بسم الله الرحمن الرحيم

دانشگاه بیرجند
دانشکده مهندسی

پایان‌نامه دوره کارشناسی ارشد مهندسی برق - الکترونیک

بهبود بازشناسی ارقام دستنویس فارسی به کمک خوشه بندی

اسماعیل میری

استاد راهنما:
دکترسید محمد رضوی

استاد مشاور:
دکتر جواد صدری

حمایت کننده:
موسسه تحقیقات ارتباطات و فناوری اطلاعات

تابستان 1390

تأییدیه هیات داوران

دانشگاه شهید بهشتی
مدیریت تحصیلات تکمیلی

صورتجلسه دفاع از پایان نامه کارشناسی ارشد

فرم شماره ۱۰

با تاییدات خداوند متعال جلسه دفاع از پایان نامه تحصیلی کارشناسی ارشد خانم / آقای اسماعیل میری
به شماره دانشجویی : ۸۸۲۳۳۱۶۰۱۶ رشته : برق گرایش : الکترونیک دانشکده:مهندسی
تحت عنوان: بهبود بازشناسی ارقام دستنویس فارسی به کمک خوشه بندی

به ارزش : ۶ واحد درساعت : ۱۷ روز : چهار شنبه مورخ:۹۰/۵/۱۲

با حضور اعضای محترم جلسه دفاع و نماینده تحصیلات تکمیلی به شرح ذیل تشکیل گردید:

سمت	نام و نام خانوادگی	رتبه علمی	اعضاء
استاد راهنمای اول	سید محمد رضوی	استادیار	
استاد راهنمای دوم			
استاد مشاور اول	جواد صدری	استادیار	
استاد مشاور دوم			
داور اول	ناصر مهرشاد	استادیار	
داور دوم	حسن فرسی	استادیار	
نماینده تحصیلات تکمیلی	محسن فرشاد	استادیار	

نتیجه ارزیابی دفاع که منوط به ارائه اصلاحات پیشنهادی توسط هیئت داوران حداکثر ظرف مدت یکماه پس از تاریخ دفاع می باشد، به شرح زیر مورد تایید قرار گرفت:

☑ قبول (با درجه : عالی و امتیاز : ۲۰) ☐ دفاع مجدد ☐ غیر قابل قبول

۱- عالی (۲۰- ۱۹) ۲- بسیار خوب (۱۸/۹۹- ۱۸) ۳- خوب (۱۷/۹۹- ۱۶) ۴- قابل قبول (۱۵/۹۹- ۱۴)

(بدیهی است عواقب آموزشی ناشی از عدم ارائه به موقع اصلاحات مزبور به عهده دانشجو می باشد)

تقدیم به

همسرم که با ایثار، حمایت‌ها وتشویق‌هایش امکان ادامه تحصیل را برایم فراهم نمود.

و

پدر و مادرم مهربانم.

تشکر و قدردانی

از اساتید محترم آقایان:
دکتر رضوی
دکتر مهرشاد
دکتر ظهیری
دکتر صدری
که در طول این دوره از محضرشان بهره بردم

چکیده

بازشناسی ارقام دستنویس فارسی یکی از شاخه های حوزه بازشناسی الگو است. تحقیقات در این زمینه چندین دهه است که آغاز شده و هنوز هم در حال تحول می باشد. این تحقیق به بهبود نرخ بازشناسی ارقام دستنویس فارسی با استفاده از خوشه بندی پرداخته است.

در این پایان‌نامه از سه طبقه‌بندی‌کننده فازی، شبکه عصبی چند لایه‌ی پرسپترون و شبکه عصبی احتمالاتی برای بازشناسی ارقام دستنویس فارسی استفاده شده است.

در یک روش داده‌های آموزش بطور کلی خوشه‌بندی و در هرخوشه نسبت به آموزش طبقه‌بندی‌کننده فازی با قوانین دهگانه اقدام شده است، در بازشناسی داده‌های آزمایش ابتدا نزدیک‌ترین خوشه به داده آزمایش ورودی مشخص شده و سپس در خوشه مزبور طبقه‌بندی‌کننده فازی بازشناسی را انجام می‌دهد. در روش دوم در هر کلاس، داده‌های آموزشی خوشه بندی می‌شوند. در این حالت تعداد کلاس‌ها برابر مجموع تعداد خوشه های کلاس‌های داده‌های آموزشی خواهد شد. در روش سوم از خوشه‌بندی فازی و شبکه عصبی استفاده شده، بدین ترتیب که هر کلاس از داده‌های آموزشی مانند روش دوم خوشه‌بندی می شود و شبکه با این کلاس‌های جدید آموزش می بیند. در روش پایانی از شبکه عصبی احتمالی با خوشه‌بندی داده های آموزش مانند روش دوم و آموزش شبکه با مراکز خوشه‌ها استفاده شده که منجر به حل مشکل نیاز به حافظه بالا، سرعت و در عین حال نرخ طبقه‌بندی بالا در این شبکه‌های عصبی می-شود. در روش اول تعداد خوشه‌های بهینه با افزایش تعدادخوشه ها تاحصول بالاترین نرخ بازشناسی بدست می آید. در سه روش دیگر برای تعیین تعداد خوشه‌های بهینه از الگوریتم بهینه‌سازی جمعیت ذرات با ذره بردار10*1حاوی تعداد خوشه‌های بهینه کلاس‌های دهگانه و تابع سازگاری نرخ بازشناسی داده های تصدیق استفاده شده است. برای بازشناسی ارقام دستنویس ویژگی‌های زونینگ، مکان مشخصه وگشتاورهای زرنیکی استفاده شده و تکنیک‌های خوشه‌بندی k-means و FCM بکار گرفته شده است .

