مرکز آموزش الکترونيکي

ارائه يک روش نوين ساختاري براي شناسايي نويسه‌هاي دستنويس تنها فارسي

پايان‌نامه براي دريافت درجه کارشناسي‌ارشد در رشته فناوري اطلاعات و ارتباطات

گرايش مخابرات امن

بابک صيرفي

استاد راهنما: دکتر محسن سرياني

استاد مشاور: دکتر احسان الله کبير

فروردين 1391

تقديم به تمامي كساني كه علم را تنها راه براي
كشف حقيقت مي‌دانند

سپاسگزاري

با سپاس و قدرداني فراوان از جناب آقاي دكتر سرياني، استاد ارجمند كه از راهنمايي‌هاي ايشان در انجام اين تحقيق بهره‌مند شدم.

همچنين از آقايان دكتر فتحي، دكتر كبير، دكتر فائز كه در انجام و نشست دفاع از اين پروژه من را ياري كردند، سپاسگزاري مي‌كنم.

بر خود لازم مي‌دانم از آقاي كسري نوايي و خانم محدثه الماسي كه در انجام اين پايان‌نامه از همكاري و حمايت‌هاي بي‌دريغشان برخوردار بودم، صميمانه تشكر نمايم.

تاييديه‌ي هيات داوران جلسه‌ي دفاع از پايان‌نامه

نام دانشكده: مركز آموزش الكترونيكي

نام دانشجو: بابک صيرفي

عنوان پايان‌نامه: ارائه یک روش نوین ساختاري براي شناسايي نويسه‌هاي دستنويس تنهاي فارسي

تاريخ دفاع:29/01/1391

رشته: فناوري اطلاعات و ارتباطات

گرايش: مخابرات امن

رديف	سمت	نام و نام‌خانوادگي	مرتبه دانشگاهي	دانشگاه	امضا
1	استاد راهنما	دكتر محسن سرياني	استاديار	علم و صنعت ايران	
2	داور داخلي	دكتر محمود فتحي	دانشيار	علم و صنعت ايران	
3	داور خارجي	دكتر احسان‌الله كبير	استاد	تربيت مدرس	

باسمه تعالي

اينجانب بابک صيرفي به شماره دانشجويي 87211306 دانشجوي رشته فناوري اطلاعات و ارتباطات گرايش مخابرات امن مقطع تحصيلي کارشناسي ارشد، تاييد مي‌نمايم که کليه نتايج اين پايان‌نامه حاصل کار اينجانب و بدون هرگونه دخل و تصرف است و موارد نسخه برداري شده از آثار ديگران را با ذکر کامل مشخصات منبع ذکر کرده‌ام. در صورت اثبات خلاف مندرجات فوق، به تشخيص دانشگاه مطابق با ضوابط و مقررات حاکم با اينجانب رفتار خواهد شد. و حق هرگونه اعتراض در خصوص احقاق حقوق مکتسب و تشخيص و تعيين تخلف و مجازات را از خويش سلب مي‌نمايم. در ضمن، مسئوليت هر گونه پاسخگويي به اشخاص، اعم از حقيقي و حقوقي و مراجع ذي‌صلاح، به عهده اينجانب خواهد بود و دانشگاه هيچ‌گونه مسئوليتي در اين خصوص نخواهد داشت.

نام و نام‌خانوادگي:

امضا و تاريخ:

مجوز بهره‌برداري از پايان‌نامه

بهره‌برداري از اين پايان‌نامه در چهارچوب مقررات كتابخانه و با توجه به محدوديتي كه توسط استاد راهنما به شرح زير تعيين مي شود، بلامانع است:

☐ بهره‌برداري از اين پايان‌نامه براي همگان بلامانع است.

☐ بهره‌برداري از اين پايان‌نامه با اخذ مجوز از استاد راهنما، بلامانع است.

☐ بهره‌برداري از اين پايان‌نامه تا تاريخ ممنوع است.

استاد راهنما:

تاريخ:

امضا:

چكيده

بازشناسي حروف و ارقام دستنويس همواره يكي از موضوعات مورد علاقه براي تحقيق بوده است. در زمينه بازشناسي حروف و ارقام دستنويس عربي و فارسي نيز، كارهاي زيادي صورت گرفته است. شناسايي حروف براي انسان كار آساني است اما اين كار براي ماشين چندان ساده نيست زيرا برخي از حروف به شدت مشابه هستند. تا كنون تحقيقات فراواني در زمينه سيستمهاي تشخيص حروف و ارقام دستنويس فارسي انجام گرفته است اما اين سيستمها تا رسيدن به نتايج ايده‌آل راه زيادي در پيش‌رو دارند. از كاربردهاي اين بازشناسي مي‌توان فرمهاي ورود اطلاعات وزراتخانه‌ها و سازمانهاي بزرگ كه متشكل از اعداد و حروف تنها فارسي مي‌باشند و نياز به شناسايي و ذخيره‌سازي در پايگاه‌هاي اطلاعاتي را دارند نام برد. هدف از اين پروژه شناسايي نويسه‌هاي دستنويس فارسي مي‌باشد كه به صورت نويسه‌هاي اسكن و بريده شده به عنوان ورودي به سيستم داده مي‌شوند. در الگوريتم پيشنهادي از تركيب دو روش ساختاري و شبكه عصبي و استفاده از منطق فازي براي شناسايي الگو استفاده شده است. به طور خلاصه پس از دريافت تصوير و انجام عمليات پيش‌پردازش بر روي آن، نويسه از لحاظ همبند بودن مورد بررسي قرار مي‌گيرد و در صورتيكه نويسه از چند جز مجزا تشكيل شده باشد، مشخص مي‌شود. سپس اجزاي متصل نويسه در نقاط خاصي همانند تيزيها، شكستگي‌ها، خميدگي‌هاي شديد از هم منفك مي‌شود. در مرحله بعد اين اجزا براي شناسايي به شبكه عصبي هدايت مي‌شوند. در نهايت نتيجه خروجي از شبكه عصبي به يک واحد ساختاري ارسال مي‌گردد. اين واحد مسئوليت شناسايي نويسه را به عهده دارد. در صورتيكه نويسه شناسايي شود به خروجي برده مي‌شود درغير اينصورت روند همبند كردن اجزاي تجزيه شده صورت مي‌گيرد و مجددا به شبكه عصبي وارد مي‌شوند. اين مكانيزم تا زمانيكه يک نويسه مناسب شناسايي شود يا تمامي حالات همبند كردن اجزاي منفك شده انجام شود، ادامه پيدا مي‌كند. نحوه تاييد يک نويسه در واحد ساختاري نيز بر اساس منطق فازي صورت مي‌گيرد. ميزان بازشناسي صحيح آزمايش شده بر روي 3 پايگاه داده هدي، IFHCDB و متاكد به طور ميانگين براي حروف 93.48% و براي ارقام 96.73% بوده است.

كلمات كليدي: شناسايي نوري نويسه ، OCR، حروف دستنويس، ارقام دستنويس، روش ساختاري، شبكه عصبي، منطق فازي ، نويسه‌هاي مجزا

فهرست مطالب

فصل اول- مقدمه 1
1-1 مقدمه 2
1-2 خواندن نوري مارک 4
1-3 خواندن نوري مارک بر پايه اسکن 4
1-4 تشخيص نويسه نوري 5
1-5 تشخيص متن تايپ شده 5
1-6 شناسايي هوشمند نويسه 7
1-7 تشخيص متن دستنويس 8
1-8 انواع سيستم‌هاي شناسايي نويسه از لحاظ نوع الگوي ورودي 8
1-9 اصول شناسايي نويسه‌هاي دستنويس 9
1-9-1 قطعه‌بندي 9
1-9-2 استخراج ويژگي‌ها 10
1-9-3 طبقه‌بندي 11
1-10 ساختار پايان‌نامه 12
فصل دوم- پيش‌پردازش و مروري بر کارهاي انجام شده 13
2-1 مقدمه 14
2-2 پيش‌پردازش 14
2-2-1 خاکستري کردن نويسه 15
2-2-2 باينري کردن تصوير 15
2-2-3 تغيير اندازه 16
2-2-4 نازک‌سازي 16
2-2-5 نويزگيري 18
2-2-6 همبندکردن 19
2-3 مروري برکارهاي انجام شده 21
2-3-1 شناسايي به روش ساختاري 21
2-3-2 شناسايي با استفاده از شبکه عصبي 24
2-3-3 شناسايي آماري 29
2-3-4 شناسايي بر اساس روش‌هاي ديگر 31
2-4 نتيجه‌گيري 34
فصل سوم- روش پيشنهادي 36
3-1 مقدمه 37
3-2 کليات روش پيشنهادي 37
3-3 شناسايي نقاط حساس و تجزيه شکل 38

3-4 شناسايي اجزاي كوچک توسط شبكه عصبي 43
3-5 بخش تشخيص 46
3-6 تعيين نتيجه نهايي 49
فصل چهارم- آزمايش‌ها و نتايج 51
4-1 مقدمه 52
4-2 معرفي پايگاه‌هاي داده تصاوير نويسه‌هاي دستنويس فارسي 52
4-2-1 پايگاه داده هدي 52
4-2-2 پايگاه داده IFHCDB 54
4-2-3 پايگاه داده CENPARMI 55
4-2-4 پايگاه داده شركت متاكد 55
4-3 درصد بازشناسايي ارقام 57
4-3-1 درصد شناسايي ارقام بر روي پايگاه داده هدي 57
4-3-2 درصد شناسايي ارقام بر روي پايگاه داده IFHCDB 58
4-3-3 درصد شناسايي ارقام بر روي پايگاه داده متاكد 59
4-4 درصد بازشناسايي حروف 60
4-4-1 درصد شناسايي حروف بر روي پايگاه داده هدي 60
4-4-2 درصد شناسايي حروف بر روي پايگاه داده IFHCDB 62
4-4-3 درصد شناسايي حروف بر روي پايگاه داده متاكد 64
4-5 نويسه‌هاي شناسايي نشده 66
4-6 نويسه‌هاي به اشتباه شناسايي شده 67
4-7 مقايسه روش پيشنهادي با ساير كارهاي انجام شده 68
فصل پنجم- جمع‌بندي و پيشنهادها 77
5-1 جمع‌بندي پايان‌نامه 78
5-2 پيشنهادها 78
مراجع 80

فهرست نمودارها و اشکال

شکل 1-1 نمونه‌اي از يک پاسخنامه تستي 4

شکل 1-2 مراحل اجراي الگوريتم شناسايي متن 7

شکل 1-3 نمونه‌اي از تصوير يک فرم ورود اطلاعات 7

شکل 1-4 بلوک دياگرام کلي شناسايي متن از طريق قطعه‌بندي نويسه‌ها 12

شکل 2- 1 تصوير اصلي به صورت رنگي 15

شکل 2- 2 تصوير خاکستري 15

شکل 2- 3 تصوير باينري 16

شکل 2- 4 نويسه پس از تغيير اندازه 16

شکل 2-5 نويسه پس از نازک‌سازي 17

شکل 2-6 خطاي AForge در نازک‌سازي 17

شکل 2-7 نويسه اصلي پيش از نازک‌سازي 18

شکل 2-8 نازک‌سازي توسط نرم افزار 18

شکل 2-9 نازک‌سازي مورد انتظار انسان 18

شکل 2-10 نویسه اصلي پیش از نازک‌سازي 18
شکل 2-11 نویسه اصلي پس از نازک‌سازي 18
شکل 2-12 نویسه اصلي پیش از نویزگیري 19
شکل 2-13 نویسه پس از نویزگیري 19
شکل 2-14 نویسه اصلي پیش از نازک‌سازي 20
شکل 2-15 نویسه پس از نازک‌سازي 20
شکل 2-16 همبند کردن نویسه پس از نازک‌سازي 20
شکل 2- 17 تجزیه حرف "ع" به عناصر NAPE 23
شکل 2-18 نقاط نرمال شده 26
شکل 2-19 ساختار شبکه عصبي مربوط به گروه A 26
شکل 2-20 نمونه‌اي از شبکه عصبي پرسپترون چند لایه 27
شکل 2-21 دیاگرام روش شناسایي پیشنهادي 28
شکل 2-22 مکانهاي مشخصه مورب 31

شكل 3-1 فلوچارت الگوريتم پيشنهادي 38
شكل 3-2 تقسيم‌بندي اول از نويسه"6" 39
شكل 3-3 تقسيم‌بندي دوم از نويسه "6" 39
شكل 3-4 قطعه‌بندي صحيح نويسه "6" 39
شكل 3-5 نقاط ابتدايي در رقم "4" 40
شكل 3-6 نقاط تقاطع در حرف "س" 40
شكل 3-7 نقطه شكستگي در رقم "8" 41
شكل 3-8 نقاط خميدگي شديد در حرف" ح" 42
شكل 3-9 تصوير وتر با زاويه بزرگتر 42
شكل 3-10 تصوير وتر با زاويه كوچكتر 42
شكل 3-11 نمونه نويسه استاندارد "2" 44
شكل 3-12 نمونه نويسه دستنويس "2" 44
شكل 3-13 تصويري از شبكه‌هاي عصبي تشخيص دهنده قطعات با خروجي بيتي 45
شكل 3-14 تصويري از شبكه عصبي تشخيص دهنده قطعات با خروجي عددي 45

شکل 3-15 تصوير کلي شبکه عصبي شناسايي قطعه 46
شکل 3-16 نمونه اول از نوشتن نويسه "2" 47
شکل 3-17 نمونه دوم از نوشتن نويسه "2" 47
شکل 3-18 نگاه کلي به تشخيص ساختاري 48
شکل 4-1 نمونه‌اي از حروف دستنويس فارسي در مجموعه داده هدي 53
شکل 4-2 نمونه‌اي از ارقام دستنويس فارسي در مجموعه داده هدي 54
شکل 4-3 نمونه‌اي از فرم‌هاي جمع‌آوري شده نويسه‌هاي دستنويس شرکت متاکد 56
شکل 4-4 درصد شناسايي ارقام به تفکيک و ميانگين بر روي پايگاه داده هدي 58
شکل 4-5 درصد شناسايي ارقام به تفکيک و ميانگين بر روي پايگاه داده IFHCDB. 59
شکل 4-6 درصد شناسايي ارقام به تفکيک و ميانگين بر روي پايگاه داده متاکد 60
شکل 4-7 درصد شناسايي حروف به تفکيک و ميانگين بر روي پايگاه داده هدي 62
شکل 4-8 درصد شناسايي حروف به تفکيک و ميانگين بر روي پايگاه داده IFHCDB..64
شکل 4-9 درصد شناسايي حروف به تفکيک و ميانگين بر روي پايگاه داده متاکد 66
شکل 4-10 نمونه‌اي از ارقام که شناسايي نشده‌اند 67
شکل 4-11 نمونه‌اي از حروف که شناسايي نشده‌اند 67

شکل 4-12 نمونه‌اي از ارقام و حروفي که به اشتباه شناسايي شده‌اند..... 68

فهرست جداول

جدول 2-1 عناصر اوليه بکار رفته در مدلهاي پلکس حروف و ارقام فارسي 22

جدول 2-2 عناصر NAPE بکار رفته در مدلهاي پلکس حروف و ارقام فارسي 22

جدول 2-3 بعضي از قواعد پلکس تعريف شده در پايگاه قواعد براي حروف و ارقام فارسي23

جدول 2-4 تقسيم حروف به 12 گروه بر اساس علائم آنها 25

جدول 2- 5 نتايج آزمايشات بر روي تعداد نرون‌هاي مختلف در لايه مخفي 29

جدول 2- 6 گروه‌بندي بر اساس تعداد اجزا 30

جدول 3-1 تقسيم‌بندي اجزا بر اساس مدل هندسي 43

جدول 4-1 درصد شناسايي ارقام به تفکيک بر روي پايگاه داده هدي 57

جدول 4-2 درصد شناسايي ارقام به تفکيک بر روي پايگاه داده IFHCDB..... 58

جدول 4-3 درصد شناسايي ارقام به تفکيک بر روي پايگاه داده متاکد 59

جدول 4-4 درصد شناسايي حروف به تفكيک بر روي پايگاه داده هدي 61

جدول 4-5 درصد شناسايي حروف به تفكيک بر روي پايگاه داده IFHCDB 63

جدول 4-6 درصد شناسايي حروف به تفكيک بر روي پايگاه داده متاكد 65

جدول 4-7 خلاصه شناسايي روش‌هاي ديگر 73

فصل اول

مقدمه

1-1 مقدمه

امروزه سيستم هاي شناسايي نوري نويسه‌ها بسيار مورد توجه قرار گرفته است. كارهاي فراواني بر روي حروف لاتين و چيني انجام شده است. اين كارها از دهه 70 ميلادي شروع شده و هنوز براي بهبود آنها تلاش مي شود. از دهه 90 بر روي نويسه‌هاي عربي كه با نويسه‌هاي فارسي به غير از چهار حرف همسان هستند، كار شده است. به نظر مي رسد براي بهبود سيستم هاي تشخيص نويسه‌ها و استفاده آن در صنعت و زندگي روزمره تحقيقات بيشتري نياز است.

به دنبال پيشرفت‌هاي حاصله در زمينه ساخت پردازنده‌هاي سريع، اثر محدود كننده دستگاه هاي ورودي متداول بر سرعت انتقال اطلاعات به داخل رايانه بيش از پيش نمود يافته است. عمل بازشناسي نوشته‌ها، شبيه‌سازي فرآيند خواندن توسط انسان بر روي ماشين است. در ساليان اخير بررسي هاي زيادي در اين زمينه صورت گرفته است و نتيجه آن ابداع الگوريتم هاي متعددي براي بازشناسي نوشته هاي الفباهاي مختلف بوده است. فناوري شناسايي متون از تشخيص قالب و الگو، هوش مصنوعي و چشم ماشيني استفاده مي‌كند تا متن‌هايي را كه به‌صورت چاپ‌شده يا دست‌نوشته بر روي كاغذهاي معمولي در اختيار داريم به متن‌هاي قابل ويرايش توسط نرم‌افزارهاي ويرايشگر متن با كامپيوتر تبديل كند.

معمولاً از اين فناوري در مواقعي استفاده مي‌شود كه فرصت زيادي براي تايپ كردن متون وجود ندارد و مي‌خواهيم به سرعت متني را كه به شكل دست‌نوشته با قلم يا خودكارهاي معمولي روي كاغذ نوشته شده است، اسكن كرده تا بتوان آن را به شكل‌هاي مختلف، ويرايش كرد. يكي ديگر از كاربردهاي فناوري تشخيص نوري نويسه‌ها در زماني است كه بخواهيم متن مقالات، كتاب‌ها يا نوشته‌هاي از قبل چاپ‌شده‌اي را كه حجم زيادي دارند اسكن كرد زيرا ذخيره‌سازي آنها در قالب فايل‌هاي تصويري مثل JPEG يا GIF امكان‌پذير نيست. كاربرد مهمتر ديگري كه مي‌توان در اين زمينه ذكر نمود امكان جستجو در متن مي‌باشد كه چنين امكاني در فرمتهاي تصويري وجود ندارد. با اين تفاسير مي‌توانيم به طور خلاصه اين گونه جمع‌بندي كنيم كه شناسايي نويسه‌ها توسط ماشين براي تبديل تصاوير گاها محو و مبهم به متن‌هاي قابل درک و ويرايش مورد استفاده قرار مي‌گيرد و اشكال گرافيكي را به نويسه‌هاي ماشيني تبديل مي‌كند، به شكلي كه فونت، اندازه، استيل، صفحه‌آرايي و تمام مختصات پاراگرافي آنها قابل ويرايش باشد.

در سال ۱۹۲۹ براي نخستين بار، گوستاو توشچک از آلمان، امتيازي براي اختراع خود با نام OCR دريافت كرد و از همان زمان، كار طراحي نرم‌افزارهاي هوشمند درک متن را آغاز نمود. اختراع او، يک دستگاه اسكنر نوري بود كه يک سري الگوهاي متني آماده را در خود ذخيره مي‌كرد و متن‌هايي كه دريافت مي‌كرد در صورت تطابق كامل و تشابه بدون كم و كاست با الگوي ابتدايي ذخيره‌شده، درک مي‌كرد و به كامپيوتر مي‌داد. ايراد

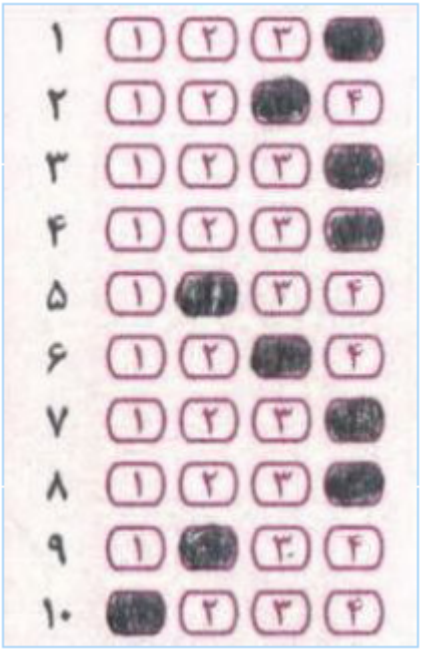
اختراع او اين بود كه يک نمونه از پيش تهيه‌شده شامل شكل كوچک و بزرگ حروف الفبا در اختيار داشت و در صورتي كه متن ارائه شده به دستگاه، كوچک‌ترين تفاوتي با الگو و نمونه آماده داشت، قادر به تشخيص نبود. با اين حال، اين كار سرآغاز ارائه اختراعاتي شد كه بعدها به توليد نرم‌افزارهاي هوشمند تشخيص متن انجاميد. در حال حاضر، اين نرم‌افزارها براي درک متون تصويري به زبان‌هاي مختلف طراحي شده‌اند و نمونه‌هاي جديد آنها بر روي بسياري از گوشي‌هاي تلفن‌همراه نصب شده است. اساس و پايه كار قلم‌هاي نوري كه همراه با برخي گوشي‌هاي تلفن همراه ارائه مي‌شوند نيز همين فناوري است. زماني كه شما براي تايپ كردن متن يک پيام كوتاه يا شماره‌گيري كردن، بدون در اختيار داشتن هيچ كيبوردي، تنها به نوشتن معمولي اعداد و حروف با دستخط خاص خودتان مبادرت مي‌كنيد، يعني در حال استفاده از ثمرات OCR هستيد، آن هم به شكلي كه يک صفحه كريستال مايع هوشمند به كمک قلم نوري، امكان درک و رديابي خطوط به ظاهر بي‌معني و ترسيم‌شده توسط شما را فراهم مي‌كند. در سال ۱۹۵۰ دومين نمونه نرم‌افزار تشخيص نوري نويسه‌ها توسط ديويد شپارد از سرويس جاسوسي ارتش آمريكا طراحي شد. شپارد كه متخصص رمزشكني و باز كردن قفل‌هاي مخفي اطلاعاتي و متني بود، با همكاري يک پروفسور ژاپني، نمونه‌اي از نرم‌افزار نوين OCR را طراحي كرد كه نيازي به الگوي ثابت از پيش تعيين شده نداشت و به راحتي مي‌توانست دست‌خط‌هاي افراد مختلف را درک كند. ويژگي اين نرم‌افزار، فهميدن و درک راحت متن‌هايي بود كه بسيار كم‌رنگ و يا توسط مداد نوشته شده بودند و يا بر اثر گذشت زمان، از وضوحشان كاسته شده بود. در آن زمان بالا بردن كيفيت و ترميم تصاوير قديمي، تار و محو شده بسيار سخت بود و به آساني انجام نمي‌گرفت. همچنين سرويس پستي ايالات متحده، از سال ۱۹۵۶ شروع به استفاده از OCR كرده تا بتواند امور مربوط به ارسال و دريافت نامه‌هاي خود در سراسر دنيا را ساماندهي كند. با استفاده از فناوري تشخيص نوري متن، ادارات پست سراسر آمريكا به جاي وارد كردن دستي يا تايپ اطلاعات مربوط به هر نامه، پاكت‌ها و بسته‌هاي پستي را اسكن مي‌كنند و اطلاعات مورد نياز نيز به شكل خودكار توسط كامپيوتر فهميده و ذخيره مي‌شود.

1-2 خواندن نوري مارک[1]

خواندن نوري مارک عمدتا در آزمونهاي چند گزينه‌اي مورد استفاده قرار مي‌گيرد. بدين ترتيب كه فرمهاي خاصي به صورت چند گزينه‌اي طراحي مي‌شود كه آزمون‌دهندگان با پركردن محل گزينه، پاسخ مورد نظر را انتخاب مي‌كنند. دستگاههاي OMR با استفاده از تاباندن نور پر بودن گزينه‌هاي پاسخنامه را بررسي مي‌كنند. در OMR بايد گزينه‌ها با

[1] Optical Mark Reader

مداد نرم و مشكي و به صورت كامل پر شوند[1]. در شكل 1-1 نمونه‌اي از يک پاسخنامه مشاهده مي‌شود.



شكل 1- 1 نمونه‌اي از يک پاسخنامه تستي [1]

1-3 خواندن نوري مارک بر پايه اسکن[2]

اين روش عمل خواندن نوري مارک را براساس نرم افزار و بر پايه اسکن كردن پاسخنامه انجام مي‌دهد. در اين روش سرعت اسکن كردن و تشخيص گزينه‌هاي پر شده بسيار بالا مي‌باشد ضمن اينكه ديگر نيازي به الزام پركردن پاسخنامه با مداد مشكي نمي‌باشد و با هر نوع قلمي امكان پذير است. سهولت، سرعت و دقت از مزاياي اين روش است و در حال حاضر در تمامي آزمونهاي روز دنيا مورد استفاده قرار مي‌گيرد[2].

