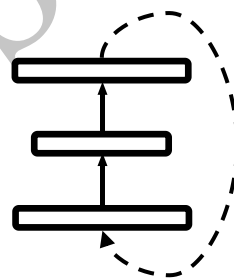
اتصال بازگشتی

گفتنی است اگر تعداد نمونه‌های تعلیمی بیش از قابلیت یادگیری شبکه باشد، شبکه عملاً خوشه‌بندی می‌کند؛ بدین ترتیب هنگامی که شبکه، دو نمونه A و B را در یک خوشه تخمین می‌زند، برای آنها یک جاذب مشترک به دست می‌آید که بر هیچ‌یک از دو نمونه منطبق نیست؛ بنابراین، خطای تخمین هر دو غیرصفر است؛ اما وقتی به یک خوشه فقط یک نمونه تخصیص می‌یابد، خطای تخمین آن می‌تواند صفر باشد و نمونه در جاذب خوشه قرار گیرد. به این ترتیب درباره حالت خوشه‌بندی چند نمونه در یک خوشه، شبکه عملاً از هر نمونه پس از جذب شدن آنها تخمینی ارائه می‌دهد؛ ولی این تخمین غیرخطی توأم با خطا است و هرچه

۱-۳- رفتار جاذب‌گونه در شبکه خودانجمنی

یک‌لایه پنهان

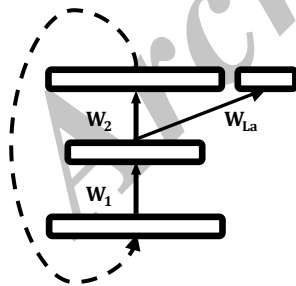
با افزایش تعداد لایه‌های غیرخطی در شبکه خودانجمنی و افزایش عمق آن، T_i ها غیرخطی‌تر می‌شوند؛ بنابراین، ایجاد تقاطع‌های بیشتری بین آنها، ممکن و در نتیجه آن، ظرفیت یادگیری جاذب‌های بیشتری فراهم می‌شود. برای مثال، اگر به شبکه یک‌لایه پنهان شکل (۱۳) با دو نورون غیرخطی در هر لایه، ۴ نمونه نشان داده شده در نمودار (۱۴-الف) تعلیم داده شود، با برقراری اتصال بازگشتی بعد از تعلیم شبکه، جاذب‌های شبکه مشاهده می‌شوند.



شکل (۱۳): ساختار شبکه عصبی خودانجمنی با اتصال بازگشتی

در نمودار (۱۴-الف) نمونه‌های تعلیم، مرز تقریبی نواحی تصمیم با ۲ نورون لایه پنهان، تعیین و نحوه همگرا شدن نمونه‌ها در فضای ورودی نشان داده شده است. در واقع آنچه رخ می‌دهد این است که نواحی تصمیم شبکه بعد از تعلیم، فضای ورودی را به ۴ قسمت افراز کرده‌اند. هر بخش از این فضا، بستر جذب یکی از الگوهای تعلیمی تلقی می‌شود که داده جدید برحسب اینکه در بستر جذب کدام جاذب باشد، به سمت یکی از نمونه‌های تعلیم جذب می‌شود.

