

تشخیص جنسیت افراد از روی چهره با استفاده از شبکه‌های عصبی پرسپترون چند لایه ای با الگوریتم پس انتشار خطا

علی یوسفی، کامبیز بدیع
(عضو هیئت علمی دانشگاه آزاد واحد همدان)
(گروه پژوهشی جامعه اطلاعاتی، مرکز تحقیقات مخابرات ایران)
(al_yo2002@yahoo.co.uk)
(k_badie@itrc.ac.ir)

چکیده

یکی از ابزارهای هوش مصنوعی شبکه‌های عصبی مصنوعی است که امروزه مورد توجه بسیاری از محققان و دانش پژوهان می‌باشد. بدلیل قابلیت‌های فراوان این ابزار در کاربردهای مختلف، مدل‌های گوناگونی برای آن طرح شده است که یکی از این مدل‌ها، شبکه‌های پرسپترون چند لایه پیشرو است که با الگوریتم پس انتشار خطا آموزش داده می‌شوند. در این مقاله هدف، استفاده از این شبکه‌ها جهت تشخیص جنسیت افراد از روی چهره می‌باشد. نتایج بدست آمده در آزمایشات انجام گرفته موید آن است که بین ۸۵٪ الی ۹۵٪ تشخیص چهره‌های ناآزموده با موفقیت انجام گرفته که نشان دهنده توانایی مناسب شبکه‌های پرسپترون چند لایه ای در حل مساله تشخیص جنسیت و مسائلی از این قبیل می‌باشد. در این روش از روش‌های پردازش چهره بعنوان پیش پردازش تصاویر استفاده شده است. همچنین با مقایسه این روش با سایر روش‌ها نظیر استفاده از حافظه خود القاء و آنالیز مولفه‌های اصلی چهره به کارایی این روش می‌توان پی برد. **کلید واژگان:** شبکه‌های عصبی پرسپترون چند لایه پیشرو، پردازش تصویر، پس انتشار خطا، آنالیز مولفه‌های اصلی، حافظه خود القاء.

۱- مقدمه

روش الگوریتمیک در پردازش اطلاعات که به محاسبه برنامه‌ریزی شده مشهور است، جهت حل مسائلی که رویه‌ای مناسب جهت حل آن استنتاج می‌شود، مناسب می‌باشد. در مسائلی که نتوان رویه‌ای مشخص ارائه داد و یا آنکه اطلاعات ما از دانش مسئله آنقدر کم باشد که راه حل دقیق مسئله ممکن نباشد، استفاده از روش الگوریتمیک مناسب نبوده و باید از روش‌های غیر الگوریتمیک بهره جست. یکی از این روش‌ها استفاده از شبکه‌های عصبی مصنوعی است که در آن به دانش قواعد حاکم بر مسئله نیاز نیست، و در حل مسائلی نظیر پردازش چهره و تشخیص آن که الگوریتم مشخصی برای آن وجود ندارد، می‌پردازد.

بدلیل عدم شناخت کافی، تشخیص مشخصه‌های چهره توسط ماشین به سادگی قابل انجام نیست. تشخیص جنسیت اشخاص از روی چهره با استفاده از ANN مشکل فوق را نداشته و استخراج ویژگی‌های مطلوب به عهده شبکه می‌باشد. در بخش‌های بعدی کارهای انجام شده در حوزه تشخیص چهره را بیان نموده و سپس آزمایشات انجام شده به روش شبکه‌های پرسپترون چند لایه با الگوریتم پس انتشار را روی تصاویر جمع آوری شده بیان نموده، در پایان به تشریح

۲- روش غیر الگوریتمیک (روش مبتنی بر استفاده از شبکه عصبی)