پایگاه‌داده‌ی مورد استفاده دراین پایان‌نامه پایگاه‌داده هدی است. این پایگاه‌داده شامل102352 نمونه است که از 60،000 نمونه برای آموزش و تصدیق طبقه‌بندی‌کننده و20،000 نمونه جهت آزمایش استفاده شده است. نتایج مربوط به این پایان نامه که در فصل چهارم گزارش شده نشان دهنده تاثیر چشمگیر خوشه‌بندی در افزایش نرخ بازشناسی ارقام دستنویس فارسی است.

کلید واژه‌ها: ارقام دستنویس، بازشناسی، خوشه بندی، شبکه عصبی چندلایه پرسپترون، شبکه عصبی احتمالاتی، طبقه‌بندی فازی، طبقه‌بندی عصبی

فهرست مطالب

عنوان **صفحه**

فهرست شکل‌ها.....ط
فهرست جدول‌ها.....ک

فصل 1- مقدمه 1
1-1- پیشگفتار 1
1-2- هدف از این پایان نامه.....1

فصل 2- مروری بر مفاهیم وکار های انجام شده 3
2-1- مرور برخی مفاهیم مرتبط با بازشناسی الگو.....3
2-1-1- بازشناسی الگو 3
2-1-2- یادگیری ماشین.....3
2-1-3- استخراج ویژگی 4
2-1-4- طبقه بندی 4
2-1-5- انتخاب ویژگی 5
2-2- بازشناسی نویسه های(حروف و ارقام)دستنویس 5
2-2-1- پیش پردازش تصویر 6
2-2-2- روشهای استخراج ویژگی 7
2-2-3- فرآیند آموزش 16
2-2-4- فرآیند بازشناسی 16
2-3- مروری بر روش‌های خوشه بندی.....16
2-3-1- تعریف خوشه بندی.....17
2-3-2- خوشه‌بندی در مقابل طبقه‌بندی.....17
2-3-3- تقسیم بندی روش های مختلف خوشه بندی.....18
2-3-4- روش‌های خوشه بندی.....23
2-3-5- بررسی تکنیک‌های اندازه‌گیری اعتبار خوشه‌ها 44
2-3-6- خوشه بندی نیمه نظارت شده 51
2-3-7- کاربرد های خوشه بندی 52
2-4- مروری بر تحقیقات انجام شده در زمینه بازشناسی ارقام 53

فصل 3- روش انجام کار.....58
3-1- مقدمه 58
3-2- روش های طبقه بندی بازشناسی ارقام دستنویس 58
3-2-1- طبقه بندی کننده فازی.....59

3-2-2- طبقه بندی کننده شبکه عصبی 59
3-2-3- الگوریتم بهینه سازی جمعیت ذرات 66
3-3- روش‌های بکار رفته برای بهبود بازشناسی ارقام دستنویس فارسی 67
3-3-1- استفاده از طبقه بندی کننده فازی وخوشه بندی **k-means** روی کل اعداد 67
3-3-2- استفاده از طبقه بندی کننده فازی وخوشه بندی **k-means** برای هر کلاس 69
3-3-3- استفاده از طبقه بندی کننده شبکه عصبی پرسپترون چندلایه وخوشه بندی فازی 69
3-3-4- استفاده از طبقه بندی کننده شبکه عصبی احتمالاتی وخوشه بندیk-means 72

فصل 4- نتایج، بحث و پیشنهادها 74
4-1- نتایج 74
4-1-1- نتایج حاصل از خوشه بندی کلی و طبقه بندی کننده فازی 74
4-1-2- نتایج حاصل از خوشه بندی کلاس‌ها و طبقه بندی کننده فازی 78
4-1-3- نتایج حاصل از خوشه بندی و طبقه بندی شبکه عصبی پرسپترون 82
4-1-4- نتایج حاصل از خوشه بندی و طبقه بندی شبکه احتمالاتی 85
4-2- جمع بندی و پیشنهادهایی برای ادامه کار 91
فهرست مراجع 92

فهرست شکل‌ها

عنوان **صفحه**

شکل 2-1 : اضافه کردن فضای سفید جهت قرار گرفتن رقم در داخل دایره واحد 12
شکل 2-2 : ویژگی های جهتی کانتور و ویژگی های نقاط خمش 13
شکل 2-3 : استخراج ویژگی زونینگ 14
شکل 2-4 : نحوه محاسبه ویژگی مکان مشخصه برای یک نقطه از تصویر 14
شکل 2-5 : نمونه‌ای از اعمال خوشه‌بندی روی یک مجموعه از داده‌ها 17
شکل 2-6 : مقایسه طبقه بندی و خوشه بندی 18
شکل 2-7 : تفاوت بین روش‌های بالا به پایین با روش‌های پایین به بالا 24
شکل 2-8 : شباهت بین دو خوشه در روش تک لینک 24
شکل 2-9 : دندو گرام 26
شکل 2-10 : شباهت بین دو خوشه در روش لینک کامل 27
شکل 2-11 : دندوگرام مثال لینک کامل 28
شکل 2-12 : شباهت بین دو خوشه در روش لینک متوسط 29
شکل 2-13 : دندوگرام مثال لینک متوسط 31
شکل 2-14 : شباهت بین دو خوشه در روش لینک متوسط گروهی 31
شکل 2-15 : مثال مراحل اجرای خوشه بندی k-means 37
شکل 2-16 : سلول دفاعی 42
شکل 2-17 : نحوه عملکرد عملگر جهش 43
شکل 2-18 : مجموعه داده‌های بکار رفته برای مقایسه ی کارایی شاخص‌های اعتبارسنجی خوشه‌ها 49
شکل 2-19 : مقادیر شاخص‌های اعتبار بر روی نتایج حاصل از خوشه‌بندی داده‌ها کاملا مجزا 49
شکل 2-20 : مقادیر مربوط به شاخص‌های اعتبار بر روی نتایج حاصل از خوشه‌بندی داده‌ها حلقوی 49
شکل 2-21 : دو حالت خوشه‌بندی درست و نادرست داده‌های با شکل دلخواه 51
شکل 2-22 : مقادیر شاخص‌های اعتبار بر روی نتایج حاصل از خوشه‌بندی داده‌ها با شکل دلخواه 51
شکل 2-23 : مقایسه ارقام انگلیسی، فارسی و هندی 55
شکل 3-1 : ساختار نرون طبیعی انسان 61
شکل 3-2 : ساختار یک نرون مصنوعی 61
شکل 3-3 : یک نمونه شبکه عصبی 8×3×8 62
شکل 3-4 : یک ساختارنوعی از شبکه عصبی احتمالاتی 64
شکل 3-5 : یک مثال از PNNدوکلاسه 65
شکل 3-6 : بلوک دیاگرام تعیین تعداد خوشه های بهینه برای بهبود طبقه بندی فازی 68
شکل 3-7 : بلوک دیاگرام طبقه بندی کننده بهبود یافته فازی با خوشه بندی 68