1-4 تشخيص نويسه نوري[3]

تشخيص نويسه نوري از مهمترين شاخه هاي بازشناسي متن است. پسوند نوري در اين عبارت در مقابل Magnetic Ink قرار داده شد تا اين روش را از روش قديمي‌تر بازشناسي نويسه ها با مركب مغناطيسي MICR متمايز كند. OCR، براي بازشناسي حروف و ارقام تنها تايپ شده مورد استفاده قرار مي‌گيرد. در اين متد تصوير سند اسكن شده به موتور بازشناسي داده مي‌شود تا نويسه معادل آن را توليد كند. با توجه به اينكه

[2] Scanned base Optical Mark Reader
[3] Optical Character Reader

در OCR نويسه‌ها چاپي هستند، لذا تشخيص آنها براي موتورهاي بازشناسي چندان كار دشواري نيست و هم اكنون الگوريتم‌هاي فراواني براي تمامي زبان‌هاي دنيا موجود مي‌باشد. در حالت كلي‌تري نرم افزارهايي وجود دارند كه يک شكل را آموزش مي‌بينند و پس از آموزش توانايي شناسايي آن شكل را دارند. در نتيجه چنين نرم افزارهايي مي‌توانند تمامي نويسه‌هاي تمامي زبان‌ها را ياد گرفته و تشخيص بدهند[3].

1-5 تشخيص متن تايپ شده

تشخيص نويسه نوري پيوسته عبارت است از بازشناسي خودكار متون موجود در تصاوير اسناد و تبديل آنها به متون قابل جستجو و ويرايش توسط رايانه. تصوير سند غالبا توسط اسكنر و يا دوربين ديجيتال توليد مي‌شود و شامل تعدادي پيكسل با رنگهاي مختلف و سطوح روشنايي گوناگون است. براي اينكه بتوان از اطلاعات نوشتاري تصوير سند استفاده كرد، بايد به نحوي نوشته هاي موجود در سند را بازشناسي كرد. با گذشت زمان و پيشرفت قابل توجه در اين زمينه، روش هاي بازشناسي دست‌نوشته و متون چاپي مطرح شدند كه دامنه كار را به كلمات و عبارات رساندند. با وجود عدم تطبيق دقيق OCR با اين موارد، اين نام براي اين روش‌ها هم استفاده شد و رواج پيدا كرد. هم اكنون OCR را بيشتر براي بازشناسي مستندات چاپي مثل صفحات كتاب ها، مجله ها و نامه هاي چاپي به كار مي برند.

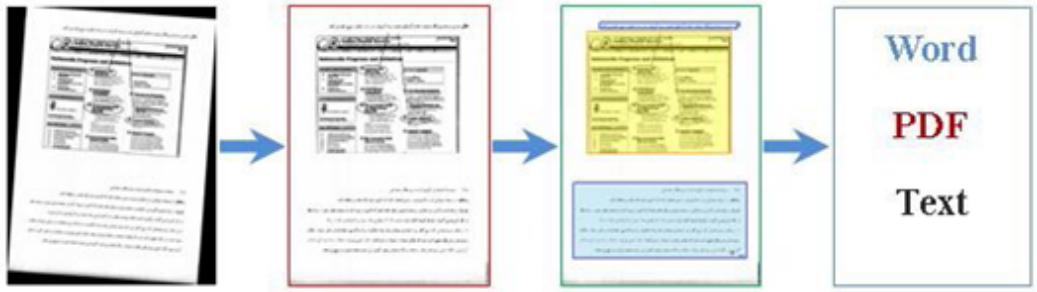
سامانه نويسه خوان مثل يک نفر ماشين نويس يا تايپيست، متن سند را مي‌خواند و آن را به قالب مناسب براي ذخيره فراهم مي‌كند. سامانه نويسه خوان، اشياء موجود در تصوير سند را كه ارقام، حروف، علائم و كلمات هستند، بازشناسي كرده و رشته‌ي متناظر با آن ها را درقالب مناسب ذخيره مي‌كند. يک فايل تصويري، حجم زيادي دارد و جستجوي متني در آن ممكن نيست. اين در حالي است كه فايل خروجي سامانه نويسه خوان بسيار كم حجم و قابل جستجو است. ضمن اينكه اطلاعات فايل متني را به راحتي مي‌توان ويرايش كرد و در جاي ديگر استفاده كرد در حالي كه اطلاعات موجود در تصوير سند قابل ويرايش نمي‌باشد[3]. سامانه‌هاي نويسه خوان مثل بسياري از سامانه هاي هوشمند ديگر، پيچيدگي زيادي دارند. پردازش تصوير و بازشناسي الگو دو پايه اصلي اين سامانه ها هستند. پيچيدگي اين سامانه‌ها براي زبانهاي گوناگون، متفاوت است. نوشتن OCR براي زبانهاي لاتين به دليل اينكه حروف آنها به طور مجزا نوشته مي‌شود از زبانهايي مثل فارسي و عربي كه حروف يک كلمه به يكديگر مي‌چسبند، آسانتر است. اين موضوع به علاوه جمعيت كم كاربران زبان فارسي، سبب شده سامانه هاي نويسه خوان كمي براي زبان فارسي داشته باشيم.

يک الگوريتم كلي براي شناسايي متن مي‌تواند به صورت زير باشد:

1. شناسايي متن در تصوير و اصلاح چرخش در صورت وجود.

2. شناسايي بلوكهاي متني و تفكيک آنها از اشكال و جداول .
3. تشخيص موقعيت كلمات در سطر .
4. تشخيص موقعيت حروف در كلمه .
5. شناسايي تک تک حروف كلمه .
6. تركيب حروف شناخته شده و ساختن كلمه معادل .

در شكل 1-2 بلوک دياگرام شناسايي متن نمايش داده شده است.



شكل 1-2 مراحل اجراي الگوريتم شناسايي متن [1]

1-6 شناسايي هوشمند نويسه[4]

اصطلاح ICR بيشتر توسط بخش تجاري جا افتاده است و در واقع شاخه‌اي از بازشناسي متن است كه به طور خاص به بازشناسي حروف و ارقام دستنويس مي‌پردازد. اطلاعات درخواستي كاربر روي فرمهاي كاغذي و در كادرهاي مجزا وارد مي‌شود. مثلا براي نوشتن نام محمد حروف "م"، "ح"، "م" و "د" در چهار خانه مربع شكل جداگانه وارد مي‌شود. براي داده آمايي(استخراج و ورود اطلاعات) خودكار اين اطلاعات، از فن‌آوري ICR استفاده مي‌شود. در اين روش پس از اسكن فرمها، موقعيت فيلدهاي مختلف (مثل نام، فاميلي، شماره ملي و غيره) تشخيص داده مي‌شود و هر فيلد براي بازشناسي حروف و ارقام موجود در آن آناليز مي‌شود[4]. در شكل 1-3 نمونه‌اي از يک فرم ورود اطلاعات نمايش داده شده است.

[4] Intelligence Character Reader

شكل 1-3 نمونه‌اي از تصوير يک فرم ورود اطلاعات [1]

1-7 تشخيص متن دستنويس

اين فناوري براي شناسايي كامل دست خط به كار مي‌رود. درواقع مراحل تشخيص خط و شناسايي كلمات در ابتدا وجود دارند. سپس سامانه، هر كلمه را به حروف تجزيه مي‌كند و از اين پس ادامه كار به ICR سپرده مي‌شود. البته در زبانهاي پيوسته همانند فارسي و عربي همواره تجزيه كلمات به حروف پاسخ خوبي نمي‌دهد و از الگوريتمهاي ديگر استفاده مي‌شود. در الگوريتم ديگري پس از شناسايي موقعيت كلمه به طور خاص خود كلمه بررسي مي‌شود و به كمک پايگاه‌داده كلمات، كلمه شناسايي مي‌شود. اين متد از جهاتي شبيه به عملكرد انسان نيز مي‌باشد. اين الگوريتم براي زبانهاي پيوسته همانند فارسي از شناسايي قابل قبولتري برخوردار است.

1-8 انواع سيستم‌هاي شناسايي نويسه از لحاظ نوع الگوي ورودي

سيستمهاي شناسايي نويسه را مي‌توان از لحاظ نوع الگوي ورودي به دو گروه اصلي تقسيم كرد: سيستم‌هاي برخط و سيستم‌هاي برون‌خط. در بازشناسي برخط، حروف در همان زمان نگارش توسط سيستم تشخيص داده مي‌شوند. اين اطلاعات معمولا توسط يک صفحه رقومي كننده اخذ مي‌شوند. در اين مدل اطلاعات بيشتري از تصوير براي موتور بازشناسي وجود دارد، همانند اطلاعات زماني، سرعت، شتاب، فشار، زمان برداشتن و گذاشتن قلم، حركت قلم و ترتيب نوشتن اجزا يک حرف. به طور مثال براي حرف "ط" ابتدا بدنه و بعد دسته نوشته مي‌شود يا براي حرف "ک" ابتدا بدنه و سپس سركش نوشته مي‌شود. براي حروف نقطه دار، نقطه در نهايت درج مي‌شود. حتي اگر حروف به اين صورت نوشته نشود مي‌توان اين قانون را براي كاربر تعريف كرد. از طرفي نوع حركت قلم نيز بسيار تاثير گذار بر روي تشخيص نويسه مي‌باشد. لذا طراحي اينگونه موتورها از پيچيدگي كمتري برخوردار است ضمن اينكه از درصد تشخيص بالايي نيز

برخوردار هستند. كاربرد اين موتورها در موبايلها براي SMS، محيطهاي نوشتاري همانند Word و حتي ايميل مي‌باشد.

در بازشناسي برون‌خط، از تصوير دو بعدي متن ورودي استفاده مي‌شود. در اين روش به هيچ نوع وسيله نگارش خاصي نياز نيست و تفسير داده‌ها مستقل از فرايند توليد آنها و تنها بر اساس تصوير متن صورت مي‌گيرد. اين روش به نحوه بازشناسي توسط انسان شباهت بيشتري دارد. در اين مدل ديگر هيچ اطلاعاتي از نحوه نوشتن و ترتيب اجزا وجود ندارد. در نتيجه طراحي و پياده سازي اين دسته موتورها از پيچيدگي بيشتري برخوردار است ضمن اينكه سرعت تشخيص آنها نيز به نسبت برخط كمتر است. كاربرد اين موتورها در خواندن فرمهاي ورود اطلاعات، فرمهاي استخدامي، خواندن كد پستي، نمره خوان مدارس و غيره مي‌باشد.

1-9 اصول شناسايي نويسه‌هاي دستنويس

براي شناسايي نويسه‌هاي دستنويس به طور كلي بايد سه مرحله اصلي انجام شود. اين سه مرحله به ترتيب عبارتند از: پيش پردازش، موتور بازشناسي و پس پردازش. منظور از پيش پردازش هموارسازي و پردازشهاي اوليه بر روي تصوير است كه باعث تميز شدن و اصلاح شدن تصوير مي‌شود. برخي از اين پيش پردازشها عبارتند از: رفع نويز، همبندكردن، نازك‌سازي، رفوكردن و نرمال‌سازي و غيره. در فصل دوم به صورت كامل درباره پيش پردازش‌ها صحبت مي‌شود. پس از پيش‌پردازش، نويسه آماده براي شناسايي مي‌باشد. اين كار توسط موتور بازشناسي انجام مي‌گيرد. موتور بازشناسي يک ماژول نرم افزاري است كه با استفاده از تكنيک‌هاي بازشناسي الگو قابليت شناسايي الگوهاي خاص را دارد. به عنوان مثال يک موتور بازشناسي ارقام دستنويس فارسي قادر است با دريافت تصوير يک رقم، آن را شناسايي كرده و كد متناظرش را برگرداند. مهمترين پارامتر يک موتور بازشناسي، دقت يا نرخ بازشناسي صحيح آن است. نرخ بازشناسي بيانگر اين است كه چند درصد نمونه‌هاي دريافتي توسط موتور بازشناسي به درستي شناخته شده‌اند. موتور بازشناسي نيز داراي يكسري بخشهاي اساسي مي‌باشد كه مهمترين آنها عبارتند از: قطعه‌بندي، استخراج ويژگيها و طبقه‌بندي.

1-9-1 قطعه‌بندي

قطعه‌بندي مرحله بسيار مهم براي سيستم‌هاي شناسايي نويسه مخصوصا در زبانهاي فارسي و عربي است كه در آنها حروف كلمات به صورت سرهم نوشته مي‌شوند. قطعه‌بندي به دو دسته تقسيم مي‌شود:

1. قطعه‌بندي بيروني كه عبارت است از تفكيک قسمتهاي مختلف تصوير نظير متن، گرافيک و خطوط و نيز جدا كردن بخشهاي مختلف متن مانند: پاراگراف‌ها، جملات يا كلمات.

2. قطعه‌بندي دروني كه منظور از آن جداسازي حروف كلمات سرهم نوشته شده در متون لاتين يا در رسم الخط‌هاي پيوسته نظير فارسي و عربي است. همچنين حروفي كه در متن اصلي جدا بوده‌اند اما به خاطر كيفيت پايين دستگاه اسكنر به هم چسبيده‌اند، توسط اين دسته از تكنيكها از يكديگر جدا مي‌شوند.

در موتوربازشناسي حروف دستنويس تنها اين بخش وجود ندارد.

1-9-2 استخراج ويژگي‌ها

اين مرحله يكي از مراحل بسيار با اهميت در سيستمهاي شناسايي نويسه است چرا كه نتايج حاصل از اين مرحله مستقيما بر روي كيفيت مرحله بازشناسي اثر مي‌گذارد. در مرحله استخراج ويژگيها به هر الگوي ورودي يک بردار ويژگي نسبت داده مي‌شود كه معرف آن الگو در فضاي ويژگي‌ها است و آن را از ديگر الگوها متمايز مي‌سازد. محتواي اين بردارهاي ويژگي شامل اطلاعات لازم براي بازشناسي مي‌باشد. انواع ويژگيها را مي‌توان به چهار دسته ويژگيهاي ساختاري، ويژگي‌هاي آماري، تبديل عمومي و تطبيق قالب و همبستگي تقسيم كرد.

1. ويژگيهاي ساختاري: ويژگيهاي ساختاري توصيف كننده يک الگو بر مبناي خواص محلي يا عمومي هندسي و توپولوژيكي آن مي‌باشند. ويژگيهاي ساختاري نسبت به نويز و خرابي تصوير بسيار مقاوم هستند ولي هميشه استخراج آنها از نويسه آسان نمي‌باشد.

2. ويژگيهاي آماري: ويژگيهاي آماري، يک الگو را به كمک يک مجموعه از خواص قابل اندازه‌گيري از الگو، توصيف مي‌كنند.

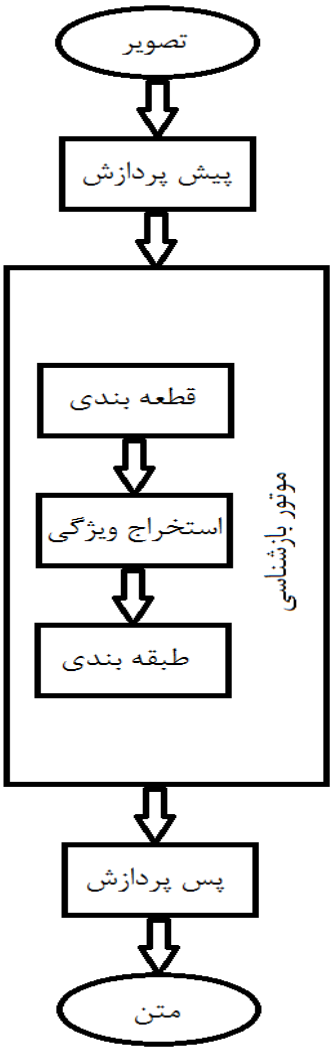
3. تبديل عمومي: رويه تبديل يعني تغيير نمايش يک پيكسل از يک الگو به يک فرم با اطلاعاتي كه ابعاد خصوصيات را كاهش داده است. يكي از ساده‌ترين تبديل‌ها نمايش اسكلت يا كانتور الگو به كمک يک زنجيره كد از جهت امتداد پيكسل‌هاي سياه آن همانند كد فريمن در هشت جهت يا كدهاي شش جهته مي‌باشد.

4. تطبيق قالب و همبستگي: در روش تطبيق قالب و همبستگي ابتدا يک الگو پيكسل به پيكسل با يک مجموعه از قالبها مقايسه مي‌شود و نياز به محاسبه ويژگي خاصي وجود ندارد.

1-9-3 طبقه‌بندي

اين مرحله شامل روشهايي براي متناظر ساختن هر يک از الگوهاي به دست آمده از مرحله استخراج ويژگي‌ها با يکي از طبقات فضاي الگوهاي مورد بحث است که از طريق پيدا کردن کمترين فاصله بردار ويژگي‌هاي هر الگوي ورودي نسبت به بردارهاي مرجع، انجام مي‌گيرد. بردارهاي مرجع، بردارهايي هستند که قبلا با استفاده از نمونه‌هاي آموزشي محاسبه شده‌اند. تکنيک‌هاي عرضه شده براي مرحله طبقه‌بندي را مي‌توان در چهار گروه تطابق قالبي، طبقه‌بندي آماري، طبقه‌بندي ساختاري و شبکه‌هاي عصبي جستجو نمود.

در برخي موارد ممکن است موتور بازشناسي در بين دو نويسه ترديد کند و نتواند به صورت نرمال نويسه واقعي را تشخيص دهد. براي رفع اين مشکل هر دو نويسه تشخيص داده شده به همراه تصوير ورودي به بخش پس‌پردازش داده مي‌شود تا در اين بخش فقط بين دو نويسه بررسي‌هاي لازم صورت گيرد و تشخيص نهايي داده شود. معمولا اين اتفاق زماني مي‌افتد که نويسه‌ها از لحاظ ساختار ظاهري شبيه به هم هستند يا به عبارت ديگر در بسياري از ويژگيهاي استخراج شده از نويسه شباهت دارند. الگوريتمهاي پس‌پردازش همانند موتوربازشناسي عمل مي‌کنند و در واقع شروع به بازشناسي مجدد نويسه مي‌نمايند فقط با اين تفاوت که نويسه نهايي را در بين نويسه‌هاي خروجي موتور بازشناسي اصلي جستجو مي‌کنند. لذا حوزه نويسه‌هاي موجود کمتر شده و شناسايي آسانتر مي‌گردد. لازم به ذکر است در سيستم‌هاي شناسايي کلمه استفاده از يک فرهنگ لغات به عنوان يک پس‌پردازش مي‌تواند بسيار مفيد باشد. در شکل 1-4 بلوک دياگرام کلي شناسايي متن از طريق قطعه‌بندي نويسه‌ها نمايش داده شده است.



شکل 1-4 بلوک دیاگرام کلي شناسايي متن از طريق قطعه‌بندي نويسه‌ها

1-10 ساختار پايان‌نامه

در ادامه اين پايان‌نامه در فصل دوم مروري بر کارهاي انجام شده در زمينه شناسايي نويسه‌هاي تنها فارسي ارايه شده است. فصل سوم به ارايه روش پيشنهادي اختصاص دارد. فصل چهارم مربوط به گزارش آزمايش‌ها و نتايج حاصل است و نهايتا در فصل پنجم جمع‌بندي و پيشنهادات آمده است.

فصل دوم

پیش‌پردازش و

مروري بر كارهاي انجام شده

2-1 مقدمه

در زمينه شناسايي نويسه‌هاي دستنويس تحقيقات بسيار زيادي انجام شده است. در زمينه نويسه‌هاي لاتين به علت ساده بودن شكل آنها درصد شناسايي بهتر گزارش شده است. در مورد شناسايي نويسه‌هاي فارسي و عربي به علت پيچيدگي شكل و همچنين تشابهات فراوان نويسه‌ها، اين درصد به خوبي نويسه‌هاي لاتين نمي باشد. در هر صورت اين تحقيقات همچنان ادامه دارند و روز به روز نتايج بهتري در اين زمينه بدست مي‌آيد. روش‌هاي ساختاري، تطابق شكل و آماري به كمک ابزاري همانند شبكه‌هاي عصبي، الگوريتمهاي تكاملي و منطق فازي روند بسيار خوبي را در اين مقوله ارائه كردند. در ادامه اين فصل ابتدا روش‌هاي معمول پيش‌پردازش معرفي شده و سپس به ارائه چند تحقيق روز پرداخته مي‌شود.

2-2 پيش‌پردازش

انسان به دليل اينكه از دوران كودكي آموزش ديده است و حجم اين آموزش بسيار بالا بوده است در كمترين زمان ممكن توانايي شناسايي يک نويسه را دارد. در طي اين آموزش خيلي از موارد به طور ناخودآگاه در ذهن ثبت مي‌شود. به طور مثال اگر رقمي نوشته شده باشد كه در كنار آن يک نقطه به عنوان نويز باشد، انسان بدون توجه به آن نويز رقم را صد در صد تشخيص مي‌دهد. همچنين انسان هيچگاه اندازه، حجم و ساير قلم نوشته برايش مهم نيست. در صورتيكه براي خواندن يک نويسه توسط نرم‌افزار حتي كوچكترين پيكسل هم تاثير گذار است. اگر نويز موجود در تصوير حذف نشود آنگاه سيستم ممكن است نويسه‌ي را به اشتباه تشخيص دهد يا حتي نتواند نويسه را تشخيص دهد. در روشهايي همانند شبكه‌هاي عصبي كه سيستم آموزش مي‌گيرد، نبايد نويسه‌ي با نويز به آن آموزش داده شود. لذا براي جلوگيري از چنين مشكلاتي قبل از شناسايي نويسه بايد تصوير هموار شود يا به اصطلاح يكسري پيش‌پردازش‌هايي صورت بگيرد.

برخي از اين پيش‌پردازشها عبارتند از:

- خاكستري كردن نويسه
- سياه و سفيد كردن نويسه
- تغيير اندازه
- نازک‌سازي
- نويزگيري
- همبند كردن

2-2-1 خاكستري كردن نويسه

منظور از خاكستري كردن، بردن تصوير از مد رنگي به مد خاكستري مي‌باشد. در حالت رنگي، رنگ هر پيكسل از تركيب سه رنگ آبي، قرمز و سبز كه بين 0 تا 255 مي‌توانند مقداردهي بشوند، حاصل مي‌شود. در مد خاكستري بايد مقدار عددي اين سه رنگ با هم برابر باشد تا رنگ خاكستري حاصل شود. لذا براي تبديل تصوير اسكن شده به خاكستري مي‌توان مقدار قرمز، سبز و آبي را از 2-1 بدست آورد.

$$\text{Gray Value}=0.3R+0.59G+0.11B \qquad (2\text{-}1)$$

كه در آن R، G و B به ترتيب مقادير رنگ‌هاي قرمز، سبز و آبي مي‌باشد.

در شكل 2-1 عكس رنگي اصلي و در شكل 2-2 تبديل عكس رنگي به خاكستري نمايش داده شده است.

شكل 2-1 تصوير اصلي به صورت رنگي

شكل 2-2 تصوير خاكستري

2-2-2 باينري كردن تصوير

پس از آنكه تصوير ورودي به مد خاكستري رفت بايد آن را به صورت باينري درآورد. براي اين كار كافيست آستانه‌اي براي 2-1 در نظر گرفت تا مرزي براي سياه يا سفيد شدن هر پيكسل بدست آورد. به صورت تجربي اين آستانه عددي بين 180 تا 200 مي‌باشد. در شكل 2-3 نمونه باينري نمايش داده شده است.

شكل 2- 3 تصوير باينري

2-2-3 تغيير اندازه

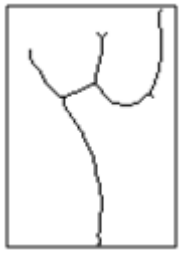
درصورتيكه نحوه شناسايي به گونه‌اي باشد كه تصوير ورودي نياز به تطابق با پايگاه‌داده داشته باشد، بايد تصوير ورودي هم اندازه با تصوير پايگاه‌داده بشود تا عمل تطبيق به درستي انجام شود. ابتدا تصوير ورودي با توجه به اولين پيكسل سياه از هر طرف قطع

مي‌شود. در واقع مختصات اولين پيكسل سياه از بالا، پائين، چپ و راست بدست مي‌آيد. سپس موقعيت هر پيكسل بسته به ميزان تغيير طول و عرض تصوير مكان‌يابي مي‌شود. در شكل 2-4 تصوير تغيير اندازه داده شده مشاهده مي‌شود.

