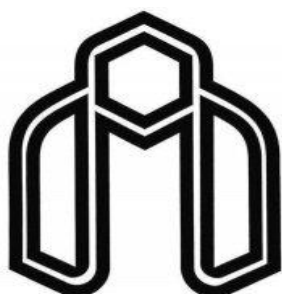


بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه صنعتی شاهرود

دانشکده مهندسی برق و رباتیک

گروه الکترونیک

تشخیص برون خط کلمات دست نوشته فارسی به کمک بلوک بندی تطبیقی

نگارش:

محمدعلیپور سراجی

استاد راهنما:

دکتر علیرضا احمدی فرد

استاد مشاور:

دکتر امیدرضا معروضی

پایان نامه جهت اخذ درجه کارشناسی ارشد

بهمن ۹۱



دیریت تحصیلات تکمیلی
فرم شماره (۶)

بسمه تعالی

شماره : ۹۰۵۰ / آ.ت.ب
تاریخ : ۹۶/۱۱/۲۹
ویرایش : -----

فرم صورتجلسه دفاع پایان نامه تحصیلی دوره کارشناسی ارشد

با تأییدات خداوند متعال و با استعانت از حضرت ولی عصر (عج) جلسه دفاع از پایان نامه کارشناسی ارشد خانم / آقای :
محمد علیپور سراجی رشته : برق
گرایش : الکترونیک
تحت عنوان : تشخیص برون خط کلمات دست نوشته فارسی بکمک بلوک بندی تطبیقی
که در تاریخ ۱۳۹۶/۱۱/۲۹ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح زیر است :

قبول (با درجه : بسیار خوب) امتیاز ۱۷۷۷۵ دفاع مجدد مردود

۱- عالی (۲۰ - ۱۹) بسیار خوب (۱۸/۹۹ - ۱۸)

۳- خوب (۱۷/۹۹ - ۱۶) ۴- قابل قبول (۱۵/۹۹ - ۱۴)

۵- نمره کمتر از ۱۴ غیر قابل قبول

| عضو هیأت داوران | نام و نام خانوادگی | مرتبه علمی | امضاء |
|---------------------------------|--------------------|------------|-------|
| ۱- استاد راهنما | محمد علیپور سراجی | استاد | |
| ۲- استاد مشاور | اسیر زینب غفرانی | استاد | |
| ۳- نماینده شورای تحصیلات تکمیلی | سنا سان ناصح | استادیار | |
| ۴- استاد ممتحن | همین برت | استاد | |
| ۵- استاد ممتحن | حمید حسن لوی | استاد | |

رئیس دانشکده :

تشکر و قدردانی:

در ابتدا از استاد عزیزم، دکتر علیرضا احمدی فرد تشکر می‌نمایم، که در به انجام رساندن این امر همواره از دانش و راهنمایی‌های ایشان بهرمنند بوده‌ام و همچنین از خانواده عزیزم تشکر می‌کنم، که در طول انجام این کار همواره پشتیبانم بوده‌اند.

تعهد نامه

اینجانب محمد علی پور دستانجی دانشجوی دوره کارشناسی ارشد رشته برق - الکترونیک دانشکده مهندسی برق و رباتیک دانشگاه صنعتی شاهرود نویسنده پایان نامه با عنوان :

سینتیزی کردن خط کلمات دستنویس در تصاویر

تحت راهنمایی آقای دکتر علیرضا احمدی متعهد می شوم :

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است .
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است .
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است .
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید .
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافت های آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است .

تاریخ : ۹۱ / ۱۱ / ۳۰

امضاء دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه ای ، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد . این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود .
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

* متن این صفحه نیز باید در ابتدای نسخه های تکثیر شده پایان نامه وجود داشته باشد .

چکیده

تشخیص خودکار متن یکی از زیر مجموعه‌های پردازش تصویر است که به طور گسترده در کتابخانه دیجیتال، خواندن آدرس پستی نامه‌ها، خواندن چک‌های بانکی، خواندن فرم‌ها و پرسشنامه‌ها استفاده می‌شود. بازشناسی متون دست‌نوشته فارسی در دو سطح کلی برون خط و بر خط انجام می‌پذیرد که سیستم‌های برون خط خود به دو دسته تقسیم می‌شود: سیستم‌های مبتنی بر جداسازی و سیستم‌های کلی نگر.

در سیستم‌های مبتنی بر جداسازی سعی می‌شود کلمات به زیر کلمات و حروف سازنده آن تقسیم شوند و از آنها برای شناسایی استفاده شود اما در سیستم‌های کلی نگر، ویژگی‌ها مستقیماً از تصویر کلی کلمه استخراج می‌شود و با مقایسه با داده‌های فرهنگ لغت، مشابه‌ترین کلمه به ورودی به عنوان خروجی در نظر گرفته می‌شود.

در این تحقیق روشی برای بازشناسی کلمات دست‌نوشته فارسی با استفاده از شکل کلی کلمه ارائه شده است. مراحل کار برای شناسایی کلمه شامل: پیش پردازش (باینری کردن، رفع شکستگی، حذف نویز)، استخراج ویژگی (گرادیان، پروفایل) و طبقه‌بند (KNN ، SVM) است.

در مرحله پیش‌پردازش برای تخمین مکان خط کرسی، تغییراتی در روش هیستوگرام افقی انجام می‌شود. تشخیص خط کرسی به منظور جداسازی نقاط از بدنه اصلی و همچنین برای نرمال سازی اجزاء بدنه اصلی ضروری است. در نتیجه فرایند پیشنهادی، از تصویر هر کلمه دو تصویر، یکی تصویر بدنه اصلی و دیگری تصویر نقاط و علائم حاصل می‌شود. برای استخراج ویژگی‌های کلمه در تصویر بدنه اصلی، پس از نرمال سازی جداگانه اجزای آن، از بلوک‌بندی تطبیقی و برای استخراج ویژگی از تصویر نقاط و علائم از

بلوک‌بندی یکنواخت استفاده می‌کنیم. هر بلوک در تصاویر فوق توسط اندازه گرادیان در ۳۲ جهت توصیف می‌شود.

نتایج حاصل با استفاده از طبقه‌بندهای k نزدیکترین همسایگی و ماشین بردار پشتیبان نشان دهنده بهبود در دقت بازشناسی روش ارائه شده است اما به علت پیش پردازش‌های صورت گرفته، روش ارائه شده به زمان بیشتری نسبت به روش‌های پیشین نیازمند است. با خوشه‌بندی نمونه‌ها به منظور کاهش فرهنگ لغت نتیجه بهبود می‌یابد. به منظور خوشه‌بندی نمونه‌ها از ویژگی‌های پروفایل و هیستوگرام افقی و الگوریتم *isodata* استفاده شده است.

کلمات کلیدی: بازشناسی کلمات دست‌نوشته فارسی، گرادیان محلی، بلوک‌بندی تطبیقی، جداسازی نقاط از کلمه، تخمین خط کرسی

لیست مقالات مستخرج از پایانامه

- "تشخیص دست نوشته فارسی به کمک بلوک بندی تلفیقی و ویژگیهای گرادیان"، اولین کنفرانس بازشناسی الگو و تحلیل تصویر، PRIA 2012
- "تشخیص کلمات دست نوشته فارسی به کمک ویژگیهای مبتنی بر بدنه و علائم"، یازدهمین کنفرانس سیستم‌های هوشمند ایران، ICIS 2013

فهرست مطالب

فصل ۱: مقدمه

۱

| | |
|---|-----------------------------------|
| ۲ | ۱-۱-مقدمه |
| ۳ | ۲-۱-داده‌های برخط و برون خط |
| ۳ | ۳-۱-اجزای اصلی یک سیستم بازشناسی |
| ۳ | ۱-۳-۱-پیش پردازش |
| ۴ | ۱-۳-۲-استخراج ویژگی |
| ۴ | ۱-۳-۳-طبقه‌بندی |
| ۴ | ۴-۱-روش‌های بازشناسی متن |
| ۷ | ۵-۱-اندازه فرهنگ لغت |
| ۷ | ۶-۱-کارایی سیستم‌های بازشناسی متن |
| ۸ | ۷-۱-معرفی ویژگی‌های خط فارسی |

۱۲

فصل ۲: مرور کارهای انجام شده

| | |
|----|---|
| ۱۳ | ۱-۲-مقدمه |
| ۱۳ | ۲-۲-بازشناسایی بر مبنای شکل کلی کلمه یا زیرکلمه |

۲۸

فصل ۳: مباحث تئوری

| | |
|----|---------------------------------|
| ۲۹ | ۱-۳-مقدمه |
| ۲۹ | ۲-۳-پیش پردازش |
| ۳۰ | ۱-۲-۳-حذف شکستگی‌های بدنه کلمات |
| ۳۲ | ۲-۲-۳-هموارسازی و حذف نویز |
| ۳۵ | ۳-۲-۳-تخمین مکان خط کرسی |
| ۳۷ | ۴-۲-۳-باینری کردن |

| | |
|----|--|
| ۴۱ | ۵-۲-۳- نازک سازی |
| ۴۵ | ۶-۲-۳- نرمال سازی |
| ۴۷ | ۳-۳- استخراج ویژگی |
| ۴۹ | ۱-۳-۳- ویژگی‌های آماری |
| ۵۵ | ۲-۳-۳- ویژگی‌های ساختاری |
| ۵۷ | ۴-۳- خوشه‌بندی |
| ۵۹ | ۱-۴-۳- روش جمع مربعات |
| ۶۵ | ۲-۴-۳- مدل‌های ترکیبی |
| ۶۶ | ۳-۴-۳- روش‌های سلسله مراتبی |
| ۷۶ | ۵-۳- ماشین بردار پشتیبان |
| ۷۸ | ۱-۵-۳- ماشین بردار پشتیبان غیر خطی |
| ۷۸ | ۲-۵-۳- ماشین بردار پشتیبان در حالت چند کلاسه |

۸۰

فصل ۴: روش پیشنهادی

| | |
|----|---|
| ۸۱ | ۱-۴- مقدمه |
| ۸۱ | ۲-۴- پیش پردازش |
| ۸۲ | ۱-۲-۴- تخمین مکان خط کرسی |
| ۸۴ | ۲-۲-۴- دلایل بروز خطا در تخمین مکان خط کرسی |
| ۸۵ | ۳-۲-۴- جداکردن نقاط و علائم از بدنه اصلی کلمات |
| ۸۶ | ۳-۴- استخراج ویژگی |
| ۸۶ | ۱-۳-۴- ویژگی‌هایی که به منظور خوشه‌بندی نمونه‌ها مورد استفاده قرار می‌گیرد |
| ۸۷ | ۲-۳-۴- ویژگی‌هایی که برای طبقه‌بندی و بازشناسی نهایی نمونه‌ها استخراج می‌گردد |
| ۸۹ | ۴-۴- طبقه‌بندی |
| ۸۹ | ۱-۴-۴- طبقه‌بند K - نزدیکترین همسایگی |
| ۹۳ | ۲-۴-۴- طبقه‌بند ماشین بردار پشتیبان |
| ۹۵ | ۵-۴- خوشه‌بندی |

| | |
|-----|-----------------|
| ۱۰۰ | ۵-۱- مقدمه |
| ۱۰۰ | ۵-۲- نتیجه گیری |
| ۱۰۳ | ۵-۳- پیشنهادات |

فهرست اشکال

- شکل ۱-۱: شکل‌های مختلف حروف فارسی با توجه به موقعیت آنها در زیر-کلمه ۱۰
- شکل ۲-۱: حروف ادغام شده در متون دست‌نوشته ۱۱
- شکل ۱-۲: (الف) تصویر کلمه با چرخش و (ب) تصویر قطبی به دست آمده [۹] ۱۴
- شکل ۲-۲: (الف) کلمه ورودی، (ب) اندازه‌گردایان و (پ) زاویه‌گردایان به دست آمده از آن [۱۰] ۱۵
- شکل ۳-۲: دو نوع ترتیب زیگزاگ برای استخراج ضرایب DCT [۱۲] ۱۷
- شکل ۴-۲: افکنش افقی متن به منظور جداسازی به خطوط [۱۵] ۲۰
- شکل ۵-۲: افکنش عمودی یک خط متنی [۱۵] ۲۱
- شکل ۶-۲: تغییرات نگارش نقاط در متون دست‌نوشته [۱۶] ۲۲
- شکل ۷-۲: نحوه استخراج ویژگی انتقال [۱۶] ۲۳
- شکل ۸-۲: مراحل بازشناسی روش ارائه شده [۱۶] ۲۴
- شکل ۹-۲: الگوریتم ارائه شده در تحقیق [۱۷] ۲۵
- شکل ۱۰-۲: نقطه نمونه p و کدهای جهتی مربوط به آن [۱۹] ۲۷
- شکل ۱-۳: شکستگی بدنه کلمات ۳۰
- شکل ۲-۳: ترمیم شکستگی کاراکتر با روش‌های مختلف [۲۰] ۳۱
- شکل ۳-۳: رسم منحنی به منظور حذف شکستگی، (الف) تصویر شکستگی، (ب) منحنی رسم شده برای ترمیم شکستگی [۲۰] ۳۲
- شکل ۴-۳: فیلترهای هموار ساز: (الف) فیلتر موزون و (ب) فیلتر استاندارد ۳۳
- شکل ۵-۳: (الف) تصویر اصلی، (ب) اعمال فیلتر استاندارد، (پ) اعمال فیلتر موزون ۳۴
- شکل ۶-۳: (الف) تصویر نویزی شده، (ب) نتیجه اعمال فیلتر میانه ۳۴
- شکل ۷-۳: نحو تخمین مکان خط کرسی، (الف) تصویر کلمه، (ب) هیستوگرام افکنش کلمه ۳۶
- شکل ۸-۳: (الف) صفحه α, β (ب) فضای پارامتری [۲۱] ۳۷
- شکل ۹-۳: پنجره ۳ در ۳ که همسایگی‌های پیکسل $P1$ را نشان می‌دهد [۲۷] ۴۲
- شکل ۱۰-۳: دو نمونه از محاسبه تابع $A(P1)$ و $B(P1)$ در: (a) $A(P1)=1, B(P1)=2$ و در (b) $A(P1)=2, B(P1)=2$ ۴۲
- شکل ۱۱-۳: در شرایط نشان داده شده حذف $P1$ موجب گسستگی در اسکلت کاراکتر می‌شود [۲۷] ۴۳
- شکل ۱۲-۳: در (a) $A(P1) \neq 1$ ، در (b) $P2.P4.P8=0$ ، در (c) $A(P2)=1$ و $P2.P4.P8 \neq 0$ ۴۴
- شکل ۱۳-۳: در (a) $A(P4) \neq 1$ ، در (b) $P2.P4.P6=0$ ، در (c) $A(P4)=1$ و $P2.P4.P6 \neq 1$ ۴۴
- شکل ۱۴-۳: نرمال سازی با روش‌های مختلف، (الف) کاراکتر نرمال نشده، (ب) *Aspect ratio preserved*، (پ) *Square root* ۴۷
- شکل ۱۵-۳: انواع نگارش در متون دست‌نویس ۴۸
- شکل ۱۶-۳: بلوک‌بندی ثابت تصویر به منظور استخراج ویژگی ناحیه‌ای ۴۹
- شکل ۱۷-۳: بلوک بندی تطبیقی، بلوک‌بندی سطری (ردیف اول، سمت چپ) بلوک‌بندی ستونی (ردیف اول، سمت راست)، بلوک‌بندی نهایی (سطر دوم) ۵۰

- شکل ۳-۱۸: عملگرهای سوپل ۵۱
- شکل ۳-۱۹: (الف) تصویر ورودی، (ب) جهت‌های گرادیان قسمت مشخص شده کاراکتر ۵۲
- شکل ۳-۲۰: الگوی کد زنجیره‌ای، (الف) الگوی کد زنجیره‌ای چهار جهتی، (ب) الگوی کد زنجیره‌ای هشت جهتی [۲۱] ۵۳
- شکل ۳-۲۱: (الف) مرز دیجیتال، (ب) نمونه برداری مجدد، (پ) مرز با کد زنجیره‌ای ۸ جهتی [۲۱] ۵۳
- شکل ۳-۲۲: نحوه استخراج ویژگی مکان مشخصه توسعه یافته [۱۴] ۵۵
- شکل ۳-۲۳: استخراج ویژگی پروفایل، (الف) تصویر اصلی، (ب) پروفایل پایین، (پ) پروفایل بالا ۵۶
- شکل ۳-۲۴: هیستوگرام پروفایل افکنش، (الف) تصویر کلمه، (ب) افکنش افقی، (پ) افکنش عمودی ۵۷
- شکل ۳-۲۵: نمودار دندوگرام ۶۷
- شکل ۳-۲۶: معیار شباهت در *Single-Link* ۶۸
- شکل ۳-۲۷: نمودار دندوگرام *singel linke* ۷۰
- شکل ۳-۲۸: معیار فاصله در *Complete-Link* ۷۱
- شکل ۳-۲۹: نمودار دندوگرام *Complete-Link* ۷۳
- شکل ۳-۳۰: معیار فاصله در *Average-Link* ۷۴
- شکل ۳-۳۱: نمودار دندوگرام *Average-Link* ۷۶
- شکل ۴-۱: حذف اجزای کوچک، (الف) تصویر اصلی، (ب) تصویر با حذف اولیه ۸۲
- شکل ۴-۲: اجزای اصلی کلمه نمونه ۸۳
- شکل ۴-۳: (الف) سطرهای مشترک اجزای اصلی، (پ) محدوده بیشترین تعداد اجزا در راستای سطرهای اصلی ۸۳
- شکل ۴-۴: دو نمونه از بهبود تخمین مکان خط کرسی توسط روش پیشنهادی ۸۳
- شکل ۴-۵: حذف اشتباه کاراکترها ۸۴
- شکل ۴-۶: انتخاب درست اجزای اصلی ۸۴
- شکل ۴-۷: جلوگیری از محدود کردن نادرست سطرها برای یافتن خط کرسی ۸۵
- شکل ۴-۸: همپوشانی ستونی اجزای کلمه ۸۶
- شکل ۴-۹: دو نمونه از جداسازی نقاط کلمات ۸۶
- شکل ۴-۱۰: حذف سطر و ستون‌های خالی تصاویر، (الف) تصویر ورودی، (ب) تصور با حذف سطر و ستون ۸۷
- شکل ۴-۱۱: نرمال سازی کلمه، (الف) تصویر بدنه اصلی، (ب) تصویر نرمال شده اجزا، (پ) تصویر نازک سازی شده پس از اعمال فیلتر گوسی ۸۸
- شکل ۴-۱۲: اعوجاج بهبود یافته توسط فیلتر گوسی ۸۸
- شکل ۴-۱۳: نمودار تغییرات دقت به ازای تعداد مختلف همسایگی در طبقه‌بند k نزدیکترین همسایگی ۹۱
- شکل ۴-۱۴: نمودار تغییرات دقت به ازای روش‌های مختلف نرمال سازی در طبقه‌بند k نزدیکترین همسایگی ۹۲
- شکل ۴-۱۵: نمودار بازشناسی به ازای تغییرات تعداد بلوک‌ها در تصویر بدنه اصلی ۹۳
- شکل ۴-۱۶: نمودار تغییرات دقت به ازای تغییرات پارامتر C ۹۴
- شکل ۴-۱۵: بلوک‌بندی تطبیقی پیشنهاد شده ۱۰۱

فهرست جداول

| | |
|--|----|
| جدول ۱-۳: توابع نگاشت نسبت ابعاد [۲۷] | ۴۶ |
| جدول ۲-۳: فاصله داده‌ها | ۶۹ |
| جدول ۳-۳: مرحله اول خوشه‌بندی به روش <i>singel linke</i> | ۶۹ |
| جدول ۴-۳: مرحله دوم خوشه‌بندی به روش <i>singel linke</i> | ۶۹ |
| جدول ۵-۳: مرحله سوم خوشه‌بندی به روش <i>singel linke</i> | ۷۰ |
| جدول ۶-۳: مرحله چهارم خوشه‌بندی به روش <i>singel linke</i> | ۷۰ |
| جدول ۷-۳: فاصله میان داده‌ها | ۷۱ |
| جدول ۸-۳: مرحله اول خوشه‌بندی به روش <i>Complete linke</i> | ۷۲ |
| جدول ۹-۳: مرحله دوم خوشه‌بندی به روش <i>Complete linke</i> | ۷۲ |
| جدول ۱۰-۳: مرحله سوم خوشه‌بندی به روش <i>Complete linke</i> | ۷۲ |
| جدول ۱۱-۳: مرحله چهارم خوشه‌بندی به روش <i>Complete linke</i> | ۷۳ |
| جدول ۱۲-۳: فاصله بین داده‌ها | ۷۴ |
| جدول ۱۳-۳: مرحله اول خوشه‌بندی به روش <i>Average linke</i> | ۷۵ |
| جدول ۱۴-۳: مرحله دوم خوشه‌بندی به روش <i>Average linke</i> | ۷۵ |
| جدول ۱۵-۳: مرحله سوم خوشه‌بندی به روش <i>Average linke</i> | ۷۵ |
| جدول ۱۶-۳: مرحله چهارم خوشه‌بندی به روش <i>Average linke</i> | ۷۶ |
| جدول ۱-۴: نتایج حاصل از تعداد بلوک‌های مختلف تصویر بدنه اصلی | ۹۰ |
| جدول ۲-۴: نتایج حاصل از روش پیشنهادی با استفاده از روش نرمال کردن <i>Aspect ratio</i> | ۹۰ |
| جدول ۳-۴: نتایج حاصل از روش پیشنهادی با استفاده از روش نرمال کردن <i>Square root of aspect ratio</i> | ۹۰ |
| جدول ۴-۴: نتایج حاصل از روش پیشنهادی با استفاده از روش نرمال کردن <i>Cubic root of aspect ratio</i> | ۹۱ |
| جدول ۵-۴: نتایج حاصل از روش پیشنهادی با استفاده از روش نرمال کردن <i>Sine of aspect ratio</i> | ۹۱ |
| جدول ۶-۴: نتایج حاصل از بلوک بندی تطبیقی بر روی تصویر [۲۹] | ۹۳ |
| جدول ۷-۴: تاثیر درجه هسته چند جمله‌ای در دقت بازشناسی ماشین بردار پشتیبان | ۹۵ |
| جدول ۸-۴: تاثیر مقدار ثابت جمع شونده هسته چند جمله‌ای در دقت بازشناسی ماشین بردار پشتیبان | ۹۵ |
| جدول ۹-۴: تاثیر پارامتر گاما، هسته RBF در دقت بازشناسی ماشین بردار پشتیبان | ۹۵ |
| جدول ۱۰-۴: مشخصات خوشه‌های خروجی | ۹۷ |
| جدول ۱۱-۴: دقت و میزان کاهش فرهنگ لغت | ۹۷ |
| جدول ۱۲-۴: طبقه‌بندی به کمک کاهش فرهنگ لغت | ۹۸ |
| جدول ۱۳-۴: بهبود سرعت با استفاده از روشهای کاهش فرهنگ لغت | ۹۸ |

فصل ۱

مقدمه

۱-۱- مقدمه

کاغذ همواره وسیله‌ای برای ارتباط بین انسان‌ها بوده است. سادگی استفاده از کاغذ، استفاده از آن برای ارتباطات و بایگانی کردن اطلاعات، موجب شده که حجم زیادی از اطلاعات کاغذی وجود داشته باشد. اما امروزه به علت گسترش محیط الکترونیکی به دلیل کم حجم بودن و بالا بودن سرعت دسترسی به اطلاعات در این محیط، علاقه به استفاده از آن برای برقراری ارتباطات و بایگانی کردن اطلاعات روز به روز بیشتر می‌شود. از این رو برای تبدیل اطلاعات کاغذی به صورت الکترونیکی باید روشی ایجاد گردد.

مسئله بازشناسی متون دست‌نوشته به علت پیچیدگی و ابهامات و همچنین اهمیت این موضوع در فرآیند خودکارسازی یکی از موضوعات مورد علاقه در زمینه بازشناسی الگو است. بازشناسی متن به معنی تبدیل اسناد متنی به شکل قابل ویرایش است که می‌تواند در زمینه پردازش کلمات، خواندن فرم‌های آماده و مبلغ چک، خواندن آدرس و دسته بندی مرسولات پستی به کار رود.

سیستم‌های شناسایی متون را می‌توان بر اساس نحوه نگارش (تایپی و دست‌نوشته)، نوع داده ورودی (برخط^۱ و برون خط^۲)، متصل بودن کاراکترها (کاراکترهای مجزا و کلمات پیوسته) به گروه‌های مختلفی تقسیم‌بندی کرد. با توجه به شرایط و نوع کاربرد، مجموعه داده‌های مورد نیاز برای آموزش و آزمایش هر یک از این سیستم‌های شناسایی متفاوت خواهد بود. این مجموعه‌ها می‌توانند از واحدهای کوچک نوشتاری مانند ارقام و حروف مجزا شروع شده و تا حد کلمه و حتی جملات و پاراگرافها توسعه یابند.

متون به دو دسته کلی تایپی و دست‌نوشته تقسیم می‌شوند. متون دست‌نوشته به آسانی و با ساده‌ترین ابزارها مانند کاغذ و قلم نوشته می‌شوند و در نتیجه از نوشتارهای تایپی متداول‌تر و فراگیرتر هستند.

¹ On-Line

² Off-Line

۲-۱- داده‌های برخط و برون خط

اولین مرحله در هر سیستم شناسایی متن، دریافت داده و تبدیل آن به شکل دیجیتال قابل استفاده در کامپیوتر است. این سیستم‌ها می‌توانند متن ورودی را به دو شکل برخط و برون خط دریافت کنند. در سیستم‌های شناسایی متون برخط، کاربر با استفاده از ابزارهای مخصوص مانند قلم و صفحه دیجیتال کننده^۱ اقدام به نوشتن متن می‌کند. در این گونه از نوشتار، اطلاعات زیادی در مورد متن، مانند تعداد پاره‌خط‌های کلمه^۲، مسیر حرکت قلم و جهت و سرعت نوشتن در دسترس است. در روش برخط، عمل شناسایی نوشتار بصورت همزمان با عمل نوشتن انجام می‌شود. با توجه به خصوصیت ذاتی سیستم‌های برخط در نحوه جمع آوری دادگان ورودی، کاربرد آنها به شناسایی متون دست‌نوشته مانند کاراکترهای مجزا، کلمات پیوسته، فرمولها یا عبارات ریاضی محدود می‌شوند.

در مقابل، سیستم‌های برون خط وظیفه شناسایی متون دست‌نوشته و یا تایپ شده بر روی کاغذ را برعهده دارند. بنابراین، این سیستم‌ها به اطلاعات زمانی موجود در دادگان برخط دسترسی نخواهند داشت. در این سیستم‌ها، متون نوشته شده و یا تایپی بوسیله یک دوربین یا پوشگر^۳ به تصاویر دیجیتالی تبدیل می‌شوند.

۳-۱- مراحل اصلی یک سیستم بازشناسی

۱-۳-۱- پیش پردازش

پردازش‌های لازم برای بهبود اولیه کیفیت تصویر مانند حذف نویز و تصحیح کجی خط کرسی و باینری کردن جزء این مرحله می‌باشد.

¹ Digitizer Tablet

² Strokes

³ Scanner

۱-۳-۲- استخراج ویژگی

در این مرحله از اطلاعات تصویری کلمه مجموعه‌ای از ویژگی‌ها جهت طبقه‌بندی در رایانه استخراج می‌شود.

۱-۳-۳- طبقه‌بندی

ویژگی‌های استخراج شده از تصویر کلمه در این مرحله جهت شناسایی وارد سیستم می‌شوند. شناسایی در این مرحله به معنی طبقه‌بندی آنها در یک مجموعه حروف و یا کلمات از پیش تعریف شده (دایره لغات^۱) می‌باشد. طبقه‌بندی می‌تواند بوسیله روش‌های متعددی مانند استفاده از شبکه عصبی، مدل مخفی مارکوف، k نزدیکترین همسایگی، ماشین بردار پشتیبان انجام شود.

۱-۴- روش‌های بازشناسی متن

سیستم‌های بازشناسی متن را با توجه به روش بازشناسی می‌توان به دو دسته کلی زیر تقسیم کرد:

- روش‌های تجزیه‌ای^۲
- روش‌های کلی‌نگر^۳

روش‌های تجزیه‌ای به جداسازی کلمات و زیر کلمات می‌پردازند و از واحدهای کوچکتر حاصل برای شناسایی کلمات استفاده می‌کنند. منظور از جداسازی، تفکیک تصویر زیر-کلمه ورودی به مولفه‌های میانی، انتهایی و مجزاست. بدلیل خطایی که در روش‌های جداسازی وجود دارد، این خروجی می‌تواند یک شبه حرف باشد. در برخی سیستم‌ها سعی می‌شود با تنظیم روش‌های جداسازی و انجام پس پردازش‌های مناسب احتمال این رخداد به حداقل برسد. اینکار خود باعث کاهش دقت سیستم می‌شود. در برخی

^۱ Lexicon

^۲ Analytical

^۳ Holistic

سیستم‌ها این شبه حرف معتبر معرفی شده و در مرحله بازشناسی تعیین تکلیف می‌شوند. [۱] از مزیت‌های این روش این است که شناسایی محدود به دایره لغات از پیش تعریف شده نمی‌شود.

روش‌های زیادی برای جداسازی حروف یا شبه حروف زیر-کلمات بکار گرفته شده است که برخی از مهمترین این روش‌ها عبارتند از:

- استفاده از هیستوگرام عمودی [۲]
- استفاده از منحنی پیرامون کلمه [۳]
- استفاده از پروفایل بالایی کلمات [۴]
- جداسازی بر اساس قواعد ساختاری یک رسم الخط خاص [۵]
- جداسازی با استفاده از محل اتصال به خط زمینه و مفهوم قطعه خط متصل [۶]
- جداسازی با نازک‌سازی حروف [۷]
- جداسازی با استفاده از برجسب زدن به منحنی بالایی زیر-کلمات [۸]

مرحله پس پردازش روش‌های جداسازی با توجه به ویژگی‌های خاص نگارش فارسی، اعمال قواعد ساختاری، ویژگی‌های تصویر و نتایج حاصل از بازشناسی و تعیین نقاط یک شبه کلمه یا حتی نتایج حاصل از شناسایی اولیه آن به اصلاح نتایج جداسازی و رفع خطاهای حاصل از آن می‌پردازد.