بر خلاف روش الگوریتمیک، که در آن غالباً از مشخصات هندسی چهره استفاده می شود، روش مبتنی بر شبکه های عصبی مستقیماً بر روی چهره و اطلاعات آن عمل می کند (شدت نور پیکسل های تصویر دو بعدی). لذا در این روش اطلاعات هندسی تصویر بطور ضمنی ذخیره می شود، بعلاوه آنکه بافت و جزئیات تصویر نیز حفظ می گردد. یکی از مشکلات این روش حساس بودن به تغییرات شرایط روشنایی محیط، جهت یابی سر و اندازه تصاویر مختلف افراد می باشد. لذا در این مقاله برای جلوگیری از بروز این مشکل یک مرحله پیش پردازش بر روی تصاویر (شامل نرمال سازی، تغییر اندازه، دوران و ...) ضروری است، این عمل می تواند به روشی نسبتاً مشخص مشابه الگوریتم های پیدا کردن خودکار موقعیت صورت در تصویر و مقیاس بندی انجام گیرد [1,2,3].

۳- دسته بندی روش های مبتنی بر شبکه عصبی جهت پردازش چهره

دو گروه عمده از شبکه های عصبی جهت شبیه سازی یادگیری تصاویر چهره مورد استفاده قرار گرفته است که گروه اول، شبکه های خودالقاء خطی می باشند که به عنوان آنالیزگر مولفه های اصلی^۱ شناخته شده اند. این شبکه ها برای کاربردهای مختلف از قبیل بازشناسی و دسته بندی تصاویر بر اساس جنسیت مورد استفاده قرار گرفته اند و در گروه دوم شبکه های خود القای سه لایه غیر خطی می باشند که با الگوریتم پس انتشار خطا^۲ آموزش دیده اند، و به عنوان شبکه های فشرده ساز شناخته شده اند. این شبکه ها طوری آموزش می بینند که ورودی خود را پس از عبور از یک کانال باریک در خروجی مجدداً بازسازی می نمایند. از لایه های میانی این شبکه بعنوان یک فشرده ساز استفاده می شود. خروجی لایه میانی به عنوان ورودیهای یک شبکه دیگر مورد استفاده قرار می گیرند که شبکه ثانویه عمل بازشناسی و دسته بندی را به عهده دارد [4,5,6].

۳-۱- شبکه های پس انتشار خطا و فشرده سازی تصویر

برخلاف شبکه های خطی خود القا که در آن واحدهای ورودی مستقیماً به واحدهای خروجی متصل هستند، شبکه های پس انتشار دارای واحدهای مخفی غیر خطی بین لایه ورودی و لایه خروجی می باشند. الگوریتم پس انتشار در ابتدا توسط Werbos پیشنهاد شد و سپس مستقلاً توسط محققان دیگری مانند Parker، Rumelhart و Hinton معرفی گردید. این الگوریتم به راحتی توسط معادلات ماتریسی قابل توصیف می باشد.

۳-۱-۱- توصیف الگوریتم

نشانه های مورد استفاده به قرار زیر می باشند:

K امین تصویر با I مشخصه توسط برداری $I \times 1$ به نام X_k نشان داده می شود. پاسخ واحدهای پردازشی مخفی به K امین تصویر به وسیله برداری $L \times 1$ به نام h_k نمایش داده می شود که L نمایش دهنده تعداد واحدهای پردازشی مخفی می باشد. پاسخ واحدهای خروجی به نام K امین تصویر توسط برداری $J \times 1$ به نام O_k نمایش داده می شود که J نمایش دهنده تعداد واحدهای خروجی می باشد. خروجی مطلوب جهت K امین تصویر برداری $J \times 1$ به d_k ، ضرایب اتصالات بین واحدهای ورودی و واحدهای مخفی بوسیله ماتریسی $L \times I$ به نام W نشان داده می شود. ضرایب اتصالات بین واحدهای مخفی و واحدهای خروجی بوسیله ماتریسی $J \times L$ به نام Z نمایش داده می شود. توابع مختلفی را می توان به عنوان تابع

فعالیت واحدهای مخفی و خروجی استفاده نمود که عمومی ترین آنها تابع Logistic می باشد که نوعی از تابع Sigmoid می باشد.

