



دانشگاه تربیت معلم سبزوار

دانشکده فنی و مهندسی

گزارش پروژه پایانی کارشناسی برق-الکترونیک

# بازشناسی ارقام دستنویس فارسی

مجری: سید مصطفی عبدالله پور

استاد راهنما: دکتر جواد حدادنیا

پاییز ۸۵

# بسم الله الرحمن الرحيم

## چکیده

در این پروژه مبانی نظری و جنبه‌های کاربردی مبحث بازشناسی نوری حروف<sup>۱</sup> (اُسی‌آر)، مورد بررسی قرار می‌گیرند و زیربخش‌ها و بلوک‌های پردازشی آن معرفی می‌گردند. همچنین خصایص و پیچیدگی‌های مختص نگارش زبان فارسی که یک نرم‌افزار «اُسی‌آر» باید آن‌ها را در عملیات پردازشی خود لحاظ نماید، بیان خواهند شد و تحقیقات داخلی انجام‌شده در زمینه «اُسی‌آر» مورد اشاره قرار خواهد گرفت. سپس تکنیک آنالیز اجزای اصلی<sup>۲</sup>، قابلیت‌های آن و شبکه عصبی پس انتشار<sup>۳</sup> خطا مورد بحث قرار می‌گیرند. در پایان یک نرم‌افزار بر اساس روش تکنیک آنالیز اجزای اصلی و شبکه عصبی جهت بازشناسی<sup>۴</sup> ارقام دست نویس فارسی پیشنهاد خواهد شد.

---

1 Optical character recognition

2 Principal component analysis

3 Back propagation neural network

4 Recognition

إِلَهِی وَالدِّیَّ

نَقَرِیْ بِه یَدِی و مَافِی

۱	فهرست مطالب
۴	فصل اول مقدمه
۴	۱-۱ شناسایی الگو
۶	۲-۱ کاربردهای بازشناسی الگو
۶	۳-۱ طرح پژوهش
۸	فصل دوم مروری بر سیستمهای OCR
۸	۱-۲ بخشهای مختلف سیستمهای ocr
۸	۱-۱-۲ بازشناسی نوری حروف
۹	۲-۱-۲ تاریخچه سیستمهای «اُسی آر»
۱۱	۳-۱-۲ تحقیقات انجام شده در داخل کشور در زمینه تولید «اُسی آر» فارسی
۱۲	۴-۱-۲ برخی ویژگیهای متون چاپی فارسی از دیدگاه پردازش رایانه‌ای
۱۴	۵-۱-۲ انواع سیستمهای «اُسی آر» از لحاظ نوع الگوی ورودی
۱۵	۶-۱-۲ معرفی بخشهای مختلف یک سیستم «اُسی آر»
۱۵	۱-۶-۱-۲ پیش پردازش
۲۰	۲-۶-۱-۲ قطعه‌بندی
۲۳	۳-۶-۱-۲ بازنمایی (استخراج ویژگی‌ها)
۲۵	۴-۶-۱-۲ طبقه‌بندی و بازشناسی (با یک یا چند طبقه‌بندی کننده)
۲۵	۵-۶-۱-۲ به کارگیری اطلاعات جانبی (پس پردازش)
۲۷	۷-۱-۲ پیشنهادهایی در جهت تسریع دستیابی به یک سیستم اُسی آر فارسی

۲۸	۲-۲ روشهای مختلف در حوزه بازشناسی اسناد
۲۸	۱-۲-۲ تبدیل سراسری ( بسط سری )
۲۸	۱-۱-۲-۲ تبدیلات فوریه
۳۰	۲-۱-۲-۲ موجکها
۳۳	۳-۱-۲-۲ تبدیل گابور
۳۶	۴-۱-۲-۲ ممانها (گشتاورها)
۳۹	۵-۱-۲-۲ بسط کارهونن لوئو
۴۰	۲-۲-۲ ویژگیهای آماری
۴۱	۳-۲-۲ ویژگیهای هندسی و توپولوژیکی
۴۵	۳-۲ تعدادی از روشهای شناسایی کاراکترهای فارسی و عربی در پایان نامه های داخل و خارج
۵۸	فصل سوم تکنیک آنالیز اجزای اصلی
۵۹	۱-۳ روشهای کاهش ابعاد
۶۰	۱-۱-۳ روشهای مبتنی بر استخراج ویژگی
۶۱	۲-۳ تکنیک آنالیز اجزا اصلی PCA
۶۲	۱-۲-۳ مفاهیم مقدماتی مورد نیاز در PCA
۶۵	۲-۲-۳ الگوریتم PCA
۷۱	فصل چهارم شبکه عصبی
۷۱	۱-۴ شبکه عصبی چیست؟
۷۴	۲-۴ چرا از شبکه های عصبی استفاده می کنیم؟
۷۵	۳-۴ شبکه عصبی MLP
۷۶	۱-۳-۴ قاعده فراگیری MLP

۹۲	فصل پنجم پیاده سازی یک نرم افزار تشخیص خودکار ارقام فارسی
۹۲	۱-۵ بلوک دیاگرام سیستم
۹۳	۲-۵ مرحله آموزش
۹۳	۳-۵ مرحله آزمایش
۹۳	۴-۵ نتایج و جداول
۹۸	فصل ششم مراجع
۹۸	مراجع لاتین
۱۰۲	مراجع فارسی

## فصل اول مقدمه

پیدایش علوم و فنون جدید، جوامع بشری را با شکلهای مختلفی از اطلاعات روبرو نموده است. سطح توسعه یک جامعه را می توان با مقدار اطلاعات و دانش تولید شده در آن ارزیابی کرد. تولید فزاینده اطلاعات به شکلهای مختلف صورت می گیرد و با درجات متفاوتی از پیچیدگی همراه میباشد. در نتیجه نیاز به سیستمهای پردازش اطلاعات بصورت روزافزون افزایش می یابد. یکی از مسائل مهم در طراحی سیستمهای مدرن اطلاعاتی، بازشناسی خودکار الگوها است [۸۲].

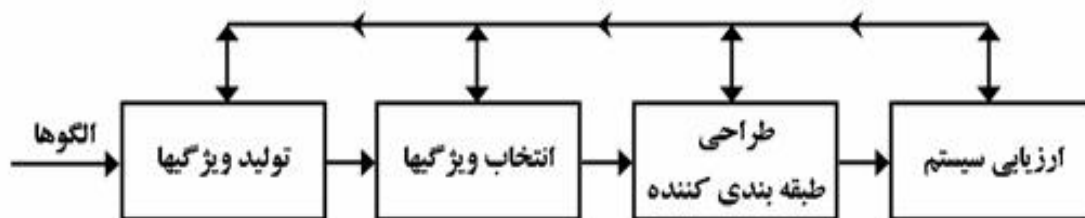
### ۱-۱ شناسایی الگو<sup>۱</sup>

شناسایی الگو، شاخه‌ای از هوش مصنوعی<sup>۲</sup> است که با طبقه‌بندی و توصیف مشاهدات سروکار دارد. شناسایی الگو به ما کمک می‌کند داده‌ها (الگوها) را با تکیه بر دانش قبلی یا اطلاعات آماری استخراج‌شده از الگوها، طبقه‌بندی نماییم. الگوهایی که می‌بایست طبقه‌بندی شوند، معمولاً گروهی از سنجش‌ها یا مشاهدات هستند که مجموعه نقاطی را در یک فضای چند بعدی مناسب تعریف می‌نمایند. یک سیستم شناسایی الگوی کامل متشکل است از یک حسگر<sup>۳</sup>، که مشاهداتی را که می‌بایست توصیف یا طبقه‌بندی شوند جمع‌آوری می‌نماید، یک سازوکار برای استخراج ویژگی‌ها<sup>۴</sup> که اطلاعات عددی یا نمادین را از مشاهدات، محاسبه می‌کند، (این اطلاعات عددی را با یک بردار بنام بردار ویژگی‌ها نمایش می‌دهند)؛ و یک نظام طبقه‌بندی یا توصیف که وظیفه اصلی طبقه‌بندی یا توصیف الگوها را با تکیه بر ویژگی‌های استخراج شده عهده‌دار است.

---

1 Pattern Recognition  
2 Artificial intelligence  
3 Sensor  
4 Feature extraction

شکل ۱-۱ نمودار بلوکی یک سیستم شناسایی الگو را نشان می‌دهد [۵۳]. همانطوری که از پیکان‌های برگشتی مشخص است، این بلوک‌ها لزوماً مستقل نیستند و بسته به نتایج حاصله گاهی لازم است که بلوک‌های اولیه مجدداً طراحی گردند تا راندمان کلی سیستم بهبود یابد.



شکل (۱-۱) نمودار بلوکی یک سیستم شناسایی الگو

نظام کلاسه بندی یا توصیف معمولاً مبتنی بر وجود یک مجموعه از الگوهایی است که قبلاً کلاسه بندی یا توصیف شده اند. این مجموعه الگوها را مجموعه آموزشی<sup>۱</sup> و قانون یادگیری منتج شده را قانون یادگیری باسرپرستی (با نظارت)<sup>۲</sup> می نامند همچنین یادگیری میتواند بصورت بدون نظارت<sup>۳</sup> باشد و این در حالی است که الگوهایی از قبل به سیستم داده نشده اند و در مقابل، سیستم خود براساس قواعد آماری الگوها، کلاسها را پایه گذاری میکند.

---

1 Training set  
2 Supervised learning  
3 Unsupervised



## ۱-۲ کاربردهای بازشناسی الگو

بازشناسی الگو در بسیاری از زمینه ها نقش کاربردی دارد [۸۲]. بازشناسی حروف، بازشناسی نویسنده، تصدیق امضاء، طبقه بندی اثر انگشت و بازشناسی گفتار نمونه هایی از این کاربردها هستند. شناسایی الگو برای تحلیل داده های پزشکی نیز بکار گرفته شده است. برای مثال تفسیر الکتروکاردیوگرام، تحلیل تصاویر اشعه X و طبقه بندی کروموزمها را میتوان نام برد. نمونه های دیگری از این کاربردها شامل طبقه بندی مناطق زراعی، مطالعه آلودگی آبها، آشکار کردن منابع زیرزمینی و پیش بینی آب و هواست. در این نوع کاربردها، تصاویر ارسال شده از ماهواره و تصاویر هوایی به کمک روشهای بازشناسی الگو تفسیر می شوند. بازرسی تصویری و بازشناسی قطعات ماشینی، از کاربردهای صنعتی شناسایی الگو هستند. تحلیل بافت، آشکارسازی هدف در سیگنالهای برگشتی رادار یا سونار<sup>۱</sup>، طبقه بندی امواج زلزله و تشخیص ذرات شیمیائی کاربردهای دیگری از بازشناسی الگو می باشند.

## ۱-۳ طرح پژوهش

در چند دهه گذشته مسأله بازشناسی الگوهای نوشتاری شامل حروف، ارقام و سایر نمادهای متداول در اسناد مکتوب شده به زبانهای مختلف، توسط گروههای مختلفی از محققین مورد مطالعه و بررسی قرار گرفته است. نتیجه این تحقیقات منجر به پیدایش مجموعه ای از روشهای سریع و تا حد زیادی مطمئن موسوم به OCR بمنظور وارد نمودن اطلاعات موجود در اسناد، مدارک، کتابها و سایر مکتوبات چاپی یا تایپ شده و حتی دستنویس<sup>۲</sup> به داخل کامپیوتر شده است [۶۸].

مسئله بازشناسی حروف الفبای فارسی سابقه ای نه چندان طولانی به همراه دارد. نخستین گزارشهای رسمی منتشر شده از تلاشهای انجام گرفته در این راه، مربوط به سالیان نخست دهه ۱۹۸۰ میلادی است [۶۸].

---

1 Sonar  
2 Handwritten

به رغم فراگیری نسبی کاربرد الفبای فارسی در میان ملل مختلف قاره آسیا، بررسیهای انجام شده در خصوص یافتن روشهایی برای بازشناسی حروف این الفبا بسیار محدود بوده است. بواسطه وجود تفاوت‌های اساسی بین نحوه نگارش کلمات فارسی و کلمات لاتین نظیر چسبیده بودن حروف سازنده یک کلمه به یکدیگر و تغییر شکل حروف بر اساس موقعیت نسبی قرارگیری آن در یک کلمه فارسی، امکان اعمال مستقیم روشهای متداول در بازشناسی حروف انگلیسی بمنظور شناسایی حروف تشکیل دهنده کلمات فارسی وجود ندارد.

اکثر کارهای انجام شده در زمینه «اُسی آر» در رابطه با متون لاتین، چینی و ژاپنی بوده است [۵۱]. نرم افزارهای تجاری «اُسی آر» لاتین در سال‌های اخیر پیشرفت کیفی قابل ملاحظه‌ای داشته‌اند. اما «اُسی آر» فارسی با وجود حجم نسبتاً وسیع تحقیقات دانشگاهی و نیاز شدید بازار تجاری به آن، هنوز هم از جایگاه مورد نظر فاصله بسیاری دارد و تاکنون هیچ سیستم «اُسی آر» کارآمدی که از نظر دقت و کیفیت محیط نرم‌افزاری، قابل مقایسه با سیستم‌های «اُسی آر» لاتین باشد، عرضه نگردیده است. در نتیجه ضرورت انجام تحقیقات بیشتر در زمینه متون فارسی و عربی کاملاً احساس می‌شود.

هدف از انجام این پروژه آزمایش توانایی تکنیک آنالیز اجزای اصلی جهت استخراج ویژگیهای مربوط به ارقام فارسی و کاربرد آن جهت شناسایی است که این امر در فصل آخر محقق شده است. امید است انجام این پروژه افق‌هایی نو را در مبحث آنالیز اسناد<sup>۱</sup> بگشاید.

---

1 Document Analysis (DA) or Document Image Analysis (DIA)

## فصل دوم مروری بر سیستمهای OCR

### ۱-۲ بخشهای مختلف سیستمهای OCR

در ادامه ماهیت سیستمهای OCR، تاریخچه و بخشهای مختلف آن مورد بررسی قرار می گیرد.

#### ۱-۱-۲ بازشناسی نوری حروف<sup>۱</sup>

اصطلاح «اُسی آر» به تکنیک‌هایی اطلاق می‌شود که در تصاویر اسکن یا فکس شده، نواحی متنی را تشخیص می‌دهند و سپس این نواحی (تصویری) را به متن قابل ویرایش تبدیل می‌نمایند. با دستگاهی به نام اسکنر می‌توان تصویر یک صفحه کاغذ را به صورت یک فایل گرافیکی (تصویری)، به رایانه ارسال و در آن ذخیره نمود. بدین ترتیب کاربر می‌تواند با یک نرم‌افزار مناسب نمایش‌دهنده تصاویر، تصویر صفحه اسکن‌شده را بر روی نمایشگر رایانه خود ملاحظه نماید یا آن را چاپ کند؛ اما قادر نخواهد بود که متن موجود در تصویر سند را ویرایش کند یا آن را مورد جستجو قرار دهد. یک نرم‌افزار «اُسی آر»، تصویر اسکن‌شده را می‌خواند، محتویات آن (شامل متن، خطوط، تصاویر، جداول، ...) را شناسایی می‌نماید، و سپس آن را به یک قالب قابل ویرایش (در واژه‌پردازها) تبدیل می‌کند. امروزه بیشتر دستگاه‌های اسکنر به نرم‌افزارهای «اُسی آر» مجهز گردیده‌اند و قادرند متن موجود در یک سند اسکن‌شده را تشخیص دهند و آن را با همان نحوه قالب‌بندی، ستون‌بندی، جدول‌بندی و نوع فونت مطابق با سند کاغذی اصلی، در قالب یک فایل متنی با قالب‌بندی مناسب ذخیره نمایند.

استفاده از سیستم‌های «اُسی آر» دو مزیت عمده دارد:

**الف.** افزایش چشمگیر سرعت دسترسی به اطلاعات؛

زیرا در متن بر خلاف تصویر، امکان جستجو و ویرایش وجود دارد.

**ب.** کاهش فضای ذخیره‌سازی؛

زیرا حجم فایل متنی استخراج‌شده از یک تصویر، معمولاً بسیار کمتر از حجم خود فایل تصویری است. چنین قابلیت‌هایی امکان استفاده گسترده از رایانه را در پردازش سریع حجم وسیعی از داده‌های مکتوب شرکت‌ها و مؤسسات مختلف (نظیر بانک‌ها، شرکت‌های بیمه، مؤسسات خدمات عمومی، اداره پست، و دیگر نهادهایی که سالانه با میلیون‌ها مورد پرداخت، دریافت و حسابرسی امور مشتریان خود مواجه‌اند فراهم می‌آورد.

## ۲-۱-۲ تاریخچه سیستم‌های «اُسی آر»

از جنبه تاریخی، سیستم‌های «اُسی آر» تا کنون سه مرحله تکاملی را پشت سر گذاشته‌اند [۶]:

**الف. مرحله تکوین (از ۱۹۰۰ تا ۱۹۸۰):** رد پای اولیه اقدامات صورت‌گرفته در زمینه بازشناسی حروف را در سال‌های اول دهه ۱۹۰۰ می‌توان یافت و آن زمانی است که «تیورینگ» دانشمند روسی بر آن بود که به افراد مبتلا به نارسایی‌های بینایی کمک نماید. اولین اختراعات ثبت‌شده در این زمینه مربوط به سال‌های ۱۹۲۹ و ۱۹ میلادی هستند [۴]. این سیستم‌ها حروف چاپی را با روش تطابق قالبی<sup>۱</sup> شناسایی می‌کردند؛ به این صورت که ماسک‌های مکانیکی مختلفی از مقابل تصویر حرف عبور می‌کردند (مکانیکی) و نور از یک سو به آن تابانده می‌شد و از سوی دیگر توسط یک آشکارساز نوری دریافت می‌گردید (اپتیک). وقتی یک انطباق کامل صورت می‌گرفت، نور به آشکارساز می‌رسید و حرف ورودی بازشناسی می‌شد. این اختراع به دلیل فناوری اپتومکانیکی مورد استفاده در آن، کاربردی نبود. تصور دسترسی به دستگاهی برای بازشناسی حروف تا دهه ۱۹۴۰ میلادی و ظهور رایانه‌های دیجیتال، به صورت یک رؤیا باقی ماند.

---

1 Template matching

اقدامات اولیه در زمینه بازشناسی حروف، بر متون چاپی یا مجموعه کوچکی از حروف و نمادهای دستنوشته که براحتی قابل تشخیص بودند، متمرکز گردیده بود. سیستم‌های بازشناسی حروف چاپی که در این مقطع زمانی عرضه شدند، عمدتاً از روش تطابق قالبی استفاده می‌نمودند که در آن، تصویر ورودی با مجموعه بزرگی از تصاویر حروف، مورد مقایسه قرار می‌گرفت. در مورد متون دستنوشته نیز الگوریتم‌های پردازش تصویر که ویژگی‌های سطح پایین<sup>۱</sup> (ویژگی‌هایی که مستقیماً و بدون اعمال هیچ تبدیلی، از تصاویر استخراج می‌شوند) را از تصاویر استخراج می‌کنند، در مورد تصاویر دوسطحی اعمال می‌شدند تا بردارهای ویژگی استخراج گردند. سپس این بردارهای ویژگی به طبقه‌بندی‌کننده‌های آماری سپرده می‌شدند.

در این دوره، تحقیقات موفق اما مقید (منظور از مقید، مفروض دانستن شرایط و پیش‌فرض‌های خاص برای کاراکترهای ورودی است)، بیشتر بر روی حروف و اعداد لاتین انجام گرفت. با این حال مطالعات چندی نیز بر روی حروف ژاپنی، چینی، عبری، هندی، سیریلیکی، یونانی و عربی در هر دو زمینه حروف چاپی و دستنوشته آغاز گردید. با ظهور صفحات رقومی‌کننده<sup>۲</sup> در دهه ۱۹۵۰ که قادر به تشخیص مختصات حرکتی نوک یک قلم مخصوص بودند، سیستم‌های «آسی‌آر» تجاری نیز امکان عرضه یافتند. این نوآوری سبب شد که محققان بتوانند در زمینه بازشناسی برخط حروف دستنوشته، فعالیت خود را آغاز نمایند. [۵۱]منبع مناسب درباره اقدامات صورت‌گرفته بر روی بازشناسی برخط حروف تا سال ۱۹۸۰ می‌باشد.

**ب. مرحله توسعه (از ۱۹۸۰ تا ۱۹۹۰):** مطالعات صورت گرفته تا قبل از سال ۱۹۸۰ بدلیل فقدان سخت‌افزارهای رایانه‌ای قدرتمند و دستگاه‌های اخذ داده‌ها با مشکل همراه بودند. در این دهه بواسطه رشد انفجارگونه فناوری اطلاعات، وضعیت بسیار مناسبی برای تحقیقات مختلف از جمله بازشناسی حروف فراهم گردید. روش‌های ساختاری به همراه روش‌های آماری در بسیاری از سیستم‌ها استفاده شدند. تحقیقات در زمینه «آسی‌آر» اساساً به توسعه روش‌های بازشناسی معطوف گردید، بی آنکه مسئله استفاده از اطلاعات معناشناختی به منظور افزایش دقت بازشناسی مورد توجه قرار گیرد. این امر سبب گردید که دقت بازشناسی (نرخ بازشناسی) از یک حد خاص فراتر نرود، که در بسیاری از کاربردهای «آسی‌آر»، قابل قبول

---

1 Low level feature

2 Digitizers

نبود. مروری بر تحقیقات و پیشرفت‌های حاصل‌شده در مورد «اُسی‌آر» در این برهه زمانی را می‌توان در [۴]، [۵۱] بترتیب برای بازشناسی برخط و برون خط، جستجو نمود.

**ج. مرحله بهبود (از ۱۹۹۰ به بعد):** در این مقطع زمانی بود که با تکوین ابزارها و تکنیک‌های پردازشی جدید، پیشرفت واقعی در سیستم‌های «اُسی‌آر» محقق گردید. در اوایل دهه ۹۰، روش‌های پردازش تصویر و بازشناسی الگو با تکنیک‌های کارآمد هوش مصنوعی ادغام گشتند. محققان، الگوریتم‌های پیچیده‌ای را در بازشناسی حروف ابداع نمودند که قادر بودند داده‌های ورودی با تفکیک‌پذیری<sup>۱</sup> بالا را دریافت کنند و در مرحله پیاده‌سازی، محاسبات بسیار زیادی را بر روی داده‌ها انجام دهند. امروزه علاوه بر وجود رایانه‌های قدرتمندتر و تجهیزات الکترونیکی دقیق‌تر مانند اسکنرها، دوربین‌ها و صفحات رقومی‌کننده، استفاده از تکنیک‌های پردازشی مدرن و توانمند همچون شبکه‌های عصبی، مدل‌های مارکوف پنهان<sup>۲</sup>، منطق‌های مجموعه فازی و مدل‌های پردازش زبان طبیعی<sup>۳</sup> امکان‌پذیر گشته است.

سیستم‌های جدید «اُسی‌آر» برون خط متون چاپی و برخط متون دستنوشته با واژگان محدود و وابسته به نویسنده، در کاربردهای محدود به نحو کاملاً رضایت‌بخشی عمل می‌کنند [۶]. اما به منظور دستیابی به هدف نهایی در شبیه‌سازی ماشینی نگارش انسانی و متون چاپی، هنوز راه درازی در پیش است.

## ۲-۱-۳ تحقیقات انجام‌شده در داخل کشور در زمینه تولید «اُسی‌آر» فارسی

با توجه به اهمیت طیف وسیع کاربردها، در سال‌های اخیر تحقیقات قابل ملاحظه‌ای در کشور در زمینه «اُسی‌آر» توسط دانشگاه‌ها، برخی نهادهای دولتی، و شرکت‌های خصوصی صورت گرفته است که متأسفانه از آمار دقیق آن‌ها اطلاعی در دست نیست. اما قدر مسلم این که برای «اُسی‌آر» متون چاپی تاکنون هیچ نرم‌افزار کارآمد «اُسی‌آر» تجاری که محصول تحقیقات داخل کشور باشد، عرضه نگردیده است.

در ادامه به برخی از تلاش‌های صورت گرفته در این زمینه اشاره می‌شود:

---

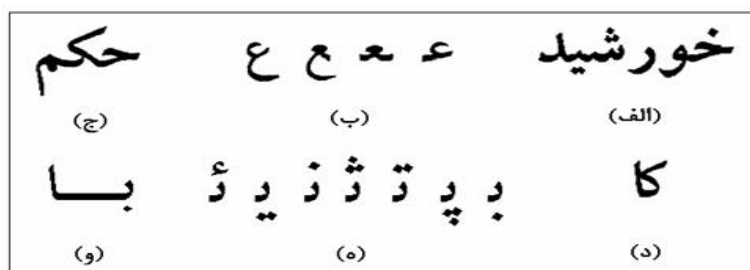
1 Resolution  
2 HMM  
3 NLP

- در حوزه تحقیقات دانشگاهی، تعداد نسبتاً زیادی پایان‌نامه (به‌خصوص در مقاطع کارشناسی ارشد و دکتری) و مقاله در این زمینه منتشر شده‌اند که نقطه تمرکز بیشتر آن‌ها، ارائه روش‌هایی به منظور قطعه‌بندی درونی، بازنمایی و بازشناسی حروف بوده است و سایر بخش‌ها شامل پیش‌پردازش، قطعه‌بندی بیرونی و پس‌پردازش کمتر مورد توجه قرار گرفته‌اند. بخش‌های مختلف پردازشی یک سیستم «اُسی‌آر» شامل پیش‌پردازش، قطعه‌بندی، بازنمایی، بازشناسی و پس‌پردازش در بخش‌های بعدی مورد بررسی قرار خواهند گرفت.

- طرح ملی «بازشناسی متون چاپی و حجم محدودی از کلمات دست‌نویس» به سرپرستی دکتر «احسان‌الله کبیر» آغاز گردید که در آن تعدادی از دانشجویان و اساتید دانشگاه‌های تربیت مدرس و صنعتی امیرکبیر در قالب پایان‌نامه‌های کارشناسی ارشد و دکتری، به انجام تحقیق پرداختند. دکتر «کبیر» پروژه‌هایی نیز با عناوین «بازشناسی متون چاپی فارسی» و «بازشناسی حروف و ارقام فارسی دست‌نویس» برای «سازمان پژوهش‌های علمی و صنعتی ایران» انجام داده است.

## ۲-۱-۴ برخی ویژگی‌های متون چاپی فارسی از دیدگاه پردازش رایانه‌ای

نگارش فارسی، ویژگی‌های منحصر به فردی دارد که آن را کاملاً از نگارش لاتین متمایز می‌سازد [۸۲]. به منظور فعالیت در حوزه «اُسی‌آر» فارسی، آگاهی از قوانین نگارشی و نحوه چاپ حروف در این زبان، امری ضروری است. در اینجا به ویژگی‌های کلی نگارش فارسی اشاره می‌شود:



شکل ۲-۱-۴. برخی از ویژگی‌های نگارش زبان فارسی: الف) کلمه خورشید از سه زیرکلمه تشکیل شده؛ ب) چهار شکل

مختلف حرف «ع» با توجه به موقعیت آن در کلمه، ج) همپوشانی دو حرف «ح» و «ک» در کلمه «حکم»؛ د) اتصال

حروف «ک» و «ا» در دو محل؛ ه) حروف متفاوت با بدنه مشابه؛ و) کشیدگی حرف «ب» در کلمه «با».

الف - متون فارسی برخلاف متون لاتین از راست به چپ نوشته می‌شوند.

ب - در کلمات فارسی برخی از حروف از یک یا دو طرف به حروف مجاور خود اتصال دارند و برخی نیز به صورت مجزا نوشته می‌شوند. در نتیجه هر کلمه ممکن است شامل یک یا چند بخش متصل باشد که «زیرکلمه» نامیده می‌شوند (شکل - الف).

چسبیده یا سرهم بودن حروف در نگارش فارسی، بازشناسی متون فارسی را برای سیستم‌های «اُسی‌آر»، نسبت به متون لاتین بسیار مشکل‌تر می‌سازد.

ج - حروف فارسی ممکن است چهار موقعیت مجزا و در نتیجه چهار شکل متفاوت نگارش داشته باشند: حروف ابتدایی، میانی، انتهایی و مجزا (شکل - ب).

چ - حروف واقع در یک کلمه ممکن است همپوشانی داشته باشند، بدین معنا که نتوان با رسم خطوط عمودی، حروف را به طور کامل از یکدیگر مجزا نمود (شکل - ج).

ح - در برخی از فونت‌ها بعضی از حروف، از یک سمت در دو محل به یکدیگر اتصال دارند (شکل - د).

خ - برخی از حروف بین یک تا سه نقطه دارند که ممکن است در بالا یا پایین بدنه حرف واقع باشند (شکل - ه).

بعضی از حروف بدنه مشابه دارند و تفاوت آن‌ها تنها در تعداد و محل قرارگیری نقاط (شکل ه) یا در وجود یک سرکش است (مانند «ک» و «گ») که در مقایسه با بدنه حروف، اندازه بسیار کوچکی دارند. این موضوع نیز یکی از مواردی است که بر پیچیدگی سیستم‌های «اُسی‌آر» فارسی می‌افزاید.

د - حروف فارسی ممکن است در بالا یا پایین بدنه دارای اعراب باشند. سه اعراب َ - ُ - ِ در زبان فارسی، اعراب‌های اصلی‌اند و اعراب ً در برخی کلمات عربی رایج در زبان فارسی دیده می‌شود (نظیر کلمات «عمداً» و «احتمالاً»). کلمات عربی دارای اعراب ُ و ِ در زبان فارسی عمومیت نیافته‌اند. هر چند کاربرد اعراب در زبان فارسی نسبت به زبان عربی بسیار محدودتر است، اما اگر کلمه‌ای نامتداول باشد یا به دلیل تشابه نگارشی آن با کلمه دیگر، تأکید بر تلفظ صحیح آن باشد، از نشانه‌های اعراب استفاده می‌شود.

ذ - در بالای بدنه یک حرف ممکن است علامت تشدید وجود داشته باشد.



ر - برخی از حروف دارای علامت همزه هستند («ئا»، «أ»، «ؤ»، «به»).

ز - حروفی که از طرف چپ قابلیت اتصال به حرف مجاور خود را دارند، ممکن است به صورت کشیده نوشته شوند (شکل - و).

بیشتر حروف فارسی (مخصوصاً حروف چسبیده) دنداندار هستند. در مواردی که کیفیت سند اصلی یا دستگاه اسکنر پایین باشد، ارتفاع دندانها نسبت به خط زمینه کوتاه می‌شود و این امر، شناسایی صحیح این حروف در مرحله قطعه‌بندی یا بازشناسی را با مشکل مواجه می‌سازد. بواسطه وجود تفاوت‌های اساسی بین نحوه نگارش فارسی و لاتین (نظیر چسبیده‌بودن حروف کلمه به یکدیگر، تغییر شکل حروف براساس موقعیت نسبی آن در کلمه فارسی، و ...)، امکان اعمال مستقیم روش‌های بازشناسی متون لاتین به منظور شناسایی متون فارسی وجود ندارد.

## ۲-۱-۵ انواع سیستم‌های «آسی آر» از لحاظ نوع الگوی ورودی

سیستم‌های «آسی آر» را می‌توان از لحاظ نوع الگوی ورودی به دو گروه اصلی تقسیم کرد:

الف. سیستم‌های برخط.