شكل 2- 4 تصوير پس از تغيير اندازه

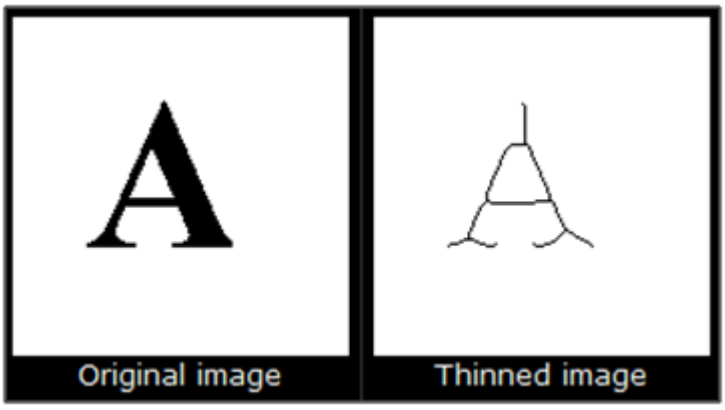
2-2-4 نازک‌سازي

نازک‌سازي مهمترين مرحله پيش پردازش براي شناسايي نويسه‌ها مي‌باشد. نازک‌سازي پردازشي است براي مينمم‌سازي پهناي يک خط در يک تصوير. به عبارت ديگر پهناي يک نويسه به يک پيكسل تغيير مي‌كند. الگوريتم‌هاي نازک سازي به دو دسته تقسيم مي‌شوند: الگوريتم‌هاي ترتيبي و الگوريتمهاي موازي. تفاوت اصلي بين آنها اين است كه در الگوريتم ترتيبي عمل بر روي يک پيكسل در لحظه اتفاق مي‌افتد و نتيجه عمليات بستگي به نتايج پردازش‌هاي قبلي دارد در صورتيكه در الگوريتم‌هاي موازي همه پيكسل‌ها همانند مي‌باشد و عملگر به يک صورت بر روي آنها اعمال مي‌شود. بيشتر الگوريتمهاي نازک‌سازي بر اساس فرسايش دور شكل عمل مي‌كنند و در واقع پس از شناسايي دور شكل به نحوي شروع به پاک كردن پيكسل‌هاي دور مي‌كنند[5]. الگوريتم‌هاي زيادي براي نازک سازي ارائه شده است كه در ادامه يكي از آنها بنام AForge معرفي مي‌شود. AForge، يک بسته نرم افزاري شامل توابع مربوط به پردازش تصوير است. از جمله توابع موجود درAForge توابع مربوط به نازک‌سازي مي‌باشد كه در بسياري از موارد به خوبي عمل مي‌كند و شكل را به درستي نازک مي‌كند. در شكل 2-5 نويسه نازک شده نمايش داده شده است.



شكل 2-5 نويسه پس از نازک‌سازي

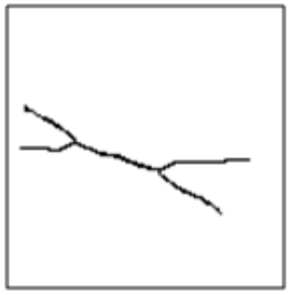
با اينكه AForge يک بسته كامل و قابل قبولي است ولي در برخي از تصاوير عمل نازک‌سازي به درستي صورت نمي‌گيرد و منجر به بوجود آمدن يكسري پاره خطهاي كوچكي كه به شكل چسبيده هستند مي‌شود. در نتيجه اين شاخک‌ها به نحوي باعث ناخواناپي تصوير براي موتور بازشناسي مي‌شوند[6]. در شكل 2-6 نويسه اوليه و نازک شده توسط AForge ديده مي‌شود كه باعث بوجود آمدن شاخكهاي اضافي شده است.



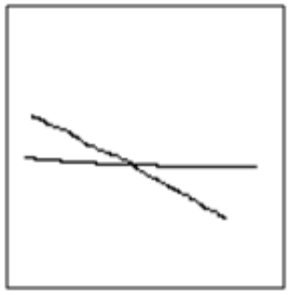
شكل 2-6 خطاي AForge در نازک‌سازي [6]

نازک‌سازي در نهايت مشكلاتي نيز دارد كه از جمله مهمترين آنها تفاوت بين نازک‌سازي توسط نرم افزار و انتظار انسان از نازک‌سازي تصوير است. در نتيجه درصورتيكه موتوربازشناسي براساس نازک شده تصوير بخواهد كار كند در چنين مواقعي پاسخ صحيح را نمي‌تواند بدهد. اشكال 2-7، 2-8 و 2-9 گوياي اين مطلب مي‌باشند.

شكل 2-7 تصوير اصلي شكل



2-8 نازک‌سازي توسط نرم افزار



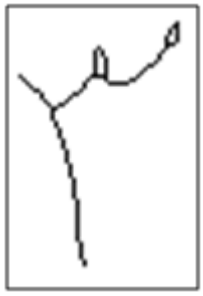
شكل2-9 نازک‌سازي مورد انتظار انسان

2-2-5 نويزگيري

در برخي از تصاوير به علت اسكن يا پارگي يک يا چند نقطه سفيد در وسط تصوير و خطوط ترسيمي بوجود مي‌آيد. در صورتيكه اين نقاط سفيد نويزگيري نشوند پس از

نازک‌سازي تصوير حاصل شامل يکسري دوايرتوخالي مي‌شود. اين دوايرتصوير را به شدت تحت تاثير قرار مي‌دهند. شکل 2-10 نويسه‌ي قبل از نازک‌سازي را نشان مي‌دهد و شکل2-11 نويسه‌ي بعد از نازک سازي را نشان مي‌دهد.

شکل 2-10 تصوير اصلي



شکل 2-11 تصوير اصلي پس از نازک‌سازي

براي رفع نويز تصوير مي‌توان از يکسري پنجره‌هاي سه وضعيتي استفاده کرد. اين پنجره‌ها به صورت آرايه دو بعدي هستند که بر روي تصوير حرکت مي‌کنند و در صورتيکه پيکسل‌هاي تصوير دقيقا مطابق با پنجره باشند آنگاه پنجره اعمال مي شود. منظور از اعمال شدن پنجره تبديل تمامي نقاط سفيد متناظر به سياه است. بدين ترتيب قسمتي از تصوير که سوراخ‌هاي سفيد دارد سياه مي‌گردند. از طرفي گاهي اوقات براساس پخش شدن جوهر خودکار يا به دليل کثيف بودن فرم يا دلايل ديگر ممکن است در تصوير نويزهاي تيره بوجود آمده باشند. اين نويزها ممکن است در خواندن تصوير تاثير به سزايي داشته باشند. براي رفع اين نويزها مي‌توان از يکسري پنجره‌هاي ديگر استفاده کرد. معمولا در تعريف اين پنجره‌ها شعاع پيکسل‌هاي سفيد بيشتر در نظرگرفته مي‌شود تا از نويز بودن آن اطمينان بيشتري حاصل شود. شکل 2-12 تصوير با نويز را نشان مي‌دهد و شکل 2-13 تصويري پس از نويزگيري را نمايش مي‌دهد.

شکل 2-12 تصوير اصلي

شکل 2-13 تصوير پس از نويزگيري

نكته قابل ذكر در مورد نقاط است كه در حروف فارسي مورد استفاده مي‌باشد. براي اينكه الگوريتم نويزگيري اين نقاط را به عنوان نويز محسوب نكند، الگوريتم‌هايي براي شناسايي نقاط وجود دارد كه بايد قبل از انجام نويزگيري اعمال شوند.

2-2-6 همبندكردن

در زمان نوشتن يک نويسه توسط خودكار ممكن است يک لحظه قطعي جوهر بوجود بيايد، در نتيجه قسمتي از نويسه نوشته نمي‌شود يا به اصطلاح گسستگي بوجود مي‌آيد. در برخي موارد استفاده از نازک‌سازي نامعتبر يا اسكن با درجه تفكيک پائين يا انتخاب نادرست آستانه در باينري كردن نيز باعث بوجود آمدن گسستگي مي‌شود. در شكل 2-14 گسستگي قبل از نازک‌سازي و در شكل 2-15 گسستگي بعد از نازک سازي نشان داده شده است.

شكل 2-14 نويسه اصلي

شكل 2-15 نويسه پس از نازک‌سازي

لذا از بوجود آمدن گسستگي نمي‌توان قطعا جلوگيري كرد پس نياز به ترميم آن مي‌باشد. براي ترميم در ابتدا نياز به شناسايي محل گسستگي مي‌باشد كه اينكار توسط گراف و كلاسهاي هم‌ارزي قابل حل مي‌باشد. پس از شناسايي گسستگي بايد دو گراف غيرهمبند بهم متصل شوند كه اينكار مي‌تواند با انتخاب دو سر نزديک گرافها بهم انجام شود كه البته همواره پاسخ صحيحي نمي‌دهد. عمل همبند كردن تصوير بايد پس از نازک‌سازي انجام بگيرد. در شكل 2-16 نويسه همبند شده بعد از نازک‌سازي نمايش داده شده است.

شكل 2-16 همبند كردن نويسه پس از نازكسازي

2-3 مروري بركارهاي انجام شده

در اين بخش مروري بر كارهاي انجام شده بر اساس روش‌ها و الگوريتم‌هاي مختلف به صورت مفصل صورت مي‌گيرد.

2-3-1 شناسايي به روش ساختاري

الف. ساختار پلكس

ساختار پلكس، فرم كلي‌تري از ارائه‌هاي چند بعدي است[7، 8] كه به دليل قابليت تعريف چندين نقطه اتصال براي يک پايانه يا عنصر اوليه، محدوديت‌هاي زبان توصيف شكل را ندارد. عنصر اصلي در يک ساختار پلكس، عنصر چند اتصالي NAPE نام دارد و با ليستي از نقاط اتصال تعريف مي‌شود. يک ساختار پلكس از اتصال چندين NAPE بوجود مي‌آيد. عناصر تشكيل دهنده يک ساختار پلكس[9] عبارتند از:

- يک ليست از عناصر NAPE
- يک ليست ازنقاط اتصالات داخلي در عناصر NAPE
- يک ليست از نقاط اتصالات خارجي پلكس براي اتصال با ديگر پلكس‌ها

فرايند شناسايي ساختاري حروف شامل[10] مراحل زير است:

1. نازک سازي.
2. تجزيه الگو به عناصر NAPE.
3. استخراج عناصر اوليه در هر NAPE توسط كدهاي زنجيره‌اي فريمن و تعبير اين كدها.
4. پس پردازش براي يافتن حالتهاي خاص.
5. انديس گذاري براي نقاط اتصال خارجي و نقاط اتصال مجازي.
6. مقايسه شي پلكس با مدلهاي از پيش تعريف شده.

در جدول 2-1 عناصر اوليه بكار رفته در مدلهاي پلكس حروف و ارقام فارسي نمايش داده شده است.

جدول 2-1 عناصر اوليه بكار رفته در مدلهاي پلكس حروف و ارقام فارسي[10]

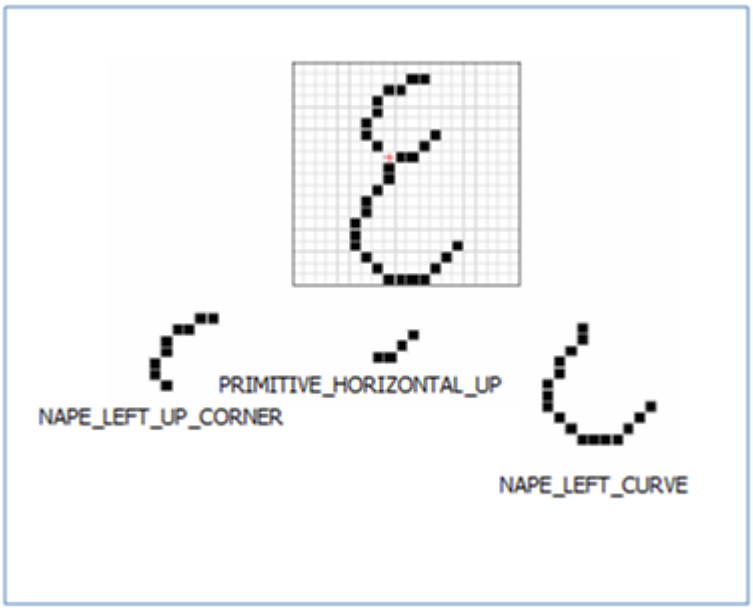
PRIMITIVE_HORIZONTAL_UP خطي با زاويه افقي بين 0 تا 45	PRIMITIVE_VERTICAL_LEFT خطي با زاويه افقي بين 45 تا 90
PRIMITIVE_HORIZONTAL_DOWN خطي با زاويه افقي بين 45- تا 0	PRIMITIVE_VERTICAL_RIGHT خطي با زاويه بين 90 تا 135
PRIMITIVE_DOT عنصري كه ايزوله است و طول آن از يک حد آستانه كمتر است	

در جدول 2-2 عناصر NAPE بكار رفته در مدل‌هاي پلكس حروف و ارقام فارسي نشان داده شده است. پس از جداسازي شكل و تجزيه آن توسط نقاط اتصال خارجي بايد قطعات مشخص و با عناصر جدول NAPE مقايسه شود تا كد مربوط به آن استخراج شود.

جدول 2-2 عناصر NAPE بكار رفته در مدلهاي پلكس حروف و ارقام فارسي[10]

NAPE_LOOP يک حلقه بسته مانند حلقه در "م" يا "ق"	NAPE_BOWL كاسه مانند منحني كاسه شكل در " ن"
NAPE_RIGHT_DOWN_CORNER گوشه سمت راست پائين مانند دندانه سمت راست در "2"	NAPE_PLATE بشقاب مانند منحني بشقاب شكل در "ب"
NAPE_LEFT_UP_CORNER گوشه سمت چپ بالا مانند دندانه بالا در "4"	NAPE_HAT كلاه مانند منحني در "ح"

براي شناسايي هر جز NAPE از كد زنجيره‌اي فريمن[11] كه شامل اطلاعات جهت مي‌باشد، استفاده مي‌شود. براي شروع از نقاطي كه به نقاط اتصال خارجي متصل هستند شروع مي‌شود و بر روي مسير شكل حركت داده مي‌شود تا كد زنجيره فريمن حاصل شود. در صورتيكه نقاط اتصال خارجي وجود نداشت از نقاط راسي(نقاط با حداكثر يک همسايگي) شروع به حركت مي‌شود. در شكل 2-17 تجزيه حرف "ع" به عناصر NAPE نمايش داده شده است.



شكل 2-17 تجزيه حرف "ع" به عناصر NAPE [10]

در نهايت پس از شناسايي NAPE ها، گرامر مربوطه نوشته مي شود. از اين پس با پارس‌كردن گرامر مي‌توان نويسه را شناسايي كرد[12]. در جدول 2-3 بعضي از قواعد پلكس تعريف شده در پايگاه قواعد براي حروف و ارقام فارسي نمايش داده شده است.

جدول 2-3 بعضي از قواعد پلكس تعريف شده در پايگاه قواعد براي حروف و ارقام فارسي[10]

<2>:<NAPE_RIGHT_DOWN_CORNER> <PRIMITIVE_VERTICAL_RIGHT> <PRIMITIVE_VERTICAL_RIGHT> (121)
<ب > : *<NAPE_PLATE> <PRIMITIVE_DOT> (41)*
<پ>: *<NAPE_PLATE> <PRIMITIVE_DOT> <PRIMITIVE_DOT> <PRIMITIVE_DOT> (4111)*
<ت> : *<NAPE_PLATE> <PRIMITIVE_DOT> <PRIMITIVE_DOT> (311)*
<ح> : *<NAPE_HAT> <NAPE_LEFT_CURVE> <PRIMITIVE_HORIZONTAL_UP> (122)*
<ج >: <NAPE_HAT> <NAPE_LEFT_CURVE> <PRIMITIVE_HORIZONTAL_UP> <PRIMITIVE_DOT>(1220, 0301)
<ش >: <NAPE_RIGHT_DOWN_CORNER> <PRIMITIVE_VERTICAL_RIGHT>

<PRIMITIVE_HORIZONTAL_UP> <PRIMITIVE_VERTICAL_UP> <NAPE_BOWL> <PRIMITIVE_DOT><PRIMITIVE_DOT> <PRIMITIVE_DOT> (12100000, 00221000, 30000111)
<ع > : *<NAPE_LEFT_DOWN_CORNER> <PRIMITIVE_HORIZONTAL_UP>* *<NAPE_LEFT_DOWN_CORNER> (122)*

2-3-2 شناسايي با استفاده از شبكه عصبي

الف. بازشناسي برخط حروف مجزاي فارسي با شبكه عصبي

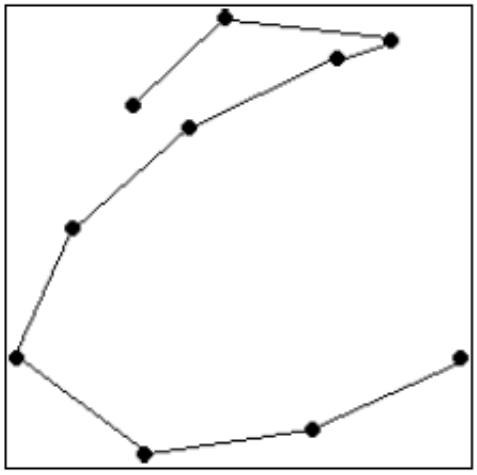
همانطور كه در فصل پيش گفته شد بازشناسي نويسه بر اساس الگوي ورودي به دو دسته برخط و برون‌خط تقسيم مي‌شود. در اين بخش نمونه‌اي از يک شناسايي برخط كه از شبكه‌عصبي بهره گرفته شده است، معرفي مي‌گردد. بعضي از سيستم‌هاي بازشناسي برخط فقط نويسه‌هاي خاص را بازشناسي مي‌كنند. از اين سيستم‌ها بيشتر روي كامپيوترهاي كوچک كه صفحه كليد كاملي ندارند و دقت بازشناسي زيادي لازم است، استفاده مي‌شود. در روش ارائه شده حروف فارسي به 12 گروه تقسيم مي‌شوند. فرض بر اين است كه نويسنده ابتدا بدنه‌اصلي حرف را در يک حركت قلم نوشته، سپس علائم آنرا مي‌نويسد[13]. در جدول 2-4 حروف فارسي به 12 گروه بر اساس علائم تقسيم بندي شده‌اند.

جدول 2-4 تقسيم حروف به 12 گروه بر اساس علائم آنها [13]

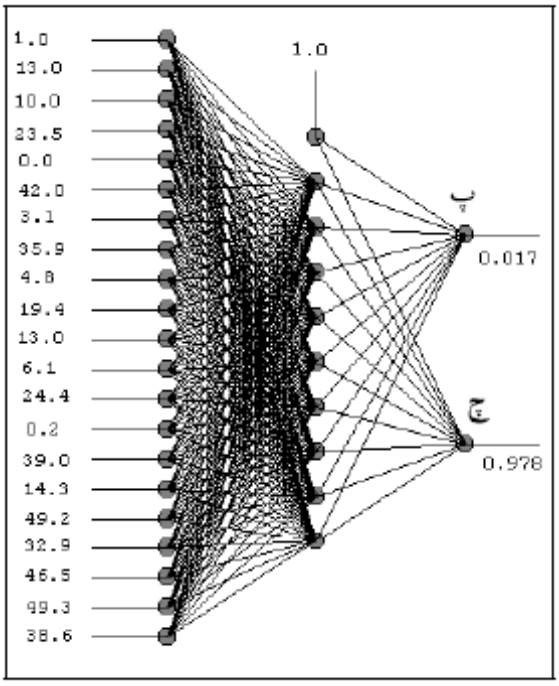
شماره گروه	حروف موجود در گروه	نوع علامت
1	آ	يک مد بالا
2	ا ء ح د ر س ص ع ل م و ه ي	بدون علامت
3	ب ج	يک نقطه پايين
4	خ ذ ز ض غ ف ن	يک نقطه بالا
5	ت ق	دو نقطه بالا
6	ث ژ ش	سه نقطه بالا

7	پ چ	سه نقطه پايين
8	ک	يک سرکش بالا
9	گ	دو سرکش بالا
10	ط	يک دسته بالا
11	ظ	يک دسته و يک نقطه بالا

بازشناسايي بدنه با شبکه عصبي انجام مي‌شود. براي هرکدام از گروه‌هاي 2 تا 8 يک شبکه عصبي سه لايه در نظر گرفته شده است. در لايه اول 21 نرون، در لايه مياني 10 نرون و در لايه آخر به تعداد کلاسهاي هرگروه نرون وجود دارد. در لايه اول يکي از نرونها بایاس و 20 نرون ديگر براي ويژگيهاي استخراج شده از بدنه حرف در نظر گرفته شده است. بدنه هر حرف از نظر نقطه به 10 و از نظر اندازه به 50x50 نرمال مي‌شود. مختصات اين 10 نقطه به عنوان ورودي شبکه عصبي داده مي‌شوند. در مرحله بعد شناسايي اجزاي کوچک نسبت به بدنه مي‌باشد. ابتدا تشخيص موقعيت جز کوچک نسبت به بدنه مي‌باشد که در بالا يا پائين قرار دارد که اين کار با استفاده از مختصات رديابي مي‌شود. سپس شناسايي نوع جز کوچک مي‌باشد که شامل يک نقطه، دو نقطه، سه نقطه، سرکش بزرگ، سرکش کوچک، دسته و مد است. اين کار توسط شبکه عصبي MLP سه لايه با الگوريتم يادگيري پس انتشار خطا صورت مي‌گيرد. اين شبکه شامل 24 ورودي و 10 نرون در لايه مخفي انجام مي‌باشد. شکل 2-18 نمايش نقاط حساس و يا به عبارتي نرمال شدن نويسه را نمايش مي‌دهد. شکل 2-19 ساختار عصبي مربوط به گروه A را نشان مي‌دهد.



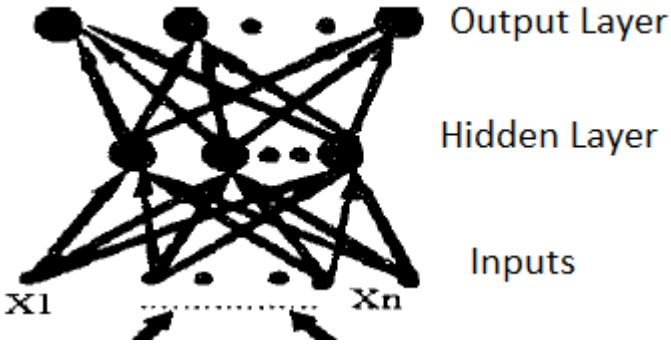
شکل 2-18 نقاط نرمال شده[13]



شكل 2- 19 ساختار شبكه عصبي مربوط به گروه A[13]

ب. شناسايي حروف دستنويس فارسي با استفاده از شبكه عصبي

در اين روش [14] شناسايي حروف دستنويس فارسي با استفاده از شبكه عصبي مورد بررسي قرار گرفته است. در ابتدا 3 پيش‌پردازش اصلي بر روي تصوير صورت مي‌گيرد كه به ترتيب عبارتند از : نويزگيري، نرمال‌سازي و نازك‌سازي. براي نويزگيري از فيلتر گوسي استفاده شده است. سپس تصاوير به ابعاد 30x30 نرمال مي‌شود و در آخر نيز عمل نازك‌سازي صورت مي‌گيرد. پس از انجام كارهاي مربوط به پيش‌پردازش، شناسايي نويسه با استفاده از شبكه عصبي صورت مي‌گيرد. شبكه عصبي پيشنهادي، پرسپترون چند لايه مي‌باشد كه از يک لايه مخفي و الگوريتم پس انتشار خطا براي آموزش در آن استفاده شده است. همچنين در اين مقاله آناليزي براي تعيين تعداد نرون‌هاي لايه مخفي صورت گرفته است. در آزمايشات مختلف مشخص شده است كه استفاده از الگوريتم پس انتشار خطا براي آموزش به نسبت ديگر الگوريتم‌ها از نرخ بهتري برخوردار مي‌باشد و به همين دليل در اين مقاله از آن استفاده شده است. در شكل 2-20 شبكه عصبي پرسپترون چند لايه مورد استفاده نمايش داده شده است.



شكل 2-20 شبكه عصبي پرسپترون چند لايه مورد استفاده در [14]

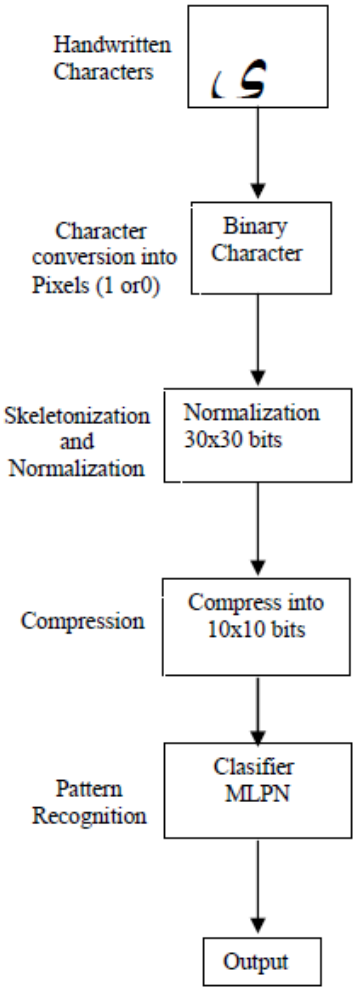
تابع فعال سازي براي نرون j به صورت (2-2) تعريف شده است كه در آن O_i خروجي واحد iام و W_{ij} وزن بين واحد i و j مي‌باشد.

$$f_j(x) = \frac{1}{1+e^{-net}} \qquad net = \sum W_{ij} O_i \tag{2-2}$$

براي به روز رساني وزن‌ها، قانون دلتا كه در آن بايد مربع اختلاف بين پاسخ شبكه و پاسخ واقعي حداقل باشد، مورد استفاده قرار گرفته است. اين قانون در (2-3) نمايش داده شده است.

$$E = \frac{1}{2}(\sum(D_{pk} - O_{pk}))^2 \tag{2-3}$$

كه در آن D_{pk} پاسخ شبكه و O_{pk} پاسخ واقعي(مقدار هدف) براي واحد kام و آموزش زوج p مي‌باشد. در شبكه استفاده شده نرخ يادگيري 0.2 و ضريب momentum، 0.1 در نظر گرفته شده است. شبكه آنقدر آموزش مي‌گيرد تا مربع خطا (E) به كمتر از 0.05 برسد.
در شكل 2-21 دياگرام مربوط به الگوريتم پيشنهادي نمايش داده شده است.