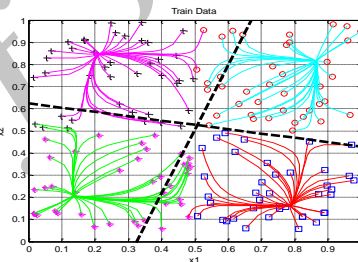
نمودار (۱۵-الف)، نمونه‌های تعلیم، مرز نواحی تصمیم که با ۲ نورون لایه پنهان تعیین شده است و نحوه همگراشدن نمونه‌ها به مرکز خوشه‌ها در فضای ورودی را نشان می‌دهد. در اینجا نیز مشاهده می‌شود ۴ جاذب در محل تقاطع‌های منحنی‌های T_1 و T_2 مطابق نمودار (۱۵-ب) تشکیل شده است. نمودار (۱۵-ج) نیز نحوه جذب شدن سایر نمونه‌های فضای ورودی که در تعلیم حضور نداشته‌اند را به سمت مرکز یکی از این مراکز خوشه‌ها نشان می‌دهد. این مدل یک فیلترسازی غیرخطی مقاوم انجام می‌دهد و نمونه‌های ورودی را به یکی از جاذب‌ها هدایت می‌کند؛ بنابراین با تعلیم مناسب الگوهای مدنظر و تشکیل جاذب‌ها سایر مؤلفه‌های ناخواسته پالایش می‌شود. یک روش برای کنترل بستر جذب‌ها تنظیم مرزهای تصمیم متناسب با مؤلفه‌های متمایزکننده مدنظر است. در [۱۴] نشان داده شده است در ساختار دو تکلیفه شبکه عصبی گلوگاه باسرپرست (SBNN) با استفاده از اطلاعات برچسب‌ها، ابرصفحات که مرزهای تقریبی مرزهای تصمیم هستند، به گونه‌ای قرار می‌گیرند که تمایز بین طبقات را حتی‌الامکان حفظ کنند. با همین روش در اینجا نیز بستر جذب‌ها با هدف جذب نمونه‌های هم‌طبقه به یک خوشه کنترل می‌شود. شکل (۱۶)، ساختار این شبکه عصبی خودانجمنی باسرپرست را نشان می‌دهد.



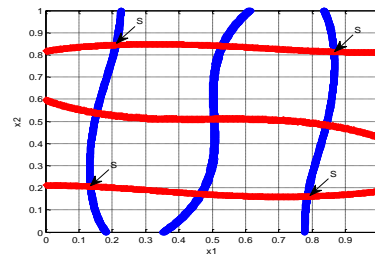
شکل (۱۶): ساختار شبکه عصبی خودانجمنی باسرپرست

نمودارهای شکل (۱۷) تأثیر بهره‌گیری از برچسب طبقات را برای شبکه خودانجمنی باسرپرست نشان می‌دهد. در اینجا نیز همان ۱۶۰ نمونه قبلی تعلیم داده شده است. با مقایسه نمودارهای (۱۵-الف) و شکل (۱۷-الف) مشاهده می‌شود بستر جذب‌ها به گونه‌ای تشکیل شده‌اند که در حد امکان نمونه‌های هم‌برچسب را در یک بستر جذب قرار دهند.

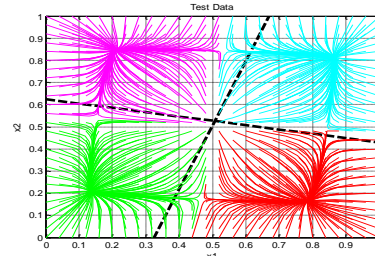
آن نمونه از نقطه جاذب خوشه دورتر باشد این خطا بیشتر خواهد بود. به عبارتی نقطه جاذب خوشه، وجه مشترک (میانگین غیرخطی) نمونه‌های خوشه است و فاصله هر نمونه از جاذب، تفاوت‌های هر نمونه با این میانگین تلقی می‌شود. نمودار (۱۵-الف) مجموعه داده‌ای شامل ۴ دسته را نشان می‌دهد که هر دسته حاوی ۴۰ نمونه است. این الگو به شبکه شکل (۱۳) با دو نورون غیرخطی در هر لایه تعلیم داده می‌شود. در اینجا تعداد نمونه‌های تعلیمی به نسبت قابلیت شبکه خیلی بیشتر است. به عبارتی، ۱۶۰ نمونه برای تعلیم داده شده است؛ در حالی که شبکه قادر است فضا را فقط به ۴ بخش تقسیم کند؛ بنابراین به صورت دادگان، خوشه‌بندی می‌کند. در ساختار این شبکه نیز جاذب‌ها شکل می‌گیرند؛ اما مرکز خوشه‌ها به جای نمونه‌ها به عنوان جاذب عمل می‌کنند.