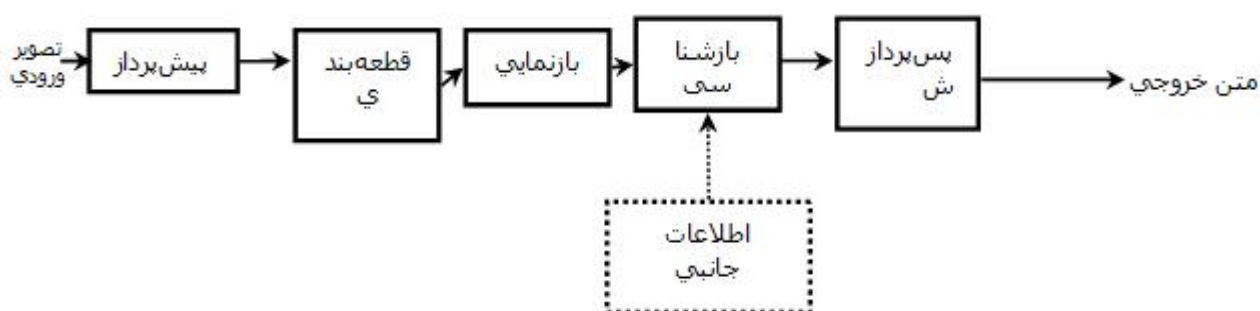
ب. سیستم‌های برون خط.

در بازشناسی برخط، حروف در همان زمان نگارش توسط سیستم تشخیص داده می‌شوند و دستگاه ورودی این سیستم‌ها یک قلم نوری است. در این روش علاوه بر اطلاعات مربوط به موقعیت قلم، اطلاعات زمانی مربوط به مسیر قلم نیز در اختیار است. این اطلاعات معمولاً توسط یک صفحه رقومی‌کننده اخذ می‌شوند. در این روش می‌توان از اطلاعات زمانی سرعت، فشار و زمان برداشتن و گذاشتن قلم روی صفحه در بازشناسی استفاده کرد.

در بازشناسی برون خط، از تصویر دوبعدی متن ورودی استفاده می‌شود. در این روش به هیچ نوع وسیله نگارش خاصی نیاز نیست و تفسیر داده‌ها مستقل از فرآیند تولید آن‌ها و تنها براساس تصویر متن صورت می‌گیرد. این روش به نحوه بازشناسی توسط انسان شباهت بیشتری دارد.

## ۶-۱-۲ معرفی بخش‌های مختلف یک سیستم «اُسی آر»

شکل ۶-۱-۲ نمودار بلوکی یک سیستم «اُسی آر» را نمایش می‌دهد. لازم به ذکر است که بسته به الگوریتم کلی به کار رفته و سطح انتظارات از عملکرد نرم‌فزار، ممکن است برخی سیستم‌ها فاقد یک یا چند مرحله از مراحل فوق باشند.



شکل (۶-۱-۲) نمودار بلوکی دیاگرام یک سیستم «اُسی آر»

در ادامه هر یک از این بلوک‌ها مورد بررسی قرار می‌گیرند.

### ۶-۱-۲-۱ پیش پردازش

کلیه اعمالی که روی تصویر صورت می‌گیرند تا موجب تسهیل در روند اجرای فازهای بعدی گردد؛ مانند دوگانی کردن تصویر، حذف نویز، هموارسازی، نازکسازی، تشخیص زبان و فونت کلمات، و نظایر این‌ها. از مجموعه این پردازش‌ها، هدف‌های زیر دنبال می‌شود [۶]:

۱. کاهش نویز

۲. نرمالیزه کردن داده‌ها

۳. فشرده‌سازی میزان اطلاعاتی که می‌بایست محفوظ بماند.

## کاهش نویز:

نویز ایجادشده بواسطه دستگاه‌های اسکنر نوری منجر به ایجاد نقطه نقطه‌های لک‌مانند، قطعه خط‌های

گسسته<sup>۱</sup>، اتصال بین خطوط، فضا‌های خالی در خطوط متن، پرشدن حفره‌های موجود در تصویر برخی

حروف، و ... می‌گردد. همچنین اعوجاج‌های مختلف شامل تغییرات محلی، منحنی‌شدن گوشه‌های حروف،

تغییر شکل<sup>۲</sup> یا خوردگی حروف<sup>۳</sup> را نیز باید در نظر داشت. قبل از مرحله بازشناسی حروف، لازم است که این

نقایص برطرف شوند. مهم‌ترین دلیل برای کاهش نویز، کم‌کردن خطا در مراحل قطعه‌بندی و بخصوص

بازشناسی می‌باشد. کاهش نویز همچنین سبب کم‌شدن اندازه فایل تصویر می‌شود که به نوبه خود، کاهش

زمان مورد نیاز برای پردازش‌ها و ذخیره‌سازی‌های آینده را در پی خواهد داشت.

نرمالیزه کردن<sup>۴</sup> داده‌ها: روش‌های نرمالیزه کردن داده‌ها به حذف تغییرات نگارشی کمک می‌کند و

داده‌های استانداردشده‌ای را نتیجه می‌دهد. روش‌های پایه نرمالیزه کردن عبارت‌اند از:

**الف.** نرمالیزه کردن کجی<sup>۵</sup> متن و استخراج خطوط زمینه<sup>۶</sup>:

به دلیل بی‌دقتی در مرحله اسکن یا بی‌دقتی نویسنده در هنگام نگارش متن دست‌نوشته، ممکن است

خطوط متن نسبت به تصویر، اندکی انحراف یا چرخش داشته باشند شکل (۱-۲-۶-۱). این وضع ممکن است

کارایی الگوریتم‌های به کار رفته در طبقات بعدی سیستم «اُسی‌آر» را تحت تأثیر قرار دهد؛ چرا که یکی از

مفروضات در بیشتر روش‌های قطعه‌بندی، کج‌نبودن تصویر متن ورودی است و در نتیجه لازم است که این

نقیصه، آشکار و تصحیح گردد. آشکارسازی خط زمینه در بسیاری از تکنیک‌های قطعه‌بندی و بازشناسی

متون فارسی، عربی و لاتین، نقش اساسی دارد. علاوه بر این، برخی از کاراکترها مانند «۹» و «g» در نگارش

---

1 Disconnected line segments

2 Dilation

3 Erosion

4 Normalization

5 Skew

6 Baselines

لاتین و «.» (نقطه) و «۰» (صفر) در نگارش فارسی را بواسطه موقعیت نسبی‌شان نسبت به خط زمینه می‌توان آشکار ساخت.



شکل (۱-۶-۱-۲) تصویر یک صفحه که کج اسکن شده است

کلیه الگوریتم‌های توسعه داده شده برای آشکارسازی کجی صفحه، بر روی صفحات متنی با ترازبندی یکنواخت، دقیق عمل می‌کنند [۴۶]. الگوریتمی کارآتر است که به واسطه حضور مواردی نظیر گرافیک، پاراگراف‌های دارای کجی متفاوت، اعوجاج‌های منحنی - خطی<sup>۱</sup> ظاهرشونده در کتاب‌های فتوکپی‌شده، نواحی وسیع پیکسل‌های سیاه نزدیک حاشیه صفحه و خطوط متنی مختصر و کوتاه، دقت آن کمتر دستخوش تغییر شود. روش‌های به کار رفته برای تصحیح کجی خطوط زمینه در متون لاتین را می‌توان به چهار گروه اصلی دسته‌بندی کرد که عبارت‌اند از [۶]:

۱. به کارگیری هیستوگرام (پروفایل تصویرنمایی)<sup>۲</sup> تصویر

۲. استفاده از روش خوشه‌بندی نزدیک‌ترین همسایه‌ها<sup>۳</sup>

۳. روش همبستگی متقابل<sup>۴</sup> بین حروف

۴. تبدیل هاف<sup>۵</sup>

اغلب پس از آشکارسازی کجی، تصویر صفحه در جهت اصلی چرخانده می‌شود تا عملیات تحلیل قالب‌بندی متن و «آسی‌آر» با سهولت و دقت بیشتری انجام پذیرد. نمونه‌برداری مجدد مورد نیاز برای این منظور که

1 Curvilinear Distortion  
2 Projection Profile  
3 Nearest neighbors clustering  
4 Cross correlation  
5 Hough transform

باید بر روی صفحات دوگانی‌شده اعمال گردد، ممکن است الگوی کاراکترها را تغییر دهد. در این حالت به جای چرخاندن تصویر می‌توان الگوریتم‌های پردازشی را به نحوی اصلاح نمود که اثر چرخش در آن‌ها لحاظ گردد [۲۶].

همچنین می‌توان تصویر سند را قبل از دوگانی‌کردن، چرخش داد یا این که مقدار چرخش را از روی انتقال‌های کوچک و بدون اعوجاج کل بلوک‌های متنی، تقریب زد [۴۶].

#### ب. نرمالیزه کردن اریب‌شدگی<sup>۱</sup>:

در متون چاپی فارسی و لاتین، کاراکترهای دارای قالب ایتالیک از راستای عمود، انحراف دارند. در متون دست‌نوشته نیز برخی از نویسندگان حروف را به صورت زاویه‌دار می‌نویسند. این پدیده با عنوان «اریب‌شدگی» شناخته می‌شود و ممکن است دقت برخی از الگوریتم‌های قطعه‌بندی یا بازشناسی را تحت تأثیر قرار دهد و و از این رو در این سیستم‌ها لازم است که در مرحله پیش‌پردازش، میزان اریب‌بودن کاراکترها شناسایی و تصحیح گردد.

#### ج. نرمالیزه کردن (تغییر مقیاس دادن<sup>۲</sup>) اندازه:

در سیستم‌های «آسی آر»، اغلب تصاویر کلمات یا حروف خیلی کوچک یا خیلی بزرگ، به یک اندازه استاندارد نرمالیزه می‌شوند تا بدین ترتیب عملیات بازشناسی، مستقل از اندازه فونت متن گردد.

---

1 Slant Normalization  
2 Scaling

## د. هموارسازی کانتور<sup>۱</sup> :

خط تشکیل دهنده مرز یک کاراکتر را کانتور آن کاراکتر گویند. در متون دست‌نوشته، به واسطه لرزش یا حرکات ناخواسته دست نویسنده در هنگام نگارش، ممکن است که کانتور حروف ناصاف شود. این وضع در سیستم‌های بازشناسی متون چاپی و دست‌نوشته نیز، به دلیل تغییر مقیاس حروف یا وجود نویز در مرحله اسکن تصاویر ممکن است ظاهر گردد. روش‌های هموارسازی کانتور، به منظور جبران این نقیصه مورد استفاده قرار می‌گیرند. به طور کلی هموارسازی کانتور، تعداد نقاط نمونه مورد نیاز برای نمایش کاراکتر را کاهش می‌دهد و در نتیجه کارایی مراحل پردازشی باقیمانده را بهبود می‌بخشد.

ر. فشرده‌سازی:

این نکته پذیرفته شده است که تکنیک‌های کلاسیک فشرده‌سازی تصاویر که تصویر را از حوزه مکانی به حوزه‌های دیگر منتقل می‌کنند، برای بازشناسی حروف مناسب نیستند. در بازشناسی حروف، عمل فشرده‌سازی نیازمند آن دسته از تکنیک‌های حوزه مکانی است که اطلاعات شکلی را حفظ می‌نمایند. دو تکنیک متعارف فشرده‌سازی، یکی تکنیک اعمال سطح آستانه<sup>۲</sup> (به منظور دوگانی یا دوسطحی کردن تصاویر سطح خاکستری) و دیگری نازک‌سازی<sup>۳</sup> می‌باشند [۶].

الف. دوگانی (دوسطحی) کردن تصویر متن:

تصاویر دیجیتالی به یکی از سه صورت- تصاویر رنگی، تصاویر سطح خاکستری (مشابه تصویر یک تلویزیون سیاه و سفید که رنگ تصویر به صورت سیاه، سفید و طیفی از رنگ‌های خاکستری ظاهر می‌شود)، و تصاویر دوگانی یا دوسطحی (مشابه تصویر یک سند فکس‌شده که رنگ پیکسل‌های تصویر، تنها سیاه یا سفید است)- می‌باشند. به منظور کاهش حجم ذخیره‌سازی مورد نیاز و افزایش سرعت و سهولت پردازش، اغلب مطلوب است که با انتخاب یک سطح آستانه، تصاویر سطح خاکستری یا رنگی را به تصاویر دوگانی تبدیل نمود [۵۶].

---

1 Contour Smoothing

2 Thresholding

3 Thinning

ب-نازک‌سازی:

با این عمل، تصویر کاراکترها به تصویری با عرض یک پیکسل تبدیل می‌شود؛ درست مثل این که کاراکترها با یک قلم نوک باریک نوشته شده باشند. نازک‌سازی در حالی که کاهش قابل‌ملاحظه‌ای در حجم داده‌ها ایجاد می‌کند، اطلاعات شکلی کاراکتر را نیز حفظ می‌نماید (شکل (۲-۶-۱-۲))



شکل (۲-۶-۱-۲) اعمال عملیات نازک‌سازی بر روی یک تصویر متنی نمونه

بازشناسی خط، زبان و فونت: بازشناسی خط، تعداد کلاس‌های مختلف نمادهایی را که باید مورد ملاحظه قرار گیرند کاهش می‌دهد. شناسایی زبان متن، به منظور به کارگیری مدل‌های متنی خاص ضرورت دارد. طبقه‌بندی فونت‌ها، کاهش تعدد شکل‌های مختلف حروف در هر کلاس که لازم است در فرایند بازشناسی لحاظ گردند را به دنبال دارد و سبب می‌شود که امر شناسایی، تنها به یک کلاس فونت محدود گردد. بازشناسی خط و زبان و فونت در کاربردهایی مانند نمایه‌سازی و دستکاری اسناد نیز مطلوب می‌باشد.

#### ۲-۶-۱-۲ قطعه‌بندی<sup>۱</sup> (جداسازی)

قطعه‌بندی مرحله‌ای بسیار مهم برای سیستم‌های «آسی آر» مخصوصاً «آسی آر» فارسی و عربی

(که حروف کلمات به صورت سرهم نوشته می‌شوند) می‌باشد. قطعه‌بندی به دو گونه تقسیم می‌شود [۶].

۱. قطعه‌بندی بیرونی<sup>۲</sup>: که عبارت است از تفکیک قسمت‌های مختلف تصویر نظیر متن، گرافیک و خطوط

و نیز جدا کردن بخش‌های مختلف متن مانند پاراگراف‌ها، جملات یا کلمات؛

---

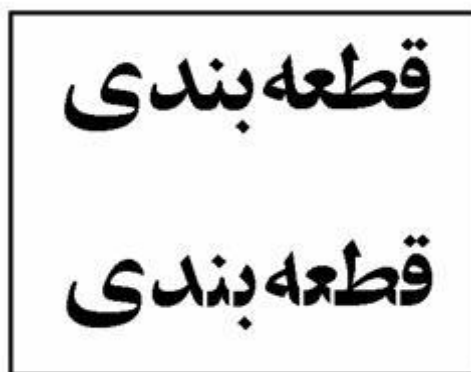
1 Segmentation

2 External Segmentation

۲. قطعه‌بندی درونی<sup>۱</sup>: که منظور از آن، جداسازی حروف کلمات مخصوصاً در مورد کلمات سر هم<sup>۲</sup>

نوشته‌شده در متون لاتین، یا در رسم‌الخط‌های پیوسته نظیر فارسی و عربی است شکل (۲-۱-۶-۳).

همچنین حروفی که در متن اصلی جدا بوده‌اند، اما به خاطر کیفیت پایین دستگاه اسکنر به هم چسبیده‌اند، توسط این دسته از تکنیک‌ها از یکدیگر جدا می‌شوند.



شکل (۲-۱-۶-۳) قطعه‌بندی<sup>۳</sup> یک کلمه به حروف

مرحله قطعه‌بندی بیرونی، بحرانی‌ترین و حساس‌ترین قسمت در حوزه تحلیل تصویر اسناد می‌باشد و یک مرحله ضروری برای سیستم‌های «اُسی‌آر» برون خط محسوب می‌شود. گرچه مبحث تحلیل اسناد با روش‌ها و تکنیک‌های خاص خود یک حوزه تحقیقاتی تا حدی متفاوت نسبت به «اُسی‌آر» است، اما تقسیم‌بندی تصویر سند به نواحی متنی و غیرمتنی، یک بخش لاینفک در نرم‌افزارهای «اُسی‌آر» به حساب می‌آید. بنابراین برای افرادی که در زمینه «اُسی‌آر» تحقیق می‌نمایند، داشتن دانش عمومی از تکنیک‌های آنالیز اسناد ضرورت دارد

نقطه تمایز اصلی میان سیستم‌های «اُسی‌آر» لاتین و فارسی برای متون چاپی، در مرحله قطعه‌بندی درونی نهفته است؛ چرا که حروف کلمات در نگارش فارسی برخلاف نگارش رسمی لاتین به صورت سرهم نوشته می‌شوند و در نتیجه ضرورت انجام صحیح این مرحله برای متون فارسی و عربی نسبت به متون لاتین، اهمیت فوق‌العاده بیشتری دارد. با وجود فعالیت‌های نسبتاً چشمگیر دهه گذشته و تنوع تکنیک‌های

---

1 Internal Segmentation

2 Cursive

3 Segmentation



عرضه شده، قطعه‌بندی متون پیوسته (بخصوص متون دست‌نوشته پیوسته) به حروف، هنوز هم یک مسئله قابل بررسی مانده است. روش‌های قطعه‌بندی حروف به سه دسته تقسیم می‌شوند:

۱. قطعه‌بندی صریح<sup>۱</sup>

۲. قطعه‌بندی ضمنی<sup>۲</sup>

۳. تکنیک‌های ادغام‌شده

در مواردی همچون متون فارسی که حروف به صورت سرهم نوشته می‌شوند، سه رویکرد مختلف در بازشناسی برون خط متون کلمات یا زیرکلمات وجود دارد [۸۶]:

۱. رویکرد مبتنی بر قطعه‌بندی کلمات

۲. رویکرد مبتنی بر بازشناسی کلمه به عنوان یک الگوی واحد

۳. رویکرد ترکیبی

در رویکرد بازشناسی مبتنی بر قطعه‌بندی، ابتدا کلمه در مرحله جداسازی به حروف یا زیرحروف، شکسته می‌شود؛ آنگاه قطعات جداشده بازشناسی می‌شوند و از کنار هم قرارگرفتن آن‌ها، کلمه شناسایی خواهد شد. روش‌های به کار گرفته‌شده در این رویکرد به دو گروه مختلف تقسیم می‌شوند:

- تقطیع کلمه به حروف

- تقطیع کلمه به زیرحروف

در گروه اول، کلمه به حروف جداسازی می‌شود و با شناسایی حروف جداشده، کلمه بازشناسی می‌گردد. در گروه دوم، کلمه به زیرحروف مثل پاره‌منحنی‌ها و ساختارهای پایه دیگر جداسازی می‌شود و با شناسایی زیرحروف‌ها و ترکیب آن‌ها، کلمه بازشناسی می‌گردد. در این رویکرد نمی‌توان در ابتدا مرز حروف را به طور کامل مشخص کرد، بلکه حروف به ترتیب از ابتدا به انتهای کلمه، بازشناسی و جداسازی می‌شوند. در هیچک از دو رویکرد نخست که مبتنی بر جداسازی هستند، به شکل کلی کلمه توجهی نمی‌شود و سعی بر آن است که با بازشناسی حروف یک کلمه، آن کلمه شناخته شود.

---

1 Explicit Segmentation  
2 Implicit Segmentation

در رویکرد بازشناسی کلمه به عنوان یک الگوی واحد، تلاشی برای تقطیع کلمه به حروف و بازشناسی حروف موجود در کلمه صورت نمی‌گیرد و کلمه در قالب یک الگو بررسی می‌گردد.

قطعه‌بندی غلط کاراکترها، عامل بسیاری از خطاهای «اُسی‌آر» است (مانند:  $rn \rightarrow m$  یا  $m \rightarrow rn$ ). میزان دقت یک الگوریتم قطعه‌بندی به سبک نگارش حروف، کیفیت دستگاه چاپ و نیز نسبت اندازه فونت به قدرت تفکیک دستگاه اسکنر (تابع گسترش نقاط و نرخ نمونه‌برداری مکانی) بستگی دارد [۴۶].

نتیجه مطلوب مرحله قطعه‌بندی، تصویری است که تنها حاوی یک کاراکتر باشد و بجز پیکسل‌های پس‌زمینه، هیچ شیء دیگری را شامل نشود. اما هنگامی که اشیای چاپی، در تصویر ورودی خیلی نزدیک به هم ظاهر شوند (مانند نقشه‌های هیدروگرافی)، این منظور همواره قابل حصول نخواهد بود. غالباً در چنین حالتی دیگر کاراکترها یا اشیای چاپی، به طور تصادفی در داخل تصویر کاراکتر قرار می‌گیرند و احتمالاً ویژگی‌های استخراج‌شده را تحریف می‌نمایند. این مورد یکی از دلایلی است که بیان می‌دارد چرا هر سیستم بازشناسی حروف، یک گزینه وازدگی دارد [۵۶]. «عزمی» بررسی جامعی بر روی روش‌های عرضه‌شده برای قطعه‌بندی کلمات در متون فارسی انجام داده است [۸۶].

## ۲-۱-۶-۳ بازنمایی<sup>۱</sup> (استخراج ویژگی‌ها<sup>۲</sup>)

این مرحله یکی از مراحل بسیار با اهمیت در سیستم‌های «اُسی‌آر» است؛ چرا که نتایج حاصل از این مرحله، مستقیماً بر روی کیفیت مرحله بازشناسی اثر می‌گذارد. در مرحله بازنمایی، به هر الگوی ورودی (کاراکتر یا کلمه بر حسب آن که رویکرد سیستم، مبتنی بر کدامیک از دو گروه «قطعه‌بندی کلمات» یا «شناسایی کلمه به عنوان یک الگوی واحد» باشد)، یک کد یا بردار ویژگی نسبت داده می‌شود که معرف آن الگو در فضای ویژگی‌ها است و آن را از دیگر الگوها متمایز می‌سازد.

در انتخاب بردارهای ویژگی لازم است موارد زیر مورد توجه قرار گیرند:

۱. بردار ویژگی هر الگو باید تا حد امکان از بردارهای ویژگی دیگر الگوها متمایز باشد

(فاصله بین بردارهای ویژگی در فضای ویژگی‌ها، حداکثر باشد).

---

1 Representation  
2 Feature extraction

۲. بردار ویژگی الگوها باید تا بیشترین حد ممکن، خصوصیات شکل و ساختار الگوها را از تصویر آنها استخراج نماید.

. تا حد امکان نسبت به نویز، تغییر اندازه و نوع فونت، چرخش، و دیگر تغییرات احتمالی الگوها دارای ثبات باشد.

۴. شرایط، نوع و خصوصیات الگوهای ورودی در انتخاب بردارهای ویژگی اثر می‌گذارند. به عنوان مثال، باید تعیین نمود که آیا حروف یا کلماتی که می‌بایست تشخیص داده شوند جهت و اندازه مشخصی دارند یا خیر، دست‌نوشته هستند یا چاپی، یا اینکه تا چه حد بوسیله نویز، مغشوش شده‌اند. همچنین گاهی کفایت می‌کند که سیستم، تنها جوابگوی گروه محدودی از الگوها (مثلاً الگوهای با اندازه یا نوع فونت از پیش مشخص شده) باشد.

۵. در مورد حروفی که به چندین شکل نوشته میشوند

لازم است که بیش از یک کلاس الگو به یک کاراکتر خاص تعلق یابد.

همانطور که عنوان شد، بازنمایی یک مرحله بسیار مهم در حصول راندمان مناسب برای سیستم‌های بازشناسی حروف است؛ ولی برای دستیابی به عملکرد بهینه، لازم است که دیگر مراحل نیز بهینه گردند و باید توجه نمود که این مراحل، مستقل از هم نیستند. یک روش خاص استخراج ویژگی‌ها، طبیعت خروجی مرحله پیش‌پردازش را به ما دیکته می‌کند یا حداقل ما را در انتخابمان محدود می‌سازد.

مراحل قطعه‌بندی و بازشناسی، دو وجه تمایز عمده میان سیستم‌های «اُسی‌آر» فارسی و لاتین می‌باشند. بواسطه وجود تفاوت‌های اساسی بین نحوه نگارش فارسی و لاتین، امکان اعمال مستقیم تکنیک‌های قطعه‌بندی و بازشناسی مربوط به سیستم‌های «اُسی‌آر» لاتین، برای متون فارسی وجود ندارد. پیچیدگی‌های مختص نگارش فارسی، بر پیچیدگی الگوریتم‌های این دو مرحله می‌افزاید. درست به همین دلیل است که بیشتر نرم‌افزارهای «اُسی‌آر» تجاری لاتین، قادر به پشتیبانی زبان فارسی و عربی نمی‌باشند.

## ۲-۱-۶-۴ طبقه‌بندی و بازشناسی (با یک یا چند طبقه‌بندی‌کننده)

این مرحله شامل روش‌هایی برای متناظر ساختن هر یک از الگوهای به دست‌آمده از مرحله استخراج ویژگی‌ها، با یکی از کلاس‌های فضای الگوهای مورد بحث است که از طریق کمینه ساختن فاصله بردار ویژگی‌های هر الگوی ورودی نسبت به یکی از بردارهای مرجع، انجام می‌گیرد. بردارهای مرجع، بردارهایی هستند که قبلاً از نمونه‌های آموزشی اخذ شده‌اند. تکنیک‌های عرضه‌شده برای این مرحله را می‌توان در روش‌های مربوط به چهار گروه عمومی مبحث شناسایی الگو، جستجو نمود [۶].

- تطابق قالبی

- تکنیک‌های آماری

- تکنیک‌های ساختاری

- شبکه‌های عصبی

چهار گروه فوق لزوماً مستقل یا مجزا از یکدیگر نمی‌باشند و گاهی می‌توان تکنیک‌های یک گروه را در میان تکنیک‌های مربوط به دیگر گروه‌ها یافت.

## ۲-۱-۶-۵ به کارگیری اطلاعات جانبی (پس‌پردازش)

در این مرحله با استفاده از اطلاعات جانبی (نظیر مجموعه لغات معتبر، اطلاعات آماری مربوط به رخداد حروف، اطلاعات دستوری و معنایی) سعی در بهبود نتایج حاصل از مرحله بازشناسی می‌گردد. تا قبل از این مرحله، هیچگونه اطلاعات معناساختی در طول مراحل پردازشی مورد استفاده قرار نگرفته بود. مرحله پیش‌پردازش سعی در «تمیز کردن» تصویر سند به نحو مقتضی دارد و به علت عدم دسترسی به اطلاعات مفهومی، ممکن است باعث حذف برخی از اطلاعات مهم تصویر گردد. فقدان اطلاعات معنایی در مرحله قطعه‌بندی می‌تواند به نتایج حادثر و غیرقابل برگشتی بینجامد؛ چرا که خروجی این مرحله، تعیین‌کننده مرز الگوهای ورودی می‌باشد. بنابراین واضح است که در صورت فراهم‌شدن اطلاعات معناساختی، دقت

نتایج بازشناسی به نحو چشمگیری افزایش می‌یابد. مروری بر تحقیقات اخیراً انجام‌شده در زمینه بازشناسی حروف نشان می‌دهد که در صورت استفاده از اطلاعات شکلی بدون به‌کارگیری اطلاعات معناشناختی، افزایش دقت قابل توجهی نخواهیم داشت. در نتیجه، یکپارچه‌کردن اطلاعات معنایی و شکلی در کلیه بلوک‌های سیستم‌های «آسی آر» به منظور بهبود نرخ بازشناسی صحیح، ضرورت دارد.

ساده‌ترین راه برای این منظور، استفاده از یک فرهنگ لغات برای اصلاح خطاهای جزئی است. این کار توسط یک برنامه واژه‌پرداز (با قابلیت خطایاب املایی) که املای کلمات را کنترل و چندین پیشنهاد برای اصلاح املای کلمات نامأنوس ارائه می‌کند، انجام می‌گیرد.

## ۷-۱-۲ پیشنهادهایی در جهت تسریع دستیابی به یک سیستم اُسی آر فارسی

همانطور که ملاحظه شد، فرآیند تبدیل یک تصویر سند به متن قابل ویرایش یا همان «اُسی آر»، پیچیده است و دستیابی به یک سیستم «اُسی آر» کارآمد فارسی، در درازمدت و با انجام تحقیقات وسیع و متمرکز محقق خواهد شد. در ادامه، راهکارهایی به منظور تسریع در دستیابی به یک سیستم «اُسی آر» فارسی پیشنهاد می‌گردد:

- برخی از عملیات پردازشی سیستم «اُسی آر» وابستگی کمتری به نوع زبان متن دارند و روش‌های پردازشی مورد استفاده در سیستم‌های «اُسی آر» لاتین، مستقیماً یا با انجام تغییرات جزئی برای آن‌ها قابل اعمال هستند. به عنوان مثال، بسیاری از الگوریتم‌های پیشنهادشده در مقالات برای عملیات پیش‌پردازش (نظیر دوگانی‌کردن تصویر، کاهش نویز، تصحیح کجی صفحه، و ...) برای هر متنی صرف نظر از نوع زبان آن قابل استفاده می‌باشند. همچنین بسیاری از ایده‌های ارائه‌شده برای قطعه‌بندی بیرونی، بازشناسی، آموزش و پس‌پردازش متون لاتین را می‌توان با انجام یکسری تغییرات، برای متون فارسی نیز استفاده نمود.

- گلوگاه یک سیستم «اُسی آر» فارسی و نقطه تمایز آن با «اُسی آر» لاتین، عملیات قطعه‌بندی درونی است که چنانچه بتوان بخوبی آن را انجام داد، بجرأت می‌توان عنوان نمود که نیمی از راه را طی نموده‌ایم. بنابراین تیمی که قرار است در زمینه «اُسی آر» فعالیت نماید، بجا است که تمرکز اصلی خود را بر این قسمت معطوف دارد.

## ۲-۲ روشهای مختلف در حوزه بازشناسی اسناد

در ادامه این بخش روشهای مختلف بازنمایی تصاویر اسناد را که به سه گروه اصلی طبقه بندی شده اند را بررسی می نماییم.

### ۲-۲-۱ تبدیل سراسری<sup>۱</sup> (بسط سری<sup>۲</sup>)

یک سیگنال پیوسته معمولاً حاوی اطلاعات بیشتری از میزان مورد نیاز برای بازنمایی جهت مقاصد کلاسه بندی است. این موضوع می تواند برای تقریبهای گسسته سیگنالهای پیوسته نیز صادق باشد. یک روش برای بازنمایی یک سیگنال عبارتست از یک بسط سری بصورت ترکیب خطی از توابع ساده تری (دارای پیچیدگی کمتر) که به نحو مناسبی تعریف شده اند. در واقع، ضرایب ترکیب خطی، یک روش کد کردن فشرده تر سیگنال می باشند که به آنها تبدیل یا بسط سری سیگنال اطلاق می شود. در زیر، معمولی ترین تبدیلهای بسط سری در حوزه سیستمهای بازشناسی حروف بررسی می گردند.

### ۲-۲-۱-۱ تبدیلات فوریه<sup>۳</sup> (توصیف کننده های فوریه<sup>۴</sup>)

تبدیل فوریه گسسته (DFT)<sup>۵</sup> از اهمیت ویژه ای در حوزه پردازش تصویر و سیگنال برخوردار است و در آن معمولاً از طیف دامنه بردار ویژگی در فضای اقلیدسی  $n$  بعدی به عنوان ویژگی استفاده می شود. این تبدیل از توابع پایه نمایی مختلط بهره می گیرد و چون از خاصیت فشرده کرده انرژی بالایی برخوردار است، راندمان کدینگ مطلوبی را ارائه می دهد. DFT چندین ویژگی دارد که در امر شاخص گذاری<sup>۶</sup> یا تطبیق الگوها مفید واقع می گردد: نخست آنکه دامنه ضرایب

---

1 Global Transformation  
2 Series Expansion  
3 Fourier Transforms  
4 Fourier Descriptors  
5 Discrete Fourier Transform  
6 Indexing

DFT نسبت به انتقال حساس نیست<sup>۱</sup> و در نتیجه با احتساب طیف دامنه به تنهایی و صرف نظر از اطلاعات فاز می توان حساسیت نسبت به تغییر مکان کاراکترها در تصویر را از بین برد. در ثانی، همبستگی حوزه زمانی را می توان بطور مؤثر با استفاده از ضرایب DFT محاسبه نمود. تبدیلات فوریه به راههای زیادی در بازشناسی حروف مورد استفاده قرار گرفته اند.