شکل 3-8 : بلوک دیاگرام تعیین خوشه های بهینه با استفاده از pso 69
شکل 3-9 : طبقه بندی فازی بهبود یافته 69
شکل 3-10 : الف: بردار هدف در حالت بدون خوشه بندی ب:: بردارهدف با خوشه بندی فازی 71
شکل 3-11 : بلوک دیاگرام بهبود طبقه بندی شبکه عصبی MLP با استفاده از خوشه بندی FCM هوشمند 71
شکل 3-12 : بلوک دیاگرام طبقه بندی کننده شبکه عصبی بهبود یافته با خوشه بندی 72
شکل 3-13 : بلوک دیاگرام بهبود طبقه بندی کننده PNN با خوشه بندی هوشمند 73
شکل 3-14 : طبقه بندی کننده PNN بهبود یافته 73
شکل 4-1 : تغییر نرخ بازشناسی داده های آموزش2 با تغییر تعداد خوشه های کلی(مکان مشخصه)..... 74
شکل 4-2 : تغییر نرخ بازشناسی داده های آموزش با تغییر تعداد خوشه های کلی(زونینگ) 75
شکل 4-3 : تغییر نرخ بازشناسی داده های آموزش با تغییر تعداد خوشه های کلی(زرنیک256) 76
شکل 4-4 : روند بهینه یابی pso در روش خوشه بندی هر رقم(زونینگ) 78
شکل 4-5 : روند بهینه یابی pso در روش خوشه بندی هر رقم(زرنیک) 79
شکل 4-6 : روند بهینه یابی pso در روش خوشه بندی هر رقم(مکان مشخصه) 79
شکل 4-7 : عملکرد pso در یافتن نرخ بازشناسی بهینه درMLP (زونینگ) 82
شکل 4-8 : عملکرد pso در یافتن نرخ بازشناسی بهینه درMLP (مکان مشخصه) 83
شکل 4-9 : تغییرات نرخ بازشناسی طبقه بندی کننده PNN با تغییر spread(زونینگ) 87
شکل 4-10 : تغییرات نرخ بازشناسی طبقه بندی کننده PNN با تغییر spread(مکان مشخصه) 87
شکل 4-11 : عملکرد pso در یافتن خوشه های بهینه(زونینگ) 87
شکل 4-12 : عملکردpso در یافتن خوشه های بهینه(مکان مشخصه) 88
شکل 4-3 : نمودار مقایسه ای نرخ بازشناسی داده آزمایشی برای حالت های مختلف ویژگی در طبقه بندی کننده فازی 89
شکل 4-14 : نمودار مقایسه ای نرخ بازشناسی داده آزمایشی برای حالت های مختلف ویژگی در طبقه بندی کننده MLP,PNN 89

فهرست جدول‌ها

عنوان ..... صفحه
جدول 2-1 : ماتریس فاصله‌ی بین 6 نمونه داده..... 24
جدول 2-2 : ماتریس فاصله‌ی بین 5 خوشه‌ی حاصل از تکرار اول 25
جدول 2-3 : ماتریس فاصله‌ی بین 4 خوشه‌ی حاصل از تکرار دوم..... 25
جدول 2-4 : ماتریس فاصله‌ی بین 3 خوشه‌ی حاصل از تکرار سوم 25
جدول 2-5 : ماتریس فاصله‌ی بین 2 خوشه‌ی حاصل از تکرار چهارم..... 26
جدول 2-6 : ماتریس فاصله‌ی بین 6 نمونه‌ی داده..... 27
جدول 2-7 : ماتریس فاصله‌ی بین 5 خوشه‌ی حاصل از تکرار اول 27
جدول 2-8 : ماتریس فاصله‌ی بین 4 خوشه‌ی حاصل از تکرار دوم..... 27
جدول 2-9 : ماتریس فاصله‌ی بین 3 خوشه‌ی حاصل از تکرار سوم 28
جدول 2-10 : ماتریس فاصله‌ی بین 2 خوشه‌ی حاصل از تکرار چهارم..... 28
جدول 2-11 : ماتریس فاصله‌ی بین 6 نمونه‌ی داده 29
جدول 2-12 : ماتریس فاصله‌ی بین 5 خوشه‌ی حاصل از تکرار اول 29
جدول 2-13 : ماتریس فاصله‌ی بین 4 خوشه‌ی حاصل از تکرار دوم..... 30
جدول 2-14 : ماتریس فاصله‌ی بین 3 خوشه‌ی حاصل از تکرار سوم 30
جدول 2-15 : ماتریس فاصله‌ی بین 2 خوشه‌ی حاصل از تکرار چهارم..... 30
جدول 2-16 : ضرائب رابطه‌ی (7) برای روش‌های متفاوت خوشه‌بندی سلسله مراتبی..... 32
جدول 2-17 : مجموعه ی علائم بکار رفته در این بخش..... 45
جدول 4-1 : کارایی نمونه های آموزش با خوشه بندی کلی (مکان مشخصه)..... 75
جدول 4-2 : کارایی نمونه های آزمایش با خوشه بندی کلی (مکان مشخصه)..... 75
جدول 4-3 : کارایی نمونه های آموزش با خوشه بندی کلی (زونینگ) 76
جدول 4-4 : کارایی نمونه های آزمایش با خوشه بندی کلی (زونینگ) 76
جدول 4-5 : کارایی نمونه های آموزش با خوشه بندی (زرنیک)..... 77
جدول 4-6 : کارایی نمونه های آزمایش با خوشه بندی کلی (زرنیک)..... 77
جدول 4-7 : نتایج بازشناسی با خوشه بندی کلی و طبقه بندی کننده فازی 77
جدول 4-8 : خوشه های بهینه خوشه بندی کلاس ها (زونینگ)..... 78
جدول 4-9 : خوشه های بهینه خوشه بندی کلاس ها (زرنیک)..... 78
جدول 4-10 : خوشه های بهینه خوشه بندی کلاس ها (مکان مشخصه) 78
جدول 4-11 : کارایی سیستم بازشناسی نمونه های آزمایش خوشه بندی کلاس‌ها(زونینگ)..... 79
جدول 4-12 : کارایی سیستم بازشناسی روی نمونه های آموزش خوشه بندی کلاس‌ها(زونینگ)..... 79
جدول 4-13 : کارایی سیستم بازشناسی نمونه های آموزش خوشه بندی کلاس‌ها(زرنیکی)..... 79