$$f(x) = \frac{1}{1 + e^{-x}} \quad \text{رابطه (۱)}$$

نحوه تنظیم وزن ها در شبکه براساس میزان خطای واحدهای خروجی (تفاضل بین خروجی شبکه و خروجی مطلوب $e_k = d_k - o_k$) و پس انتشار آن از لایه خروجی به سمت لایه ورودی انجام می گیرد. با نشانه های معرفی شده در فوق می توان فرمول تصحیح وزن ها را به شرح زیر نوشت:

$$h_k = f(Wx_k) \quad \text{رابطه (۲)}$$

$$O_k = f(Zh_k) \quad \text{رابطه (۳)}$$

$$e_k = (d_k - o_k) \quad \text{رابطه (۴)}$$

$$\delta_{\text{output},k} = f'(Zh_k) \cdot (e_k) = o_k \cdot (1 - o_k) \cdot (d_k - o_k) \quad \text{رابطه (۵)}$$

$$Z^{(t+1)} = Z^{(t)} + \eta \delta_{\text{output},k} h_k^T = Z^{(t)} + \Delta Z^{(t+1)} \quad \text{رابطه (۶)}$$

که در آن δ_{output} سیگنال خطای سلول های خروجی و f' مشتق تابع فعالیت می باشد. برای لایه های میانی نیز می توانیم از فرمول زیر استفاده کنیم:

$$\delta_{\text{hidden},k} = f'(Wx_k) \cdot (Z^T \delta_{\text{output},k}) \quad \text{رابطه (۷)}$$

$$= h_k \cdot (1 - h_k) \cdot (Z^T \delta_{\text{output},k})$$

$$W^{(t+1)} = W^{(t)} + \eta \delta_{\text{hidden},k} x_k^T \quad \text{رابطه (۸)}$$

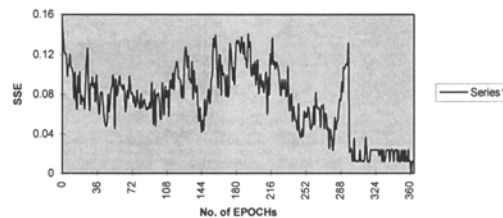
$$= W^{(t)} + \Delta W^{(t+1)}$$

۴- آزمایشات و نتایج شبیه سازی شده

در این مرحله به جمع آوری تصاویر و اعمال آن به شبکه های عصبی با استفاده از نرم افزار C و مشاهده نتایج مبادرت شده است.

۴-۱- آزمایش اول تشخیص جنسیت

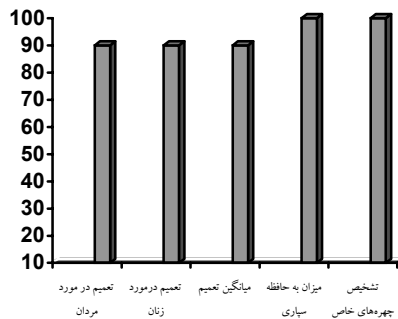
در مورد اول شبکه دارای ۴۰۹۶ ورودی (۶۴×۶۴)، ۲۰۰ واحد پردازشی (نرون) در لایه میانی اول، ۲۰ واحد پردازشی (نرون) در لایه میانی دوم و ۱ خروجی بود، این شبکه از ۳۶۵ مرتبه نمایش کل مجموعه آموزشی توانست ۱۰۰٪ مجموعه الگوی آموزشی را به حافظه بسپارد (یاد بگیرد)، نرخ یادگیری در این آزمون ۰/۵۵ و ضریب گشتاور ۰/۲۵ بود. نمودار خطای این شبکه در نمودار (۱) آورده شده است.



نمودار (۱)

در مرحله تست از تعداد ۲۰ عکس نمایش داده نشده به شبکه، جهت سنجش قدرت تعمیم شبکه استفاده شد. از این تعداد ۱۰ عکس متعلق به زنان و ۱۰ عکس متعلق به مردان می باشد. نتایج حاصله چنین بود که ۹۰٪ از کل تصاویر درست تشخیص داده شدند که از میزان باقیمانده ۵٪ از تصاویر آقایان و ۵٪ از تصاویر خانمها اشتباه تشخیص داده شد.