روشهای مختلف کد کردن منحنی پیرامون در موارد زیادی برای بازشناسی حروف استفاده شده اند. در این میان توصیف کننده های فوریه نرمالیزه که نسبت به اندازه، انتقال و چرخش ثابت هستند، کاربرد بیشتری دارند. از این ویژگیها در بازشناسی حروف لاتین و کاربردهای پردازش تصویر استفاده زیادی شده است. از جمله کارهایی که با استفاده از این روش بر روی متون فارسی و عربی صورت گرفته است، عبارتند از [۸۶]:

الف – استخراج مستقیم توصیف کننده های فوریه از منحنی پیرامونی حروف.

ب – استفاده از کدهای فریمین در کد کردن منحنی پیرامونی و استخراج ضرایب فوریه از آن. چنانچه منحنی پیرامونی حرف را به صورت یک منحنی دو بعدی با مختصات  $x(n)$  و  $y(n)$  در نظر بگیریم، توصیف کننده های فوریه با استفاده از تبدیل فوریه گسسته بصورت زیر تعریف می

شوند:

$$u(n) = x(n) + jy(n)$$

$$u(n) = \frac{1}{N} \sum_{k=0}^{N-1} a(k) \exp\left(\frac{j2\pi kn}{N}\right), \quad 0 \leq n \leq N-1$$

$$a(n) = \frac{1}{N} \sum_{k=0}^{N-1} u(k) \exp\left(\frac{-j2\pi kn}{N}\right), \quad 0 \leq n \leq N-1$$

در معادله فوق  $a(n)$  ضرایب فوریه را نشان می دهد. همانطور که گفته شد، معمولاً در کاربردهای بازشناسی تنها اندازه این ضرایب استفاده می شود.

---

1 Translation invariant



اخیراً تبدیلات زیرباند و موجک گسسته (DWT)<sup>۲</sup> در عرصه کاربردهای کدینگ، شاخص گذاری تصاویر<sup>۳</sup> و بازشناسی حروف با مقبولیت فراوانی مواجه شده‌اند. در این تبدیلهای، یک تصویر بطور پیاپی از مجموعه‌ای از فیلترهای پایین‌گذر و بالا‌گذر عبور نموده و خروجی فیلترها یک چندی<sup>۴</sup> می‌شود تا نرخ داده‌ها ثابت باقی بماند. در DWT، تنها خروجی پایین‌گذر به صورت پیاپی فیلتر می‌گردد.

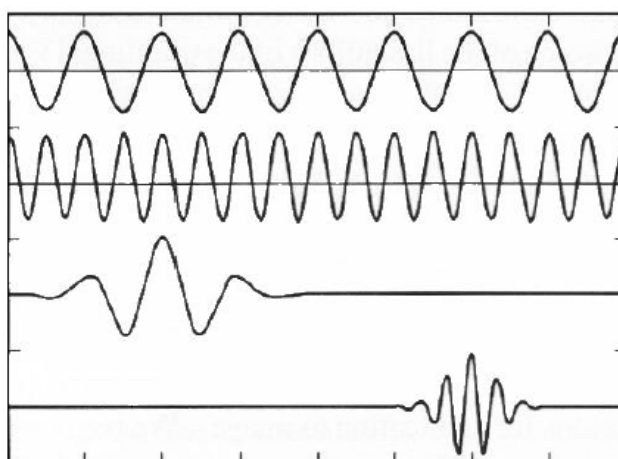
در تبدیل فوریه که نوعی از تبدیل با تابع پایه ارتوگونال (متعامد)<sup>۵</sup> می‌باشد، از «موج»<sup>۶</sup> های سینوسی استفاده می‌شود که در هر دو جهت دارای دوره نامحدود هستند. بطور مشابه در حالت گسسته، بردارهای پایه تبدیل DFT نیز دارای دوره نامحدود می‌باشند و بنابراین این تبدیل دارای ویژگی «دوره - محدود»<sup>۷</sup> نمی‌باشد. اما در عمل به سیگنالهای گذرای بر می‌خوریم که تنها در یک دوره کوتاه زمانی غیر صفر هستند. مثلاً بسیاری از خصوصیات مهم تصاویر مانند لبه، در یک مکان مشخص قرار دارند. این ترکیبات به هیچیک از توابع پایه تبدیل فوریه شبیه نیستند و در نتیجه توابع مذکور ابزار بهینه‌ای برای آنالیز سیگنالها و تصاویری که حاوی مؤلفه‌های گذرا یا متمرکز در یک زمان یا مکان هستند، نمی‌باشند. البته می‌دانیم که تبدیل فوریه هر تابع تحلیلی (حتی یک سیگنال گذرای باریک) را بصورت مجموعی از سینوسها نمایش دهد. اما ضرایب بدست آمده از تبدیل، معیار فشرده‌ای از اطلاعات تصویر نخواهند بود.

برای رفع این نقیصه ریاضیدانان و مهندسين به کوششهایی دست زده‌اند که در آنها توابع پایه با دوره محدود بکار برده می‌شوند. به این توابع پایه «موجک» و به تبدیلاتی که بر اساس آنها هستند، «تبدیلات موجک» می‌گویند. شکل (۲-۱-۲-۲) تفاوت بین موج و موجک را نشان می‌دهد.

---

1 Wavelets  
2 Discrete Wavelet Transform  
3 Image Indexing  
4 Decimate  
5 Orthogonal  
6 Wave  
7 Compactly supported

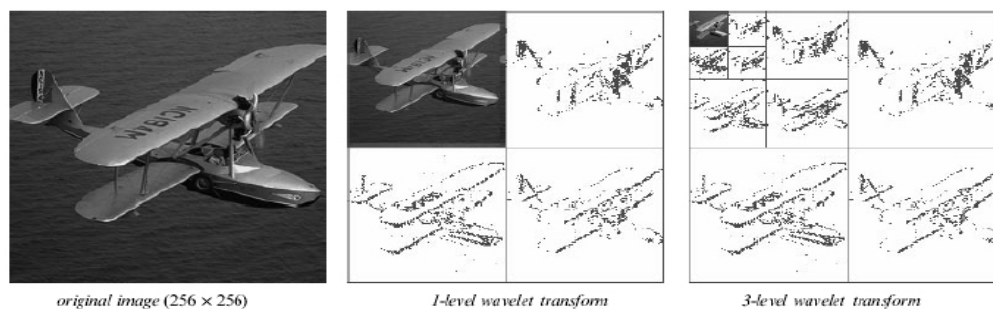
دهد. دو منحنی بالایی ، امواج سینوسی با فرکانسهای مختلف و دو منحنی پایینی ، دو موجک با فرکانس و موقعیت متفاوت می باشند .



شکل (۲-۱-۲-۲)

برای بسیاری از سیگنالها ، تحلیل فوریه ابزاری بینهایت مفید محسوب می شود ؛ زیرا پی بردن به محتوای فرکانسی سیگنالها از اهمیت ویژه ای برخوردار است . پس به چه دلیل ما نیازمند تکنیکهای دیگری همچون تحلیل موجک می باشیم ؟ تحلیل فوریه یک نقطه ضعف اساسی دارد و آن اینکه در فرایند تبدیل ک سیگنال به حوزه فرکانس ، اطلاعات زمانی یا مکانی آن از بین می رود . به عبارت دیگر ، با نگاه کردن به تبدیل ف.ریه یک سیگنال نمی توان اظهار داشت که یک رویداد خاص ، در چه برهه زمانی و یا در چه مکانی از تصویر بوقوع پیوسته است . چنانچه تغییرات یک سیگنال در زمان یا مکان قابل ملاحظه نباشد ( سیگنال ایستادن ) ، این نقطه ضعف اهمیت چندانی ندارد . لیکن بیشتر سیگنالهای مورد توجه ، در بسیاری نقاط ( مانند نقاط ابتدایی و انتهایی ) از ماهیت غیر ایستادن یا گذرا برخوردارند . این نقاط اغلب مهمترین قسمت سیگنال می باشند که تحلیل فوریه ابزار مناسبی برای آشکارسازی آنها محسوب نمی شود .

تبدیل موجک یک تکنیک بسط سری است که این امکان را بوجود می آورد که بتوان سیگنال را در درجات مختلف تفکیک پذیری نمایش داد. قطعات مختلف تصویر سند که ممکن است مربوط به حروف یا کلمات باشند، متناظر با سطوح مختلف تفکیک پذیری بوسیله ضرایب موجک نشان داده می شوند. بازنمایی سیگنال در فرم آنالیز چند تفکیکه (MRA)<sup>1</sup> با درجه تفکیک پذیری پایین می تواند در مقایسه با MRA با درجه تفکیک پذیری بالا، تغییرات محلی حادث شده در دست نوشته ها را بنحو مؤثری به خود جذب نماید. لیکن بازشناسی در تفکیک پذیری پایین ممکن است باعث شود که جزئیات با اهمیت برای مرحله بازشناسی از دست برود. شکل پایین تبدیل موجک یک تصویر نمونه را نشان می دهد که در آن، ضرایب زیر یک سطح آستانه حذف گردیده اند.



تبدیل موجک انواع مختلفی دارد که از توابع پایه متفاوتی برخوردارند؛ نظیر: Haar، Mexican، Meyer، Morlet، Symlets، Coiflets، Biorthogonal، Daubechies و hat و غیره. از آنجائیکه توابع پایه موجک Haar شکل گسسته دارند، تصور می شود که این موجک برای کار با تصاویر باینری کاراکترها، بر سایر موجکها ارجحیت داشته باشد. در اینجا از ذکر ریاضیات تبدیلات موجک بدلیل حجم نسبتاً زیاد آن صرف نظر می شود. اخیراً Mowlaei [44] و همکارانش روشی بمنظور بازشناسی کدهای پستی و اسامی شهرهای ایران که بر روی بسته های پستی بصورت دست نویس نوشته می شوند، پیشنهاد نموده اند. بردارهای ویژگی این روش، از طریق بکارگیری تبدیل موجک گسسته با موجک اساسی Haar محاسبه می گردند. در مرحله

---

1 Multi-Resolution Analysis

پیش پردازش این الگوریتم ، ابتدا محدوده هر حرف مشخص گردیده ، تصویر حرف به ابعاد  $64 \times 64$  پیکسل تغییر اندازه می یابد و در نهایت عمل نازکسازی انجام می گیرد . در مرحله استخراج ویژگیها ، تبدیل موجک سه سطحی بر روی تصویر نازک شده اعمال می شود و ضرایب باند LL از سطح سوم استخراج می گردند . در انتها ماتریس ضرایب باند LL بصورت برداری درآمده ، یک فیلتر پایین گذر بمنظور هموارسازی بیشتر ضرایب ، به آن اعمال می شود . برای آموزش سیستم ، یک شبکه عصبی MLP با قانون پس انتشار خطا مورد استفاده قرار می گیرد . راندمان این روش برای حروف و ارقام دست نویس فارسی بترتیب  $92/33$  و  $91/81$  درصد گزارش شده است .

## ۲-۱-۳ تبدیل گابور<sup>۱</sup>

این تبدیل تغییر یافته تبدیل فوری پنجره شده است که در آن پنجره مورد استفاده بجای آنکه یک اندازه گسسته داشته باشد ، بوسیله یک تابع گوسی تعریف می گردد . به بیان دیگر ، تبدیل گابور مشابه تبدیل موجک است که در آن توابع پایه طبیعت گوسی دارند و در نتیجه این تبدیل در جانمایی زمانی - فرکانسی<sup>۲</sup> بهینه است . موجک گابور از جهت حداقل سازی عدم قطعیت<sup>۳</sup> دو بعدی توأم بین حوزه های مکان و فرکانس ، یک تبدیل بهینه می باشد . این موجک را می توان به عنوان آشکارسازهای جهت دار و تطبیق پذیر با مقیاس ، بمنظور آشکارسازی خطوط و لبه ها در تصاویر مورد استفاده قرار داد . همچنین از خواص آماری این تبدیل می توان جهت تعیین ساختار و محتوای بصری تصاویر بهره گرفت . ویژگیهای تبدیل گابور در چندین کاربرد آنالیز تصاویر شامل کلاسه بندی و قطعه بندی بافتها ، بازشناسی تصاویر ، بازشناسی حروف ، ثبت تصاویر و مسیریابی حرکت بکار رفته اند . به رغم غیر متعامد بودن این خانواده موجکها ، تابع گابور تنها تابعی است که می تواند حد تئوری دقت تفکیک توأم اطلاعات در هر دو حوزه مکان و فرکانس را حاصل نماید .

---

1 Gabor Transform

2 Time-frequency localization

3 Uncertainty

یک موجک گابور دو بعدی عبارت است از یک تابع گوسی مدوله شده با یک تابع سینوسی مختلط. این تابع را می توان بوسیله فرکانس تابع سینوسی ( $W$ ) و انحراف معیارهای  $\sigma_x$  و  $\sigma_y$  پوشش تابع گوسی بصورت زیر بیان نمود :

$$\psi(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + 2\pi j W x\right] \quad (3-1-2-2)$$

علیرغم این شکل کلی ، هیچ تعریف استاندارد و دقیقی از تابع گابور دو بعدی وجود ندارد و اشکال مختلفی از این تابع در مقالات ظاهر شده است . بیشتر این اختلافات به اندازه های مختلف پوش تابع گوسی و فرکانس تابع سینوسی مربوط می شود . بر اساس تئوری موجک ، تابع گابور که بنحو مناسبی نرمالیزه شده باشد ، می تواند به عنوان یک موجک مادر برای تولید خانواده ای از موجکهای گاور غیر متعامد بکار رود .

Ma و Manjunath در [40] بمنظور شاخص گذاری و بازیابی تصاویر بافت گونه ، بانکی از فیلترهای گابور نرمالیزه شده با پاسخ dc صفر را مورد استفاده قرار داده اند . این فیلترها بنحوی طراحی شده اند که نمایش آنها ، درجه افزونگی پایینی را شامل شود . موجکهای گابور بکار رفته در اینجا ، از طریق انبساط<sup>۱</sup> و چرخش تابع  $\psi(x, y)$  در (3-1-2-2) ایجاد شده اند که تابع مولد آنها

$$\begin{aligned} \psi_{mn}(x, y) &= a^{-m} \psi(x', y'), \quad a > 1, \quad m, n = \text{Integers} \quad : \text{ به فرم زیر می باشد} \\ x' &= a^{-m} [(x - x_0) \cos(\theta) + (y - y_0) \sin(\theta)] \\ y' &= a^{-m} [-(x - x_0) \sin(\theta) + (y - y_0) \cos(\theta)] \end{aligned}$$

که در آن  $(x_0, y_0)$  مرکز فیلتر در حوزه مکان ،  $\theta = n\pi/K$  و  $S$  و  $K$  بترتیب تعداد کل مقایسهها<sup>۲</sup> و راستاهای<sup>۳</sup> در نظر گرفته شده می باشد . فاکتور  $a^{-m}$  در معادله فوق به این دلیل ظاهر شده است که از مستقل بودن انرژی نسبت به  $m$  اطمینان حاصل شود .

---

1 Dilation  
2 scales  
3 orientations

جهت حصول اطمینان از اینکه بانک فیلتری طراحی شده را بتوان یک خانواده مناسب از موجکهای گابور دو بعدی در نظر گرفت ، فیلترهای  $\psi(x,y)$  بایستی شرط پذیرش « محدودیت انرژی » را برآورده سازند که بر پایه آن ، تبدیل فوریه این فیلترها باید توابع میان گذر خالص با پاسخ dc صفر باشد . این شرط با صفر نمودن پاسخ فرکانسی فیلترها در فرکانس صفر ( $\psi(0,0)=0$ ) و یا با اضافه نمودن یک عدد ثابت به قسمت حقیقی ( زوج ) فیلترها تحقق می یابد که بدین وسیله از عدم پاسخ فیلترها به نواحی با شدت ثابت اطمینان حاصل می شود . بحث فوق ، روابط زیر را بمنظور محاسبه پارامترهای  $\sigma_v$  و  $\sigma_u$  فیلتر نتیجه می دهد [40] :

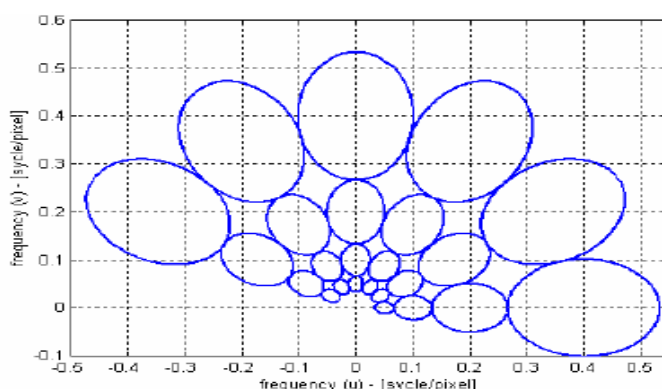
$$a = \left( \frac{U_h}{U_l} \right)^{\frac{1}{s-1}} \text{ و } \sigma_u = \frac{(a-1)U_h}{(a+1)\sqrt{2 \ln 2}} \quad \sigma_v = \tan\left(\frac{\pi}{2K}\right) \times \sqrt{\frac{U_h^2}{2 \ln 2} - \sigma_u^2}$$

در این مقاله طراحی جهتها و مقیاسها مبتنی بر سه اصل زیر بوده است :

۱- تفکیک یکنواخت جهتها : کلیه فیلترهای متعلق به یک مقیاس خاص می بایست انحراف استاندارد زاویه ای ( $\sigma_v$ ) یکسانی داشته و اختلاف در جهت گیری آنها نسبت به یکدیگر یکسان باشد .

۲- تفکیک نمایی مقیاسها : پهنای فیلترها باید با افزایش فاصله از مرکز صفحه ( $u,v$ ) بصورت نمایی افزایش یابد .

۳- پوشش پیوسته فضای فرکانسی : طراحی فیلترها می بایست بنحوی باشد که کانترهای نصف توان مجاور در هر دو سمت محورهای مقیاس و جهت با یکدیگر تماس داشته باشند . شکل زیر بانک فیلترهای موجک گابور را مطابق طراحی فوق نمایش می دهد .



از تبدیل گابور در چند مقاله بمنظور بازشناسی حروف چینی و لاتین استفاده گردیده است . Hue و همکارانش [24] الگوریتمی بمنظور قرائی ارقام پلاک خودروها با استفاده از فیلترهای گابور معرفی نموده اند . ابتدا مجموعه ای از فیلترهای گابور<sup>۱</sup> ، بردار ویژگی تصویر محاسبه می گردد . آزمایش نشان می دهد که این الگوریتم قادر به تشخیص مناسب حروف حتی در زمینه های بسیار نویزی است . راندمان این روش ۶۶-۹۶ درصد برآورد شده است .

GE و همکارانش [20] بتازگی یک موتور OCR برای شناسایی دستنوشته های زبان چینی ابداع نموده اند که از تکنیکهای مختلفی چون ویژگیهای گابور و مشتقات مکانی آنها به عنوان ویژگیهای خام ، آنالیز مبین خطی ( LDA )<sup>۲</sup> بمنظور استخراج ویژگیها و کاهش ابعاد فضای ویژگی ، CDHMM ها برای مدل کردن کاراکترهای چینی در راستای هر دو جهت افقی و عمودی ، و معیار حداقل خطای کلاسه بندی جهت آموزش مدل بهره گرفته اند . در زمان تدوین این گزارش ، موتور OCR آنها از یک فرهنگ لغت با ۴۶۱۶ کاراکتر ( شامل ۴۵۱۶ کاراکتر چینی ساده ، ۶۲ کاراکتر الفبایی - عددی<sup>۳</sup> و ۳۸ علامت و نماد نقطه گذاری ) پشتیبانی می کند . با بکارگیری ۱۳۸۴۸۰۰ نمونه کاراکتر جهت آموزش سیستم بازشناسی کننده ، دقت بازشناسی ۹۶/۳۴ درصد حاصل گردیده است و این در حالی است که بهترین سیستم OCR چینی موجود در بازار ، روی این مجموعه از راندمان ۹۴/۰۷ درصد برخوردار بوده است .

## ۲-۱-۲-۴ ممانها (گشتاورها)<sup>۴</sup>

ممانها بطور وسیعی در کاربردهای پردازش تصویر بمنظور توصیف شکل هندسی اشیاء تصویری گوناگون مورد استفاده قرار گرفته اند . این کمیتها خواص هندسی بنیادی ( مانند سطح ، مرکز جرم، ممان انرسی ، میزان اریب بودن و درجه تیزی نقطه اوج منحنی<sup>۵</sup> ) یک توزیع را نشان می -

---

1 Gabor Dominant Orientation Matrix

2 Linear Discriminant Analysis

3 Alphanumeric

4 Moments

5 kurtosis

دهند . همچنین ممانها را می توان بمنظور تقریب زدن تابع چگالی احتمال<sup>۱</sup> ( هیستوگرام ) تصاویر بکار برد . در اینجا خاطر نشان می شود که PDF سطوح خاکستری ( شدت روشنایی ) یک تصویر ، همان هیستوگرام آن می باشد ؛ جز اینکه با یک تغییر مقیاس ، سطح کل زیر منحنی آن برابر واحد است .

ممانهای نرمالیزه ویژگیهای مناسبی برای بازشناسی حروف هستند و در موارد زیادی بکار گرفته شده اند . مزیت این ویژگیها غیر حساس بودن آنها نسبت به چرخش ، تغییر مکان و اندازه حروف است . ممانهای مختلف نظیر ممانهای مرکزی<sup>۲</sup> ، ممانهای منظم<sup>۳</sup> ، ممانهای لژاندار<sup>۴</sup> و ممانهای شبه زرنیکه<sup>۵</sup> نمایش فشرده ای از تصویر سند را ارائه می دهند . ممانها بصورت بسط سریها در نظر گرفته می شوند ؛ زیرا تصویر اولیه را می توان از ضرایب ممانها بطور کامل بازسازی نمود .

گشتاورهای هندسی و نرمالیزه بصورت زیر تعریف می شوند :

$$\mu_{pq} = \sum_x (x - \bar{x})^p (y - \bar{y})^q f(x, y)$$

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\frac{p+q+2}{2}}}$$

در روابط فوق ،  $\mu_{pq}$  گشتاور مرکزی و  $\eta_{pq}$  گشتاور مرکزی نرمالیزه از مرتبه p و q است .

در [7] الگوریتمی برای بازشناسی حروف دستنویس فارسی با استفاده از این ویژگیها ارائه شده است . در این الگوریتم ، حروف ورودی با استفاده از خواصی مانند تعداد قطعات و وجود یا عدم وجود تک نقطه به پنج گروه طبقه بندی می شوند . در هر گروه ، کلاس حرف ورودی با استفاده از طبقه بندی بیز و ویژگیهای گشتاوری تعیین می شود . در نهایت در یک مرحله پس پردازش ، با بکار بردن برخی ویژگیهای ساختاری ، میزان خطا در کلاسههای مشابه کاهش داده می شود . میزان بازشناسی درست برای ۳۳۰۰ نمونه از حروف مجزای دستنویس که توسط ۱۰۰ نویسنده مختلف نوشته شده حدود ۹۱ درصد گزارش شده است .

---

1 Probability Density Function (PDF)

2 Central moments

3 Regular moments

4 Legendre moments

5 Pseudo-Zernike



در [3] برای شناسایی حروف عربی ، ابتدا بدنه آنها جدا شده و سپس هیستوگرامهای عمودی و افقی مربوط به تصویر بدنه محاسبه شده است [۸۲] . با محاسبه گشتاورهای هندسی یک بعدی برای هیستوگرامهای عمودی و افقی ، مجموعه ای از ۹ ویژگی نرمالیزه در بازشناسی بکار گرفته شده اند . برای طبقه بندی ، از یک تابع ممیز مرتبه ۲ ( نوع خاصی از طبقه بندی کننده بیز ) استفاده شده است . میزان بازشناسی برای سه قلم مختلف حدود ۹۸٪ گزارش شده است .

در [2] بازشناسی حروف عربی با استفاده از ویژگیهای گشتاوری و شبکه عصبی مطرح است [۸۲] . سیستم بازشناسی دارای سه مرحله مختلف است . ابتدا کلمات با استفاده از هیستوگرام عمودی به حروف شکسته می شوند . در مرحله بعد شش ویژگی گشتاوری مختلف برای بازشناسی حروف استخراج می شود . در نهایت یک شبکه عصبی پرسپترون با الگوریتم آموزشی پس انتشار خطا برای بازشناسی بکار برده شده است .

از گشتاورهای هندسی مرکزی نرمالیزه و ثابت های گشتاوری در [۸۰] برای بازشناسی حروف تایپ شده با یک ماشین تحریر استفاده شده است . در سیستم طراحی شده ، بعد از استخراج ویژگیهای نرمالیزه از طبقه بندی کننده بیز برای تشخیص حروف و علائم استفاده شده است . تحقیق انجام شده نشان داده است که میزان بازشناسی برای یک نوع قلم در حدود ۱۰۰ درصد است و ویژگیهای استخراج شده به نوع قلم حساس هستند .

در [۷۶] برای بازشناسی حروف دستنویس (نسخ) ، بعد از مرحله جداسازی ابتدا نقاط و علائم از بدنه حروف جدا شده و با استفاده از گشتاورهای لژاندار از مرتبه ۴ بازشناسی شده اند . برای بازشناسی بدنه حروف از گشتاورهای شبکه زرنیکه مرتبه ۴ بعنوان ویژگی و از یک شبکه عصبی پرسپترون برای طبقه بندی استفاده شده است . میزان بازشناسی برای ۵۰۰ کلمه دستنویس که توسط ۵ خطاط معروف ایرانی نوشته شده است ۹۵٪ بوده است . دلیل بیشتر خطاها عدم امکان جداسازی حروف در بعضی از کلمات مانند « محمد » است .

فائز و میرزاخانی در [۸۳] از گشتاورهای شبه زرنیکه و شبکه عصبی پرسپترون برای بازشناسی حروف و ارقام مجزای دستنویس نیز استفاده شده است که دقت بدست آمده برای اعداد ۹۸٪ و برای حروف ۹۹٪ گزارش شده است .

## ۵-۱-۲-۲ بسط کارهونن لوئو (K-L)<sup>۱</sup>

تبدیل K-L بر اساس خواص آماری یک تصویر عمل می کند و در واقع یک آنالیز بردار ویژه است . توابع پایه این تبدیل ، مقادیر ویژه ماتریس همبستگی تصویر می باشند . این تبدیل سعی در کاهش ابعاد مجموعه ویژگیها بوسیله ایجاد ویژگیهای جدیدی که ترکیبات خطی از ویژگیهای اولیه می باشند ، دارد . تبدیل K-L حداکثر فشردگی انرژی را انجام می دهد و از لحاظ آماری تنها تبدیل بهینه محسوب می شود . از آنجائیکه توابع پایه KLT با هر دسته از تصاویر قابل تطبیق می باشند ، با تصویر کردن<sup>۲</sup> تصاویر در فضای K-L و محاسبه ضرایب KLT ، راندمان بالایی حاصل خواهد شد . بسط K-L در چندین زمینه شناسایی الگو از جمله بازشناسی چهره مورد استفاده قرار می گیرد . همچنین از این تبدیل در سیستم OCR مؤسسه ملی استانداردها و فناوری (NIST) بمنظور بازشناسی دستنوشته های فرم گونه بکار گرفته می شود . نقطه ضعف این تبدیل ، حجم محاسباتی بالای آن است و بدلیل برخورداری از الگوریتمهای محاسباتی بسیار سنگین ، استفاده از ویژگیهای K-L در سیستمهای OCR رایج نمی باشد . لیکن با افزایش قدرت محاسباتی ماشینهای محاسب ، این تبدیل طی چند سال آتی به یک ویژگی محقق پذیر مبدل خواهد گردید . به عنوان نمونه ، این تبدیل در [52] بمنظور بازشناسی حروف عربی و در [27] جهت شناسایی حروف چینی مورد استفاده واقع گردیده است .

---

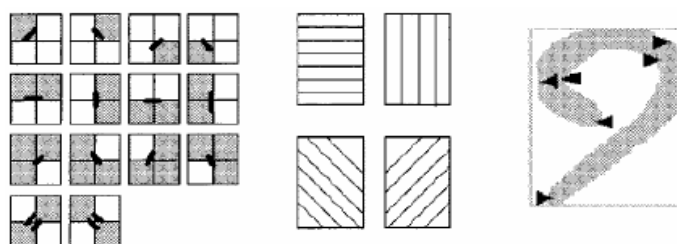
1 Karhunen-Loeve  
2 projection

## ۲-۲-۲ ویژگیهای آماری

بازنمایی یک تصویر سند بوسیله توزیع آماری آن ، تا اندازه ای تغییرات نگارشی را مورد توجه قرار می دهد . گرچه این روش بازنمایی ، بازسازی تصویر اولیه را امکان پذیر نمی سازد ، اما از آن در جهت کاهش ابعاد مجموعه ویژگیها استفاده می شود که بدین وسیله سرعت بالاتر و پیچیدگی کمتری حاصل می گردد . عمده ویژگیهای آماری مورد استفاده در بازشناسی حروف بشرح زیر می باشند :

### الف - ناحیه بندی<sup>۱</sup>

در این روش ، قاب در برگیرنده کاراکتر به چندین ناحیه همپوشان یا غیر همپوشان تقسیم می شود . چگالی نقاط یا ویژگیهایی از نواحی مختلف پس از آنالیز بمنظور بازنمایی آن کاراکتر بکار می روند به عنوان مثال ، ویژگیهای جهتی کانتور کاراکتر ، جهت کانتور کاراکتر را اندازه گیری می کند . این ویژگیها با تقسیم آرایه تصویر به نواحی مستطیلی و قطری ، و محاسبه هیستوگرامهای کدهای زنجیره ای<sup>۲</sup> در این نواحی تولید می گردند . به عنوان مثال دیگر می توان به ویژگیهای نقاط خم<sup>۳</sup> اشاره نمود که نشان دهنده نقاط دارای خمیدگی زیاد<sup>۴</sup> ، نقاط انتهایی<sup>۵</sup> و نقاط انشعاب<sup>۶</sup> می باشند . شکل (۲-۲-۲) ویژگیهای جهتی کانتور و ویژگیهای نقاط خم را نشان می دهد .



شکل (۲-۲-۲)

- 
- 1 Zoning
  - 2 Chain codes
  - 3 Bending points
  - 4 High curvature
  - 5 End points
  - 6 Fork points

## ب - تلاقیها و فواصل<sup>۱</sup>

یک ویژگی معمول ، تعداد دفعات تلاقی یک کانتور با یک قطعه خطی در یک جهت تعیین شده است . در [4] قاب دربرگیرنده کاراکتر به مجموعه ای از نواحی در جهات مختلف تقسیم بندی می شود و سپس تعداد گذرهای سیاه در هر ناحیه ، با توان دو کد می گردد . در تحقیق دیگری ([42]) مکان و تعداد دفعات گذر از پیکسلهای پس زمینه به پیش زمینه در امتداد خطوط عمودی از تصویر کلمه کد می شود . همچنین فاصله قطعه خطوط از یک مرز مشخص مانند قسمت بالایی و پایینی قاب کلمه را می توان به عنوان ویژگیهای آماری بکار برد . این ویژگیها بطور ضمنی تصریح می نمایند که یک سطح آستانه افقی در بالا ، پایین و مرکز دست نوشته نرمالیزه شده وجود دارد و تعداد دفعاتی که آن کاراکتر از آن سطح آستانه عبور می کند ، مقدار ویژگی آن کاراکتر محسوب می شود . هدف اصلی در اینجا تشخیص قسمت‌های بالارونده و پایین رونده در تصویر کاراکتر است .

