

## بازشناسی حروف چاپی با استفاده از ویژگیهای منحنی پیرامونی<sup>x</sup>

رضا عزمی<sup>\*</sup>، احسان... کبیر<sup>\*\*</sup>، کامبیز بدیع<sup>\*\*\*</sup>

<sup>\*</sup>گروه کامپیوتر، دانشگاه الزهرا  
<sup>\*\*</sup>بخش مهندسی برق، دانشگاه تربیت مدرس  
<sup>\*\*\*</sup>مرکز تحقیقات مخابرات ایران

### چکیده

در این مقاله دو الگوریتم مختلف برای شناسایی حروف چاپی فارسی با قلمهای متنوع ارائه شده است. الگوریتم اول از کدهای فریمین مربوط به کانتور حروف به عنوان ویژگی و از مجموعه‌ای از اتوماتون‌های آماری برای طبقه‌بندی استفاده می‌کند. الگوریتم دوم از تبدیل هاف فازی برای استخراج ویژگی از کانتور حروف و از شبکه عصبی پرسپترون برای طبقه‌بندی استفاده می‌کند. بازشناسی نهایی حروف با توجه به نقاط و علائم آنها و استفاده از ویژگیهای ساختاری مناسب انجام می‌شود.

این دو الگوریتم روی مجموعه‌ای از حروف شامل ۱۱۵۰۰ نمونه از ده قلم مختلف آزموده شده و به ترتیب ۹۷/۱۳ و ۹۸/۳۲ درصد از حروف به درستی بازشناسی شده‌اند.

کلمات کلیدی: بازشناسی حروف چاپی، استخراج ویژگی، منحنی پیرامونی، اتوماتون آماری، تبدیل هاف فازی، شبکه عصبی، پرسپترون.

### ۱- مقدمه

پیشرفت تکنولوژی کامپیوتر و کاهش زمان پردازش باعث شده است که بازشناسی نوری حروف، که یکی از شاخه‌های مورد توجه شناسایی الگو است، کاربرد بیشتری پیدا کند. حجم بسیاری از اطلاعات موجود بصورت مکتوب است. خودکار کردن عملیات مربوط به پردازش این اطلاعات احتیاج به یک سیستم واسط برای وارد کردن و بازشناسی آنها دارد. این وظیفه را الگوریتمهای شناسایی حروف در قالب سیستمهای بازشناسی نوری حروف به انجام می‌رسانند.

درباره بازشناسی اتوماتیک حروف لاتین، چینی و ژاپنی تحقیقات زیادی صورت گرفته و سیستم‌های تجاری نیز به بازار ارائه شده‌اند [۱، ۲، ۳]. برای بازشناسی حروف فارسی و عربی نیز تحقیقاتی صورت گرفته است [۴-۲۴]. تفاوت‌های ساختاری مهمی مانند اتصال حروف در کلمه، وجود نقاط و علائم در بالا یا پایین حروف و همپوشانی حروف با یکدیگر که بین نگارش فارسی و لاتین وجود دارد، کاربرد مستقیم روشهای شناسایی متون لاتین برای خط فارسی را با مشکل روبرو می‌کند.

در این مقاله دو الگوریتم مختلف به منظور شناسایی مجموعه حروف چاپی فارسی با قلم‌های متنوع مطرح می‌شود. این حروف بعد از اجرای الگوریتم جداسازی ایجاد شده‌اند [۲۵، ۲۶ و ۲۷]. الگوریتم اول از کدهای فریمین مربوط به کانتور

بیرونی حروف به‌عنوان ویژگی و از اتوماتون‌های آماری برای طبقه‌بندی استفاده می‌کند. الگوریتم دوم از تبدیل هاف فازی در مرحله استخراج ویژگی و از شبکه عصبی در مرحله طبقه‌بندی استفاده می‌کند. هر دو الگوریتم ویژگیهای ورودی خود را از تمام یا بخشی از کانتور بیرونی حروف، با توجه به موقعیت حرف در کلمه، استخراج می‌کنند. چون کانتور بیرونی قبلاً در مرحله جداسازی استخراج شده است، زمان لازم برای اجرای الگوریتم بازشناسی کاهش می‌یابد.

در بخش ۲ گروههای مختلف حروف و نحوه انتخاب کانتور مناسب آنها بررسی شده است. بخش‌های ۳ و ۴ به ترتیب الگوریتم‌های اول و دوم را مطرح می‌کنند. در بخش ۵ چگونگی استفاده از ویژگیهای ساختاری برای بازشناسی نهایی حروف توضیح داده شده است. بخشهای ۶ و ۷ به آزمایش و ارزیابی دو الگوریتم و نتیجه‌گیری نهایی می‌پردازند.

### ۲- خصوصیات حروف و نحوه انتخاب کانتور مناسب آنها

۳۲ حرف الفبای فارسی با توجه به موقعیت آنها در کلمه اشکال مختلفی به خود می‌گیرند. بر این اساس آنها را به چهار گروه حروف ابتدایی، میانی، انتهایی و مجزا تقسیم می‌کنند. برای اینکه تعداد کلاسهای مورد نظر در مرحله بازشناسی کاهش یابد، با توجه به امکان شناسایی نقاط و علائم، می‌توان تنها بدنه حروف را در نظر

<sup>x</sup> این تحقیق در چارچوب طرح ملی بازشناسی متون چاپی و حجم محدودی از کلمات دستنویس انجام شده است.