جدول 4-14 : کارایی سیستم بازشناسی نمونه های آزمایش خوشه بندی کلاس‌ها(زرنیکی).....79
جدول 4-15 : کارایی سیستم بازشناسی نمونه های آزمایش خوشه بندی کلاس‌ها(مکان مشخصه).....81
جدول 4-16 : کارایی سیستم بازشناسی نمونه های آموزش خوشه بندی کلاس‌ها(مکان مشخصه).....81
جدول 4-17 : نتایج به دست آمده در حالات مختلف خوشه بندی و طبقه بندی کننده فازی.....81
جدول 4-18 : خوشه ها بهینه خوشه بندی کلاس‌ها (زونینگ).....82
جدول 4-19 : خوشه های بهینه خوشه بندی کلاس‌ها (مکان مشخصه).....82
جدول 4-20 : کارایی سیستم بازشناسی نمونه های آزمایش درMLP(زونینگ250 تکرار).....83
جدول 4-21 : کارایی سیستم بازشناسی نهایی نمونه های آموزش درMLP (زونینگ250 تکرار).....83
جدول 4-22 : کارایی سیستم بازشناسی نمونه های آموزش درMLP(مکان مشخصه).....84
جدول 4-23 : کارایی سیستم بازشناسی آزمایش درMLP(مکان مشخصه).....84
جدول 4-24 : نتایج بازشناسی با استفاده از شبکه عصبی MLP.....84
جدول 4-25 : خوشه های بهینه خوشه بندی کلاس ها (زونینگ).....85
جدول 4-26 : خوشه های بهینه خوشه بندی کلاس ها (مکان مشخصه).....85
جدول 4-27 : کارایی سیستم بازشناسی نمونه های آزمایش درPNN(زونینگ).....85
جدول 4-28 : جدول کارایی سیستم بازشناسی روی نمونه های آموزش در PNN(زونینگ).....86
جدول 4-29 : جدول کارایی سیستم بازشناسی روی نمونه های آموزش در PNN(مکان مشخصه).....86
جدول 4-30 : جدول کارایی سیستم بازشناسی روی نمونه های آزمایش در PNN(مکان مشخصه).....86
جدول 4-31 : نتایج بازشناسی با استفاده از شبکه عصبی احتمالاتی.....88
جدول 4-32 : نتایج روشهای مختلفی که در زمینه بازشناسی ارقام دستنویس انجام شده است.....89

فصل 1 - مقدمه

1 - 1 - پیشگفتار

امروزه در فرم‌های رسمی و غیر رسمی، چک‌های بانکی، قراردادهای اقتصادی، انواع تجارت‌های بانکی، فرم‌های گزارش مالی، فرم‌های گزارش آماری و مقادیر ریالی روی صورت‌حساب‌ها حتی با وجود ابزارهای الکتریکی که در دسترسند و می‌توانند به راحتی برای تایپ مورد استفاده قرار گیرند هنوز مردم از اعداد دستنویس استفاده می‌کنند. ابزارهای مورد استفاده در تشخیص این اعداد دستنویس بعلت شباهت آنها گاها دچار خطا می‌شوند که ممکن است این خطا ها مشکلات جبران ناپذیری ایجاد کنند، لذا میدان تحقیق روی سیستم های با خطای کمتر هنوز باز است. بازشناسی ارقام دستنویس فارسی یکی از زمینه-های پردازش تصویر و بازشناسی الگو و یکی از مسائل مهم در حوزه بازشناسی نوری حروف است که تحقیقات گسترده‌ای را به خود اختصاص داده و هنوز هم در حال تحول است. روش‌های مختلفی برای بازشناسی استفاده شده است که از جمله آنها می توان به روش‌های آماری، ساختاری، شبکه های‌عصبی و فازی اشاره کرد.

در زمینه بازشناسی ارقام دستنویس،کارهای زیادی صورت گرفته است. در یک تحقیق از مکان‌های مشخصه، طبقه بندی بیز، و زنجیره مارکف برای بازشناسی ارقام دستنویس استفاده شده است [1]. در [2] بازشناسی نهایی ارقام دستنویس فارسی به کمک روش‌های فازی انجام شده است. در تحقیقی دیگر از الگوریتم تطابق شکل و طبقه بندی نزدیکترین فاصله از نماینده یا نماینده‌های هر کلاس برای بازشناسی ارقام دستنویس فارسی استفاده شده است [3]. در [4] از طبقه بندی توسط شبکه عصبی و ترکیب سه ویژگی مکان مشخصه، گرادیان بهبود یافته و کریش برای بازشناسی ارقام استفاده شده است. در[5] از شبکه عصبی و ترکیب ده طبقه بندی کننده دو کلاسی برای بازشناسی ارقام دستنویس فارسی استفاده شده است.