## ج - هیستوگرامها ( تصویر نمایها<sup>۲</sup> )

کاراکترها را می توان بوسیله هیستوگرام مقدار سطوح خاکستری پیکسل‌هایشان بر روی خطوط در جهت‌های متفاوت ، بازنمایی کرد . این روش نمایانند کاراکترها ، یک سیگنال یک بعدی از تصویر دو بعدی ایجاد می کند که می تواند به عنوان ویژگی معرف تصویر کاراکتر مورد استفاده قرار گیرد .

## ۲-۳ ویژگیهای هندسی<sup>۳</sup> و توپولوژیکی<sup>۴</sup>

خصوصیات مختلف سراسری و محلی کاراکترها را می توان بوسیله ویژگیهای هندسی و توپولوژیکی که از تحمل پذیری بالایی نسبت به اعوجاجها و تغییرات سبک نگارشی برخوردارند ، بازنمایی نمود . این روش بازنمایی ممکن است برخی دانشها در مورد ساختار شیئی را نیز کد نماید و یا امکان دارد اطلاعاتی را فراهم نماید که بواسطه آن بتوان تشخیص داد که چه نوع از اجزاء ساختار

---

1 Crossings and Distances

2 Projections

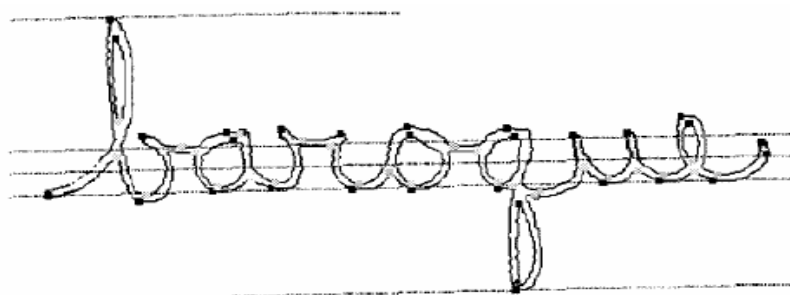
3 Geometrical

4 Topological

شیء را تشکیل می دهند . نمونه های متعدد از روشهای بازنمایی هندسی و توپولوژیکی را میتوان در گروههای زیر جای داد :

### الف - ویژگیهای توپولوژیکی

در این گروه از روشهای استخراج ویژگیها ، یک ساختار از پیش تعریف شده در حرف یا کلمه مورد جستجو قرار می گیرد . تعداد یا موقعیت نسبی این ساختارها در کاراکتر ، یک روش بازنمایی توصیفی را نتیجه می دهد . ساختارهای اولیه عمومی زیرحرفها می باشند که بدنه یک حرف را تشکیل می دهند . این ساختارهای اولیه می توانند بسادگی خطوط (1) و کمانها (C) که از اجزاء اصلی حروف لاتین می باشند و نیز به پیچیدگی منحنیها و دندانه های تشکیل دهنده حروف چینی و عربی باشند . برخی از این ویژگیها عبارتند از : نقاط حدی ( کرانه )<sup>1</sup> ، نقاط ماکزیمم و مینیمم ، رئوس<sup>2</sup> بالا و پایین یک حد آستانه ، روزنه های واقع در طرف راست (C) ، چپ (S) ، بالا (u) و پایین (n) ، نقاط تقاطع (X) ، نقاط انشعاب (T) ، انتهای خطی شکل (J) ، حلقه ها (O) ، جهت یک زیرحرف از یک نقطه خاص ، انحنای بین دو نقطه ، نقطه های مجزا ، شکستگی بین دو نقطه ، تقارن حرف ، منحنیهای افقی در بالا یا پایین ، پاره خطهای بین دو نقطه ، پاره خطهای صعودکننده ، نزول کننده و میانی ، ارتباط بین پاره خطوطی که یک زیرحرف را تشکیل می دهند و غیره . شکل (۳-۲-۲) برخی از ویژگیهای توپولوژیکی را نمایش می دهد .



شکل (۳-۲-۲)

---

1 Extreme points  
2 cusp

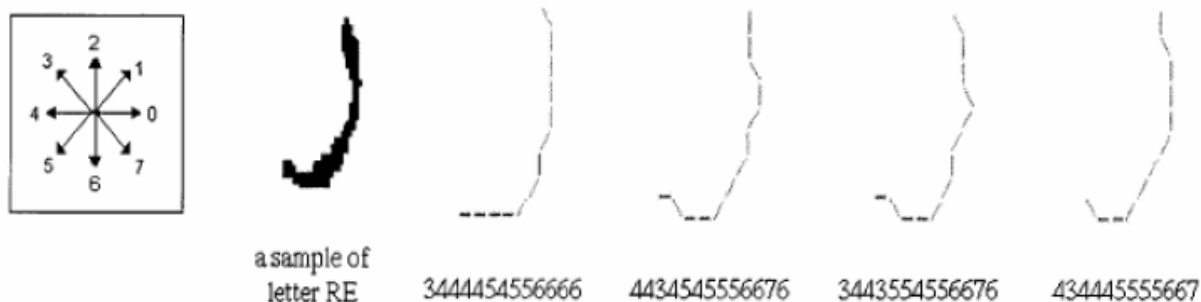
## ب - ویژگیهای هندسی

تعدادی از این ویژگیها عبارتند از : نسبت عرض و ارتفاع قاب دربردارنده کاراکتر ، فاصله نسبی بین نقطه پایانی و نقطه دارای کمترین عرض ، فواصل نسبی افقی و عمودی بین اولین و آخرین نقطه ، فاصله بین دو نقطه ، طولهای قابل تطبیق بین دو پاره خط ، پهنای یک زیرحرف ، بخشهای عمده بالا و پایی کلمات ، طول کلمات و انحناء یا تغییرات انحناء .

ویژگیهای هندسی اندازه گیری شده میتوان با یک مجموعه هندسی مناسبتر و فشرده تر از ویژگیها تقریب زد . از جمله این روشها ، تقریب چند ضلعی کاراکترهای نازک شده است . نمونه دقیقتر روش تقریب چند ضلعی ، روش بازنمایی cubic spline می باشد .

## ج - کدگذاری

از جمله معمولترین روشهای کدگذاری ، کد زنجیره ای « فریمن »<sup>۱</sup> است . این کد اساساً از طریق نگاشت زیرحرفهای یک کاراکتر به یک فضای پارامتری دو بعدی که از کدها تشکیل شده است ، بدست می آید . شکل پایین روش تشکیل این کد را برای حرف «ر» نشان می دهد . نمونه های مختلفی از کدگذاری زنجیره ای وجود دارد به عنوان مثال در [23] قاب کاراکتر با یک پنجره متحرک در جهت چپ به راست پوشانده می شود و هر ناحیه توسط کد زنجیره ای کد می گردد .



## د - گرافها و درختها

1 Freeman

در ابتدا کلمات یا حروف به مجموعه ای از قسمتهای بنیادی توپولوژیکی نظیر پاره خطها ، حلقه ها نقاط تقاطع و غیره تقسیم بندی می شوند . سپس این قسمتهای بنیادی توسط گرافهای منتسب شده<sup>۱</sup> یا پیوندی<sup>۲</sup> بازنمایی می گردند . دو نوع روش بازنمایی تصویر بوسیله گراف وجود دارد . روش از مختصات شکل حروف استفاده می کند . در روش دوم ، گره ها متناظر با زیرحرفها می باشند . همچنین می توان از درختها بمنظور بازنمایی کلمات یا حروف به همراه مجموعه ای از ویژگیها که با یکدیگر ارتباط سلسله مراتبی دارند ، بهره جست .

---

1 Attributed  
2 Relational

## ۳-۲ تعدادی از روشهای شناسایی کاراکترهای فارسی و عربی در پایان نامه های داخل و خارج

در این بخش به تعدادی از روشهای شناسایی حروف فارسی و عربی که در پایان نامه های کارشناسی ارشد و دکترای داخل کشور<sup>۱</sup> و مقالات داخلی و خارجی معرفی شده اند، اشاره می کنیم.

«عزمی» در [۸۶و۸۲] سیستمی جهت بازشناسی متون چاپی فارسی با استفاده از رویکرد ترکیبی ( بکارگیری هر دو رویکرد مبتنی بر جداسازی کلمات و مبتنی بر بازشناسی کلمه بعنوان یک الگوی واحد ) ارائه نموده است. در این روش ابتدا خطوط متن از یکدیگر جدا شده و ضخامت قلم تعیین می شود. سپس زیر کلمات با توجه به هیستوگرام عمودی تصویر متن جدا می گردند. بعد از آن نوار زمینه با محاسبه هیستوگرام افقی تعیین می شود. سپس با کمک یک الگوریتم اصلاحی، نوار زمینه برای هر زیر کلمه جداگانه تصحیح می گردد. در تعیین دقیق نوار زمینه از منحنی پیرامونی زیر کلمات استفاده می شود. نقاط جداسازی حروف، با استفاده از کانتور بالایی کلمات تعیین می شوند. نقاط جداسازی با اعمال قواعدی در قالب یک دیاگرام حالت روی منحنی پیرامونی کلمات مشخص می گردند. در یک آزمون روی چند صفحه متن چاپی شامل حدود ۱۱ هزار حرف با قلمهای مختلف حدود ۹۹ درصد از حریف بدرستی جداسازی شده اند. برای بازشناسی حروف با قلمهای گوناگون، دو الگوریتم مختلف ارائه شده است. الگوریتم اول از کدهای فریم کانتور حروف به عنوان ویژگی و از یک اتوماتون آماری برای طبقه بندی استفاده می نماید. الگوریتم دوم از تبدیل هاف در مرحله استخراج ویژگی و از یک شبکه عصبی پس از انتشار خطا در مرحله طبقه بندی استفاده می کند. الگوریتم های فوق روی مجموعه ای از حروف شامل ۱۱۵۰۰ نمونه از دو قلم مختلف آزموده شده اند و میزان بازشناسی به ترتیب ۹۷/۱۳ و ۹۸/۳۲ درصد بوده

---

۱- بیشتر این پایان نامه ها از طریق وب سایت مرکز اطلاعات و مدارک علمی ایران (<http://www.irandoc.ac.ir>) در دسترس قرار گرفته اند.



است. در این تحقیق برای بکارگیری شکل کلی زیر کلمات در بازشناسی آنها، از دیکشنری تصویری استفاده شده است. به این منظور، سه دیکشنری مختلف با استفاده از ویژگیهای مکان مشخصه، توصیف کننده های فوری و برجسب های کانتور بالایی طراحی شده و آزمایش های مختلفی روی آنها انجام گرفته است. در مرحله نهایی این تحقیق سیستمی برای بازشناسی کلمات فارسی با یک رویکرد ترکیبی ارائه شده است. در این سیستم با جداسازی و بازشناسی حروف شاخص زیر کلمه، رویکرد مبتنی بر جداسازی و با در نظر گرفتن یک دیکشنری تصویری برای بدنه زیرکلمات، رویکرد بازشناسی کلمه بعنوان یک الگوی واحد در نظر گرفته شده است. نوع و موقعیت نقاط و علائم نیز در بازشناسی زیر کلمات بکار گرفته شده اند. اطلاعات آماری درباره احتمال رخداد متوالی حروف شاخص و امتیاز حاصل از بازشناسی آنها به کمک مدل مخفی مارکوف (HMM) و یک الگوریتم ویتربی<sup>۱</sup> تغییر یافته برای بازشناسی زیر کلمات بکار گرفته شده اند. سیستم ترکیبی نهایی روی یک متن با چهار قلم مختلف آزموده شده و نتایج با جزئیات مختلف ارائه شده است.

«**مسروری**» در [۸۸] روشهایی برای شناسایی کلمات دستنویس فارسی در یک مجموعه محدود بدون قطعه بندی کلمات ارائه نموده است. این روشها در سطح کلمات یا زیر کلمات عمل می کنند. همچنین روشی برای تعمیم الگوریتم DTW به دو بعد معرفی شده و برای مقایسه کلمات بکار رفته است. برای آزمودن این روشها، از یک مجموعه محدود از نام ۵۰۰ شهر پر جمعیت ایران استفاده شده است. نتایج این پایان نامه نشان می دهد که از بین روشهای ارائه شده، روش DTW دو بعدی، بیشترین میزان بازشناسی را دارد. در آزمایش این روش برای مجموعه نامهای ۱۰۰، ۳۰۰ و ۵۰۰ شهر ایران، حدود ۸۸ درصد از کلمات به درستی شناسایی شده اند.

«**مسروری**» در [۸۹] روشی برای بازشناسی حروف مجزای دستنویس فارسی ارائه نموده است. وی ۳۳ کلاس برای حروف فارسی در نظر گرفته است. بازشناسی حروف در دو مرحله انجام

---

1 Viterbi

می گیرد. در مرحله اول به کمک یک طبقه بندی کننده فازی « قاعده - پایه » که قوانین آن به وسیله مجموعه نمونه های آموزشی یاد گرفته می شوند، حروف به ۸ گروه تقسیم می گردند. ویژگی های مورد استفاده در این مرحله با روش مکانهای مشخصه در راستاهای ۹۰ و ۴۵ درجه از تصویر باینری حرف بدست می آید. در مرحله دوم، به وسیله طبقه بندی کننده های درخت تصمیم که برای هر گروه به طور خاص طراحی شده اند، بازشناسی نهایی حروف در هر گروه انجام می گیرد. در این مرحله، ویژگی های به کار رفته عموماً ساختاری هستند. مجموعه نمونه های جمع آوری شده شامل ۴۹۵۰ نمونه ( ۱۵۰ نمونه از هر حرف ) که توسط نویسندگان های مختلف با سطوح تحصیلاتی متفاوت نوشته شده اند. ۳۳۰۰ نمونه از این مجموعه ( ۱۰۰ نمونه از هر حرف ) برای یادگیری و ۱۶۵۰ نمونه دیگر برای آزمایش به کار برده شده است که طبق نتایج گزارش شده، ۹۲ درصد از این نمونه های آزمایش بدرستی بازشناسی شده اند.

« مقسمی » در [۹۱] یک روش ساختاری جهت شناسایی حروف تایپی فارسی معرفی نموده است. کد زنجیره ای که جهت حرکت قلم فرضی هنگام نوشتن حروف را مشخص می سازد، به عنوان ویژگی در نظر گرفته شده است. کد زنجیره ای از روی محیط مرزی حروف نازک شده بدست می آید. مسائل مهم در این روش عبارتند از جداسازی حروف و کاهش طور کد زنجیره ای جهت جداسازی از روش مجموع فواصل نقاط تا خط پایه به همراه بررسی حرکت افقی روی خط مبنا استفاده می شود. جهت کاهش طول بردار ویژگی، روشی تحت عنوان روش سیستم خبره به کار گرفته شده است. الگوریتم مستقل از اندازه حروف است. همچنین این روش را بدون هیچگونه تغییری می توان جهت شناسایی حروف تایپی انگلیسی، ارقام دستنویس فارسی و انگلیسی به کار برد. جهت طبقه بندی نهایی از ترکیب بانک اطلاعاتی و شبکه عصبی استفاده شده است.

« اسدی » در [۶۳] یک روش بازشناسی حروف تایپی فارسی بصورت ساختاری و با پارامتری فرض کردن حروف پیشنهاد نموده است. در مرحله استخراج ویژگیها از ویژگیهای ساختاری درصد نقاط سیاه بالای خط زمینه، تعداد حفره های حروف، تعداد قطعات تشکیل دهنده حرف، نوع

اتصال حرف به حروف راست و چپ و غیره استفاده گردیده است. این ویژگیها بگونه ای مورد استفاده قرار می گیرند که بتوانند نماینگر حروف مختلف بصورت یکتا باشند. با استفاده از جداول و درخت های تصمیم گیری، روشی جهت بازشناسی حروف تاییبی فارسی ارائه می گردد. مسائل مهم در این پایان نامه عبارتند از: جداسازی حروف که بر اساس موقعیت خط زمینه صورت می پذیرد، استخراج ویژگیهای مناسب و بازشناسی حروف فارسی که با استفاده از بانک های اطلاعاتی و درخت های تصمیم گیری انجام می گیرد. روش دیگر بازشناسی که در این تحقیق مورد استفاده قرار گرفته است، استفاده از Padalines شبکه عصبی چند جمله ایهای جداکننده است. در این روش تعدادی از ویژگیهای استخراج شده تبدیل به ثابت هایی برای چند جمله ایهای مختلف می-گردد و برای حالات مختلف حروف یک یا چند جمله ای ارائه می گردد. الگوریتم ها مستقل از اندازه حروف می باشند. روشهای ارائه شده در این پایان نامه بدون هیچگونه تغییری می تواند برای بازشناسی حروف تاییبی لاتین نیز بکار رود. برای طبقه بندی نهایی در هر دو روش بازشناسی نیز از بانک های اطلاعاتی استفاده شده است.

« دهقانی » [۷۲] از الگوریتمی مبتنی بر مدل پنهان مارکوف (HMM) در بازشناسی متون فارسی بهره برده است. در مرحله پیش پردازش جهت فائق آمدن بر نویزهای موجود در تصاویر ورودی، چندین عمل پیش پردازشی از قبیل فیلتر کردن میانه ای، باینری کردن، اعمال فیلترهای موفولوژیکی، نرمالیزه سازی و متمرکز کردن انجام شده است. در مرحله استخراج مشخصه ها، چهار نوع بردار مشخصه بر اساس روش RPCT از هر کاراکتر استخراج می شود. سیستم پیشنهادی از دو قسمت تشکیل شده است. برای هر کاراکتر در مرحله آموزش چند مدل مارکوف متناظر با بردارهای مشخصه مختلف ساخته می شود. در مرحله بازشناسی، نتیجه تشخیص هر کدام از طبقه بندی کننده ها جهت تشکیل نتیجه نهایی بازشناسی، نتیجه تشخیص هر کدام از طبقه بندی کننده ها جهت تشکیل نتیجه نهایی بازشناسی با یکدیگر ترکیب می شوند.

در این تحقیق نرخ بازشناسی ۹۳ درصد برای داده های آموزش و ۷۱ درصد برای داده های آزمایش گزارش شده است .

« بحری » [۶۵] از روش عصبی فازی برای شناسایی حروف دستنویس جداگانه استفاده نموده است . ابتدا الگوریتم GLVQF که یک الگوریتم خودساز مانده می باشد ، بکار گرفته شده است تا الگوی وزنها ( که بطور نظری برابر الگوی ورودیها می باشد ) بدست می آید و سپس واریانس این ورودیها نسبت به این وزنها محاسبه شده و توأم با الگوی وزنها مناطق تعلق فازی به طور مثلثی و دوزنقه ای را بدست می دهد . این مناطق در نورونهای فازی لایه دوم شبکه به خاطر سپرده شده و رویهم مقدار تعلق ورودیهای مورد شناسایی نسبت به دسته وزنها را مشخص می سازند . سپس این اطلاعات بدست آمده توسط لایه سوم که یک لایه پرسپترون می باشند فازی زدائی می گردند . در این لایه فازی زدائی بهینه شده است که با وجود تعداد قوانین کم فازی نتایج قابل قبولی را بدست می دهد .

« ثانی » [۶۹] از یک روش مبتنی بر شبکه عصبی بمنظور بازشناسی حروف دست نویس فارسی استفاده نموده است . ورودی سیستم فرمهایی است که شامل جداول مختلفی می باشند و حروف به صورت دست نویس در این جداول وارد شده اند . ابزار تهیه فرم بصورت دلخواه نیز عنوان پیش نیاز این پروژه پیاده سازی شده است . ورودی شبکه BPN الگوی بیتی انتخابی از تصویر حروف مورد نظر می باشد و خروجی سیستم یکی از ۳۲ کلاس الفبای فارسی است . سیستم پیشنهادی قادر است نزدیکی الگوی ورودی به هر یک از کلاسها را بصورت احتمالات بیان کند .

« صدوقی یزدی » در [۷۷] به مبحث دو سطحی کردن تصویر و جداسازی تصویری کلمات چاپی فارسی و لاتین پرداخته است . در بخشی از این تحقیق ، پس از بررسی و آزمایش روشهای مختلف دو سطحی کردن تصاویر ، الگوریتمی ارائه شده است که روش خوشه یابی با دو میانگین و روش محلی مورین را ترکیب می کند . این الگوریتم در مقابل نایکنواختی روشنایی و وجود لکه های موضعی در تصویر مقاوم است . بخش دیگر این تحقیق درباره جداسازی تصویری کلمات

فارسی و لاتین تمایز خوبی ایجاد می کند ، انتخاب شده اند . برای جداسازی کلمات ، یک طبقه بندی کننده حداقل فاصله از متوسط هر کلاس بکار رفته است . پارامترهای الگوریتم طبقه بندی با استفاده از یک مجموعه تمرین شامل حدود ۲۰۰۰ کلمه فارسی و لاتین محاسبه شده است . در یک آزمایش برای یک مجموعه آزمون شامل حدود ۲۰۰۰ کلمه دیگر ۹۵ درصد از کلمات به درستی طبقه بندی شدند .

« شاه حسینی » [۷۵] روشی برای شناسایی حروف دستنویس فارسی با استفاده از شبکه های عصبی معرفی نموده است . شبکه عصبی مورد استفاده ، یک طبقه بندی کننده سه لایه پرسپترون با قانون یادگیری پس انتشاری می باشد . روشهای مناسبی برای بالا بردن کارایی الگوریتم آموزش شبکه ارائه شده است . سه نوع گشتاور هندسی به عنوان ویژگی مورد استفاده قرار گرفته اند . مرحله عمده الگوریتم بازشناسی ، گروه بندی اولیه حروف در ۱۷ گروه است. گروههای هفده گانه بر اساس نتایج حاصل از طبقه بندی حروف توسط یک شبکه عصبی بر روی ۳۳ کلاس تعریف شده اند . هر گروه شامل یک ، دو ، سه یا چهار حرف است که با توجه به ویژگیهای به کار رفته در این طبقه بندی با یکدیگر اشتباه می شوند . برای گروه بندی حروف شبکه ای با ۲۲ گره ورودی به تعداد ویژگیهای گشتاوری ، ۳۵ گره لایه پنهان و ۱۷ گره خارجی بکار گرفته شده است . در مرحله دیگری از الگوریتم بازشناسی حروف هر گروه با شبکه هایی با ابعاد کوچکتر شناسایی می شوند . مجموعه نمونه حروف در این تحقیق شامل ۱۶۵۰ نمونه آموزش و ۱۵۶۵ نمونه آزمون است . از این نمونه ها با یک روبشگر ( اسکن ) دستی با دقت ۲۰۰ نقطه در اینچ تصویربرداری شده است . ۹۷/۸۲ درصد از نمونه های آموزش و ۹۰/۳۵ درصد از نمونه های آزمون به درستی دسته بندی شده اند . متوسط میزان شناسایی حروف در گروههای هفده گانه ۸۹/۴۸ درصد می باشد .

« تیمساری » [۶۸] روشی برای شناسایی حروف تشکیل دهنده کلمات در متون تایپ شده فارسی ارائه نموده است . روش پیشنهادی بر اساس اعمال عملگرهای مورفولوژیکی بر روی تصویر باینری کلمه و بازشناسی حروف بر مبنای تصمیم گیری روی این ویژگیها پایه ریزی شده است . در

این روش ، اعمال جداسازی و تشخیص حروف تشکیل دهنده یک کلمه بطور همزمان صورت می گیرند . روند اجرای الگوریتم بدین صورت است که ابتدا با محاسبه پروجکشن ( projection ) افقی و سپس قائم ، متن به خطوط و خطوط به زیر کلمات تفکیک می شوند . سپس نقطه های موجود در حروف سازنده زیر کلمه استخراج شده ، متعاقب آن عمل نازکسازی روی تصویر باقیمانده انجام می پذیرد . برای انجام این کار از الگوریتمی سریع بر مبنای عملگر مورفولوژیکی Hit/Miss استفاده شده است . استخراج ویژگیهای لازم برای شناسایی حروف ، توسط اعمال عملگر Hit/Miss با المانهای ساختاری مناسب روی اسکلت تصویر صورت می گیرد .تصمیم گیری نهایی بر روی ویژگیهای ساختاری بدست آمده و تشخیص حروف ، بر اساس یکسری قواعد از پیش تعیین شده انجام می گیرد .

« نمازی » [۹۴] الگوریتمی بر مبنای منطق فازی و ترکیب آن بال شبکه عصبی بمنظور بازشناسی حروف چاپی فارسی با قلمهای متفاوت نموده است . در این سیستم تصاویر توسط اسکنر و با دقت ۳۰۰ نقطه بر اینچ وارد کامپیوتر می شوند و سپس نازک می شوند . در این روش از گشتاورهای شبه زرنیکه به عنوان ویژگی اصلی تصویر ورودی استفاده شده است . پس از فازی کردن این گشتاورها توسط توابع عضویت مربوطه ، چند شبکه عصبی با توپولوژی های متفاوت برای طبقه بندی حروف آموزش داده شده اند و در آخر ، طبقه بندی کننده نهایی با استفاده از این شبکه ها اقدام به تشخیص الگوهای ورودی می نماید . آموزش و تست این سیستم با استفاده از حدود ۳۷۰۰ نمونه چاپ شده انجام شده است که نسبت نمونه های آموزشی بیش از ۳ به یک می باشد . دقت بدست آمده در این سیستم بیش از ۹۹/۸ درصد می باشد . پراکندگی دقت سیستم در کلاسهای مختلف حدود ۲ درصد است . همچنین در این تحقیق نشان داده شده است که در شرایط مساوی ، قدرت تفکیک الگوهای ورودی در محیطهای مبهم برای شبکه فازی ، بالاتر از شبکه عادی می باشد .

«اردشیر بهرستانی» در [۶۲] سعی در بازشناسی حروف مجزای دستنویس فارسی که توسط افراد مختلف نوشته شده اند ، داشته است . فرض بر این بوده که حروف با دقت ۲۰۰ نقره در اینج تصویربرداری و با کد MSP ذخیره شده اند . در این تحقیق ابتدا یک مرحله پیش پردازش برای رفع بریدگی و حفره و نویز صورت می گیرد . سپس با اجرای الگوریتم برچسب زنی مؤلفه ها ، نقطه و سرکش موجود ، از بدنه اصلی حرف جدا و نوع آن تشخیص داده می شود . در نهایت با گروه بندی بخشهای سطرهای مختلف ماتریس تصویر حرف ، زیر حروف استخراج و مورد شناسایی قرار می گیرند . برای طبقه بندی حروف از یک درخت تصمیم استفاده گردیده است . برای یک مجموعه متشکل از ۶۱ نمونه از هر حرف ، میزان بازشناسی ۹۱/۵ درصد و برای مجموعه ای دیگر متشکل از ۴۰ نمونه از هر حرف ، بازشناسی صحیح ۹۲/۱۲ درصد گزارش گردیده است .

«مقدم تبریزی» [۹۰] روشهایی برای شناسایی و استخراج سمبل های مربوط به نقشه های جغرافیایی و یا فنی از متن نقشه ، و یا بطور کلی استخراج یک شکل و یا سمبل از تصویری که ممکن است حاوی خطوط ، منحنی ها و یا نوشته باشد ، پیشنهاد نموده است . هدف از این تحقیق ، یعنی در مکانیزه کردن روند تبدیل نقشه های دستی و تصویری به نقشه های کامپیوتری برای سیستمهای GIS می باشد . برای نیل به این منظور ، از سه روش مبتنی بر شبکه عصبی هاپفیلد ، شبکه عصبی پرسپترون چند لایه و الگوریتم ژنتیک استفاده شده است . در روش اول کاهش انرژی شبکه منجر به رفع اغتشاشات شده و بدین طریق سمبل و سپس تمامی سمبل شناسایی می گردد . در روش سوم ، معرفی تابع هدفی خاص و بهینه سازی آن با روش الگوریتم ژنتیک اساس شناسایی تصویر است . نتایج حاصل از این تحقیق نشان می دهد که روش مبتنی بر شبکه عصبی پرسپترون چندلایه ، بر دو روش دیگر برتری دارد .

«عباسیان» در [۸۱] الگوریتمی برای بازشناسی بر خط حروف فارسی ارائه نموده است . این الگوریتم ۶۱ حرف را بر اساس ویژگیهای پویا و ایستای آنها و با استفاده از یک درخت تصمیم باینری که به شکل تجربی طراحی شده است ، بازشناسی می کند . در این الگوریتم ابتدا با

استخراج ویژگی‌های زمانی و مکانی لازم، مجموعه‌ای از الگوهای پایه در ساختار حرف‌بازشناسی شده‌اند. این الگوهای پایه عبارتند از: پاره‌خطهای افقی، عمودی و مورب، منحنی بسته، زیر حرف S مانند و کمانهایی یا گودی رو به چپ یا راست. در الگوریتم‌بازشناسی نهایی، ابتدا حروف بر اساس تعداد بخشهای آنها به چهار گروه اولیه تقسیم می‌شوند و سپس حروف هر گروه با استفاده از یک درخت تصمیم‌گیری شده است.  $93/7$  درصد از نویسه‌های یک مجموعه آزمایش شامل ۴۵ نمونه از هر حرف بدرستی شناخته شده و  $1/9$  درصد از آنها وازده شده‌اند. در یک آزمایش دیگر برای حروف مجزای دستنویس بر روی مجموعه‌ای شامل ۵۰ نمونه از هر حرف،  $94/7$  درصد از حروف شناخته شده و  $2/36$  درصد وازده شده‌اند. برای بازشناسی ارقام، درخت تصمیم‌گیری طراحی شده است که برای تمرین آن، از یک مجموعه با ۱۰ نمونه از هر رقم استفاده شده است. با این الگوریتم  $94/8$  درصد از ارقام یک مجموعه آزمایش شامل ۲۵ نمونه از هر رقم، به درستی شناخته شده و  $0/8$  درصد از آنها وازده شده‌اند.

«رضوی» در [۷۳] روشی برای خواندن اتوماتیک فرمهای انتخاب درس ارائه نموده است. در هر فرم، حروف و ارقام بصورت جداگانه در داخل مستطیلهایی نوشته می‌شوند. یک الگوریتم بازشناسی حروف و ارقام که از روش N-tuple و اطلاعات پیش زمینه‌ای در سطح حرف (طول کد حرفی و موقعیت مکانی هر کاراکتر) استفاده می‌کند، طراحی شده است. برای تشخیص نهایی کد درس از روش جستجوی واژه نامه استفاده می‌شود. اثر تعداد N-tuple ها و چگونگی انتخاب آنها و تعداد نمونه‌های تمرین در میزان بازشناسی مورد بررسی شده است. در بهترین حالت (انتخاب ۸۵-3 tuple ناهمپوشان) ۱۰۰ نمونه تمرین در هر کلاس و استفاده از روش N-tuple اصلاح شده برای نمونه آزمایش در هر کلاس،  $9/80$  درصد از حروف و ۸۱ درصد از ارقام به درستی بازشناسی شدند. برای آزمایش، ۱۳۴ فرم تکمیل شده، شامل ۴۵۲ درس انتخابی، از ۶۷ دانشجو جمع‌آوری شد و از این فرمها با درجه تفکیک ۲۰۰ نقطه در اینچ، به صورت باینری تصویربرداری و در کاراکتر در ابعاد  $16 \times 16$  نرمالیزه شد. در الگوریتم نهایی خواندن فرم، میزان



بازشناسی درست برای کد درس ۹۵/۱ درصد ، میزان خطا ۰/۸۸ درصد و میزان وازدگی ۳/۹۸ درصد است .