اما در روش‌های کلی‌نگر، کلمات از روی شکل عمومی آنها و بدون انجام عمل تقطیع تشخیص داده می‌شوند. در این روش، هر کلمه در فرهنگ کلمات به عنوان یک کلاس جداگانه در نظر گرفته می‌شود. روش‌های کلی‌نگر یکی از متداولترین روش‌های شناسایی کلمات در کاربردهایی با مجموعه واژگان محدود و ثابت هستند. از آنجاییکه در این روش، هر کلمه از فرهنگ لغت به عنوان یک کلاس مجزا در نظر گرفته شده است، نمونه‌های لازم برای آموزش مدل مربوط به هر کلمه باید به اندازه کافی جامع باشند. به این

معنی که نمونه‌های آموزش باید بتوانند کلیه تاثیرات متقابل بین حروف یک کلمه را در شیوه‌های مختلف نوشتن مدل کنند. در برخی از کاربردها مانند شناسایی آدرسهای پستی که در آن فرهنگ کلمات بزرگ و یا متغیر است، توانایی روش‌های کلی‌نگر برای ایجاد تمایز بین کلاس‌های مختلف بدلیل افزایش شباهت بین آنها، به شدت کاهش می‌یابد. بعلاوه، از نظر عملی، در چنین کاربردهایی جمع‌آوری نمونه‌های لازم از هر کلاس برای آموزش سیستم شناسایی، بسیار دشوار است. بنابراین هنگامیکه تعداد نمونه‌های آموزش موجود در هر کلاس به اندازه کافی زیاد نباشد، باید از روش‌های تجزیه‌ای به جای روش‌های کلی‌نگر برای آموزش سیستم شناسایی استفاده کرد. به طور کل تعداد نمونه‌های لازم، بستگی به پارامترهایی مانند اندازه ویژگی‌های^۱ استخراج شده، دقت مورد انتظار از سیستم شناسایی، پیچیدگی طبقه‌بندی کننده^۲ (بر اساس تعداد پارامترهایی که باید در آن تخمین زده شود) و احتمال خطای بین کلاس‌ها دارد.

مشکل اصلی در رویکرد مبتنی بر بازشناسی بدون جداسازی و استفاده از شکل کلی کلمه، تعداد زیاد کلاس‌هاست. در مسائل شناسایی الگو، افزایش تعداد کلاس به معنی کاهش میزان دقت در بازشناسی و بالا رفتن زمان پردازش سیستم است. زیر سیستم فرهنگ لغت می‌تواند با خوشه‌بندی اولیه زیرکلمات در کاهش تعداد کلاس‌ها کمک موثری کند. به این ترتیب که برای هر کلمه و یا زیر-کلمه ورودی، با استفاده از ویژگی‌های استخراج شده از شکل کلی آن، خوشه و یا خوشه‌های نزدیک به آن در فرهنگ لغت انتخاب شده و مسئله شناسایی به نمونه‌های این خوشه‌ها محدود می‌شود. فرهنگ لغت تنها روش برای انجام دسته‌بندی اولیه بر روی کلمات و یا زیرکلمات نیست این دسته‌بندی می‌تواند بر اساس ویژگی‌های ساختاری کلمات بوده و از مفاهیم نحوی برای اینکار استفاده کند. در نهایت آنچه اهمیت دارد این است که بتوان یک دسته‌بندی اولیه و با دقت بالا برای کاهش تعداد کلاس‌ها ایجاد کرد.

¹ Feature

² Classifier

۱-۵- اندازه فرهنگ لغت

اندازه فرهنگ کلمات، یکی از پارامترهای مهم یک سیستم شناسایی دست‌نوشته است. معمولاً سیستم‌های شناسایی نوشتار، بمنظور کاهش ابهامات موجود در کلمات ورودی و افزایش درصد شناسایی، از یک مجموعه به نام فرهنگ کلمات استفاده می‌کنند. در حقیقت، فرهنگ کلمات مجموعه‌ای از واژگان است که از سیستم شناسایی انتظار می‌رود بتواند آنها را به درستی تشخیص دهد.

اگرچه تعریف معینی برای اندازه فرهنگ کلمات وجود ندارد، اما در حالت کلی تقسیم‌بندی به صورت زیر است:

- فرهنگ کلمات کوچک (شامل ده‌ها کلمه)
- فرهنگ کلمات متوسط (شامل صدها کلمه)
- فرهنگ کلمات بزرگ (شامل هزاران کلمه)
- فرهنگ کلمات خیلی بزرگ (شامل صدها هزار کلمه)

اندازه فرهنگ کلمات با توجه به کاربرد سیستم بازشناسی تغییر می‌کند. به عنوان مثال، در سیستم‌های بازشناسی ارقام و حروف، فرهنگ کلمات کوچک می‌باشد. سیستم در شناسایی مبالغ نوشته شده به حروف روی چک‌های بانکی دارای فرهنگ کلمات متوسط می‌باشد. سیستم بازشناسی اسامی شهرها در کاربرد خواندن آدرس‌های پستی، با یک فرهنگ کلمات بزرگ روبرو می‌باشیم.

۱-۶- کارایی سیستم‌های بازشناسی متن

در هر سیستم بازشناسی متن معیارهایی برای ارزیابی کارایی سیستم وجود دارد مانند نرخ بازشناسی که برابر است با تعداد حروف، زیر-کلمه و یا کلماتی که به درستی بازشناسی می‌شوند به تعداد کل نمونه‌هایی که برای آزمایش سیستم استفاده شده‌اند و یا نرخ خطا که برابر تعداد حروف، زیر-کلمه و یا کلماتی است

که نسبت به کل داده‌های آزمایش به اشتباه بازشناسی شده‌اند. خطای بازشناسی ممکن است در مراحل مختلف فرایند بازشناسی اتفاق بیفتد به عنوان مثال خطا ممکن است مرتبط با وجود نویز در تصویر، روشنایی ناهمگن، کجی کلمه و مواردی از این قبیل باشد. اگر اشکالات اشاره شده در مرحله پیش پردازش اصلاح نشوند، ممکن است در مراحل بعدی موجب ایجاد خطا در سیستم گردند. اگر سیستم دارای مرحله جداسازی به منظور تجزیه کلمات به واحدهای کوچکتر باشد و این جداسازی به صورت درست انجام نگیرد، نرخ بازشناسی مناسبی به دست نمی‌آید. در مرحله استخراج ویژگی انتخاب ویژگی‌هایی که شکل کلمه و یا زیر-کلمه را به خوبی توصیف کند، در بالا بردن دقت بازشناسی موثر خواهد بود.

۷-۱- معرفی ویژگی‌های خط فارسی

آشنایی با قواعد نگارش فارسی در انتخاب روش‌های مناسب برای بازشناسی متون نقش اساسی دارد. در ادامه به قواعد مهم نگارش فارسی اشاره می‌شود.

- زبان فارسی از ۳۲ حرف تشکیل شده است. بر خلاف نوشتار لاتین، متون فارسی هم در حالت تایپی و هم در حالت دست‌نویسته به صورت پیوسته و از راست به چپ نوشته می‌شود. نگارش فارسی مانند نگارش لاتین بر اساس کلمه است، به این معنی که بین کلمات معمولاً فاصله وجود دارد. در کلمات فارسی، بیشتر حروف از دو طرف به حروف مجاور خود می‌چسبند. بنابراین هر کلمه می‌تواند شامل یک یا چند بخش باشد که به آن‌ها زیر-کلمه^۱ می‌گویند. هر کلمه از یک یا چند زیر-کلمه تشکیل می‌شود و هر زیر-کلمه از یک یا چند حرف چسبیده به هم تشکیل می‌شود. مطابق قواعد زبان فارسی باید بین کلمات فاصله وجود داشته باشد و بین زیر-کلمات فاصله اضافه وجود نداشته باشد، ولی در بسیاری از موارد این قواعد رعایت نمی‌شود.

¹ Subword

- حروف فارسی بسته به موقعیت آنها در زیر-کلمه می‌توانند تا چهار شکل متفاوت داشته باشند: حروف ابتدایی، میانی، انتهایی و مجزا. شکلهای مختلف حروف الفبای فارسی در شکل ۱-۱ نشان داده شده است.
- ۱۷ حرف از ۳۲ حرف فارسی نقطه دارند. تعداد نقاط این حروف بین یک تا سه متغیر است. نقاط ممکن است در بالا، پایین یا داخل بدنه حرف قرار گیرند. به بخشی از حرف یا زیر-کلمه که نقاط آن حذف شده باشند، بدنه حرف یا زیر-کلمه گفته می‌شود.
- حروف در یک کلمه فارسی می‌توانند حتی بدون اتصال با هم، همپوشانی عمودی داشته باشند. معمولاً حروف (ا، د، ذ، ر، ز، ژ، ع، و) با حرف مجاور خود دارای همپوشانی هستند. علاوه بر این حروف، قسمتهای مکمل حروف (آ، ک) نیز می‌توانند بخشی از حروف مجاور خود را بپوشانند. این ویژگی در نوشتار فارسی عمل تقطیع کلمات به حروف تشکیل دهنده آنها را مشکل می‌کند.
- تفاوت شکل برخی از حروف مانند {ب، پ، ت، ث} تنها در تعداد یا محل قرار گرفتن نقاط آنهاست.
- اندازه تمام حروف فارسی یکسان نیستند. مثلاً حروف (ب) و (س) در حالت جدا اندازه بزرگتری نسبت به حروف (د) و (ه) دارند. این تنوع در اندازه حروف، کار قطعه بندی حروف را مشکل می‌کند.
- در برخی از شیوه‌های نوشتاری زبان فارسی، دو یا چند حرف کنار هم می‌توانند به گونه‌ای با هم ترکیب شوند که شکل حاصل شباهتی به حروف تشکیل دهنده آن ندارد. به این ترکیب، حروف ادغام شده^۱ می‌گویند. چنین مواردی نه تنها در نوشتار دست‌نویس، بلکه در متون تایپی نیز وجود دارد. متداولترین ترکیب در متون تایپی، ادغام دو حرف (ل) و (ا) بصورت (لا) است. در نوشته‌های

¹ *Ligature*

دست‌نوشته فارسی نیز به خاطر زیبایی بصری نوشتار و همچنین سلیقه نویسنده، شکل کلی بعضی از حروف کنار هم، بکلی تغییر می‌کند. نمونه‌هایی از این حروف ادغام شده در شکل ۱-۲ نشان داده شده است.

| انتها | میان | ابتدا | مجزا | | انتها | میان | ابتدا | مجزا | |
|-------|------|-------|------|----|-------|------|--------|--------|----|
| ص | ص | ص | ص | ۱۷ | ا | --- | ا یا آ | ا یا آ | ۱ |
| ض | ض | ض | ض | ۱۸ | ب | ب | ب | ب | ۲ |
| ط | ط | ط | ط | ۱۹ | پ | پ | پ | پ | ۳ |
| ظ | ظ | ظ | ظ | ۲۰ | ت | ت | ت | ت | ۴ |
| ع | ع | ع | ع | ۲۱ | ث | ث | ث | ث | ۵ |
| غ | غ | غ | غ | ۲۲ | ج | ج | ج | ج | ۶ |
| ف | ف | ف | ف | ۲۳ | چ | چ | چ | چ | ۷ |
| ق | ق | ق | ق | ۲۴ | ح | ح | ح | ح | ۸ |
| ک | ک | ک | ک | ۲۵ | خ | خ | خ | خ | ۹ |
| گ | گ | گ | گ | ۲۶ | د | --- | --- | د | ۱۰ |
| ل | ل | ل | ل | ۲۷ | ذ | --- | --- | ذ | ۱۱ |
| م | م | م | م | ۲۸ | ر | --- | --- | ر | ۱۲ |
| ن | ن | ن | ن | ۲۹ | ز | --- | --- | ز | ۱۳ |
| و | --- | --- | و | ۳۰ | ژ | --- | --- | ژ | ۱۴ |
| ه | ه | ه | ه | ۳۱ | س | س | س | س | ۱۵ |
| ی | ی | ی | ی | ۳۲ | ش | ش | ش | ش | ۱۶ |

شکل ۱-۱: شکل‌های مختلف حروف فارسی با توجه به موقعیت آنها در زیر-کلمه

لا ء کا ل لہ لہ لہ لہ لہ لہ

لاها کا را ان بی به ره

شکل ۲-۱: حروف ادغام شده در متون دست‌نوشته

ساختار پایان نامه:

در این تحقیق روشی برای بازشناسی کلمات دست‌نوشته فارسی با استفاده از شکل کلی کلمات ارائه شده است. بدین منظور در فصل دوم تحقیقات انجام شده در زمینه بازشناسی کلمات دست‌نوشته و تایپی، فارسی و عربی که از شکل کلی کلمه استفاده می‌کنند به طور مختصر مورد بررسی قرار می‌گیرند. سپس در فصل سوم مباحث تئوری در این زمینه شامل روش‌های مربوط به پیش پردازش، استخراج ویژگی و خوشه‌بندی را توضیح می‌دهیم به این ترتیب که ابتدا روشهای پیش پردازش مورد بررسی قرار می‌گیرند سپس روشهای استخراج ویژگیهای متداول در این زمینه و نحوه محاسبه آنها را توضیح می‌دهیم و در نهایت روشهای خوشه‌بندی نمونه‌ها که به منظور کاهش فرهنگ لغت مورد استفاده قرار می‌گیرند بررسی می‌شوند. در فصل چهارم و پنجم به توضیح روش ارائه شده و نتایج آن با استفاده از طبقه‌بندهای مختلف و روشهای کاهش فرهنگ لغت در مقایسه با روش‌های مشابه می‌پردازیم.

فصل ۲

مرور کارهای انجام شده

۲-۱- مقدمه

در این فصل تحقیقات انجام شده در زمینه بازشناسی کلمات دست‌نوشته و تایپی، فارسی و عربی که از شکل کلی کلمه استفاده می‌کنند به طور مختصر مورد بررسی قرار می‌گیرند. همانطور که در فصل قبل نیز اشاره شده است در روشهای کلی نگر، کلمات از روی شکل عمومی آنها و بدون انجام عمل تقطیع تشخیص داده می‌شوند که یکی از متداولترین روش‌های شناسایی کلمات در کاربردهایی با مجموعه واژگان محدود و ثابت هستند.

۲-۲- بازشناسایی بر مبنای شکل کلی کلمه یا زیر کلمه

در این روش‌ها کل تصویر کلمه یا زیر-کلمه به صورت یک الگو در نظر گرفته می‌شود و ویژگی‌های لازم برای طبقه‌بندی تصویر استخراج می‌شوند. این ویژگی‌ها می‌توانند ساختاری و یا آماری باشند. از ویژگی‌های ساختاری می‌توان به تعداد و نوع پاره‌خط‌های مختلف (عمودی، افقی، بالا رونده‌ها و پایین رونده) اشاره کرد. گشتاورها و مکان مشخصه نمونه‌هایی از ویژگی‌های آماری است. در اینجا به توضیح مختصر بعضی تحقیقاتی که در زمینه بازشناسی کلمات بر اساس شکل کلی آنها و ویژگی‌های به کار رفته در این زمینه انجام شده‌اند، می‌پردازیم.

استخراج ویژگی به نحوی که نسبت به چرخش و مقیاس تصاویر کلمه پایدار باشند یکی از مسائل ضروری در تحلیل تصاویر متنی است. برومند نیا در تحقیقی به معرفی دیدگاه موثر برای بازشناسی کلمات دست‌نوشته پرداخته است که نسبت به چرخش و تغییرات مقیاس پایدار است [۹]. در این دیدگاه برای استخراج ویژگی از کلمات فارسی، از تبدیل موجک^۱ استفاده شده است. به منظور تغییر ناپذیر بودن ویژگی استخراج شده نسبت به چرخش و اندازه کلمات از تصویر قطبی نرمالیزه شده آنها استفاده شده است. تبدیل قطبی نسبت به اندازه و جا به جایی مقاوم و با چرخش تصویر، به صورت ردیفی انتقال

¹ *M-band packet Wavelet Transform*

می‌یابد (شکل ۱-۲). ضرائب موجک خروجی نسبت به اندازه و چرخش پایدارند. برای هر زیر باند از این ضرایب موجک مجموعه‌ای از ویژگی‌های انرژی محلی محاسبه می‌شود و ویژگی‌ها از ضرائب موجک این زیر باندها استخراج می‌شود.



شکل ۱-۲: (الف) تصویر کلمه با چرخش و (ب) تصویر قطبی به دست آمده [۹]

در این تحقیق برای مقایسه بردارهای ویژگی استخراج شده از کلمات ورودی و کلمات فرهنگ لغت از معیار فاصله ماهالانویس^۱ استفاده شده است. نرخ بازشناسی برای این ویژگی ۹۲٪ گزارش شده که نسبت به روش‌های مشهور فوریه و گشتاورهای زرنیکی که در برابر چرخش پایدارند ۱۲٪ افزایش پیدا کرده است.

در مرجع [۱۰] *sagheer* روشی را برای بازشناسی کلمات اردو با استفاده از شکل کلی آنها ارائه کرده است. زبان اردو همانند زبان فارسی و عربی است اما تعداد حروفش بیشتر می‌باشد. در اینجا پس از پیش پردازش به منظور حذف نویز و نرمال کردن اندازه تصاویر، سعی بر استخراج ویژگی با استفاده از ویژگی ساختاری و گرادیان مبتنی بر شکل کلی کلمه هستند و نیازی به جداسازی کلمات به اجزای سازنده آنها نمی‌باشد، شده است.

بعد از انجام پردازش‌های اولیه دو نوع ویژگی استخراج می‌شود :

¹ Mahalanobis

ویژگی گرادیان: یک ویژگی جهتی است که از گرادیان تصویر خاکستری^۱ استخراج می‌شود. نحوه استخراج این ویژگی به این صورت است که ابتدا اندازه تصاویر خاکستری نرمال می‌شود سپس با اعمال فیلتر روبرتز یا سوبل یا کریش، اطلاعات گرادیان برای هر پیکسل تصویر استخراج می‌شود. در این تحقیق از عملگر روبرتز استفاده شده است. سپس تصویر به بلوک‌هایی با اندازه ثابت تقسیم می‌شود و در هر بلوک اندازه گرادیان در جهت‌های کوانتیزه شده با هم جمع می‌شوند. شکل ۲-۲ تصویر یک کلمه و ویژگی‌های آن را نمایش می‌دهد. [۱۰]



شکل ۲-۲: (الف) کلمه ورودی، (ب) اندازه گرادیان و (پ) زاویه گرادیان به دست آمده از آن [۱۰]

ویژگی ساختاری: در این تحقیق از ویژگی پروفایل بالایی که معرف ساختار بیرونی کلمه است، به عنوان ویژگی ساختاری استفاده شده است.

برای طبقه‌بندی ۵۷ کلاس از یک پایگاه داده شامل ۵۷ کلمه از طبقه‌بند ماشین بردار پشتیبان، SVM استفاده شده است و نرخ بازشناسی ۹۷٪ گزارش شده است.

در تحقیق دیگری که توسط *Alkhateeb* انجام گرفته، روشی برای بازشناسی کلمات دست‌نوشته عربی که شامل سه مرحله پیش پردازش، استخراج ویژگی و طبقه‌بندی است، ارائه شده است [۱۱]. در مرحله پیش پردازش به منظور کاهش تغییرات بین تصاویر دست‌نوشته، اندازه تصاویر نرمال سازی می‌شوند سپس برای استخراج ویژگی تصاویر به بلوک‌هایی با همپوشانی و اندازه ثابت تقسیم بندی شده و از هر بلوک مجموع تعداد پیکسل‌های پیش زمینه به تعداد کل پیکسل‌های بلوک محاسبه می‌شود. اندازه هر بلوک ۱۲ پیکسل

^۱ Grayscale Image

و میزان همپوشانی بلوک‌ها ۲ پیکسل انتخاب شده که این مقادیر به صورت تجربی حاصل شده است. در مرحله طبقه‌بندی تصاویر از طبقه‌بند K نزدیکترین همسایگی با معیار فاصله اقلیدوسی استفاده شده که نتایج حاصل بر روی ۳۲۴۹۲ تصویر از ۹۳۷ کلاس، ۰۴/۷۶٪ گزارش شده است.

Alkhateeb در تحقیق دیگری، تاثیر طبقه‌بند شبکه عصبی NN و k نزدیکترین همسایگی را در بازشناسی کلمات دست‌نویس عربی بررسی کرده است [۱۲]. الگوریتم پیشنهادی شامل سه مرحله است: پیش پردازش، استخراج ویژگی، طبقه‌بندی.

مرحله پیش پردازش شامل: آستانه گذاری برای باینری کردن تصاویر اسکن شده، حذف نویز برای بهبود تصاویر، تصحیح شیب و کجی^۱ و نازک سازی^۲ می‌باشد. در مرحله استخراج ویژگی، پنج نوع ویژگی مبتنی بر شکل کلی کلمات استخراج می‌شود:

۱- بلوک‌های هم پوشانی^۳ (OB): ابتدا لبه‌های تصویر توسط فیلتر بالا گذر عمودی و افقی پربویت^۴

استخراج می‌شود. تصویر فیلتر شده به بلوک‌هایی با اندازه 12×12 با دو پیکسل هم‌پوشانی تقسیم می‌شود. از هر بلوک میانگین تعداد پیکسل‌های سیاه محاسبه می‌شود.

۲- DCT :^۵ ضرایب DCT به کاهش اضافات در تصویر کمک می‌کند. ضرایب DCT تصویر $F(x,y)$

توسط DCT دو بعدی محاسبه می‌شود:

$$f(u,v) = \alpha(u)\alpha(v) \sum_{i=1}^{I-1} \sum_{j=1}^{J-1} f(i,j) \cos\left[\frac{(2i+1)u}{2I}\right] \cos\left[\frac{(2j+1)v}{2J}\right] \quad (1-2)$$

$f(u,v)$ تصویر کلمه است.

¹ Slope And Slant Correction

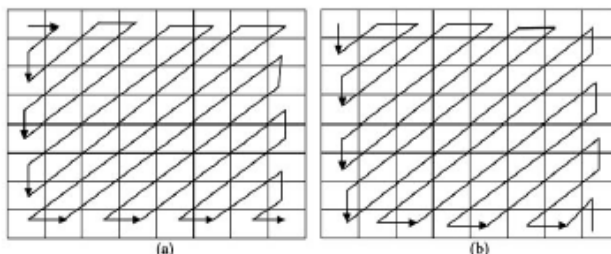
² Thinning

³ Overlapped Blocks

⁴ Prewitt

⁵ Discrete Cosine Time

۳- بعد از اعمال تبدیل DCT بر روی تصویر، ۱۰۰ ضریب مهم DCT با ترتیب زیگزاگ به عنوان بردار ویژگی استفاده می‌شود. به دلیل متمرکز شدن بیشترین انرژی در سمت چپ بالای تصویر ترتیب زیگزاگ انتخاب شده و به دو صورت انجام می‌شود (شکل ۲-۳)



شکل ۲-۳: دو نوع ترتیب زیگزاگ برای استخراج ضرایب DCT [۱۲]

۴- DCT بلوکی ($b-DCT$): ابتدا تصویر به بلوک‌هایی تقسیم می‌شود سپس از هر بلوک تعدادی از ضرایب DCT استخراج می‌شود.

۵- گشتاورهای نامتغییر: از روی تصویر نازک شده لگاریتم گشتاورهای نامتغییر هیو [۱۳] به صورت زیر استخراج می‌شوند.

$$\eta_{pq} = \sum_i \sum_j x^p y^q f(i, j) \quad (2-2)$$

پارامترهای p, q مرتبه گشتاور را نشان می‌دهند.

۶- DWT : تبدیل موجک گسسته دو بعدی برای استخراج این ویژگی‌ها استفاده می‌شود. ضرایب موجک اطلاعات خوبی در زمینه مشخصات مکانی و فرکانسی تصویر ارائه می‌دهند. در این تحقیق هر تصویر به چهار سطح از تبدیل موجک تجزیه می‌شود.

در این تحقیق از دو طبقه‌بند شبکه عصبی و k نزدیکترین همسایه بر روی پایگاه داده *IFN/ENIT* استفاده شده است. تعداد کلاس‌ها در این پایگاه داده ۹۳۷ متناسب با تعداد کلمات و شامل ۳۲۴۹۴ تصویر دست‌نوشته از این کلمات است.

تعداد لایه‌های خروجی در شبکه عصبی، به دلیل زیاد بودن تعداد کلاس‌ها، برابر با معادل باینری آنها می‌باشد.

بهترین نرخ بازشناسی برای بردار ویژگی مبتنی بر ضرایب *DCT* همراه با شبکه عصبی گزارش شده است. خسروی در [۱۴] دو رویکرد برای بازشناسی متون فارسی با استفاده از شکل کلی آنها ارائه کرده است. رویکرد اول بر بازشناسی زیر کلمات، بدون حذف نقاط و علائم آنها استوار است و رویکرد دوم مبتنی بر شکل بدنه زیرکلمات است که از حذف نقاط و علائم زیر کلمات حاصل می‌شود. در هر دو رویکرد در مرحله آموزش، زیرکلمات مجموعه آموزش خوشه‌بندی می‌شوند که برای خوشه‌یابی از الگوریتم *ISODATA* استفاده شده و مراکز اولیه خوشه‌ها توسط یک الگوریتم خوشه‌یابی سلسله‌مراتبی تهیه شده‌اند. در مرحله خوشه‌یابی از ویژگی میانگین بلوکی و معیار فاصله بلوک شهری^۱ استفاده شده است. در روش خوشه‌یابی سلسله‌مراتبی پس از استخراج ویژگی در مرحله اول فاصله بین تمام نمونه‌ها محاسبه می‌شود و در مرحله بعد از روی بردارهای فاصله، نمونه‌های نزدیک به هم پیدا می‌شوند و با یکدیگر ترکیب می‌شوند تا تشکیل یک خوشه دهند. در دوره‌های^۲ بعد بردارهای میانگین خوشه‌های جدید به عنوان ورودی در نظر گرفته شده و مراحل اول و دوم تکرار می‌شوند.

برای محاسبه ویژگی میانگین بلوکی، پس از تقسیم بندی تصویر به بلوک‌هایی با اندازه ثابت، میانگین پیکسل‌های پیش‌زمینه هر بلوک محاسبه شده و در کنار هم قرار می‌گیرد و در نهایت اندازه بردار حاصل به ۱ نرمال می‌شود. فاصله بلوک شهری نیز به صورت زیر محاسبه می‌شود: [۱۴]

¹ City block

² Epoch

$$cb = \sum_{i=1}^n |f(i) - g(i)| \quad (3-2)$$

در معادله بالا f و g دو بردار n بعدی و cb فاصله این دو بردار بر اساس معیار فاصله بلوک شهری است. در مرحله بازشناسی پس از یافتن خوشه‌های نزدیک به زیر کلمه ورودی، ویژگی‌های جدیدی که جزئیات بیشتری از زیر کلمه را نشان می‌دهند، استخراج می‌شوند تا نزدیکترین زیر کلمه به نمونه ورودی بدست آید. بدین منظور از مکان مشخصه و هیستوگرام گرادیان استفاده شده است. در مورد بازشناسی زیر کلمات بدون نقطه پس از یافتن بدنه زیر کلمه سه روش برای بازشناسی نهایی استفاده شده است: در روش اول، پس از یافتن بدنه، نقاط را به جای خود برگردانده و پس از استخراج ویژگی‌های مجدد با حضور نقاط، در بین جایگزین‌ها نزدیکترین (مشابه‌ترین) گزینه انتخاب می‌گردد. در روش دوم بدنه زیر کلمات که در بین جایگزین‌ها مشترک است، حذف شده و استخراج ویژگی، تنها با حضور نقاط به دست می‌آید و بازشناسی بر اساس آن صورت می‌گیرد. در روش سوم پس از یافتن بدنه زیر کلمات اطلاعات نقاط زیر کلمه ورودی مانند تعداد و مکان آنها نسبت به خط کرسی، با زیر کلمات جایگزین مقایسه می‌شوند و بر این اساس زیر کلمه نهایی تعیین می‌شود.

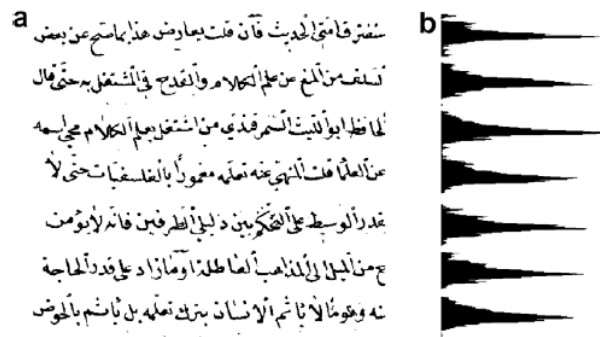
هر دو روش نتایج قابل قبولی روی تصاویر تمیز ارائه می‌دهند به طوری که روش بانقطه دقتی حدود ۹۴٪ و روش بدون نقطه دقتی حدود ۹۳٪ در سطح کلمه ارائه می‌دهند. اما در برخورد با تصاویر کم کیفیت و نویزی دچار افت دقت می‌شوند که این کاهش در برخی موارد بسیار شدید است.

Aghbari روشی جدید بر اساس شکل کلی برای طبقه‌بندی و بازیابی متون دست‌نوشته تاریخی عربی ارائه کرده است [۱۵]. مراحل بازشناسی به این ترتیب است که در ابتدا با استفاده از یک آستانه کلی تصویر سطح خاکستری متون به تصویر باینری تبدیل می‌شود. مقدار این آستانه از میانگین گرفتن حداکثر شدت و حداقل شدت مقدار پیکسل‌ها در سطح خاکستری تصویر به دست می‌آید. پس از تبدیل تصویر

خاکستری به تصویر باینری به منظور حذف نویز از فیلتر میانه^۱ با اندازه ۳×۳ استفاده شده است. در مرحله بعد به کمک فیلتر لاپلاسیین سعی به از بین بردن شکستگی‌های بدنه کلمات شده است و در مرحله پایانی پیش پردازش، تصویر متنی نازک سازی می‌شوند.

جداسازی متون شامل سه مرحله است:

- تصویر متنی با استفاده از افکنش افقی به تصویر خطوط متنی تبدیل می‌شوند به این ترتیب که فواصل بین پیکها در نمودار افکنش افقی که بیانگر خطوط کرسی است محاسبه می‌شود. مقدار این میانگین محدوده هر خط را با توجه به خط کرسی آن مشخص می‌کند. نمونه‌ای از محاسبه افکنش افقی محاسبه شده برای متن در شکل ۲-۴ نشان داده شده است.