1 - 2 - هدف از این پایان نامه

یکی از مهمترین مسائلی که دراستفاده از یک سیستم بازشناسی ارقام دستنویس مورد توجه است ضرورت بالا بودن نرخ بازشناسی است. در حوزه زبان فارسی، به دلیل شباهت زیادی که ارقام به هم دارند و همچنین تفاوت در شیوه رسم آنها، ایجاد یک سیستم بازشناسی با دقت قابل قبول برای استفاده عملی با مشکلاتی مواجه است. از این رو توسعه روش‌هایی برای بهبود دقت در آنها ضروری است.در این پایان-نامه تحقیقاتی پیرامون چگونگی بهبود نرخ بازشناسی به کمک خوشه بندی صورت گرفته است. پایگاه داده مورد استفاده در این تحقیق یک پایگاه داده بزرگ از ارقام دستنویس فارسی بنام پایگاه داده هدی [6] است. تعداد کل ارقام این پایگاه داده 102352 رقم است که 60000 از آن به عنوان نمونه‌های آموزش وتصدیق، 20000 به عنوان نمونه‌های آزمایش و 22352 رقم آن هم به عنوان نمونه های باقیمانده در نظر گرفته شده‌اند.

طبقه‌بندی کننده‌های مورد استفاده در این تحقیق عبارتند از طبقه‌بندی‌کننده فازی [2] که به دو صورت مورد استفاده قرار گرفته است، در یک روش بعد از خوشه‌بندی کلی نمونه‌های آموزشی پایگاه داده، در هر خوشه با استفاده از ویژگی‌های استخراج شده ارقام، توابع عضویت گوسی برای مدل فازی بدست آمده است. مدل فازی مورد استفاده در این اینجا یک مدل ممدانی با 10 قاعده است. هر قاعده برای یک رقم اختصاص داده شده است. در مرحله آزمایش ابتدا نزدیک ترین خوشه به داده آزمایش انتخاب شده و سپس طبقه بندی در این خوشه صورت می‌گیرد.

در روش بعدی ابتدا داده‌های آموزشی هر رقم با استفاده از برچسب جداسازی شده، سپس داده های آموزشی هرعدد بصورت مجزا خوشه بندی می‌شود. که این به مفهوم ایجاد زیر کلاس‌های متعدد برای کلاس‌های دهگانه ویا بعبارت دیگر افزایش کلاس‌ها از ده کلاس قبلی به تعداد مجموع خوشه های ایجاد شده بعد از خوشه بندی است. تعداد خوشه‌های بهینه توسط الگوریتم بهینه سازی جمعیت ذرات(pso)[1] تعیین می شود که در اینجا ذره مورد استفاده یک بردار 10*1 حاوی تعدادخوشه ها در کلاس‌های دهگانه و تابع سازگاری نرخ بازشناسی داده‌های تصدیق است. در این حالت با توجه به زیرکلاس‌های ایجاد شده، تعداد قواعد طبقه‌بندی‌کننده فازی به تعداد خوشه‌های بهینه انتخابی بستگی دارد. در مرحله آزمایش، از این مدل فازی با تعداد کلاس معادل تعدادکل خوشه های بهینه برای طبقه‌بندی نمونه‌های آزمایشی پایگاه داده استفاده می‌شود.

طبقه‌بندی‌کننده دیگر مورد استفاده شبکه عصبی چند لایه پرسپترون است که با ترکیب با یک خوشه‌بندی فازی[2] که روی هر رقم از مجموعه آموزش بصورت جداگانه انجام شده(شبیه روش دوم فازی) به نتایج بازشناسی مطلوب تری دست یافته است. در این روش بردار هدف به کمک درجه عضویت خوشه‌ها ساخته می‌شود. و در نهایت طبقه‌بندی‌کننده پایانی شبکه عصبی احتمالاتی[3] است که از مراکز خوشه‌های ارقام آموزشی (بدست آمده از خوشه بندی مانند روش قبل) برای آموزش شبکه استفاده کرده و در نهایت یک نرخ بازشناسی قابل قبول در داده های آزمایش ارائه داده است.

همانطور که گفته شد در این پایان نامه با استفاده از خوشه بندی هوشمند سعی شده نرخ بازشناسی افزایش یابد. نتایج و مقایسه های ارائه شده در پایان تحقیق نشان می دهد هدف تحقیق برآورده گردیده و روش‌های انتخاب شده منجر به افزایش نرخ بازشناسی در حد قابل قبولی می شود.

ساختار بقیه فصل های این پایان نامه به این صورت است که در فصل دوم به مرور برخی مفاهیم مرتبط با بازشناسی الگو، خوشه بندی و مروری بر تحقیقات انجام شده در زمینه بازشناسی ارقام اختصاص دارد. در فصل سوم روش‌های طبقه بندی وخوشه‌بندی استفاده شده جهت بهبود بازشناسی تشریح شده و فصل چهارم به عنوان فصل آخر به بررسی نتایج، تحلیل و مقایسه آنها و پیشنهادهایی برای ادامه کار پرداخته است.

[1] PSO
[2] Fuzzy c-means
[3] PNN

فصل 2 - مروری بر مفاهیم وکار های انجام شده

2 - 1 - مرور برخی مفاهیم مرتبط با بازشناسی الگو

2 - 1 - 1 - بازشناسی الگو

بازشناسی الگو امری است که در زندگی روزمره ما به وفور رخ می‌دهد. از آن جمله می‌توان به بازشناسی و تشخیص صدای افراد مختلف، بازشناسی چهره افرادی که با آنها روبرو می‌شویم و یا تشخیص نوع دستخط یا امضاء افراد مختلف اشاره کرد. اگرچه بازشناسی الگو امری پیچیده است ولی بدلیل قابلیت های بالای مغز انسان به آسانی و در حداقل زمان انجام می‌شود. ما به دلیل آنکه همواره و به طور ناخودآگاه انواع مختلفی از این بازشناسی را انجام می‌دهیم، از پیچیده بودن این فرآیند بی‌خبریم. اما اهمیت توانایی انسان در بازشناسی الگو هنگامی مشخص می‌شود که انسان سعی می‌کند ساده ترین نوع بازشناسی را که براحتی توسط مغز انجام می‌شود، توسط رایانه انجام دهد [3].