«رفیعی کراچی» در [۷۴] یک روش جداسازی حروف برای متون تایپ شده فارسی پیشنهاد نموده است . این روش جداسازی مستقل از رسم الخطها می باشد و بر اساس پروفیل بالایی تصویر زیر کلمه انجام می گیرد . الگوریتم جداسازی حروف شامل سه مرحله پردازش اولیه ، پیدا کردن محل اتصال حروف و پس پردازشها می باشد . پیش پردازشها در بر گیرنده جداسازی خطوط متن ، محاسبه پهنای قلم ، پیدا کردن نوار زمینه و جداسازی زیر کلمات می باشند . در مرحله دوم ابتدا پروفیل بالایی تصویر زیر کلمه بدست می آید . در این مرحله بعضی از حروف به زیر حروف می شکنند . از پس پردازشها برای بازسازی این حروف استفاده می شود . پس پردازشها شامل قواعدی مبتنی بر رفتار بینایی انسان در جداسازی حروف و قواعد نگارشی حروف فارسی می باشند . متون نمونه بکار رفته در این تحقیق شامل ۱۳۴۴۲ حرف در ۲۶ رسم الخط مختلف هستند . این متون با دقت ۲۰۰ نقطه در اینچ تصویربرداری شده اند . در این مجموعه که ۶۷/۳۱ درصد از حروف به یکدیگر متصل هستند که ۹۹/۱۸ درصد از آنها بدرستی از یکدیگر جدا شده اند .

«امیری» [۶۴] روشی برای جداسازی متن چاپی فارسی از تصاویر گرافیکی ارائه نموده است . جدولها ، روندناها<sup>۱</sup> و نقشه ها از جمله تصویرهای گرافیکی هستند . این روش از سه مرحله تشکیل شده است . در مرحله اول عناصر گرافیکی شاخص با توجه به برخی ویژگیهای تصویری آنها جدا می شوند . در مرحله دوم متنهای چسبیده به عناصر گرافیکی آشکار و جدا می گردند . در مرحله سوم ، با استفاده از طبقه بندی حداقل فاصل ، عناصر گرافیکی که در تصویر متن خروجی مرحله دوم باقی مانده اند ، جدا می شوند . در این تحقیق از یک مجموعه تمرین شامل ۴۲ تصویر و یک مجموعه آزمایش شامل ۳۰ تصویر استفاده شده است . برای مجموعه آزمایش ۹۱/۱ درصد از عناصر

متن و ۸۷/۷ درصد از عناصر گرافیک درست طبقه بندی شده اند . منابع خطا و میزان تأثیر هر یک از آنها بررسی شده اند و پیشنهادهایی برای کاهش آنها ارائه شده است .

« مرتضی پور » [۸۷] الگوریتمی برای قطعه بندی برخط کلمات دستنویس فارسی ارائه نموده است . منظور از قطعه بندی در اینجا یافتن قطعاتی است که بوسیله آنها بتوان کلیه کلمات را ایجاد کرد ؛ بدین معنی که مجموعه کلمات نوشته شده توسط افراد مختلف به اجزاء ساده تری تقسیم می شوند تا برای بازشناسی مورد استفاده قرار گیرند. این الگوریتم بر اساس یافتن نقاط اصلی در یک کلمه و مقایسه شیب و زوایای پاره خطهای متصل کننده این نقاط به یکدیگر ، قطعه بندی را انجام می دهد . برای جمع آوری کلمات دستنویس از صفحه رقومی کننده با قدرت تفکیک ۱۲۵ نقطه در هر اینچ و سرعت نمونه برداری ۳۳۳ نمونه در ثانیه استفاده گردیده است . در طراحی الگوریتم از یک مجموعه تمرین با ۲۰ کلمه و از هر کلمه ۲ نمونه دستنویس افراد مختلف استفاده شده است . در آزمایش این الگوریتم ، از یک مجموعه آزمایش با ۲۰ کلمه و از هر کلمه ۲ نمونه از همان افراد استفاده شده است . در این آزمایش برای ۳ کلمه ، قطعه بندی درست انجام نشد .

« نظام آبادی پور » در [۹۳] با اصلاح الگوریتم « عزمی » [۸۲] که مبتنی بر کانتور بالایی است ، یک الگوریتم جداسازی برای متون چاپی قدیمی ارائه نموده است . برای حل مشکل نایکخواختی کرسی خط ، روشی برای تعیین نوار زمینه پیشنهاد گردیده و با اصلاح روش برچسب زنی کانتور بالایی و تکمیل قواعد جداسازی ، دقت الگوریتم برای متون قدیمی افزایش یافته است . نرخ جداسازی درست حروف با استفاده از این الگوریتم برای یک مجموعه آزمایش که از متون قدیمی برگزیده شده است ، بدون هیچ پس پردازشی ۹۷ درصد گزارش شده است . برای جداسازی نقاط چسبیده به بدنه حروف نیز روشی ارائه شده است .

« نحوی » در [۹۲] برای بعضی از مراحل اتوماسیون پست و بازشناسی آدرسهای پستی فارسی روشهایی پیشنهاد نموده است . با توجه به تفاوتهای ساختاری آدرسهای دستنویس و چاپی ، سیستم های جداگانه ای برای پردازش آنها لازم منظور شده است . در این تحقیق با استفاده از

ویژگیهای مناسب و روش حداقل فاصله از متوسط هر کلاس آدرسهای دستنویس از چاپی جدا و سپس بر اساس مجموعه ای از ویژگیهای مناسب و تصمیم گیریهای سلسله مراتبی ، کد پستی از آدرس های چاپی و دستنویس استخراج می شود . ۵۰ آدرس دستنویس و ۵۰ آدرس چاپی با ۱۲ قلم گوناگون به عنوان مجموعه تمرین استفاده شده است . مجموعه آزمایش شامل ۱۵۰ آدرس دستنویس و ۱۵۰ آدرس چاپی متفاوت از مجموعه تمرین است که با درجه تفکیک ۲۰۰ نقطه در اینچ و به شکل باینری تصویربرداری شده اند . در تفکیک آدرسهای چاپی و دستنویس آزمایش نشان می دهد که ۹۸ درصد از آدرسهای مجموعه تمرین و ۶۷/۹۷ درصد از آدرسهای مجموعه آزمایش به درستی طبقه بندی شده اند . برای مجموعه تمرین ، در ۹۸ درصد از آدرسهای چاپی و ۸۲ درصد از آدرسهای دستنویس ، کد پستی به درستی استخراج شده اند . در مجموعه آزمایش ، در ۹۶/۸۸ درصد از آدرسهای چاپی و ۸۷/۶۷ درصد از آدرسهای دستنویس کد پستی به صورت صحیح استخراج شده است .

« بنی اسدی » در [۶۶] تشخیص حروف دست نویس فارسی به وسیله سیستم ها بپرسید نور و فازی را مورد توجه قرار داده است . نقش منطق فازی در این پایان نامه مدیریت عدم قطعیت موجود در اطلاعات ورودی است . در واقع مرحله استخراج خصوصیات در این پایان نامه به کمک منطق فازی و بر اساس محاسبه درجه عضویت هر حرف در هفت مجموعه فازی مختلف انجام میگردد . در مرحله طبقه بندی نهایی شبکه های عصبی استفاده شده است . شبکه عصبی مورد استفاده شبکه پیش خور چند لایه (MLP) میباشد که به صورت ۴لایه شبیه شده است .

Mowiaei ، Faez و Haghigat [44] روشی به منظور بازشناسی کد های پستی و اسامی

شهر های ایران که بر روی بسته های پستی به صورت دستنویس نوشته می شوند،

پیشنهاد نموده اند .

برد های ویژگی این روش، از طریق بکارگیری تبدیل موجک گسته (DWT) با موجک اساسی Haar محاسبه می گردند . در مرحله پیش پردازش این الگوریتم ، ابتدا محدوده هر حرف مشخص

گردید، تصویر حرف به ابعاد  $64 \times 64$  پیکسل تغییر اندازه می یابد و در نهایت عمل نازک سازی انجام میگردد. در مرحله استخراج ویژگیها، تبدیل موجک سه سطحی بر روی تصویر نازک شده اعمال میشود و ضرایب باند LL از سطح سوم استخراج می گردند. در انتها ماتریس ضرایب باند LL به صورت برداری در آمده، یک فیلتر پایین گذر به منظور هموار سازی بیشتر ضرایب، به آن اعمال می شود. برای آموزش سیستم، یک شبکه عصبی MLP با قانون پس انتشار خطا مورد استفاده قرار می گیرد. راندمان این روش برای حروف و ارقام دست نویس فارسی بترتیب ۹۲/۳۳ و ۹۱/۸۱ درصد گزارش شده است.

Korsheed و Clocksin [29] الگوریتمی برای بازشناسی کلمات عربی بدون تقطیع آنها<sup>۱</sup> طراحی نموده اند که در آن ابتدا هر کلمه به یک تصویر قطبی نرمالیزه تبدیل می شوند و سپس یک تبدیل فوری دو بعدی به تصویر قطبی اعمال می گردد. طیف منتج شده نیست به دوران، تغییر اندازه و تغییر مکان کلمات، غیر حساس می باشد. مرحله بازشناسی با بکارگیری تابع فاصله اقلیدسی نرمالیزه شده انجام می گیرد. جهت آزمایش این الگوریتم، بیش از ۱۷۰۰ کلمه در ۴ فونت مختلف مورد استفاده قرار می گیرد. کلمات در جهتهای مختلف (۰ تا ۱۸۰ درجه) و اندازه های متفاوت (۱۸ تا ۴۸pt) در نظر گرفته می شوند. نرخ خطا برای این سیستم در بهترین حالت (با بکارگیری ۱۶ باند فرکانسی پایینی) در حدود ۶ درصد برآورد شده است.

Menhaj و adab [41] الگوریتمی بمنظور تقطیع و بازشناسی حروف فارسی در فونتها و اندازه های متفاوت با استفاده از شبکه عصبی چند لایه با تغذیه رو به جلو<sup>۲</sup> پیشنهاد نموده اند. فیدبک اتوماتیک بین مراحل تقطیع و بازشناسی بمیزان زیادی، خطاهای ناشی از تقطیع را حذف می نماید. مرحله بازشناسی، از شبکه های عصبی MLP با توصیف کننده های فوری به بهره می گیرد. راندمان این روش برای متون copy/paste شده ۱۰۰ درصد و برای متون اسکن شده با درجه تفکیک dpi ۴۰۰، ۹۳ درصد گزارش شده است.

---

1 Segmentation free  
2 Feed forward

## فصل سوم تکنیک آنالیز اجزای اصلی

### ۳-۱ روشهای کاهش ابعاد

پیشرفتهای بوجود آمده در جمع آوری داده و قابلیت‌های ذخیره سازی در طی دهه‌های اخیر باعث شده در بسیاری از علوم با حجم بزرگی از اطلاعات روبرو شویم. محققان در زمینه‌های مختلف مانند مهندسی، ستاره شناسی، زیست شناسی و اقتصاد هر روز با مشاهدات بیشتر و بیشتری روبرو می‌شوند. در مقایسه با بسترهای داده‌ای قدیمی و کوچکتر، بسترهای داده‌ای امروزی چالشهای جدیدی در تحلیل داده‌ها بوجود آورده‌اند. روشهای آماری سنتی به دو دلیل امروزه کارائی خود را از دست داده‌اند. علت اول افزایش تعداد مشاهدات<sup>۱</sup> است، و علت دوم که از اهمیت بالاتری برخوردار است افزایش تعداد متغیرهای مربوط به یک مشاهده می‌باشد.

تعداد متغیرهایی که برای هر مشاهده باید اندازه گیری شود ابعاد داده نامیده می‌شود. عبارت "متغیر"<sup>۲</sup> بیشتر در آمار استفاده می‌شود در حالی که در علوم کامپیوتر و یادگیری ماشین بیشتر از عبارات "ویژگی"<sup>۳</sup> و یا "صفت"<sup>۴</sup> استفاده می‌گردد.

بسترهای داده‌ای که دارای ابعاد زیادی هستند علیرغم فرصتهایی که به وجود می‌آورند، چالشهای محاسباتی زیادی را ایجاد می‌کنند. یکی از مشکلات داده‌های با ابعاد زیاد اینست که در بیشتر مواقع تمام ویژگیهای داده‌ها برای یافتن دانشی که در داده‌ها نهفته است مهم و حیاتی نیستند. به همین دلیل در بسیاری از زمینه‌ها کاهش ابعاد داده یکی از مباحث قابل توجه باقی مانده است.

---

1 observations  
2 variable  
3 feature  
4 attribute

روشهای کاهش ابعاد داده به دو دسته تقسیم می‌شوند:

- **روشهای مبتنی بر استخراج ویژگی:** این روشها یک فضای چند بعدی را به یک فضای

با ابعاد کمتر نگاشت می‌کنند. در واقع با ترکیب مقادیر ویژگیهای موجود، تعداد کمتری ویژگی بوجود می‌آورند بطوریکه این ویژگیها دارای تمام (یا بخش اعظمی از) اطلاعات موجود در ویژگیهای اولیه باشند. این روشها به دو دسته‌ی خطی و غیر خطی تقسیم می‌شوند که موضوع مورد بحث در همین حوزه قرار می‌گیرد.

- **روشهای مبتنی بر انتخاب ویژگی:** این روشها سعی می‌کنند با انتخاب زیرمجموعه‌ای از

ویژگیهای اولیه، ابعاد داده‌ها را کاهش دهند.

در این فصل کمتر به اثباتهای ریاضی پرداخته شده و بیشتر به مفاهیم و کاربرد روشها توجه شده است. در تهیه‌ی مطالب این فصل سعی شده است با ارائه‌ی مثالهای مناسب، خواننده در درک بهتر مفاهیم مربوطه یاری شود.

### ۳-۱-۱ روشهای مبتنی بر استخراج ویژگی

همانطور که اشاره شد روشهای مبتنی بر استخراج ویژگی، یک فضای چند بعدی را به یک فضای با ابعاد کمتر نگاشت می‌دهند. این روشها به دو دسته‌ی خطی و غیرخطی تقسیم می‌شوند. روشهای خطی که ساده‌ترند و فهم آنها راحت‌تر است بدنبال یافتن یک زیرفضای تخت عمومی<sup>۱</sup> هستند. اما روشهای غیرخطی که مشکل‌ترند و تحلیل آنها سخت‌تر است بدنبال یافتن یک زیرفضای تخت محلی<sup>۲</sup> می‌باشند.

از روشهای خطی می‌توان به DFT، DWT، PCA و FA اشاره کرد

از روشهای غیرخطی نیز می‌توان به موارد زیر اشاره کرد:

Principal Curves

Self Organizing Maps

Vector Quantization

Genetic and Evolutionary Algorithms

Regression

مسئله‌ی کاهش ابعاد داده را بطور ریاضی می‌توان به اینصورت بیان کرد: یک متغیر تصادفی  $p$ -بعدی  $\mathbf{x} = (x_1, \dots, x_p)^T$  داریم. می‌خواهیم متغیر  $k$ -بعدی  $\mathbf{s} = (s_1, \dots, s_k)^T$  را به گونه‌ای پیدا کنیم که اولاً  $k \leq p$  باشد و ثانیاً  $\mathbf{s}$  محتویاتی که در  $\mathbf{x}$  وجود دارد را بر اساس معیاری خاص دارا باشد. روشهای خطی سعی می‌کنند هر یک از این  $k$  مؤلفه را از ترکیب خطی  $p$  مؤلفه‌ی اولیه بدست آورند.

$$s_i = w_{i,1}x_1 + \dots + w_{i,p}x_p, \quad \text{for } i = 1, \dots, k, \quad \text{or}$$

$$\mathbf{s} = \mathbf{W}\mathbf{x},$$

---

1 Global flat subspace

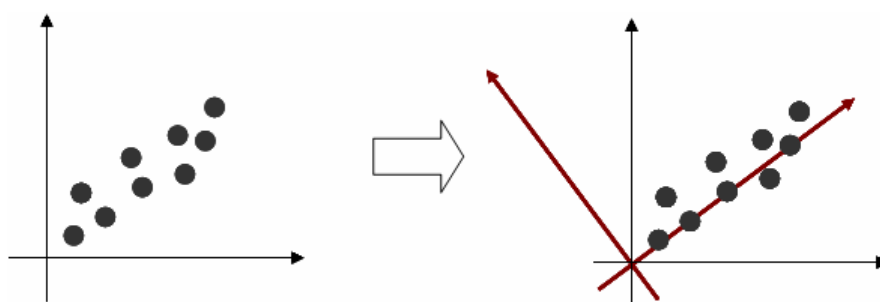
2 Locally flat subspace

که  $W_{k \times p}$  ماتریس وزنهای نگاشت خطی می‌باشد.

تکنیک PCA بهترین روش برای کاهش ابعاد داده به صورت خطی می‌باشد. یعنی با حذف ضرایب کم‌اهمیت بدست آمده از این تبدیل، اطلاعات از دست رفته نسبت به روشهای دیگر کمتر است.

### ۲-۳ تکنیک آنالیز اجزا اصلی (PCA)

در این روش محورهای مختصات جدیدی برای داده‌ها تعریف شده و داده‌ها براساس این محورهای مختصات جدید بیان می‌شوند. اولین محور باید در جهتی قرار گیرد که واریانس داده‌ها ماکسیمم شود (یعنی در جهتی که پراکندگی داده‌ها بیشتر است). دومین محور باید عمود بر محور اول به گونه‌ای قرار گیرد که واریانس داده‌ها ماکسیمم شود. به همین ترتیب محورهای بعدی عمود بر تمامی محورهای قبلی به گونه‌ای قرار می‌گیرند که داده‌ها در آن جهت دارای بیشترین پراکندگی باشند. در شکل زیر این مطلب برای داده‌های دو بعدی نشان داده شده است.



شکل (۱-۳) انتخاب محورهای جدید برای داده‌های دو بعدی



روش PCA به نامهای دیگری نیز معروف است. مانند:

- Karhunen Loeve Transform (KLT)
- Hotelling Transform
- Empirical Orthogonal Function (EOF)

قبل از اینکه به جزئیات این روش بپردازیم ابتدا مفاهیم ریاضی و آماری مرتبط با این روش را بطور مختصر بیان می‌کنیم. این مفاهیم شامل انحراف از معیار استاندارد، کواریانس، بردارهای ویژه و مقادیر ویژه می‌باشد.

### ۱-۲-۳ مفاهیم مقدماتی مورد نیاز در PCA

#### مفاهیم آماری

فرض کنید  $X$  رشته‌ای از مقادیر است. میانگین این مقادیر از رابطه زیر بدست می‌آید.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

انحراف از معیار نیز از رابطه زیر محاسبه می‌شود.

$$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)}}$$

علت اینکه در مخرج رابطه فوق از عبارت  $n-1$  استفاده شده (و نه  $n$ ) اینست که فرض شده  $X$  شامل تمام مقادیر موجود نیست بلکه تعدادی از این مقادیر انتخاب شده‌اند و در  $X$  قرار گرفته‌اند. یعنی  $X$  مجموعه نمونه است و نه کل داده‌ها. با این فرض اگر از  $n-1$  در رابطه فوق استفاده شود، انحراف از معیار بدست آمده به انحراف از معیار داده‌های واقعی نزدیکتر خواهد بود نسبت به اینکه از  $n$  استفاده شود.

با به توان ۲ رساندن انحراف از معیار، واریانس بدست می‌آید.

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)}$$

معیارهایی که در بالا ذکر شد فقط اطلاعات مربوط به یک بُعد را ارائه می‌کنند و دانشی در مورد ارتباط بین ابعاد مختلف به ما نمی‌دهند. با استفاده از کواریانس می‌توانیم ارتباط بین ابعاد مختلف داده‌ها را پیدا کنیم. فرض کنید یک رشته دیگر از اعداد داریم که آن را با  $Y$  نشان می‌دهیم. کواریانس بین  $X$  و  $Y$  از رابطه زیر بدست می‌آید.

$$\text{cov}(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)}$$

مقداری که از رابطه بالا بدست می‌آید در بازه  $[-1, 1]$  قرار خواهد داشت که یکی از سه حالت زیر را بوجود می‌آورد:

اگر مقدار بدست آمده مثبت باشد آنگاه  $X$  و  $Y$  با هم افزایش یا کاهش می‌یابند.

اگر مقدار بدست آمده منفی باشد آنگاه با افزایش  $X$  مقدار  $Y$  کاهش می‌یابد و بالعکس.

اگر مقدار بدست آمده صفر باشد آنگاه  $X$  و  $Y$  از یکدیگر مستقلند.

کواریانس بین تمامی ابعاد داده‌ها را می‌توان دوبه‌دو محاسبه کرده و در یک ماتریس ذخیره کرد. به

این ماتریس، ماتریس کواریانس می‌گویند. ماتریس کواریانس یک ماتریس مربعی متقارن است. مثلاً

اگر سه بعد به نامهای  $x$ ،  $y$  و  $z$  داشته باشیم، ماتریس کواریانس آنها برابر است با:

$$C = \begin{pmatrix} \text{cov}(x, x) & \text{cov}(x, y) & \text{cov}(x, z) \\ \text{cov}(y, x) & \text{cov}(y, y) & \text{cov}(y, z) \\ \text{cov}(z, x) & \text{cov}(z, y) & \text{cov}(z, z) \end{pmatrix}$$

## مفاهیم جبر ماتریسها

در این بخش مفهوم بردار ویژه و مقادیر ویژه را بیان می‌کنیم. همانطور که می‌دانید برای اینکه بتوان دو ماتریس را در یکدیگر ضرب کرد، آن دو باید از نظر اندازه با هم سازگار باشند. بردارهای ویژه نوع خاصی از ضرب ماتریسها را ارائه می‌کنند. به مثال زیر توجه کنید.

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 11 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \\ 8 \end{pmatrix} = 4 \times \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

در مثال اول بردار بدست آمده مضرب صحیحی از بردار اولیه نیست. اما در مثال دوم، بردار بدست آمده چهار برابر بردار اولیه می‌باشد. ماتریس  $2 \times 2$  که در این دو بردار ضرب کرده‌ایم را می‌توان یک ماتریس تبدیل<sup>۱</sup> در نظر گرفت که با ضرب آن در یک بردار می‌توان اندازه و راستای آن بردار را تغییر داد. در میان تمام بردارهایی که می‌توان ماتریس تبدیل را در آنها ضرب کرد، بردارهایی وجود دارد که پس از تبدیل راستای آنها تغییر نمی‌کند و فقط اندازه آنها ممکن است عوض شود، مانند بردار  $[3;2]$  در مثال فوق. این بردارها را بردارهای ویژه می‌نامند. برای هر بردار ویژه یک مقدار ویژه نیز وجود دارد که بیان می‌کند اندازه آن بردار (و تمام بردارهای دیگر که در راستای آن بردار هستند) پس از تبدیل، چند برابر خواهد شد. در مثال فوق مقدار ویژه برای بردار  $[3;2]$  و البته تمام بردارهای هم راستا با آن مانند  $[6;4]$  برابر با 4 می‌باشد.

بردارهای ویژه و مقادیر ویژه فقط برای ماتریسهای مربعی معنی پیدا می‌کنند. یک ماتریس  $n \times n$  می‌تواند دارای  $n$  بردار ویژه باشد. به منظور استاندارد کردن بردارهای ویژه، پس از یافتن بردارهای ویژه اندازه آنها را به گونه‌ای تغییر می‌دهند تا طول آنها برابر با یک شود. مثلاً برای بردار  $[3;2]$  داریم:

---

1 Transformation Matrix

$$\sqrt{3^2 + 2^2} = \sqrt{13} \quad \begin{pmatrix} 3 \\ 2 \end{pmatrix} \div \sqrt{13} = \begin{pmatrix} 3/\sqrt{13} \\ 2/\sqrt{13} \end{pmatrix}$$

ویژگی مهم بردارهای ویژه اینست که آنها برهم عمودند. مثلاً ماتریس تبدیل زیر را در نظر بگیرید.

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

با ضرب ماتریس  $A$  در هر بردار می‌توان قرینه آن بردار نسبت به خط  $y=0$  را بدست آورد. بردارهای ویژه‌ی این ماتریس عبارتند از  $[1;0]$  و  $[0;1]$ . مقادیر ویژه‌ی این بردارها نیز به ترتیب 1- و 1 می‌باشد. همانطور که گفتیم این دو بردار ویژه برهم عمودند.

بدست آوردن بردارهای ویژه برای ماتریسهای بزرگتر از  $3 \times 3$  کار نسبتاً دشواری است. این کار توسط یک الگوریتم بازگشتی انجام می‌شود که توضیح آن خارج از حوصله‌ی این گزارش است. برای اطلاعات بیشتر در این رابطه می‌توانید به [۹۶] مراجعه کرد.

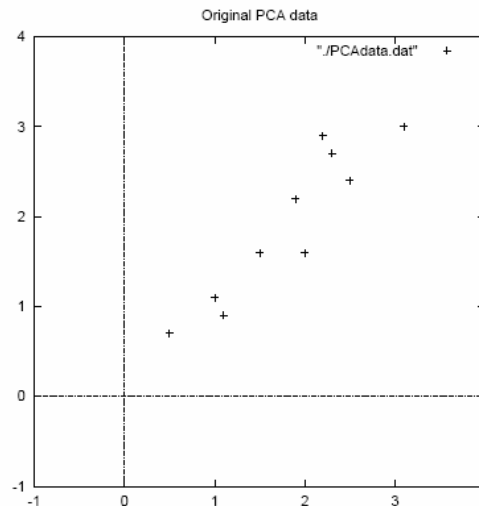
### ۲-۲-۳ الگوریتم PCA

در این بخش الگوریتم PCA را با ذکر یک مثال توضیح می‌دهیم [۹۶].

#### مرحله ۱- انتخاب داده

در اینجا ما قصد داریم PCA را بر روی یک مجموعه داده‌ی دو بعدی اعمال کنیم. این داده‌ها در شکل زیر نشان داده شده است.

$x$	$y$
2.5	2.4
0.5	0.7
2.2	2.9
1.9	2.2
3.1	3.0
2.3	2.7
2	1.6
1	1.1
1.5	1.6
1.1	0.9



شکل (۳-۲) داده‌های دوبعدی اولیه که قرار است PCA بر روی آنها اعمال شود

## مرحله ۲- کم کردن میانگین از داده‌ها

در این مرحله، میانگین هر بُعد را از مقادیر آن بُعد کم می‌کنیم تا میانگین داده‌ها در هر بُعد صفر شود.

## مرحله ۳- محاسبه‌ی ماتریس کواریانس

ماتریس کواریانس را به طریقی که در بالا ذکر شد برای داده‌ها بدست می‌آوریم. برای مثال ما این ماتریس، یک ماتریس  $2 \times 2$  است:

$$cov = \begin{pmatrix} .616555556 & .615444444 \\ .615444444 & .716555556 \end{pmatrix}$$

## مرحله ۴- محاسبه‌ی بردارهای ویژه و مقادیر ویژه

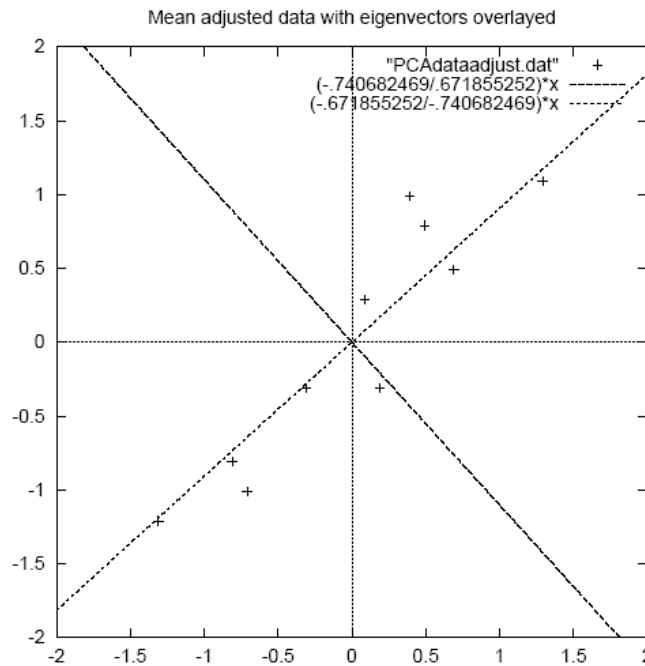
اکنون بردارهای ویژه و مقادیر ویژه‌ی ماتریس کواریانس را محاسبه می‌کنیم. طبق قضایای جبر خطی، یک ماتریس متقارن  $n \times n$  دارای  $n$  بردار ویژه‌ی مستقل و  $n$  مقدار ویژه می‌باشد. همچنین

یک ماتریس نیمه قطعی مثبت، دارای مقادیر ویژه‌ی غیر منفی است. این مقادیر برای مثال ما برابر است با:

$$eigenvalues = \begin{pmatrix} .0490833989 \\ 1.28402771 \end{pmatrix}$$

$$eigenvectors = \begin{pmatrix} -.735178656 & -.677873399 \\ .677873399 & -.735178656 \end{pmatrix}$$

توجه داشته باشید که این دو بردار ویژه به گونه‌ای انتخاب شده‌اند که طول هر دو برابر با 1 باشد. اما این دو بردار چه چیزی به ما می‌دهد؟ ما راستای این دو بردار را در شکل ۳-۳ نشان داده‌ایم. همانطور که می‌بینید یکی از این دو بردار در جهتی قرار گرفته است که داده‌ها در آن جهت بیشترین پراکندگی را دارند. بردار دیگر نیز عمود بر بردار اول است. و اما مقادیر ویژه چه چیزی ارائه می‌دهند؟ در این مثال برداری که در راستای بیشترین پراکندگی داده‌ها قرار گرفته دارای مقدار ویژه 1.284 و بردار دیگر دارای مقدار ویژه 0.049 می‌باشد. در واقع مقادیر ویژه میزان پراکندگی داده‌ها در راستای بردار ویژه‌ی مربوطه را نشان می‌دهد. می‌توان گفت بردار ویژه‌ای که دارای بزرگترین مقدار ویژه است مؤلفه‌ی اصلی<sup>۱</sup> داده‌های موجود می‌باشد.



شکل (۳-۳) داده‌های نرمالسازی شده (با کم شدن میانگین) به همراه بردارهای ویژگی ماتریس کواریانس

## مرحله ۵- انتخاب مؤلفه‌ها و ساختن Feature Vector

در این مرحله مفهوم کاهش ابعاد داده وارد می‌شود. بردارهای ویژه‌ای که در مرحله‌ی قبل بدست آوردیم را بر اساس مقادیر ویژه‌ی آنها از بزرگ به کوچک مرتب می‌کنیم (توجه داشته باشید که مقادیر ویژه‌ی ماتریس کواریانس همگی بزرگتر یا مساوی صفر هستند). بدین ترتیب مؤلفه‌های داده‌ها از پر اهمیت به کم اهمیت مرتب می‌شوند. در اینجا اگر بخواهیم ابعاد داده‌ها را کاهش دهیم می‌توانیم مؤلفه‌های کم اهمیت را حذف نماییم. البته این کار با از دست دادن مقدار کمی اطلاعات همراه است.