شکل ۲-۴: افکنش افقی متن به منظور جداسازی به خطوط [۱۵]

- در مرحله دوم با توجه به فاصله بین کلمات، خطوط متنی به کلمات مجزا تقسیم می‌شوند. بدین منظور هیستوگرام عمودی خطوط محاسبه شده و بر اساس تعداد ستون‌های صفر ایجاد شده، کلمات هر خط استخراج می‌شوند. نمونه‌ای از این هیستوگرام در شکل ۲-۵ ترسیم شده است.

¹ Median



شکل ۲-۵: افکنش عمودی یک خط متنی [۱۵]

• در این مرحله نیز همانند مرحله‌ی پیش، از فاصله بین اجزای کلمه برای جدا کردن زیرکلمات استفاده شده است. برای از بین بردن همپوشانی بین اجزا کلمه، کلمه قبل از این مرحله کشیده می‌شود.

به منظور استخراج ویژگی از هر زیر کلمه پروفایل بالا، پروفایل پایین و هیستوگرام افکنش به عنوان ویژگی‌های آماری و تعداد نقاط بالا و پایین خط کرسی و نسبت طول نقاط به طول بدنه زیر کلمه به عنوان ویژگی‌های ساختاری به دست آمده‌اند. در نتیجه برای هر کلمه، خروجی مرحله استخراج ویژگی، تعدادی بردار ویژگی متناسب با تعداد زیر کلمات آن است که به ترتیب در کنار یکدیگر قرار می‌گیرند. لذا طول بردار ویژگی در کلمات مختلف متفاوت خواهد بود. با توجه به اینکه کلمات پایگاه داده مورد بررسی حداکثر شامل ۵ زیر کلمه‌اند برای یکسان سازی طول بردار ویژگی کلماتی که کمتر از ۵ زیر کلمه دارند با بردارهای ویژگی صفر به طول حداکثر نگاشت می‌شوند.

نتایج به دست آمده از روش بالا با استفاده از شبکه عصبی چند لایه $3/90\%$ گزارش شده است این در حالی است که اگر از شکل کلی کلمات بدون جداسازی برای بازشناسی استفاده شود این مقدار به $5/79\%$ کاهش می‌یابد. [۱۵]

مظفری تحقیقی برای بازشناسی کلمات دست‌نوشته فارسی با استفاده از شکل کلی آن انجام داده است که در این تحقیق از تعداد و موقعیت نقاط برای کاهش فرهنگ لغت به معنی کاهش تعداد کلمات داده‌های آموزش که در بازشناسی کلمه ورودی به کار می‌روند، استفاده شده است. [۱۶] در زبان فارسی و عربی تعدادی حروف وجود دارند که تفاوت آنها تنها در موقعیت و تعداد نقاط است و بدنه آنها کاملاً یکسان

است. نقاط در دست‌نوشته با تغییرات زیادی نگارش می‌شوند نحوه‌های مختلف نگارش نقاط در شکل ۲-۶ نشان داده شده است.

| | Printed | Handwritten |
|------------|---------|-------------|
| Single dot | · | · |
| Double dot | ∴ | ∴ |
| Triple dot | ⋯ | ⋯ |

شکل ۲-۶: تغییرات نگارش نقاط در متون دست‌نوشته [۱۶]

به منظور جدا کردن نقاط از کلمه ابتدا خط کرسی کلمه تخمین زده می‌شود. که این مهم با استفاده از نقاط مینیم محلی کانتور کلمه و رگرسیون خطی بر روی این نقاط تقریب زده می‌شود. نقاط، کوچکترین اجزای در کلمات فارسی و عربی محسوب می‌شوند. لذا در مرحله اول جداسازی نقاط، اجزای کوچکتر از میانگین اجزای کلی کلمه به عنوان کاندیدهای اولیه نقاط در نظر گرفته می‌شوند. از آنجایی که بعضی کاراکترها مانند (ر، د، و) در دست‌نوشته مانند ۳ نقطه به هم پیوسته نگارش می‌شوند و از لحاظ معیار مساحت معمولاً در شرایط نقاط قرار می‌گیرند برای حذف کردن آنها از نقاط دو معیار در نظر گرفته شده است.

(۱) کاراکترهای ذکر شده بر روی خط کرسی و یا نزدیک به آن نگارش می‌شوند این در حالی است

که ۳ نقطه با مقداری فاصله در بالا و یا پایین خط کرسی نگارش می‌شوند.

(۲) میزان هم پوشانی ستونی، ۳ نقطه با اجزای دیگر کلمه بیشتر از میزان هم پوشانی کاراکترهای

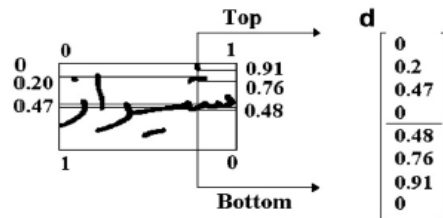
مجزاست.

پس از استخراج نقاط به منظور توصیف کلمه، هر مجموعه نقاط با یک عدد که بیان کننده تعداد نقاط و

یک حرف که بیان کننده موقعیت آنها نسبت به خط کرسی است، بیان می‌شود.

با استفاده از روش تطبیق رشته دمرا- لون اشتین^۱ کد توصیفگر کلمه ورودی با تمام کدهای توصیفگر کلمات موجود در فرهنگ لغت مقایسه می‌شود و k تا از بهترین کلمات کاندید انتخاب می‌شود. مقدار k وابسته به نرخ دقت بازشناسی و درجه کاهش فرهنگ لغت است.

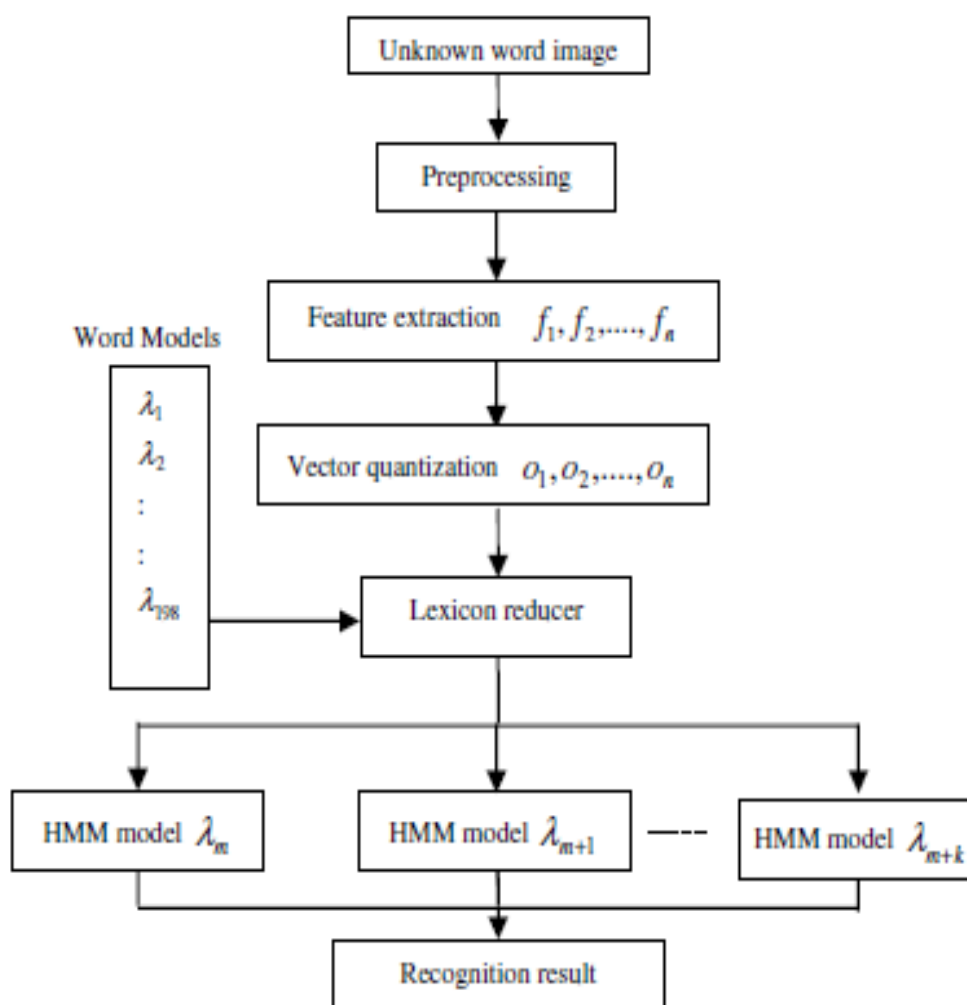
برای هر کلمه موجود در فرهنگ لغت یک مدل HMM گسسته با استفاده از ویژگی انتقال ساخته می‌شود. ویژگی انتقال همانگونه که در شکل ۲-۷ نشان داده شده است، مکان و تعداد انتقال‌ها از پیکسل سیاه به سفید را در هر ستون، از بالا به پایین و از پایین به بالا، محاسبه می‌کند.



شکل ۲-۷: نحوه استخراج ویژگی انتقال [۱۶]

در مرحله بازشناسی کلمه تست ورودی تنها مدل HMM کلماتی شرکت می‌کند که در مرحله کاهش فرهنگ لغت انتخاب شده‌اند. روند کلی این تحقیق در شکل ۲-۸ نشان داده شده است.

¹ Demerau levenshtein

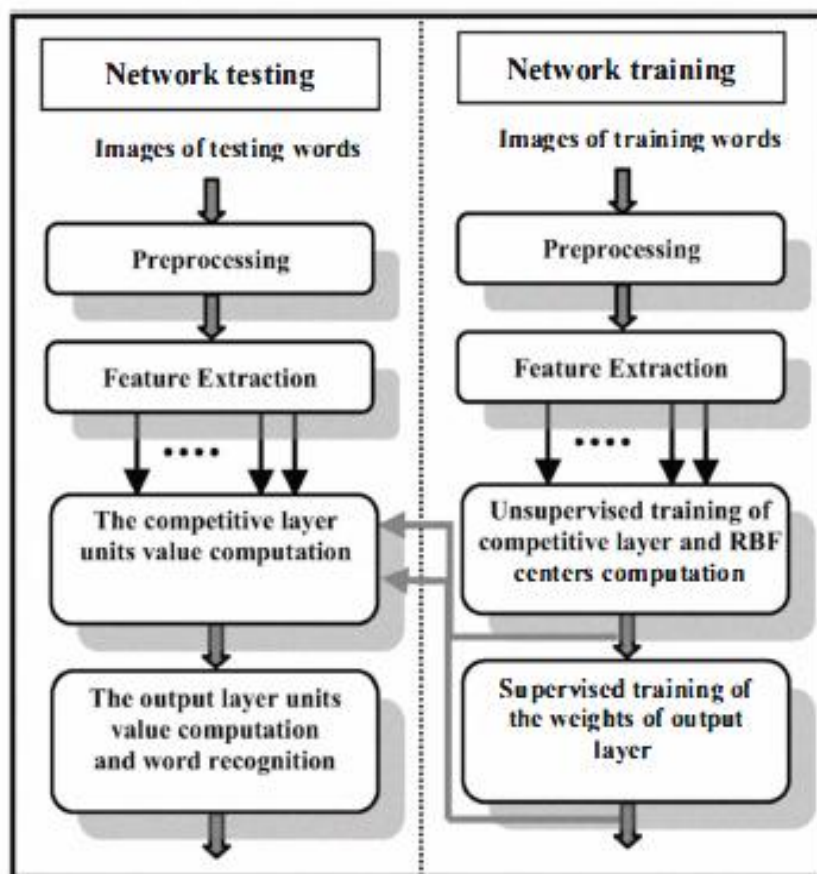


شکل ۲-۸: مراحل بازشناسی روش ارائه شده [۱۶]

این الگوریتم بر روی یک پایگاه داده شامل ۱۷۰۰۰ تصویر از اسامی ۲۰۰ شهر ایران آزمایش شده است. نرخ بازشناسی در انتخاب اول بدون کاهش فرهنگ لغت ۶۰/۲٪ و با کاهش فرهنگ لغت ۷۳/۶۱٪ گزارش شده است.

بهمنی در تحقیقی از ضرایب موجک استخراج شده از پروفایل تصویر هموار شده در چهار جهت استاندارد به عنوان ویژگی استفاده کرده است [۱۷]. دلیل انتخاب این ویژگی را ساختار ویژه زبان فارسی بیان کرده است. در زبان فارسی، شکل حروف بسته به موقعیتش در کلمه پروفایل‌های مختلفی تولید می‌کند. بنابراین ویژگی پروفایل تقریباً تمام ویژگی یک کلمه در زبان فارسی را بیان می‌کند. در مرحله بازشناسی

از شبکه عصبی RBF ^۱ که برای معماری بهینه آن الگوریتم ژنتیک بکار رفته، استفاده شده است. بلوک
 دیگرام الگوریتم پیشنهاد شده در این تحقیق در شکل ۲-۹ آمده است. این روش بر روی اسامی
 دست‌نویس ۳۰ نام رایج فارسی آزمایش شده است. برای هر نام ۱۱۰ نمونه وجود دارد. نرخ بازشناسی در
 اولین انتخاب ۹۶٪ و در انتخاب دوم ۹۷٪ گزارش شده است. [۱۷]



شکل ۲-۹: الگوریتم ارائه شده در تحقیق [۱۷]

در مرجع [۱۸]، ابراهیمی یک روش دو مرحله‌ای برای بازشناسی زیر-کلمات چاپی فارسی ارائه کرده است. زیر-کلمات چاپی با استفاده از ویژگی مکان مشخصه و روش k -میانگین، به ۳۰۰ خوشه تقسیم بندی می‌شوند. از میانگین ویژگی‌های هر خوشه به عنوان نماینده آن خوشه استفاده می‌شود. برای یک زیر-

¹ Radial basis Function

کلمه در مرحله اول با استفاده از ویژگی مکان مشخصه و فاصله اقلیدوسی از میانگین خوشه‌ها، طبقه‌بندی اولیه به ۳۰۰ خوشه انجام می‌شود و ۱۰ خوشه نزدیکتر انتخاب می‌شوند. از آنجا که ویژگی مکان مشخصه که برای خوشه‌بندی استفاده می‌شود شکل کلی زیر کلمات را توصیف می‌کند پس از خوشه‌بندی زیر کلمات، نمونه‌های داخل هر خوشه از نظر شکل کلی شبیه به هم هستند، بنابراین در مرحله دوم باید از ویژگی که جزئیات شکل این زیر-کلمات را توصیف می‌کند، استفاده شود. از این رو برای طبقه‌بندی در داخل خوشه‌ها از توصیفگرهای فوریه منحنی پیرامون زیر-کلمات استفاده شده است.

مجموعه داده‌های آموزش شامل زیر-کلمات متداول فارسی برای چهار قلم لوتوس، میترا، زر و یاقوت در سه اندازه ۱۰، ۱۲ و ۱۴ است. برای ارزیابی طبقه‌بندی از مجموعه‌ای شامل ۵۰۰ زیر-کلمه استفاده شد. با احتساب اولین انتخاب، پنج انتخاب اول و ده انتخاب اول به ترتیب ۷۱/۴٪، ۹۵٪ و ۹۸/۲٪ از این زیر-کلمات به درستی طبقه‌بندی شدند. در مرحله پس پردازش از نوع و ترتیب زیر-کلمات برای بهبود بازشناسی آنها استفاده شده است. که برای بازشناسی مجموعه ۵۰۰ زیر-کلمه، در انتخاب اول به دقت بازشناسی ۹۲،۶۰٪ دست یافت.

روشی دو مرحله‌ای برای بازشناسی حروف مجزای فارسی توسط *Alai* پیشنهاد شده است [۱۹]. در مرحله اول کاراکترهای مشابه خوشه‌بندی می‌شوند، ۳۲ کاراکتر به ۸ گروه خوشه‌بندی شده‌اند. در این مرحله استخراج ویژگی بر اساس میانگین بلوکی صورت می‌گیرد. بدین منظور ابتدا در مرحله پیش پردازش، تصویر ورودی باینری می‌شود سپس برای بهبود نتیجه و کاهش حساسیت به اندازه و مکان، تصاویر به اندازه ۴۹×۴۹ نرمال می‌شوند. پس از آن تصویر به ۴۹ بلوک ۷×۷ بدون همپوشانی و اندازه ثابت تقسیم‌بندی شده و در هر بلوک تعداد پیکسل‌های پیش زمینه محاسبه می‌شود. در مرحله دوم گروه‌هایی که شامل بیش از یک کاراکتراند برای بازشناسی نهایی انتخاب می‌گردند. ویژگی به کار رفته در این مرحله *chain-code* است که معرف خوبی از شکل حروف و ضخامت آنهاست. برای استخراج این ویژگی ابتدا

کانتور بدنه تصویر استخراج می‌گردد و بعد از بلوک‌بندی با همپوشانی تصویر، این بار ۴۹ بلوک 9×9 ، *chain-code* در ۸ جهت محاسبه می‌شود. این تحقیق به جای بیان ویژگی در ۸ جهت، همانگونه که در شکل ۲-۱۰ نشان داده است، ویژگی‌ای ساده بر اساس ۴ جهت اصلی عمودی، افقی، قطری و غیر قطری پیشنهاد می‌کند.



شکل ۲-۱۰: نقطه نمونه p و کدهای جهتی مربوط به آن [۱۹]

به منظور طبقه‌بندی نیز از ماشین بردار پشتیبان چند کلاسه استفاده شده است. روش پیشنهادی بر روی پایگاه داده *IFHCDB* که شامل ۳۶۶۸۲ کاراکتر آموزش و ۱۵۳۳۸ کاراکتر تست می‌باشد به دقت ۹۶/۶۸٪ دست یافته است [۱۹].

فصل ۳

مباحث تئوری

۳-۱- مقدمه

در این فصل به بررسی مراحل مختلف سیستم بازشناسی می‌پردازیم.

یک سیستم بازشناسی شامل مراحل زیر است:

- پیش پردازش: پردازش‌های لازم برای بهبود اولیه کیفیت تصویر، مانند رفع کجی و حذف نویز در این مرحله انجام می‌شود.
- استخراج ویژگی: برای معرفی هر یک از اجزای تصویر به شکل قابل فهم برای سیستم بازشناسی، مشخصات هر یک از اجزای تصویر در این مرحله استخراج می‌شوند.
- طبقه‌بندی: به کمک ویژگی‌های استخراج شده از مرحله قبل، کلمه در این مرحله بازشناسی می‌شود.

در قسمت آخر این فصل نیز به طور مختصر مباحث مربوط به خوشه‌بندی داده‌ها که در زمینه کاهش فرهنگ لغت به منظور بازشناسی کلمات استفاده می‌شود مرور می‌کنیم.

۳-۲- پیش پردازش

هدف از پیش پردازش کاهش نویز و افزایش وضوح تصویر ورودی است. یکسان سازی و یکنواخت کردن تصاویر متنی که در یک سیستم بازشناسی کاملاً ضروری است، مرحله پیش پردازش را یکی از مراحل مهم فرایند بازشناسی قرار داده است که به صورت مستقیم بر روی مراحل بعدی که شامل استخراج ویژگی و طبقه‌بندی می‌شود، اثر می‌گذارد. در نتیجه پیش پردازش، مرحله‌ای برای ارتقاء کیفیت استخراج ویژگی از طریق بالا بردن کیفیت تصویر است که می‌تواند شامل مراحل از قبیل هموار سازی و حذف نویز، یافتن خط کرسی و انحراف آن، باینری کردن، جداسازی متون به کلمه و یا کلمه به حروف باشد. که در زیر به طور مختصر توضیح داده می‌شوند.

۳-۲-۱- حذف شکستگی‌های بدنه کلمات

وجود شکستگی در بدنه کاراکترها موجب ایجاد اشتباه در مرحله طبقه‌بندی می‌شود این به علت استخراج ویژگی نامناسب می‌باشد. همچنین اگر سیستم بازشناسی دارای مرحله مجزاسازی باشد وجود شکستگی باعث ایجاد خطا در این مرحله می‌شود. منظور از شکستگی در بدنه کلمات همانطور که در شکل ۳-۱ نشان داده شده است اعوجاج و پارگی‌هایست که باعث تبدیل یک جز یکپارچه به دو جز جدا می‌شود.



شکل ۳-۱: شکستگی بدنه کلمات

یکی از مراحل اصلی برای حذف شکستگی، پیدا کردن مکان شکستگی در کاراکترهاست. در این تحقیق این شکستگی‌ها بر اساس فاصله بین نقاط پایانی اجزای تصویر تعریف می‌شوند به این ترتیب که اگر فاصله بین نقاط پایانی از دو جز مختلف از حد آستانه مشخص کمتر باشد آن دو جز به عنوان اجزای یک جز یکپارچه که بر اثر شکستگی از هم جدا شده‌اند شناسایی می‌شوند. [۲۰]

پس از شناسایی این اجزا برای متصل کردن آنها به یکدیگر می‌توان از سه روش زیر استفاده کرد.

۱. یکی از اجزای مربوط به شکستگی به اندازه فاصله بین دو جز حرکت داده شده تا به هم متصل شوند. میزان این جابجایی به اندازه زیر است.

$$\Delta x = x_{cpR} - x_{cpL} \quad (۱-۳)$$

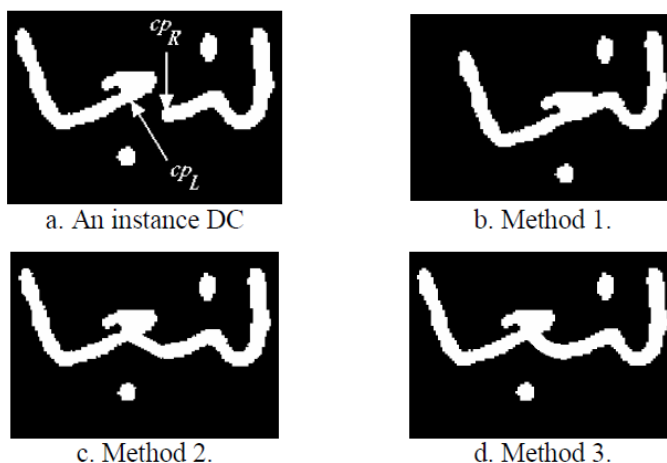
$$\Delta y = y_{cpR} - y_{cpL} \quad (۲-۳)$$

در معادلات بالا cp_R معرف پیکسل پایانی جزء سمت راست و cp_L معرف پیکسل پایانی جزء سمت چپ مربوط به شکستگی است. Δx و Δy نیز به ترتیب میزان جابجایی در راستای افقی و عمودی است.

۲. دو جزء مربوط به شکستگی با یک خط مستقیم به یکدیگر متصل می‌شوند.

۳. دو جزء با یک منحنی به یکدیگر متصل می‌شوند. که این منحنی می‌تواند با استفاده از روش‌های درون یابی غیر خطی^۱، گوسی^۲ و یا چند جمله‌ای^۳ رسم گردد.

نمونه‌ای از شکستگی که با سه روش بالا حذف گردیده در شکل ۳-۲ نشان داده شده است.



شکل ۳-۲: ترمیم شکستگی کاراکتر با روش‌های مختلف [۲۰]

در روش چند جمله‌ای منحنی به وسیله نقاطی که از معادله زیر به دست می‌آیند رسم می‌گردد.

$$F(x) = \sum_{i=0}^n p_i x^i \quad (3-3)$$

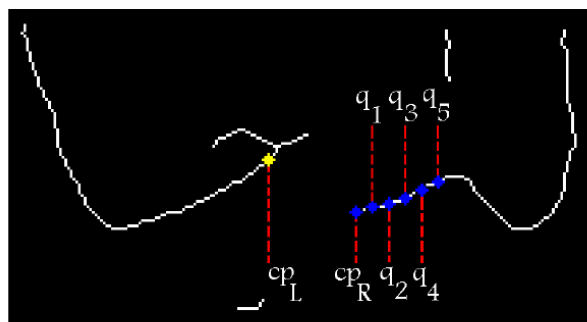
که $(x, f(x))$ مختصات مربوط به مکان هر پیکسل در منحنی می‌باشند و p_i ضریب نام چند جمله‌ای است. برای به دست آوردن منحنی دقیقتر و هموارتر باید نمونه‌های زیادی برای به دست آوردن ضرایب وجود

¹ Nonlinear

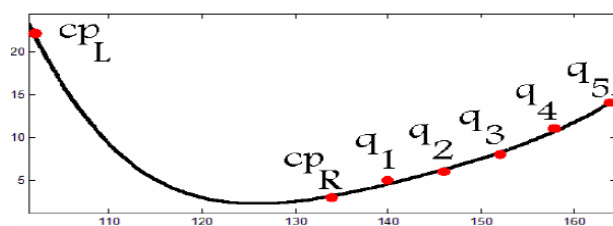
² Gaussian

³ Ploynomial

داشته باشد این در حالی است که ما تنها دو نقطه ابتدایی و انتهایی شکستگی را در اختیار داریم. این نقاط اضافی را از اسکلت بدنه قسمت سمت راست شکستگی به دست می‌آوریم. به این ترتیب که پس از عبور از هر ۶ پیکسل بدنه کاراکتر، مشخصات پیکسل بعدی به مجموعه نقاط اضافه می‌شود. این فرایند را ۵ بار تکرار می‌کنیم حال ضرایب چند جمله‌ای باید به نحوی به دست آید که از این ۷ نقطه (با توجه به دو نقطه ابتدایی و پایانی) عبور کند. علت انتخاب سمت راست شکستگی برای به دست آوردن نقاط بیشتر نیز ماهیت نگارش از سمت راست به چپ زبان فارسی است. روند انتخاب نقاط نمونه و رسم منحنی در شکل ۳-۳ نشان داده شده است.



الف

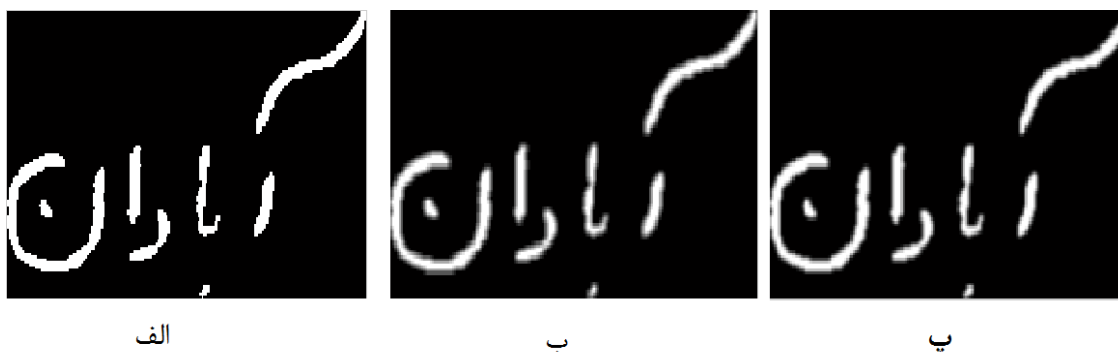


ب

شکل ۳-۳: رسم منحنی به منظور حذف شکستگی، (الف) تصویر شکستگی، (ب) منحنی رسم شده برای ترمیم شکستگی [۲۰]

۳-۲-۲- هموارسازی و حذف نویز

این مرحله از پیش پردازش به منظور حذف نویز و شکستگی‌های کوچک در بدنه تصویر کلمه انجام می‌گیرد. کاهش نویز می‌تواند از طریق هموارسازی تصویر، به کمک فیلتر خطی و غیر خطی انجام شود.



شکل ۳-۵: (الف) تصویر اصلی، (ب) اعمال فیلتر استاندارد، (پ) اعمال فیلتر موزون

فیلترهای مرتبه آماری^۱، فیلترهای مکانی غیر خطی اند که پاسخ آنها مبتنی بر رتبه بندی پیکسل‌های موجود در ناحیه تصویر است که توسط فیلتر در بر گرفته شده است سپس مقدار پیکسل‌های مرکزی با مقدار تعیین شده توسط نتیجه رتبه بندی، جایگزین می‌شود. معروف ترین فیلتر در این دسته، فیلتر میانه^۲ است که مقدار پیکسل را با میانه مقادیر شدت در همسایگی آن پیکسل جایگزین می‌کند. شکل ۳-۶ نتیجه اعمال فیلتر میانه بر تصویر نویزی شده را نشان می‌دهد.



شکل ۳-۶: (الف) تصویر نویزی شده، (ب) نتیجه اعمال فیلتر میانه

¹ Order Statistic Filters

² Median Filter

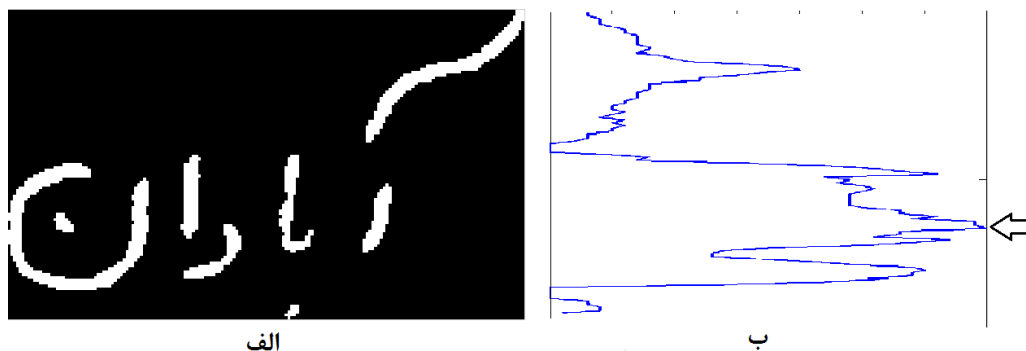
۳-۲-۳- تخمین مکان خط کرسی

خط کرسی به خط مجازی که کلمه بر روی آن نگارش می‌شود اطلاق می‌گردد. متون فارسی در امتداد خط کرسی که ناحیه‌ای با اکثریت پیکسل‌های پیش زمینه است نگارش می‌شود. بسیاری از روش‌های بازشناسی از یافتن خط کرسی کلمه به عنوان قسمتی از فرآیند بازشناسی استفاده می‌کنند که می‌تواند یکی از مراحل ضروری قبل از استخراج ویژگی باشد. یافتن خط کرسی می‌تواند برای اصلاح چرخش، جدا کردن نقاط و علائم و ... به کاربرده شود که تشخیص صحیح آن می‌تواند تفاوت زیادی در کارایی سیستم بازشناسی داشته باشد. در زیر چند روش که برای یافتن موقعیت خط کرسی به کار می‌رود شرح داده می‌شود:

۳-۲-۳-۱- هیستوگرام افکنش^۱

این روش بر این اساس که بیشترین میزان پیکسل‌های کلمه باید بر روی خط کرسی قرار داشته باشند عمل می‌کند. به این ترتیب که نمودار فراوانی پیکسل‌های پیش زمینه به صورت افقی محاسبه می‌شود (در هر سطر تعداد پیکسل‌های پیش زمینه بدست می‌آید) و در نهایت موقعیت خط کرسی با یافتن بیشترین مقدار این هیستوگرام بدست می‌آید. نحو یافتن خط کرسی با این روش در شکل ۳-۷ ترسیم شده است.