الف



ب



ج

شکل (۱۵): (الف)، نمونه‌های تعلیم، مرز نواحی تصمیم و نحوه جذب ورودی‌ها به مراکز خوشه‌ها و (ب)، تقاطع‌های منحنی‌های T_1 و T_2 و (ج)، نحوه جذب نمونه‌های آزمون به مراکز خوشه‌ها در شبکه خودانجمنی با اتصال بازگشتی

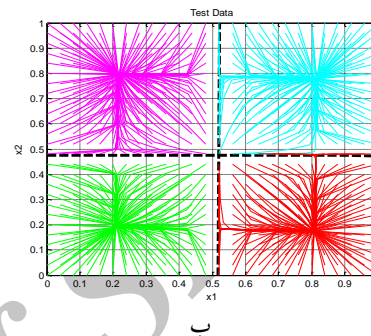
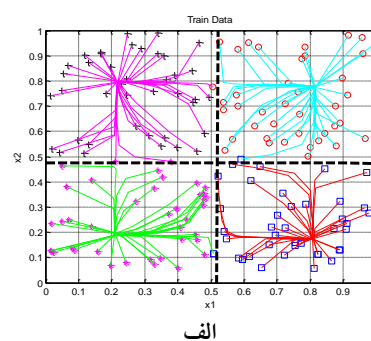
دو مجموعه از این دادگان استفاده شده است. یکی با عنوان پایگاه داده CK+ که شامل تصاویر مربوط به ۹۶ نفر با ۷ حالت احساسی مختلف است. این مجموعه در کل حاوی ۲۸۱۶ تصویر است. دادگان دیگر، پایگاه داده کوچک CK+ است که شامل بخشی از این دادگان است. این داده نیز شامل ۲۸۱ تصویر از ۱۰ نفر با سه حالت خستگی، لبخند و تعجب است.

۲-۲- ارزیابی تأثیر عمق شبکه خودانجمنی در شکل‌گیری جاذب‌ها

در این بخش، تأثیر عمق ساختار در یادگیری جاذب‌ها در شبکه‌های عصبی خودانجمنی ارزیابی شده است. بدین منظور، جاذب‌های دو شبکه خودانجمنی کم‌عمق و عمیق بعد از یادگیری خودانجمنی تصاویر پایگاه داده CK+ بررسی شده است. برای شبکه خودانجمنی یک لایه پنهان، ۱۰۰۰ نورون در لایه پنهان آن و همچنین تعداد نورون‌های لایه‌های پنهان شبکه خودانجمنی عمیق نیز ۱۰۰۰-۴۰۰-۱۰۰۰ قرار داده شده است. برای پیش‌تعلیم شبکه خودانجمنی عمیق نیز از روش پیش‌تعلیم لایه‌به‌لایه [۱۷] استفاده شده است. برای تعیین جاذب‌های هر مدل از نمونه‌های تعلیم داده شده به مدل‌ها استفاده شد. در نتیجه آن برای شبکه یک لایه پنهان، ۱۲ نقطه ثابت پایدار مجانبی حاصل شد که در شکل (۱۸) نشان داده شده‌اند. این جاذب‌ها هیچ‌یک از تصاویر تعلیم داده شده نیستند. با توجه به حجم بالای تصاویر تعلیم داده شده و ساختار یک لایه پنهان شبکه، این شبکه قادر به یادگیری هر یک از نمونه‌ها در یک ناحیه تصمیم مجزا نیست. در نتیجه، نمونه‌ها را خوشه‌بندی می‌کند و برای تمام نمونه‌هایی که در یک ناحیه تصمیم قرار دارند یک جاذب مشترک ارائه می‌دهد. این جاذب مشترک، نماینده خوشه مربوطه است.



شکل (۱۸): نقاط تعادل پایدار مجانبی شبکه خودانجمنی یک لایه پنهان تعلیم داده شده با داده CK+ برای نقاط شروع از نمونه‌های تعلیم



شکل (۱۷): (الف)، نمونه‌های تعلیم، مرز نواحی تصمیم و نحوه جذب ورودی‌ها به مراکز خوشه‌ها و (ب)، نحوه جذب نمونه‌های آزمون به مراکز خوشه‌ها در شبکه عصبی خودانجمنی با سرپرست