امروزه، با پیشرفت علوم رایانه و بالا رفتن سرعت تجزیه و تحلیل برنامه‌های رایانه‌ای، پیشرفت زیادی در زمینه اجرای الگوریتم‌های بازشناسی الگو به کمک رایانه صورت گرفته است. این الگوریتم ها در زمینه‌های گوناگونی به کار رفته اند. برای مثال در زمینه تشخیص صداهای مختلف، بازشناسی چهره افراد و یا تشخیص حروف و ارقام دستنویس از الگوریتم های بازشناسی الگو استفاده شده است [3].

امروزه روش های بازشناسی الگو، به عنوان یک شاخه از یادگیری ماشین[1]، کاربردهای فراوانی در زمینه های مختلف علمی و صنعتی پیدا کرده اند.

2 - 1 - 2 - یادگیری ماشین

یادگیری ماشینی یک شاخه مهم از گرایش هوش مصنوعی است که هدف آن تعلیم یک ماشین است بطوریکه بتواند تجربیات و نمونه‌های موجود را یاد بگیرد. حاصل این یادگیری ایجاد یک مدل طبقه-بندی است که بر اساس آن ماشین می‌تواند نمونه‌هایی را که در آینده می‌بیند و مشابه نمونه های موجوداند در کلاس مناسب خود قرار دهد [7].

روشهای یادگیری به دونوع تقسیم می شوند:

الف) یادگیری همراه با نظارت: در روش یادگیری همراه با نظارت، هر الگو از مجموعه داده با یک برچسب کلاسی همراه است. هدف این است که بر اساس نمونه های موجود، مدل طبقه‌بندی را طوری بسازیم که بتواند نمونه‌هایی را که تاکنون ندیده است با کمترین خطا در کلاس مربوط به خودشان دسته‌بندی کند.

[1] Mechine lerning

ب) یادگیری بدون نظارت: در یادگیری بدون نظارت، الگوها برچسب کلاسی ندارند و بر اساس شباهتشان در دسته های یکسان قرار می گیرند.

2-1-3- استخراج ویژگی

برای استخراج ویژگی تعریف‌های متعددی بیان شده است. بعنوان مثال یک تعریف می گوید: استخراج ویژگی استخراج اطلاعاتی از داده خام است که برای اهداف طبقه‌بندی بیشترین تناسب را داشته باشند، بدین ترتیب که تغییرات الگو در درون هر کلاس کمینه و در بین کلاس‌های مختلف بیشینه شود.

واضح است که روش‌های مختلف استخراج ویژگی، بسته به نوع داده ها و مسئله بازشناسی، این نیاز را تا حدی برآورده می‌کنند. یک روش استخراج ویژگی که برای یک کاربرد موفق است، ممکن است برای کاربرد دیگری چندان مناسب نباشد [8].

بر اساس تعریفی دیگر، استخراج ویژگی یک نگاشت از فضای الگوها به فضای ویژگی‌هاست که نتیجه آن پیدایش یک بردار ویژگی x است.

انتخاب روش استخراج ویژگی به عنوان یک فاکتور مهم در کارایی سیستم بازشناسی الگو مطرح است. برای شناسایی حروف و ارقام، ویژگی‌های ناحیه ای[1][9]، مکان مشخصه[2][4]، گشتاورهای هندسی[3][4]، گشتاورهای زرنیکی[4][10]، توصیفگرهای فوریه، ثابت‌های گشتاوری [11][5]، هیستوگرام نما و توصیف گراف پیشنهاد شده اند.

2-1-4- طبقه بندی

یکی از مراحل اصلی فرآیند بازشناسی الگو، طبقه بندی است. مرحله طبقه بندی شامل روش‌هایی برای متناظر ساختن هر یک از الگوهای به دست آمده از مرحله استخراج ویژگی، با یکی از کلاس‌های الگو است. این کار با کمینه ساختن فاصله بردار ویژگی های هر الگوی ورودی نسبت به یکی از بردارهای مرجع، انجام می گیرد. بردارهای مرجع، بردارهایی هستند که قبلا از نمونه‌های آموزشی اخذ شده‌اند. بر اساس [12] تکنیک های عرضه شده برای مرحله طبقه‌بندی را می توان در روش‌های مربوط به چهار گروه عمومی مبحث بازشناسی الگو جستجو نمود.

1- تطابق قالبی 2- تکنیک های آماری 3- تکنیک های ساختاری 4- شبکه های عصبی

چهار گروه فوق لزوما مستقل یا مجزا از یکدیگر نیستند و گاهی می توان تکنیک‌های یک گروه را در میان تکنیک‌های مربوط گروه‌های دیگر یافت. در یک سیستم بازشناسی الگو، می‌توان از یک یا چند

[1] Zoning
[2] Loci
[3] Geometric Moments
[4] Zernike Moments
[5] Invariant Moments

طبقه‌بندی‌کننده استفاده کرد. طبقه بندی بر مبنای فاصله [13]، شبکه های عصبی [14]، svm [15] و روش‌های فازی از روش‌های رایج برای طبقه بندی الگوها اند.

2-1-5- انتخاب ویژگی

انتخاب ویژگی یکی از عوامل موثر در کاهش یا افزایش نرخ بازشناسی الگو است. انتخاب نوع ویژگی، نیازمند دانش اولیه در مورد الگوهاست. انتخاب ویژگی باید به گونه ای باشد که ویژگی‌ها، وجه مشترک الگوهای متعلق به یک کلاس و وجه تمایز الگوهای کلاس‌های مختلف باشند. توانمندی ویژگی برای جداسازی نمونه های کلاس‌های مختلف، معیار انتخاب آن است.