کاری که باید در این مرحله انجام دهیم ایجاد یک Feature Vector است که در واقع ماتریسی از بردارها می‌باشد. این ماتریس شامل آن بردارهای ویژگی‌ای است که ما می‌خواهیم آنها را نگه داریم.

$$FeatureVector = (eig_1 \ eig_2 \ eig_3 \ .... \ eig_n)$$

اگر همه‌ی بردارهای ویژگی را در این ماتریس قرار دهیم، هیچ اطلاعاتی از دست نخواهد رفت و دوباره می‌توانیم دقیقاً همان داده‌های اولیه را بدست آوریم. در ادامه‌ی مثال فوق، Feature Vector را برابر با مقدار زیر در نظر می‌گیریم.

$$FeatureVector = \begin{pmatrix} -.677873399 \\ -.735178656 \end{pmatrix}$$

**مرحله ۶-** بدست آوردن داده‌های جدید

در آخرین مرحله از PCA فقط باید ترانهاده‌ی ماتریس Feature Vector که در مرحله‌ی قبل بدست آوردیم را در ترانهاده‌ی داده‌های نرمالسازی شده ضرب نماییم.

$$FinalData = RowFeatureVector \times RowDataAdjust$$

که RowFeatureVector ماتریسی است که بردارهای ویژه در سطرهاى آن به ترتیب مقادیر ویژه از بالا به پایین قرار گرفته‌اند، و RowDataAdjust ماتریسی است که شامل داده‌هایی است که میانگین هر بُعد از آن بُعد کم شده است. در این ماتریس، داده‌ها در ستونهای آن ذخیره شده و هر سطر آن مربوط به یک بُعد است. در مثال ذکر شده بدلیل اینکه ما فقط یکی از بردارهای ویژه را انتخاب کردیم داده‌های بدست آمده از PCA، داده‌های یک بُعدی می‌باشند.

$x$
-.827970186
1.77758033
-.992197494
-.274210416
-1.67580142
-.912949103
.0991094375
1.14457216
.438046137
1.22382056

شکل (۳-۴) داده‌های بدست آمده از تبدیل PCA با انتخاب مهمترین بردار ویژگی



با استفاده از رابطه‌ی زیر می‌توانیم مقادیری که از تبدیل PCA بدست آورده‌ایم را به داده‌های اولیه که مقدار میانگین از آنها کم شده بازگردانیم.

$$RowDataAdjust = RowFeatureVector^{-1} \times FinalData$$

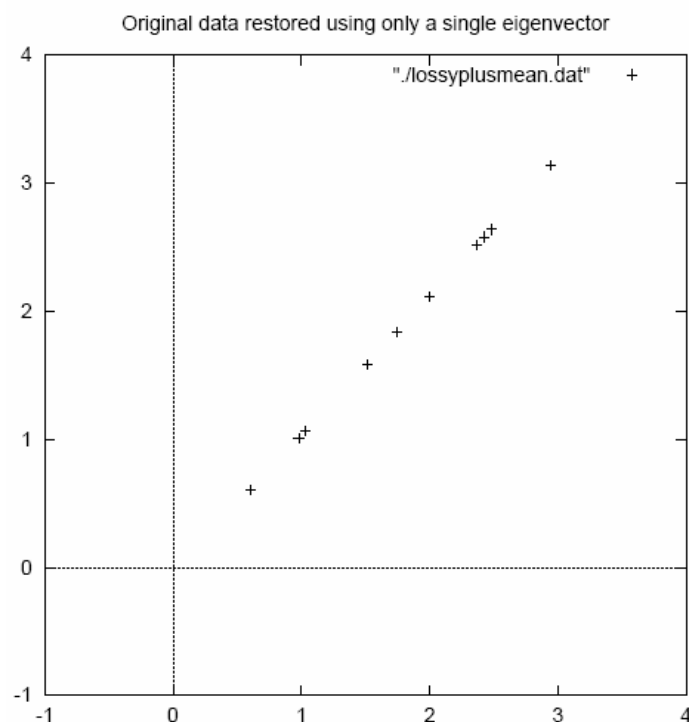
بدلیل اینکه ماتریس RowFeatureVector حاوی بردارهای ویژه‌ی یکه است، معکوس آن با ترانزپوز آن برابر است. بنابراین:

$$RowDataAdjust = RowFeatureVector^T \times FinalData$$

با اضافه کردن میانگین، داده‌های اولیه را خواهیم داشت:

$$RowOriginalData = (RowFeatureVector^T \times FinalData) + OriginalMean$$

در شکل زیر داده‌هایی که پس از تبدیل PCA بازیابی شده‌اند را مشاهده می‌کنید. همانطور که می‌بینید مقدار کمی اطلاعات از بین رفته است و باعث شده داده‌ها همگی در امتداد یک خط راست قرار گیرند.



شکل (۳-۵) داده‌های بازیابی شده از تبدیل PCA با انتخاب مهمترین بردار ویژگی

## فصل چهارم شبکه عصبی

### ۴-۱ شبکه عصبی چیست؟

شبکه‌های عصبی را می‌توان با اگماض زیاد، مدل‌های الکترونیکی از ساختار عصبی مغز انسان نامید. مکانیسم فراگیری و آموزش مغز اساساً بر تجربه استوار است. مدل‌های الکترونیکی شبکه‌های عصبی طبیعی نیز بر اساس همین الگو بنا شده‌اند و روش برخورد چنین مدل‌هایی با مسائل، با روش‌های محاسباتی که به‌طور معمول توسط سیستم‌های کامپیوتری در پیش گرفته شده‌اند، تفاوت دارد. می‌دانیم که حتی ساده‌ترین مغزهای جانوری هم قادر به حل مسائلی هستند که اگر نگوییم کامپیوترهای امروزی از حل آنها عاجز هستند، حداقل در حل آنها دچار مشکل می‌شوند. به عنوان مثال، مسائل مختلف شناسایی الگو، نمونه‌ای از مواردی هستند که روش‌های معمول محاسباتی برای حل آنها به نتیجه مطلوب نمی‌رسند. درحالی‌که مغز ساده‌ترین جانوران به راحتی از عهده چنین مسائلی بر می‌آید. تحقیقات در این زمینه نشان داده است که مغز، اطلاعات را همانند الگوها<sup>۱</sup> ذخیره می‌کند. فرآیند ذخیره‌سازی اطلاعات به صورت الگو و تجزیه و تحلیل آن الگو، اساس روش نوین محاسباتی را تشکیل می‌دهند. این حوزه از دانش محاسباتی<sup>۲</sup> به هیچ وجه از روش‌های برنامه‌نویسی سنتی استفاده نمی‌کند و به جای آن از شبکه‌های بزرگی که به صورت موازی آرایش شده‌اند و تعلیم یافته‌اند، بهره می‌جوید.

یک شبکه عصبی مصنوعی<sup>۳</sup> ایده‌ای است برای پردازش اطلاعات که از سیستم عصبی زیستی الهام گرفته شده و مانند مغز به پردازش اطلاعات می‌پردازد. عنصر کلیدی این ایده، ساختار جدید سیستم پردازش اطلاعات است. این سیستم از شمار زیادی عناصر پردازشی فوق العاده بهم پیوسته

---

1 pattern

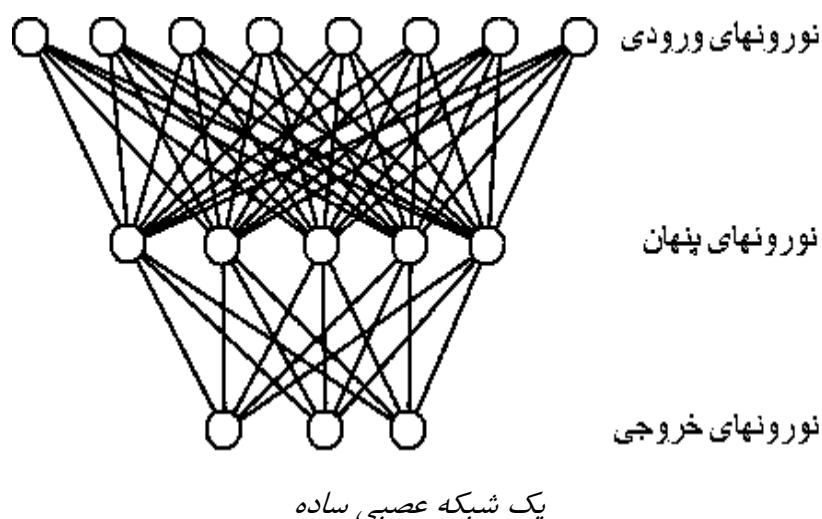
2 computation

3 Artificial Neural Network (ANN)

تشکیل شده<sup>۱</sup> که برای حل یک مسأله با هم هماهنگ عمل می کند. ANN ها، نظیر انسانها، با مثال یاد می گیرند. یک ANN برای انجام وظیفه ای مشخص، مانند شناسایی الگوها و دسته بندی

اطلاعات، در طول یک پروسه یاد گیری، تنظیم می شود. در سیستم های زیستی یاد گیری با تنظیماتی در اتصالات سیناپسی که بین اعصاب قرار دارد همراه است این روش ANN ها نیز است. یک شبکه عصبی مصنوعی شامل تعداد زیادی از نورونهاست که بر اساس یک الگوی اتصال به هم پیوند یافته اند. نورونها معمولاً به سه دسته تقسیم میشوند: نورونهای ورودی، که ورودی را برای پردازش دریافت میکنند. نورونهای خروجی که حاوی نتیجه پردازش هستند و نورونهای میانی که به نام نورونهای پنهان<sup>۲</sup> نامیده میشوند.

این یک مثال ساده از شبکه عصبی است:



نورونهای ورودی از طریق منابع خارجی فعال میشوند. هر نورون ورودی پس از فعال شدن مقدار فعال سازی خود را به نورونهای پنهانی که به آنها متصل است میفرستد. هر کدام از نورونهای پنهان با توجه به مقادیری که از نورونهای همسایه خود دریافت میکند، فعال بودن خود را تعیین میکند و اگر فعال بشود مقدار فعال سازی خود را به نورونهای خروجی یا به لایه دیگری از نورونهای پنهان

---

1 Neurons  
2 Hidden Neurons

می فرستد و به طور پیوسته و همزمان تمام نورونها به کار خود ادامه میدهد و به این ترتیب سیگنالهای تولید شده توسط نورونهای ورودی در تمام شبکه پخش میشوند تا وضعیت فعال بودن نورونهای خروجی مشخص شود. الگوی فعال شدن یک شبکه توسط وزن بین نورونهای آن مشخص میشود. وزن میتواند مثبت یا منفی باشد. یک وزن منفی نشاندهنده عدم تمایل نورون مقصد نسبت به فعال شدن در صورت فعال بودن نورون مبدا است. فعال بودن هر نورون توسط یک تابع ساده فعال سازی<sup>۱</sup> انجام میشود. توابع فعال سازی مختلفی وجود دارند اما همگی بر یک پایه بنا شده اند. تابع تمام ورودیهای نورون (اگر نورون مبدا فعال باشد اندازه ورودی آن به نورون مقصد به اندازه وزن بین آن دو است)، را با هم جمع میکند. حاصل جمع معمولاً کمی تغییر می یابد. مثلاً به عددی بین ۰ و ۱ تبدیل میشود و سپس اگر این مجموع از عددی خاص بیشتر بود، نورون فعال میشود. شبکه‌ای که در بالا معرفی شد معروف به شبکه feed forward است. در این شبکه ورودی از لایه‌ی ورودی دریافت شده و سپس از یک یا چند لایه پنهان میگذرد تا به لایه خروجی برسد. هیچ یالی درون یک لایه یا از لایه جلوتر به لایه عقب تر وجود ندارد. مدلهای واقعی تر از مغز بسیار پیچیده تر هستند. آنها شامل تعداد بسیار زیادی از لایه های پنهان، و اتصالات<sup>۲</sup> برگرداننده<sup>۳</sup> هستند که سیگنال را از لایه های جلویی به لایه های عقبی میفرستد. این برگرداندن برای توصیف کردن خصوصیات<sup>۳</sup> مثل حافظه کوتاه مدت<sup>۳</sup> ضروری هستند.

---

1 Activation Function  
2 Recurrent Connections  
3 Short Term Memory

## ۴-۲ چرا از شبکه های عصبی استفاده می کنیم؟

شبکه های عصبی ، با قابلیت قابل توجه در استنتاج معانی از داده های پیچیده یا مبهم ، برای استخراج الگوها و شناسایی روشهایی که آگاهی از آنها برای انسان و دیگر تکنیک های کامپیوتری بسیار پیچیده و دشوار است به کار گرفته می شوند. یک شبکه عصبی تربیت یافته می تواند به عنوان یک متخصص در مقوله اطلاعاتی ای که برای تجزیه تحلیل به آن داده شده به حساب آید. از این متخصص می توان برای برآورد وضعیت های دخواه جدید و جواب سؤال های "چه می شد اگر " استفاده کرد.

مزیت های دیگر آن شامل موارد زیر می شود :

۱. **یادگیری انطباق پذیر:** قابلیت یاد گیری نحوه انجام وظایف بر پایه اطلاعات داده شده

برای تمرین و تجربه های مقدماتی .

۲. **سازماندهی توسط خود:** یک ANN می تواند سازماندهی یا ارائه اش را ، برای اطلاعاتی

که در طول دوره یادگیری در یافت می کند، خودش ایجاد کند.

۳. **عملکرد بهنگام<sup>۱</sup>:** محاسبات ANN می تواند بصورت موازی انجام شود، و سخت

افزارهای مخصوصی طراحی و ساخته شده است که می تواند از این قابلیت استفاده کند.

**تحمل اشتباه بدون ایجاد وقفه در هنگام کد گذاری اطلاعات :** خرابی جزئی یک شبکه

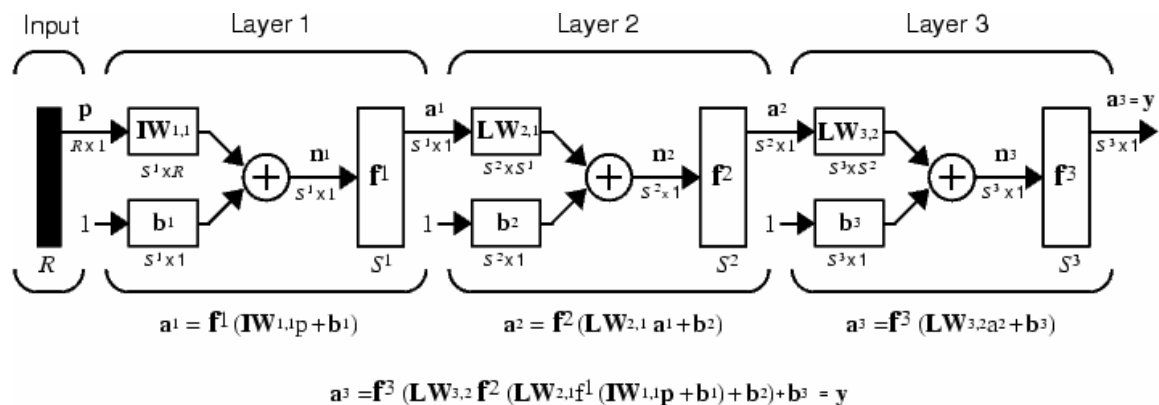
منجر به تنزل کارایی متناظر با آن می شود اگر چه تعدادی از قابلیت های شبکه ممکن است

حتی با خسارت بزرگی هم باقی بماند.

### ۳-۴ شبکه عصبی MLP<sup>۱</sup>

بیشتر روشهایی که جهت دست یابی به کارائی بالا در بازشناسی کاراکترها ارائه شده است عملیات شناسائی را با استفاده از راه حلهای آماری ، شبکه های عصبی و ترکیبی از شبکه های عصبی و روشهای آماری ، انجام می دهند . اینگونه روشها را می توان با بررسی پارامترهایی از قبیل ، سرعت، کارائی ، نیازمندیهای حافظه ای ، امکان پیاده سازی بر روی سیستمهای دیجیتال و در شبکه های عصبی سرعت یادگیری ارزیابی کرد . در این میان شبکه های عصبی بیشتر از سایرین جالب توجه است . چون تعادل مناسبی را بین کارائی ، سرعت و امکانات سخت افزاری بر قرار می کند . MLP نیز نمونه ای از شبکه های عصبی است که جهت دسته بندی اطلاعات مورد استفاده قرار می گیرد.

در شکل زیر نمونه ای از یک شبکه MLP مشاهده می نمایید



شبکه MLP همانطور که از اسمش بر می آید مجموعه ای از نرونها است که در لایه های مختلفی پشت سر هم قرار گرفته اند . مقادیر ورودی پس از ضرب در وزنهای موجود در گذر گاههای بین لایه ها به نرون بعدی رسیده ودر آنجا باهم جمع شده وپس از عبور از تابع شبکه مربوطه خروجی نرون را تشکیل می دهند.در پایان خروجی بدست آمده با خروجی مورد نظر مقایسه شده و خطای

<sup>1</sup> Multilayer perceptron

بدست آمده جهت اصلاح وزنهای شبکه به کار می رود، که این امر اصطلاحاً آموزش شبکه عصبی نامیده می شود، که در ادامه می آید.

#### ۴-۳-۱ قاعده فراگیری MLP

قاعده فراگیری پرسپترون چند لایه را «قاعده کلی دلتا»<sup>۱</sup> یا «قاعده پس انتشار»<sup>۲</sup> گویند. این عناوین در سال ۱۹۸۶ توسط رومل هارت، مک کلند و ویلیامز پیشنهاد شد [۹۶].

این افتخار به آن ها تعلق می گیرد که اولین گروهی بودند که نه تنها قاعده فراگیری پرسپترون را به طور مستقل کشف کردند بلکه با ترکیب آن ها پرسپترون چند لایه ای را ایجاد کرده و مورد مطالعه قرار دادند. کتاب آن ها به نام «پردازش توزیع شده موازی»<sup>۳</sup> هنوز یکی از مهم ترین کتاب های این حوزه علمی است.

نحوه عمل پرسپترون چند لایه ای مشابه پرسپترون تک لایه ای است. بدین صورت که الگویی به شبکه عرضه می شود و خروجی آن محاسبه می گردد، مقایسه خروجی واقعی و خروجی مطلوب باعث می گردد که ضرایب وزنی شبکه تغییر یابد به طوری که در دفعات بعد خروجی درست تری حاصل شود. قاعده فراگیری روش میزان کردن ضرایب وزنی شبکه را بیان می کند.

وقتی به شبکه آموزش ندیده ای الگویی را عرضه می کنیم، خروجی های تصادفی تولید می کند. ابتدا باید تابع خطایی را تعریف کنیم که تفاوت خروجی واقعی و خروجی مطلوب را نشان دهد. چون خروجی مطلوب را می دانیم این نوع فراگیری را «فراگیری با سرپرست»<sup>۴</sup> می نامیم. برای موفق شدن در آموزش شبکه باید خروجی آن را به تدریج به خروجی مطلوب نزدیک کنیم. به عبارت دیگر باید میزان تابع خطا را به طور دائم کاهش دهیم. برای این منظور ضرایب وزنی خطوط ارتباطی واحد ها با استفاده از قاعده کلی دلتا میزان می شود. قاعده دلتا مقدار تابع خطا را محاسبه

---

1 Delta rule

2 Backpropagation

3 Parallel Distributed Processing

4 Supervised learning

کرده و آن را به عقب از یک لایه به لایه پیشین انتشار می دهد. عبارت «پس انتشار» به این علت است. ضرایب وزنی هر واحد جداگانه میزان می شود و بدین صورت میزان خطا کاهش می یابد. این عمل در مورد واحد های لایه خارجی ساده است زیرا خروجی واقعی و مطلوب آن ها را می دانیم، ولی در مورد لایه میانی چندان روشن نیست. این گمان می رود که ضرایب وزنی واحدهای پنهان که به واحدهای خروجی با میزان خطایی بزرگ مرتبط هستند باید بیش تر از واحدهای پنهان که به واحد های مرتبط آن ها خروجی تقریباً صحیحی دارند تغییر یابد. در واقع ریاضیات نشان می دهد که ضرایب واحد ها باید به تناسب میزان خطای واحدی که به آن متصل اند تغییر کند. بنابراین می توان با انتشار خطا به عقب ضرایب وزنی خطوط ارتباطی تمام لایه ها را به درستی میزان کرد. به این طریق تابع خطا کاهش ، شبکه آموزش می یابد.

### الگوریتم پرسپترون چند لایه ای

الگوریتم پرسپترون چند لایه ای که از قاعده آموزش پس انتشار استفاده می کند در زیر شرح داده شده است. این الگوریتم به توابع غیر خطی نیاز دارد که به طور پیوسته قابل مشتق گیری

باشند. به عبارت دیگر توابع باید هموار باشند. ما استفاده از تابع سیگموید،  $f(net) = \frac{1}{1 + e^{-knet}}$  را

به علت سادگی مشتق آن را انتخاب کرده ایم.

الگوریتم آموزش پرسپترون چند لایه ای

۱. مقادیر اولیه ضرایب وزنی و آستانه ها را انتخاب کنید

تمام وزن ها و آستانه ها را برابر با اعداد کوچک تصادفی قرار دهید.

۲. ورودی ها و خروجی های مطلوب را به شبکه عرضه کنید.

ورودی  $x_p = x_0, x_1, x_2, \dots, x_{n-1}$  و خروجی هدف  $T_p = t_0, t_1, t_2, \dots, t_{m1}$  را به شبکه عرضه کنید. n

تعداد عناصر بردارهای ورودی و مقدار عناصر بردارهای خروجی است. ضریب وزنی  $W_0$  را برابر

با منفی مقدار آستانه،  $\theta$ ، و مقدار  $X_0$  را برابر ۱ قرار دهید. اگر مسئله مود نظر ما مسئله تداعی

باشد  $X_p$  و  $T_p$ ، نمایانگر دو بردار تداعی شونده هستند. در مسئله طبقه بندی تمام عناصر  $T_p$



برابر با صفر قرار داده می شود مگر یکی از عناصر که برابر با ۱ است و آن طبقه ای را نشان می دهد که  $X_p$  در آن قرار دارد.

۳. خروجی را محاسبه کنید.

هر لایه مقادیر زیر را محاسبه کرده و به لایه بعدی انتقال می دهد.

$$y_{pj} = \left[ \sum_{i=0}^{n-1} w_{ij} x_i \right] \text{ و } f(net) = \frac{1}{1 + e^{-knet}}$$

۴. ضرایب وزنی را میزان کنید. ابتدا از لایه خارجی شروع کنید و به عقب برگردید.

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_{pj} o_{pj}$$

$w_{ij}(t)$  نشان دهنده ضرایب وزنی از گره  $i$  به گره  $j$  در زمان  $t$ ،  $\eta$  ضریب بهره و  $\delta_{pj}$  نمایانگر خطای

مربوط به الگوی  $p$  در گره  $j$  است

در مورد واحد های لایه خارجی

$$\delta_{pj} = k o_{pj} (1 - o_{pj}) (t_{pj} - o_{pj})$$

در مورد واحد های لایه پنهان

$$\delta_{pj} = k o_{pj} (1 - o_{pj}) \sum_k \delta_{pk} w_{jk}$$

در حالی که عمل جمع در مورد  $k$  واحد واقع در لایه بعد از واحد  $j$  صورت می گیرد.

### تجسم رفتار شبکه

حال که قاعده کلی دلتا را به طور اجمال و به صورت ریاضی مورد بحث قرار دادیم لازم است روشی را برای تجسم آنچه در شبکه می گذرد به کار گیریم. تحلیل ریاضی شبکه ها رهیافت آسوده و مفیدی را برای تصور رفتار شبکه ها فراهم می آورد. همان طور که دیدیم، شبکه تابع خطا یا انرژی

$$E_p = \frac{1}{2} \sum (t_{pj} - o_{pj})^2$$

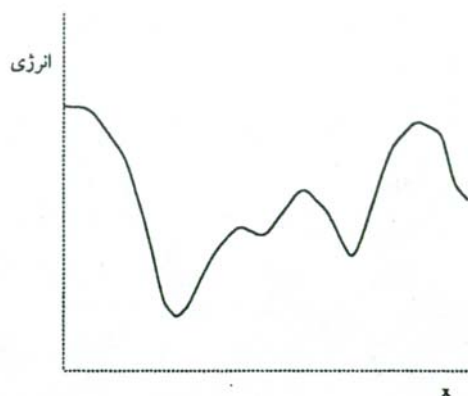
را محاسبه می کند. این تابع نمایانگر انحراف خروجی واقعی شبکه از

خروجی مطلوب است. اختلاف های زیاد حاکی از انرژی زیاد و اختلاف های کم حاکی از انرژی کم است. چون خروجی شبکه به ضرایب وزنی بین واحدها و ورودی آن ها مربوط است، انرژی نیز تابع

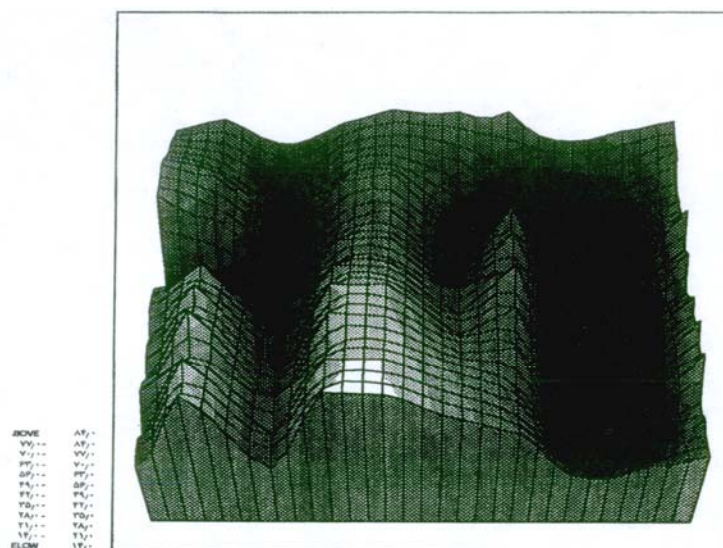
ضرایب وزنی و ورودی شبکه خواهد بود. می توانیم نمودار تابع انرژی را طوری ترسیم کنیم که نحوه اثر تغییرات ضرایب وزنی بر مقدار انرژی را برای یک الگوی ثابت نشان دهد. برای مثال اگر یک شبکه پیچیده را در نظر بگیریم و تغییرات تابع انرژی آن را بر حسب تغییرات در یکی از ضرایب وزنی آن ترسیم کنیم ممکن است شکلی مانند ۴-۱ به وجود آید.

اگر این بحث را به تغییرات دو ضریب وزنی تعمیم دهیم، آن گاه تعداد محورهای ضرایب به دو افزایش یافته و نمودار به دست آمده تابع انرژی به صورتی خواهد بود که مثلاً در شکل ۲-۲ آمده است.

با دو محور ضرایب وزنی و یک محور انرژی، نموداری سه بعدی خواهیم داشت. اگر ضریب وزنی دیگری نیز تغییر کند آن گاه محور چهارمی خواهیم داشت که ترسیم آن دشوار است. به طور کلی می توانیم تمام ضرایب وزنی شبکه را تغییر دهیم. بدیهی است



تابع انرژی در یک بعد بر حسب تغییرات یکی از ضرایب وزنی برای یک الگوی ثابت.



تابع انرژی در دو بعد. به دره سمت راست توجه کنید اگر از یک نقطه میانی در جلوی صفحه شروع کنید ممکن است بر حسب جهتی که هر بار انتخاب می کنید از دره سمت راست پایین رفته و به عمق دره برسید و یا مجدداً قله ای را دور زده و به نقطه دیگری منتهی شوید. ضمناً توجه کنید که دره سمت چپ چندین گودال دارد. این نقاط کمینه محلی ممکن است جواب مسئله را به خود جلب کرده و مانع از رسیدن به نقاط عمیق تر شوند.

شکل (۴-۱)

شکل (۴-۲)

که تعداد آن ها بسیار زیاد باشد در نتیجه تابع انرژی<sup>۱</sup> خواهیم داشت که قابل ترسیم نیست. لیکن بهتر است این تابع را تا حد امکان به صورت صفحه ای سه بعدی تصور کنیم ولی بخاطر داشته باشیم که صفحه انرژی<sup>۲</sup> در واقع چند بعدی است. درک مسائل در ابعاد بالاتر با قیاس مسائل در فضای سه بعدی که تصور آن راحت است ساده تر می گردد. صفحه انرژی چون چشم اندازی موج پوشیده از تپه ها و دره ها، چاله ها و کوه ها است. چاله ها در واقع نقاط کم ترین انرژی و قله ها نقاط بیش ترین انرژی اند.

قاعده کلی دلتا سعی می کند تابع خطای  $E$  را با میزان کردن ضرایب وزنی شبکه کمینه کند به طوری که ضرایب به دست آمده منطبق بر پایین ترین نقطه صفحه انرژی گردند.

برای این منظور از روشی به نام کاهش گرادیان<sup>۳</sup> استفاده می شود.

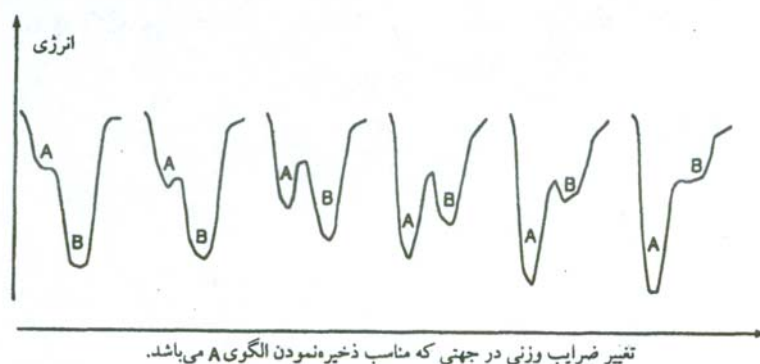
در این روش مقدار تابع انرژی محاسبه می شود، سپس ضرایب وزنی در جهت بیش ترین کاهش گرادیان تغییر می کنند. این روش هنگامی که صفحه انرژی ساده باشد رسیدن به جواب را تضمین می کند. هر یک از گودال ها و یا حوضچه ها نمایانگر یکی از جواب های احتمالی مسئله است. این نقاط که اصطلاحاً حوضچه های ربایش<sup>۴</sup> نامیده می شوند نمایانگر آن دسته از ضرایب وزنی است که خروجی مطلوب را در خروجی ورودی عرضه شده به شبکه تولید می کند. به خاطر داشته باشید که این حوضچه ها در واقع چند بعدی است لیکن آن ها را فقط به صورت سه بعدی می توانیم ترسیم کنیم.

می توانیم صفحه چشم انداز انرژی را به صورت ورقه لاستیکی بزرگ قابل انعطاف تصور کنیم که در ابتدا صاف است. حوضچه های ربایش با قرار دادن گلوله های سنگین بر این ورقه لاستیکی ایجاد می شود و سطح ورقه لاستیکی گود شده و چاله هایی به وجود می آید. کف چاله ها نمایانگر نقاط کم ترین انرژی و جواب هایی است که شبکه فرا گرفته است.