شكل 2-21 دياگرام روش شناسايي پيشنهادي [14]

آزمايشات انجام گرفته بر روي 250 نمونه حروف جمع‌آوري شده از 25 نفر در گروه‌هاي سني مختلف زن و مرد صورت گرفته است. جدول 2-5 نتايج اين آزمايشات با تعداد نرون‌هاي متفاوت در لايه مخفي را نشان مي‌دهد.

جدول 2-5 نتايج آزمايشات بر روي تعداد نرون‌هاي مختلف در لايه مخفي[14]

Input of the MLPN	No. of hidden units	No. of iteration	Training time (s)	Recognition Accuracy (%)	
				Training Data	Test Data
30x30	12	200	1625	100	80
	24	**200**	**3125**	**100**	**85**
	36	200	4750	100	80

همانطور كه در جدول 2-5 نشان داده شده است افزايش تعداد نرون در لايه مخفي درصد شناسايي را بهتر نكرده است و 24 نرون در لايه مخفي آمار بهتري به نسبت 36 نرون در لايه مخفي داشته است.
در نهايت نيز درصد شناسايي 85% ارائه شده است.

2-3-3 شناسايي آماري

الف. يک روش آماري براي شناسايي حروف فارسي

در اين روش[15] فرض بر اين است كه بريدگيهاي نامطلوب در ساختار حروف وجود ندارد و نقاط از بدنه حروف جدا هستند. ابتدا حروف بر اساس بخش‌هاي موجود در آن به 5 گروه تقسيم مي‌شوند، سپس بازشناسي در هر گروه بر اساس ويژگي‌هاي گشتاوري مناسب با آن گروه و با استفاده از طبقه بندي بيز صورت مي‌گيرد. در نهايت پس‌پردازش لازم در مورد كلاس‌هاي مشابه انجام مي‌شود. تعيين كلاس حرف ورودي در 3 مرحله گروه بندي، بازشناسي اصلي در گروه و پس‌پردازش براي كلاس‌هاي مشابه صورت مي‌گيرد.

- گروه‌بندي

با استفاده از الگوريتم برچسب زدن به مولفه ها بر اساس همسايگي 8 بر مبناي تعداد عناصر تصوير و با حذف آنهايي كه از يک مقدار آستانه كوچكتر باشند، گروه‌بندي صورت مي‌پذيرد. در جدول 2-6 نحوه گروه‌بندي بر اساس تعداد اجزا نشان داده شده است.

جدول 2-6 نحوه گروه‌بندي بر اساس تعداد اجزا [15]

گروه1	گروه2	گروه3	گروه4	گروه5
ا ح د	ب ج	پ ت	پ ت	پ
ر س ص	خ ذ	ث چ	ث چ	ث
ط ع ک	ز ض	ژ ش	ژ ش	چ

گ ل م	ظ غ	ق ک	ظ ق	ژ
ل و ه	ف ن	گ ط		ش
ي				

- بازشناسي اصلي در گروه

از گشتاورهاي نرمال نسبت به مكان استفاده شده است. اين ويژگيها طبق روابط زير مي‌باشند. در 2-4 نحوه محاسبه گشتاور مركزي و در 2-5 نحوه محاسبه گشتاور نرمال از مرتبه p+q مشخص شده است.

$$\mu_{pq} = \sum_{(m,n)\epsilon R} \sum (m-\overline{m})^p (n-\overline{n})^q \qquad (2\text{-}4)$$

$$m_{pq} = \frac{\mu_{pq}}{\mu_{00}^{(\frac{p+q+2}{2})}} \qquad (2\text{-}5)$$

در طبقه بندي اصلي از قاعده بيز استفاده شده است. بردار ميانگين و ماتريس كواريانس براي هر كلاس با استفاده از نمونه هاي تمرين تخمين زده شده است.

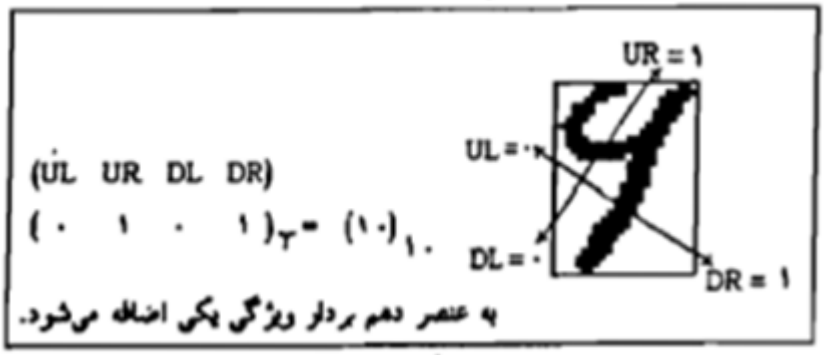
- پس‌پردازش براي كلاس‌هاي مشابه

به منظور كم كردن ميزان خطا، پس‌پردازشهاي معيني در مجموعه كلاس‌هاي مشابه صورت گرفته است. براي هر كلاس از تكنيک خاصي استفاده شده است به طور مثال براي تفكيک "ح" و"ع" از مكان‌هاي مشخصه در [16] استفاده شده است.

2-3-4 شناسايي بر اساس روش‌هاي ديگر

الف. روش مكانهاي مشخصه و زنجيره ماركف

در اين روش ابتدا ويژگي‌هاي تصوير با استفاده از مكان‌هاي مشخصه به روش عمودي-افقي و روش مورب استخراج مي‌شود. تصوير به يک ماتريس 9x9 تقسيم بندي مي‌شود سپس براي هر سلول تعداد تقاطع‌ها به صورت عمودي-افقي و مورب بدست مي‌آيد. بدين ترتيب يک بردار ويژگي براي هر سلول براي مكان‌هاي مشخصه عمودي- افقي و مورب بدست مي‌آيد[17]. شكل 2-22 به خوبي بيانگر نمايش مكان‌هاي مشخصه مورب مي‌باشد.



شكل 2-22 مكانهاي مشخصه مورب [17]

در نهايت يک بردار ويژگي اصلي به طول 81 در نظر گرفته مي‌شود كه با توجه به بردارهاي بدست آمده در متدهاي مكان‌هاي مشخصه مقداردهي مي‌شود. روش مقداردهي بردار ويژگي اصلي بدين صورت است كه بردار ويژگي مورب يا عمودي 4 بعدي است و در مبناي 3 فرض مي‌شود. آنرا به مبناي 10 برده، عدد بدست آمده انديس بردار ويژگي اصلي مي‌باشد سپس يک واحد به محل انديس بدست آمده اضافه مي‌شود. طبقه بندي بيز با فرض وابستگي مولفه‌هاي بردار ويژگي[18] به صورت زير است:

مولفه هاي بردار ويژگي، يک زنجيره ماركف مرتبه 1 را تشكيل مي‌دهند. با استفاده از بردارهاي ويژگي نمونه‌هاي آموزشي براي هر كلاس يک مدل ماركف بدست مي‌آيد. مدل ماركف براي هر كلاس داراي 10 حالت و 81 مشاهده است. سپس احتمال گذر از حالت i در مشاهده t به حالت j در مشاهده t+1 محاسبه مي‌شود. با ضرب كردن بردار ويژگي نمونه مجهول در احتمال‌هاي گذر متناظر با مشاهدات (2-6) براي مدل ماركف هر كلاس، احتمال اينكه نمونه مجهول در آن كلاس مشاهده شود، بدست مي‌آيد.

ماتريس ويژگي نمونه مجهول — ماتريسهاي ويژگي متوسط

$$X = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad X_1 = \begin{bmatrix} \frac{3}{4} & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{3}{4} \end{bmatrix} \quad X_2 = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{3}{4} & \frac{1}{4} \end{bmatrix} \tag{2-6}$$

$$P(x|\omega_1) = \frac{1}{4} * \frac{1}{4} * \frac{3}{4} = \frac{3}{64}$$
$$P(x|\omega_2) = \frac{3}{4} * \frac{3}{4} * \frac{1}{4} = \frac{9}{64} \qquad x \in \omega_2$$

ب. کاربرد تطابق شکل در بازشناسي ارقام

کارهاي زيادي در زمينه توصيف شکل بر اساس کانتور انجام شده است[16]. چنگ و يان[19] ويژگي‌هايي را از کانتور ارقام استخراج کرده‌اند. اين ويژگيها عبارتند از: توصيفگر فوريه کانتور، ميانگين و واريانس فاصله پيکسل‌هاي روي کانتور تا مرکز ثقل آن و همچنين طول و مساحت نرمال شده کانتور. در روش ديگري چاکراوارتي و کمپلا [20] بر اساس نقاط بحراني، توصيفي رياضي از کانتور حروف و ارقام ارائه دادند. نقطه بحراني، نقطه‌اي است که مشتق اول تابع کانتور در آن صفر است. در اين روش کانتور هر شکل بر اساس نقاط بحراني به وسيله گرافي نمايش داده مي‌شود. اين روش که بر کانتور شکل مبتني است سه مرحله دارد:

1. حل مساله تناظر بين نقاط کانتور شکل اول با نقاط کانتور شکل دوم.

2. استفاده از اين تناظر براي تعريف تبديل هندسي که دو شکل را بر روي هم قرار دهد.

3. محاسبه فاصله بين دو شکل به صورت مجموع خطاهاي تطابق نقاط متناظر دو شکل و جمله‌اي که ميزان پيچيدگي اين تبديل هندسي را نشان مي‌دهد.

در اين روش براي هر نقطه نمونه برداري شده بر روي کانتور شکل، توصيفگري بر اساس توزيع مکاني نقاط ديگر کانتور بدست مي‌آيد. اين نمونه‌گيري بر اساس لبه‌ياب Canny مشخص مي‌شود. تعداد نقاط نمونه برداري شده بر روي کانتور، بسته به دقت وکاربرد متفاوت است. هرچه تعداد اين نقاط بيشتر باشد توصيف دقيقتري از شکل بدست مي‌آيد ولي به همان نسبت حجم کار برنامه افزايش مي‌يابد. اگر مجموعه بردارهايي را که از يک نقطه بر روي کانتور به تمام نقاط ديگر کانتور ترسيم مي‌شود در نظر بگيريم، بردارهايي را داريم که توصيف تمام شکل را نسبت به نقطه مرجع نشان مي‌دهند. براي هر نقطه بر روي شکل اول مي‌توان هيستوگرام را بر اساس فاصله و زاويه نسبي پاره خط واصل آن نقطه با n-1 نقطه ديگر روي کانتور تعريف کرد. براي محاسبه اين هيستوگرام از مختصات لگاريتمي قطبي استفاده مي‌شود. براي تعيين ميزان شباهت دو شکل ابتدا بر اساس توصيفگرها، تناظري يک به يک بين نقاط نمونه برداري شده روي کانتور شکل اول با نقاط کانتور شکل دوم بدست مي‌آيد. براي پيدا کردن اين تناظر از دو روش رياضي مجاري و يونکر مي توان استفاده کرد. جمع فواصل بين نقاط متناظر در دو شکل، معياري براي عدم شباهت آنهاست. در نهايت تبديلي تعريف مي‌شود که نقاط کانتور شکل اول را بر روي شکل دوم قرار مي‌دهد. ميزان پيچيدگي اين تبديل، معيار ديگري براي عدم شباهت دو شکل است[21].

ج. استخراج ويژگي توسط گراديان بهبود يافته در ارقام تنها فارسي

ويژگي گراديان با استفاده از جهت‌هاي محلي، به خوبي مشخصه‌هاي نمونه هاي داخل يک کلاس و تغييرات بين کلاسي را مدل مي‌کند. اين ويژگي براي تصاوير خاکستري توسعه يافته و براي استخراج آن، لازم است تصاوير نرمال شوند. از اين رو براي استفاده در مورد تصاوير دودويي، ابتدا بايد آنها را خاکستري کرده و به اندازة نرمال درآورد.

نحوة استخراج ويژگي گراديان روشنايي به ترتيب زير است: ابتدا تصوير ورودي به اندازه (مثلا40x40) نرمال مي‌شود. سپس با اعمال يک فيلتر ، اطلاعات گراديان براي هر پيکسل به دست مي‌آيد. اين فيلتر مي‌تواند عملگرهايي مثل روبرتز، سوبل و يا کريش باشد. عملگر کريش از اين نظر که مولفة گراديان در چهار راستاي افقي، عمودي، قطر اصلي و قطر فرعي مي دهد، از بقيه پيچيده‌تر است. عملگرهاي روبرتز و سوبل تنها دو مولفة گراديان در اختيار مي گذارند.

براي افزايش سرعت استخراج ويژگي گراديان، به گونه‌اي که کارايي آن نيز حفظ شود، سه تغيير عمده در روش استخراج ويژگي داده شده است که با اين تغييرات علاوه بر افزايش سرعت استخراج تا دو برابر، دقت بازشناسي نيز به ميزان 0.032% افزايش يافت .اين روش با 4 تغيير انجام مي‌شود:

الف) به جاي توليد هشت تصوير و نمونه گيري از تک تک آنها، روش نمونه‌گيري به گونه‌اي تغيير داده شده که تنها با يک بار جاروب کردن کل تصوير از بالا به پايين و چپ به راست، بتوان ويژگي‌ها را استخراج کرد. به اين ترتيب که تصوير به 16 بلوك تقسيم و در هر بلوك مقادير شدت گراديان در هر يك از راستاهاي تعريف شده با هم جمع مي‌شوند. در نتيجه در هر بلوك به تعداد راستاهاي گراديان تعريف شده ويژگي وجود دارد، که تعداد کل ويژگي‌ها از حاصلضرب تعداد بلوك‌ها در تعداد راستاها به دست مي‌آيد.

ب) اعمال فيلتر روبرتز به تصوير ورودي .اين فيلتر جهت کاهش حجم محاسباتي و نيز به دست دادن تعداد زواياي کم براي تصاوير دوسطحي استفاده شده است.

ج) تقسيم کردن تصوير ورودي به 4x4 بلوک مساوي که منجر به 16 بلوک 10x10 مي‌شود.

د) در هر بلوک مقدار شدت گراديان به ازاي هر ۹ جهت ذکر شده با هم جمع مي شود. به ازاي جهت تعريف نشده، چون شدت گراديان همواره صفر مي‌شود و مقدار خاصي به دست نمي‌دهد، گراديان برابر 0.1 تعريف شده که اين مقدار به صورت تجربي و با توجه به مقادير گراديان براي ساير جهت‌ها به دست آمده است .

2-4 نتيجه‌گيري

همانطور كه ملاحظه شد بدون شک پيش‌پردازش براي تصوير، قبل از شناسايي نويسه الزامي است. در اين فصل پيش‌پردازش‌هاي خاكستري، سياه و سفيد، تغيير اندازه، نازک‌سازي، نويزگيري و همبندكردن بررسي شد و در هركدام نحوه عملكرد و نقاط ضعف آن ذكر شد. با توجه به مهم بودن پيش‌پردازش بايد الگوريتم‌هاي مورد استفاده كمترين ضعف را داشته باشند. به طور مثال در زمينه نازک‌سازي و همبندكردن شكل هنوز مشكلات فراواني وجود دارد كه بايد رفع گردد.

همچنين در اين فصل الگوريتم‌ها و روشهاي متفاوتي جهت بازشناسي نويسه‌هاي فارسي و عربي بررسي شدند. هركدام از آنها داراي معايب و مزاياي از لحاظ درصد تشخيص، سرعت، نيازمند به پيش‌پردازش‌هاي متفاوت و غيره بودند. در مجموع نمي‌توان به طور صريح مشخص كرد كه چه روشي از ديگر روش‌ها بهتر است ولي نتيجه كلي كه مي‌توان گرفت اين است كه الگوريتم و روشي بهتر عمل خواهد كرد كه به روش حل مسئله انسان نزديكتر باشد. به نظر مي‌آيد توسط فرمول‌هاي پيچيده رياضي همانند قواعد بيز، ماركف و فوريه بتوان درصد خوبي از تشخيص را بدست آورد ولي كاملا واضح است كه انسان براي تشخيص از هيچكدام از فرمولهاي پيچيده رياضي استفاده نمي‌كند.

در فصل آتي درباره روش پيشنهادي و نحوه عملكرد الگوريتم شناسايي نويسه صحبت خواهد شد.

فصل سوم

روش پیشنهادي

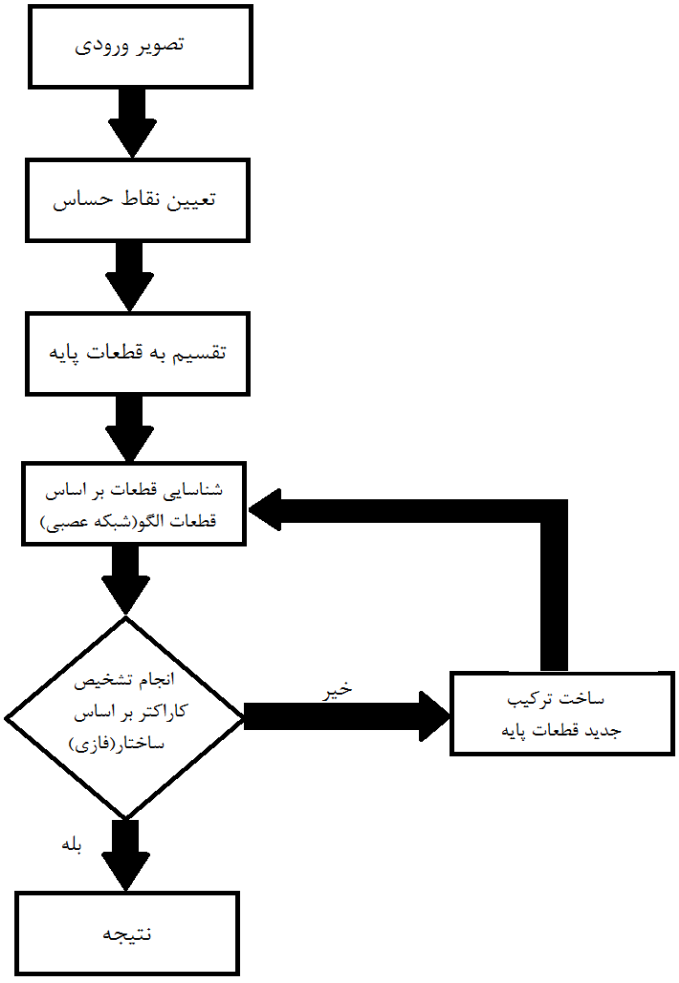
3-1 مقدمه

در اين فصل روشي نوين براي شناسايي نويسه‌هاي تنها دستنويس فارسي ارائه مي‌گردد. در اين روش پس از اعمال تمامي پيش‌پردازش‌هاي لازم بر روي تصوير كه در فصل دوم بيان شد، نويسه به موتور بازشناسي ارسال مي‌گردد. موتور بازشناسي ارائه شده در اين فصل با استفاده از تركيب روش‌هاي آموزشي و ساختاري به كمک ابزارهاي شبكه عصبي و منطق فازي به شناسايي نويسه‌هاي دستنويس تنها فارسي مي‌پردازد. در واقع اين موتور بازشناسي را مي‌توان روشي نوين براساس تركيبي از روش هاي ساختاري و شبكه عصبي مبتني بر منطق فازي براي شناسايي حروف و ارقام دستنويس فارسي معرفي كرد. در ادامه به توضيح هر قسمت از اين موتور بازشناسي پرداخته خواهد شد.

3-2 كليات روش پيشنهادي

پيش از ورود تفصيلي به جزئيات روش پيشنهادي، كليات آن را مد نظر قرار مي‌دهيم. آنچه در اين بخش آمده، بدون ذكر جزئيات، بدون لحاظ نمودن دقت لازم و بدون در نظر گرفتن «پيش‌پردازش» يا «پس‌پردازش» مي‌باشد.

در اين روش، ابتدا با در نظر گرفتن يک سري قوانين و شرايط، نويسه دستنويس به قطعات مختلف تقسيم مي‌شود (هر قطعه شامل اطلاعات مكاني نقاطي خواهد بود كه آن را تشكيل داده‌اند)؛ اين تقسيم‌بندي از طريق تعيين نقاطي از شكل – موسوم به نقاط حساس – صورت مي‌گيرد. پس از آن، هر قطعه، توسط يک شبكه عصبي شناسايي مي‌شود؛ اين شناسايي بر اساس قطعات الگويي كه از پيش تعريف و به شبكه عصبي آموزش داده شده انجام خواهد شد. (هر قطعه الگو، يكي از اجزاء تشكيل دهنده «ساختار» نويسه خواهد بود). در نهايت اطلاعات به دست آمده به يک بخش تشخيصي وارد مي‌شود كه در آن بر اساس الگوي ساختاري تعريف شده براي هر نويسه (بر اساس نوع و ويژگي‌هاي قطعات تشكيل دهنده آن)، شباهت دستنويس به هريک از نويسه‌ها مردود يا مورد قبول واقع مي‌شود. در صورت عدم شناسايي نويسه ورودي توسط تمام الگوها در اين مرحله، برخي از قطعات شناسايي شده نويسه، بر اساس قوانيني معيني با يكديگر تركيب شده و قطعات جديد همين مسير را طي مي‌كنند. اين چرخه تا زمان رسيدن به نتيجه – و يا اتمام حالات ممكن – ادامه مي‌يابد. در شكل 3-1 فلوچارت الگوريتم پيشنهادي نمايش داده شده است.

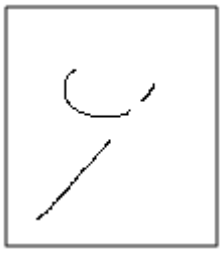


شکل 3-1 فلوچارت الگوريتم پيشنهادي

3-3 شناسايي نقاط حساس و تجزيه شكل

نقاط حساس به نقاطي از نويسه گفته مي‌شود كه مرز جدا كننده قطعات الگوي آن نويسه باشند. براي مثال عدد "6" را در نظر مي‌گيريم. يک "6" ايده‌آل را مي‌توان به چند صورت تقسيم‌بندي نمود. براي مثال دو روش زير را براي تعريف "6" ايده‌آل در نظر مي‌گيريم:

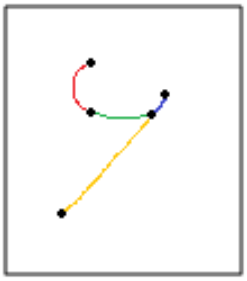
1- تركيبي از يک «نيم‌دايره» و دو «خط راست» (شكل 3-2)
2- تركيبي از دو «ربع دايره» و دو «خط راست» (شكل 3-3)



شكل 3-2 تقسيم‌بندي اول از نويسه"6"

شكل 3-3 تقسيم‌بندي دوم از نويسه "6"

اما براي اينكه بتوانيم «نقاط حساس» نويسه "6 "را بدست آوريم، بايد بدانيم كه قطعات الگوي ما چه قطعاتي هستند. اين قطعات در بخش‌هاي بعد به طور كامل معرفي مي‌شوند ولي براي پيشبرد بحث فرض مي‌كنيم خط و ربع دايره (در جهات مختلف) جزو قطعات الگو هستند. در اين صورت واضح است كه تركيب شماره 1 صحيح نيست و بنابراين بايد نقاط حساس را به گونه‌اي اختيار نماييم تا قطعات حاصله، تطابق هرچه بيشتري با قطعات الگو داشته باشند. در شكل 3-4 قطعه‌بندي صحيح نويسه "6" نمايش داده شده است.

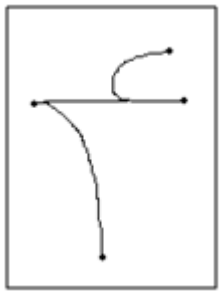


شكل 3-4 تصوير قطعه‌بندي صحيح نويسه "6"

نتيجه خلاصه‌اي كه از مباحث فوق حاصل مي‌گردد اين است كه «الگوريتم يافتن نقاط حساس، تابعي از شكل قطعات الگو است». روش يافتن نقاط حساس نويسه براي الگوهاي تعيين شده در اين مقاله به شرح زير است:

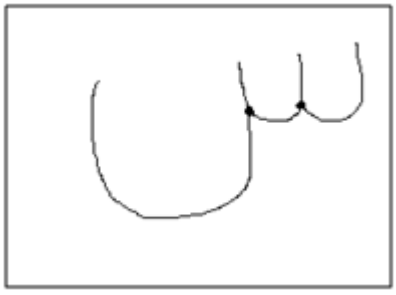
نقاط حساس، خود به چهار گونه نقاط ابتدايي، نقاط تقاطع، نقاط شكستگي و نقاط كمان تقسيم مي‌شوند.

نقاط ابتدايي: به نقاطي گفته مي‌شود كه تنها يک نقطه در همسايگي خود داشته باشند كه از روي نويسه (البته نويسه نازک‌سازي شده) به خوبي به دست مي‌آيند. اين نقاط براي نويسه "4" در شكل 3-5 نمايش داده شده‌اند.



شكل 3-5 نقاط ابتدايي در عدد "4 "

نقاط تقاطع: نقاطي هستند كه بيش از دو نقطه در همسايگي آن‌ها وجود داشته باشد كه از روي نويسه به خوبي به دست مي‌آيند. اين نقاط براي نويسه "س" در شكل 3-6 نمايش داده شده‌اند.