¹ Projection Histogram



شکل ۳-۷: نحو تخمین مکان خط کرسی، (الف) تصویر کلمه، (ب) هیستوگرام افکنش کلمه

هیستوگرام افکنش برای کلمه‌هایی با طول زیاد و یا مجموعه کلمات کنار هم (جمله) تخمین خوبی از موقعیت خط کرسی می‌دهد در حالی که تخمین این روش برای کلمات مجزا با طول کم معمولاً نادرست است اما به علت پیچیدگی محاسباتی کم و سرعت بالا، برای تخمین خط کرسی متون کاربرد فراوانی دارد. *Bagdanov* و *Kanai* از این روش برای اصلاح کجی تصویر متنی استفاده کرده‌اند. [۲۲]

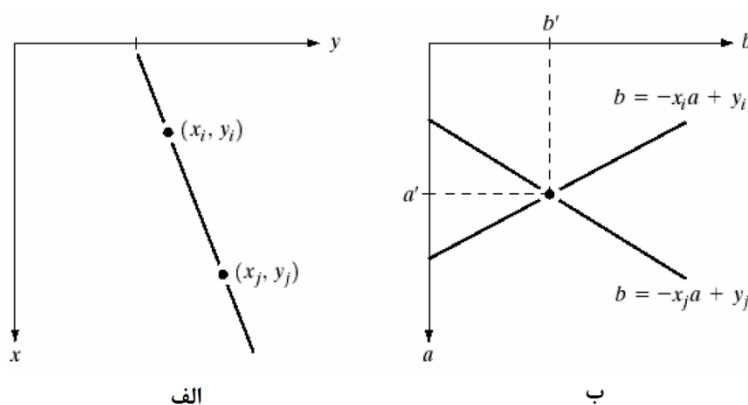
۳-۲-۳-۲- تبدیل هاف^۱

هر نقطه در فضای هاف متناظر با یک خط با زاویه و فاصله مشخص از مبدا در فضای کارترین می‌باشد مقادیر نقاط در فضای هاف بیان کننده این است که چه تعداد از نقاط داده‌ها بر روی یک خط قرار می‌گیرند. لذا این تبدیل می‌تواند برای یافتن خطوط مستقیم بر روی مجموعه‌ای از داده‌ها به کار رود. نقاط در فضای هاف با مقدار عددی بالا به این معنی است که بسیاری از نقاط داده‌ها بر روی خط مورد نظر در فضای کارترین قرار دارند. بنابراین از این تبدیل می‌توان برای یافتن خط کرسی که در آن تعداد زیادی از پیکسل‌های پیش زمینه بر روی خطی مستقیم قرار دارند استفاده کرد. [۲۳]

همانگونه که در شکل ۳-۸ نشان داده شده است، در تبدیل هاف، نقطه (x_i, y_i) و همه خطوطی که از آن عبور می‌کنند مطلوب ما می‌باشند. بی‌نهایت خط از نقطه (x_i, y_i) عبور می‌کنند که هر یک از آنها بر روی

^۱ Hough Transform

خط $y_i = ax_i + b$ قرار دارند. با نوشتن این تساوی به صورت $b = -x_i a + y_i$ و در نظر گرفتن صفحه ab (تحت عنوان فضای پارامتری)، این تساوی برای زوج ثابت (x_i, y_i) به صورت یک خط در (a', b') صفحه ab را قطع می‌کند که محل تقاطع خط شامل نقاط (x_i, y_i) و (x_j, y_j) در صفحه ab می‌باشد. در واقع تمامی نقاط موجود بر این خط دارای خطوطی در فضای پارامتری می‌باشند که همگی از (a', b') می‌گذرند. [۲۱]



شکل ۳-۸: الف) صفحه xy (ب) فضای پارامتری [۲۱]

۳-۲-۳- دیدگاه نزدیکترین همسایگی^۱

در این روش ابتدا اجزا متصل تصویر کلمه بدست می‌آیند سپس زاویه خطی که هر مرکز جز را به مراکز نزدیکترین جز همسایگی‌اش متصل می‌کند محاسبه می‌گردد و با محاسبه میانگین این زوایا، زاویه خط کرسی بدست می‌آید. که می‌تواند مستقیماً برای اصلاح کجی تصویر متنی نیز به کار برده شود. [۲۴]

۳-۲-۴- باینری کردن

هدف از این مرحله تبدیل تصویر سطح خاکستری^۱ ورودی به شکل دو سطحی^۲ می‌باشد. تصویر خاکستری، تصویری است که نقاط آن مقادیر بین سفید و سیاه را دارا هستند و تصویر دو سطحی،

^۱ Nearest Neighbour

تصویری است که نقاط آن تنها سیاه و سفید هستند. بسیاری از روش‌های بازشناسی از شکل دو سطحی تصاویر ورودی استفاده می‌کنند. برای دو سطحی کردن تصویر ورودی می‌توان از یک سطح آستانه ثابت یا یک سطح آستانه وقتی استفاده نمود. روش‌های مورد نظر نباید باعث ایجاد نویز در مرز و داخل حروف، چسبیدگی، خوردگی یا جداسازی آنها شود. برای تصاویر متنی که کاراکترها دارای سطح خاکستری ثابتی هستند که آنها را از پس‌زمینه مجزا می‌کند، روش آستانه کلی یک روش کاراست. نمودار فراوانی^۳ این تصاویر دارای دو میانگین است. لذا دسته‌بندی هر پیکسل تنها به سطح خاکستری آن بستگی دارد و به ویژگی محلی وابسته نیست از میان روش‌های آستانه گذاری کلی روش *otsu* [۲۵] یک روش سریع و خوب برای بدست آوردن آستانه کلی است.

در روش‌های آستانه گذاری کلی بر اساس آستانه تعیین شده، پیکسل‌های تصویر به دو کلاس پیش زمینه و پس زمینه تقسیم بندی می‌شوند. روش *otsu* مسئله پیدا کردن آستانه مناسب را با مینیمم کردن خطای طبقه‌بندی پیکسل‌ها پیش زمینه به عنوان پس زمینه و بر عکس، حل می‌کند.

برای یک تصویر با محدوده خاکستری $G = \{0, 1, 2, \dots, L-1\}$ محدوده پیکسل‌های پیش زمینه و پس زمینه می‌توانند با دو کلاس $C_0 = \{0, 1, 2, \dots, t\}$ و $C_1 = \{t+1, t+2, \dots, L-1\}$ نمایش داده شوند. به عنوان یک معیار ساده و مناسب برای مجزاسازی کلاس‌ها می‌توان از ماتریس پراکندگی میان کلاس‌ها^۴ و ماتریس پراکندگی داخلی کلاس^۵ استفاده کرد. به این ترتیب که آستانه بهینه t^* به وسیله ماکزیمم کردن یکی از توابع زیر به دست می‌آید.

¹ *Grayscale*

² *Binary*

³ *Histogram*

⁴ *Between Class Scatter Matrices*

⁵ *Whithen Class Scatter Matrices*

$$\lambda = \frac{\sigma_B^2}{\sigma_W^2}, \quad \eta = \frac{\sigma_B^2}{\sigma_T^2}, \quad \kappa = \frac{\sigma_T^2}{\sigma_W^2}. \quad (4-3)$$

که در معادلات بالا σ_W^2 نشان دهنده واریانس داخلی کلاس، σ_B^2 نشان دهنده واریانس بین کلاس‌ها و σ_T^2 معرف واریانس کلی است. به عنوان مثال در اینجا مقدار η را برای به دست آوردن t^* بهینه ماکزیمم می‌کنیم.

$$P_i = n_i/n, \quad w_0 = \sum_{i=0}^t P_i, \quad w_1 = w - w_0, \quad (5-3)$$

$$\mu_T = \sum_{i=0}^{L-1} iP_i, \quad \mu_t = \sum_{i=0}^t iP_i, \quad \mu_0 = \frac{\mu_t}{w_0}, \quad \mu_1 = \frac{\mu_T - \mu_t}{1 - \mu_0}, \quad (6-3)$$

$$\sigma_T^2 = \sum_{i=0}^{L-1} (i - \mu_T)^2 P_i, \quad \sigma_B^2 = w_0 w_1 (\mu_1 - \mu_0)^2. \quad (7-3)$$

در معادلات بالا n_i نشان دهنده i امین مقدار هیستوگرام فراوانی تصویر است که این مقدار برابر با تعداد پیکسل‌ها با سطح خاکستری i در تصویر خواهد بود. $n = \sum_{i=0}^{L-1} n_i$ برابر تعداد کل پیکسل‌های تصویر و $p_i = n_i/n$ احتمال وقوع سطح خاکستری i ام است. برای انتخاب آستانه t^* تصویر ورودی، احتمال کلاس‌ها W_0 و W_1 نشان دهنده مساحت اشغال شده توسط پیکسل‌های شی و پس زمینه است. مقدار η در محدوده $[0,1]$ قرار دارد که محدوده پایین آن هنگامی به دست می‌آید که تصویر از لحاظ سطح خاکستری غیر یکنواخت باشد و محدوده بالا نشان دهنده تصویر، تنها با دو سطح خاکستری است. [۲۵]

آستانه گذاری محلی

اسناد متنی واقعی معمولا دارای پس زمینه پیچیده، رنگی و در نتیجه زمینه مشکل برای جدا کردن کاراکترها از پس زمینه‌اند. این در حالی است که روش‌های آستانه گذاری کلی می‌توانند اشیا را از یک پس زمینه ساده و یکنواخت با سرعت بالا جدا کنند. اما روش‌های آستانه گذاری محلی به قیمت زمان پردازش بیشتر می‌تواند تغییرات پس زمینه را حذف کند.

در روش محلی *Bensen's* ماکزیمم و مینیمم سطح خاکستری اطراف هر پیکسل به دست می‌آید و از مقدار میانگین این دو مقدار به عنوان آستانه پیکسل استفاده می‌کند. [۲۶]

$$g(x, y) = (F_{\max}(x, y) + F_{\min}(x, y))/2, \quad (۸-۳)$$

$$b(x, y) = \begin{cases} 1 & \text{if } f(x, y) < g(x, y), \\ 0 & \text{otherwise.} \end{cases} \quad (۹-۳)$$

برای حذف علائم با شدت سطح خاکستری نزدیک به پیکسل‌های شی در پس زمینه وایانس محلی را می‌توان به صورت زیر محاسبه کرد.

$$c(x, y) = F_{\max}(x, y) - F_{\min}(x, y). \quad (۱۰-۳)$$

طبقه‌بندی پیکسل‌های پیش زمینه $f(x, y)$ می‌تواند با مقایسه واریانس محلی نسبت به آستانه بازبینی شود. هر پیکسل بر طبق شرایط زیر دسته بندی می‌شود.

$$b(x, y) = \begin{cases} 1 & \text{if } (f(x, y) < g(x, y) \text{ and } c(x, y) > c^*) \text{ or} \\ & (f(x, y) < f^* \text{ and } c(x, y) \leq c^*), \\ 0 & \text{otherwise.} \end{cases} \quad (۱۱-۳)$$

که c^* و f^* در معادلات بالا توسط روش *otsu* بر روی هیستوگرام $c(x, y)$ و $g(x, y)$ به دست می‌آیند. این روش آستانه گذاری محلی برای تصاویر متنی با پس زمینه غیر یکنواخت با تغییرات آهسته، مناسب است. [۲۶]

۳-۲-۵- نازک سازی^۱

اسکلت الگو، شکل کلی الگو را نشان می‌دهد که می‌توان ویژگی‌های مهمی مانند نقاط شکست، نقاط برخورد و ... را از آن استخراج کرد. نازک سازی به معنی کاهش پیکسل‌های الگو بدون تاثیرگذاری بر روی شکل کلی آن است به طوری که پس از حذف این پیکسل‌ها الگو هم چنان باید قابل بازشناسی باشند. از این رو اسکلت الگو باید ویژگی‌های زیر را دارا باشد:

- به اندازه ممکن (یک پیکسل) نازک باشد.
- یکپارچه و متصل به هم باشد (نازک سازی باعث شکست در بدنه الگو نمی‌شود)
- در مرکزیت بدنه الگو قرار داشته باشد.
- ایجاد شاخک اضافی نکند.

روش‌های موجود بر اساس شرایط بالا عمل کرده و هنگامی که ویژگی‌های بالا برآورده شد، الگوریتم خاتمه می‌یابد.

الگوریتم‌های نازک سازی

روش‌های زیادی برای استخراج اسکلت الگوهای دیجیتال طراحی شده‌اند. در اینجا ما به توضیح دو روش که در این زمینه استفاده می‌شوند می‌پردازیم. ابتدا روش *Hilditch's* [۲۷] و بعد روش *Zhang-Suen* [۲۸]. این روش‌ها بسیار ساده‌اند. برای توضیح این روش‌ها ما پیکسل‌های پس زمینه را سفید (صفر) و پیکسل‌های پیش زمینه را سیاه (یک) در نظر می‌گیریم.

¹ *Thinning*

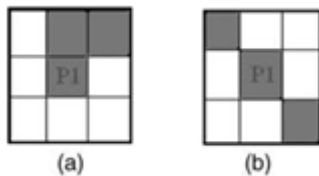
همانطور که در شکل ۳-۹ نشان داده شده ۸ همسایگی پیکسل P_1 در نظر گرفته می‌شوند تا تعیین شود که پیکسل P_1 باید حذف شود و یا اینکه به عنوان اسکلت الگو نگه داشته شود. هشت همسایگی P_1 را به صورت ساعتگرد مرتب کرده و دو تابع $A(P_1)$ و $B(P_1)$ را به صورت زیر تعریف می‌کنیم.

- تعداد تغییر حالت از ۰ به ۱ $A(P_1)$
- تعداد پیکسل‌های سیاه در همسایگی نقطه مورد ($B(P_1) = P_2 + P_3 + P_4 + P_5 + P_6 + P_7 + P_8 + P_9$)
(نظر)

| | | |
|----|----|----|
| P9 | P2 | P3 |
| P8 | P1 | P4 |
| P7 | P6 | P5 |

شکل ۳-۹: پنجره ۳ در ۳ که همسایگی‌های پیکسل P_1 را نشان می‌دهد [۲۷]

همانطور که در بالا ذکر شده در تعریف این معادلات پیکسل‌های سیاه برابر با یک و سفید برابر صفر در نظر گرفته شده است. در شکل ۳-۱۰ این توابع برای دو نمونه محاسبه شده‌اند.



شکل ۳-۱۰: دو نمونه از محاسبه تابع $A(P_1)$ و $B(P_1)$: در (a) $A(P_1)=1, B(P_1)=2$ و در (b) $A(P_1)=2, B(P_1)=2$

در تصویر الگو پیکسل‌هایی که شرایط زیر را دارا باشند برای بدست آوردن اسکلت حذف می‌شوند.

$$2 \leq B(P_1) \leq 6 \quad (۳-۱۲)$$

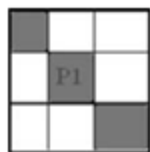
$$A(P_1)=1 \quad (13-3)$$

$$A(P_2) \neq 1 \text{ یا } P_2 \cdot P_4 \cdot P_8 = 0 \quad (14-3)$$

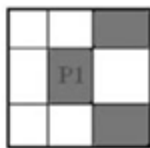
$$A(P_4) \neq 1 \text{ یا } P_2 \cdot P_4 \cdot P_6 = 0 \quad (15-3)$$

روش زمانی خاتمه میابد که دیگر تغییری در تصویر حاصل نشود (پیکسل‌های سیاه دیگری را نتوان حذف کرد) حال شرایطی که در بالا برای حذف پیکسل آمده است توضیح داده می‌شوند.

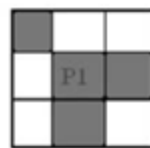
- شرط اول: که خود شامل دو شرط است. اول اینکه تعداد پیکسل‌های سیاه همسایگی P_1 از دو بزرگتر است که نشان دهنده این است که پیکسل مورد نظر یک نقطه پایانی نیست (هر پیکسل با یک پیکسل سیاه در همسایگی‌اش یک نقطه پایانی است)، شرط دوم که تابع در شرایطی کوچکتر و یا مساوی ۶ خواهد بود که پیکسل مورد نظر باید یک پیکسل مرزی باشد.
- شرط دوم: این شرایط برای بازبینی اتصال به کار می‌رود. به عنوان مثال همانطور که شکل ۱۱-۳ نشان می‌دهد در صورتی که $A > I$ اسکلت الگو مجزا شده و یکپارچگی خود را از دست می‌دهد.



$A(P_1)=2$



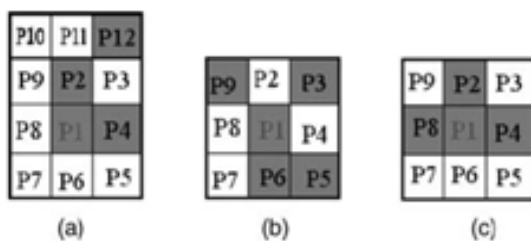
$A(P_1)=2$



$A(P_1)=3$

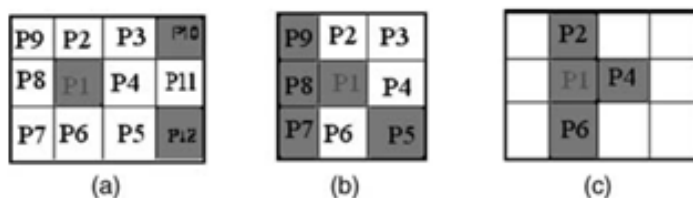
شکل ۱۱-۳: در شرایط نشان داده شده حذف P_1 موجب گسستگی در اسکلت کاراکتر می‌شود [۲۷]

- شرایط سوم: این شرایط همانطور که در شکل ۱۲-۳ نشان داده شده است توسط این روش حذف نخواهد شد.



شکل ۳-۱۲: (a) $A(P_1) \neq 1$ ، در (b) $P_2 \cdot P_4 \cdot P_8 = 0$ ، و (c) $A(P_2) = 1$ و $P_2 \cdot P_4 \cdot P_8 \neq 0$

• شرایط چهارم: این شرط نیز در شکل ۳-۱۳ نشان داده شده است که با این روش حذف نخواهد شد.



شکل ۳-۱۳: (a) $A(P_4) \neq 1$ ، (b) $P_2 \cdot P_4 \cdot P_6 = 0$ ، و (c) $A(P_4) = 1$ و $P_2 \cdot P_4 \cdot P_6 \neq 0$

۳-۲-۵-۲ - روش Zhang-Suen [۲۸]

این روش دو مرحله دارد. در هر مرحله نقاط بدنه ناحیه شناسایی می‌شوند. نقاط بدنه به صورت نقاطی که مقدار یک داشته و حداقل یک همسایگی با ارزش صفر داشته باشند، تعریف می‌شوند.

در مرحله اول از روش اگر شرایط زیر ارضا شود نقطه P به عنوان نقطه بدنه تعریف می‌گردد.

Condition 1: $2 \leq B(P_1) \leq 6$;

Condition 2: $A(P_1) = 1$;

Condition 3: $P_2 \cdot P_4 \cdot P_6 = 0$; (۳-۱۶)

Condition 4: $P_4 \cdot P_6 \cdot P_8 = 0$.

در مرحله دوم نیز شرط اینکه پیکسل جز نقاط بدنه باشد باید شرایط زیر برآورده شود.

Condition 1: $2 \leq B(P_1) \leq 6$ (the same condition as step 1); (۱۷-۳)

Condition 2: $A(P_1) = 1$ (the same condition as step 1);

Condition 3: $P_2 \cdot P_4 \cdot P_8 = 0$;

Condition 4: $P_2 \cdot P_6 \cdot P_8 = 0$.

در این روش مرحله اول بر روی تمامی پیکسل‌های تصویر اعمال می‌شود اگر شرایط این مرحله برآورده شد، نقطه برای حذف علامت گذاری می‌شود. پس از پایان این مرحله نقاط مورد نظر حذف می‌شوند سپس مرحله دوم درست همانند مرحله اول بر روی تصویر اعمال می‌گردد. اعمال مرحله اول و دوم پس از یکدیگر به صورت متناوب تکرار می‌شود تا زمانی که دیگر تغییری در تصویر ایجاد نشود.

۳-۲-۶- نرمال سازی

نرمال سازی اجزا به معنی تغییر اندازه تصاویر ورودی به ابعاد استاندارد، یکی از مهمترین مراحل پیش پردازش در بازشناسی است. هدف از نرمال سازی آسان کردن استخراج ویژگی، کاهش تغییرات شکل اجزا و در نتیجه افزایش دقت طبقه‌بندی است. در اینجا به معرفی چند روش نرمال سازی خطی و غیر خطی که در زمینه بازشناسی موثرند می‌پردازیم.

اگر H_1 و W_1 را به عنوان طول و عرض تصویر اصلی و H_2 و W_2 را به عنوان طول و عرض تصویر نرمال شده و L را به عنوان اندازه استاندارد برای تصاویر در نظر بگیریم آنگاه نسبت ابعاد کاراکتر در تصویر اصلی و تصویر نرمال شده به صورت زیر تعریف می‌شود.

$$R_1 = \frac{\min(W_1, H_1)}{\max(W_1, H_1)}, \quad (۱۸-۳)$$

$$R_2 = \frac{\min(W_2, H_2)}{\max(W_2, H_2)}, \quad (۱۹-۳)$$

که R_1 و R_2 همیشه در محدوده $[0,1]$ قرار خواهد داشت.

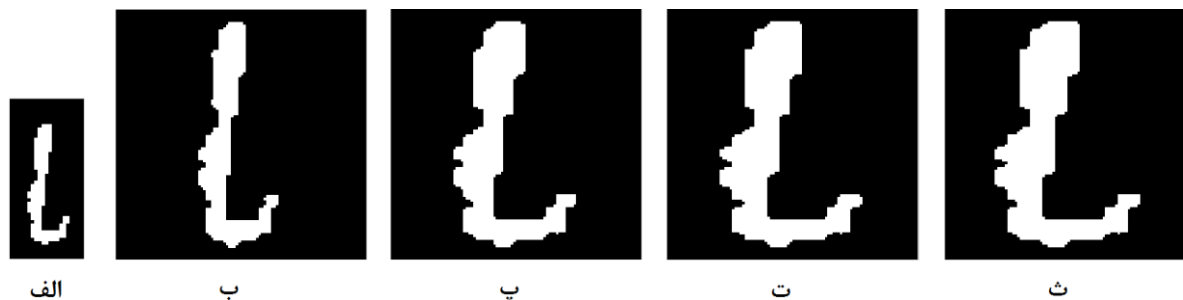
نسبت ابعاد کاراکتر نرمال شده R_2 به صورت تطبیقی بر اساس نسبت ابعاد کاراکتر اصلی R_1 توسط یکی از توابع نگاشت زیر محاسبه می‌شود.

| Method | Function |
|-----------------------------|--|
| Fixed aspect ratio | $R_2 = 1$ |
| Aspect ratio preserved | $R_2 = R_1$ |
| Square root of aspect ratio | $R_2 = \sqrt{R_1}$ |
| Cubic root of aspect ratio | $R_2 = \sqrt[3]{R_1}$ |
| Sine of aspect ratio | $R_2 = \sqrt{\sin(\frac{\pi}{2} R_1)}$ |

جدول ۳-۱: توابع نگاشت نسبت ابعاد [۲۷]

طول و یا عرض بزرگتر در تصویر اصلی به مقدار L نرمال شده و طول و یا عرض کمتر با توجه به تابع نگاشت و نسبت ابعاد R_2 به دست آمده، به مقدار مشخص نرمال می‌شود بعد از آن برای ایجاد تصویر خروجی استاندارد در ابعاد $L \times L$ ، تصویر حاصل از مرحله پیش در مرکز تصویری به طول $L \times L$ و هم سطح با پس زمینه تصویر اصلی قرار می‌گیرد. [۲۷]

شکل ۳-۱۴ کاراکتر 'با' با ابعاد 45×21 از کلمه 'آبادان' است که با توابع نگاشت جدول ۳-۱ به اندازه 100×100 نرمال شده است را نشان می‌دهد.



شکل ۳-۴: نرمال سازی با روش های مختلف، (الف) کاراکتر نرمال نشده، (ب) *Aspect ratio preserved*، (پ) *Square root*،
Sine of aspect ratio (ث)، *Cubic root of aspect ratio* (ت)، *of aspect ratio*

۳-۳- استخراج ویژگی

هدف از استخراج ویژگی بدست آوردن مشخصه های اصلی الگوها است که یکی از دشوارترین و اصلیتترین مراحل سیستم بازشناسی است. در انتخاب بردار ویژگی باید موارد زیر رعایت گردد:

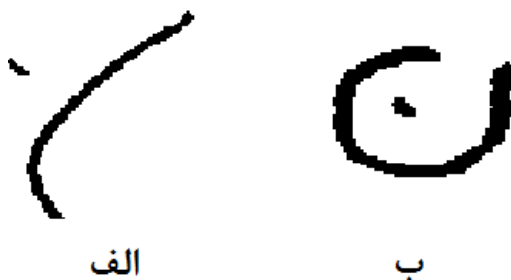
(۱) بردار ویژگی الگو در هر کلاس باید متمایز کننده نمونه های آن کلاس از نمونه های کلاس های دیگر باشد به این معنی که فاصله بین بردارهای ویژگی نمونه ها دو کلاس در فضای ویژگی ها، حداکثر باشد.

(۲) ویژگی های انتخاب شده باید خصوصیات شکل و ساختار الگو را تا حد ممکن از تصاویر الگوهای آموزش استخراج کند.

(۳) نسبت به نویز، تغییر اندازه، چرخش و دیگر تغییرات احتمالی تصاویر پایداری قابل قبولی داشته باشند.

(۴) شرایط و نوع الگوهای ورودی در انتخاب بردار ویژگی تاثیر گذار است به عنوان مثال دست نوشته و یا تایپی بودن متون، میزان حضور نویز و یا زمینه کاربرد سیستم می تواند در انتخاب بردار ویژگی موثر باشد.

۵) اختصاص دادن بردار ویژگی‌های متفاوت به الگوهای مربوط به یک کلاس که تنوع زیادی در نحوی نوشتن دارند می‌تواند در تشخیص الگوهای کلاس موثر باشند. شکل ۳-۱۵ یک کاراکتر نمونه در این زمینه را نشان می‌دهد.



شکل ۳-۱۵: تنوع نگارش در متون دست‌نویس

برای ارزیابی میزان کارایی ویژگی‌های انتخاب شده می‌توان معیارهای مانند حساسیت نسبت به نویز، تغییر اندازه، سرعت و آسانی استخراج آنها را در نظر گرفت. به طور کل ارزیابی روش‌های استخراج ویژگی را می‌توان به معیارهای زیر دسته بندی کرد.

- معیارهای مربوط به شکل:

نویز: میزان حساسیت به فاصله‌ها، ضربه‌ها و غیره.

اعوجاج: میزان حساسیت به تغییرات محلی از قبیل کشیدگی، از بین رفتن گوشه‌ها و غیره.

انتقال: حساسیت نسبت به حرکت افقی و یا عمودی کلی تصویر و یا اجزایش.

چرخش: حساسیت نسبت به تغییر مرکز ثقل تصویر.

- معیارهای کاربردی:

سرعت محاسبه ویژگی

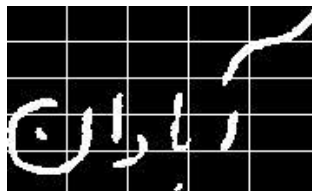
عدم نیاز به روش‌های اضافی

۳-۳-۱- ویژگی‌های آماری^۱

این نوع ویژگی‌ها بر اساس توزیع آماری نقاط تصاویر استخراج می‌گردند و پایداری مناسبی نسبت به تغییر شکل و اعوجاج دارند که چند نمونه از این ویژگی‌ها در زیر توضیح داده می‌شوند.

۳-۳-۱-۱- ناحیه‌ای^۲

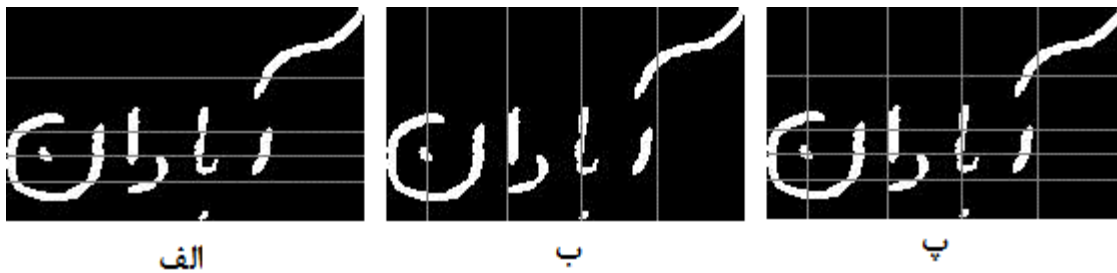
تصویر ورودی به چندین ناحیه دارای همپوشانی یا غیر همپوشانی با مساحت برابر یا متغیر [۲۹] تقسیم می‌شود و از هر ناحیه ویژگی‌های مشخصی استخراج می‌گردد. نحوی تقسیم بندی تصویر به بلوک‌های ثابت در شکل ۳-۱۶ نشان داده شده است. در تقسیم بندی به صورت تطبیقی، اندازه بلوک‌ها در تصویر برابر نمی‌باشد. همانطور که در شکل ۳-۱۷ نشان داده شده است ابتدا بلوک بندی در راستای سطرها انجام می‌شود. بدین نحو که تصویر به تعداد مشخصی نوار افقی تقسیم می‌شود بطوریکه تعداد پیکسل‌های یکسانی در هر نوار قرار گیرد. سپس همین فرایند در راستای ستون‌ها انجام شده و در نهایت از تلاقی نوارهای حاصل از این دو بلوک بندی تقسیم بندی نهایی تصویر صورت می‌گیرد.