همچنین این شبکه در تشخیص ۱۰ تصویر مربوط به زنان و مردان دارای شرایط خاص (ریش و سبیل و عینک در مورد آقایان و آرایش چهره و استفاده از عینک در مورد خانمها) ۱۰۰٪ دسته‌بندی درستی ارائه نمود در شکل (۱) نمونه‌هایی از تصاویر خانمها و آقایان دارای ویژگی خاص نمایش داده شده است. در نمودار (۲) نتایج حاصله از آزمایش مشاهده می‌گردد.



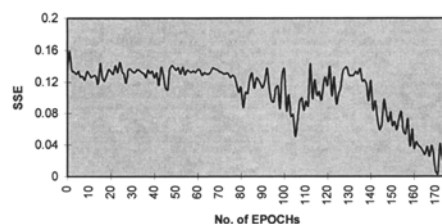
نمودار (۲)



شکل (۱)

۲-۴- آزمایش دوم تشخیص جنسیت

در مورد دوم که مشابه آزمایش قبلی است، تعداد ورودیها ۴۰۹۶ تعداد خروجی ۱، تعداد واحدهای لایه میانی اول ۱۲۸ و لایه میانی دوم ۱۶ می‌باشد، نرخ یادگیری و ضریب گشتاور در این آزمون به ترتیب ۰/۷ و ۰/۶ انتخاب شده است. شبکه پس از ۱۷۵ بار تکرار همگرا شده و توانسته است که تمامی نمونه‌های آموزشی را یاد گیرد. در مرحله تست نیز از میان ۲۰ تصویر نادیده، ۱۶ مورد را درست طبقه‌بندی نمود (۹۰٪) که از این میان، ۸۵٪ آقایان و ۹۵٪ خانمها مشمول دسته‌بندی صحیح واقع شده‌اند. در مورد نمونه‌های دارای ویژگیهای خاص سیستم قادر به تشخیص ۹۰٪ موارد شد. خطای سیستم در نمودار (۳) نمایش داده شده است.



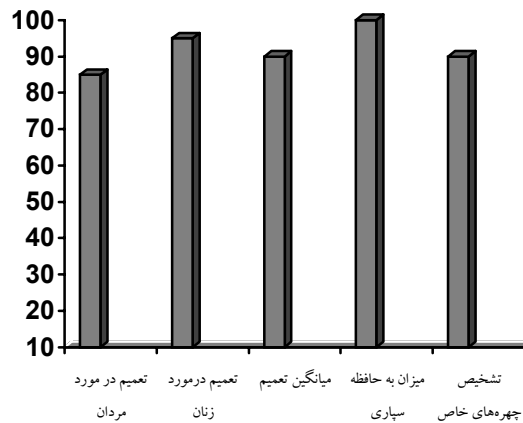
نمودار (۳)



شکل (۲)

دو نکته مهم در این آزمایش به چشم می‌خورد:

- ۱- در مواردی که شبکه قادر به تشخیص تصاویر نبود (در مرحله تست) اکثراً تصاویر خاص و مشترک بودند، گاهی چشم انسان نیز در تشخیص اینگونه تصاویر دچار اشتباه می‌شد.
- ۲- در مرحله تست این دو شبکه، از تصاویر کودکان نیز استفاده شده که در هر سیستم از ۵ تصویر مربوط به کودکان ۱ تصویر نادرست تشخیص داده می‌شد. در نمودار (۴) نتایج حاصل از آزمایش فوق ملاحظه می‌شود.