2-2- بازشناسی نویسه های(حروف و ارقام)دستنویس

بازشناسی نویسه های دستنویس یکی از مسائل مهم در حوزه بازشناسی الگو و فرآیندی است که در آن زبان نوشته شده در شکل فضایی به شکل قابل فهم توسط رایانه تبدیل می شود.

بازشناسی نویسه‌های دستنویس چندین دهه است که مورد مطالعه قرار گرفته است. تاکنون، تحقیقات وسیعی روی آن به انجام رسیده و هنوز از بعضی جهات به عنوان یکی از مسائل باز مطرح است. بازشناسی نویسه‌های دستنویس در واسطه بین انسان و ماشین کاربردهای گوناگونی دارد. معیارهای گوناگونی برای ارزیابی عملکرد یک روش بازشناسی نویسه‌های دستنویس (دست نوشته‌ها) وجود دارد. از جمله این معیارها می توان به اندازه ارقام، استقلال از سبک نوشتار، قابلیت اطمینان و سرعت بازشناسی اشاره نمود. بنا به دلایلی همچون سبک نوشتاری خاص افراد مختلف، بافت (محدوده) ارقام، ابزار و واسطه‌های نوشتاری متفاوت، تصاویر اسکن شده دستخوش تغییراتی مثل کجی و اندازه متفاوت ارقام و ضربه هایی با پهنا و شکل های متفاوت می شوند. ازاین رو بازشناسی ارقام دستنویس مشکل است. توسعه روش‌های کارآمد جهت بازشناسی نویسه‌های دستنویس می‌تواند درکاربردهای آن راهگشا باشد. بزرگ‌ترین چالش در این حوزه، تنوع شیوه‌های رسم نویسه‌ها است. بازشناسی نویسه های دستنویس از لحاظ نوع الگوی ورودی به دو دسته اصلی تقسیم شده است:

1- **بازشناسی بر خط**[1]

2- **بازشناسی برون خط**[2]

روش های بازشناسی برخط، روش‌های اند که بطور مستقیم با ماشین در تعامل‌اند. در این روش‌ها حروف در همان زمان نگارش توسط سیستم تشخیص داده می‌شوند و دستگاه ورودی این سیستم ها یک قلم نوری است. در این روش‌ها علاوه بر اطلاعات مربوط به موقعیت قلم، اطلاعات زمانی مربوط به مسیر قلم نیز در اختیار است. این اطلاعات معمولا توسط یک صفحه رقومی کننده اخذ می‌شوند. در این روش-

[1] Online
[2] Offline

ها می‌توان از اطلاعات زمانی سرعت، فشار و زمان برداشتن و گذاشتن قلم روی صفحه در بازشناسی استفاده کرد.

بازشناسی برون خط خود به دو دسته چاپی و دستنویس تقسیم می شود. در بازشناسی برون خط، از تصویر دوبعدی متن ورودی استفاده می‌شود. در این روش‌ها به هیچ نوع وسیله نگارش خاصی نیاز نیست و تفسیر داده‌ها مستقل از فرآیند تولید آنها و تنها بر اساس تصویر متن صورت می‌گیرد. این روش‌ها به نحوه بازشناسی توسط انسان شباهت بیشتری دارد.

در بازشناسی برون خط با تصاویر سرو کار داریم و در بازشناسی برخط با ضربه‌های قلم سروکار داریم که دارای اطلاعات زمان‌بندی هستند. بنابراین به خاطر وجود اطلاعات زمان‌بندی اضافی بازشناسی کاراکترهای برخط آسانتر از بازشناسی کاراکترهای برون‌خط دستنویس است.

روش‌های زیادی برای بازشناسی حروف و ارقام استفاده شده‌اند. تحقیق ما بر روی بازشناسی ارقام دستنویس فارسی است. روش های سنتی بازشناسی نویسه های دستنویس به چهار مرحله پیش پردازش تصویر، استخراج ویژگی تصویر، فرآیند آموزش و فرآیند شناسایی تقسیم می‌شوند [16].

2-2-1- پیش پردازش تصویر

پیش‌پردازش شامل کلیه اعمالی است که بر روی یک تصویر صورت می‌گیرد تا موجب تسهیل در روند اجرای فازهای بعدی گردد. برخی از این اعمال شامل دوسطحی‌کردن تصویر، حذف نویز، هموارسازی و نازک سازی است. از مجموعه این پردازش ها اهداف زیر حاصل می شوند که در ادامه به توضیح هر یک از آنها می‌پردازیم [12].

- **کاهش نویز**
- **نرمالیزه‌کردن داده ها**
- **فشرده سازی میزان اطلاعاتی که باید محفوظ بماند**

2-2-1-1- کاهش نویز

نویز ایجاد شده به واسطه دستگاه های اسکنر نوری باعث ایجاد نقطه های لک مانند، قطعه خط‌های گسسته[1]، اتصال بین ارقام و فضاهای خالی بین ارقام می‌شود. همچنین اعوجاج های دیگری مثل تغییر شکل[2] یا خوردگی ارقام[3] وجود دارند. بهتر است این نقایص قبل از مرحله بازشناسی برطرف شوند. این کار باعث کاهش خطای بازشناسی وکاهش اندازه فایل تصویر می شود. با کاهش اندازه فایل تصویر، زمان مورد نیاز برای پردازش داده‌ها و همچنین حجم حافظه لازم برای ذخیره سازی کاهش می‌یابد.

[1] Disconnected line segments
[2] Dilation
[3] Erosion

2-2-1-2- نرمالیزه کردن داده ها

روش‌های نرمالیزه کردن[1] داده ها به حذف تغییرات نگارشی کمک می کند و داده های استاندارد شده‌ای را نتیجه می‌دهد. از روش‌های نرمالیزه کردن داده‌ها، دو روش نرمالیزه کردن اندازه وهموارسازی کانتور را بررسی می‌کنیم.