شکل ۳-۱۶: بلوک بندی ثابت تصویر به منظور استخراج ویژگی ناحیه‌ای

¹ Statistical

² Zoning



شکل ۳-۱۷: بلوک بندی تطبیقی، بلوک بندی سطری (ردیف اول، سمت چپ) بلوک بندی ستونی (ردیف اول، سمت راست)، بلوک بندی نهایی (سطر دوم)

۱) میانگین بلوکی

سادترین ویژگی ناحیه‌ای محسوب می‌شود که در آن پس از تقسیم بندی تصویر، در هر ناحیه تراکم نقاط پیش زمینه محاسبه می‌شود و به عنوان ویژگی آن ناحیه در کنار ویژگی ناحیه‌های دیگر، بردار ویژگی تصویر کلمه را تولید می‌کنند. [۱۴]

۲) ویژگی‌های جهتی^۱ - گرادیان روشنایی

ویژگی گرادیان، یک ویژگی آماری است که به نتایج خوبی در زمینه بازشناسی ارقام و حروف دست یافته است. این ویژگی از هیستوگرام گرادیان تصویر سطح خاکستری استخراج می‌گردد. از آنجا که معمولا تصویر خروجی مرحله پیش پردازش باینری می‌باشند که دارای لبه‌های شدیدی است و برای معرفی الگوها با این ویژگی مناسب نمی‌باشد در مرحله اول باید تصویر را به سطح خاکستری تبدیل کنیم تا لبه‌های نرم و صافی بدست آیند برای انجام این کار می‌توان از فیلتر میانگیر گیر و یا فیلتر گوسی استفاده کرد. این فیلترها با هموار کردن تصویر باعث توصیف بهتر در تغییرات شدت روشنایی تصویر می‌شوند.

پس از تبدیل تصویر باینری به تصویر خاکستری، باید اطلاعات گرادیان تصویر استخراج گردد. برای انجام این کار می‌توان از فیلترهای سوبل، روبرتز و یا کریش استفاده کنیم. از این میان عملگر کریش به علت

¹ Direction

محاسبه مولفه گرادیان در چهار جهت عمودی، افقی، قطر اصلی و قطر فرعی پیچیدگی بیشتری نسبت به دو عملگر دیگر دارد. عملگر روبرتز و سوبل مولفه گرادیان را در دو جهت عمودی و افقی محاسبه می‌کنند که عملگر سوبل به علت استفاده از همسایگی بزرگتر نسبت به عملگر روبرتز، نتایج بهتری می‌دهد. عملگر سوبل برای محاسبه اطلاعات گرادیان هر پیکسل در دو جهت عمودی و افقی در شکل ۳-۱۸ نشان داده شده است. [۱۰]

| | | |
|----|---|---|
| -1 | 0 | 1 |
| -2 | 0 | 2 |
| -1 | 0 | 1 |

| | | |
|----|----|----|
| 1 | 2 | 1 |
| 0 | 0 | 0 |
| -1 | -2 | -1 |

شکل ۳-۱۸: عملگرهای سوبل

با استفاده از این عملگر، گرادیان $g(x,y)=[g_x, g_y]$ در موقعیت (x,y) به صورت زیر محاسبه می‌شود.

$$g_x(x, y) = f(x+1, y-1) + 2f(x+1, y) + f(x+1, y+1) - f(x-1, y-1) - 2f(x-1, y) - f(x-1, y+1) \quad (۲۰-۳)$$

$$g_y(x, y) = f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1) - f(x-1, y-1) - 2f(x, y-1) - f(x+1, y-1)$$

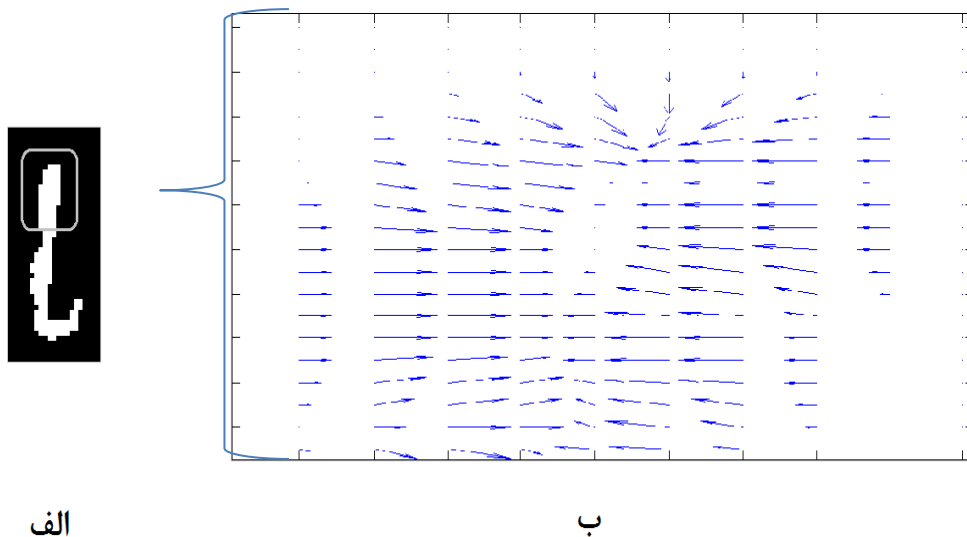
که g_x مولفه گرادیان در جهت افقی و g_y مولفه گرادیان در جهت عمودی است. حال به وسیله این دو مولفه، شدت و جهت گرادیان به صورت زیر محاسبه می‌شوند.

$$s(x, y) = \sqrt{g_x^2 + g_y^2} \quad \text{اندازه} \quad (۲۱-۳)$$

$$\theta(x, y) = \tan^{-1}(g_y / g_x) \quad \text{جهت} \quad (۲۲-۳)$$

برای استخراج ویژگی، جهت‌های گرادیان به ۸، ۱۶ و یا ۳۲ جهت کوانتیزه می‌شوند. پس از به دست آوردن زوایا برای کوانتیزه کردن آن می‌توان جهت گرادیان به نزدیکترین جهت پایه و یا به دو جهت پایه در برگیرنده آن نگاشت شود و در هر جهت اندازه‌های گرادیان با هم جمع می‌شوند. لذا بردار ویژگی

استخراج شده از هر بلوک تصویر برابر تعداد جهت‌های کوانتیزه در نظر گرفته شده، خواهد بود. در شکل ۳-۱۹ تصویر قسمتی از کلمه 'آبادان' و قدرت و جهت گرادیان حاصل از آن را نشان می‌دهد.

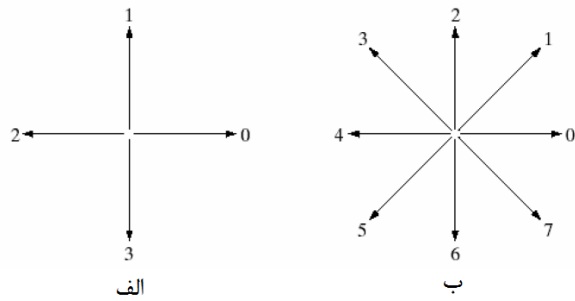


شکل ۳-۱۹: (الف) تصویر ورودی، (ب) جهت‌های گرادیان قسمت مشخص شده کاراکتر

۳) ویژگی‌های جهت‌ی-کدهای زنجیره‌ای^۱ [۱۹]

رمزهای زنجیره‌ای برای نمایش مرز توسط دنباله‌ی متصل از قطعات خط مستقیم با طول و جهت معین به کار می‌رود. به طور معمول جهت‌ها به ۴ یا ۸ قسمت تقسیم می‌شوند. جهت هر بخش با استفاده از طرح شمارگذاری که در شکل ۳-۲۰ نشان داده شده است رمز گذاری می‌شود.

¹ Chain Code



شکل ۳-۲۰: الگوی کد زنجیره‌ای، الف) الگوی کد زنجیره‌ای چهار جهتی، ب) الگوی کد زنجیره‌ای هشت جهتی [۲۱]

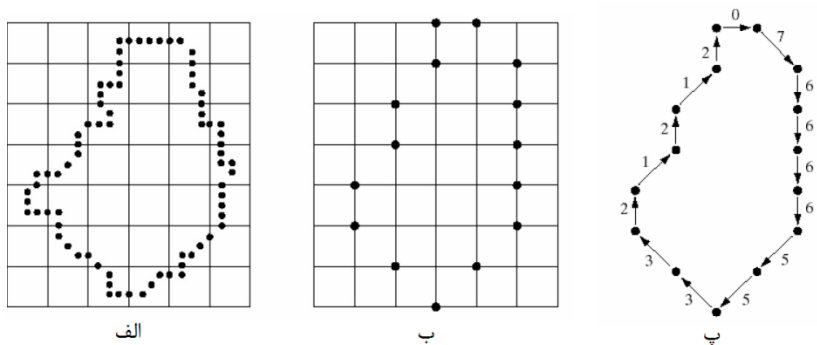
رمز زنجیره‌ای می‌تواند با دنبال کردن با ردیابی مرز، به عنوان مثال در جهت ساعتگرد و تشخیص جهت‌های آنها (وضعیت هر پیکسل نسبت به پیکسل قبلی) طبق الگوهای ذکر شده محاسبه شوند. این روش به دو دلیل اصلی قابل قبول نیست:

(۱) زنجیره حاصل می‌تواند طولانی باشد.

(۲) هر خرابی کوچک در امتداد مرز به خاطر نویز و یا ناحیه بندی ناقص، می‌تواند باعث ایجاد

تغییراتی در رمز شود که ممکن است مربوط به خواص اصلی شکل مرز نباشد.

روشی که اغلب برای حل این مشکلات به کار می‌رود، نمونه برداری مجدد مرز توسط انتخاب فضای بزرگتر از تصویر است که در شکل ۳-۲۱ نشان داده شده است.



شکل ۳-۲۱: الف) مرز دیجیتالی، ب) نمونه برداری مجدد، پ) مرز با کد زنجیره‌ای ۸ جهتی [۲۱]

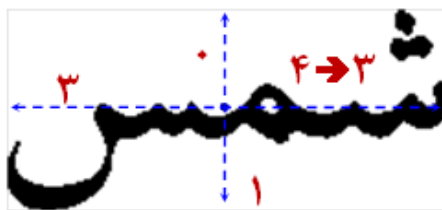
رمز زنجیره‌ای مرز، به نقطه شروع بستگی دارد. اما می‌تواند نسبت به نقطه شروع نرمال سازی شود. نقطه شروع را دوباره تعریف می‌شود به طوری که دنباله حاصل از اعداد، یک مقدار مینیمم صحیح را ایجاد کند. چرخش را نیز می‌توان با استفاده از تفاضل اول زنجیره رمز به جای خود رمز، نرمال کرد. این تفاضل، با شمارش تعداد تغییرات جهت دو عنصر همجوار از رمز در جهت پادراستگرد، به دست می‌آید. به عنوان مثال تفاضل اول رمز زنجیره‌ای ۴ جهتی ۱۰۱۰۳۲۲ برابر با ۳۱۳۳۰۳۰ است. از سمت چپ جابجایی ۱ به ۰ معادل ۳ تغییر، ۰ به ۱ یک تغییر و به همین ترتیب بقیه دنباله تفاضل می‌شوند.

۳-۱-۲- ویژگی مکان مشخصه^۱

این ویژگی توصیف خوبی از حفره‌ها، تعداد دندان‌ها و بالا و پایین رونده‌ها را ارائه می‌کند. نحوی استخراج آن به این ترتیب است که، به ازای هر پیکسل پس زمینه در چهار جهت راست، بالا، چپ و پایین حرکت می‌کنیم و تعداد عبور از سفید به سیاه را می‌شماریم. جهت جلوگیری از افزایش تعداد ویژگی‌ها، اگر تعداد عبور از سفید به سیاه بیش از دو بود، آن را به دو محدود می‌کنیم. به این ترتیب در هر جهت سه حالت برای تعداد عبور از سفید به سیاه اتفاق می‌افتد؛ ۰، ۱ و ۲. بنابراین هر پیکسل می‌تواند ۸۱ حالت $(۳ \times ۳ \times ۳ \times ۳)$ در چهار جهت داشته باشد. لذا طول بردار ویژگی در این روش برابر ۸۱ خواهد بود که به ازای اتفاق افتادن یک حالت برای یک پیکسل، مولفه ویژگی متناظر با آن حالت یکی افزوده می‌شود. [۱۴]

اما با توجه به اینکه در حروف فارسی دندان‌ها زیاد داریم، احتمال اینکه در جهت‌های راست و چپ، تعداد گذر از سفید به سیاه بیش از دو باشد بسیار است از این رو ویژگی مکان مشخصه توسعه یافته تعریف شده است که تعداد عبور از چهار جهت به ۳ محدود شده است که تعداد ویژگی‌ها را به ۲۵۶ افزایش می‌دهد. شکل ۳-۲۲ نحوی استخراج این ویژگی را نشان می‌دهد.

¹ *Characteristic Loci*



$$(1303)_{4,3} = 3 + 0 \times 4 + 3 \times (4 \times 3) + 1 \times (4 \times 3 \times 4) = 87$$

شکل ۳-۲۲: نحوه استخراج ویژگی مکان مشخصه توسعه یافته [۱۴]

۳-۲-۳- ویژگی‌های ساختاری^۱

در این روش‌ها ویژگی‌هایی که ساختار توپولوژیک و هندسی یک نماد را توصیف می‌کنند، استخراج می‌شوند. بعضی از ویژگی‌های معمول مورد استفاده در این روش، پاره‌خطها، نقاط پایانی، تلاقی بین خطوط، خمیدگیها، حلقه‌ها و نسبت ابعاد می‌باشند. ویژگی‌های ساختاری، در برابر نویز و تغییر شکل الگو مقاومت بالایی دارند اما استخراج آنها ساده نمی‌باشد.

۳-۲-۳-۱- پروفایل

ویژگی پروفایل، نمای کلمه دیده شده از طرفین تصویر را نشان می‌دهند. برای استخراج این ویژگی، تعداد پیکسل‌های (فاصله) بین لبه الگو و مرز تصویر در هر جهت محاسبه می‌شود. اندازه بردار ویژگی حاصل در این روش برای هر کاراکتر برابر با محیط کاراکتر خواهد بود. ویژگی پروفایل به خوبی شکل بیرونی الگو را توصیف می‌کند اما نسبت به چرخش الگو حساس می‌باشد. نمونه‌ای از نحوه استخراج ویژگی با این روش در شکل ۳-۲۳ زیر نشان داده شده است [۳۱].

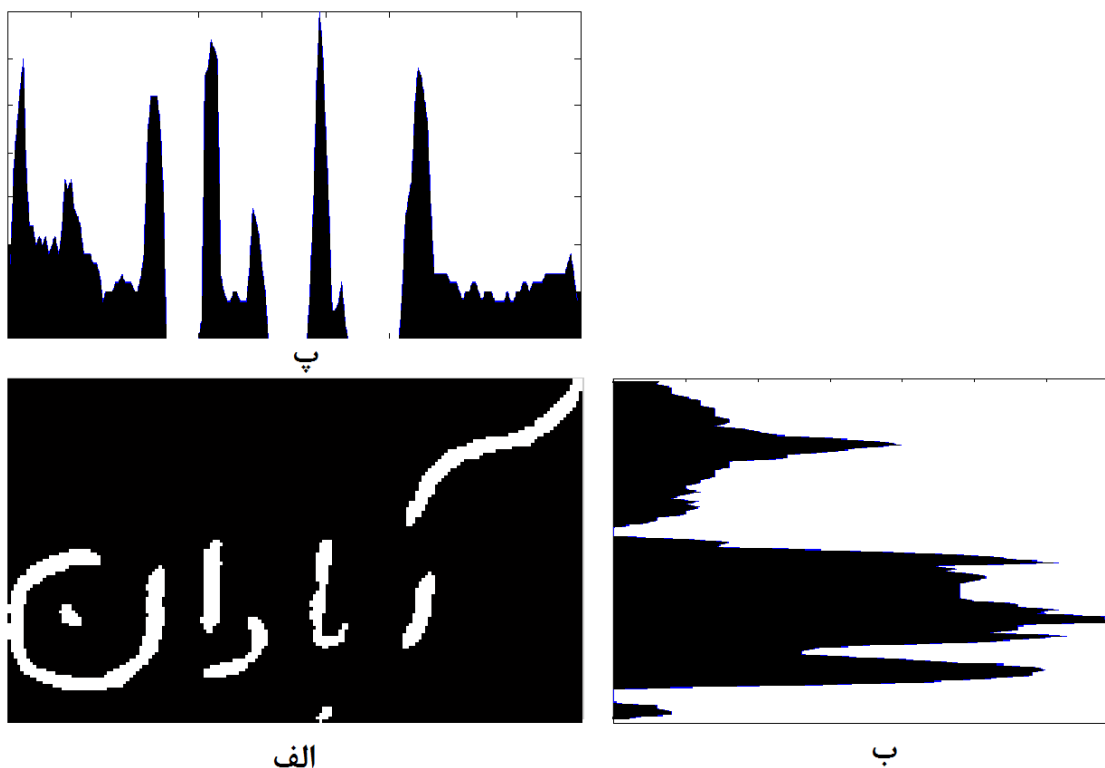
¹ Structural



شکل ۳-۲۳: استخراج ویژگی پروفایل، (الف) تصویر اصلی، (ب) پروفایل پایین، (پ) پروفایل بالا

۳-۲-۲- هیستوگرام پروفایل افکنش

این ویژگی تعداد پیکسل‌های پیش زمینه در هر سطر و ستون تصویر الگو را به طور جداگانه محاسبه می‌کند به نوبت و تغییرات ایجاد شده در شکل الگو حساس نمی‌باشد اما به چرخش الگو حساس می‌باشد. نحو استخراج ویژگی به این روش در شکل ۳-۲۴ زیر برای الگوی نمونه ترسیم شده است.



شکل ۳-۲۴: هیستوگرام پروفایل افکنش، (الف) تصویر کلمه، (ب) افکنش افقی، (پ) افکنش عمودی

۳-۴ - خوشه‌بندی^۱

خوشه‌بندی را می‌توان به عنوان مهم‌ترین مسئله در یادگیری بدون نظارت در نظر گرفت. در خوشه‌بندی سعی می‌شود تا داده‌ها به خوشه‌هایی تقسیم شوند که شباهت بین داده‌های درون هر خوشه حداکثر و شباهت بین داده‌های درون خوشه‌های متفاوت حداقل شود. بر خلاف طبقه‌بندی که در آن هر داده به یک طبقه (کلاس) از پیش مشخص شده تخصیص می‌یابد، در خوشه‌بندی هیچ اطلاعی از کلاس‌های موجود برای داده‌ها در دست نیست و به عبارتی خوشه‌ها مستقیماً از ویژگی خود داده‌ها استخراج می‌شوند.

¹ Clustering

فرض کنید فرهنگ لغت l شامل n کلمه باشد، کاهش فرهنگ لغت به صورت یک زیر مجموعه l' از l تعریف می‌شود، برای بررسی کارایی فرهنگ لغت، متغیر تصادفی A را به این صورت تعریف می‌کنیم که اگر مجموعه دوم شامل کلمه مورد نظر باشد $A=I$ و در غیر این صورت صفر خواهد بود لذا برای معرفی میزان کاهش در فرهنگ لغت، متغیر تصادفی R را به صورت زیر تعریف می‌کنیم. [۱۶]

$$R = \frac{l-l'}{l} \quad (23-3)$$

لذا کارایی سیستم نیز به صورت زیر تعریف می‌شود:

دقت کاهش فرهنگ لغت:

$$\alpha = E(A) \quad (24-3)$$

میزان کاهش فرهنگ لغت:

$$\rho = E(R) \quad (25-3)$$

روش‌های خوشه‌بندی را می‌توان از چندین جنبه تقسیم‌بندی کرد: [۳۲]

۱. خوشه‌بندی انحصاری^۱ و خوشه‌بندی با هم‌پوشی^۲

در روش خوشه‌بندی انحصاری پس از خوشه‌بندی هر داده دقیقاً به یک خوشه تعلق می‌گیرد مانند روش خوشه‌بندی K -Means. ولی در خوشه‌بندی با هم‌پوشی پس از خوشه‌بندی به هر داده یک درجه تعلق بازاء هر خوشه نسبت داده می‌شود. به عبارتی یک داده می‌تواند با نسبت‌های متفاوتی به چندین خوشه تعلق داشته باشد. نمونه‌ای از آن خوشه‌بندی فازی است.

۲. خوشه‌بندی سلسله مراتبی^۳ و خوشه‌بندی مسطح^۴ [۳۲]

در روش خوشه بندی سلسله مراتبی، به خوشه‌های نهایی بر اساس میزان عمومیت آنها ساختاری سلسله مراتبی نسبت داده می‌شود. مانند روش *Single Link*. ولی در خوشه‌بندی مسطح تمامی

¹ Exclusive or Hard Clustering

² Overlapping or Soft Clustering

³ Hierarchical

⁴ Flat

خوشه‌های نهایی دارای یک میزان عمومیت هستند مانند *K-Means*. به ساختار سلسله مراتبی حاصل از روش‌های خوشه‌بندی سلسله مراتبی دندوگرام (*Dendogram*) گفته می‌شود.

با توجه با اینکه روش‌های خوشه‌بندی سلسله مراتبی اطلاعات بیشتر و دقیق‌تری تولید می‌کنند برای تحلیل داده‌های با جزئیات پیشنهاد می‌شوند ولی از طرفی چون پیچیدگی محاسباتی بالایی دارند برای مجموعه داده‌های بزرگ روش‌های خوشه‌بندی مسطح پیشنهاد می‌شوند.

سه روش کلی که به توضیح آن می‌پردازیم عبارت‌اند از: [۳۲]

- روش جمع مربعات^۱: که به مینیمم کردن معیار خطای جمع مربعات می‌پردازد. روش‌هایی مانند *fuzzy c-means*, *k-means*.
- روش‌های ترکیبی^۲: که در آن تابع چگالی خوشه‌ها به صورت مجموع چند چگالی بیان می‌شود.
- روش‌های سلسله مراتبی: که در آن خوشه‌ها از ماتریس تشابه داده‌ها به دست می‌آید.

۳-۴-۱- روش جمع مربعات [۳۲]

در خوشه‌بندی به روش‌های جمع مربعات، داده‌ها به دسته‌هایی که معیارهای از پیش تعریف شده‌ای را بهینه می‌کند، تقسیم می‌شوند. این روش‌ها در انتخاب معیارهای بهینه سازی خوشه‌ها متفاوت‌اند که تعدادی از این معیارها در زیر توضیح داده می‌شوند.

¹ *Sum Of Squares*

² *Mixture Models*

معیارهای خوشه‌بندی [۳۲]

اگر $x_1 \dots x_n$ را n نمونه داده در نظر بگیریم آنگاه ماتریس همبستگی^۱ این داده‌ها به صورت زیر محاسبه می‌شود.

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n (x_i - m)(x_i - m)^T \quad (۲۶-۳)$$

m میانگین نمونه‌ها به صورت $m = \frac{1}{n} \sum_{i=1}^n x_i$ است. با فرض داشتن g خوشه، آنگاه ماتریس پراکندگی داخلی کلاس‌ها یا ماتریس پراکندگی داخلی گروه‌ها به صورت زیر محاسبه می‌شود.

$$S_W = \frac{1}{n} \sum_{j=1}^g \sum_{i=1}^n z_{ji} (x_i - m_j)(x_i - m_j)^T \quad (۲۷-۳)$$

در معادله بالا اگر x_i متعلق به خوشه j ام باشد آنگاه $z_{ij} = 1$ ، در غیر این صورت صفر است. m_j میانگین و n_j تعداد اعضای خوشه j ام است. ماتریس پراکندگی بین کلاس‌ها نیز به صورت زیر محاسبه می‌شود.

$$S_B = \hat{\Sigma} - S_W = \sum_{j=1}^g \frac{n_j}{n} (m_j - m)(m_j - m)^T \quad (۲۸-۳)$$

روش‌های موجود از ماتریس‌های بالا به منظور بهینه‌سازی خوشه‌ها استفاده می‌کنند. معیارهایی از قبیل:

۱. مینیمم کردن $Tr(S_W)$

۲. مینیمم کردن $|S_W| / |\hat{\Sigma}|$

۳. ماکزیمم کردن $Tr(S_W^{-1} S_B)$

۴. مینیمم کردن $Tr(\hat{\Sigma}^{-1} S_W)$

چند نمونه روش‌های خوشه‌بندی با استفاده از معیارهای ذکر شده عبارت‌اند از:

¹ Covariance Matrix

۳-۴-۱-۱- روش k -means [۳۲]

این روش علی‌رغم سادگی آن یک روش پایه برای بسیاری از روش‌های خوشه‌بندی دیگر (مانند خوشه‌بندی فازی) محسوب می‌شود. این روش، روشی انحصاری است. برای این الگوریتم شکل‌های مختلفی بیان شده است. ولی همه آنها دارای روالی تکراری هستند که برای تعدادی ثابت از خوشه‌ها سعی در تخمین موارد زیر دارند:

بدست آوردن نقاطی به عنوان مراکز خوشه‌ها این نقاط در واقع همان میانگین نقاط متعلق به هر خوشه هستند و نسبت دادن هر نمونه داده به یک خوشه که آن داده کمترین فاصله تا مرکز آن خوشه را دارا باشد.

الگوریتم روش k -means به صورت زیر است.

۱. در ابتدا تعداد خوشه‌ها مشخص می‌شود.
۲. مقدار میانگین خوشه‌ها به صورت تصادفی مقدار دهی اولیه می‌شود.
۳. داده‌ها بر اساس این میانگین‌ها خوشه‌بندی می‌شوند.
۴. میانگین جدید خوشه‌ها محاسبه می‌شود.
۵. اگر خوشه‌بندی نمونه‌ها تغییر نکرد از الگوریتم خارج شده و در غیر این صورت به مرحله سه می‌رود.

لذا می‌توان تابع هزینه زیر را برای روش k -means تعریف کرد که الگوریتم سعی در مینیمم کردن آن را دارد.

$$J_{MSE} = \sum_{i=1}^c \sum_{x \approx \omega_i} |x - \mu_i|^2 \quad (۳-۲۹)$$

تابع فوق زمانی مینیمم می شود که نمونه‌ها به نزدیکترین خوشه‌ها تعلق یابند یعنی نمونه‌ها کمترین فاصله با میانگین‌ها را داشته باشند.

اشکال این روند این است که در هر بار اجرا، برنامه به جواب بهینه همگرا نمی شود و به انتخاب نقاط اولیه حساس است و نکته دیگر اینکه تعداد خوشه‌ها باید از قبل مشخص باشد.

۳-۴-۱-۲- روش Fuzzy C-Means

FCM یک روش بر مبنای میزان شباهت بین یک نمونه با هر کدام از نقاط میانی خوشه‌ها یعنی C است.

تابع هدف این الگوریتم بصورت زیر تعریف می شود: [۳۲]

$$J_{FCM}(U, v, X) = \sum_{k=1}^n \sum_{i=1}^c (\mu_{ik})^m \|x_k - v_i\|^2 \quad 1 \leq m < \infty \quad (30-3)$$

که در آن μ_{ik} مقدار عضویت فازی نمونه K ام در کلاس i ام است، $X = \{x_1, \dots, x_n\}$ یک مجموعه داده متناهی در R^d است که $1 \leq k \leq n$ و یک بردار d بعدی است. $V = \{v_1, \dots, v_c\}$ مجموعه نقاط میانی کلاس‌هاست که در آن $1 \leq i \leq c$ و $v_i \in R^d$. مقادیر عضویت باید مثبت باشد و در معادله زیر صدق کنند:

$$\sum_{i=1}^c \mu_{ik} = 1 \quad (31-3)$$

تابع J زمانی مینیمم می شوند که مقادیر عضویت یک نمونه بالا باشد و فاصله آن به نقطه میانی یک کلاس نزدیک باشد. مقادیر عضویت پایین زمانی بوجود می آیند که فاصله آنها با مرکز زیاد باشد. پارامتر m مقدار وزن است که $m > 1$ است و درجه فازی بودن را در توابع عضویت مشخص می کند به گونه‌ای که

وقتی m به یک نزدیک می‌شود، توابع عضویت سخت تر می‌شوند (خوشه بندی سخت) و به توابع باینری می‌رسند. با افزایش m ، توابع عضویت بیشتر فازی می‌شوند.

$$\mu_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{\|x_k - v_i\|}{\|x_k - v_j\|} \right)^{2/(m-1)}} \quad \forall i, k$$

$$v_i = \frac{\sum_{k=1}^n (\mu_{ik})^m x_k}{\sum_{k=1}^n (\mu_{ik})^m}$$

(۳۲-۳)

v_i نشان دهنده مرکز کلاس i ام است.

۳-۴-۱-۳-روش Isodat [۳۳]

این روش مانند روش k -means با کمی تفاوت برای بدست آوردن تعداد خوشه‌ها به صورت خودکار است. این روش نیاز به مقداری تعاریفات اولیه دارد که شامل:

۱. مینیمم مقدار نمونه‌ها که در هر خوشه می‌تواند قرار گیرد N_{\min_EX}

۲. تعداد تقریبی مورد نظر خوشه‌ها N_D

۳. ماکزیمم پارامتر پراکندگی برای جدا کردن خوشه‌ها σ_s^2

۴. ماکزیمم فاصله جدایی برای ترکیب کردن D_{MERGE}

۵. ماکزیمم تعداد خوشه‌ها که می‌توان ترکیب کرد N_{MERGE}

این روش در یک روند تکرار شونده کار می‌کند به این ترتیب که:

۱. الگوریتم k -means اجرا می‌شود.

۲. خوشه‌هایی که معیار عدم شباهت را برآورده کنند از هم جدا می‌شوند.

۳. خوشه‌هایی که معیار شباهت را برآورده کنند با هم ترکیب می‌شوند.

۴. به مرحله اول می رود.

جزئیات این الگوریتم در زیر آورده شده است:

۱. تعداد کلاسها N_c مقدار دهی اولیه می شود و N_c داده اول به ترتیب به عنوان مراکز خوشهها در نظر گرفته می شوند.

۲. هر داده به خوشه با نزدیکترین مرکز تعلق می گیرد.

• اگر خوشه بندی دادهها پس از یک بار اتمام برنامه، تغییر نکرد از الگوریتم خارج می شویم.

۳. خوشههایی که دارای تعداد نمونههای کمتر از N_{\min_EX} هستند حذف می شوند.