۲- پیاده‌سازی‌ها

۲-۱- دادگان

در این مقاله از پایگاه داده CK+ استفاده شده است. پایگاه داده CK+ توسعه‌یافته پایگاه داده کوهن-کند است [۱۵]. در پایگاه داده CK+ دنباله‌ای از تصاویر برای ۶ حالت احساسی افراد موجود است که از حالت خستگی تا حالت مدنظر تغییر می‌کنند؛ البته در این مقاله تصاویر نیمه اول هر یک از این دسته‌ها حذف شدند تا تصویری نگه داشته شوند که واقعاً حالت احساسی مربوطه را نمایان می‌کنند. این حالت‌های احساسی شامل حالت‌های شاد، متعجب، خشمگین، ترسیده، متنفر و ناراحت هستند. همچنین برای هر تصویر چهره، برداری از نشانه‌ها وجود دارد که از آن برای تراز کردن چهره‌ها استفاده می‌شود. در اینجا چهره‌ها طوری تراز شده‌اند که چشم‌ها در یک راستای افقی قرار بگیرند [۱۶]. اطلاعات اضافه اطراف تصاویر نیز به‌نحوی حذف شده است که همه جزئیات مهم در تصویر موجود باشند. در نهایت، ابعاد همه تصاویر به ۵۰×۵۰ تبدیل شدند.



شکل (۲۱): مسیر نمونه برداری شده رفتار نوسانی حول تصویر ۱، تصویر اول از سمت چپ در سطر اول، نقطه شروع است. تصاویر از سطر ۸ به بعد مجدداً تکرار شده‌اند.



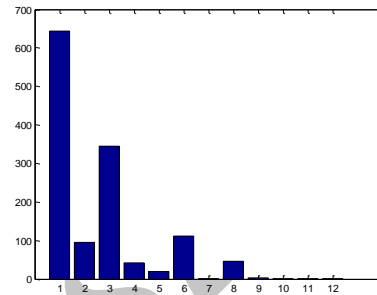
شکل (۲۲): نقاط تعادل پایدار مجانبی شبکه خودانجمنی عمیق تعلیم داده شده با داده CK+ برای نقاط شروع از نمونه‌های تعلیم.

شکل (۲۳) نیز نحوه جذب شدن برای دو تا از تصویرهای ورودی را نشان می‌دهد. همان‌طور که مشاهده شد با افزایش تعداد لایه‌های شبکه، امکان ذخیره جاذب‌های بیشتری فراهم شد؛ اما به نظر می‌رسد برای یادگیری تمام نمونه‌های تعلیم به‌عنوان جاذب به ساختار بسیار بزرگ‌تری نیاز باشد.



شکل (۲۳): مسیر نمونه برداری شده جذب به نقاط تعادل پایدار مجانبی شبکه عمیق با شروع از ۲ نمونه تعلیم، تصویر اول از سمت چپ، نقطه شروع است.

از مجموع ۲۸۱۶ نمونه تعلیم داده شده، ۱۳۱۳ نمونه جذب این جاذب‌ها شدند که توزیع آنها نیز در نمودار شکل (۱۹) نشان داده شده است. همان‌طور که مشاهده می‌شود این توزیع یکنواخت نیست و برای برخی جاذب‌ها نسبت به سایرین، نمونه‌های بسیار بیشتری در بستر جذبشان قرار گرفته‌اند.



شکل (۱۹): توزیع نمونه‌های جذب شده به نقاط تعادل پایدار شبکه خودانجمنی یک لایه پنهان

رفتار شبکه برای ۱۵۰۳ نمونه دیگر درخور توجه است. مشاهده شد با شروع از این نمونه‌ها در نهایت شبکه حول یکی از تصاویر نشان داده شده در شکل (۲۰) رفتار نوسانی خواهد داشت. شکل (۲۱) این رفتار نوسانی را حول تصویر ۱ شکل (۲۰) نشان می‌دهد.



شکل (۲۰): نقاط تعادلی که شبکه یک لایه پنهان تعلیم داده شده با داده CK+ به‌ازای برخی شرایط اولیه حول آنها رفتار نوسانی دارد.

در ادامه، وزن‌های شبکه یک لایه پنهان، وزن‌های پیش‌تعلیم یافته شبکه خودانجمنی عمیق فرض شد و بعد از پیش‌تعلیم سایر وزن‌ها تعلیم یک پارچه آن صورت گرفت. سپس رفتار شبکه به‌ازای نقطه شروع از ورودی‌های تعلیم داده شده بررسی شد. مشاهده شد شبکه خودانجمنی عمیق دارای ۵۶ جاذب مطابق شکل (۲۲) خواهد بود که برخی از آنها نیز دقیقاً تصاویر تعلیم هستند.