الف)نرمالیزه کردن اندازه

ممکن است تصاویر ارقام خیلی کوچک یا خیلی بزرگ باشند. معمولا این تصاویر به یک اندازه استاندارد نرمالیزه می شوند تا عملیات بازشناسی، مستقل از اندازه فونت متن شود. به این روش نرمالیزه کردن(تغییر مقیاس دادن[2]) اندازه گویند.

ب)هموارسازی کانتور

خط تشکیل دهنده مرز یک کاراکتر را کانتور آن کاراکتر گویند. در متون دستنویس، به واسطه لرزش یا حرکات ناخواسته دست نویسنده درهنگام نگارش، ممکن است کانتور حروف ناصاف شود. این وضعیت، درسیستم‌های بازشناسی متون دستنویس ممکن است به دلیل تغییر مقیاس ارقام یا وجود نویز در مرحله اسکن تصاویر ظاهر شود. هموارسازی کانتور، تعداد نقاط نمونه مورد نیاز برای نمایش کاراکتر را کاهش داده و کارایی سایر مراحل پردازشی را بهبود می بخشد.

2-2-1-3- فشرده سازی

یکی از رایج ترین روش‌هایی که برای فشرده سازی به کار می‌رود نازک‌سازی می‌باشد. باعمل نازک-سازی، تصویر کاراکترها به تصویری با عرض یک پیکسل تبدیل می شود. درست مانند اینکه کاراکترها با یک قلم نوک باریک نوشته شده باشند. نازک سازی علاوه بر این که کاهش قابل ملاحظه ای در حجم داده ها ایجاد می کند، اطلاعات شکلی کاراکتر را نیز حفظ می کند.

2-2-2- روش‌های استخراج ویژگی

همان گونه که در بخش‌های قبل گفته شد، مرحله استخراج ویژگی[3] به دلیل اینکه نتایج حاصل از آن مستقیما بر روی کیفیت و نرخ مرحله بازشناسی اثر می‌گذارد در بازشناسی الگو بسیار مهم است. در مرحله استخراج ویژگی، به هر الگوی ورودی، یک بردار ویژگی نسبت داده می‌شود که معرف آن الگو در فضای ویژگی‌هاست و آن را از سایر الگوها متمایز می‌سازد. در ادامه به توضیح برخی ویژگی‌های استخراج شده از حروف و ارقام می‌پردازیم. به طور کلی روش های مختلف استخراج ویژگی از تصاویر اسناد به سه گروه اصلی تقسیم می شوند. [16] تبدیل سراسری[4] (بسط سری[5])، آماری و هندسی

[1] Normalization
[2] Scaling
[3] Feature extraction
[4] Global Transformation
[5] Series Expansion

2-2-2-1- تبدیل سراسری (بسط سری)

یک سیگنال پیوسته معمولا حاوی اطلاعات بیشتر از میزان مورد نیاز برای استخراج ویژگی جهت اهداف طبقه بندی است. یک روش برای استخراج ویژگی یک سیگنال، بسط سری به صورت ترکیب خطی از توابع ساده تری (دارای پیچیدگی کمتر) که به نحو مناسبی تعریف شده اند، می باشد. به عبارتی ضرایب ترکیب خطی، سیگنال را به صورت فشرده تری کد می‌کنند. در زیر به برخی از تبدیل ها و روش های بسط سری اشاره می‌کنیم.

الف) تبدیل فوریه

تبدیل فوریه گسسته[1] (DFT) در زمینه پردازش تصویر و سیگنال اهمیت فوق العاده ای دارد. در آن معمولا از طیف دامنه بردار ویژگی در فضای اقلیدسی n بعدی به عنوان ویژگی استفاده می شود.DFT چندین ویژگی دارد که در امر شاخص گذاری[2] یا تطبیق الگوها مفید است: اول اینکه دامنه ضرایبDFT نسبت به انتقال حساس[3] نیست و در نتیجه با احتساب طیف دامنه به تنهایی و صرف نظر از اطلاعات فاز می توان حساسیت نسبت به تغییر مکان کاراکترها را در تصویر از بین برد. دوم اینکه همبستگی حوزه زمانی را می توان به طور موثر با استفاده از ضرایب DFT محاسبه نمود. تبدیل‌های فوریه به شکل‌های متفاوتی در بازشناسی حروف مورد استفاده قرار گرفته اند.

روش‌های مختلف کدکردن منحنی پیرامون، بارها در بازشناسی حروف استفاده شده است. در این میان توصیف کننده‌های فوریه نرمالیزه که نسبت به اندازه، انتقال و چرخش ثابت اند، کاربرد بیشتری دارند [12].

ب) DCT

DCT یکی از تبدیل‌های یکتا است که شبیه FFT است. از این تبدیل معمولا برای کاربردهای فشرده سازی استفاده می شود. به عنوان مثال در استاندارد تصویر JPEG از این تبدیل برای فشرده سازی استفاده شده است.

از مهمترین خواص DCT این است که بیشترین اطلاعات سیگنال را در ضرایب معدودی که در فرکانس های پایین قرار دارند ذخیره می کند. همین طور در مورد تصویر هم بیشترین اطلاعات تصویر در ماتریس گوشه بالا- سمت چپ ماتریس خروجی ذخیره می شود. می توانیم از همین اطلاعات به عنوان ویژگی استفاده کنیم. برای این کار باید تصویر ورودی نرمال شود زیرا ابعاد ماتریس های خروجی باید یکسان باشد تا بتوان اطلاعات را از موقعیت های یکسانی استخراج کرد. برای مثال می توانیم تصاویر ارقام خاکستری ورودی را به 32×32 نرمالیزه کنیم.

بیشترین اطلاعات تصویر در تبدیل DCT در ابتدای ماتریس تبدیل قرار دارد. این که چند ضریب از ماتریس خروجی را به عنوان ویژگی در نظر بگیریم سوالی است که غالبا مطرح می شود. در کاربردهای

[1] Discrete Fourier Transform
[2] Indexing
[3] Translation invariant