• دادههای آن به خوشه ای که کمترین فاصله را با مرکز آن دارند، وارد می شوند.

• در نتیجه مقدار N_c نیز کاهش میابد.

۴. برای هر خوشه k

• میانگین محاسبه می شود.

• میانگین فاصلهها تا مراکز خوشهها محاسبه می شود.

$$d_{AVG} = \frac{1}{N} \sum_{k=1}^{N_c} N_k d_k \quad (3-33)$$

$$d_k = \frac{1}{N_k} \sum_{x \in \omega_k} |x - \mu_k| \quad (3-34)$$

• واریانس محورها محاسبه شده و محور با بیشترین میزان واریانس را بدست می آوریم n^*

۵. برای هر خوشه k که $\sigma_k^2(n^*) > \sigma_s^2$ ، آنگاه اگر $N_k > 2N_{\min_EX} + 1$ و یا $N_c < N_D/2$

• خوشه به دو خوشه با میانگینهای μ_{k1} و μ_{k2} تقسیم می شود.

$$\mu_{k1}(n^*) = \mu_k(n^*) + \varepsilon \sigma_k(n^*) \quad \circ$$

$$\mu_{k2}(n^*) = \mu_k(n^*) - \varepsilon \sigma_k(n^*) \quad \circ$$

$$0 < \varepsilon < 1 \quad \circ$$

- در نتیجه مقدار N_c افزایش میابد.
- داده‌های خوشه تجزیه شده به دو خوشه، بر اساس میانگین‌های جدید بدست آمده، دسته‌بندی می‌شوند.

$$N_c > 2N_D \quad \text{اگر } ۶.$$

- فاصله تمامی مراکز خوشه‌ها بدست میاید $D_{ij} = d(\mu_i, \mu_j)$
- D_{ij} به صورت کاهشی مرتب می‌شوند.
- برای هر جفت خوشه، اگر هیچ خوشه‌ای ترکیب نشد و فاصله $D_{ij} < D_{MERGE}$ و تعداد جفت خوشه‌هایی که در حلقه ترکیب می‌شوند بیشتر از N_{MERGE} نباشند انگاه:

$$\circ \text{ خوشه } i \text{ و } j \text{ با هم ترکیب می‌شوند.}$$

$$\circ \text{ مرکز جدید آنها به صورت زیر به دست می‌آید.}$$

$$\mu' = \frac{N_i \mu_i + N_j \mu_j}{N_i + N_j} \quad (۳۵-۳)$$

$$\circ \text{ در نتیجه } N_c \text{ کاهش میابد.}$$

به مرحله یک باز می‌گردیم.

۳-۴-۲- مدل‌های ترکیبی [۳۲]

در خوشه‌بندی به روش ترکیب مدل‌ها، هر خوشه توسط توزیع احتمالاتی مختلفی مدل می‌شود. این توزیع‌های احتمالاتی ممکن است از یک خانواده باشند اما در مقادیر پارامترهای توزیع متفاوت خواهند

بود. اساساً در توزیع مدل مخلوط، چگالی احتمال^۱ به صورت ترکیبی از چگالی اجزا به شکل زیر بیان می‌شود.

$$p(x) = \sum_{i=1}^g \pi_i p(x; \theta_i) \quad (36-3)$$

در معادله بالا π_i احتمال مجموع است که $\sum_{i=1}^g \pi_i = 1$ و $p(x; \theta_i)$ تابع احتمال است که به بردار پارامترهای θ_i وابسته است. بنابراین در این معادله سه دسته از پارامترها باید تخمین زده شود، مقدار π_i ، اجزای بردار θ_i و مقدار g که تعداد توزیع‌های برای مدل کردن هر خوشه است.

توزیع‌های مخلوط مختلفی و روش‌های زیادی برای تخمین پارامترهای آنها وجود دارد به عنوان مثال مدل مخلوط گوسی^۲، که در آن چگالی احتمالاتی به صورت زیر بیان می‌شود.

$$p(x) = \sum_{i=1}^g \pi_i p(x; \Sigma_i, \mu_i) \quad (37-3)$$

که در معادله بالا μ_i و Σ_i به ترتیب میانگین و ماتریس همبستگی جز i ام، توزیع نرمال است که به صورت زیر بیان می‌شود.

$$P(x; \Sigma_i, \mu_i) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma_i|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i)\right\} \quad (38-3)$$

در این رابطه p ابعاد بردار ویژگی x است. الگوریتم EM^3 کاربرد زیادی برای تخمین پارامترهای GMM دارد.

۳-۴-۳ روش‌های سلسله مراتبی [۳۲]

روش‌های سلسله مراتبی از روش‌های متداول در بیان ساختار داده‌هاست که در آن به نمونه‌ها بر اساس عضویت، ساختار سلسله مراتبی نسبت داده می‌شود. درخت سلسله مراتبی همانطور که در شکل ۳-۲۵

¹ Probability Function

² GMM (Gaussian Mixture Model)

³ Expectation Maximisation

نشان داده شده است یک مجموعه تو در تو از خوشه‌هاست که به صورت یک نمودار درختی یا دندوگرام نشان داده می‌شود.



شکل ۳-۲۵: نمودار دندوگرام

دو روش برای محاسبه درخت سلسله مراتبی وجود دارد

۳-۴-۳-۱- بالا به پایین یا تقسیم کننده^۱

در این روش در ابتدا کل داده‌ها به عنوان یک خوشه در نظر گرفته می‌شوند و در هر مرحله داده‌هایی که شباهت کمتری دارند به خوشه‌های جدا تقسیم می‌شوند و این روند ادامه می‌یابد تا تمامی خوشه‌ها تنها شامل یک داده شوند. [۳۲]

۳-۴-۳-۲- پایین به بالا یا متراکم کننده^۲

بر خلاف حالت قبل در ابتدا هر داده به عنوان یک خوشه مستقل در نظر گرفته می‌شود و بعد در هر مرحله خوشه‌ها با میزان شباهت بیشتر با هم ادغام شده و این روند تکرار می‌شود تا تمامی داده‌ها در یک خوشه قرار گیرند. [۳۲]

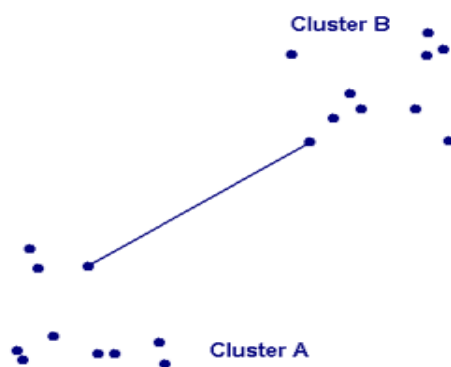
¹ Divisive Methods

² Agglomerative Methods

خوشه بندی با روش *Single-Link*

روشی ساده است که به آن نزدیکترین همسایگی^۱ نیز گفته می‌شود در این روش فاصله کمتر به عنوان میزان شباهت بیشتر شناخته می‌شود به این معنی که در خوشه‌ها همانطور که در شکل ۳-۲۶ نشان داده شده است میزان شباهت برابر است با کمترین فاصله‌ای که میان اعضای دو خوشه وجود دارد. [۳۲]

$$d_{AB} = \min_{i \in A, j \in B} d_{ij} \quad (3-39)$$



شکل ۳-۲۶: معیار شباهت در *Single-Link*

به این ترتیب که در ابتدا هر داده به عنوان یک خوشه در نظر گرفته می‌شود و بعد دو خوشه که کمترین فاصله را دارند به عنوان یک خوشه در نظر گرفته می‌شوند در این مرحله فاصله خوشه جدید با خوشه‌های قبلی به صورت کمترین فاصله بین ۲ عضو آن با سایر خوشه‌ها در نظر گرفته می‌شود و این روند ادامه می‌یابد تا تمامی داده‌ها در یک خوشه قرار گیرند. در زیر یک مثال برای این روش ذکر شده است.

فرض می‌کنیم ۶ نمونه داریم که ماتریس فاصله بین آنها در جدول زیر آورده شده است.

¹ Nearest Neighbour

جدول ۳-۲: فاصله داده‌ها

| داده‌ها | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ |
|---------|---|---|----|----|----|-----|
| ۱ | ۰ | ۴ | ۱۳ | ۲۴ | ۱۲ | ۸ |
| ۲ | | ۰ | ۱۰ | ۲۲ | ۱۱ | ۱۰ |
| ۳ | | | ۰ | ۷ | ۳ | ۹ |
| ۴ | | | | ۰ | ۶ | ۱۸ |
| ۵ | | | | | ۰ | ۸,۵ |
| ۶ | | | | | | ۰ |

طبق جدول بالا دو نمونه ۵ و ۳ نزدیکترین داده‌ها از لحاظ فاصله هستند و لذا بیشترین میزان تشابه را دارند، پس با هم ترکیب شده و تشکیل یک خوشه را می‌دهند حال خوشه‌ها با توجه به معیار *singel linke* دارای فاصله به شکل زیر خواهند شد.

جدول ۳-۳: مرحله اول خوشه‌بندی به روش *singel linke*

| خوشه‌ها | ۱ | ۲ | (۵ و ۳) | ۴ | ۶ |
|---------|---|---|---------|----|-----|
| ۱ | ۰ | ۴ | ۱۲ | ۲۴ | ۸ |
| ۲ | | ۰ | ۱۰ | ۲۲ | ۱۰ |
| (۵ و ۳) | | | ۰ | ۶ | ۸,۵ |
| ۴ | | | | ۰ | ۱۸ |
| ۶ | | | | | ۰ |

همانطور که در جدول بالا مشخص شده فاصله خوشه جدید با خوشه‌های قبلی برابر کمترین فاصله داده‌های آن از داده‌های سایر خوشه‌هاست. لذا در این مرحله نیز داده‌های ۱ و ۲ با هم ترکیب شده و تشکیل یک خوشه را می‌دهند.

جدول ۳-۴: مرحله دوم خوشه‌بندی به روش *singel linke*

| خوشه‌ها | (۲ و ۱) | (۵ و ۳) | ۴ | ۶ |
|---------|---------|---------|----|-----|
| (۲ و ۱) | ۰ | ۱۰ | ۲۲ | ۸ |
| (۵ و ۳) | | ۰ | ۶ | ۸,۵ |
| ۴ | | | ۰ | ۱۸ |
| ۶ | | | | ۰ |

در این مرحله نیز خوشه‌های شامل ۵ و ۳ و خوشه شامل داده ۴ کمترین فاصله را داشته و با هم ترکیب می‌شوند.

جدول ۳-۵: مرحله سوم خوشه‌بندی به روش *singel linke*

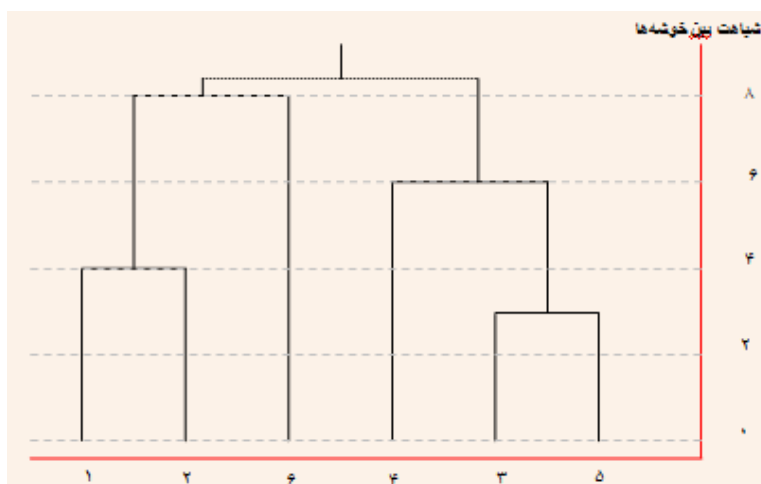
| | | | |
|-----|-------------|---------|-------------|
| ۶ | (۵ و ۳ و ۴) | (۲ و ۱) | خوشه‌ها |
| ۸ | ۱۰ | ۰ | (۲ و ۱) |
| ۸,۵ | ۰ | | (۵ و ۳ و ۴) |
| ۰ | | | ۶ |

داده ۶ نیز به خوشه شامل داده‌های (۲ و ۱) می‌پیوندد.

جدول ۳-۶: مرحله چهارم خوشه‌بندی به روش *singel linke*

| | |
|-------------|-------------|
| (۵ و ۳ و ۴) | خوشه‌ها |
| ۸,۵ | (۶ و ۲ و ۱) |

بر اساس نتایج بالا نمودار دندوگرام به صورت شکل ۳-۲۷ بدست می‌آید.

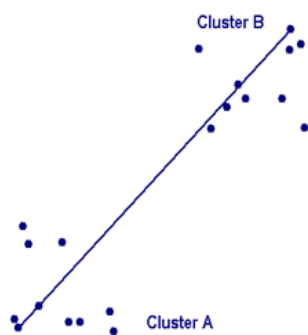


شکل ۳-۲۷: نمودار دندوگرام *singel linke*

خوشه‌بندی با روش Complete-Link

به این روش دورترین همسایگی^۱ نیز اطلاق می‌شود که در آن، ابتدا هر داده به عنوان یک خوشه در نظر گرفته می‌شود و بعد دو خوشه که کمترین فاصله را با هم دارند ترکیب می‌شوند و در مرحله بعد فاصله خوشه جدید با خوشه‌های قبلی همانگونه که در شکل ۳-۲۸ نشان داده شده است به صورت دورترین فاصله بین دو عضو آن و دیگر خوشه‌ها تعریف می‌شود و این روند تا زمانی ادامه می‌یابد که تمامی داده‌ها در یک خوشه قرار گیرند. به طور کلی در این روش فاصله میان خوشه‌ها به صورت زیر تعریف می‌شود. [۳۲]

$$d_{AB} = \max_{i \in A, j \in B} d_{ij} \quad (۴۰-۳)$$



شکل ۳-۲۸: معیار فاصله در Complete-Link

حال همان مثال بالا را با روش Complete-Link خوشه بندی می‌کنیم.

جدول ۳-۷: فاصله میان داده‌ها

| داده‌ها | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ |
|---------|---|---|----|----|----|-----|
| ۱ | ۰ | ۴ | ۱۳ | ۲۴ | ۱۲ | ۸ |
| ۲ | | ۰ | ۱۰ | ۲۲ | ۱۱ | ۱۰ |
| ۳ | | | ۰ | ۷ | ۳ | ۹ |
| ۴ | | | | ۰ | ۶ | ۱۸ |
| ۵ | | | | | ۰ | ۸,۵ |
| ۶ | | | | | | ۰ |

^۱ Farthest Neighbour

مانند حالت قبل خوشه‌های ۳ و ۵ کمترین میزان فاصله را دارند لذا با هم ترکیب شده و تشکیل یک خوشه واحد را می‌دهند.

جدول ۳-۸: مرحله اول خوشه‌بندی به روش *Complete linke*

| خوشه‌ها | ۱ | ۲ | (۵و۳) | ۴ | ۶ |
|---------|---|---|-------|----|----|
| ۱ | ۰ | ۴ | ۱۳ | ۲۴ | ۸ |
| ۲ | | ۰ | ۱۱ | ۲۲ | ۱۰ |
| (۵و۳) | | | ۰ | ۷ | ۹ |
| ۴ | | | | ۰ | ۱۸ |
| ۶ | | | | | ۰ |

همانطور که در جدول بالا مشخص است در این روش فاصله خوشه جدید با خوشه‌های قبلی برابر بیشترین فاصله اعضای آن با خوشه‌های دیگر است. پس از پیدا کردن فاصله‌های جدید در این مرحله نیز مشخص می‌شود که خوشه شامل داده ۲ کمترین فاصله را با خوشه شامل داده ۱ دارد.

جدول ۳-۹: مرحله دوم خوشه‌بندی به روش *Complete linke*

| خوشه‌ها | (۲و۱) | (۵و۳) | ۴ | ۶ |
|---------|-------|-------|----|----|
| (۲و۱) | ۰ | ۱۳ | ۲۴ | ۱۰ |
| (۵و۳) | | ۰ | ۷ | ۹ |
| ۴ | | | ۰ | ۱۸ |
| ۶ | | | | ۰ |

در این مرحله نیز خوشه‌های شامل (۵و۳) و داده ۴ کمترین فاصله را دارند لذا با هم ترکیب می‌شوند.

جدول ۳-۱۰: مرحله سوم خوشه‌بندی به روش *Complete linke*

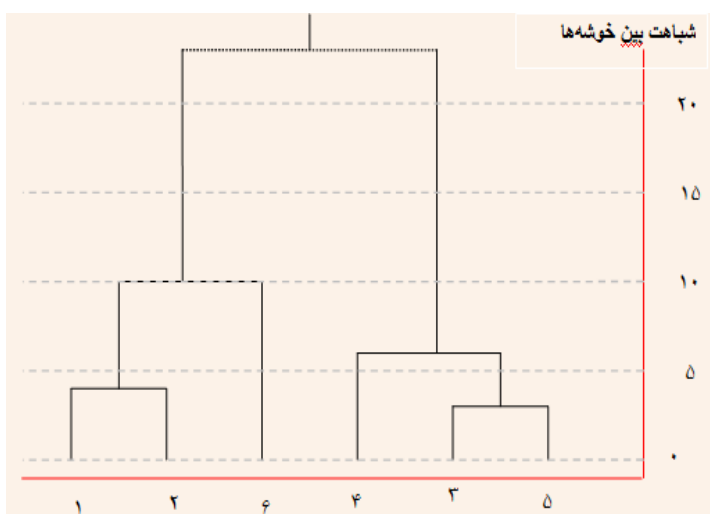
| خوشه‌ها | (۲و۱) | (۵و۴و۳) | ۶ |
|---------|-------|---------|----|
| (۲و۱) | ۰ | ۲۴ | ۱۰ |
| (۵و۴و۳) | | ۰ | ۱۸ |
| ۶ | | | ۰ |

و به همین ترتیب در مرحله بعد خوشه‌های شامل داده ۶ و (۲ و ۱) با هم ترکیب می‌شوند.

جدول ۱۱-۳: مرحله چهارم خوشه‌بندی به روش *Complete linke*

| خوشه‌ها | (۶ و ۲ و ۱) | (۵ و ۴ و ۳) |
|-------------|-------------|-------------|
| (۶ و ۲ و ۱) | ۰ | ۲۴ |
| (۵ و ۴ و ۳) | ۰ | ۰ |

و به این ترتیب نمودار دندوگرام نیز به صورت شکل ۲۹-۳ بدست می‌آید.



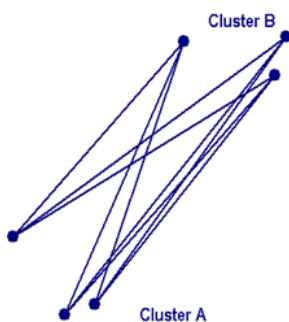
شکل ۲۹-۳: نمودار دندوگرام *Complete-Link*

خوشه‌بندی با روش *Average-Link*

در این روش به منظور کاهش حساسیت خوشه‌بندی به نویز، برای محاسبه میزان شباهت میانگین فاصله بین تمامی اعضای دو خوشه در نظر گرفته می‌شود. ابتدا هر داده به عنوان یک خوشه در نظر گرفته می‌شود.

شود، دو خوشه نزدیک از لحاظ فاصله با هم ترکیب می‌شوند و در مرحله بعد فاصله خوشه جدید با خوشه‌های قبلی، به صورت میانگین فاصله ۲ عضو جدید با اعضای سایر خوشه‌ها محاسبه می‌شود و دوباره خوشه‌های نزدیک به هم ترکیب می‌شوند این روند دوباره تکرار می‌شود تا تمامی نمونه‌ها در یک خوشه قرار گیرند. در این روش فاصله بین دو خوشه همانطور که در شکل ۳-۳۰ نیز نشان داده شده است به صورت زیر محاسبه می‌شود. [۳۲]

$$d_{AB} = \frac{\sum_{i \in A, j \in B} d_{ij}}{N_A N_B} \quad (41-3)$$



شکل ۳-۳۰: معیار فاصله در Average-Link

این بار مثال بالا را با روش Average-Link خوشه بندی می‌کنیم.

جدول ۳-۱۲: فاصله بین داده‌ها

| داده‌ها | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ |
|---------|---|---|----|----|----|-----|
| ۱ | ۰ | ۴ | ۱۳ | ۲۴ | ۱۲ | ۸ |
| ۲ | | ۰ | ۱۰ | ۲۲ | ۱۱ | ۱۰ |
| ۳ | | | ۰ | ۷ | ۳ | ۹ |
| ۴ | | | | ۰ | ۶ | ۱۸ |
| ۵ | | | | | ۰ | ۸,۵ |
| ۶ | | | | | | ۰ |

مرحله اول کاملا مانند دو روش قبل است.

جدول ۳-۱۳: مرحله اول خوشه‌بندی به روش *Average linke*

| خوشه‌ها | ۱ | ۲ | (۵و۳) | ۴ | ۶ |
|---------|---|---|-------|-----|------|
| ۱ | ۰ | ۴ | ۱۲,۵ | ۲۴ | ۸ |
| ۲ | | ۰ | ۱۰,۵ | ۲۲ | ۱۰ |
| (۵و۳) | | | ۰ | ۶,۵ | ۸,۷۵ |
| ۴ | | | | ۰ | ۱۸ |
| ۶ | | | | | ۰ |

همانطور که دیده می‌شود معیار فاصله در این روش نه دورترین فاصله و نه نزدیکترین آن است و به صورت میانگین فاصله تا تک تک داده‌های خوشه‌های دیگر تعریف می‌شود.

جدول ۳-۱۴: مرحله دوم خوشه‌بندی به روش *Average linke*

| خوشه‌ها | (۲و۱) | (۵و۳) | ۴ | ۶ |
|---------|-------|-------|-----|------|
| (۲و۱) | ۰ | ۱۱,۵ | ۲۳ | ۹ |
| (۵و۳) | | ۰ | ۶,۵ | ۸,۷۵ |
| ۴ | | | ۰ | ۱۸ |
| ۶ | | | | ۰ |

خوشه شامل داده ۴ و داده‌های (۵و۳) با هم ترکیب می‌شوند.

جدول ۳-۱۵: مرحله سوم خوشه‌بندی به روش *Average linke*

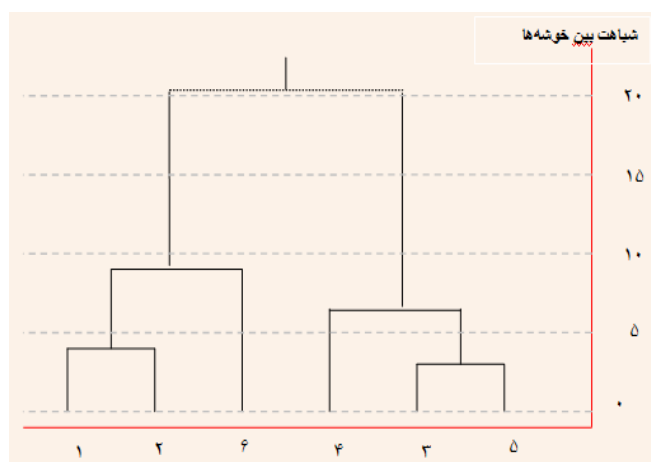
| خوشه‌ها | (۲و۱) | (۵و۴و۳) | ۶ |
|---------|-------|---------|----|
| (۲و۱) | ۰ | ۲۳ | ۹ |
| (۵و۴و۳) | | ۰ | ۱۸ |

و در گام آخر

جدول ۳-۱۶: مرحله چهارم خوشه‌بندی به روش *Average linke*

| خوشه‌ها | (۶ و ۲) | (۵ و ۴ و ۳) |
|-------------|---------|-------------|
| (۶ و ۲) | ۰ | ۲۴ |
| (۵ و ۴ و ۳) | | ۰ |

در نهایت نمودار دندوگرام حاصل در شکل ۳-۳۱ رسم شده است.



شکل ۳-۳۱: نمودار دندوگرام *Average-Link*

۳-۵ - ماشین بردار پشتیبان [۳۴]

فرض کنید تعدادی از بردارهای ویژگی یا الگوهای آموزشی بصورت $\{x_1, x_2, \dots, x_N\}$ داریم که هر کدام یک بردار ویژگی d بعدی بوده و دارای برچسب y_i است که برای دو کلاس به ترتیب ۱ و -۱ می‌باشد. هدف حل یک مساله طبقه‌بندی دو کلاسه بصورت بهینه می‌باشد. فرض کنید این دو کلاس را با تابع تمایز $f(x)$ و با یک ابر صفحه H^1 با معادله زیر بخواهیم از هم جدا کنیم:

$$H: w \cdot x + b = 0$$

(۳-۴۲)

¹ Hyperplane

بردار وزن w ، بردار عمود بر ابر صفحه جدا کننده و b مقدار بایاس^۱ است و منظور از $w \cdot x$ حاصل ضرب داخلی می‌باشد. فرض کنید داده‌های دو کلاس جدایی پذیر باشند و بردارهای ویژگی مرزی کلاس اول روی ابر صفحه H^+ و بردارهای ویژگی مرزی کلاس دوم روی ابر صفحه H^- قرار گیرند. ابر صفحات H^+ و H^- به صورت زیر تعریف می‌شوند.

$$H^+ : w \cdot x + b = 1 \quad (43-3)$$

$$H^- : w \cdot x + b = -1 \quad (44-3)$$

الگوهایی که بر روی ابر صفحات H^+ و H^- قرار می‌گیرند، بردار پشتیبان نامیده می‌شوند. ناحیه بین دو ابر صفحه را حاشیه یا ناحیه مرزی^۲ گویند که برابر $2/\|w\|$ خواهد بود. طراحی ابر صفحه بهینه بر این استوار است که با شرط طبقه‌بندی درست نمونه‌ها، عرض ناحیه مرزی حداکثر شود، یعنی $2/\|w\|$ ماکزیمم شود و در نتیجه $\|w\|$ منیمم گردد. که این دو شرط به صورت زیر تعریف می‌شود.

$$\begin{aligned} & \underset{w,b}{\text{minimize}} \left(\frac{\|w\|^2}{2} \right) & (3) & \quad (45-3) \\ & \text{Subject to: } & y_i(w \cdot x_i + b) \geq 1 \text{ for } i = 1, 2, \dots, N \end{aligned}$$

¹ Bias

² Margin

۳-۵-۱- ماشین بردار پشتیبان غیر خطی

می‌توان بردار ورودی x را با یک تبدیل غیر خطی به یک فضا با بعد زیاد انتقال داده و در آن فضا حاصلضرب داخلی را انجام داد و ثابت کرد که اعمال این هسته در فضای ورودی با بعد کم، محاسبات را به شدت کاهش می‌دهد. چند نمونه، توابع هسته^۱ در زیر معرفی شده‌اند. [۳۵]

هسته چند جمله‌ای^۲:

$$K(x, y) = (g * x \cdot y + \text{coef})^q \quad (۴۶-۳)$$

$Coef$ نشان دهنده عدد ثابت جمع شونده و q درجه چند جمله ایست

هسته^۳ rbf :

$$K(x, y) = \exp(-g \times (x - y)^2) \quad (۴۷-۳)$$

هسته سیگموئید^۴:

$$K(x, y) = \tanh(g \times x \cdot y + \text{coef}) \quad (۴۸-۳)$$

۳-۵-۲- ماشین بردار پشتیبان در حالت چند کلاسه

ماشین بردار پشتیبان یک طبقه بندی کننده دودویی است. بنابراین در حالتی که بیش از دو کلاس وجود داشته باشد نمی‌توان مستقیماً از آن استفاده کرد. در حالت کلی برای استفاده از طبقه‌بند دودویی در حالت چند کلاسه باید ابتدا چند طبقه‌بندی کننده دودویی طراحی شود، طبقه‌بندی نهایی با استفاده از ادغام اطلاعات طبقه‌بندی کننده‌های دودویی انجام می‌گیرد. برای استفاده از ماشین بردار پشتیبان در حالت

¹ Kernel

² Polynomial

³ Radial Basis Function

⁴ Sigmoid

چند کلاسه این تکنیک کلی با دو روش متفاوت اجرا می‌شود. برای مثال فرض کنید k کلاس وجود داشته باشد. در روش اول که آن را روش یکی در مقابل بقیه^۱ می‌نامند k طبقه‌بندی کننده دودویی که هر کدام برای جدا کردن یک کلاس از بقیه کلاس‌ها می‌باشد طراحی می‌شود. بنابراین در این حالت برای آموزش هر طبقه‌بندی کننده از تمام نقاط آموزشی k کلاس استفاده می‌گردد. هر نمونه با تمام k طبقه‌بندی کننده حاصل طبقه‌بندی می‌شود. در نهایت نمونه با شماره طبقه‌بندی کننده‌ای که دارای بیشترین مقدار $w.x+b$ است برچسب زده خواهد شد. در روش دوم که آن را دسته‌بندی کردن زوج-زوج^۲ یا یکی در مقابل یکی می‌نامند، برای تمامی جفت‌های مختلف k کلاس، یک طبقه‌بندی کننده طراحی می‌شود. سپس هر نمونه ورودی با تمام این طبقه‌بندی کننده‌ها طبقه‌بندی شده و در هر طبقه‌بندی به کلاس برنده یک رای داده می‌شود. کلاسی که حداکثر آرا را آورده است به عنوان کلاس نمونه ورودی در نظر گرفته خواهد شد. [۳۶]

¹ *One versus the Rest*

² *Pairwise Classification*

فصل ۴

روش پیشنهادی

۴-۱- مقدمه

در این تحقیق روشی برای بازشناسی کلمات دست‌نوشته فارسی ارائه کردیم که بدین منظور از شکل کلی کلمه برای استخراج ویژگی استفاده شده است. نتایج بدست آمده از روش پیشنهادی نیز بروی پایگاه داده ایرانشهر شامل ۱۳۰۰۰ نمونه آموزش و ۲۵۰۰ نمونه آزمایش حاصل گردیده است. این پایگاه داده شامل ۳۳ تصویر از اسامی ۵۰۲ شهر با دستخط متفاوت می‌باشد. در این فصل به توضیح روش پیشنهادی می‌پردازیم.