شکل (۲۵): نقاط تعادل پایدار مجانبی شبکه خودانجمنی عمیق
تعلیم داده شده با داده کوچک CK+ با روش پیش‌تعلیم
لایه‌به‌لایه برای نقاط شروع از نمونه‌های تعلیم.

همچنین، شکل‌گیری این جاذب‌های نسبتاً مشابه به دلیل وجود چند تصویر از یک حالت فرد در دادگان است که شبکه برای آنها نواحی تصمیم مجزا ایجاد کرده است. بار دیگر شبکه خودانجمنی عمیق با روش لایه‌به‌لایه باسرپرست پیش‌تعلیم شد و بعد از تعلیم یک‌پارچه، جاذب‌های آن به‌زای شروع از نقاط تعلیم تعیین شدند که شکل (۲۶) این جاذب‌ها را نشان می‌دهد. در اینجا در طی پیش‌تعلیم حفظ اطلاعات متمایزکننده حالت مدنظر قرار داده شده است.



شکل (۲۶): جاذب‌های شبکه خودانجمنی عمیق تعلیم داده
شده با داده کوچک CK با روش پیش‌تعلیم لایه‌به‌لایه
باسرپرست برای نقاط شروع از نمونه‌های تعلیم.

شکل (۲۷) نیز مسیر نمونه‌برداری شده جذب به یکی از این نقاط تعادل پایدار را نشان می‌دهد.



شکل (۲۷): مسیر نمونه‌برداری شده جذب به یکی از نقاط
تعادل پایدار مجانبی شبکه خودانجمنی عمیق تعلیم داده شده با
داده کوچک CK+ با روش پیش‌تعلیم لایه‌به‌لایه باسرپرست،
تصویر اول از سمت چپ، نقطه شروع است.

برای بررسی تأثیر اطلاعات برجسب در طی مراحل پیش‌تعلیم در نحوه شکل‌گیری جاذب‌ها، ماتریس اختلاط^۳ بین حالت‌های مختلف برای شبکه پیش‌تعلیم شده با هریک از روش‌ها ارزیابی شده است. این ماتریس‌های اختلاط در جدول (۱) و جدول (۲) آورده شده‌اند.

۲-۳- ارزیابی تأثیر روش پیش‌تعلیم با دو روش پیش‌تعلیم در شکل‌گیری جاذب‌ها

با توجه به اینکه در آزمایش قبل مشاهده شد به دلیل حجم بالای نمونه‌های تعلیم فقط تعداد کمی از آنها نقاط تعادل شبکه خودانجمنی عمیق نیز شده‌اند. در این بخش از دادگان کوچک CK+ استفاده می‌شود. همان‌طور که قبلاً گفته شده است این دادگان شامل تصاویر مربوط به سه حالت خشتی، تعجب و لبخند برای ده نفر هستند که تصاویر خشتی این ده نفر در شکل (۲۴) نشان داده شده است.



شکل (۲۴): تصاویرهای خشتی برای هریک از افراد داده کوچک
CK+.

ساختار شبکه خودانجمنی عمیق نیز مانند قسمت قبل است فقط تعداد نورون‌های لایه گلوگاه به ۲۰۰ افزایش داده شده است. برای پیش‌تعلیم آن از دو روش پیش‌تعلیم لایه‌به‌لایه [۱۷] و لایه‌به‌لایه باسرپرست [۱۴] استفاده شد. مراحل پیش‌تعلیم نقش تعیین‌کننده‌ای در تنظیم ابرصفحات دارند و از سوی دیگر این ابرصفحات هستند که مرزهای تقریبی بستر جذب‌ها را تعیین می‌کنند؛ بنابراین با جهت‌دهی به نحوه تنظیم آنها در مراحل پیش‌تعلیم در چگونگی تشکیل بستر جذب جاذب‌ها مداخله می‌شود. برای شبکه پیش‌تعلیم داده شده با روش لایه‌به‌لایه مشاهده شد که با شروع از ۲۸۱ نمونه تعلیم، در نهایت خروجی به ۴۰ نقطه ثابت پایدار مجانبی همگرا می‌شود که در شکل (۲۵) نشان داده شده است. همان‌طور که مشاهده می‌شود تمام این نقاط ثابت دقیقاً روی نمونه‌های تعلیم تشکیل شده‌اند و فقط یکی از آنها کمی فاصله دارد. به عبارتی با کاهش نمونه‌های تعلیم امکان ذخیره آنها به عنوان نقاط ثابت پایدار در شبکه فراهم شده است. هرچند برخی از این جاذب‌ها ظاهراً مشابه هستند، بررسی دقیق‌تر آنها نشان می‌دهد تفاوت اندکی با هم دارند و این تفاوت اندک در فضای ورودی به اختلاف بیشتر در لایه میانی شبکه نیز منجر شده است؛ برای مثال، دو تصویر پنجم و ششم در سطر دوم در فضای ورودی، میانگین مجذور اختلاف ۰/۰۰۴ را دارند؛ در حالی که در لایه میانی این مقدار ۰/۰۵۴۳ است.