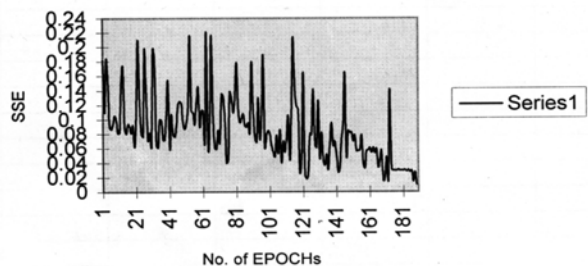
نمودار (۴)

در سری دوم آزمایش‌ها جهت مقایسه میزان تاثیر گذاری تعداد نمونه‌های آموزش در قدرت تعمیم شبکه به آزمایشاتی با مجموعه آزمایشی متشکل از ۳۰ تصویر (۱۵ تصویر مرد و ۱۵ تصویر زن) انجام شد که در نتیجه در مجموع کاهش تعداد تصاویر مجموعه آموزشی، کاهش قابلیت تعمیم شبکه را در پی داشت که در اینجا به دو مورد آن اشاره می‌گردد.

۳-۴- آزمایش سوم تشخیص جنسیت

در مورد اول شبکه‌های با ۴۰۹۶ ورودی، ۲۲۰ واحد میانی اول، ۲۰ واحد میانی دوم و ۱ خروجی می‌باشد که این شبکه پس از ۱۹۸ تعداد تکرار، همگرا شده و تمامی تصاویر ورودی را به حافظه سپرد. از بین ۲۰ تصویر مجموعه تست تنها قادر به دسته‌بندی ۸۲/۵٪ آنها به طور صحیح (۸۵٪ از تصاویر خانم‌ها و ۸۰٪ از تصاویر آقایان)، همچنین در خصوص تصاویر با ویژگی خاص شبکه تنها به ۹۰٪ تصاویر پاسخ صحیح داد. نرخ یادگیری در این آزمایش ۰/۸ و ضریب ممان ۰/۵ انتخاب شده است. در نمودار ۵ خطای سیستم در مرحله آموزش مشاهده می‌شود.

Sex classification (Picture Size = 64x64 , Training samples = 30)



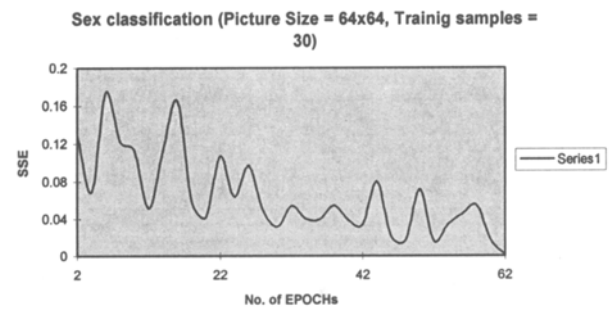
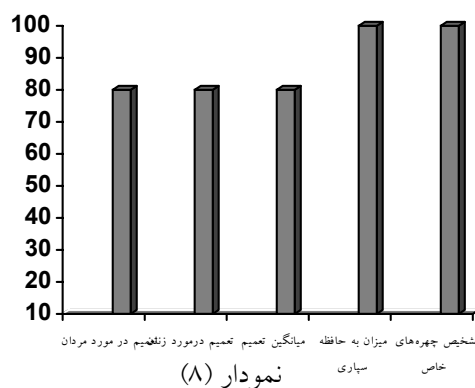
نمودار (۵)

۴-۴- آزمایش چهارم تشخیص جنسیت

در این آزمایش که مشابه به آزمایش پیشین می‌باشد، شبکه‌ای با ۴۰۹۶ ورودی، ۱۸۰ نرون میانی اول، ۲۰ نرون میانی دوم و ۱ خروجی شبیه سازی شد که پس از ۶۲ تکرار، همگرا شد. ولی تنها قادر به تشخیص صحیح ۸۲/۵٪ تصاویر موجود در فایل تست بود. (۸۰٪ در مورد آقایان و ۸۵٪ در مورد زنان)

این شبکه نیز در مورد تصاویر خاص تنها قادر به دسته‌بندی ۹۰٪ موارد بود. نرخ یادگیری در این آزمایش نیز ۰/۸ و ضریب همان ۰/۵ انتخاب شده است. در نمودار ۷ خطای مربوط به شبکه در آزمایش فوق و نمودار ۸ نتایج حاصل از آزمایش

مشاهده می شود.