۴-۲- پیش پردازش:

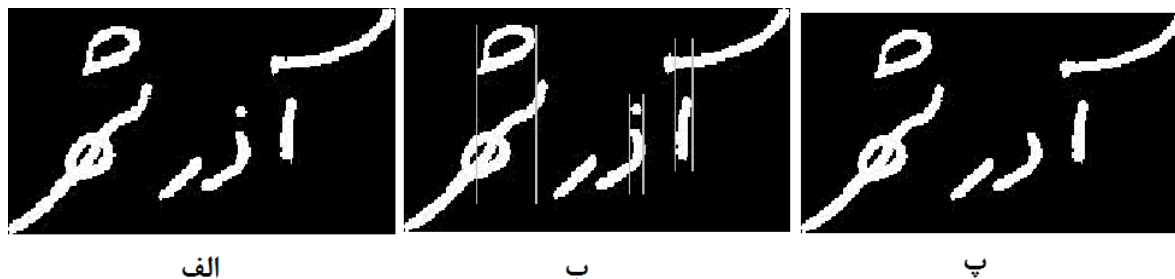
همانطور که در فصول گذشته نیز توضیح داده شده است این مرحله یکی از مراحل ضروری برای بالا بردن کیفیت تصویر ورودی است. در این تحقیق پیش پردازش‌های صورت گرفته به این ترتیب است که پس از اعمال فیلتر میانه به منظور حذف نویز، برای کاهش شکستگی‌های بدنه کلمات از روش درون یابی چندجمله‌ای استفاده شده است. در گام بعدی به دلیل استفاده از روش استخراج ویژگی ناحیه‌ای به روش تطبیقی، به منظور افزایش شباهت بین بلوک‌های متناظر در تصاویر یک کلاس و ماهیت بلوک‌بندی تطبیقی، نقاط و علائم که در کلمات دست‌نوشته فارسی با اندازه و مکان مختلف نگارش می‌شوند عاملی مزاحم در ایجاد تشابه بین بلوک‌های متناظر از تصاویر یک کلاس در نظر گرفته شده‌اند لذا سعی بر جدا کردن آنها از تصویر بدنه اصلی گرفته‌ایم.

به منظور جدا کردن نقاط و علائم از تصویر بدنه اصلی کلمات در ابتدا به تخمین مکان خط کرسی پرداخته‌ایم. برای تخمین خط کرسی از روش افکنش افقی استفاده شده است اما همانطور که در فصل قبل نیز توضیح داده‌ایم این روش برای جملات و کلمات با طول زیاد با خطای کم همراه است. با توجه به اینکه طول کلمات مورد مطالعه کوتاه است این روش تخمین خوبی از مکان خط کرسی ندارد. لذا

تغییراتی برای بهبود این روش اعمال نموده‌ایم. بر این اساس مراحل یافتن خط کرسی به صورت زیر پیشنهاد شده است:

۴-۲-۱- تخمین مکان خط کرسی

۱. در ابتدا سعی در حذف اولیه نقاط با استفاده از همپوشانی ستونی و نسبت مساحت آنها صورت می‌گیرد، به این ترتیب که اگر اجزایی از کلمه به طور کامل از لحاظ ستونی با یک جزء دیگر همپوشانی داشته و نسبت مساحت آنها از آستانه مورد نظر کمتر باشد، جزء کوچکتر حذف می‌شود، فرایند این مرحله در شکل ۴-۱ برای تصویر کلمه "آذرشهر" نشان داده شده است. همانطور که در تصاویر نیز مشاهده می‌شود سه جزء (سه نقطه مربوط به زیر کلمه "شهر"، نقطه "ذ" و الف زیر "~") از لحاظ ستونی کاملاً همپوشانی دارند از بین این سه جزء، تنها نقطه "ذ" در شرایط آستانه مساحت صدق کرده و حذف می‌شود. (نسبت مساحت در این تحقیق برابر یک سوم در نظر گرفته شده است)



شکل ۴-۱: حذف اجزای کوچک، الف) تصویر اصلی، ب) تصویر با حذف اولیه

۱. حال در بین اجزای کلمه، اجزایی که دارای بیشترین میزان ارتفاع هستند را به عنوان جزء اصلی در نظر می‌گیریم. منظور از جزء اصلی اجزایی هستند که فرض می‌شود بر روی خط کرسی قرار دارند. حداکثر تعداد این اجزاء برابر ۲ جزء در نظر گرفته شده است که باید با نیمه پایین تصویر، سطر مشترک داشته باشند. شکل ۴-۲ اجزای اصلی کلمه نمونه را نشان می‌دهد.



شکل ۴-۲: اجزای اصلی کلمه نمونه

۲. سطرهای مشترک بین اجزای اصلی محاسبه شده و همانطور که در شکل ۴-۳ نشان داده شده است از بین این سطرها آنهایی که بیشترین تعداد اجزاء جدا را در تصویر اولیه کلمه دارند بدست می‌آید.



شکل ۴-۳: الف، ب) سطرهای مشترک اجزای اصلی، پ) محدوده بیشترین تعداد اجزاء در راستای سطرهای اصلی
 ۳. در نهایت از میان سطرهای باقی مانده، سطرهای که بیشترین مقدار پیکسل‌های پیش زمینه را دارد به عنوان خط کرسی در نظر گرفته می‌شود.

نتایج اعمال روش بالا بر روی داده‌های آزمایش، پایگاه داده ذکر شده نشان دهنده ۷٪ بهبود در تخمین مکان خط کرسی نسبت به روش افکنش افقی است. چند تصویر کلمه که با روش پیشنهادی مکان خط کرسی آنها تصحیح شده است در شکل ۴-۴ نشان داده شده است.



شکل ۴-۴: دو نمونه از بهبود تخمین مکان خط کرسی توسط روش پیشنهادی

۴-۲-۲- دلایل بروز خطا در تخمین مکان خط کرسی

۱. حذف بعضی از اجزا اصلی کلمه به علت مساحت کوچک و همپوشانی کامل ستونی در مرحله اول فرایند، مانند شکل ۴-۵ که برای کلمه 'آباده' رخ داده و در آن حرف الف حذف شده است در حالی که باید در این کلمه جزء اجزای اصلی باشد. آستانه نسبت مساحت برای رفع این مشکل لحاظ شده است.



شکل ۴-۵: حذف اشتباه کاراکترها

مرحله اول از فرایند تخمین مکان خط کرسی به این علت اجرا می‌شود که نقاط در مرحله مربوط به، بدست آوردن سطرها با بیشترین تعداد اجزای جدا در محدوده سطرهای اجزای اصلی به عنوان، یک جز مجزا حاضر نباشند.

۲. علائمی که به علت ارتفاع زیاد، به نادرست جز اصلی در نظر گرفته می‌شوند. این مورد اکثرا در علامت '~' حرف 'آ' و ۳ نقطه متصل رخ می‌دهد، که برای جلوگیری از بروز آن، شرط اشتراک با نیمه پایین تصویر برای انتخاب اجزای اصلی را اعمال نموده‌ایم. در نمونه شکل ۴-۶ تنها اعمال این شرط باعث تخمین درست مکان خط کرسی شده است.



شکل ۴-۶: انتخاب درست اجزای اصلی

۱. علائمی که با حضور قسمتی از بدنه خود در سطرهای مشترک بین اجزای اصلی باعث تولید سطرهایی با تعداد اجزای زیاد در این محدوده می‌شوند که، تعداد سطرها برای یافتن مکان خط کرسی را به نادرست محدود می‌کنند. این مورد همانطور که در شکل ۴-۷ نیز نشان داده شده است، اکثراً به علت وجود قسمتی از بدنه کاراکتر '~' در محدوده سطرهای مشترک به وجود می‌آید. برای حل این مشکل نیز ماکزیمم تعداد اجزایی که در ابتدای محدوده مشترک وجود دارند را نادیده می‌گیریم به عبارت دیگر ماکزیمم مقدار اجزا در محدوده سطرهای مشترک باید بعد از یک مقدار کمتر از خود واقع شده باشد.



شکل ۴-۷: جلوگیری از محدود کردن نادرست سطرها برای یافتن خط کرسی

۴-۲-۳- جدا کردن نقاط و علائم از بدنه اصلی کلمات:

حال به منظور جدا کردن نقاط و علائم از بدنه اصلی معیار فاصله تا خط کرسی را در نظر گرفته [۱۶] و اجزایی که با ۱۰ سطر اطراف خط کرسی، سطر مشترک نداشته باشد به عنوان نقاط و علائم در نظر گرفته می‌شوند. اما در این مرحله به علت نزدیکی برخی نقاط به خط کرسی مانند نقطه 'ن' کلمه 'آبادان' جداسازی به درستی صورت نمی‌گیرد. برای رفع این مشکل مرحله اول از فرایند تخمین خط کرسی دوباره انجام می‌شود. در شکل ۴-۸ هر دو نقطه کاراکتر 'با' و 'ن' به علت فاصله تا خط کرسی حذف می‌شوند اگر این فرایند (حذف این دو نقطه) صورت نپذیرد همانگونه که در شکل نیز نشان داده شده است طبق روش بالا از بدنه اصلی جدا می‌شوند.



شکل ۴-۸: همپوشانی ستونی اجزای کلمه

در شکل ۴-۸ اگر در کاراکتر 'آ' نیز همپوشانی کامل ستونی بین '~' و 'ا' نیز وجود داشته باشد، شرط نسبت مساحت مانع از حذف کاراکتر 'ا' از بدنه اصلی می‌شود.

لذا از هر تصویر ورودی دو تصویر حاصل می‌شود، یکی تصویر بدنه اصلی کلمه و دیگری تصویر نقاط و علائم آن که در شکل ۴-۹ چند نمونه از این جداسازی نشان داده شده است.



الف

ب

شکل ۴-۹: دو نمونه از جداسازی نقاط کلمات

۴-۳- استخراج ویژگی

در این تحقیق به علت استفاده از روش‌های کاهش فرهنگ لغت، استخراج ویژگی شامل دو مرحله می‌باشد:

۴-۳-۱- ویژگی‌هایی که به منظور خوشه‌بندی نمونه‌ها مورد استفاده قرار می‌گیرد

ویژگی‌های مورد استفاده در خوشه‌بندی باید کلیات تصاویر کلمات را به خوبی مدل کند لذا از پرفایل بالا و پایین، افکنش عمودی و تعداد عبور از پیکسل‌های پیش زمینه به پس زمینه در هر ستون برای

معرفی تصاویر استفاده شده است. تصاویر استفاده شده در این مرحله تصاویر اصلی کلمات هستند که نقاط آن جداسازی نشده است و تنها در مرحله پیش پردازش پس از اعمال فیلتر میانه به منظور کاهش نویز، به اندازه 200×200 با استفاده از روش نرمال سازی خطی، نرمال شده‌اند. به این منظور همانطور که در شکل ۴-۱۰ نشان داده شده است، قبل از نرمال سازی تصاویر، فضای خالی تصاویر حذف می‌شوند.



شکل ۴-۱۰: حذف سطر و ستون‌های خالی تصاویر، الف) تصویر ورودی، ب) تصور با حذف سطر و ستون

بردار ویژگی حاصل از تصاویر کلمات به منظور خوشه‌بندی، ۳۵۰ بعدی می‌باشد.

۴-۳-۲- ویژگی‌هایی که برای طبقه‌بندی و بازشناسی نهایی نمونه‌ها استخراج می‌گردد. در این مرحله از استخراج ویژگی به منظور ایجاد تناظر بیشتر میان بلوک‌های تصاویر یک کلاس و همچنین کاهش حساسیت به مکان اجزای تصویر که در کلمات دست‌نوشته تغییرات زیادی دارد، به جای تغییر در نحوی بلوک‌بندی، اجزای بدنه اصلی را به گونه‌ای نرمال می‌کنیم که تا حد امکان در بلوک‌های متناظر از تصاویر مختلف یک کلاس قرار گیرند. به این ترتیب که اجزای تصویر بدنه اصلی را با روش‌های مختلف نرمال سازی، به اندازه مشخص (64×64) نرمال کرده و به همان ترتیبی که در تصویر بدنه اصلی قرار گرفته‌اند در کنار یکدیگر قرار می‌دهیم. در طی این فرایند به علت تغییر اندازه اجزا، اندازه قلم در تصویر بدنه اصلی متغیر می‌شود از این رو برای یکسان سازی اندازه قلم، اجزای بدنه اصلی را نازک سازی می‌کنیم اما به علت غیر یکنواخت بودن لبه اجزا ناشی از تغییر اندازه آنها، نازک سازی مستقیم بر روی

بدنه اصلی باعث تولید شاخک و نويز می‌شود که در جهت‌های گرادیان که در مرحله بعد به عنوان ویژگی استخراج می‌شود بسیار تاثیر گذار است لذا برای جلوگیری از ایجاد شاخک و نويز در تصویر نازک شده ابتدا تصویر کلمه را با استفاده از فیلتر گوسی، یکنواخت می‌کنیم و سپس به کمک آستانه مناسب تصویر را دو سطحی می‌کنیم. مراحل انجام فرایند فوق در شکل ۴-۱۱ برای نمونه 'آبادان' نشان داده شده است.



شکل ۴-۱۱: نرمال سازی کلمه، الف) تصویر بدنه اصلی، ب) تصویر نرمال شده اجزاء، پ) تصویر نازک سازی شده پس از اعمال فیلتر گوسی

نمونه از اعوجاجی که به علت نازک سازی مستقیم بر روی بدنه کلمات به وجود آمده است در شکل ۴-۱۲ نشان داده شده است.



شکل ۴-۱۲: اعوجاج بهبود یافته توسط فیلتر گوسی

لازم به توضیح است که به علت ماهیت استخراج ویژگی تطبیقی در این مرحله که در آن برای بلوک‌بندی تصویر میزان پیکسل‌های پیش زمینه در نظر گرفته می‌شوند، از بین بردن سطر و ستون‌های خالی میان اجزاء، تغییری در ویژگی‌های استخراج شده به وجود نمی‌آورد.

حال از تصویر بدنه اصلی نرمال شده و تصویر نقاط و علائم ویژگی‌های مناسبی استخراج می‌کنیم. در این تحقیق تصویر بدنه اصلی به کمک بلوک‌بندی تطبیقی تجزیه شده و از هر بلوک اندازه‌گرادیان در ۳۲ جهت محاسبه می‌شود این ۳۲ ویژگی با استفاده از بردار وزن [۱ ۴ ۶ ۴ ۱] به ۱۶ ویژگی کاهش می‌یابند. در گام بعدی به منظور استخراج ویژگی از تصویر نقاط و علائم، برای کاهش حساسیت به تغییر مکان نقاط اندازه بلوک‌ها را نسبت به حالت قبل افزایش می‌دهیم تا در حد ممکن نقاط یکسان از تصاویر مختلف یک کلاس در بلوک‌های متناظری قرار گیرند. همچنین برای کاهش حساسیت به نحوی نوشتن و اندازه قلم آنها از بلوک بندی به روش یکنواخت استفاده می‌کنیم همانند حالت قبل از هر بلوک ۱۶ ویژگی استخراج می‌شود.

۴-۴ - طبقه‌بندی

نتایج حاصل از اعمال روش فوق بر روی پایگاه داده ذکر شده در ابتدا با استفاده از طبقه‌بند k نزدیکترین همسایگی به ازای تعداد همسایگی‌های مختلف و معیارهای فاصله متفاوت و سپس با استفاده از طبقه‌بند ماشین بردار پشتیبان در جداول زیر آورده شده است.

۴-۴-۱ - طبقه‌بند k - نزدیکترین همسایگی

در ابتدا نتایج حاصل از بلوک بندی تصویر با تعداد متفاوت بلوک را بررسی می‌کنیم. تصویر بدنه اصلی با استفاده از بلوک‌بندی تطبیقی و تصویر نقاط و علائم با استفاده از بلوک‌بندی ثابت تقسیم بندی شده‌اند. نتایج حاصل از این تغییرات در جدول ۴-۱ نشان داده شده است. در این آزمایش تصویر نقاط و علائم به ۹ بلوک 3×3 تقسیم بندی می‌شود و کارایی روش پیشنهادی برای تعداد بلوک‌های مختلف تصویر بدنه اصلی سنجیده می‌شود.

جدول ۱-۴: نتایج حاصل از تعداد بلوک‌های مختلف تصویر بدنه اصلی

| فاصله \cosine | فاصله بلوک شهری | فاصله اقلیدوسی | طول بردار ویژگی | تعداد بلوک‌ها |
|-----------------|-----------------|----------------|-----------------|---------------|
| ٪۳۶/۸ | ٪۳۷/۵۶ | ٪۳۱/۶۴ | ۲۸۸ | ۳×۳ |
| ٪۳۸/۷۱ | ٪۴۱/۹۶ | ٪۳۶/۶۵ | ۴۰۰ | ۴×۴ |
| ٪۷۶/۷۱ | ٪۷۹/۱۳ | ٪۷۸/۴۱ | ۴۶۴ | ۴×۵ |
| ٪۷۶/۲۷ | ٪۷۷/۷۵ | ٪۷۷/۰۷ | ۴۶۴ | ۵×۴ |
| ٪۷۹/۶۴ | ٪۷۷/۴۷ | ٪۷۸/۰۷ | ۵۴۴ | ۵×۵ |
| ٪۳۷/۸۹ | ٪۴۱/۷۹ | ٪۴۰/۷۹ | ۷۲۰ | ۶×۶ |

همانطور که در جدول ۱-۴ نیز مشهود است مطلوب ترین حالت در بلوک‌بندی تصاویر بدنه اصلی به تعداد فرد ۵×۵ حاصل شده است. حال در گام بعدی اعمال روش‌های مختلف نرمال کردن اجزا را بر روی تصویر بدنه اصلی بررسی می‌کنیم. تمامی نتایج زیر حاصل از بلوک بندی تصویر بدنه اصلی به ۲۵×۵ به صورت تطبیقی و بلوک‌بندی تصویر نقاط و علائم به ۹×۳ به صورت ثابت است.

جدول ۲-۴: نتایج حاصل از روش پیشنهادی با استفاده از روش نرمال کردن *Aspect ratio*

| فاصله \cosine | فاصله بلوک شهری | فاصله اقلیدوسی | |
|-----------------|-----------------|----------------|-------------------|
| ٪۷۹/۶۴ | ٪۷۷/۴۷ | ٪۷۸/۰۷ | نزدیکترین همسایگی |
| ٪۷۹/۶۳ | ٪۷۷/۹۵ | ٪۷۹/۶۵ | سه همسایگی اول |
| ٪۸۰/۷۳ | ٪۷۸/۷۳ | ٪۷۹/۵۱ | چهار همسایگی اول |
| ٪۷۹/۱۵ | ٪۷۷/۹۶ | ٪۷۸/۸۱ | پنج همسایگی اول |

جدول ۳-۴: نتایج حاصل از روش پیشنهادی با استفاده از روش نرمال کردن *Square root of aspect ratio*

| فاصله \cosine | فاصله بلوک شهری | فاصله اقلیدوسی | |
|-----------------|-----------------|----------------|-------------------|
| ٪۷۷/۷۱ | ٪۷۹/۳۳ | ٪۷۸/۵۴ | نزدیکترین همسایگی |
| ٪۷۹/۳۵ | ٪۷۹/۶۱ | ٪۷۸/۴۹ | سه همسایگی اول |
| ٪۷۹/۸۷ | ٪۷۹/۹۳ | ٪۷۹/۴۵ | چهار همسایگی اول |
| ٪۷۹/۱۵ | ٪۷۸/۸۹ | ٪۷۸/۴۰ | پنج همسایگی اول |

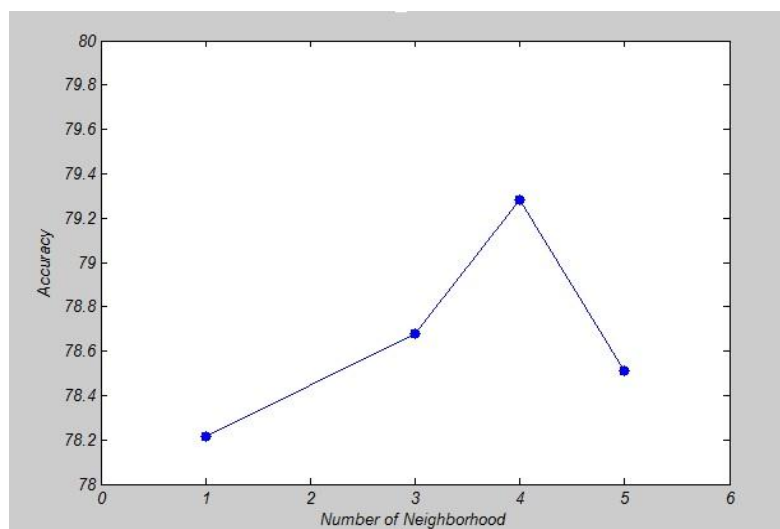
جدول ۴-۴: نتایج حاصل از روش پیشنهادی با استفاده از روش نرمال کردن *Cubic root of aspect ratio*

| فاصله <i>cosine</i> | فاصله بلوک شهری | فاصله اقلیدوسی | |
|---------------------|-----------------|----------------|-------------------|
| ٪.۷۷/۷۲ | ٪.۷۸/۱۳ | ٪.۷۷/۶۴ | نزدیکترین همسایگی |
| ٪.۷۸/۲۷ | ٪.۷۸/۷۳ | ٪.۷۸/۰۴ | سه همسایگی اول |
| <u>٪.۷۹/۲۳</u> | ٪.۷۸/۸۵ | ٪.۷۸/۴۵ | چهار همسایگی اول |
| ٪.۷۸/۷۱ | ٪.۷۸/۶۹ | ٪.۷۷/۸۷ | پنج همسایگی اول |

جدول ۴-۵: نتایج حاصل از روش پیشنهادی با استفاده از روش نرمال کردن *Sine of aspect ratio*

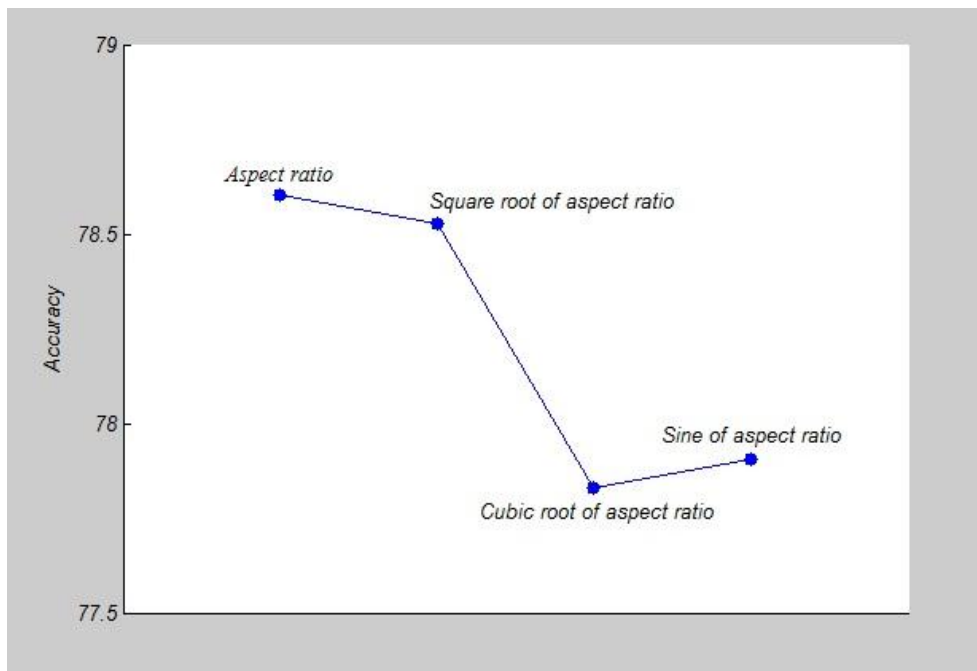
| فاصله <i>cosine</i> | فاصله بلوک شهری | فاصله اقلیدوسی | |
|---------------------|-----------------|----------------|-------------------|
| ٪.۷۷/۴۵ | ٪.۷۸/۱۷ | ٪.۷۸/۱۰ | نزدیکترین همسایگی |
| ٪.۷۸/۳۱ | ٪.۷۸/۸۱ | ٪.۷۸/۲۹ | سه همسایگی اول |
| ٪.۷۸/۶۳ | <u>٪.۷۹/۲۵</u> | ٪.۷۸/۷۷ | چهار همسایگی اول |
| ٪.۷۸/۰۳ | ٪.۷۸/۵۶ | ٪.۷۷/۹۱ | پنج همسایگی اول |

شکل ۴-۱۳ نمودار تغییرات حاصل از چهار روش نرمال سازی فوق به ازای مقادیر مختلف همسایگی را نشان می‌دهد به ازای هر همسایگی میانگین خروجی چهار روش نرمال سازی محاسبه شده است.



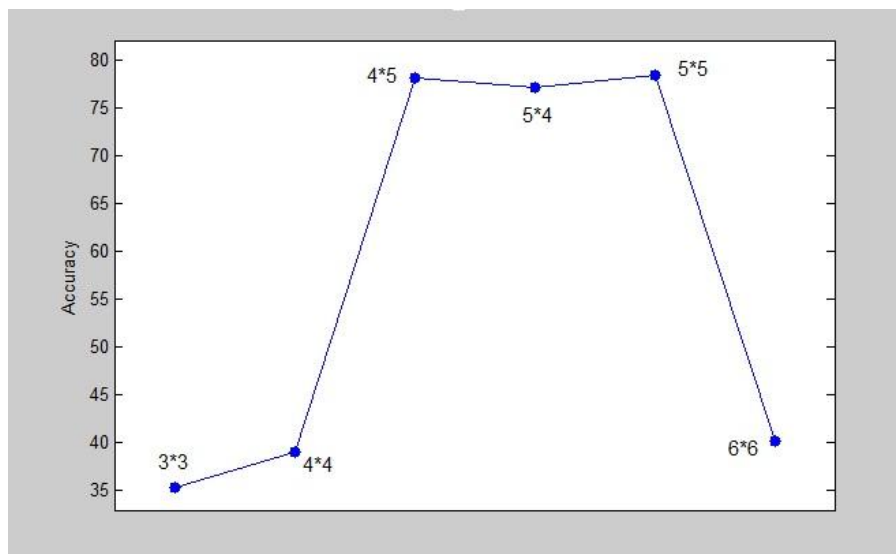
شکل ۴-۱۳: نمودار تغییرات دقت به ازای تعداد مختلف همسایگی در طبقه‌بند k نزدیکترین همسایگی

شکل ۴-۱۴ نمودار تغییرات دقت بازشناسی به ازای چهار روش نرمال سازی و معیار نزدیکترین همسایگی را نشان می‌دهد. برای هر یک از روش‌های نرمال سازی میانگین سه معیار فاصله محاسبه شده است.



شکل ۴-۱۴: نمودار تغییرات دقت به ازای روش‌های مختلف نرمال سازی در طبقه‌بند k نزدیکترین همسایگی

شکل ۴-۱۵ نمودار تغییرات بازشناسی نمونه‌های آزمایش نسبت به تقسیم بندی تصویر بدنه اصلی به تعداد بلوک‌های مختلف را نشان می‌دهد.



شکل ۴-۱۵: نمودار بازشناسی به ازای تغییرات تعداد بلوک‌ها در تصویر بدنه اصلی

به عنوان مقایسه نتایج حاصل از اعمال بلوک‌بندی تطبیقی بر روی تصویر کلمه در جدول ۴-۶ آورده شده است. [۲۹]

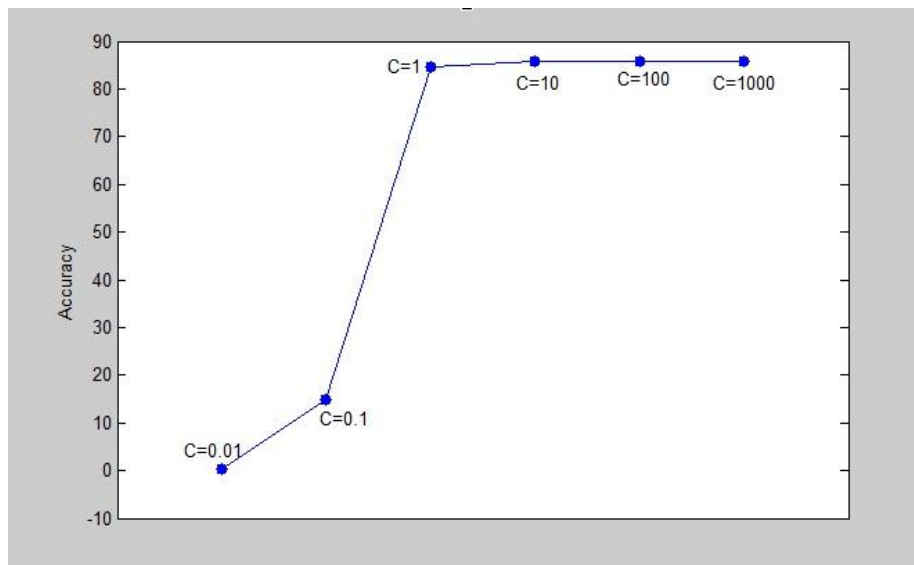
جدول ۴-۶: نتایج حاصل از بلوک بندی تطبیقی بر روی تصویر [۲۹]

| فاصله <i>cosine</i> | فاصله بلوک شهری | فاصله اقلیدوسی | |
|---------------------|-----------------|----------------|-------------------|
| ٪۷۳/۳۶ | ٪۷۲/۵۹ | ٪۷۰/۳ | نزدیکترین همسایگی |
| ٪۷۴/۸ | ٪۷۴/۰۸ | ٪۷۲/۱۵ | سه همسایگی اول |
| ٪۷۴/۷۶ | <u>٪۷۵/۴۴</u> | ٪۷۲/۴۰ | چهار همسایگی اول |
| ٪۷۵/۲۴ | ٪۷۵/۱۶ | ٪۷۳/۱۵ | پنج همسایگی اول |

۴-۴-۲- طبقه‌بند ماشین بردار پشتیبان

در طبقه بندی با استفاده از ماشین بردار پشتیبان، مسئله با استفاده از هسته‌های گوناگون و پارامترهای متفاوت حل شده و نتایج آن در زیر آورده شده‌اند. بدین منظور از کتابخانه *libsvm* استفاده شده است. نرم افزار *libsvm* دو حالت کلی برای حل مسئله دسته‌بندی دارد: *C-SVM* و *nu-SVM*. تفاوت این دو در

چگونگی بیان مسئله بهینه سازی و همچنین پارامترهای مورد تنظیم است. نمودار تغییرات متغیر C در بازشناسی نمونه‌های آزمایش در شکل ۴-۱۶ نشان داده شده است.