در ادامه به منظور رفع این مسئله، تصاویر لبخند فرد ۸م به عنوان داده نامناسب، حذف و پیش‌تعلیم با سرپرست مجدد انجام شد. نتایج این ارزیابی در جدول (۳) آورده شده است.

جدول (۳): ماتریس اختلاط حالت‌ها برای جاذب‌های شبکه خودانجمنی عمیق با روش پیش‌تعلیم لایه‌به‌لایه با سرپرست با حذف نمونه نامناسب

برچسب حالت جاذب‌ها		
تعب	لبخند	خنثی
تعب	لبخند	خنثی
۱۴/۲۹٪	۲۲/۸۶٪	۶۰٪
۰٪	۹۰/۴۳٪	۹/۵۷٪
۹۲/۳۱٪	۰٪	۷/۶۹٪

در یک مقایسه کلی، متوسط عناصر روی قطره‌های این ماتریس‌های اختلاط با در نظر گرفتن تعداد نمونه‌های هر دسته برای DBNN با روش لایه‌به‌لایه ۵۲/۶۷٪ است؛ درحالی‌که با جهت‌دهی به نحوه تشکیل جاذب‌ها با روش لایه‌به‌لایه با سرپرست این میزان تا ۸۱/۱۴٪ و با حذف تصاویر لبخند فرد ۸م و پیش‌تعلیم لایه‌به‌لایه با سرپرست تا ۸۷/۲۷٪ رشد داشته است. همچنین در مقایسه نتایج جدول (۲) و درک‌های انسانی از حالت‌های احساسی [۱۸]، همبستگی‌هایی مشاهده می‌شود؛ به طوری‌که هر دو در اختلاط بین حالت‌های مختلف نقاط مشترکی دارند؛ به گونه‌ای‌که برای هر دو، اختلاط‌های تصاویر خنثی با لبخند نسبت به اختلاط با تعجب و لبخند با خنثی نسبت به اختلاط با تعجب بیشتر هستند.

۳- جمع‌بندی

یکی از کاراترین قابلیت‌های مغز انسان، عملکرد آن به عنوان حافظه انجمنی است. این حافظه پیونددهنده بین دانش ورودی با سایر اطلاعات است. بازیابی اطلاعات در این حافظه از یک حالت دلخواه - بازنمایی گذشته از یک تصویر بینایی - آغاز می‌شود و نهایتاً فعالیت مغز به یک حالت دیگر همگرا می‌شود. این حالت، وضعیت پایدار است که قبلاً مغز آن را به خاطر سپرده است [۱۹]. در این مقاله کوششی اولیه در راستای مدل‌کردن این توانمندی با استفاده از شبکه‌های خودانجمنی استخراج‌کننده مولفه‌های

جدول (۱): ماتریس اختلاط حالت‌ها برای جاذب‌های شبکه خودانجمنی عمیق با روش پیش‌تعلیم لایه‌به‌لایه

برچسب حالت جاذب‌ها		
تعب	لبخند	خنثی
تعب	لبخند	خنثی
۲۰٪	۱۱/۴۳٪	۶۸/۵۷٪
۵۰/۳۹٪	۳۸/۷۶٪	۱۰/۸۵٪
۶۳/۲۴٪	۲۳/۹۳٪	۱۲/۸۲٪