نمودار (۷)

۵- نتیجه گیری

مدل عصبی پس انتشار خطا که بر خلاف مدل حافظه خود القا از واحدهای پردازش غیرخطی استفاده می کند، دامنه عملکرد وسیع تری نسبت به روش حافظه خود القا از خود نشان می دهد، نتایج حاصله در مقایسه با نتایج شبکه های حافظه خود القا نشان دهنده قدرت بالا در تعمیم و به حافظه سپاری در تشخیص جنسیت از روی چهره دارد، همچنین استفاده از الگوریتم های بهبود یافته پس انتشار خطا باعث خواهد شد که بر خلاف شبکه های خود القا مساله توقف در کمینه های محلی و نیز تسریع در فرایند آموزش از بوجود آمدن وضعیت برآزش بیش از حد و آموزش بیش از حد جلوگیری نموده که به نوبه خود باعث افزایش تعمیم در تشخیص جنسیت تصاویر می باشد. (جدول ۱).

۶- مراجع

- [1] Abdi, H., Valentin, D., Edelman B. G. and O'Toole, J.A. ‘ ‘More about the difference between men and women: Evidence form linear neural networks and the principal component approach’ ’, perception, vol.24,pp.539-562,1995.
- [2] Bouattour, H., Soulie, F.F and Viennet, E. “Neural nest for human face recognition” , In proceedings of International Conference on Neural networks, vol. 3,pp. 700-704, 1992.
- [3] Valentin, D., Abdi, H., O'Toole, A.J. and Cottrell. W.G. “Connectionist models of face processing: A survey” , pattern recognition, vol. 27,pp.1209-1230,1994.
- [4] Abdi, H. ‘ ‘A generalized approach for connectionist autoassociative memories: Interpretation, implications and illustration for face processing’ ’ , Artificial intelligence and cognitive sciences, Demongeot, J. ed., Manchester university press, Manchester, 1988.
- [5] Gutta, S. and Wechsler, H. “Face recognition using hybrid classifier systems” , In proceedings of International Conference on Neural networks, vol.2,pp. 1017-1022,1996.
- [6] Parodi, G. and Passaggio, F. “Size – adaptive neural network for image compression” , In Proceedings of International Conference on Neural networks , vol. 3 ,pp. 945-947,1994.

جدول (۱)- نتایجهای حاصله از آزمایشهای تعیین جنسیت

سایز تصاویر ورودی	تعداد نمونه‌های آموزشی در مرحله یادگیری	ساختار شبکه خروجی، میانی دوم، میانی اول، ورودی	نرخ یادگیری	ضریب ممان	تکرار تا زمان همگرا شدن	میزان به حافظه سپاری	قابلیت تشخیص تصاویر ویژه	میزان تعمیم		میانگین قدرت تعمیم
								مرد	زن	
۶۴×۶۴	۳۰	۴۰۹۶-۲۲۰-۲۰-۱	%۸۰	%۵۰	۱۸۹	%۱۰۰	%۹۰	%۸۲/۵		۸۵ %
								%۸۰	%۸۵	
		۴۰۹۶-۱۸۰-۲۰-۱	%۸۰	%۵۰	۶۲	%۱۰۰	%۱۰۰	%۸۲/۵		
								%۸۰	%۸۵	
	۴۰	۴۰۹۶-۱۲۸-۱۶-۱	%۷۰	%۶۰	۱۷۵	%۱۰۰	%۹۰	%۹۰		۹۰ %
								%۸۵	%۹۵	
		۴۰۹۶-۲۰۰-۲۰-۱	%۵۵	%۲۵	۳۶۵	%۱۰۰	%۱۰۰	%۹۰		
								%۹۰	%۹۰	
۸۶/۲ %										