شکل ۴-۱۶: نمودار تغییرات دقت به ازای تغییرات پارامتر C

نتایج حاصل برای هسته‌های مختلف:

هسته خطی: این هسته هیچ پارامتری نداشته و لذا یک حالت اجرا دارد. طبقه‌بندی به کمک این هسته به دقت بازشناسی ۸۵/۴۷٪ بر روی نمونه‌های آزمایش و میزان بازشناسی ۱۰۰٪ بر روی نمونه‌های آموزش دست یافته است. این در حالی است که بازشناسی نمونه‌های آزمایش با کمک بلوک‌بندی تطبیقی بر روی تصویر کلی کلمه به کمک این هسته به دقت ۸۳/۳۶٪ دست یافته است.

هسته چند جمله‌ای: این هسته دارای دو پارامتر است. پارامتر اول درجه چند جمله‌ای است که مسئله به ازای چند جمله‌ای با درجه ۱، ۲، ۳ و ۴ حل شده و جوابهای زیر حاصل شده است.

جدول ۴-۷: تاثیر درجه هسته چند جمله‌ای در دقت بازشناسی ماشین بردار پشتیبان

| درجه | ۱ | ۲ | ۳ | ۴ |
|---|--------|---------------|--------|--------|
| دقت شناسایی داده‌های آزمایش | ٪۸۴/۵۹ | <u>٪۸۵/۵۱</u> | ٪۸۴/۴۷ | ٪۸۲/۶۷ |
| دقت شناسایی داده‌های آموزش | ٪۹۷/۷۵ | ٪۹۹/۹۱ | ٪۱۰۰ | ٪۱۰۰ |
| دقت شناسایی بلوک بندی تطبیقی بر روی تصویر کلی کلمه [۲۹] | ٪۸۳/۲۸ | ٪۸۳/۰۸ | ٪۸۲/۳۲ | ٪۷۹/۹۲ |

برای پارامتر دوم هسته چند جمله‌ای که عدد ثابت جمع شوند است از سه مقدار ۰، ۱- و ۱ استفاده شده است.

جدول ۴-۸: تاثیر مقدار ثابت جمع شونده هسته چند جمله‌ای در دقت بازشناسی ماشین بردار پشتیبان

| مقدار ثابت | -۱ | ۰ | ۱ |
|---|--------|--------|---------------|
| دقت شناسایی داده‌های آزمایش | ٪۸۰/۳۹ | ٪۸۴/۴۷ | <u>٪۸۵/۲۷</u> |
| دقت شناسایی داده‌های آموزش | ٪۹۸/۹۳ | ٪۱۰۰ | ٪۱۰۰ |
| دقت شناسایی بلوک بندی تطبیقی بر روی تصویر کلی کلمه [۲۹] | ٪۸۱/۱۲ | ٪۸۲/۳۲ | ٪۸۲/۸۸ |

هسته *RBF*: این هسته دارای یک پارامتر گاما است که نتایج حاصل از این هسته با مقادیر گامای مختلف در جدول ۴-۹ آورده شده است.

جدول ۴-۹: تاثیر پارامتر گاما، هسته *RBF* در دقت بازشناسی ماشین بردار پشتیبان

| گاما | ابعاد بردار ویژگی/۱ | γ^2 (ابعاد بردار ویژگی)/۱ | ۱ |
|---|---------------------|----------------------------------|--------|
| دقت شناسایی داده‌های آزمایش | <u>٪۸۴/۵۱</u> | ٪۷/۶۱ | ٪۱۲/۳۶ |
| دقت شناسایی داده‌های آموزش | ٪۹۹/۴۲ | ٪۱۰۰ | ٪۱۰۰ |
| دقت شناسایی بلوک بندی تطبیقی بر روی تصویر کلی کلمه [۲۹] | ٪۷۷/۶۴ | ٪۱۱/۳۶ | ٪۱۱/۳۶ |

۴-۵ - خوشه بندی

در روش‌های بازشناسی با استفاده از کاهش فرهنگ لغت، در مرحله آموزش ویژگی‌های مناسبی از تصاویر کلمه استخراج می‌شوند و با استفاده از روش‌های خوشه‌یابی، تصاویر پایگاه داده به چندین خوشه تقسیم

می‌شوند پس از آن در مرحله بازشناسی، در گام اول بر اساس ویژگی‌های استفاده شده در خوشه‌یابی، نزدیکترین خوشه‌ها به کلمه ورودی به دست می‌آیند و در گام بعدی ویژگی‌های مناسبتری که تصاویر کلمات را بهتر توصیف می‌کنند از ورودی استخراج شده و با ویژگی‌های متناظر از تصاویر خوشه‌های باقی مانده به منظور بازشناسی کلمه مقایسه می‌شوند.

مرحله خوشه‌یابی یکی از بخشهای مهم در فرآیند بازشناسی است که در آن تصاویر کلمات به تعدادی خوشه تقسیم می‌شوند. برای تعیین خوشه‌ها از الگوریتم *ISODATA* استفاده شده است.

در مرحله بازشناسی ابتدا خوشه‌های نزدیک به داده ورودی بدست می‌آید و سپس تصاویر کلمات ورودی به دو تصویر نقاط و علائم و تصویر بدنه اصلی تقسیم می‌شوند و از هر کدام ویژگی‌های گرادینان مانند حالت قبل استخراج می‌شود و از میان تصاویر خوشه‌های مورد نظر بازشناسی نهایی کلمات صورت می‌گیرد.

مراکز خوشه‌ها به عنوان معرف هر خوشه با بردار ویژگی تصویر ورودی مقایسه می‌شود و تعدادی از خوشه‌های نزدیک به نمونه ورودی انتخاب می‌شوند. کلاس‌های مرتبط با اسامی شهرهایی که حداقل ۲ نمونه از آنها در خوشه‌های انتخاب شده وجود داشته باشند به عنوان کاندید به منظور بازشناسی نهایی کلمه ورودی انتخاب می‌شوند. بر این اساس دقت بازشناسی کلمه و میزان کاهش کلاس‌ها، به انتخاب تعداد خوشه‌های نزدیک به نمونه ورودی وابسته است به این ترتیب که اگر تعداد خوشه‌های انتخاب شده را افزایش دهیم میزان کاهش فرهنگ لغت، کاهش می‌یابد در حالی که دقت بازشناسی افزایش می‌یابد. به عنوان فاصله در خوشه‌بندی از معیار فاصله بلوک شهری استفاده شده است و برای محاسبه مراکز اولیه خوشه‌ها، نمونه‌ها توسط روش سلسله مراتبی به ۵ خوشه تقسیم بندی شده‌اند. مشخصات خوشه‌های خروجی در جدول ۴-۱۰ بیان شده است.

جدول ۴-۱۰: مشخصات خوشه‌های خروجی

| تعداد خوشه‌های اولیه | تعداد خوشه‌های نهایی | میانگین تعداد تصاویر در هر خوشه | میانگین تعداد کلاس‌ها در هر خوشه | میانگین حداقل تعداد تصاویر یک کلاس در خوشه‌ها | میانگین حداکثر تعداد تصاویر یک کلاس در خوشه‌ها |
|----------------------|----------------------|---------------------------------|----------------------------------|---|--|
| ۵ | ۷۷ | ۱۶۹ | ۶۵ | ۱ | ۱۹ |

جدول ۴-۱۱ میزان کاهش فرهنگ لغت و دقت آن را به صورت تابعی از تعداد خوشه‌های نزدیک انتخاب شده به نمونه ورودی بیان می‌کند.

جدول ۴-۱۱: دقت و میزان کاهش فرهنگ لغت

| تعداد خوشه‌های نزدیک | میزان کاهش فرهنگ لغت | میزان دقت کاهش فرهنگ لغت | کارایی سیستم |
|----------------------|----------------------|--------------------------|--------------|
| ۱ | ٪۹۲/۷۱ | ٪۶۵/۳۱ | ٪۶۰/۵۴ |
| ۲ | ٪۸۶/۹۸ | ٪۸۲/۰۳ | ٪۷۱/۳۴ |
| ۳ | ٪۸۱/۴۰ | ٪۸۹/۷۲ | ٪۷۳/۰۳ |
| ۴ | ٪۷۸/۲۶ | ٪۹۳/۸ | ٪۷۳/۴ |
| ۵ | ٪۷۴/۶۲ | ٪۹۵/۹۶ | ٪۷۱/۶ |
| ۱۰ | ٪۶۲/۱ | ٪۹۹/۲۸ | ٪۶۱/۶۵ |

همانطور که در جدول ۴-۱۱ مشخص شده است مطلوب ترین حالت در انتخاب ۴ خوشه اول از همسایگی نمونه ورودی رخ می‌دهد که بر این اساس میزان بازشناسی نهایی بر اساس کاهش فرهنگ لغت در جدول ۴-۱۲ به ازای ۴ خوشه نزدیک نمونه ورودی آورده شده است.

جدول ۴-۱۲: طبقه‌بندی به کمک کاهش فرهنگ لغت

| فاصله <i>cosine</i> | فاصله بلوک شهری | فاصله اقلیدوسی | |
|---------------------|-----------------|----------------|-------------------|
| ٪۷۸/۷۵ | ٪۷۸/۲۹ | ٪۷۷/۴۹ | نزدیکترین همسایگی |
| ٪۷۹/۸۴ | ٪۷۸/۶۸ | ٪۷۹/۷۶ | سه همسایگی اول |
| ٪۸۱/۳۲ | ٪۷۹/۴۵ | ٪۸۰/۰۲ | چهار همسایگی اول |
| ٪۸۰/۱۵ | ٪۷۹/۱۴ | ٪۷۷/۶۸ | پنج همسایگی اول |

این در حالی است که میانگین بهبود سرعت روش برای محاسبه خروجی در جدول ۴-۱۳ آورده شده است. ستون سمت چپ برای هر معیار فاصله، میانگین زمان برای محاسبه هر خروجی بدون اعمال خوشه‌بندی است.

جدول ۴-۱۳: بهبود سرعت با استفاده از روشهای کاهش فرهنگ لغت

| فاصله <i>cosine</i> | | فاصله بلوک شهری | | فاصله اقلیدوسی | | |
|---------------------|--------|-----------------|--------|----------------|--------|-------------------|
| ۰/۳۲۹۲ | ۰/۱۹۹۸ | ۱/۱۱۸۲ | ۰/۰۷۹۱ | ۰/۱۱۱ | ۰/۰۷۷۱ | نزدیکترین همسایگی |
| ۰/۳۳۸۸ | ۰/۲۰۰۳ | ۱/۱۰۶۵ | ۰/۰۷۹۴ | ۰/۱۲۰۸ | ۰/۰۷۷۳ | سه همسایگی اول |
| ۰/۳۳۶۲ | ۰/۱۹۹۳ | ۱/۱۹۵۹ | ۰/۰۸۰۳ | ۰/۱۱۹۵ | ۰/۰۷۷۳ | چهار همسایگی اول |
| ۰/۳۳۷ | ۰/۲۰۰۲ | ۱/۱۵۷۵ | ۰/۰۷۹۴ | ۰/۱۲۱۷ | ۰/۰۷۷۳ | پنج همسایگی اول |

فصل ۵

نتیجه‌گیری و پیشنهادات

۵-۱- مقدمه

در این فصل مراحل تحقیق که شامل پیش پردازش، استخراج ویژگی و طبقه‌بندی است به ترتیب به طور مختصر مرور شده و نتیجه‌گیری در مورد کارهای انجام شده و تاثیرات تغییر پارامترهای طبقه‌بندی بکار رفته، بیان می‌شود و سپس در مورد پیشنهادات برای کارهای آتی بحث می‌گردد. بعضی از روشهای انجام شده نیز که به نتایج مطلوب دست نیافته‌اند، توضیح داده می‌شوند.

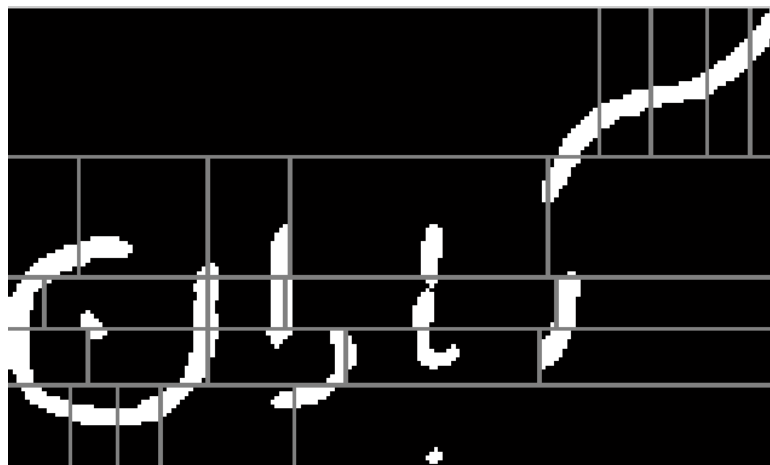
۵-۲- نتیجه‌گیری

سیستم‌های بازشناسی متن با توجه به روش بازشناسی به دو دسته کلی زیر تقسیم کرد:

- روش‌های تجزیه‌ای
- روش‌های کلی‌نگر

در این تحقیق با استفاده از روش‌های کلی‌نگر سعی در شناسایی کلمات دست‌نوشته فارسی صورت گرفته است. در این راستا به منظور توصیف کلمات از ویژگی‌های گرادیان روشنایی که به صورت تطبیقی از تصاویر کلمات استخراج می‌گردد، استفاده شده است. در روش تطبیقی برای ایجاد تشابه بیشتر میان بلوک‌های یکسان از تصاویر مختلف یک کلاس، به جای بلوک‌بندی تصویر به تعدادی بلوک‌ها با اندازه برابر، بلوک‌بندی بر اساس میزان پیکسل‌های تصویر صورت می‌گردد که در ابتدا بلوک‌بندی در راستای سطرها انجام می‌شود. بدین نحو که تصویر به تعداد مشخصی نوار افقی تقسیم می‌شود بطوریکه تعداد پیکسل‌های یکسانی در هر نوار قرار گیرد. سپس همین فرایند در راستای ستون‌ها انجام شده و در نهایت از تلاقی نوارهای حاصل از این دو بلوک‌بندی تقسیم‌بندی نهایی تصویر صورت می‌گیرد. در این تحقیق سعی بر ایجاد تشابه بیشتر میان بلوک‌های تصویر با اعمال تغییراتی بر روی نحوه‌ی بلوک‌بندی تطبیقی صورت گرفت. این تغییرات به این نحو اعمال گردید که به جای بلوک‌بندی تصویر در دو راستای سطری و ستونی که در هر بلوک سطری و ستونی میزان پیکسل‌های برابری وجود داشته و از تلاقی آنها بلوک‌بندی نهایی

صورت پذیرد، پس از بلوک‌بندی در راستای سطری (و یا ستونی)، بلوک‌های سطری (و یا ستونی) حاصل به طور مجزا به چند بلوک ستونی (و یا سطری) به نحوی تقسیم شوند که در هر بلوک حاصل میزان برابری از پیکسل‌ها قرار گیرد. نمونه‌ای از این بلوک‌بندی در شکل ۵-۱ نشان داده شده است. همانطور که در این تصویر نیز مشخص است در بلوک‌بندی با این روش در تمامی بلوک‌ها قسمتی از کلمه وجود دارد. نتایج این نحوه‌ی بلوک‌بندی هر چند نسبت به حالت بلوک‌بندی ثابت بهبود دارد اما به نتایج بلوک‌بندی تطبیقی دست نیافته است از این رو در این تحقیق بلوک‌بندی به صورت تطبیقی صورت گرفته است.



شکل ۵-۱: بلوک‌بندی تطبیقی پیشنهاد شده

بلوک‌بندی تطبیقی برای ایجاد تشابه میان بلوک‌های تصویر پیشنهاد شده است. لذا در گام اول از این تحقیق نقاط و علائم که در کلمات دست‌نوشته فارسی با اندازه و مکان متفاوت نگارش می‌گردد عاملی مزاحم در ایجاد تشابه میان بلوک‌های استخراج شده در نظر گرفته شده و سعی در جدا کردن آنها از تصویر اصلی کلمه شده است. برای جدا کردن نقاط و علائم نیز از معیار فاصله تا خط کرسی استفاده شده است. برای تخمین مکان خط کرسی از روش هیستوگرام افقی با اعمال تغییراتی به منظور بهبود تخمین آن استفاده شده است. پس از جداسازی نقاط و علائم از تصویر، هر تصویر به دو تصویر نقاط و علائم و تصویر بدنه اصلی تجزیه می‌شود. حال به منظور استخراج ویژگی از تصویر بدنه اصلی از بلوک‌بندی به

صورت تطبیقی و از تصویر نقاط و علائم از بلوک‌بندی ثابت استفاده شده است. علت استفاده از بلوک‌بندی ثابت در تصاویر مربوط به نقاط و علائم نیز تنوع مکان و اندازه نگارش آنهاست که برای جلوگیری از عوامل یاد شده در کارایی روش، تعداد تقسیم‌بندی این تصاویر به بلوک‌های ثابت از تعداد بلوک‌های بدنه اصلی کمتر در نظر گرفته شده است به این امید که نقاط یکسان از تصاویر مختلف یک کلاس در بلوک‌های متناظر قرار بگیرند. نتایج حاصل از اعمال روش ارائه شده با استفاده از دو طبقه‌بند k نزدیکترین همسایگی و ماشین بردار پشتیبان نشان دهنده بهبود کارایی نسبت به اعمال بلوک‌بندی تطبیقی بر روی کل تصویر اصلی بدون جداسازی است. نتایج حاصل از تغییر تعداد بلوک‌های تصویر بدنه اصلی نشان دهنده این است که در حالتی که تصویر بدنه اصلی به ۵ بلوک، هم در راستای سطرها و هم در راستای ستون‌ها تقسیم می‌شود کارایی روش نسبت به حالت‌های دیگر افزایش می‌یابد. پایگاه داده استفاده شده جهت بررسی کارایی روش ارائه شده نیز شامل ۱۳۰۰۰ نمونه آموزش و ۲۵۰۰ نمونه آزمایش از نام‌های ۵۰۳ شهر ایران است.

در استفاده از طبقه‌بند k نزدیکترین همسایگی افزایش تعداد همسایگی‌های نمونه ورودی به منظور بازشناسی کلمه باعث بهبود نتایج حاصل می‌شود این افزایش تعداد همسایگی تا چهار همسایگی اول به بهینه‌ترین حالت خود دست می‌یابد. سپس استفاده از معیارهای فاصله متفاوت نشان دهنده بهبود دقت بازشناسی به ازای معیار فاصله \cosine است. در طبقه‌بندی به کمک ماشین بردار پشتیبان نیز تاثیر تغییرات پارامترهای مدل و استفاده از توابع هسته مختلف در کارایی روش پیشنهادی مورد بررسی قرار گرفته شده است که بر این اساس بهینه‌ترین حالت در هنگامی است که از هسته چند جمله‌ای با درجه دو و مقدار ثابت جمع شونده یک، استفاده شده است. در حالت کلی استفاده از ماشین بردار پشتیبان باعث بهبود نتایج حاصل نسبت به طبقه‌بند k نزدیکترین همسایگی شده است. طبقه‌بند ماشین بردار پشتیبان در اکثر مدل‌های تعریف شده، به نتایج بسیار خوبی در یادگیری داده‌های آموزش دست یافته

است اما عمومیت آن برای داده‌های آزمایش تا حدی کاهش می‌یابد که این خود می‌تواند به علت تعدد کلاس‌ها باشد.

در انتها با استفاده از روش‌های خوشه‌بندی تصاویر به تعدادی خوشه تقسیم می‌شوند که این امر موجب بهبود در زمان و نتایج حاصل از طبقه‌بند k نزدیکترین همسایگی نسبت به حالتی که از تمامی تصاویر آموزش برای شناسایی نمونه ورودی استفاده می‌شود، گردیده است.

۵-۳- پیشنهادات

در تحقیق انجام شده به منظور جداسازی نقاط و علائم از تصویر کلمات از معیار فاصله تا خط کرسی استفاده شده است که برای بهبود تخمین مکان خط کرسی تغییراتی بر روش هیستوگرام افقی اعمال گردید. خطا در تخمین مکان خط کرسی موجب ایجاد خطا در تعریف نقاط و علائم می‌گردد لذا ارائه روشی کارا تر برای بهبود تخمین مکان خط کرسی و یا تشخیص نقاط و علائم به روش استخراج ویژگی از اجزای جدای تصویر و طبقه‌بندی آنها به دو گروه علائم و اجزای کلمه می‌تواند در بهبود نتایج حاصل مفید واقع گردد.

همچنین در زمان نرمال سازی اجزای بدنه اصلی می‌توان از اطلاعات مربوط به ترتیب قرار گرفتن اجزا کنار هم نیز برای تصحیح نتایج بازشناسی کمک گرفت.

در مرحله نرمال سازی اجزای بدنه اصلی می‌توان از محدوده ارتفاع اجزا نیز در انتخاب اندازه نرمال شده خروجی استفاده کرد به این معنی که بین اجزا با محدوده ارتفاعات مختلف تفاوت در اندازه، ایجاد کرد در این تحقیق تلاشهایی برای اعمال این روش صورت گرفته است اما در نتایج بهبودی حاصل نگردید.

- [۱] ر. عزمی، احسان ا. کبیر، ک. بدیع، "بازشناسی متون چاپی فارسی"، رساله دکتری، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، ۱۳۷۸.
- [2] Amin, A. and Masini, "Machine recognition of multifont printed Arabic texts", *Proc. 8th Int. Conf. on Patt. Recog, Paris, pp. 392-395, 1986.*
- [3] Margner, V., "START-A system for the recognition of Arabic Printed texts", *Proc. 11th Int. Conf. on Pattern Recognition , pp. 561-564, 1992.*
- [۴] رفیعی، شعبانعلی و کبیر، احسان اله، "شکستن کلمات تایپ شده فارسی به حروف در رسم الخطهای مختلف"، مجموعه مقالات سومین کنفرانس الکترونیک، ص ۹۸-۱۰۴، دانشگاه شیراز، مهر ۱۳۷۴.
- [5] Parhami, B. and M. Taraghi, " Automatic recognition of printed Farsi texts", *Pattern Recognition, Vol. 14, pp. 395-403, 1981.*
- [۶] فهیمی، حمید و حاتم، احمد، "ارائه یک روش ساختاری جهت تشخیص حروف در متون تایپ شده فارسی"، مجموعه مقالات سومین کنفرانس مهندسی برق ایران (کامپیوتر)، ص ۱۹۰-۱۹۷، دانشگاه علم و صنعت ایران، تهران، اردیبهشت ۱۳۷۴.
- [7] El-Khaly, F. and M. Sid-Ahmad, "Machine recognition of optically captured machine printed Arabic text", *Pattern Recognition, Vol. 23, pp1207-1214, 1990.*
- [8] R. Azmi, E. kabir, "A New Segmentation Technique for Omnifont Farsi texts", *Pattern Recognition Letters 22, pp.97-104, 2001.*
- [9] Broumandnia, J. Shanbehzadeh, M. Rezakhah Varnoosfaderani, " Persian/arabic handwritten word recognition using M-band packet wavelet transform ", *Image and Vision Computing 26,pp 829-842 , 2008.*
- [10] Malik Waqas Sagheer, Chun Lei He, Nicola Nobile, Ching Y. Suen, " Holistic Urdu Handwritten Word Recognition Using Support Vector Machine", *pp1977-1980, 2010IEEE.*
- [11] Jawad H Alkhateeb , Fouad Khelifi , Jianmin Jiang , Stan S Ipson "A New Approach For Off-Line Handwritten Arabic Word Recognition Using Knn Classifier" *IEEE International Conference on signal and Image Processing Applications, pp.191-194, 2009*
- [12] Jawad H. AlKhateeb, Jianmin Jiang, Jinchang Ren, Fouad Khelifi and Stan S. Ipson, " Multiclass Classification of Unconstrained Handwritten Arabic Words Using Machine Learning Approaches " *The Open Signal Processing Journal , Vol. 2, pp. 21-28, 2009.*
- [13] M. K. Hu, "Visual Pattern Recognition by Moment Invariants", *IRE Transaction on Information Theory, Vol. 49, pp. 179-187, Feb. 1962*
- [۱۴] خسروی، خ، کبیر، ا.، "ارزیابی روش‌های بازشناسی متون فارسی بر مبنای شکل کلی زیر-کلمات" نشریه مهندسی برق و کامپیوتر

- [15] Zaher AL Aghbari , Salama Brook "HAH manuscripts: A holistic paradigm for classifying and retrieving historical Arabic handwritten documents", *Expert Systems with Applications*, pp.10942-10951, 2009 ELSEVIER
- [16] Saeed Mozaffari, Karim Faez, Volker Ma`rgner, Haikal El-Abed , " Lexicon reduction using dots for off-line Farsi/Arabic handwritten word recognition " *Pattern Recognition Letters* 29, pp.724-734, 2008
- [17] Zahra bahmani, Fatemh Alamdar, Reza Azmi, Saman Haratizadeh , " Off-Iine Arabic/Farsi Handwritten Word Recognition Using RBF Neural Network and Genetic algorithm " , pp.352-357, 2010 IEEE
- [۱۸] ا. ابراهیمی، ا. کبیر، (۱۳۸۳). یک روش دو مرحله‌ای برای بازشناسی زیر-کلمات چاپی. نشریه مهندسی برق و مهندسی کامپیوتر ایران سال ۲-شماره ۲: ۵۷-۶۲
- [19] Alireza Alaei, P.Nagabhushan, Umapada Pal "A New Tow-stage Scheme for the Recognition of Persian Handwritten Character" *12th International Conference on Frontiers in handwriting Recognition* ,pp. 130-135, 2010.
- [20] Majid Ziaratban and Karim Faez "Detection and Compensation of Undesirable Discontinuities within the Farsi/Arabic Subwords", *The International Arab Journal of Information Technology*, Vol. 8, No. 3, pp. 293-301, July 2011.
- [21] Gonzalez, Rafael C. Woods, Richard E. "Digital Image Processing" Third Edition, Pearson Prentice Hall, 2008.
- [22] J. Kanai and A.D. Bagdanov" Projection Profile Based Skew Estimation Algorithm for Jbig Comperssed Image". *IJDAR*, 1(1):pp.43-51, 1998.
- [23] B. Yu and A.K. Jain" A robust and fast Skew detection Algorithm for Generic Documents" *Pattern Recognition*, 29(10):pp.1599-1629, October 1996.
- [24] A.Hashizume, P.S. Yeh, and A. Rosenfeld "A Method of detecting the Orientation of Aligned Components" *Pattern Recognition Letters*, 4:pp.125-132, 1986.
- [25] N. Otsu. A threshold selection method from gray-scale histogram. *IEEE Transactions on System, Man, and Cybernetics*.9, pp.62-66, 1979.
- [26] J. Bernsen. Dynamic Thersholding of gray level images. In *Processdings of 8th International Conference on Pattern Recognition*. Paris, France, Vol.2, pp.1251-1255, 1986.
- [27] Mohamed Cheriet, Nawwaf Kharma , Cheng -lin liu , ching Y.suen , " character recognition system, A Guide for Students and Practioners " , wiley 2007
- [28] T. Y. Zhang and C. Y. Suen " A fast parallel algorithm for thinning digital pattern " *Communication of the ACM*. 27(3), pp. 236-239, 1984.

[۲۹] الهام بایسته و علیرضا احمدی فرد " بازشناسی کلمات دست نوشته فارسی با استفاده از ویژگی

گردادین تطبیقی و طبقه بند ماشین بردار پشتبان چند کلاسه " پذیرفته شده در هفدهمین کنفرانس

ملی انجمن کامپیوتر ایران ۱۳۹۰

- [30] Afshin Ebrahimi, Ehsanollah Kabir "A prictorial dictionary for printed Farsi subwords". *Pattern Recognition Letters* 29, pp.656-663, 2008.
- [31] Toni M. Rath and R. Manmatha, " Word Image Matching Using Dynamic Time Warping", *Proceeding of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.

- [32] Andrew R. Webb "Statistical Pattern Recognition" John Wiley & Sons, Chapter 10, pp. 361-406, 2002.
- [33] Ball, G. H. And D. I. Hall "Some Fundamentation Concepts and Synthesis Procedures for Pattern Recognition Preprocessors" Proc. Int'l Conf. Microwaves, Circuit Theory, and Information Theory, Tokyo, Japan, pp. 281-297, 1964.
- [34] V. Vapnik, "The nature of statistical learning theory", 2nd edition, Springer, 1999
- [35] C. Cortes, V. Vapnik, "Support-Vector Networks", Machine Learning, Vol. 20, pp. 273-297, 1995.

[۳۶] احمد کشاورز و حسن قاسمیان یزدی، "یک الگوریتم سریع مبتنی بر ماشین بردار پشتیبان برای طبقه‌بندی تصاویر ابر طیفی با استفاده از همبستگی مکانی"، نشریه مهندسی برق و مهندسی کامپیوتر ایران، ۱۳۸۴.

Abstract

Automatic recognition of text is a subset of image processing, which widely used in digital library, reading e-mail, reading check bank, reading forms and questionnaires . manuscript text recognition can be performed in two general states: analytical method and holistic method .

In systems based on analytical method, words are divided into their components, sub words and letters which are used to identify word; but in holistic method features are extracted directly from the whole word image and by comparing them with the data dictionary (lexicon), the word which is similar to the input is considered as the output.

In Preprocessing stage for estimating the location of baseline, the changes are performed in horizontal histogram method. This process is needed for two reasons.

First to isolate the main body of the word and second for normalizing the main components of word. This lead to decompose a word into two parts main body image and diacritics & dots image. In order to describe the image of word After normalizing each component ,we propose to divide the image of main body using adaptive blocking. For second image part (diacritics and dots) we propose using constant blocking. Each block in each of these two images are described using image gradient in 32 directions. The result of experiments shows that the recognition rate of proposed method is improved using k nearest neighbor and support vector machine classifiers. However, because of preprocessing the proposed method require more time than previous methods. The clustering of samples in order to lexicon reduction will improve the results so Horizontal histogram and profile feature with using isodata algorithm are used to cluster samples .

Keywords: Farsi handwriting word recognition; gradient feature; adaptive blocking; baseline estimation



Shahrood University of Technology

Faculty of Electrical and Robatic

***Offline Farsi Handwritten Word Recognition Using Adaptive Image
Partitioning***

Mohammad Alipour Saraji

Supervisor:

Dr. Alireza Ahmadifard

Assistant:

Dr. Omidreza Marouzi

***FOR THE DEGREE OF
MASTER OF SCIENCE***

January 2013