شكل 3-6 نقاط تقاطع در حرف "س"

نقاط شكستگي: نقاط شكستگي نقاطي هستند كه اولا جزو نقاط «ابتدايي» و «تقاطع» نبوده و ثانيا در جهت عمومي حركت شكل، پيش و پس از آن‌ها تغيير زيادي مي‌كند. يک نقطه شكستگي براي عدد "8" در شكل 3-7 نمايش داده شده است.

شكل 3-7 نقطه شكستگي در عدد "8"

براي يافتن اين نقاط تعريف 3-1و 3-2 را ارائه مي‌دهيم:

$$m(a,b,c) = a + c - 2b \qquad (3\text{-}1)$$

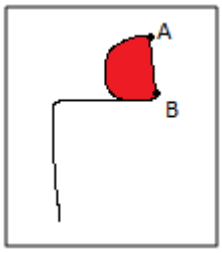
$$dB_i = m(X_{i-\alpha},\ X_i, X_{i+\alpha})^2 + m(Y_{i-\alpha}, Y_i, Y_{i+\alpha})^2 \qquad (3\text{-}2)$$

در محاسبه dB، Xها و Yها مولفه‌هاي نقاط نويسه، i ايندكس نقاط متصل (با توجه به اينكه نقاط همسايه در صورت امكان داراي ايندكس‌هاي پشت‌سرهم باشند) و α ميزان حساسيت تصميم‌گيري را تعيين مي‌كنند. در اين شرايط، به وضوح ديده مي‌شود كه بزرگ شدن dB_i با افزايش ميزان شكستگي در نقطه iام (با تعريفي كه از شكستگي ارائه شد) رابطه مستقيم دارد.

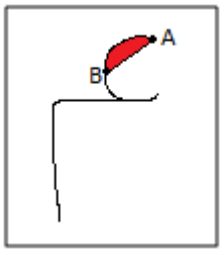
نقاط كمان: نقاط كمان به نقاطي گفته مي‌شود كه اولا جزو نقاط «ابتدايي»، «تقاطع» و «شكستگي» نبوده و ثانيا انتخاب آن‌ها به عنوان مرز جداكننده، باعث تشكيل قطعات الگو در شكل مي‌شود – و البته بدون انتخاب آن‌ها اين امر ميسر نيست. اين نقاط معمولا نقاط مياني كمان‌هاي طولاني با ميزان خميدگي بالا هستند. در شكل 3-8 يک نقطه شكستگي و دو نقطه كمان در حرف "ح" نمايش داده شده است.

شكل 3-8 تصوير نقاط كمان و يک نقطه شكستگي در حرف "ح"

براي يافتن اين نقاط از نسبت مساحت محصور بين كمان و وتر و طول وتر استفاده مي‌كنيم. در هر دو شكل زير وتر AB (با طول |AB|) مساحتي برابر با S را بين خود و كمان AB محصور نموده است.



شكل 3-9 وتر با زاويه بزرگتر



شكل 3-10 وتر با زاويه كوچكتر

به راحتي مي‌توان نشان داد كه:

1- در شكلي مانند شكل 3-9 (كه وتر زاويه بزرگتري از كمان را پوشش داده است) نسبت |S/|AB بيشتر از شكل 3-10 است.
2- اين نسبت در يک كمان ايده‌آل در وضعي كه A و B بسيار نزديک به هم باشند داراي حدي برابر صفر است و با دور شدن A و B به صورت يكنوا افزايش مي‌يابد.

با اين حساب، چنانچه براي نسبت |S/|AB يک آستانه مشخص (بر حسب |AB|) در نظر بگيريم مي‌توان نقاط كمان را شناسايي و تعيين نمود.

3-4 شناسايي اجزاي كوچک توسط شبكه عصبي

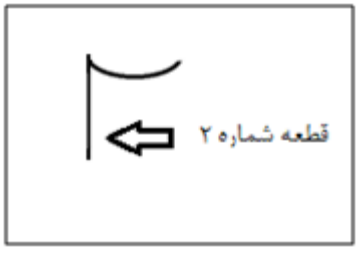
پس از آنكه نقاط حساس كه در بخش 3-2 ذكر شد مشخص شدند، مي‌توان نويسه را به اجزاي كوچكتري تجزيه نمود. در نتيجه شروع و انتهاي هر جز تشكيل شده از نقاط حساس خواهد بود. هر جز را مي‌توان در يک دسته هندسي قرار داد. يک دسته از اين اجزا به صورت خط‌هاي راست و دسته ديگر به صورت منحني مي‌باشند. در هر دو دسته چرخش اجزا در زواياي مختلف درنظر گرفته شده است. در جدول 3-1 كليه اجزا ديده مي‌شوند و به هر يک، شماره يكتايي براي شناسايي داده مي‌شود.

جدول 3-1 تقسيم‌بندي اجزا بر اساس مدل هندسي

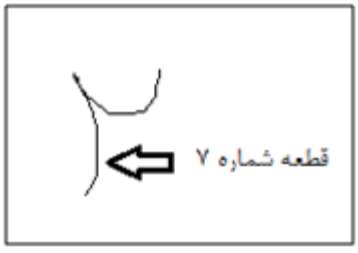
4	3	2	1
8	7	6	5

12	11	10	9

نكته اينكه المان‌هاي ذكر شده تنها براي اعداد و البته براي بدنه حروف كاربرد دارند اما براي شناسايي حروف به المان‌هاي مربوط به نقطه‌ها نيز نياز مي‌باشد. به بيان ديگر مي‌توان گفت يک نويسه متشكل از به هم پيوستن اجزاي هندسي مي‌باشد كه در جدول 3-1 نمايش داده شد. پس از تجزيه نياز مي‌باشد كه هر جز به تنهايي شناسايي شود. در طرح ارائه شده براي شناسايي اجزاي كوچک استفاده از شبكه عصبي پيشنهاد مي‌شود. در اين طرح بر خلاف كارهاي انجام شده در زمينه شناسايي نويسه‌هاي دستنويس به كمک شبكه عصبي كه عمل بازشناسي را كاملا بر عهده شبكه عصبي قرار مي‌دهند، شبكه عصبي فقط همانند يک ماشين براي شناسايي 12 جز هندسي مورد استفاده قرار مي‌گيرد. نكته‌اي كه در اين باره وجود دارد، تشابه بسياري از قطعات مذكور، با يكديگر است كه اين تشابه نه تنها يک مولفه تاثيرگذار در تعيين حساسيت تشخيص اجزاء توسط شبكه عصبي است، بلكه مهمتر از آن، بخشي از واقعيت سيستم را توضيح مي‌دهد. بدين معني كه جايگزين شدن يكي از اين قطعات با قطعات مشابه آن - در هنگام نوشتن نويسه دستنويس - بسيار محتمل است و بنابراين حتي اگر در جايي از نويسه دستنويس، مثلا قطعه شماره 7 وجود داشته باشد و شبكه عصبي نيز، آن را به درستي تشخيص دهد، ممكن است مقصود نويسنده قطعه شماره 2 باشد كه بنا به دلايلي ازقبيل عادت و يا سرعت نوشتن، به اين شكل درآمده است. در شكل 3-11 نمونه‌اي استاندارد از نويسه "2" و در شكل 3-12 نمونه‌اي دستنويس از نويسه "2" نمايش داده شده است.



شكل 3-11 نويسه استاندارد 2

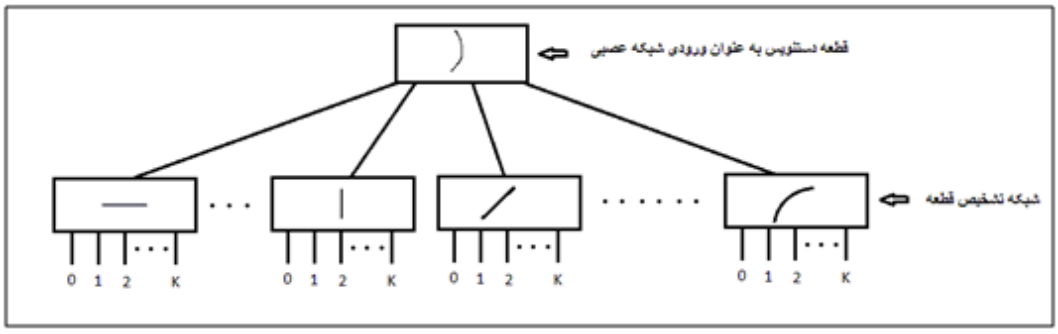


شكل 3-12 نويسه دستنويس 2

يكي از روش‌ها براي حل اين موضوع اين است كه تغييراتي در خروجي شبكه عصبي ايجاد گردد. بدين ترتيب كه شبكه عصبي، به جاي اينكه در نهايت يكي از قطعات را از ميان قطعات 12گانه برگزيند، داراي خروجي‌اي به شكل برداري 12 بعدي باشد كه در هر بعد، ميزان شباهت تصوير قطعه مورد نظر با شكل مرجع قطعه را تعيين مي‌كند (بر

اساس آموزش‌هاي داده شده)؛ براي حل مشكل فوق‌الذكر، روش ديگري نيز (منتج به بردار 12 بعدي با همين تعريف) در قسمت پيشنهادات ارائه شده كه ميزان آموزش كمتري را نياز دارد و البته دقت آن نيز به مراتب پايين‌تر از روش ارائه شده در اين بخش است.

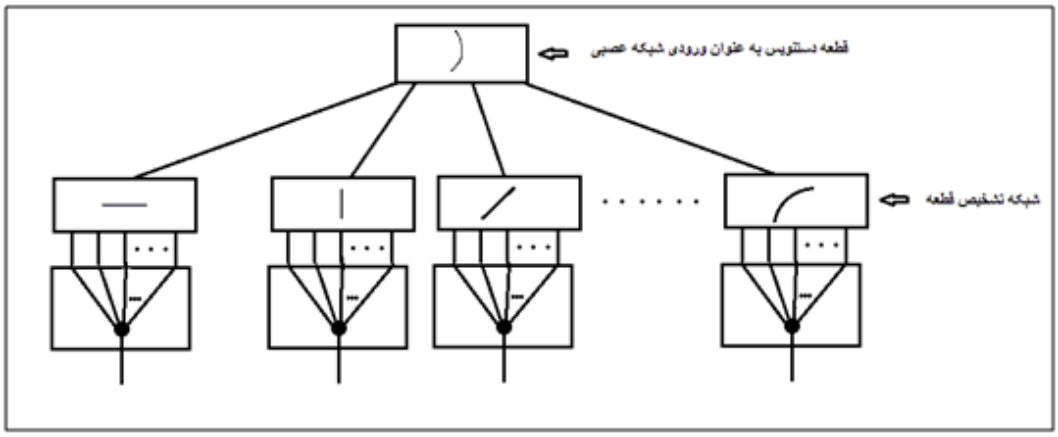
براي اجراي اين طرح، تصوير هر قطعه در ابتدا به 12 شبكه عصبي مختلف از نوع MLP وارد مي‌شوند كه هريک از آن‌ها داراي $k+1$ خروجي **بيتي** هستند (K عددي صحيح است كه ميزان حساسيت سيستم ناميده مي‌شود و در آزمايشات نهايي عدد 5 براي آن برگزيده شده است). در شكل 3-13 تصويري از اين ساختار نمايش داده شده است.



شكل 3-13 تصويري از شبكه‌هاي عصبي تشخيص دهنده قطعات (باخروجي بيتي)

هريک از شبكه‌ها مختص شناسايي يک قطعه خاص هستند و به گونه‌اي طراحي شده و آموزش داده مي‌شوند كه به ازاي هر تصوير ورودي، تنها يكي از خروجي‌ها داراي مقدار يک و ساير‌ين داراي مقدار صفر باشند (براي اين كار از شبيه‌سازي مكانيزم شبكه‌هاي عصبي رقابتي از نوع «تک برنده» استفاده شده است). هر خروجي نماينده عددي بين صفر و يک است كه ميزان شباهت تصوير با قطعه مورد نظر اين شبكه را تشريح مي‌كند. اولين آن‌ها صفر، آخرين آن‌ها يک و خروجي‌هاي مابين به ترتيب نمايانگر اعداد $1/k/2$، k تا k/k هستند.

خروجي‌هاي هريک از اين 12 شبكه، به عنوان ورودي به يک شبكه مجزاي ديگر داده مي‌شوند (به ازاي هركدام يک شبكه وجود دارد) كه هركدام داراي $k+1$ ورودي و يک خروجي از نوع عدد حقيقي هستند. در شكل 3-14 تصوير كامل اين روابط نمايش داده شده است.

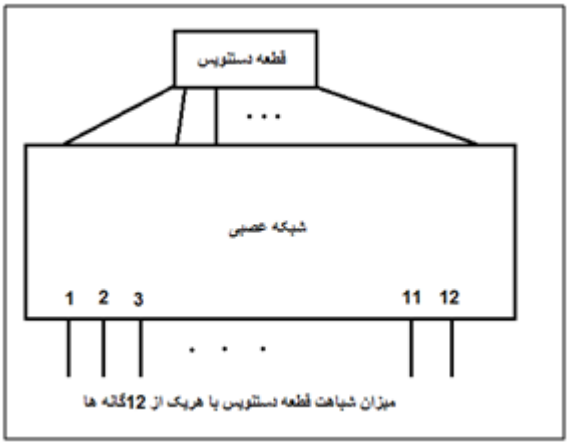


شكل 3-14 تصويري از شبكه عصبي تشخيص دهنده قطعات (با خروجي عددي)

در هريک از اين شبکه‌ها تمام ورودي‌ها به طور مستقيم به تنها خروجي شبكه متصل مي‌شوند و وزن روي هر اتصال عددي است كه ورودي متناظر، آن را نمايندگي مي‌كند. بدين ترتيب هر شبكه در خروجي خود رابطه (3-3) را ارائه مي‌كند كه مي‌توان آن را به عنوان ميزان شباهت تصوير به قطعه متناظر با همان شبكه تعبير نمود.

$$\sum_{i=0}^{k} \frac{i}{k} x_i \qquad (3\text{-}3)$$

در نتيجه مباحث مذكور، يک شبكه عصبي با 12 خروجي تعريف مي‌شود كه با دريافت تصوير يک بخش از دستنويس (متناظر يک قطعه)، ميزان شباهت آن را به هريک از قطعات در خروجي مشخص مي‌كند. در شكل 3-15 تصوير كلي شناسايي قطعه به ازاي هر يک از 12گانه‌ها نمايش داده شده است.

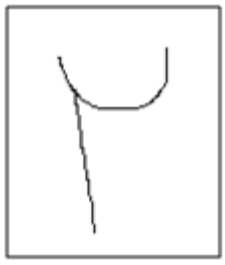


شكل 3-15 تصوير كلي شبكه عصبي شناسايي قطعه

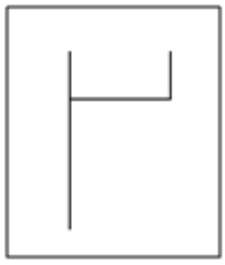
پس از آنكه مولفه‌هاي هر يک از اجزا مشخص گرديد، نتيجه حاصله به واحد ساختاري براي بررسي نهايي و شناسايي نويسه ارسال مي‌شود. در بخش 3-5 به معرفي و نحوه شناسايي نويسه توسط واحد ساختاري پرداخته خواهد شد.

3-5 بخش تشخيص

پس از آنكه نقاط حساس در نويسه شناسايي شدند و نويسه به اجزاي كوچک هندسي تجزيه شد، هر جز به شبكه عصبي ارسال مي‌گردد تا از لحاظ نوع هندسي مشخص شود. شبكه عصبي هر جز را به عنوان ورودي دريافت كرده و نوع آن را مشخص مي‌كند. در نهايت نتيجه شناسايي خود را به عنوان ورودي به «واحد تشخيص» ارسال مي‌كند. در اين بخش به توضيح و نحوه عملكرد «واحد تشخيص» مي‌پردازيم. وظيفه اصلي واحد ساختاري شناسايي نويسه به لحاظ ساختار آن مي‌باشد. بدين معنا كه اين واحد بر اساس ساختار نويسه سعي در شناسايي آن دارد. در واحد ساختاري براي هر نويسه الگوي ساختاري در نظر گرفته شده است. براي برخي از نويسه‌ها چندين الگوي ساختاري پياده‌سازي مي‌شود و اين به آن دليل است كه در نحوه نوشتن دستنويس آن‌ها بيش از يک الگو وجود دارد. به طور مثال در نوشتن نويسه "2" مي‌توان دو صورت مختلف را در نظر گرفت. اين دو شكل متفاوت در شكل‌هاي 3-16 و 3-17 نمايش داده شده است.



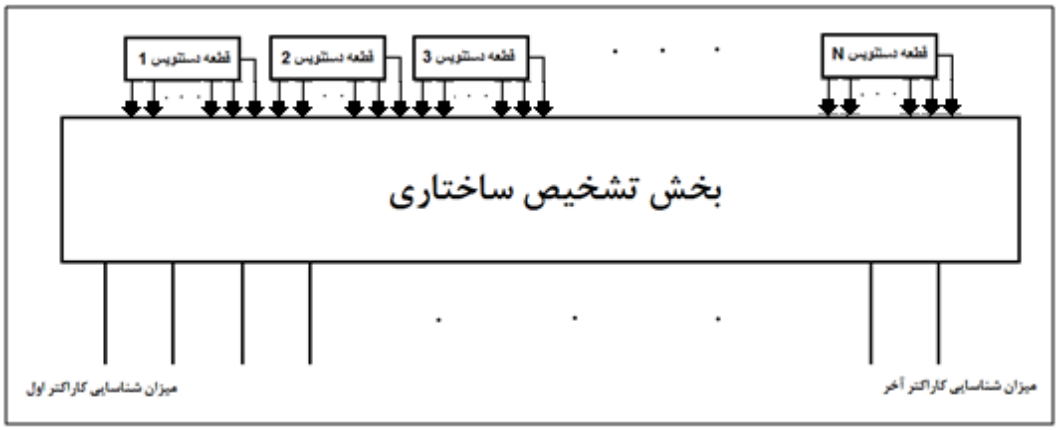
شكل 3-16 نمونه اول از نوشتن "2"



شكل 3-17 نمونه دوم از نوشتن "2"

همانطور كه ديده مي‌شود، شكل 3-16 از اجزاء شماره 6، 4 و 2 و شكل 3-17 از اجزاء 2، 1، 2 و 2 تشكيل شده است.

بخش تشخيص، در يک نگاه كلي، تبديلي است كه مجموعه كامل قطعات شكل و نتيجه شناسايي آن‌ها (از طريق شبكه عصبي) را به يک بردار N بعدي مي‌برد كه هر بعد آن نماينده ميزان شباهت نويسه اصلي – حاصل از اجتماع تمام قطعات - با نويسه مورد نظر همان بردار است. در شكل 3-18 اين موضوع نمايش داده شده است.



شکل 3-18 نگاه کلي به تشخيص ساختاري

به طور مثال در هنگام شناسايي اعداد، خروجي بخشِ تشخيص، يک بردار 10 بعدي خواهد بود (البته در حالتي که هر عدد را تنها با يک مدل نوشتار در نظر بگيريم) که هر خروجي به ترتيب به عنوان ميزان شباهت نويسه با الگوي عدد 0، عدد 1 تا عدد 9 تعبير خواهد شد.

براي بخشِ تشخيص، از روش ساختاري با رويکرد فازي استفاده مي‌شود. بدين معني که **اولا** شباهت تصوير به هر نويسه، بر اساس ويژگي‌هاي ساختاري خاص آن نويسه – نظير محل قرارگيري، جهت و اندازه قطعات، زاويه‌هاي مطلق قطعات و زواياي نسبي آن‌ها با يکديگر و ساير ويژگي‌هايي از اين دست – تعيين مي‌شود، **ثانيا** وضعيت‌هاي محتمل براي هريک از ويژگي‌هاي مذکور از طريق متغيرهاي زباني تعريف شده و ترکيبات فازي آن‌ها در تعيين شرايط مورد استفاده قرار مي‌گيرد. نکته اساسي بحث فازي اين بخش، اين است که نه تنها ويژگي‌هاي هندسي منتسب به يک قطعه، بلکه ماهيت وجودي يک قطعه (نوع آن) نيز به صورت فازي بيان مي‌شود (بر اساس نتايج حاصل از خروجي‌هاي شبکه عصبي). به عبارت دقيق‌تر، قطعات 12گانه، خود مجموعه‌هاي فازي هستند که ويژگي «نوع» يک قطعه را بيان مي‌کنند و هريک از درجه‌بندي‌هاي مربوط به حساسيت سيستم حکم يک متغير زباني را براي اين مجموعه خواهند داشت.

3-6 تعيين نتيجه نهايي

براي تعيين نتيجه نهايي، در ابتدا هر خروجي (ميزان شباهت تصوير به هر نويسه)، از طريق مقايسه با يک مرز مشخص، ويژگي‌اي به نام «ردي» يا «قبولي» در مورد نويسه متناظر خود را به ما خواهد داد. به عبارت ديگر مي‌توان عمليات defuzzy کردن را از

طريق مقياس قرار دادن يک آستانه انجام داد. اما نکته اين است که آستانه مطلوب براي هر نويسه با نويسه ديگر تفاوت دارد. بنابراين ويژگي «ردي» يا «قبولي» هر نويسه از طريق مقايسه مقدار خروجي، با آستانه تعيين شده براي همان نويسه به دست مي‌آيد.

با توجه به مطالب بالا، پس از مشخص شدن نتايج توسط بخشِ تشخيص، دو حالت عمومي محتمل است: اول اينکه هيچ‌يک از خروجي‌ها نتوانند مقدار «قبولي» را ارائه کنند؛ دوم اينکه حداقل يک خروجي مقدار «قبولي» داشته باشد. در ذيل چگونگي پردازش هريک از اين حالات را بررسي مي‌کنيم:

اول: همه حالات «مردود» - اين حالت دلالت بر عدم شناسايي نويسه بر اساس قطعه‌بندي فعلي دارد. به عبارت ديگر در اين حالت نتيجه اين است که «به فرمي که در اين مرحله، قطعات شکل دستنويس را تعيين کرده‌ايم، هيچ معادلي براي آن يافت نشده است». نتيجه بلافصل اين جمله آن است که «محتمل است با تعويض قطعات بتوان معادل مناسبي براي شکل يافت». به همين دليل در ازاي گرفتن اين جواب، الگوريتم بايد ترکيب ديگري از قطعات را بوجود آورد. روش کار، بر اين اساس است که در هنگام تعيين قطعات اوليه، آن‌ها را با تدابير گفته شده بوجود مي‌آوريم که در نتيجه، قطعات در ريزترين حالت ممکن استخراج مي‌شود. به اين صورت، مي‌توان با بررسي تمام ترکيبات ممکن، در نهايت به حالتي که مورد نظر نويسنده بوده دست يافت – توجه به اين نکته ضروري است که بسياري از ترکيبات، با توجه به فاصله قطعات، خود به خود حذف مي‌شوند و از اين نظر تعداد ترکيبات، زياد نخواهد شد.

دوم: يک يا چند حالت «قبولي» - در اين صورت الگوريتم به جواب خود رسيده است و با توجه به ماکزيمم عدد موجود در خروجي، نويسه نهايي مشخص خواهد شد. ابهام در اين مورد آن است که «چنانچه که در اين مرحله، عمليات قطع نشود، امکان آن وجود دارد که سيستم در ترکيبات بعدي نيز به جوابي دست پيدا کند؛ آيا حذف اين احتمال شانس خطا را تحت تاثير قرار مي‌دهد؟». پاسخ اين است که مي‌توان عمليات را تا پايان دنبال نمود اما به طور تجربي ديده شده است که افزايش دقت در اين صورت بسيار ناچيز بوده و در عوض سرعت تشخيص بسيار کم‌تر خواهد شد.

فصل چهارم

آزمایش‌ها و نتایج

4-1 مقدمه

در اين فصل ابتدا به معرفي پايگاه‌هاي داده معروف و تاييد شده در زمينه تصاوير ارقام و حروف دستنويس فارسي پرداخته مي‌شود. در [22] و [23] پايگاه‌هاي داده ديگري نيز معرفي گرديده‌اند. سپس مروري مفصل بر آزمايش‌هاي انجام شده روي اين پايگاه‌هاي داده و نتايج آنها خواهد شد. اين نتايج به صورت كلي و تفكيكي و بر اساس جداول و نمودارهاي گرافيكي ارائه مي‌شود. پس از آن يک جمع‌بندي كلي بر روي تمامي پايگاه‌هاي داده و نتايج آنها ارايه خواهد شد. در ادامه به نويسه‌هايي كه شناسايي نشده‌اند يا به اشتباه شناسايي شده‌اند و دلايل آنها پرداخته مي‌شود. در انتها نيز به مقايسه نتيجه كارهاي انجام شده در زمينه شناسايي نويسه‌هاي دستنويس فارسي و روش پيشنهادي پرداخته خواهد شد.

4-2 معرفي پايگاه‌هاي داده تصاوير نويسه‌هاي دستنويس فارسي

در اين بخش سه پايگاه داده معروف و مورد تاييد كه اكثر محققان در زمينه شناسايي نويسه‌هاي دستنويس فارسي آنها را به براي مقايسه نتايج كارهاي خود انتخاب مي‌كنند، معرفي مي‌گردد. همچنين يک پايگاه داده كه توسط مولف و همكاران تهيه گرديده است، معرفي مي‌شود.