جدول (۲): ماتریس اختلاط حالت‌ها برای جاذب‌های شبکه خودانجمنی عمیق با روش پیش‌تعلیم لایه‌به‌لایه با سرپرست

برچسب حالت جاذب‌ها		
تعب	لبخند	خنثی
تعب	لبخند	خنثی
۲۰٪	۳۷/۱۴٪	۴۲/۸۶٪
۶۷/۹۸٪	۸۲/۹۵٪	۱۰/۰۸٪
۹۰/۶٪	۸/۵۵٪	۰/۸۵٪

هر سطر از این جدول‌ها نشان می‌دهند چه درصدی از نمونه‌های با برچسب آن سطر، جذب جاذب‌هایی با برچسب مشخص شده برای ستون مربوطه می‌شوند؛ در نتیجه، عناصر روی قطر نشان‌دهنده درصد نمونه‌هایی هستند که جذب جاذب‌هایی با برچسب مشابه خود شده‌اند. با مقایسه عناصر روی قطرهای این ماتریس‌های اختلاط مشاهده می‌شود با روش پیش‌تعلیم لایه‌به‌لایه با سرپرست برای حالت‌های لبخند و تعجب میزان جذب به جاذب‌هایی با برچسب مشابه به طور معناداری افزایش می‌یابد. علت کاهش این میزان برای حالت خنثی، تصویر لبخند ضعیف فرد ۸م در دادگان است. همان‌طور که در شکل (۲۸) مشاهده می‌شود برای فرد ۸م شدت حالت لبخند در تصویر ضعیف است و تصویر به حالت خنثی نزدیک است؛ بنابراین، هنگامی که در پیش‌تعلیم لایه‌به‌لایه با سرپرست برای این تصویر، برچسب لبخند تعریف می‌شود، به دلیل شباهت حالت این تصویر به حالت خنثی، برخی از تصاویر خنثی جذب جاذب‌هایی با برچسب لبخند می‌شوند.



شکل (۲۸): تصویر حالت لبخند فرد ۸م از داده کوچک CK+

- [4] Seow M.-J. and Asari V.K., "Learning a Perceptual Manifold", in Proceedings of the ninth international conference on cognitive and neural systems, 2005.
- [5] Seow M.-J. and Asari V.K., "Recurrent Neural Network as a Linear Attractor for Pattern Association", IEEE Transactions on Neural Networks, Vol. 17, No. 1, pp. 246-250, 2006.
- [6] Nejadgholi I., Seyyedsalehi S.A., and Chartier S., "A Brain-Inspired Method of Facial Expression Generation Using Chaotic Feature Extracting Bidirectional Associative Memory", Neural Processing Letters, Vol. 46, pp. 943-960, 2017.
- [7] Trappenberg T.P., "Continuous Attractor Neural Networks", Recent developments in biologically inspired computing, pp. 398-425, 2003.
- [8] Saul L.K. and Jordan M.I., "Attractor Dynamics in Feedforward Neural Networks", Neural computation, Vol. 12, No. 6, pp. 1313-1335, 2000.
- [9] Diamantaras K.I. and Kung S.Y., Principal Component Neural Networks: Wiley New York, 1996.
- [10] Chartier S. and Boukadoum M., "A Bidirectional Heteroassociative Memory for Binary and Grey-level Patterns", IEEE Transactions on Neural Networks, Vol. 17, No. 2, pp. 385-396, 2006.
- [11] Amini N. and Seyyedsalehi S.A., "Manipulation of attractors in feed-forward autoassociative neural networks for robust learning", Electrical Engineering (ICEE), 2017 Iranian Conference on, pp. 29-33, 2017.
- [12] Seyyedsalehi S.Z. and Seyyedsalehi S.A., "Improvement of Autoassociative Neural Networks by Attractor Dynamics", ICBME2012.
- [13] Kulenovic M.R. and Merino O., Discrete Dynamical Systems and Difference Equations with Mathematica: Chapman and Hall/CRC, 2010.
- [14] Seyyedsalehi S.Z. and Seyyedsalehi S.A., "Simultaneous Learning of Nonlinear Manifolds Based on the Bottleneck Neural Network", Neural Processing Letters, Vol. 40, No. 2, pp. 191-209, 2014.
- [15] Lucey P., Cohn J. F., Kanade T., Saragih J., Ambadar Z., and Matthews I., "The Extended Cohn-Kanade Dataset (CK+): A Complete Dataset for Action Unit and Emotion-specified Expression", in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 94-101, 2010.
- [16] Mohammadian, A., Aghaeinia, H., & Towhidkhah, F. "Video-based Facial Expression Recognition by Removing the Style Variations", Image Processing, IET, Vol. 9, No. 7, pp. 596-603, 2015.
- [17] Seyyedsalehi S.Z. and Seyyedsalehi S.A., "A Fast and Efficient Pre-training Method Based on Layer-by-layer Maximum Discrimination for Deep Neural Networks", Neurocomputing, Vol. 168, pp. 669-680, 2015.
- [18] Calder A., Rhodes G., Johnson M., and Haxby J., Oxford Handbook of Face Perception, Oxford University Press, 2011.