---

1 Energy function  
2 Energy landscape  
3 Gradient descent  
4 Basins of attraction

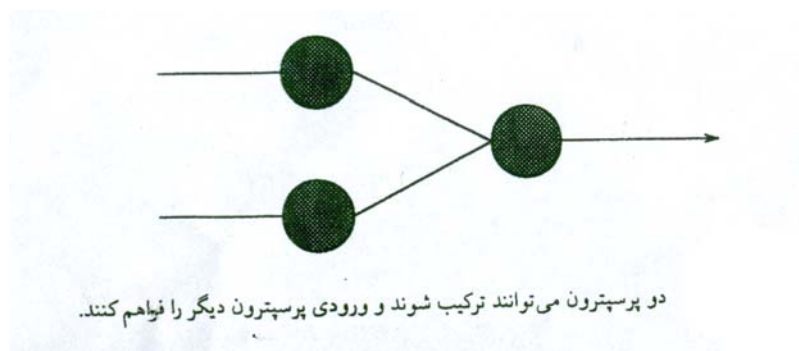
می‌توانیم فضایی چند بعدی را تصور کنیم که هر محور آن نماینده یکی از ضرایب وزنی باشد. در آن صورت هر نقطه در فضا نمایانگر ترکیبی ویژه از ضرایب وزنی شبکه می‌باشد. به این فضا طبعاً فضای ضرایب وزنی گویند. مثال فوق نشان داد که چگونه می‌توانیم تغییرات انرژی را هنگامی که ضرایب وزنی تغییر می‌کند تصور کنیم. لیکن به همان طریق می‌توانیم تغییرات انرژی را برای یک دسته ضرایب ثابت در حالی که بردارهای ورودی تغییر می‌کردند تصور کنیم. در آن صورت هر نقطه از فضای ضرایب وزنی معرف یک صفحه چشم اندازی انرژی ویژه می‌باشد در حالی که متغیرهای فضا الگوهای مختلف و مقدار انرژی مربوط به آن‌هاست. نمودار این رفتار در شکل ۳-۴ آمده است.



شکل (۳-۴)

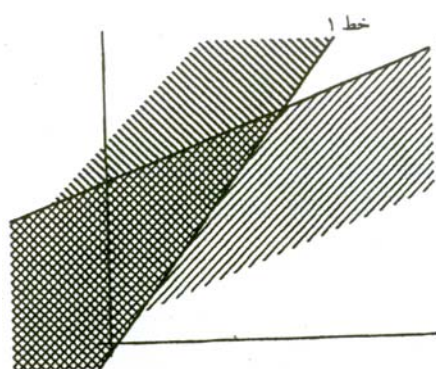
بسیاری از ویژگی‌های پرسپترون چند لایه ای با استفاده از شیوه بیان صفحه انرژی بهتر قابل درک می‌باشند.

پرسپترون تک لایه ای تنها قادر است که یک صفحه تفکیک کننده طبقات ایجاد کند. شبکه ای را شامل سه پرسپترون به صورتی که در شکل ۴-۴ آمده است در نظر بگیرید.



شکل (۴-۴)

اگر مقدار آستانه واحد لایه دوم طوری تعیین شود که هنگامی که هر دو واحد ورودی فعال هستند آن واحد فعال شود شبکه ای خواهیم داشت که عمل AND منطقی را انجام می دهد. چون هر کدام از واحدهای لایه اول خطی را در فضای الگوها تعریف می کند و واحد لایه دوم با ترکیب این دو خط فضا را طبقه بندی می کند. اگر یکی از واحدهای لایه اول طوری تنظیم شود که هنگامی که الگوی ورودی بالای خط مربوط به آن واحد است فعال شود و خروجی ۱ تولید کند و واحد دیگر طوری تنظیم شود که هنگامی که الگوی ورودی زیر خط مربوط به آن است فعال شود، آن گاه لایه دوم همان طور که در شکل ۴-۵ نشان داده شده تنها هنگامی فعال می شود و جواب ۱ تولید می کند که الگوی ورودی بالای خط اول و زیر خط دوم باشد.

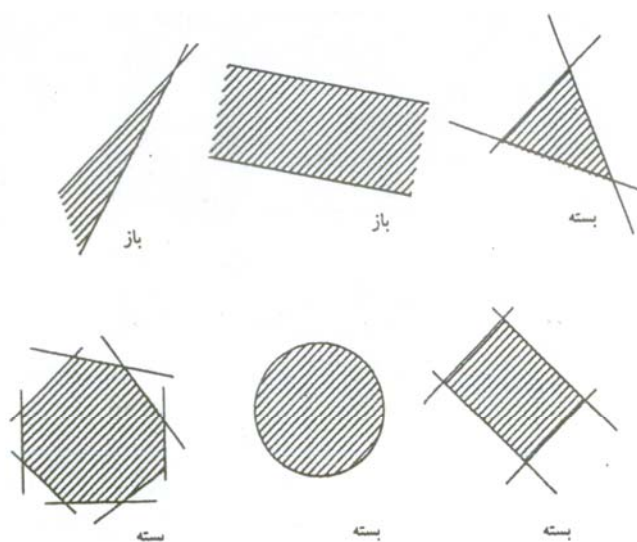


شکل (۴-۵)

ترکیب پرسپترون ناحیه  
تصمیم از ترکیب دو پرسپترون و یک  
پرسپترون دیگر به وجود می آید.

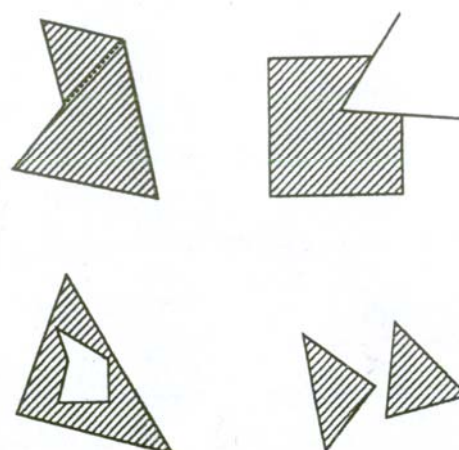
در لایه اول می توان بیش از دو واحد به کار گرفت که در این صورت تفکیک فضا با ترکیب بیش از ۲ خط صورت می گیرد. ناحیه هایی را که به این طریق ایجاد می شود ناحیه های محدب می نامند. ناحیه محدب قسمتی از فضا است که هر نقطه از آن را بتوان با یک قطعه خط به یکدیگر وصل کرد به طوری که آن قطعه خط محدوده فضا را قطع نکند. ناحیه ها می توانند بسته یا باز باشند. یک ناحیه بسته محدوده ای به دور خود دارد مانند مثلث و دایره. ناحیه باز محدوده ای ندارد مانند فضای محصور بین دو خط موازی. شکل ۴-۶ مثال هایی را از ناحیه های محدب باز و بسته نشان می دهد.

افزایش تعداد پرسپترون ها در لایه اول امکان می دهد که اضلاع بیش تر و بیش تری را تعریف کنیم. از بحثی که داشتیم نتیجه می شود که تعداد اضلاع ناحیه تفکیک شده از فضا حداکثر به اندازه تعداد پرسپترون ها در لایه اول است و ناحیه به دست آمده به هر صورت محدب است. لیکن اگر لایه ای جدید اضافه کنیم، پرسپترون این لایه به جای خط ناحیه های محدب را به عنوان ورودی دریافت می کنند. ترکیب این فضا ها همان طور که در شکل ۴-۷ نشان داده شده دیگر لزوماً محدب نیست. در این ترکیب ناحیه های محدب ممکن است همدیگر را قطع کنند، بر یکدیگر منطبق گردند یا از یکدیگر جدا باشند و بدین طریق اشکال دلخواهی را ایجاد کنند.



شکل (۴-۶)

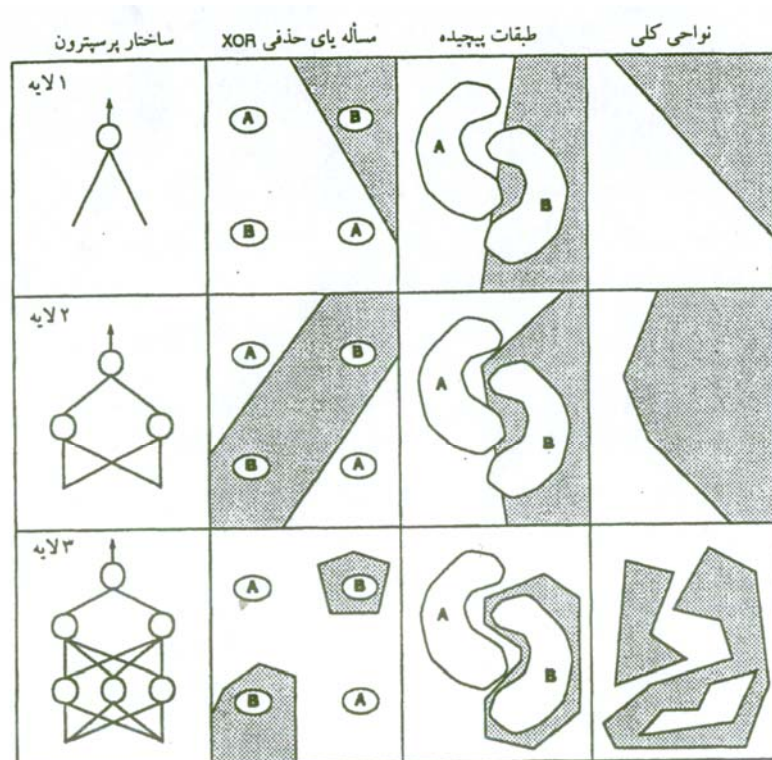
مثال هایی از ناحیه های محدب باز و بسته.



شکل (۴-۷)

مثال هایی از ناحیه دلخواه که از ترکیب ناحیه های محدب ایجاد شده است.

بنابراین با استفاده از سه لایه پرسپترون می توان هر شکل پیچیده ای را ایجاد کرد و انواع مختلف طبقات را تفکیک کرد. پیچیدگی اشکال محدود به تعداد گره ها خواهد بود. زیرا هر گره سطح جداگانه ای را در فضا ایجاد می کند. توانایی ما در ایجاد اشکال دلخواه بدان معنی است که تنها سه لایه برای تفکیک هر نوع فضایی مناسب می کند و هرگز نیازی به لایه های بیش تر نیست، قضیه ای که به قضیه کولموگوروف<sup>۱</sup> معروف است. این قضیه را می توان با ریاضیات پیچیده اثبات کرد لیکن ما در این جا تنها به ذکر آن اکتفا می کنیم. شکل ۴-۸ توانایی پرسپترون ها را در تفکیک فضاهای دلخواه نشان می دهد.



شبکه های عصبی و توانایی تفکیک فضاهای داخلی.

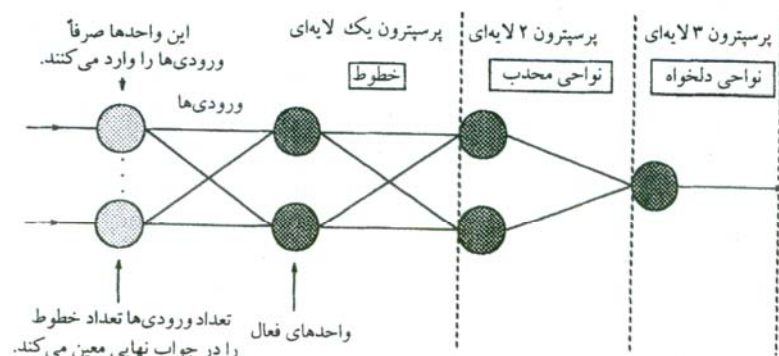
شکل (۴-۸)

متون شبکه های عصبی در مورد تعریف تعداد لایه های شبکه هماهنگ نیست. عده ای از نویسندگان تعداد لایه های ضرایب وزنی متغیر و عده ای دیگر تعداد لایه های گره ها به عنوان

1 Kolmogrov Theorem

تعداد لایه های شبکه تعریف می کنند. این عدم هماهنگی اغلب باعث سردرگمی است زیرا گره های اولین لایه یعنی لایه ورودی تنها مقادیر ورودی را به گره های لایه بعد منتقل می کند و خود هیچ عمل جمع و آستانه ای را انجام نمی دهد. این سردرگمی هنگامی بیش تر می شود که بعضی اصولاً گره های لایه ورودی را به طور کلی از قلم انداخته و آن را در شبکه ترسیم نمی کنند. با توضیح زیر سعی می شود مطلب تا اندازه ای روشن شود: یک شبکه چند لایه ای ورودی هایی را دریافت می کند. ورودی ها در لایه ای که هیچ گونه عمل جمع و آستانه ای را انجام نمی دهد توزیع می گردد، به طوری که هر گره فقط یک ورودی را دریافت می کند. چون هر گره تنها یک ورودی دارد عمل جمع ورودی گره ها در این مرحله اصولاً موضوعیت نخواهد داشت. این ورودی ها سپس از طریق لایه ای از خطوط اتصال با ضرایب وزنی قابل تنظیم به لایه ای از گره های پرسپترون مانند منتقل می گردند. در این لایه هر گره ورودی های خود را جمع و عمل آستانه را انجام می دهد. این لایه قادر است که در فضای الگوهای خطوطی را ایجاد کند. خروجی این لایه از طریق لایه دیگری از خطوط اتصال با ضرایب وزنی قابل تنظیم به لایه دیگری از گره های پرسپترون منتقل می گردند. خروجی این لایه است که نواحی محدبی را در فضا ایجاد می کند. خروجی این لایه نیز از طریق لایه دیگری از خطوط اتصال به لایه دیگری از واحدهای پرسپترون مانند منتقل می شود و مجدداً عمل جمع و آستانه در هر گره از این لایه صورت می پذیرد. خروجی این لایه است که نهایتاً هر شکل دلخواهی را در فضای الگوها تفکیک می کند. با شمارش تعداد لایه های ضرایب وزنی قابل تنظیم (یا تعداد لایه های پرسپترون مانند فعال) شبکه ما شبکه ای سه لایه ای تعریف خواهد شد. لیکن اگر لایه ورودی گره های غیر فعال را نیز در نظر بگیریم شبکه ما شبکه ای چهار لایه ای تعریف می شود. روند کلی پیروی از تعریف اول است زیرا گویاتر به نظر می رسد. خلاصه این بحث در شکل ۴-۹ آمده است.





شکل (۴-۹) نحوه تشکیل محدوده‌های فضا توسط تعداد مختلف لایه‌های پرسپترون.

تابع غیر خطی سیگموئید در پرسپترون‌های چند لایه ای خطوط مستقیم سطوح تصمیم در فضا را به منحنی‌های صاف تبدیل می‌کند، به صورتی که فضا‌های ایجاد شده اکنون با منحنی‌های صافی محصور می‌شوند، لیکن موضوع پیچیدگی اشکال و نیاز به شبکه‌های دو لایه ای و سه لایه ای کماکان باقی می‌ماند.

می‌توانیم موضوع طبقه بندی الگوها را به طریق دیگر مورد بحث قرار دهیم. هر الگوی ورودی باید به یکی از طبقه‌های مورد نظر ما تعلق داشته باشد. بنا براین هر الگوی ورودی به طبقه ای معین نگاشته می‌شود. این نگاشت را می‌توان به صورت تابعی در نظر گرفت که الگوهای ورود را به طبقات خروجی صحیح تبدیل می‌کند. شبکه ای را به خوبی آموزش دیده می‌نامیم که بتواند این عمل نگاشت را به درستی انجام دهد. در واقع هر تابعی را می‌توان صرف نظر از میزان پیچیدگی آن با حد اکثر سه لایه پرسپترون نمایش داد. ورودی‌ها از سه لایه ورودی، پنهان و خروجی می‌گذرند، همان‌طور که قبلاً ذکر شد این قضیه به قضیه کولموگوروف معروف است این قضیه از اهمیت بالایی برخوردار است زیرا ثابت می‌کند که آنچه را می‌توان با چهار لایه یا بیشتر پرسپترون انجام داد با سه لایه پرسپترون نیز قابل حصول است لیکن متأسفانه قضیه کولموگوروف هیچ اشاره ای به تعداد گره‌ها در شبکه و نحوه اتصال آنها و یا تنظیم ضرایب وزنی بین گره‌ها نمی‌کند.

## برخی توانایی‌ها و ضعف‌های شبکه‌های عصبی

### تعمیم دهی<sup>۱</sup>

یکی از ویژگی‌های عمده شبکه‌های عصبی توانایی آن‌ها در تعمیم دهی است، به صورتی که می‌توانند الگوهای را نیز که در جریان آموزش مشاهده نکرده‌اند به درستی طبقه‌بندی کنند. پرسپترون پند لایه‌ای این عمل را با فراگیری ویژگی‌های برجسته الگوهای ورودی در جریان آموزش و کد کردن آن‌ها در واحدهای درونی خود انجام می‌دهد. بدین صورت الگوهای ناشناخته بر اساس اشتراک ویژگی‌های برجسته آن‌ها با سایر الگوها طبقه‌بندی می‌شوند. در نتیجه آموزش توسط نمونه‌ها امکان پذیر می‌شود، زیرا تنها لازم است مجموعه محدودی از الگوها به شبکه آموزش داده شود. شبکه الگوهای مشابه را با تعمیم خاصیت‌هایی که فرا گرفته است طبقه‌بندی می‌کند. لیکن ورودی‌های متفرقه و ناخواسته نیز بر اساس شباهت صرف آن‌ها با الگوهای اصلی طبقه‌بندی می‌شوند. این خاصیت تعمیم است که شبکه‌های پرسپترون چند لایه‌ای را نسبت به سایر روش‌های الگوها و سیستم‌های خبره در مقابله با مسائل توانمندتر می‌کند. به طور کلی، شبکه‌های عصبی در درون یابی<sup>۲</sup> عملکرد خوبی دارند لیکن در برون یابی<sup>۳</sup> چندان قوی نیستند. آن‌ها می‌توانند الگوهای را که قبلاً دیده‌اند تشخیص دهند و الگوهای میانی را که نیز ندیده‌اند در حیطه توانایی خود قرار دهند ولیکن الگوهای که خارج از گستره الگوهای آموزش داده اولیه شبکه است را نمی‌توانند به خوبی طبقه‌بندی کنند. زیرا نمونه‌های چندانی را برای مقایسه در اختیار ندارد. به عبارت دیگر اگر الگوهای تازه‌ای بین دو الگوی آشنا قرار گرفته باشد شبکه آن را به عنوان نمونه‌ای از الگوی قبلی که شبکه دیده است مشابهت نداشته باشد طبقه‌بندی به خوبی انجام می‌گیرد.

---

1 Generalization

2 Interpolation

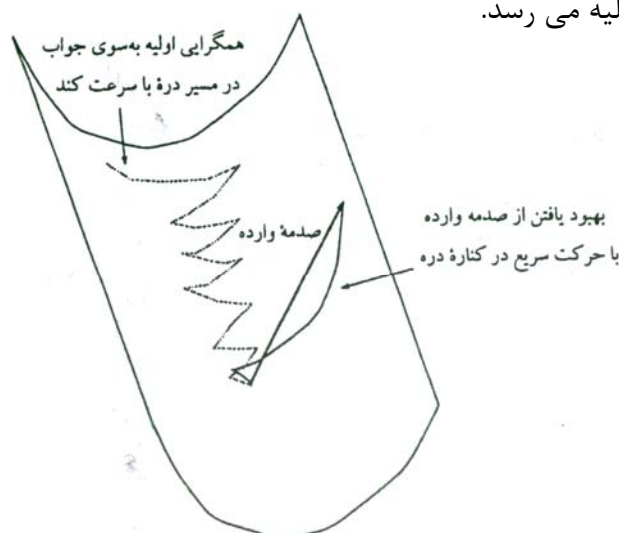
3 Extrapolation

## تحمل نقص<sup>۱</sup>

پرسپترون های چند لایه ای ذاتاً در مقابل نقایص احتمالی چشم پوش هستند، زیرا عناصر تشکیل دهنده آنها به صورت موازی و توزیع شده عمل می کنند و هر گره به طور جداگانه در خروجی نهایی شبکه مشارکت دارد و اگر گره ای و یا ضرایب وزنی آن گره حذف شود یا صدمه ببیند توان شبکه تا اندازه ای از نظر کیفی کاهش خواهد یافت، لیکن با توجه به طبیعت توزیعی اطلاعات میزان صدمات وارده باید بسیار شدید باشد تا خروجی شبکه به طور محسوسی تضعیف شود.

پرسپترون ها نسبت به اغتشاش<sup>۲</sup> نیز چشم پوش هستند، زیرا قادرند نمونه های آموخته شده را به نمونه های مغشوش تعمیم دهند.

صدمات وارده به شبکه به علت از دست دادن چند گره و یا وجود اغتشاش در داده های آموزشی را معمولاً می توان با آموزش مجدد بر طرف کرد و در این موارد بهبود اغلب سریع می باشد شکل ۴-۱۰ این مطلب را به خوبی نشان میدهد. همگرایی به سوی جواب اولیه در سطح دره صورت می گیرد و بنا براین بسیار کند است. صدمه وارده شبکه را متلاطم می کند، لیکن شبکه به سرعت در شیبی تند به جواب صحیح نزدیک می شود، بنابراین وقتی آموزش مجدد صورت می گیرد شبکه در طول شیب تند حرکت کرده و به جواب اولیه می رسد.



شکل (۴-۱۰)

چگونه می توان صدمه وارده به شبکه را به سرعت بهبود بخشید.

1 Fault tolerance

2 Noise

## مشکلات آموزش

شبکه گاه در یک جواب پایدار متوقف می شود که خروجی درستی را ارائه نمی دهد . در این موارد تابع انرژی شبکه به یک کمینه محلی<sup>۱</sup> رسیده است . به عبارت دیگر شبکه در هر جهت در بستر انرژی حرکت کند مقدار انرژی آن از موقعیت کنونی آن بیشتر است. ممکن است تنها یک تکان کافی باشد که شبکه را از گیر موقعیت کنونی رهایی بخشیده آن را به سوی کمینه عمیق تری هدایت کند لیکن شبکه هیچ گونه آگاهی نسبت به این موضوع ندارد زیرا آموزش آن با تعقیب تابع انرژی در تند ترین شیب ممکن تا رسیدن به عمق یک چاه محلی صورت می گیرد و در این وضعیت دیگر در هیچ جهتی انرژی کاهش نمی یابد. برای کاهش تعداد این موارد راه کارهایی وجود دارد که در زیر آمده است.

## کاهش ضریب بهره<sup>۲</sup>

اگر نرخ که طبق آن ضرایب وزنی تغییر می یابند به تدریج کاهش یابد آن گاه الگوریتم نزول در شیب می تواند به جواب های بهتری دست یابد اگر ضریب بهره در ابتدا بزرگ فرض شود میزان حرکت به سوی جواب در ضرایب وزنی و فضای انرژی زیاد خواهد بود به تدریج که نرخ ضریب بهره کاهش می یابد و حرکت در شیب کمتر و کم تر می شود ضرایب وزنی شبکه به حالت کم ترین انرژی رسیده و در آنجا آرامش می یابد بدون اینکه ناخواسته از آن حالت بیرون آیند این روش به شبکه امکان می دهد که از حالت های کمینه محلی در ابتدا اجتناب کند و نهایتاً امید می رود در کمینه های عمیق تری بدون نوسانات شدید قرار گیرد به هر صورت کاهش ضریب بهره به معنی افزایش زمان همگرایی شبکه خواهد بود .

---

1 Local Minimum

2 Gain term

## افزایش تعداد گره های داخلی

می توان تصور کرد که کمینه محلی هنگامی اتفاق می افتد که دو یا چند طبقه متمایز به صورت یک طبقه واحد دسته بندی شوند . این وضع باعث ضعف در تعیین نمادهای داخلی شبکه خواهد بود بنا براین افزایش تعداد گره های داخلی امکان می دهد که ورودی ها بهتر کد شوند و امکان بروز کمینه محلی را کاهش می دهد.

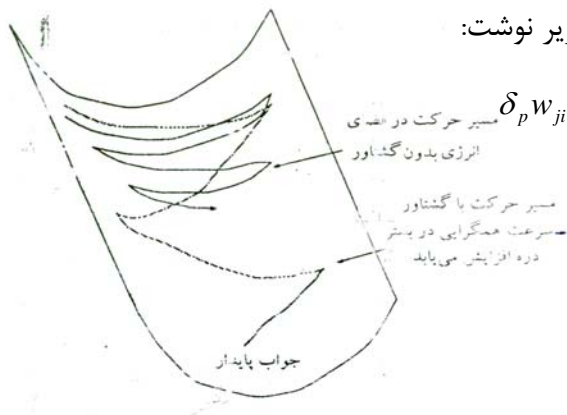
## عبارت گشتاور<sup>۱</sup>

تحرك بیشتری را می توان در میزان تغییر ضرایب وزنی شبکه با اضافه نمودن عبارت گشتاور در معادله تنظیم ضرایب ایجاد کرد این عبارت در ابتدا باعث تغییرات زیاد در ضرایب وزنی می شود لیکن با کاهش مقدار تغییرات به علت عبارت گشتاور تغییراتی را فراتر از افزایش های محلی در تابع انرژی جهت روند کلی کاهش انرژی باعث می شود . گشتاور در تسریع همگرایی در شیب های کم عمق بسیار موثر است زیرا باعث می شود شبکه در جهت رسیدن به جواب در مسیر سرازیری سرعت گیرد بستر انرژی ممکن است دارای دره های تنگ و طولانی با شیب کم باشد که نهایتاً به نقطه کمینه ختم می شوند همگرایی در طول این دره ها کند است زیرا شیبی که باید طی شود بسیار کند است و الگوریتم معمولاً به صورتی که در شکل ۱۱-۲ نشان داده شده در حالی که بسوی جواب حرکت می کند در عرض دره ها نوسان می کند . در این حالت تسریع در حرکت بدون افزایش احتمال پریدن و گذشتن از نقطه کمینه دشوار است . افزایش عبارت گشتاور در این موارد

بسیار موثر است . عبارت گشتاور را می توان به صورت زیر نوشت:

$$\delta_p w_{ji}(t+1) = w_{ji}(t) + \eta \delta_{pj} o_{pi} + \alpha (w_{ji}(t) - w_{ji}(t-1))$$

در حالی که  $\alpha$  عامل گشتاور و  $0 < \alpha < 1$  است.



شکل (۱۱-۴)

## افزایش اغتشاش

با افزایش اغتشاش الگوریتم نزول در شیب به لرزش در آمده و از خط تند ترین شیب خارج می شود. این عمل اغلب باعث می شود که سیستم از گرفتار شدن در کمینه های محلی رهایی یابد مزیت این روش در این است که به زمان محاسباتی چندان بیشتری احتیاج ندارد و بنابراین سرعت آن از الگوریتم نزول در شیب چندان کند تر نیست.

## سایر مشکلات آموزش

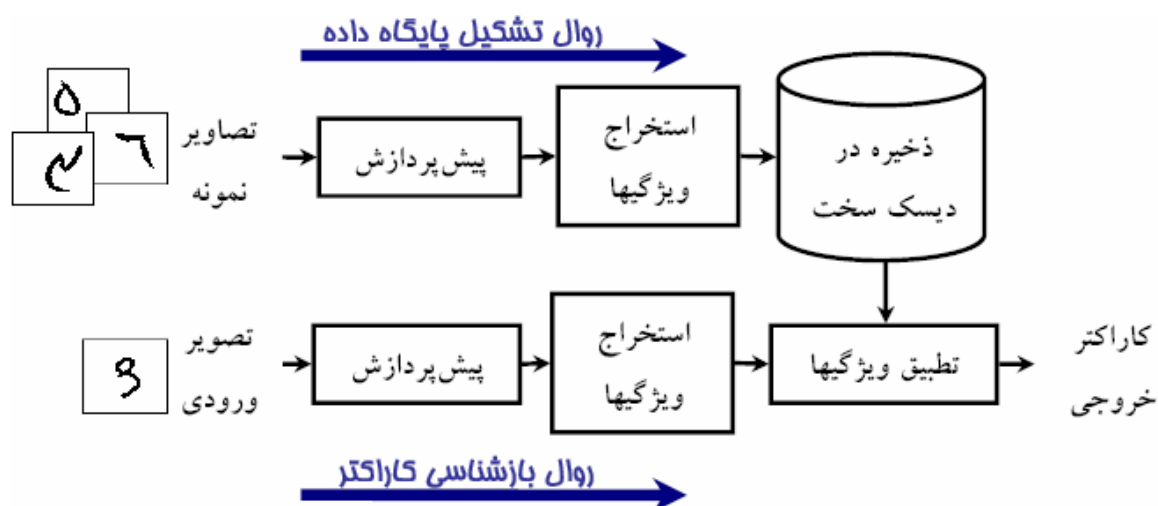
از انتقادات عمده ای که بر پرسپترون چند لایه ای وارد می شود یکی این است که آموزش آن نیاز به تکرار فراوان الگو های ورودی و تکرار محاسبات پس انتشار خطای هر الگو دارد، تا اینکه نهایتاً شبکه به وضع پایداری برسد. همگرایی در بسترهای پیچیده در روش نزول در شیب به علت پیچیدگی سطح انرژی ذاتاً کند است. افزون عبارت گشتاور، همان طور که بحث شد، اغلب همگرایی را سرعت می بخشد. روش دیگر تغییر ضریب بهره  $\eta$  است. راه کار دیگری که به ممانعت از ایجاد کمینه های محلی کاذب کمک می کند استفاده از مشتق دوم در الگوریتم نزول در شیب است. لیکن افزایش میزان دقت در تعیین مسیر نزول انرژی که با این راه حل به دست می آید با افزایش میزان پیچیدگی محاسباتی آن جبران می شود.

## فصل پنجم پیاده سازی یک نرم افزار تشخیص خودکار ارقام فارسی

پس از آشنایی با مفاهیم تکنیک آنالیز اجزای اصلی و شبکه عصبی در فصلهای گذشته حال به بررسی روش استفاده شده جهت شناسایی ارقام دست نویس فارسی در این پروژه می پردازیم. این برنامه به زبان برنامه نویسی MATLAB نوشته شده است و قادر به شناسایی ارقام فارسی با دقت و سرعت بالایی می باشد. در ادامه به شرح این برنامه می پردازیم.

### ۵-۱ بلوک دیاگرام سیستم

بلوک دیاگرام سیستم پیشنهادی در شکل زیر نشان داده شده است.



ابتدا برای هر رقم ۸۶۰ تصویر دستنویس از دانشجویان دانشگاههای تهران تهیه شد. این ارقام پس از باینری شدن و حذف نویز از لحاظ اندازه همگی به اندازه  $40 \times 40$  نرمالیزه شده، در یک ماتریس ۴ بعدی ذخیره شدند و بعنوان دیتابیس مورد استفاده قرار گرفتند.

## ۵-۲ مرحله آموزش<sup>۱</sup>:

ابتدا تکنیک آنالیز اجزای اصلی به بخشی از دیتا بیس مورد نظر اعمال شد، که در این مورد برای هر رقم از ۲۰۰ نمونه جهت آموزش استفاده شد. پس از انتخاب ۲۰۰ تصویر مختلف از هر رقم که تقریباً تمام حالات مورد نظر به آن را پوشش می داد و ذخیره آنها در یک ماتریس خاص، تکنیک آنالیز اجزای اصلی به این ماتریس اعمال شد و ماتریس انتخابی از بردارهای ویژه آن بدست آمد که بردار ویژگی نامیده می شود. حال پس از اینکه هریک از تصاویر مرحله آموزش از ماتریس میانگین کم شدند،

تحت ماتریس بردارهای ویژه نگاشت یافته و ویژگی مربوط به آنها حاصل می شود که این ویژگیها جهت آموزش شبکه عصبی در یک ماتریس ذخیره شده و استفاده شده است. شبکه عصبی استفاده شده در این پروژه یک شبکه عصبی MLP ۳ لایه با خاصیت پس انتشار خطا و روش آموزش مبتنی بر کمینه کردن مجموع مربعات خطا بود.

## ۵-۳ مرحله آزمایش<sup>۲</sup>:

در این مرحله پس از اینکه مقادیر مربوط به ماتریس هر تصویر از ماتریس میانگین کم شد، این ماتریس تحت ماتریس بردار ویژگی نگاشت یافته و بعنوان ورودی شبکه عصبی آموزش دیده ما استفاده می شود.