4-2-1 پايگاه داده هدي

مجموعه ارقام و حروف دستنويس هدي [24] كه اولين مجموعه‌ي بزرگ دستنويس فارسي است، مشتمل بر 102353 نمونه دست‌نوشته ارقام و 88351 نمونه دست‌نوشته حروف به صورت سياه و سفيد است. اين مجموعه طي انجام يک پروژهي كارشناسي ارشد درباره بازشناسي فرم‌هاي دستنويس تهيه شده است. داده هاي اين مجموعه از حدود 12000 فرم ثبت نام آزمون سراسري كارشناسي ارشد سال 1384 و آزمون كارداني پيوستهي دانشگاه جامع علمي كاربردي سال 1383 استخراج شده است. درجه تفكيک نمونه‌ها در اين مجموعه داده 200 نقطه بر اينچ مي‌باشد. در شكل 4-1 نمونه‌اي از حروف و در شكل 4-2 نمونه‌اي از ارقام موجود در اين پايگاه داده نمايش داده شده است.

شكل 4-1 نمونه‌اي از حروف دستنويس فارسي در مجموعه داده هدي [24]

شكل 4-2 نمونه‌اي از ارقام دستنويس فارسي در مجموعه داده هدي[24]

4-2-2 پايگاه داده IFHCDB[5]

در بخشي از اين تحقيق از پايگاه داده IFHCDB كه شامل حروف و ارقام فارسي است و توسط دانشگاه صنعتي اميركبير تهيه گرديده، استفاده شده است[25]. پايگاه داده IFHCDB توسط دانشكده مهندسي برق در سال 1385 تهيه شده است. اين پايگاه در حقيقت يک زير مجموعه از يک پايگاه داده عظيم است كه به عنوان يک پروژه تحقيقاتي با حمايت دبيرخانه عالي شوراي اطلاع رساني، كارگروه خط و زبان فارسي و توسط شركت انديشه نرم‌افزار پايا جمع‌آوري شده است. هدف اصلي از تهيه اين پايگاه داده كمک به محققان براي ايجاد روش‌ها و الگوريتم‌هاي جديد در زمينه شناسايي خودكار نويسه‌هاي دستنويس فارسي است. اين مجموعه مشتمل بر 17740 نمونه دست‌نوشته ارقام و 52380 نمونه دست‌نوشته حروف به صورت خاكستري است. درجه تفكيک نمونه‌ها در اين مجموعه داده 300 نقطه بر اينچ مي‌باشد.

4-2-3 پايگاه داده CENPARMI[6]

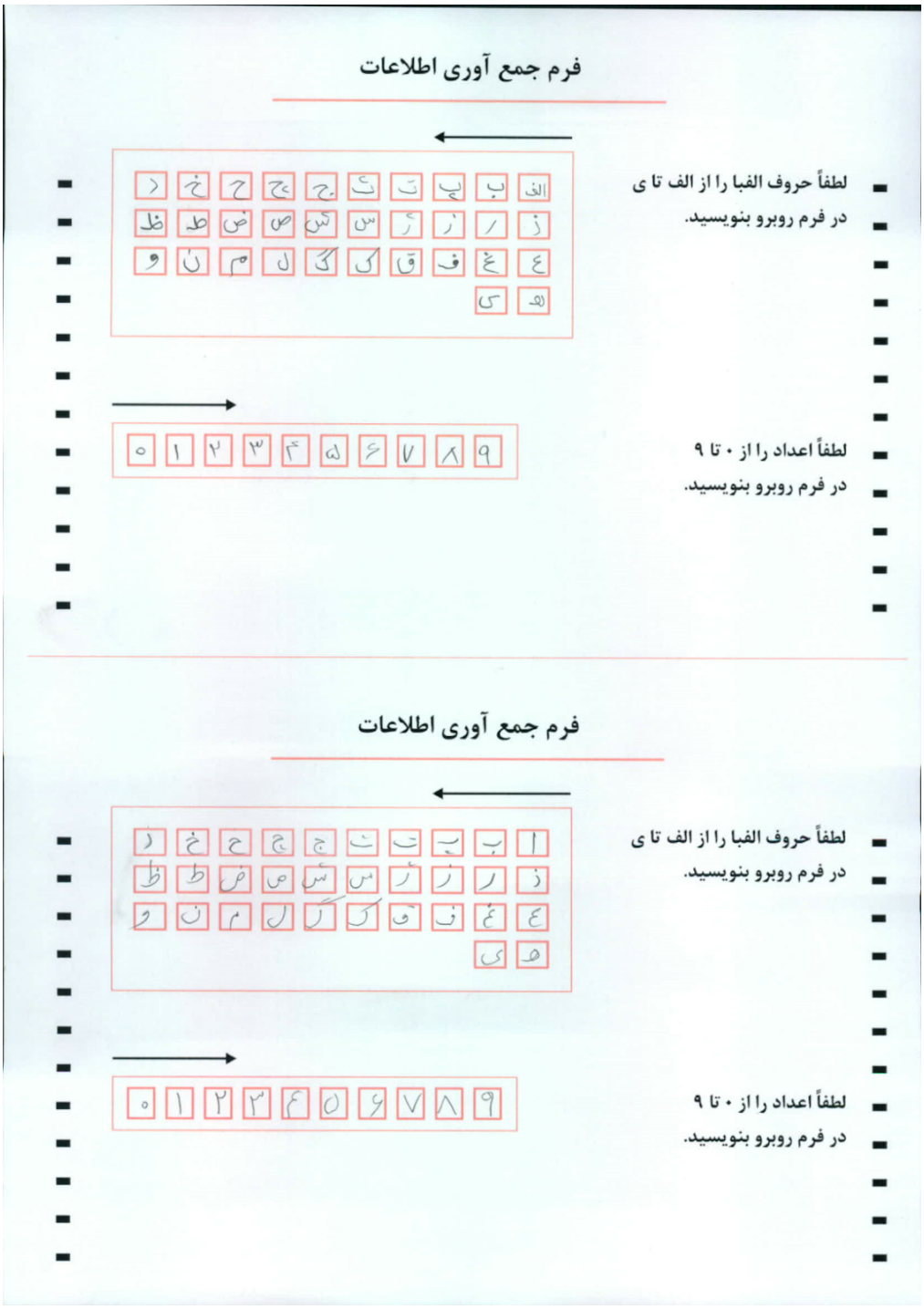
پايگاه داده CENPARMI [26] متشكل از حروف و اعداد دستنويس فارسي است كه توسط دانشگاه كونكورديايِ كانادا، مركز تحقيقات بين‌المللي در دانشكده علوم كامپيوتر جمع آوري شده است. نمونه‌هاي جمع‌آوري شده توسط 175 نفر در سنين و جنسيت‌هاي مختلف نوشته شده‌اند. تعداد نمونه دست‌نوشته ارقام 18000 و

[5] Isolated Farsi Handwritten Character Database
[6] Center for Pattern Recognition and Machine Intelligence

دست‌نوشته حروف 11900 نمونه به صورت رنگي مي‌باشد. درجه تفكيک نمونه‌ها در اين مجموعه داده 300 نقطه بر اينچ مي‌باشد. براي استفاده از اين مجموعه بايد هزينه آن به دانشگاه پرداخته شود تا مجوز استفاده داده شود.

4-2-4 پايگاه داده شركت متاكد

پايگاه داده شركت متاكد[27] با همكاري شركت سيتكس[28] در سال 1389 تهيه گرديده است. تمامي تصاوير نويسه‌هاي جمع‌آوري شده از افراد زير 12 سال (مقطع دبستان) مي‌باشند تا نويسه‌ها كاملا به صورت مبتدي و بي‌نظم باشند. اين روش باعث دقت بيشتري در شناسايي و نتايج نهايي موتورهاي بازشناسي مي‌گردد. پايگاه تصاوير شركت متاكد شامل 8328 نمونه حروف و 12703 نمونه اعداد دستنويس به صورت سياه و سفيد مي‌باشد. درجه تفكيک نمونه‌ها در اين مجموعه داده 200 نقطه بر اينچ مي‌باشد. نمونه‌اي از فرم‌هاي جمع‌آوري شده از مدارس در شكل 4-3 نمايش داده شده است. نويسه‌ها توسط نرم‌افزار بريده و در داخل پايگاه داده‌ها ذخيره شده است.



شكل 4-3 نمونه‌اي از فرم‌هاي جمع‌آوري شده نويسه‌هاي دستنويس توسط شركت متاكد

4-3 درصد بازشناسايي ارقام

در اين بخش درصد شناسايي ارقام توسط روش پيشنهادي به تفكيک هر رقم به صورت جدول و نمودار گرافيكي و بر اساس آزمايش بر روي پايگاه‌هاي داده مختلف بررسي مي‌گردد.

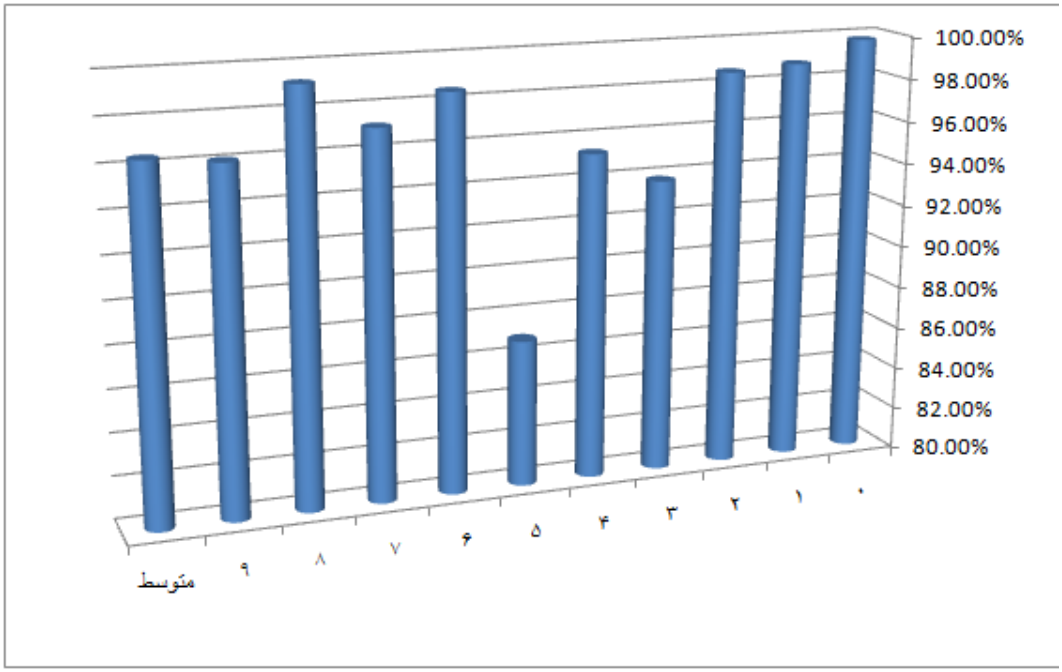
4-3-1 درصد شناسايي ارقام بر روي پايگاه داده هدي

در جدول 4-1 درصد شناسايي ارقام بر روي پايگاه داده هدي به تفكيک هر رقم نمايش داده شده است.

جدول 4-1 درصد شناسايي ارقام به تفكيک بر روي پايگاه داده هدي

رقم	0	1	2	3	4
درصد	%99.80	%98.84	%98.64	%93.83	%95.32
رقم	5	6	7	8	9
درصد	%86.91	%98.53	%97.16	%99.24	%96.11

به طور متوسط ميزان شناسايي ارقام بر روي پايگاه داده هدي 96.44% مي‌باشد. درشكل 4-4 درصد شناسايي ارقام بر روي پايگاه داده هدي به تفكيک ارقام و همچنين ميانگين نمايش داده شده است.



شكل 4-4 درصد شناسايي ارقام به تفكيک و ميانگين بر روي پايگاه داده هدي

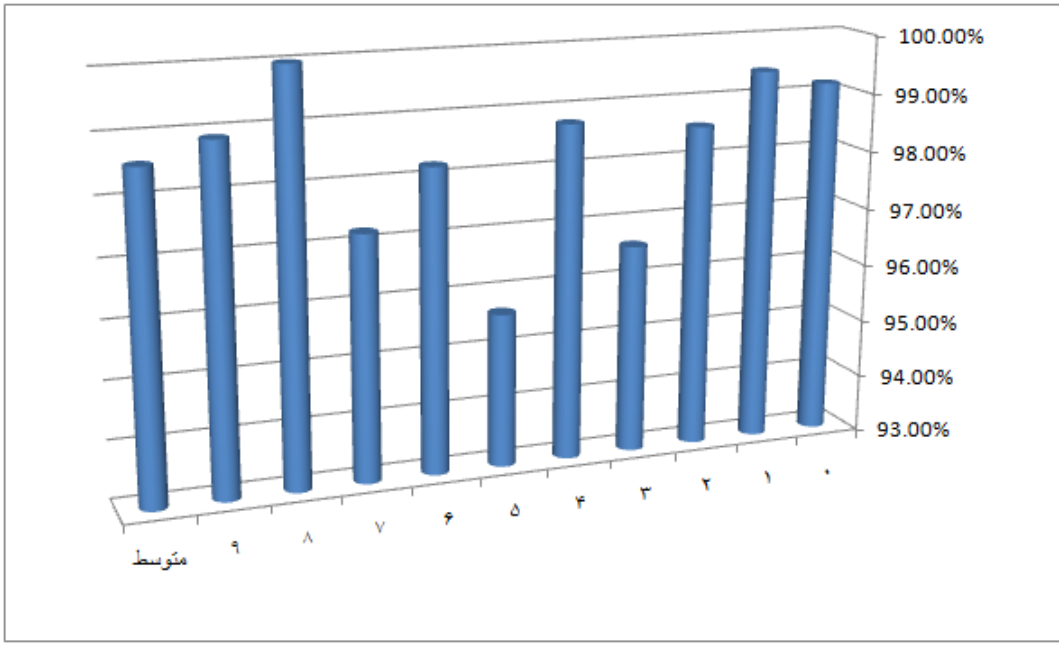
4-3-2 درصد شناسايي ارقام بر روي پايگاه داده IFHCDB

در جدول 4-2 درصد شناسايي ارقام بر روي پايگاه داده IFHCDB به تفكيک هر رقم نمايش داده شده است.

جدول 4-2 درصد شناسايي ارقام به تفكيک بر روي پايگاه داده IFHCDB

رقم	0	1	2	3	4
درصد	%99.16	%99.43	%98.56	%96.60	%98.76
رقم	5	6	7	8	9
درصد	%95.65	%98.22	%97.22	%100	%98.90

به طور متوسط ميزان شناسايي ارقام بر روي پايگاه داده IFHCDB، %98.56 مي‌باشد. در شكل 4-5 درصد شناسايي ارقام بر روي پايگاه داده IFHCDB به تفكيک ارقام و همچنين ميانگين نمايش داده شده است.



نمودار 4-5 درصد شناسايي ارقام به تفكيک و ميانگين بر روي پايگاه داده IFHCDB

4-3-3 درصد شناسايي ارقام بر روي پايگاه داده متاكد

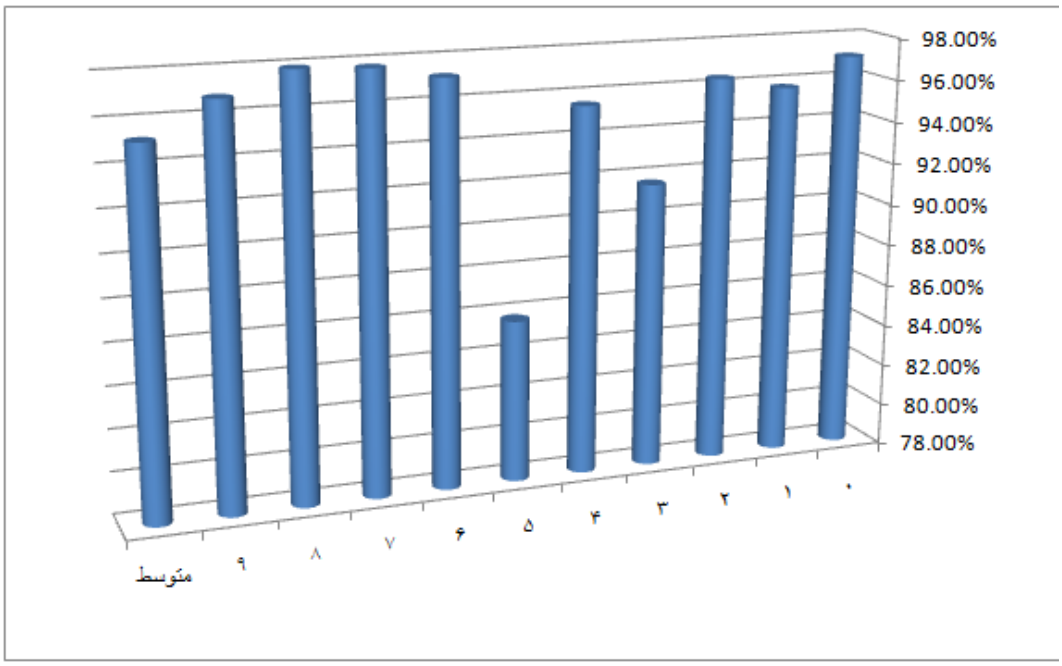
در جدول 4-3 درصد شناسايي ارقام بر روي پايگاه داده متاكد به تفكيک ارقام نمايش داده شده است.

جدول 4-3 درصد شناسايي ارقام به تفكيک بر روي پايگاه داده متاكد

رقم	0	1	2	3	4
درصد	%97.08	%95.77	%96.39	%91.64	%95.53

رقم	5	6	7	8	9
درصد	%85.74	%97.19	%97.78	%97.92	%96.84

به طور متوسط ميزان شناسايي ارقام بر روي پايگاه داده متاكد 95.19% مي‌باشد. در شكل 4-6 درصد شناسايي ارقام بر روي پايگاه داده متاكد به تفكيك ارقام و همچنين ميانگين نمايش داده شده است.



نمودار 4-6 درصد شناسايي ارقام به تفكيك و ميانگين بر روي پايگاه داده متاكد

در نهايت ميانگين كلي شناسايي ارقام در مجموع هر 3 پايگاه داده برابر 96.73% مي‌باشد.

4-4 درصد بازشناسايي حروف

در اين بخش درصد شناسايي حروف توسط روش پيشنهادي به تفكيك حرف به صورت جدول و نمودار گرافيكي و بر اساس آزمايش بر روي پايگاه‌هاي داده مختلف بررسي مي‌گردد.

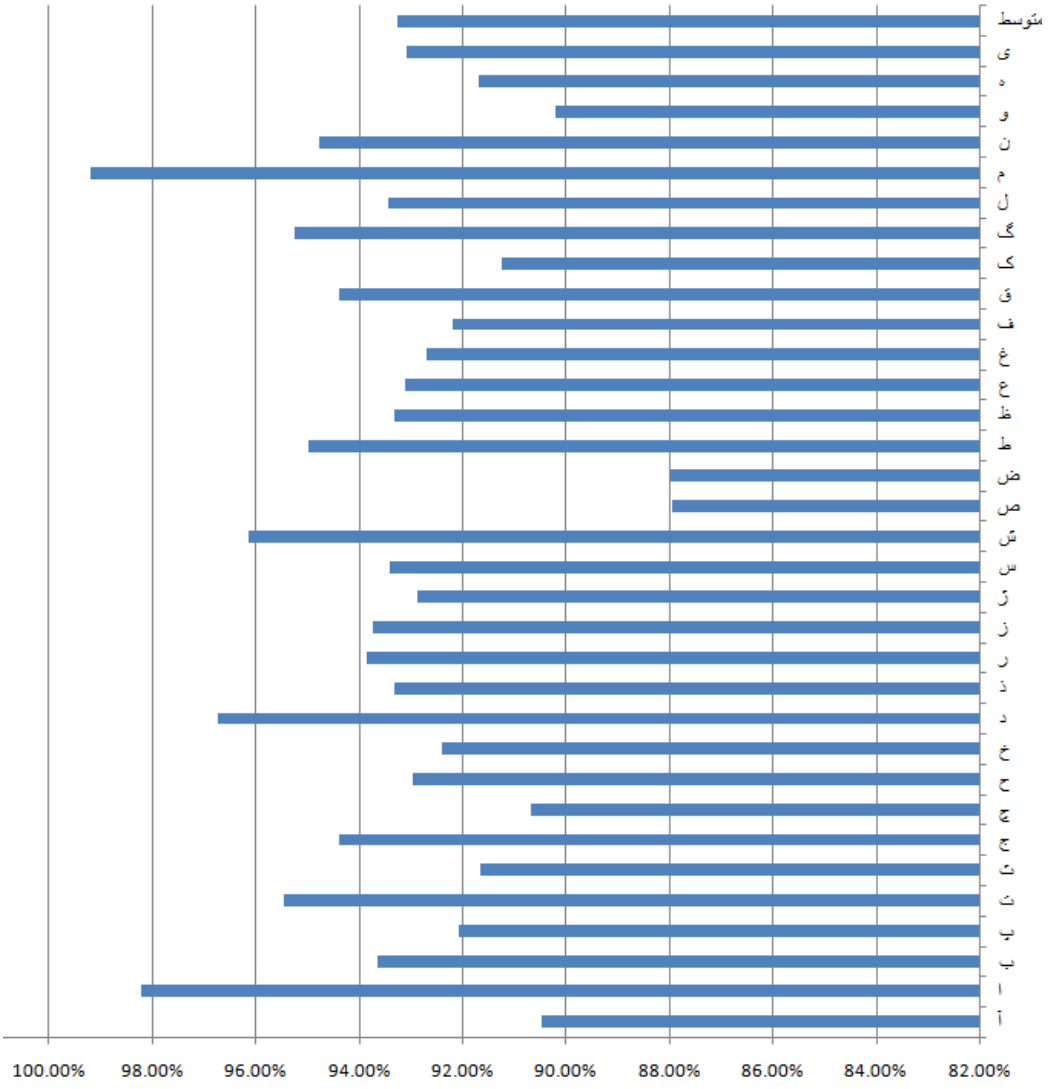
4-4-1 درصد شناسايي حروف بر روي پايگاه داده هدي

در جدول 4-4 درصد شناسايي حروف بر روي پايگاه داده هدي به تفكيك حرف نمايش داده شده است.

جدول 4-4 درصد شناسايي حروف به تفكيک بر روي پايگاه داده هدي

ث	ت	پ	ب	ا	آ	حرف
%91.67	%95.45	%92.06	%93.65	%98.23	%90.48	درصد
ذ	د	خ	ح	چ	ج	حرف
%93.33	%96.74	%92.41	%92.97	%90.67	%94.40	درصد
ص	ش	س	ژ	ز	ر	حرف
%87.94	%96.14	%93.42	%92.89	%93.73	%93.86	درصد
ف	غ	ع	ظ	ط	ض	حرف
%92.18	%92.71	%93.10	%93.33	%95.00	%87.96	درصد
ن	م	ل	گ	ک	ق	حرف
%94.77	%99.20	%93.44	%95.26	%91.25	%94.40	درصد
			ي	ه	و	حرف
			%93.08	%91.70	%90.20	درصد

به طور متوسط ميزان شناسايي حروف بر روي پايگاه داده هدي 93.26% مي‌باشد. در شكل 4-7 درصد شناسايي حروف بر روي پايگاه داده هدي به تفكيک حرف و همچنين ميانگين نمايش داده شده است.



شكل 4-7 درصد شناسايي حروف به تفكيک و ميانگين بر روي پايگاه داده هدي

4-4-2 درصد شناسايي حروف بر روي پايگاه داده IFHCDB

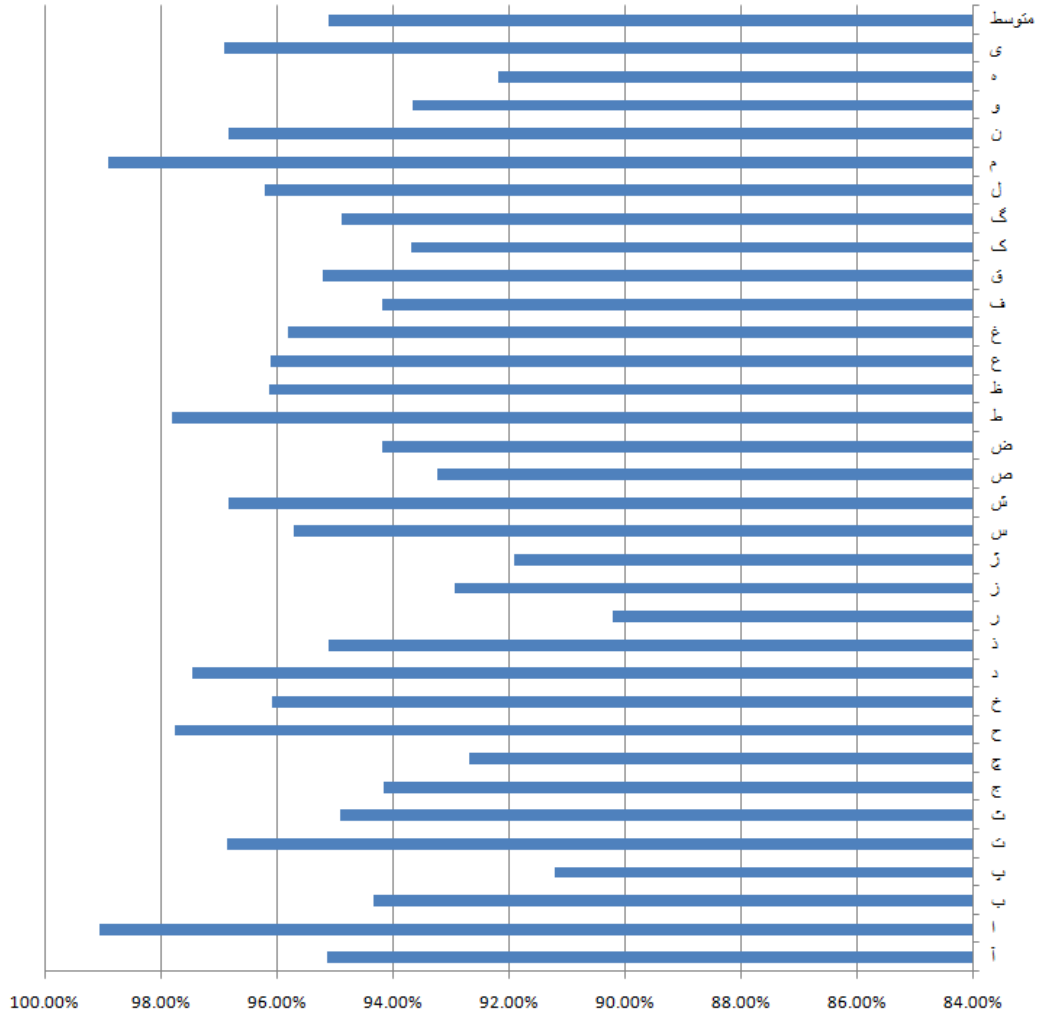
در جدول 4-5 درصد شناسايي حروف بر روي پايگاه داده IFHCDB به تفكيک حرف نمايش داده شده است.