اساسی صورت گرفت. با تحلیل‌های انجام‌شده نشان داده شد برای یک نورون با تابع غیرخطی تانژانت هیپربولیک با برقراری این اتصال بازگشتی، همواره در محل نمونه‌تعلیم داده شده و قرینه آن دو جاذب ایجاد می‌شود؛ اما در صورتی که تابع غیرخطی، سیگموئید باشد برای $x_1 < 0.211$ جاذب، تشکیل و برای مقادیر کمتر از آن نمونه‌تعلیم داده شده دافع می‌شود و در مرز بستر جذب‌های دو جاذب نوسانی قرار می‌گیرد. همچنین، مشاهده شد با افزودن بایاس در روابط برای مقادیر مختلف آن، رفتار نورون تغییر خواهد کرد و لزوماً قانون کلی برای رفتار آن وجود ندارد.

در ادامه، رفتار شبکه خودانجمنی تک‌لایه با اتصال بازگشتی تحلیل شده است و نشان داده شد برخلاف شبکه تک‌نورونی که با برقراری شرایط حتماً در محل نمونه‌تعلیم داده شده یک جاذب پایدار ایجاد می‌شود، برای شبکه تک‌لایه چندبعدی آرایش نمونه‌های تعلیم در نوع تعادلشان نقش دارد. همچنین، آزمایشات نشان دادند با افزایش تعداد لایه‌های شبکه خودانجمنی، قابلیت ذخیره جاذب‌های بیشتری فراهم می‌شود. همچنین، مشاهده شد جهت‌دهی به نحوه تشکیل جاذب‌ها با روش لایه‌به‌لایه باسرپرست در شبکه خودانجمنی به‌عنوان روشی برای کنترل مرز بستر جذب‌ها مؤثر واقع می‌شود. در مقایسه نتایج این روش و درک‌های انسانی از حالت‌های احساسی نیز مشاهده شد هردو در اختلاط بین حالت‌های مختلف نقاط مشترکی دارند؛ به‌گونه‌ای که برای هردو، اختلاط‌های تصاویر خنثی با لبخند نسبت به اختلاط با تعجب و لبخند با خنثی نسبت به اختلاط با تعجب بیشتر هستند.

مراجع

- [1] Hou Y., Yao L., and He P., "Robust Nonlinear Dimension Reduction: A Self-organizing Approach", in Fuzzy Systems and Knowledge Discovery, ed: Springer, pp. 67-72, 2005.
- [2] Huang D., Yi Z., and Pu X., "Manifold-based Learning and Synthesis", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 39, No. 3, pp. 592-606, 2009
- [3] Moghadam S.M. and Seyyedsalehi S.A., "Nonlinear analysis and synthesis of video images using deep dynamic bottleneck neural networks for face recognition", Neural Networks, Vol. 105, pp. 304-315, 2018.

and Applications, Vol. 19, No. 2, pp. 337-342, 2010.

[19] Seow M.-J., Asari V.K., and Livingston A., "Learning as a Nonlinear Line of Attraction in a Recurrent Neural Network", Neural Computing

¹ Hypersurface

² Cohn-Kanade

³ Confusion Matrix

Archive of SID