## ۵-۴ نتایج

پس از آموزش شبکه عصبی مربوطه با ۲۰۰۰ نمونه ( ۲۰۰ نمونه از هر رقم)، جهت آزمایش نیز از ۲۰۰۰ نمونه (از هر رقم ۲۰۰ نمونه) استفاده کردیم که درصد کل خروجیهای درست ۹۵٪ بوده است. زمان لازم جهت تشخیص یک رقم به طور میانگین ۹۰ میلی ثانیه بوده است. در ادامه نمونه هایی از تصاویر را مشاهده و پس از آن نتایج بدست آمده برای دیگر شبکه ها را بررسی می کنیم.

---

1 Train  
2 Test

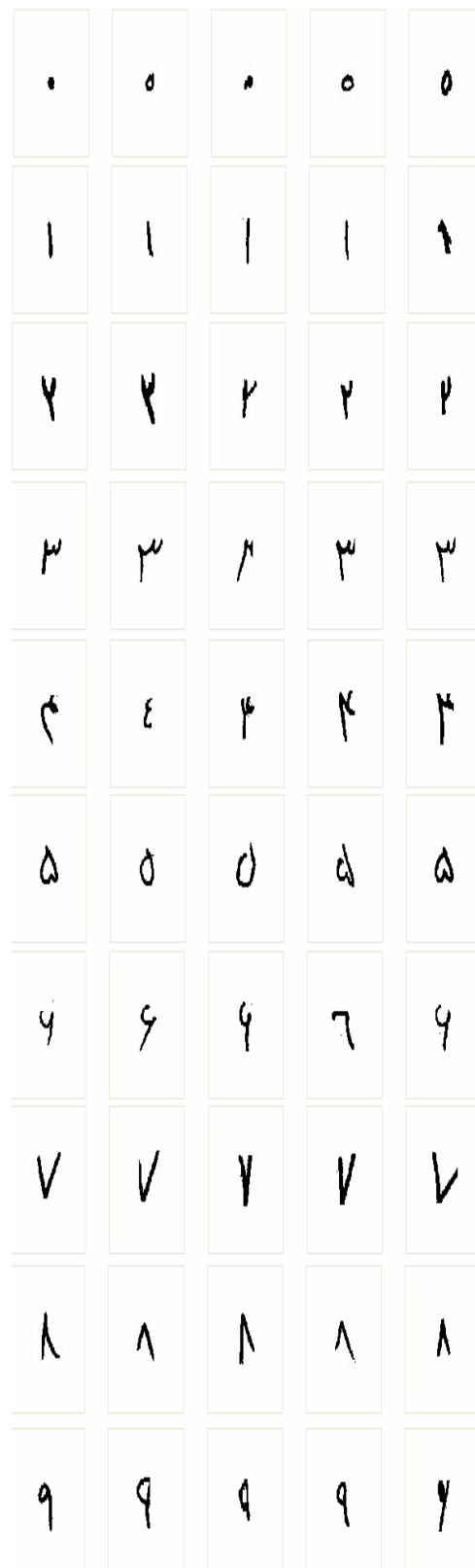


در زیر بخشی از تصاویر مربوط به مرحله آموزش در شکل (۵-۱) و بخشی از تصاویر مربوط به مرحله آزمایش را در شکل (۵-۲) نشان داده شده اند.

شکل (۵-۲)



شکل (۵-۱)

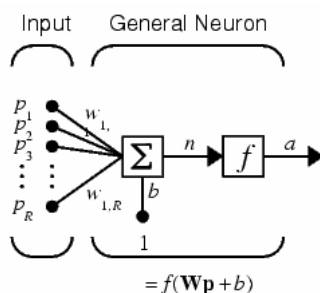


در ادامه به بعضی نتایج بدست آمده برای شبکه عصبی های مختلف اشاره شده است.

سعی شده فقط حالات مربوط به درصد های بالای ۹۰ ذکر شود.

تعداد لایه های شبکه استفاده شده	تعداد node هر لایه	تابع شبکه	تعداد تصویر train جهت هر رقم	درصد خروجی از ۲۰۰ تصویر تست
۳	10-300 -450	'logsig' 'logsig' 'logsig'	۱۰۰	۹۰
۳	10-300 -450	'logsig' 'logsig' 'logsig'	۲۰۰	۹۵,۲
۳	10-500 -450	'logsig' 'logsig' 'logsig'	۲۰۰	۹۱
۳	10-300 -450	'logsig' 'logsig' 'logsig'	۵۰۰	۹۷
۳	10-300 -500	'logsig' 'logsig' 'logsig'	۵۰۰	۹۶,۹

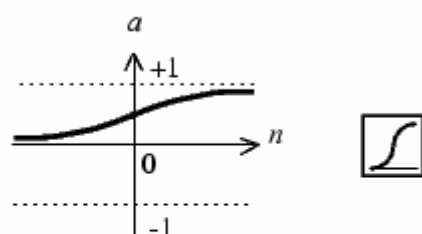
تعداد لایه های شبکه استفاده شده	تعداد node هر لایه	تابع شبکه	تعداد تصویر train جهت هر رقم	درصد خروجی از ۲۰۰ تصویر تست
۴	10-70 -180-250	'logsig' 'logsig' 'logsig' 'logsig'	۲۰۰	۸۹,۳
۴	10-70 -180-260	'logsig' 'logsig' 'logsig' 'logsig'	۲۰۰	۹۲,۶
۴	10-70 -180-300	'logsig' 'logsig' 'logsig' 'logsig'	۲۰۰	۹۳,۵
۴	10-70 -200-300	'logsig' 'logsig' 'logsig' 'logsig'	۲۰۰	۹۴,۵
۴	10-70 -240-300	'logsig' 'logsig' 'logsig' 'logsig'	۲۰۰	۹۳,۶
۴	10-70 -240-350	'logsig' 'logsig' 'logsig' 'logsig'	۲۰۰	۹۳,۶
۴	10-70 -200-300	'logsig' 'logsig' 'logsig' 'logsig'	۵۰۰	۹۶,۶



تابع شبکه ( $f$ )، در یک نرون نمونه در شکل بالا نشان داده شده است

منظور از 'logsig'، تابع شبکه logarithm sigmoid است که شکل

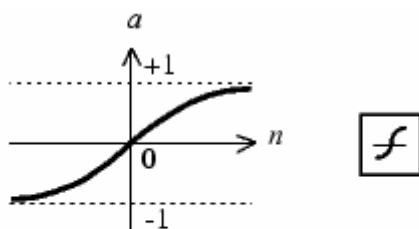
آن در پایین ملا حظه می شود.



Log-Sigmoid Transfer Function

منظور از 'tansig'، تابع شبکه tangent sigmoid است که شکل آن در پایین ملا حظه می

شود.



Tan-Sigmoid Transfer Function

یادآور می شود که :

۱. مقادیر حدودی تعداد nod های هر شبکه به طور تجربی بدست آمده است.

۲. تعداد eigenvector های استفاده شده به طور تجربی ، ۵۰ تا اول بوده است.

۳. مشاهدات عملی ثابت کرده که تابع شبکه logsig بهترین نتیجه را ارائه می دهد.

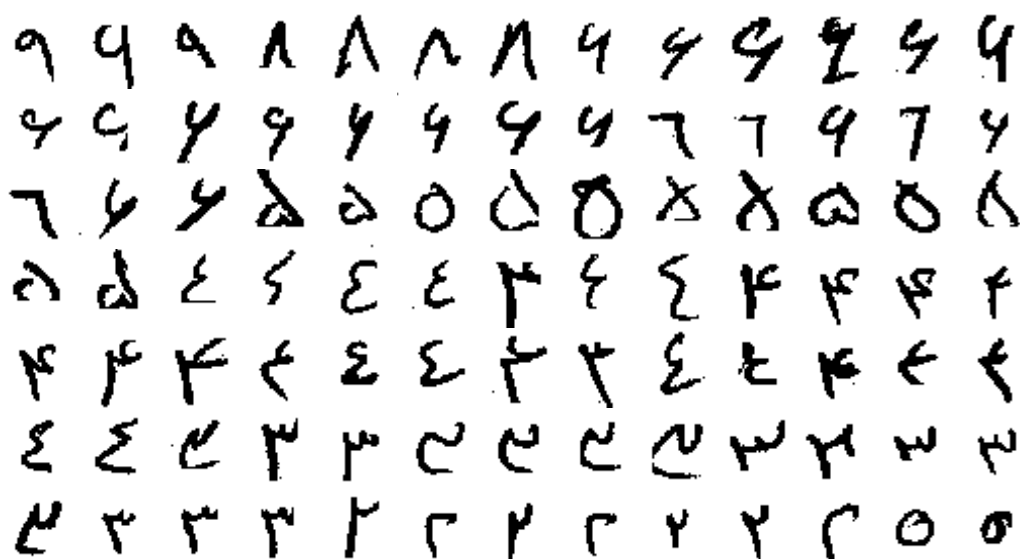
و کاربرد تابع شبکه tansig به تنهایی یا با دیگر توابع عملکرد ضعیفتری را ارائه می دهد.

۴. لازم به ذکر است الگوریتم استفاده شده در آموزش شبکه همان الگوریتم توضیح داده شده در

فصل چهارم، یعنی الگوریتم gradient descent بوده است.

در ادامه برخی تصاویر را که نرم افزار مورد نظر در شناسایی آنها دچار مشکل شده ، مورد بررسی

قرار می دهیم.



همانطور که ملاحظه می شود اشتباه نرم افزار مربوطه در برخی موارد منطقی و قابل بخشش !

است،

ولی در بعضی موارد دیگر ملاحظه می شود که نرم افزار مربوطه بدلیل نحوه آموزش اعمال شده

دچار اشتباه شده است، چون در زمان آموزش شبکه عصبی مزبور ، داده های آموزش از همین

دیتابیس انتخاب شده بودند و تعدادی داده نویزی به شبکه مزبور آموزش داده شده است.

در جدول زیر درصد شناسایی هر کدام از ارقام آمده است.

رقم	۰	۱	۲	۳	۴	۵	۶	۷	۸	۹
درصد	۹۸,۵	۱۰۰	۹۳	۸۹	۸۶	۹۲	۸۸	۹۹,۵	۹۷	۹۸,۵

## فصل ششم مراجع

### مراجع لاتین

- 1 Abulhaiba S. H., Mahmood S. A., Green R. J., "Recognition of handwritten cursive Arabic characters", IEEE Trans. On PAMI, Vol. 16, No. 6, pp. 664-671, 1994.
- 2 Altuwaijir M. and Bayoumi M., "Arabic text recognition using neural networks", Proc. Int. Symp. On Circuit and Systems (ISCAS'94), pp. 415-418, 1994.
- 3 Al-Usefi H. and Udpa S.S., "Recognition of Arabic characters", IEEE PAMI, Vol. 14, No. 8, pp. 853-857, 1992.
- 4 Arica N. and Yarman-Vural F.T., "One dimensional representation of two dimensional information for HMM based handwritten recognition", Pattern Recognit. Lett., vol. 21, No. 6-7, pp. 583-592, 2000.
- 5 Arica Nafiz, Vural Fatoş Yarman, "A new scheme for off-line handwritten connected digit recognition", 14th Int. Conf. Pattern Recognit., Brisbane, Australia, pp. 1127-1131, 1998.
- 6 Arica Nazif, Yamin-Vural Fatos T., "An overview of character recognition based focused on off-line handwriting", IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews, Vol. 31, No. 2, May 2001.
- 7 Azmi R., Kabir E., "A recognition algorithm for hand-printed farsi characters", Proc. OF ICT96, Istanbul, Turkey, pp. 852-855, 1996.
- 8 Badie K., Shimora M., "Machine recognition of Arabic cursive script", Pattern Recognition in Practice, Proc. of International Workshop, Amsterdam, Netherlands, pp. 315-323, 1980.
- 9 H. Baird, "Document image defect models", IAPR Workshop on Syntactic and Structural Pattern Recognition, pages 38-46, Murray Hill, NJ, June 1990.
- 10 Baird H.S., Jones S.E., and Fortune S.J., "Image segmentation by shape-directed covers", Proc. of the 10th International Conference on Pattern Recognition, Atlantic City, NJ, June, 17-21, 1990.
- 11 Bozinovic R. M. and Srihari S. N., "Off-line cursive script word recognition", IEEE Trans. Pattern Anal. Machine Intell., vol. 11, pp. 68-83, Jan. 1989.
- 12 Casey R. G. and Lecolinet E., "A survey of methods and strategies in character segmentation", IEEE Trans. Pattern Anal. Machine Intell., vol. 18, pp. 690-706, July 1996.
- 13 Casey R., Ferguson D., Mohiuddin K. and Walach E., "Intelligent Forms Processing System", Machine Vision and Applications, vol. 5, No. 3, pp. 143-155, 1992.
- 14 Chen M. Y., Kundu A., and Zhou J., "Off-line handwritten word recognition using a hidden Markov model type stochastic network", IEEE Trans. Pattern Anal. Machine Intell., vol. 16, pp. 481-496, May 1994.
- 15 Correia S.E.N., De Carvalho J.M., Sabourin, R., "Human-perception handwritten character recognition using wavelets", Proceedings XV Brazilian Symposium on Computer Graphics and Image Processing, pp.404, Oct. 2002.

- 16 Correia S.E.N., De Carvalho J.M., Sabourin, R., "On the performance of wavelets for handwritten numerals recognition", 16th International Conference on Pattern Recognition, Vol. 3, pp. 127 -130, Aug. 2002.
- 17 Dengel A., Bleisinger R., Hoch R., Fein F., Hones F., "From Paper to Office Document Standard Representation", IEEE Computer, vol. 25, No. 7, pp. 63-67, 1992.
- 18 El-Sheikh T.S., Guindi R. M., "Computer recognition of Arabic cursive script", Pattern Recognition 21, pp. 293-302, 1988.
- 19 Esposito F., Malerba D., Semeraro G., Annese E., and Scafuro G., "An experimental page layout recognition system for office document automatic classification: an integrated approach for inductive generalization", IEEE International Conference on Pattern Recognition (Atlantic City, N J), pp. 557-562, 1990.
- 20 GE Yong, HUO Quiang, FENG Zhi-Dan, "Offline recognition of handwritten Chinese characters using Gabor features, CDHMM modeling and MCE training", IEEE 2002.
- 21 Gilloux E., "Hidden Markov models in handwriting recognition", in Fundamentals in Handwriting Recognition. ser. NATO ASI F, S. Impedovo, Ed. New York: Springer-Verlag, vol. 124, 1994.
- 22 Guillevic D. and Suen C. Y., "Cursive script recognition: A sentence level recognition scheme", in Proc. Int. Workshop Frontiers in Handwriting Recognit., pp. 216–223, 1994.
- 23 Guillevic D. and Suen C. Y., "HMM–KNN word recognition engine for bank cheque processing", in Proc. 14th Int. Conf. Pattern Recognit., Brisbane, Australia, 1998, pp. 1526–1529.
- 24 Hu Peifeng, Zhao Yannan, Yang Zehong, Wang Jiaqin, "Recognition of gray character using gabor filters", Proceedings of the Fifth International Conference on Information Fusion, Vol.1 , pp. 419 -424, July 2002.
- 25 Huang Lei, Huang Xiao, "Multiresolution recognition of offline handwritten Chinese characters with wavelet transform", Sixth International Conference on Document Analysis and Recognition, pp. 631 -634, Sept. 2001.
- 26 Jain A. K. and Yu B., "Document representation and its application to page decomposition", IEEE Trans. Pattern Anal. Machine Intell., vol. 20, pp. 294–308, Mar. 1998.
- 27 Jinying Chen, Yijiang Jin, Shaoping Ma, "A learning algorithm detecting the similar Chinese characters' boundary based on unequal-contraction of dimension", Proceedings of the 3rd World Congress on Intelligent Control and Automation, vol.4, pp. 2765 -2769, July 2000.
- 28 Kasturi R., O’Gorman L., and Govindaraju V., "Document Image Analysis: A Primer", Sadhana, special issue on Document Processing, accepted, in press, 2002.
- 29 Khorsheed M.S., Clocksin W.F., "Multi-font Arabic word recognition using spectral features", 15th International Conference on Pattern Recognition, vol. 6, pp. 3574 - 3577, June 2000.
- 30 Kornai András, Mohiuddin K.M., Connell Scott D., "Recognition of cursive writing on personal checks", in Proc. Int. Workshop Frontiers Handwriting Recognit., Essex, U.K., pp. 373–378, 1996.

- 31 Kunango T., Haralick R., and Phillips I., "Nonlinear local and global document degradation models", *Int. J. Imaging Syst. Technol.*, vol. 5, no. 4, pp. 274–282, 1994
- 32 Lam L., Lee S. W., and C. Y. Suen, "Thinning methodologies—A comprehensive survey", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 869–885, Sept. 1992.
- 33 Lam Louisa, Suen Ching Y., "An evaluation of parallel thinning algorithms for character recognition", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 914–919, Sept. 1995.
- 34 Lee S.W., Lee D. J., and Park H. S., "A new methodology for gray-scale character segmentation and recognition", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, pp. 1045–1050, Oct. 1996.
- 35 Madhvanath S., Kim G., and Govindaraju V., "Chaincode contour processing for handwritten word recognition", *IEEE Pattern. Anal. Machine Intell.*, vol. 21, pp. 928–932, Sept. 1999.
- 36 Mahmoud S. A., Abuhabia I., and R. J. Green, "Skeletonization of Arabic characters using clustering based skeletonization algorithm", *Pattern Recognit.*, vol. 24, no. 5, pp. 453–464, 1991.
- 37 Mahmoud S., "Arabic Character Recognition using Fourier descriptors and character encoding", *Pattern Recognition*, Vol. 27, No. 6, pp. 815-824, 1994.
- 38 Mandal S., Chowdhury S.P., Das A.K. and Chanda B., "Automated Detection and Segmentation of Table of Contents Page from Document Images", *7th International Conference on Document Analysis and Recognition (ICDAR'03)*, pp. 398-402, Aug. 2003.
- 39 Mandal S., Chowdhury S.P., Das A.K. and Chanda B., "Automated Detection and Segmentation of Table of Contents Page and Index Pages from Document Images", *12th International Conference on Image Analysis and Processing*, pp. 213-218, Sept. 2003.
- 40 Manjunath B.S. and W. Ma, "Texture features for browsing and retrieval of image data", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 8, pp. 837-842, August 1996.
- 41 Menhaj M.B., Adab M., "Simultaneous segmentation and recognition of Farsi/Latin printed texts with MLP", *International Joint Neural Networks: IJCNN '02*, Vol. 2, pp. 1534 -1539, May 2002.
- 42 Mohamed M. and Gader P., "Handwritten word recognition using segmentation-free hidden Markov modeling and segmentation based dynamic programming techniques", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, pp. 548–554, May 1996.
- 43 Mori S., Suen C. Y., and Yamamoto K., "Historical review of OCR research and development", *Proc. IEEE*, vol. 80, pp. 1029–1057, July 1992.
- 44 Mowlaei A., Faez K., Haghighat A.T., "Feature extraction with wavelet transform for recognition of isolated handwritten Farsi/Arabic characters and numerals", *14th International Conference on Digital Signal Processing*, Vol. 2, pp. 923-926, July 2002.
- 45 Nagy George, Seth Sharad, Viswanathan Mahesh, "A Prototype Document Image Analysis System for Technical Journals", *Computer*, vol 25 no 7, pp 10-22, July 1992.
- 46 Nagy, G., "Twenty years of document image analysis in PAMP", *IEEE Transactions*

- on Pattern Analysis and Machine Intelligence, Vol. 22, pp. 38-62, Jan 2000.
- 47 Niblack W., *An Introduction to Digital Image Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1986.
- 48 O'Gorman Lawrence, "The Document Spectrum for Structural Page Layout Analysis", IEEE Trans. Pattern Analysis and Machine Intelligence, Nov., 1993.
- 49 O'Gorman Lawrence, Kasturi Rangachar, *Document Image Analysis*, IEEE Computer Society Press, Los Alamitos, 1995.
- 50 Srihari Sargur, Lam Stephen, Govindaraju Venu, Srihari Rohini and Hull Jonathan, "Document Understanding: Research Directions", Technical report CEDAR-TR-92-1, State University of New York at Buffalo: Center of Excellence for Document Analysis and Recognition (CEDAR), May 1992.
- 51 Suen C. Y., Tappert C. C., and Wakahara T., "The state of the art in online handwriting recognition", IEEE Trans. Pattern Anal. Machine Intell., vol. 12, pp. 787-808, Aug. 1990.
- 52 Tayel M., Shalaby H., Saleh, H., "Winner-take-all neural network for visual handwritten character recognition", Thirteenth National Radio Science Conference: NRSC '96, pp. 239 -249, March 1996.
- 53 Theodoridis Sergios, Koutroumbas Konstantinos, Smith Ricky, *Pattern Recognition*, Academic Press, 1st edition, January 15, 1999.
- 54 Tomio, Yamashita Akio, Itoh Nobuyasu, Kobayashi Yoshinao, Katoh Shin, Toyokawa Kazuharu and Takahashi Hiroyasu, "DRS: a workstation-based document recognition system for text entry", IEEE Computer Vol. 25, No. 7, pp. 67-71, 1992.
- 55 Trier Ø. D. and Jain A. K., "Goal directed evaluation of binarization methods," IEEE Trans. Pattern Anal. Machine Intell., vol. 17, pp. 1191-1201, Dec. 1995.
- 56 Trier O. D., Jain A. K. and Taxt T., "Feature extraction methods for character recognition - A survey", Pattern Recognition 29, pp. 641-662, 1996.
- 57 Wang Xuwen, Ding Xiaoqing, Liu Changsong, "Gray-scale character image recognition based on fuzzy DCT transform features", 15th International Conference on Pattern Recognition, Vol. 2 , pp. 235 -238, Sept 2000.
- 58 Wieser Johann and Pinz Axel, "Layout and analysis: Finding text, titles, and photos in digital images of newspaper pages", Proceedings of the 2nd ICDAR, Int. Conf. Document Analysis and Recognition, Japan, pages 774-777, IEEE press, 1993.
- 59 Wong K.Y., Casey R.G., and Wahl F.M., "Document analysis system", IBM J. Research Development, Vol. 6, No. 6, pp. 647-656, Nov 1982.
- 60 Yang Y. and Yan H., "An adaptive logical method for binarization of degraded document images", Pattern Recognit., vol. 33, no. 5, pp. 787-807, 2000.
- 61 Yanikoglu B. A. and Sandon P., "Recognizing off-line cursive handwriting", Proc. Int. Conf. Comput. Vision Pattern Recognit., Seattle, WA, 1994, pp. 397-403.
- 96 Lindsay I Smith, A tutorial on Principal Components Analysis, 2002.



## مراجع فارسی

- ۶۲ اردشیر بهرستانی غلامرضا، "شناسایی ساختاری حروف دستنویس فارسی"، پایان نامه کارشناسی ارشد، به راهنمایی: احسان الله کبیر، دانشکده مهندسی برق، دانشگاه تربیت مدرس، ۱۳۷۲
- ۶۳ اسدی سعید، "بازشناسی حروف فارسی با استفاده از شبکه عصبی چند جمله ایهای جداکننده"، پایان نامه کارشناسی ارشد، به راهنمایی: فتحی محمود، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی اصفهان، ۱۳۷۷
- ۶۴ امیری امیررضا، "استخراج متن چاپی از تصاویر گرافیکی"، پایان نامه کارشناسی ارشد، به راهنمایی: کبیر احسان الله، دانشکده مهندسی برق، دانشگاه تربیت مدرس، ۱۳۸۱
- ۶۵ بحری پیمان، "شناسایی حروف دستنویس فارسی به کمک شبکه عصبی فازی"، پایان نامه کارشناسی ارشد، به راهنمایی: فتحی محمود، دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، ۱۳۷۶
- ۶۶ بنی اسدی امیرعلی، "تشخیص حروف دست نویس فارسی به وسیله سیستم هایبیرید نور و فازی"، پایان نامه کارشناسی ارشد، به راهنمایی: ساداتی ناصر، دانشکده مهندسی برق، دانشگاه صنعتی شریف، ۱۳۷۳
- ۶۷ پارسی اصفهانی امیر، "بررسی روشهای ارائه اطلاعات بصورت تمام تصویر در صفحات وب"، مرکز اطلاعات و مدارک علمی ایران، بهار ۱۳۷۹
- ۶۸ تیمساری بیژن، "بازشناسی حروف در کلمات تایپ شده فارسی با استفاده از روش مورفولوژی"، به راهنمایی: فهیمی حمید، دانشکده مهندسی برق، دانشگاه صنعتی اصفهان، ۱۳۷۱
- ۶۹ ثانی رویا، "بازشناسی حروف دست نویس فارسی با استفاده از شبکه عصبی مصنوعی"، پایان نامه کارشناسی ارشد، به راهنمایی: فهیمی مهرداد، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شریف، ۱۳۷۳
- ۷۰ رادمهر مهدی، "درک پیکربندی هندسی یک صفحه متن چاپی"، پایان نامه کارشناسی ارشد، به راهنمایی: احسان الله کبیر، دانشکده مهندسی برق، دانشگاه تربیت مدرس، ۱۳۷۲
- ۷۱ رحمتی محمد، زارعی ناصر، "شناسایی حروف تاییپ فارسی با استفاده از تبدیل خودهمبستگی مورفولوژی "سومین کنفرانس انجمن کامپیوتر ایران دانشگاه علم و صنعت ایران، تهران، دی ۱۳۷۶
- ۷۲ دهقانی علیرضا، "بازشناسی حروف مجزای دستنویس فارسی با استفاده از مدل پنهان مارکف با چگالی پیوسته و ایده ترکیب چند سیستم خبره"، پایان نامه کارشناسی ارشد، به راهنمایی: مسندی شیرازی محمدعلی، دانشکده مهندسی برق، دانشگاه شیراز، ۱۳۷۹
- ۷۳ رضوی محمد، "خواندن اتوماتیک فرمهای انتخاب درس"، پایان نامه کارشناسی ارشد، به راهنمایی: کبیر احسان الله، دانشکده مهندسی برق، دانشگاه تربیت مدرس، ۱۳۷۵
- ۷۴ رفیعی کراچی شعبانعلی، "شکستن کلمات تایپ شده به حروف در رسم الخطهای مختلف"، پایان نامه کارشناسی ارشد، به راهنمایی: کبیر احسان الله، دانشکده مهندسی برق، دانشگاه تربیت مدرس، ۱۳۷۳
- ۷۵ شاه حسینی علی، "شناسایی حروف دستنویس فارسی با استفاده از شبکه عصبی"، پایان نامه کارشناسی ارشد، به راهنمایی: کبیر احسان الله، دانشکده مهندسی برق، دانشگاه تربیت مدرس، ۱۳۷۴
- ۷۶ شیرعلی شهرضا محمد حسن، "تشخیص کلمات و ارقام دستنویس فارسی به کمک شبکه های

- عصبی"، رساله دکتری، دانشکده برق، دانشگاه صنعتی امیرکبیر، ۱۳۷۴
- ۷۷ صدوقی یزدی هادی، "پیش پردازش برای بازشناسی متون فارسی"، پایاننامه کارشناسی ارشد، به راهنمایی: کبیر احسان الله، دانشکده مهندسی برق، دانشگاه تربیت مدرس، ۱۳۷۵
- ۷۸ صفابخش رضا، آنالیز اسناد - تصحیح کجشدگی اسناد، طرح ملی "بازشناسی متن فارسی"، گروه (ب) بهار ۱۳۷۸
- ۷۹ صفابخش رضا، تصحیح کج شدگی اسناد، طرح ملی "بازشناسی متن فارسی"، گروه (ب)، گزارش سه ماهه اول، خرداد ۱۳۷۸
- ۸۰ صفابخش رضا و دست پاک وحدت، "شناسایی مستقل از چرخش و اندازه حروف تایی فارسی" مجموعه مقالات اولین کنفرانس مهندسی برق ایران دانشگاه صنعتی امیرکبیر، تهران، اردیبهشت ۱۳۷۲
- ۸۱ عباسیان کریم، "بازشناسی برخط نویسه های فارسی"، پایان نامه کارشناسی ارشد، به راهنمایی: کبیر احسان الله، دانشکده مهندسی برق، دانشگاه تربیت مدرس، ۱۳۷۶
- ۸۲ عزمی رضا، "بازشناسی متون چاپی فارسی"، پایان نامه دکتری، به راهنمایی: کبیر احسان الله، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، ۱۳۷۸
- ۸۳ فائز کریم، ختن زاد علیرضا و شیرعلی شهرضا محمدحسن، "تشخیص حروف و ارقام دستنویس فارسی با استفاده از گشتاورهای شبکه زرنیکه و بکمک شبکه های عصبی"، دانشگاه علم و صنعت - مجموعه مقالات سومین کنفرانس مهندسی برق ایران، ۸۴ فهیمی حمید، حاتم احمد، "ارائه یک روش ساختاری جهت تشخیص حروف در متون تایپ شده فارسی"، مجموعه مقالات سومین کنفرانس مهندسی برق ایران (کامپیوتر)، ص ۹۷
- ۱۹۰، دانشگاه علم و صنعت ایران، اردیبهشت ۱۳۷۴
- ۸۵ کبیر احسان الله، بهاری کورش، احمدزاده محمدرضا، "بازشناسی متون تایپ شده فارسی"، مجموعه مقالات اولین کنفرانس مهندسی برق ایران دانشگاه صنعتی امیرکبیر، تهران، اردیبهشت ۱۳۷۲
- ۸۶ کبیر احسان الله، عزمی رضا، طرح ملی "بازشناسی متن فارسی"، گزارش شش ماهه اول و دوم، ۱۳۷۸
- ۸۷ مرتضی پور حمیدرضا، "قطعه بندی بر خط کلمات دستنویس فارسی"، پایاننامه کارشناسی ارشد، به راهنمایی: کبیر احسان الله، دانشکده مهندسی برق، دانشگاه تربیت مدرس، ۱۳۷۸
- ۸۸ مسروری کیوان، "شناسایی برون خط کلمات دستنویس فارسی در یک مجموعه محدود"، پایان نامه دکتری، به راهنمایی: کبیر احسان الله، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، ۱۳۷۹
- ۸۹ مسروری کیوان، "بازشناسی حروف دستنویس فارسی با روش فازی، پایان نامه کارشناسی ارشد، به راهنمایی: کبیر احسان الله، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، ۱۳۷۳
- ۹۰ مقدم تبریزی کاوه، "تشخیص الگوی تصویری در حیط پیچیده"، پایاننامه کارشناسی ارشد، به راهنمایی مشیری بهزاد، دانشکده فنی، دانشگاه تهران، ۱۳۷۵
- ۹۱ مقسمی حمیدرضا، "تشخیص حروف تایی فارسی با روش ساختاری"، پایان نامه کارشناسی

- ارشد، به راهنمایی: فتحی محمود، دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، ۱۳۷۶
- ۹۲ نحوی منوچهر، "استخراج کدپستی از آدرسهای تایپی و دستنویس"، پایان نامه کارشناسی ارشد، به راهنمایی: کبیر احسانالله، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، ۱۳۷۶
- ۹۳ نظام آبادی پور حسین، "پیش پردازش متون چاپی فارسی برای جداسازی حروف"، پایان نامه کارشناسی ارشد، به راهنمایی: کبیر احسان الله، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، ۱۳۷۹
- ۹۴ نمازی مهدی، "شناسایی حروف تایپی فارسی با قلمهای متفاوت به کمک شبکه عصبی فازی"، پایان نامه کارشناسی ارشد، به راهنمایی: فائز کریم، دانشکده مهندسی برق، دانشگاه صنعتی امیرکبیر، ۱۳۷۳
- ۹۵ مقدمه ای بر شبکه های عصبی، آر.بیل و تی. جکسون؛ ترجمه دکتر محمود البرزی، انتشارات دانشگاه صنعتی شریف، ۱۳۸۰