جدول 4-5 درصد شناسايي حروف به تفكيک بر روي پايگاه داده IFHCDB

حرف	آ	ا	ب	پ	ت	ث

درصد	%95.12	%99.05	%94.32	%91.19	%96.86	%94.90
حرف	ج	چ	ح	خ	د	ذ
درصد	%94.14	%92.67	%97.75	%96.08	%97.44	%95.09
حرف	ر	ز	ژ	س	ش	ص
درصد	%90.21	%92.92	%91.89	%95.70	%96.83	%93.23
حرف	ض	ط	ظ	ع	غ	ف
درصد	%94.18	%97.81	%96.13	%96.11	%95.81	%94.18
حرف	ق	ک	گ	ل	م	ن
درصد	%95.19	%93.67	%94.87	%96.20	%98.90	%96.82
حرف	و	ه	ي			
درصد	%93.65	%92.17	%96.90			

به طور متوسط ميزان شناسايي حروف بر روي پايگاه داده IFHCDB %95.09 مي‌باشد. در شكل 4-8 درصد شناسايي حروف بر روي پايگاه داده IFHCDB به تفكيک حرف و همچنين ميانگين نمايش داده شده است.



شکل 4-8 درصد شناسايي حروف به تفکيک و ميانگين بر روي پايگاه داده IFHCDB

4-4-3 درصد شناسايي حروف بر روي پايگاه داده متاکد

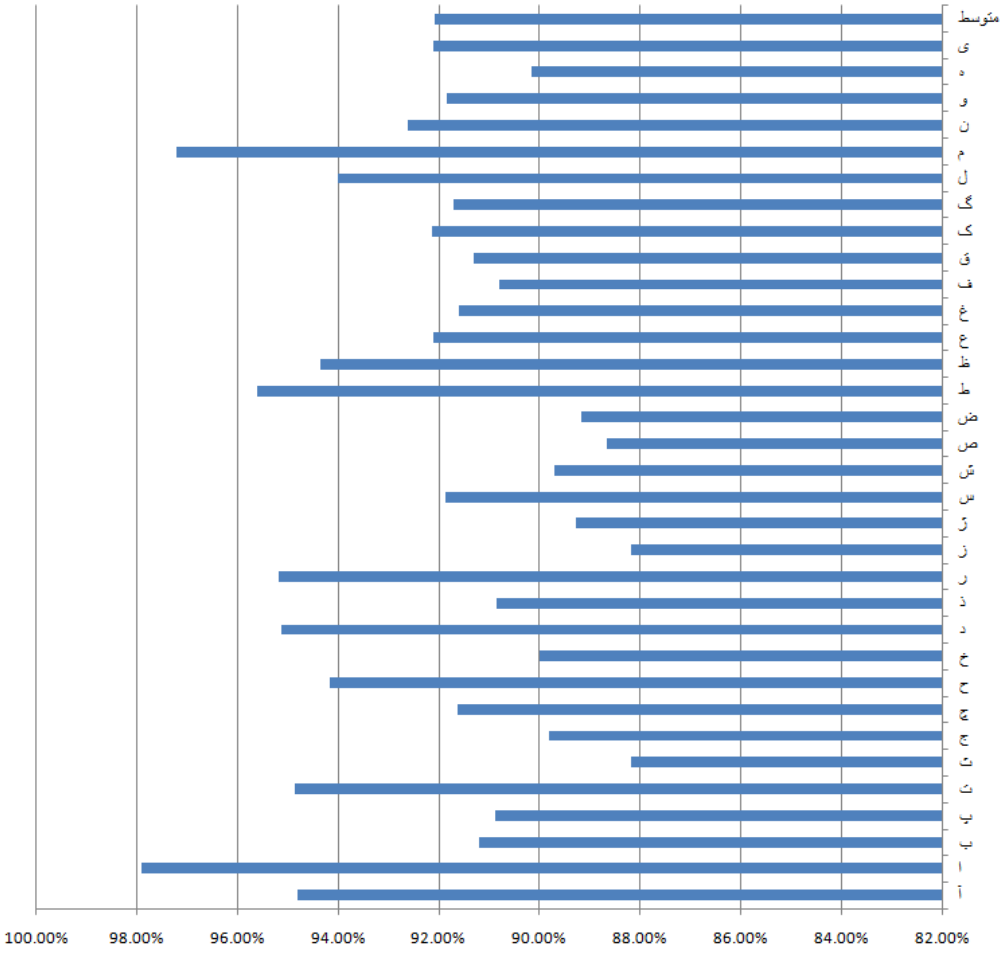
در جدول 4-6 درصد شناسايي حروف بر روي پايگاه داده متاکد به تفکيک حرف نمايش داده شده است.

جدول 4-6 درصد شناسايي حروف به تفکيک بر روي پايگاه داده متاکد

حرف	آ	ا	ب	پ	ت	ث

درصد	%94.81	%97.91	%91.19	%90.89	%94.86	%88.18
حرف	ج	چ	ح	خ	د	ذ
درصد	%89.80	%91.62	%94.16	%90.01	%95.12	%90.86
حرف	ر	ز	ژ	س	ش	ص
درصد	%95.19	%88.19	%89.28	%91.86	%89.70	%88.66
حرف	ض	ط	ظ	ع	غ	ف
درصد	%89.17	%95.61	%94.36	%92.11	%91.61	%90.80
حرف	ق	ک	گ	ل	م	ن
درصد	%91.31	%92.13	%91.71	%94.02	%97.21	%92.61
حرف	و	ه	ي			
درصد	%91.85	%90.17	%92.12			

به طور متوسط ميزان شناسايي حروف بر روي پايگاه داده متاکد 92.09% مي‌باشد. در شکل 4-9 درصد شناسايي حروف بر روي پايگاه داده متاکد به تفکيک حرف و همچنين ميانگين نمايش داده شده است.



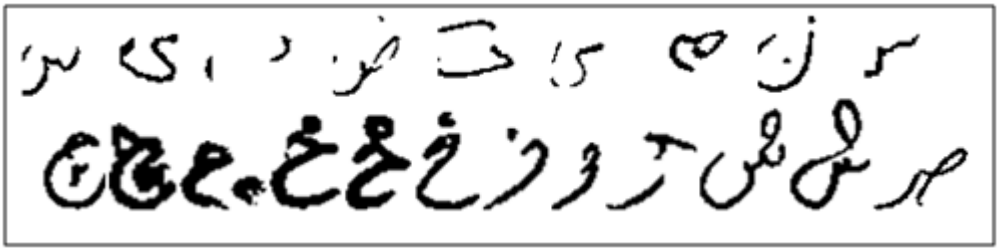
شكل 4-9 درصد شناسايي حروف به تفكيک و ميانگين بر روي پايگاه داده متاكد

در نهايت ميانگين كلي شناسايي حروف در مجموع هر 3 پايگاه داده برابر 93.48% مي‌باشد.

4-5 نويسه‌هاي شناسايي نشده

موتور بازشناسي پيشنهادي در برخي از موارد كه نويسه به درستي نوشته نشده است يا به دليل نويز زياد در تصوير كه حتي توسط پيش‌پردازش رفع نگرديده است، قادر به شناسايي نويسه نمي باشد. در شكل 4-10 نمونه‌اي از تصاوير ارقام و در شكل 4-11 نمونه‌اي از حروف كه شناسايي نشده‌اند، نمايش داده شده است.

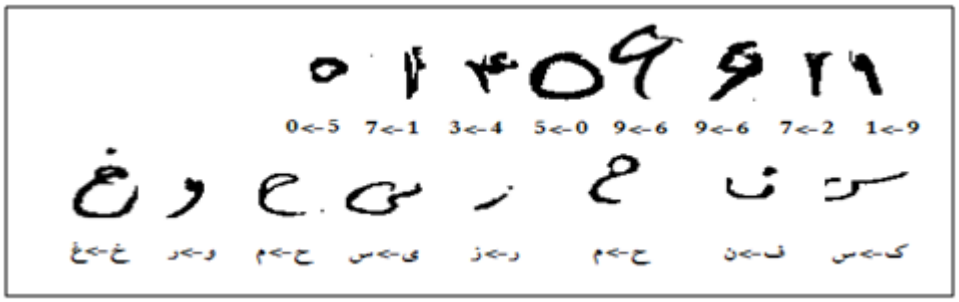
شكل 4-10 نمونه‌اي از ارقام كه شناسايي نشده‌اند



شكل 4-11 نمونه‌اي از حروف كه شناسايي نشده‌اند

4-6 نويسه‌هاي به اشتباه شناسايي شده

در برخي از مواقع امكان دارد نويسه توسط موتوربازشناسي تشخيص داده بشود ولي اين شناسايي درست نباشد. از دلايل آن مي‌توان به شباهت بسيار زياد نويسه‌هاي همانند نام برد. در نتيجه كاربر در زمان نوشتن به علت سرعت بالا و دقت كم تمركزي خوبي بر روي نوشتن نويسه نمي‌كند و اين باعث شناسايي اشتباه مي‌شود. گاهي اوقات اين شباهت‌ها به قدري زياد است كه انسان نيز قادر به تشخيص صحيح نمي‌باشد و لذا از نويسه‌هاي قبل و بعد براي شناسايي كلمه‌اي كه نويسه در آن مي‌باشد، استفاده مي‌كند تا نويسه را شناسايي كند. نازك‌سازي نويسه نيز يكي ديگر از دلايل تشخيص نادرست نويسه مي‌باشد. در شكل 4-12 نمونه‌اي از ارقام و حروفي كه به اشتباه شناسايي شده‌اند نمايش داده شده است.



شكل 4-12 نمونه‌اي از ارقام و حروفي كه به اشتباه شناسايي شده‌اند

4-7 مقايسه روش پيشنهادي با ساير كارهاي انجام شده

در اين بخش به نتايج كارهاي انجام شده در زمينه شناسايي نويسه‌هاي دستنويس فارسي پرداخته خواهد شد. در [29] از يک روش آماري براي شناسايي ارقام فارسي استفاده شده است. در اين روش استخراج ويژگي‌هاي اصلي به روش مكان‌هاي مشخصه مورب و شناسايي نهايي توسط طبقه‌بندي بيز صورت گرفته است. پايگاه داده‌اي متشكل از 12778 رقم با دقت 200 نقطه بر اينچ كه 60% آن براي آموزش در نظر گرفته شده، به عنوان مجموعه آزمايشي استفاده شده است. نتايج شناسايي صحيح در اين روش 95.82% گزارش شده است. در [15] يک روش شناسايي آماري براي حروف دستنويس فارسي ارائه شده است. در اين روش با فرض حذف بريدگي‌هاي نامطلوب و جدا بودن نقاط از بدنه، حروف به 5 گروه تقسيم‌بندي شده‌اند. با استفاده از طبقه‌بندي بيز و شناسايي ويژگيهاي گشتاوري مناسب و در آخر با استفاده از پس‌پردازشهاي لازم، شناسايي صورت گرفته است. در يک مجموعه شامل 6600 نمونه حروف كه 50% آن براي آموزش استفاده گرديده، ميزان بازشناسي صحيح 90.64% گزارش شده است. در [30] بازشناسي حروف دستنويس فارسي با روش فازي ارائه شده است. در اين روش بازشناسي در دو مرحله صورت گرفته است. 1- طبقه‌بندي كننده فازي "قاعده- پايه" با يادگيري اتوماتيک حروف كه ويژگيهاي استخراج شده در اين مرحله از روش مكان مشخصه بدست مي‌آيند. 2- طبقه‌بندي كننده‌هاي درخت تصميم بايناري با ويژگي‌هاي متنوع ساختاري. در يک مجموعه با 4950 نمونه كه 3200 نمونه براي آموزش استفاده شده است، نتيجه شناسايي 92% گزارش شده است. در [31] يک روش مبتني بر كدهاي فركتالي براي شناسايي حروف و اعداد گسسته دستنويس آورده شده است. هنگاميكه اين كدها به مجموعه بلوک‌هاي دامنه و برد در يک تصوير دلخواه اوليه چند بار متوالي اعمال گردند باعث بوجود آوردن تصويري خواهند شد كه به نويسه اصلي شباهت زيادي دارد. مجموعه كدهاي فركتالي به عنوان ويژگي‌هاي مربوط به هر حرف و عدد به شبكه عصبي توابع پايه شعاعي و شبكه عصبي پرسپترون چند لايه داده مي‌شود.

در یک مجموعه حروف متشکل از 6080 نمونه حرف، 3200 نمونه و در یک مجموعه ارقام متشکل از 3840 رقم، 2240 نمونه براي آموزش در نظر گرفته شده است. نتايج بازشناسي صحيح براي حروف توسط توابع پايه شعاعي 91.70% و براي پرسپترون 91.37% و براي ارقام توسط توابع پايه شعاعي 90.90% و براي پرسپترون 82.26% گزارش شده است. در [32] كاربرد تركيب طبقه‌بندها در بازشناسي ارقام دستنويس فارسي آورده شده است. تركيب طبقه‌بندها از يک الگوريتم يادگيري دو مرحله‌اي بهره مي‌گيرد. از تصوير هر رقم دستنويس يک بردار ويژگي با 81 مولفه استخراج مي‌شود. به روش تحليل مولفه‌هاي اصلي، يک بردار ويژگي با 15 مولفه انتخاب شده و به سه شبكه عصبي پرسپترون با تعداد نرونهاي متفاوت در لايه پنهان ارسال مي‌گردد. بازشناسي به صورت مستقل در هر طبقه صورت مي‌گيرد. در نهايت خروجي اين سه شبكه عصبي به يک شبكه عصبي پرسپترون با يک لايه مخفي جهت تركيب ارسال مي‌گردد. در يک مجموعه ارقام متشكل از 2430 نمونه رقم با دقت 300 نقطه بر اينچ، 1900 نمونه جهت آموزش و 530 نمونه جهت آزمايش قرار گرفته‌اند. نتايج حاصل از بازشناسي صحيح در اين روش 91% اعلام شده است. در [10] يک روش ساختاري براي شناسايي حروف و ارقام فارسي ارائه گرديده است. اين روش مبتني بر استخراج ساختار پلكس و همچنين روشي براي توصيف الگوهايي كه داراي عناصر غيرمتصل هستند، آورده شده است. اطلاعي از پايگاه داده‌اي كه بر روي آن آزمايش صورت گرفته است، موجود نمي‌باشد ولي نرخ بازشناسي 95% اعلام گرديده است. در [16] كاربرد تطابق شكل در بازشناسي ارقام دستنويس فارسي ذكر شده است. براي هر نقطه نمونه بر روي كانتور شكل، توصيفگري بر اساس توزيع مكاني نقاط ديگر كانتور بدست مي‌آيد. جمع فواصل بين نقاط متناظر دو شكل معياري براي عدم شباهت آنها است. در يک آزمايش بر روي مجموعه‌اي از 1288 رقم، ميزان 89.9% بازشناسي صحيح حاصل شده است. در [33] دو ويژگي سريع و كارآمد براي بازشناسي ارقام دستنويس فارسي آمده است. اين ويژگي‌ها كه آنها را گراديان بهبود يافته و و هيستوگرام گراديان ناميده‌اند، مبتني بر ويژگي گراديان روشنايي بوده و به ترتيب براي تصاوير خاكستري و باينري توسعه يافته‌اند. اين روش بر روي پايگاه داده هدي با آموزش 60000 نمونه و آزمايش 20000 نمونه تست شده كه حاصل نتايج آن به ترتيب 99.02% براي روش اول و 98.80% براي روش دوم بدست آمده است. در [34] يک روش بازشناسي براي حروف دستنويس فارسي با MOMENT INVARIANTS آورده شده است. در اين روش از گشتاورهاي زرنيكي[7]، شبه زرنيكي [8] و Legendre استفاده شده است. براي طبقه‌بندي نيز از شبكه عصبي ART2 استفاده شده است. در مجموعه‌اي از حروف متشكل از 16000 نمونه كه 10000 نمونه آن براي آموزش بوده است، ميزان بازشناسي صحيح 96.92% گزارش شده است. در [35]

[7] - Zernike Moment
[8] - Pseudo Zernike

شناسايي ارقام فارسي/عربي با استفاده از ماشينهاي بردار پشتيبان بيان شده است. اين روش هر رقم را از 4 ديدگاه متفاوت در نظر گرفته كه در هر ديدگاه 16 ويژگي مختلف يعني در مجموع 64 ويژگي مختلف را استخراج و تركيب مي‌كند. از ماشينهاي بردار پشتيبان براي طبقه‌بندي كلاسهاي مختلف استفاده شده است. اين روش بر روي پايگاه داده CENPARMI با 7390 نمونه آموزشي بر روي 3035 نمونه آزمايش شده و نتيجه 94.41% ميزان بازشناسي صحيح را در بر داشته است. در [36] نيز همانند [5] بررسي و مقايسه‌اي بين سيستمهاي شناسايي كاركترهاي دستنويس عربي صورت گرفته است. در [37] شناسايي نويسه‌هاي عربي با استفاده از شبكه عصبي و آناليز آماري ارائه شده است. شبكه عصبي با استفاده از الگوريتم حداقل ميانگين مربعات آموزش ديده است. در سيستم پيشنهادي هر نويسه به صورت ماتريس باينري به عنوان ورودي داده مي‌شود و سيستم شروع به استخراج ويژگيها مي‌كند. اين روش همواره سعي در پايين آوردن ميانگين مجذور خطا دارد. در اين مقاله ادعا شده است كه سيستم پيشنهادي ميزان موفقيت بالاتري از سيستمهاي مشابه دارد ولي هيچ آماري از درصد شناسايي صحيح و نام پايگاه داده ذكر نشده است. در [38] شناسايي نويسه‌هاي دستنويس عربي با استفاده از چند طبقه‌بندي بر اساس فرم حرف بيان شده است. در اين روش از 5 طبقه‌بندي DQDA، DLDA، LDA، QDA و PCA براي استخراج 25 ويژگي استفاده شده است. اين طبقه‌بندها به صورت خطي و درجه دوم پياده‌سازي شده‌اند. براي آزمايش از پايگاه داده‌اي كه توسط 48 نفر در سنين و جنسيت‌ها مختلف جمع‌آوري شده است، استفاده گرديده است. نتايج بازشناسي صحيح براي خطي 84% و براي درجه دوم 87% بدست آمده است. در [39] يک ركورد جديد در شناسايي ارقام دستنويس فارسي بر روي پايگاه‌هاي داده CENPARMI و IFHCDB بدست آمده است. در اين روش از تركيب 6 نوع روش طبقه بندي CFPC، PNC، DLQDF، MQDF، MLP و SVM استفاده شده است. اين روش بر روي تصاوير خاكستري كار مي‌كند و تصاوير را در ابعاد 35x35 نرمال‌سازي مي‌نمايد. ميزان شناسايي صحيح در اين مقاله 99.42% ذكر شده است. در [14] شناسايي نويسه‌هاي دستنويس فارسي با استفاده از شبكه عصبي ارائه شده است. در اين روش از يک شبكه عصبي پرسپترون چند لايه با يک لايه مخفي و الگوريتم يادگيري خطاي پس از انتشار استفاده شده است. تصاوير در ابعاد 30x30 نرمال مي‌شوند سپس به عنوان ورودي به شبكه عصبي ارسال مي‌گردند. در اين مقاله هيچ اشاره‌اي به پايگاه داده مورد استفاده يا تعداد نمونه‌هاي مورد آزمايش اشاره نشده است. آمار ارائه شده براي ميزان شناسايي صحيح 85% گزارش شده است. در [40] كه مربوط به مسابقات شناسايي نويسه‌هاي دستنويس فارسي/عربي مي‌باشد، 4 روش ارائه شده با هم مقايسه مي‌شوند. اين روشها به ترتيب براي CEMATER-JU، MDLSTM، REGIM و ECA مي‌باشد كه در شناسايي ارقام بهترين نتيجه بازشناسي براي ECA با 95.9% و در شناسايي حروف بهترين بازشناسي براي MDLSTM با 91.85% ذكر شده

است. در [41] شناسايي اعداد دستنويس فارسي با استفاده از ساختار پراكندگي و نمايندگيهاي كامل بر پايه شبكه عصبي مورد بحث و بررسي قرار گرفته است. الگوريتم K-SVD براي آموزش شبكه عصبي مورد استفاده قرار گرفته است. ميزان باز شناسي صحيح در اين روش 99.12% اعلام شده است. در [42] شناسايي اعداد دستنويس عربي توسط دو روش پيادهسازي و مقايسه شده است. روش اول با استفاده از SVM و طبقهبندي چند مقايسهاي و روش دوم بر اساس مدل مخفي ماركوف پيادهسازي شده است. در يک پايگاه دادهاي با 21120 نمونه تصوير از ارقام، 75% جهت آموزش و 25% براي آزمايش استفاده شده است. نتيجه شناسايي صحيح به روش اول 99.83% و به روش دوم 99.00% درصد گزارش شده است. در [43] يک موتوربازشناسي براي شناسايي اعداد چاپي فارسي بر اساس استنتاج فازي- عصبي ارائه شده است. در اين روش حد 98% بازشناسي صحيح اعلام شده است. در [44] يک روش جديد پيادهسازي FPGA براي شناسايي اعداد دستنويس فارسي آورده شده است. در اين روش از يک شبكه عصبي MLP دو لايه با 16 نرون در لايه مخفي استفاده شده است. براي پيادهسازي از چند جمع و تفريق كننده كه باعث بالابردن سرعت پردازش ميشود استفاده گرديده است. تصاوير به اندازههاي 40x40 نرمال ميشوند سپس به عنوان ورودي به شبكه عصبي ارسال ميشوند. ميزان شناسايي صحيح در اين روش 96% اعلام شده است. در [45] چهار چوبي جديد براي شناسايي اعداد دستنويس فارسي و لاتين آورده شده است. اساس اين روش بر مقياس انحناي ويژگيهاي فضايي ميباشد. در مقاله ارائه شده براي طبقهبندي از ماشينهاي بردار پشتيبان استفاده شده است. الگوريتم جديدي بر اين اساس به نام خوشهبندي ارائه شده است. براي استخراج ويژگيها از PCA و براي تجزيه و تحليل خطي از الگوريتم LDA استفاده گرديده است. با آموزش 660 نمونه و آزمايش بر 200 نمونه، درصد شناسايي صحيح 95.6% اعلام شده است. در [46] شناسايي دستنويس اعداد فارسي بر اساس تلفيق سيستمهاي خبره RBF ارائه شده است. در اين روش 4 شبكه عصبي RBF براي استخراج 81 ويژگي و يک شبكه عصبي RBF براي جداسازي بين ورودیهاي 4 شبكه عصبي ديگر پيادهسازي شده است. 6000 نمونه از پايگاه داده هدي جهت آموزش و 2000 نمونه براي آزمايش استفاده شده است. ميزان شناسايي صحيح 93.5% اعلام شده است. در اين مقاله ذكر شده است كه اين روش در مقايسه با تلفيق سيستمهاي شبكه عصبي MLP، 3.75% بهبود داشته است. در [47] شناسايي اعداد فارسي چاپي بر اساس المانهاي هندسه مركزي و شبكه فازي- عصبي Min-Max ارائه شده است. ميزان شناسايي صحيح 99.16% اعلام گرديده است. در [48] شناسايي حروف و ارقام دستنويس عربي با استفاده از توابع عطفي فازي ارائه شده است. در اين روش از اندازهگيري عدم تجانس موثر براي مقايسه اشكال چند ضلعي در تجزيه و تحليل شكل استفاده شده است. نكته مورد توجه در اين روش پيچيدگي محاسباتي بالا ميباشد. در مجموعهاي متشكل از

1948 حرف، 70% آن و در مجموعه‌اي متشكل از 70000 رقم، 85% جهت آموزش در نظر گرفته شده است. ميزان بازشناسي صحيح براي حروف 98% و براي ارقام 97.18% گزارش شده است. در [49] طبقه‌بندي چند سطحي در شناسايي حروف دستنويس عربي مطرح شده است. در اين مقاله به شناسايي هر 4 مدل حرف يعني حروف ابتدايي كه در ابتداي كلمه، حروف مياني كه در وسط كلمه، حرف انتهايي كه در انتهاي كلمه و حرف مجزا كه به تنهايي ظاهر مي‌شوند، پرداخته شده است. در اين روش 100 كلاس جهت طبقه‌بندي پياده‌سازي شده است. پايگاه داده‌اي كه جهت آموزش و آزمايش بايد باشد، معرفي نشده است و فقط از ميزان 75% نمونه براي آموزش و 25% نمونه براي آزمايش صحبت شده است. ميزان بازشناسي صحيح براي حروف ابتدايي 84%، حروف مياني 89%، حروف انتهايي 85% و حروف مجزا 93% اعلام شده است. در [50] شناسايي اعداد دستنويس فارسي با استفاده از منطق فازي مورد بحث واقع شده است. در اين روش ابتدا به قطعه‌بندي نويسه‌ها سپس استخراج ويژگيهاي آماري، ساختاري و هندسي شكل پرداخته شده و در نهايت با استفاده از منطق فازي به شناسايي نويسه پرداخته شده است. در يک پايگاه داده متشكل از 5000 نمونه كه توسط نويسندگان مقاله جمع‌آوري شده است، ميزان بازشناسي صحيح 95% براي ورودي‌هاي با محدوديت و 74% براي ورودي‌هاي بدون هيچ محدوديتي بدست آمده است. در [51] دو روش براي شناسايي حروف دستنويس فارسي ارائه شده است. در روش اول ابتدا پيش‌پردازش، نويزگيري و باينري كردن تصوير صورت مي‌گيرد. سپس براي شناسايي نويسه به تحليل موجک در مرزهاي نويسه و استخراج ويژگي‌هاي لازم پرداخته مي‌شود. اين روش با استفاده از شبكه عصبي پياده‌سازي مي‌شود. درروش دوم حروف به 5 گروه با توجه به تعداد قطعات آن طبقه‌بندي مي‌شود. سپس با استخراج ويژگي‌هاي لحظه‌اي مناسب در هر گروه، طبقه‌بندي بر اساس قانون بيز صورت مي‌گيرد. در آخر براي شناسايي نهايي در هر دو روش از پس‌پردازش‌هاي لازم و استخراج ويژگيهاي ساختاري و آماري نويسه و درخت تصميم‌گيري استفاده مي‌شود. پايگاه داده‌اي براي آزمايش در اين روش معرفي نشده است. به طور خلاصه مي‌توان موارد ذكر شده را در جدول 4-7 نمايش داد.

جدول 4-7 خلاصه شناسايي روشهاي ديگر

ردیف - مرجع	الگوریتم شناسایی	سال انجام	پایگاه داده و تعداد نمونه آن	درصد شناسایی	تعداد آموزش	نوع نویسه

1[29]	آماري	1373	مخصوص به خود- 12778	%95.82	%60	رقم
2[15]	آماري	1373	مخصوص به خود- 6600	%90.64	%50	حرف
3[30]	فازي	1374	مخصوص به خود- 4950	%92	3200	حرف
4[31]	شبکه عصبي	1383	مخصوص به خود- 6080	%91.70	3200	حرف
5[31]	شبکه عصبي	1383	مخصوص به خود- 3840	%90.90	2240	رقم
6[32]	شبکه عصبي	1383	مخصوص به خود- 2430	%91	1900	رقم
7[10]	ساختاري	1383	نا مشخص	%95	X	رقم
8[16]	تطابق شکل	1384	مخصوص به خود- 1288	%89.9	X	رقم
9[33]	شبکه عصبي	1384	هدي	%99.02	60000	رقم
10[34]	شبکه عصبي	1997	مخصوص به خود- 16000	%96.92	10000	حرف
11[35]	SVM	2003	CENPARMI	%94.41	7390	رقم
12[37]	شبکهعصبي - آماري	2007	نا مشخص	X	X	رقم
13[38]	شبکه عصبي	2008	مخصوص به خود-	%87	X	حرف
14[39]	شبکه عصبي	2009	IFHCDB، CENPARMI	%99.42	X	رقم
15[14]	شبکه عصبي	2009	نا مشخص	%85	X	حرف
16[41]	شبکه عصبي	2009	نا مشخص	%99.12	X	رقم
17[42]	SVM	2009	مخصوص به خود- 21120	%99.83	%75	رقم
18[44]	شبکه عصبي	2010	نا مشخص	%96	X	رقم
19[45]	SVM	2010	مخصوص به خود- 660	%95.6	220	رقم

46]20[	شبكه عصبي	2010	6000 نمونه هدي	%93.50	2000	رقم
48]21[	فازي	2010	مخصوص به خود- 1948	%98	%70	حرف
48]22[	فازي	2010	مخصوص به خود- 70000	%97.18	%85	رقم
49]23[	شبكه عصبي	2011	نا مشخص	%93	%75	حرف
50]24[	فازي- آماري	2011	مخصوص به خود-5000	%95	X	رقم
51]25[	شبكه عصبي	2011	نا مشخص	X	X	حرف
26روش پيشنهادي	ساختاري- فازي	2012	IFHCDB، هدي، متاكد 132796	%96.73	بدون آموزش	رقم
27 روش پيشنهادي	ساختاري- فازي	2012	IFHCDB، هدي، متاكد	%93.48	بدون آموزش	حرف

در مقايسه روش‌هاي توضيح داده شده و روش پيشنهادي شايد نتوان قياس درستي را بيان نمود زيرا اولا در مقايسه دو كار با هم، بايد نتيجه آزمايشات بر روي يک پايگاه داده مشترک صورت بپذيرد كه در مقالات معرفي شده در بالا به غير از چندتا از آنها مابقي بر روي پايگاه‌هاي داده مخصوص به خود نتيجه را اعلام كرده‌اند. ثانيا اكثر طرح‌هاي ارائه شده بر اساس آموزش سيستم مي‌باشد و با توجه به اينكه بخشي از پايگاه داده را كه آموزش داده شده است نبايد به عنوان مجموعه آزمايش قرار داد، لذا مقدار كمي از داده‌ها براي آزمايش باقي مي‌مانند. همانطور كه ذكر شده است به طور متوسط 65% از كل داده‌هاي پايگاه داده مورد آموزش قرار مي‌گيرد و 35% باقيمانده براي آزمايش استفاده مي‌شوند. اين درحاليست كه در سيستم‌هاي مبتني بر شناسايي ساختاري، سيستم هيچ آموزشي نمي‌بيند و در نتيجه كل داده‌هاي پايگاه داده مي‌توانند براي آزمايش مورد استفاده قرار گيرند. بديهي است كه درصد شناسايي بدست آمده بر روي كل پايگاه داده ارزشي درست‌تر و دقيق‌تر نسبت به آزمايش بر روي 25% داده‌هاي پايگاه داده دارد. به عبارت ديگر هر چه تعداد نمونه‌هاي آزمايش بيشتر باشد، ميزان بازشناسي صحيح‌تر خواهد بود. به طور مثال اگر 10 نمونه آزمايشي را به يک موتور بازشناسي بدهيم و موتور بازشناسي هر 10 نمونه را به درستي تشخيص دهد بايد نتيجه گرفت كه

ميزان بازشناسي صحيح در اين موتور بازشناسي 100% است كه واضح است چنين نتيجه‌گيري صحيح نمي‌باشد.

در روش ارائه شده در اين پايان‌نامه براي شناسايي ارقام در پايگاه‌هاي داده هدي تمامي 102353 نمونه، IFHCDB تمامي 17740 نمونه و متاكد تمامي 12703 نمونه يعني در مجموع 132796 نمونه مورد آزمايش قرار گرفته‌اند و درصد شناسايي 96.73% بدست آمده است. به طور مشابه اين روند براي حروف نيز بوده است.

در مقايسه روش پيشنهادي با ساير كارها بر روي پايگاه‌داده‌هاي مشترک نتيجه بدست آمده بدين ترتيب بود كه براي شناسايي ارقام بر روي پايگاه‌داده هدي در[33] ميزان 99.02% و در [46] ميزان 93.50% گزارش شده است. اين در حاليست كه روش پيشنهادي بر روي كل داده‌هاي هدي 96.44% مي‌باشد. در [39] ميزان 99.42% در مجموع پايگاه‌هاي‌داده IFHCDB و CENPARMI اعلام شده است و در روش پيشنهادي ميزان 98.56% بر روي تمامي داده‌هاي IFHCDB مي‌باشد. در زمينه شناسايي حروف اعلامي بر اساس پايگاه داده مشترک پيدا نشد.

در نوع ديگري از مقايسه، ميانگين كلي درصد بازشناسي صحيح از تمامي مقالات ذكر شده در شناسايي ارقام و حروف محاسبه گرديد. ميانگين كلي ديگر كارها در شناسايي ارقام 95.14% و در شناسايي حروف 91.64% مي‌باشد. اين در حاليست كه ميانگين كلي در روش پيشنهادي براي شناسايي ارقام 96.73% و براي شناسايي حروف 93.48% مي‌باشد.

در نهايت با توجه به ميزان بازشناسي صحيح روش‌هاي ذكر شده، مي‌توان گفت كه روش پيشنهادي به نسبت روش‌هاي بررسي شده از ميزان بازشناسي خوبي برخوردار مي‌باشد.

فصل پنجم

جمع‌بندي و پيشنهادها

5-1 جمع‌بندي پايان‌نامه

بازشناسي حروف و ارقام دستنويس همواره يكي از موضوعات مورد علاقه براي تحقيق بوده است. كارهاي بسيار زيادي در اين زمينه صورت گرفته است. با توجه به پيچيدگي حروف و ارقام فارسي، راه حل‌هاي شناسايي نويسه‌هاي دستنويس براي اين زبان مشكل به نظر مي‌آيد. تكنيك‌هاي گوناگوني براي شناسايي الگو وجود دارد كه مهمترين آنها عبارتند از: روش ساختاري، شبكه عصبي، آماري، ژنتيک و منطق فازي. هر كدام از اين روش‌ها داراي مزايا و معايب مربوط به خود مي‌باشد. به طور مثال شبكه‌هاي عصبي از سرعت بهتري به نسبت روشهاي ساختاري برخوردار هستند زيرا قبلا آموزش ديده‌اند در حاليكه روش ساختاري براي هر نويسه ورودي ويژگي‌هاي ساختاري آن را در همان لحظه استخراج مي‌كند.

در اين پايان‌نامه اكثر كارهايي كه در اين زمينه پياده‌سازي شده، معرفي گرديده است. همچنين يک روشي نوين پيشنهاد داده شد كه به نسبت ديگر كارهاي انجام شده به صورت ميانگين از بازدهي بهتري برخوردار است. در روش پيشنهادي از تركيب سه روش شبكه عصبي، ساختاري و منطق فازي استفاده شده است. شناسايي بخش‌هاي زمان‌گير به شبكه عصبي سپرده شده و از طرفي از روش ساختاري مبتني بر منطق فازي براي استخراج ويژگيها و شناسايي نهايي استفاده شد.

روش ارائه شده بر پايه تكنولوژي .Net و با زبان C# پياده‌سازي گرديده است.

5-2 پيشنهادها

با توجه به تحقيقات وسيع انجام شده و در حال انجام مشخص مي‌شود كه كار بر روي اين زمينه هنوز در ابتداي مسير خود است و تا رسيدن به نتايج ايده‌آل راه فراواني باقيمانده است. يكي از مهمترين بخشهاي شناسايي نويسه‌ها، پيش پردازش مي‌باشد. كار بر روي اين بخش بايد با بازدهي بسيار بالا و دقيق صورت بگيرد. بحث نويزگيري از جمله كارهاي مهم است كه بايد به درستي صورت گيرد، مخصوصا در مورد حروف به دليل آنكه داراي نقطه مي‌باشند و ممكن است كه نقطه به اشتباهي به عنوان نويز شناسايي شود. اين خطا در انسان بسيار كمتر مي‌باشد و دليل آن ساير اطلاعاتي است كه انسان از بقيه محيط كسب مي‌كند.

بحث ديگر مربوط به نازک‌سازي نويسه مي‌باشد. علي‌رغم اينكه الگوريتمهاي فراواني در اين زمينه ارائه شده است ولي همچنان مشكلات فراواني در زمينه نازک‌كردن نويسه وجود دارد كه در متن نيز به آنها اشاره شد. به طور كلي به نظر نمي‌آيد كه عمل نازک‌سازي براي شناسايي نويسه، عمل صحيحي باشد زيرا انسان اين كار را انجام نمي‌دهد ولي با توجه به اينكه نمي‌توان قابليت‌هاي انسان را با ماشين مقايسه كرد گاهي مجبور به انجام اين كار مي‌شويم. گاهي اوقات ذات الگوريتم طوري طراحي مي‌شود كه فقط بر روي نويسه‌هاي نازک شده مي‌تواند كار كند. الگوريتم ارائه شده در اين پايان‌نامه نيز بر اساس نازک كردن نويسه عمل شناسايي را انجام مي‌دهد.

مورد مهم ديگر سرعت شناسايي نويسه مي‌باشد كه بستگي به تكنيک به كار برده دارد. همچنين با توجه به چند هسته‌اي شدن پردازنده كامپيوترها مي‌توان براي افزايش سرعت از تكنيک چند نخي استفاده نمود كه جدا از دقت در پيادەسازي برنامه‌هاي چند نخي در مورد برنامه‌هاي شناسايي الگو بايد امكان پيادەسازي آن در الگوريتم وجود داشته باشد. پيشنهاد ديگر براي بالا بردن سرعت استفاده از زبان‌هاي سطح پايين و نزديک به سطح ماشين همانند C++ و بدون هيچگونه وابستگي به چهارچوبي مثل .Net مي‌باشد.

مراجع

1. http://www.farsiocr.ir/ocr-tutorials/47-ocr-related-defenitions, 89/02/12.

2. http://www.citexco.com/SOMRDesc.aspx, 88/11/15.

3. http://www.farsiocr.ir/ocr-tutorials/47-ocr-related-defenitions?start=2 , 89/02/13.

4. http://www.farsiocr.ir/ocr-tutorials/47-ocr-related-defenitions?start=1, 89/02/14.

5. K. Jumari, M. Ali, "A Survey and Comparative Evaluation of Selected Off-Line Arabic Handwritten Character Recognition Systems", Journal Technology, pp: 1-18, June 2002.

6. http://www.aforgenet.com/ , 89/02/20.

7. H. Bunke and B. Haller, "Syntactic Analysis of Context-Free Plex Languages for Pattern Recognition", in Structured Document Image Analysis, 1978.

8. K. Chan, D. Yeung, "Recognizing on-line handwritten alphanumeric characters through flexible structural matching", Pattern Recognition, Vol 32, pp 1099-1114, 1999.

9. R. C. Gonzalez, M. G. Thomason, "Syntactic Pattern Recognition An Introduction", Addison-Wesley Publishing, 1978.

10. م. رحمتي، م. زيبائي فرد، "يک روش ساختاري براي شناسايي حروف و ارقام دست نويس فارسي"، سومين کنفرانس ماشين بينايي و پردازش تصوير ايران 1383.

11. H. Freeman, "Computer processing of line drawing images", ACM Computing Surveys, 6(1): 57- 98, March 1974.

12. H. Bunke and B. Haller, "A parser for context free Plex grammars", In M. Nagl, editor, Proceedings of the 15th International Workshop on Graph-Theoretic Concepts in Computer Science- WG'89, number 411 in Lecture Notes in Computer Science, pages 136- 150. Springer- Verlag, 1989.

13. س. م. رضوي و ا. کبير، "بازشناسي بر خط حروف مجزاي فارسي با شبکه عصبي"، ارائه شده درسومين کنفرانس ماشين بينايي و پردازش تصوير ايران 1383.

14. R. Ghoraie, M. Farajpoor, "Handwritten Farsi Character Recognition using Artificial Neural Network", International Journal of Computer Science and Information Security. Vol. 4, No. 1& 2 2009.

15. ر. عزمي، ا. كبير، " شناسايي آماري حروف دستنويس فارسي"، مجموعه مقالات كنفرانس مهندسي برق ايران 1373.

16. ع. درويش، ا. كبير، ح. خسروي، "كاربرد تطابق شكل در بازشناسي ارقام دستنويس فارسي"، مجله فني و مهندسي دانشگاه تربيت مدرس 1384.

17. ا.كبير، ک.بهاري، م.احمدزاده، "بازشناسي متون تايپ شده فارسي"، مجموعه مقالات كنفرانس مهندسي برق ايران. جلد دوم، 1372.

18. J. T. Tou, R. C. Gonzalez, "Pattern Recognition Principles", Addison Wesley Publishing Co. Reading, Massachusetts 1974.

19. D. Cheng, H. Yan, "Recognition of Handwritten Digits Based on Contour Information". Pattern Recognition, Vol. 31, No. 3, pp. 235- 255, 1999.

20. S. Chakravarthy, B. kompella, "The Shape of handwritten characters", Pattern Recognition Letters, Vol. 24, pp. 1901- 1913, 2003.

21. A. Amin, "Off-Line Arabic Character Recognition: The State of the Art", Pattern Recognition, Vol. 31: No. 5; 1998.

22. H. Khosravi, E. Kabir, "Introducing a very large dataset of handwritten Farsi digits and a study on their varieties", Pattern Recognition Letters Vol. 28, 1133–1141, 2007.

23. H. Alamri, J. Sadri, Y. Suen, N. Nobile, "A Novel Comprehensive Database for Arabic Off-Line Handwriting Recognition", CENPARMI (Center for Pattern Recognition and Machine Intelligence), Computer Science and Software Engineering Department, Concordia University, 2008.

24. حسين خسروي "بازشناسي ارقام و حروف دستنويس فارسي در فرمهاي ثبت نام آزمون سراسري"، پايان نامه كارشناسي ارشد الكترونيك، دانشگاه تربيت مدرس، 1384.

25. S. Mozaffari, K. Faez, F. Faradji, M. Ziaratban and S.M. Golzan "A comprehensive isolated Farsi/Arabic character database for handwritten OCR research", Proc. 10th International Workshop on Frontiers in Handwriting Recognition (IWFHR) , La Baule, France , pp 385-389, 2006.

26. F. Solimanpour, J. Sadri, C.Y. Suen, “Standard databases for recognition of handwritten digits, numerical strings, legal amounts, letters and dates in Farsi language,” Proc. 10th IWFHR, La Baule, France, 2006.

27. http://www.MetaCodeResearch.com 1389.

28. http://www.CitexCo.com 1389.

29. ح. نفيسي، ا. كبير، "شناسايي ارقام دستنويس فارسي"، مجموعه مقالات دومين كنفرانس مهندسي برق ايران 1373.

30. ک. مسروري، ا. كبير، " بازشناسي حروف دستنويس فارسي با روش فازي"، مجموعه مقالات سومين كنفرانس سالانه مهندسي برق ايران 1374.

31. س. مظفري، ر. صفابخش، " شناسايي حروف و اعداد دستنويس فارسي مبتني بر شبكه عصبي توابع پايه شعاعي و ويژگيهاي فركتالي"، مجموعه مقالات سومين كنفرانس ماشين بينايي و پردازش تصوير ايران 1383.

32. ح. نبوي، ر. ابراهيمپور، ا كبير، " كاربرد تطبيق طبقه بندها در بازشناسي ارقام دستنويس فارسي"، مجموعه مقالات سومين كنفرانس ماشين بينايي و پردازش تصوير ايران 1383.

33. ا.كبير، ح. خسروي، " معرفي دو ويژگي سريع و كارآمد براي بازشناسي ارقام دستنويس فارسي"، مجله فني و مهندسي دانشگاه تربيت مدرس 1384.

34. M. Dehghan, K. Faez, "Farsi handwritten, character recognition with moment invariants", Proc. of 13th international conference of digital signal processing, Vol. 2 pp. 507-510, 1997.

35. J. Sadri, Y. Suen, D. Bui, "Application of Support Vector Machines for Recognition of Handwritten Arabic Persian Digits", 2nd Machine Vision and Image Processing Conference, K. N. Toosi Univ. of Tech, Tehran, Iran, Feb, 2003.

36. M. Lorigo, Member, IEEE, V. Govindaraju, Senior, IEEE, " Off-line Arabic Handwriting Recognition A Survey", Appear In IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006.

37. Ahmad M. Sarhan, and Omar I. Al Helalat, "Arabic Character Recognition using Artificial Neural Networks and Statistical Analysis", Proceedings of World Academy of Science, Engineering and Technology, Vol. 21, MAY 2007.

38. Gheith A. Abandah, Khaled S. Younis, "Handwritten Arabic Character Recognition Using Multiple Classifiers Based on Letter Form", In Proc. 5th IASTED Int'l Conf. on Signal Processing, Pattern Recognition, & Applications (SPPRA), Innsbruck, Austria, Feb 2008.

39. C. Lin Liu, Y. Suen," A New Benchmark on the Recognition of Handwritten Bangla and Farsi Numeral Characters", Pattern Recognition, Vol. 42, No. 12, 2009.

40. S. Mozaffari, H. Soltanizadeh, "ICDAR 2009 Handwritten Farsi Arabic Character Recognition Competition", 10th International Conference on Document Analysis and Recognition, 2009.

41. W. M. Pan, T. D. Bui, C. Y. Suen, "Isolated Handwritten Farsi numerals Recognition Using Sparse And Over-Complete Representations", 10th International Conference on Document Analysis and Recognition, 2009.

42. A. Mahmud, M. Awaida, "Recognition of Off-Line Handwritten Arabic (Indian) Numerals Using Multi-Scale Features and Support VECTOR MACHINES VS. HIDDEN MARKOV MODELS", the Arabian Journal for Science and Engineering, Volume 34, Number 2B, 2009.

43. G. A. Montazer, H. Q. Saremi, V. Khatibi, "A Neuro-fuzzy Inference Engine for Farsi Numeral Characters Recognition", Published in Expert Systems with Applications, Vol. 37, No. 9, pp : 6327-6337, 2010.

44. M. Moradi, M. A. Pourmina, F. Razzazi, "A New Method of FPGA Implementation of Farsi Handwritten Digit Recognition", European Journal of Scientific Research. Vol. 39 No. 3, pp: 309- 315, 2010.

45. A. Behrad, M. Khoddami, M. Salehpour, "A Novel Framework for Farsi and Latin Script", Journal of Automatic Control, University of Belgrade, Vol. 20:17-25, 2010.

46. R. Ebrahimpour, A. Esmkhani, S. Faridi, "Farsi handwritten digit recognition based on mixture of RBF experts", IEICE Electronics Express , Vol. 7, No. 14, 1014- 1019, 2010.

47. H. R. Boveiri, "Persian Printed Numeral Characters Recognition Using Geometrical Central Moments and Fuzzy Min-Max Neural Network", International Journal of Signal Processing Vol. 6, No. 2, 2010.

48. M. T. Parvez, Sabri A. Mahmud, "Arabic Handwritten Alphanumeric Character Recognition using Fuzzy Attributed Turning Functions", FAHR 2010: Workshop in Frontiers in Arabic Handwriting Recognition, in conjunction with 20th International Conference in Pattern Recognition (ICPR), pp. 9-14, August 22-26, Istanbul, Turkey, 2010.

49. R. I. Zaghloul, E. F. AlRawshedeh, D. M. K. Bader, "Multilevel Classifier in Recognition of Handwritten Arabic Characters", Journal of Computer Science Vol. 7, No. 4, pp: 512-518, 2011.

50. M. Sharifzadeh, S. Alirezaee, "Recognition of Farsi Handwritten Numbers Using the Fuzzy Method", International Journal of Image Processing (IJIP), Vol. 5, No. 5, 2011.

51. M. R. Jenabzadeh, R. Azmi, B. Pishgoo, S. Shirazi, "Two Methods for Recognition of Hand Written Farsi Characters", International Journal of Image Processing (IJIP), Vol. 5,No. 4, 2011.

Abstract

This paper proposes a hybrid algorithm of neural networks and fuzzy logics. The algorithm goes to distinguish discrete Farsi handwriting characters. It divides handwriting characters to some segments and achieving this state, it uses a sequence of rules and circumstances. Segmentation, itself, has some sensitive points. The values of sensitive points are used as input for a neural network. What is made stable by the neural network as the true answer will be sent to the final part of the recognition algorithm. In the final part, a structural pattern is defined which, using fuzzy logics, determines how much the handwriting character is like to one specific character. Based on the value of fuzzy similarity, all characters except one will be refused. If there is no character respective to handwriting, then some of its segments merge with each other based on the rules. The merged object again passes the steps of algorithms. Algorithm continues the iterations until the result will be taken or the characters will be finished. Applying the algorithm on three databases, the authors' database, IFHCDB and HODA, the ratio of 93.48% for letters and 96.73% for digits gained as efficiency of recognition.

Keywords: Optical Character Recognition (OCR), discrete letters, handwriting digits, structural method, Neural Networks, fuzzy logics.

Iran University of Science and Technology

A New Structural Method for Farsi Isolated Handwritten Character Recognition

A Thesis Submitted in Partial Fulfillment of the Requirement for the Degree of Master of Science in Information Technology

By

Babak Sairafi

Supervisor

Dr. M. Soryani

Advisor

Dr. E. Kabir

April 2012