گرفت. شکل ۱ گروههای مختلف حروف را نشان می‌دهد. در زیر به بررسی نحوه انتخاب کانتور مناسب در هر گروه می‌پردازیم.

**حروف ابتدایی**  
این گروه شامل ۲۵ حرف است که با حذف نقاط و علائم تعداد کلاسهای این گروه به ۱۲ کلاس بدنه کاهش می‌یابد. در مرحله جداسازی حروف، کانتور بالایی این حروف و نقاط جداسازی آنها تعیین شده است. در مرحله بازشناسی، کانتور پایینی این حروف نیز در نظر گرفته می‌شود شکل ۲ (الف).

**حروف میانی**  
این گروه شامل ۲۵ حرف است که با حذف نقاط و علائم تعداد کلاسها به ۱۲ کاهش می‌یابد. برای این حروف در مرحله بازشناسی از همان کانتور بالایی که در مرحله جداسازی استخراج شده است، استفاده می‌شود شکل ۲ (ب).

**حروف انتهایی**  
در این گروه ۳۲ حرف و ترکیب دو حرفی "لا" وجود دارد. تعداد کل کلاسهای بدنه در این گروه ۱۹ است. در این گروه کانتور پایینی نیز برای بازشناسی در نظر گرفته می‌شود شکل ۲ (ج).

**حروف مجزا**  
تعداد کلاسها در این گروه مانند حروف انتهایی ۳۲ است و تعداد کلاسهای بدنه در این گروه نیز ۱۹ است. در این گروه کل کانتور بدنه در مرحله بازشناسی در نظر گرفته می‌شود شکل ۲ (د).

ب پ ت ث ج چ ح خ س ش ص ض ط ظ ع غ ف ق ک گ ل م ن ه ی	گروه حروف ابتدایی تعداد حروف ۲۵ تعداد بدنه ها ۱۲
ب پ ت ث ج چ ح خ س ش ص ض ط ظ ع غ ف ق ک گ ل م ن ه ی	گروه حروف میانی تعداد حروف ۲۵ تعداد بدنه ها ۱۲
ا ب پ ت ث ج چ ح خ س ش ص ض ر ز ژ س ش ص ض ط ظ ع غ ف ق ک گ ل م ن و ه ی لا	گروه حروف انتهایی تعداد حروف ۳۲ تعداد بدنه ها ۱۹
ا ب پ ت ث ج چ ح خ س ش ص ض ر ز ژ س ش ص ض ط ظ ع غ ف ق ک گ ل م ن و ه ی لا	گروه حروف مجزا تعداد حروف ۳۲ تعداد بدنه ها ۱۹

شکل ۱- گروههای مختلف حروف



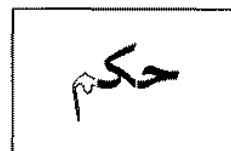
(ب)



(د)



(الف)



(ج)

شکل ۲- نحوه محاسبه کانتور برای

(الف) - حروف ابتدایی، (ب) - حروف میانی، (ج) - حروف انتهایی، (د) - حروف مجزا



استفاده شده است. میزان بازشناسی برای ۵۰۰ کلمه دستنویس که توسط ۵ خطاط معروف ایرانی نوشته شده است، ۹۵٪ بوده است. از گشتاورهای شبه زرنیکه بهمراه شبکه عصبی پرسپترون برای بازشناسی حروف و ارقام مجزای دستنویس نیز استفاده شده است که دقت بدست آمده برای اعداد ۹۸٪ و برای حروف ۹۹٪ گزارش شده است [۳۰].

در مقاله [۳۱] سیستمی برای بازشناسی کلمات چاپی عربی با استفاده از شبکه عصبی پیشنهاد شده است. ابتدا مؤلفه‌های متصل<sup>۲</sup> تصویر تعیین می‌شوند. سپس ویژگیهای عمومی مانند تعداد این مؤلفه‌ها، تعداد ماکزیمم‌ها در هر مؤلفه، تعداد و محل نقاط و علائم و ویژگیهای هندسی دیگری از شکل کلی کلمه استخراج می‌شود. از یک شبکه عصبی سه لایه برای بازشناسی کلمات استفاده می‌شود.

#### ۴-۱ استخراج ویژگیها

از تبدیل هاف برای استخراج ویژگی در کاربردهای مختلفی از بازشناسی الگو استفاده شده است. گونه های مختلف اصلاح شده و فازی این تبدیل برای افزایش سرعت، کاهش اثر نویز و استخراج ویژگیهای ساختاری معرفی شده اند [۳۲-۳۷]. در این تحقیق ویژگیهای ورودی با اعمال تبدیل هاف فازی بر روی تصویر کانتور حرف استخراج می شوند. علت استفاده از این تبدیل این است که می خواهیم از اندازه، زاویه و مکان نسبی پاره خطهای کانتور حرف برای شناسایی آن استفاده کنیم و این کار را تبدیل هاف بدون حساسیت به نویز موجود روی کانتور انجام میدهد.

در فضای تبدیل محدوده  $\theta$  بین  $-90^\circ$  و  $+90^\circ$  درجه است و محدوده  $\rho$  با استفاده از داده‌های ورودی از رابطه  $\pm \sqrt{2}D$  بدست می‌آید، که در آن  $D$  قطر مستطیل در بر گیرنده حرف ورودی است. محدوده  $\rho$  به پنج قسمت مساوی و محدوده  $\theta$  به هفت قسمت مساوی تقسیم می‌شود. سلولها بصورت  $(\rho_i, \theta_j)$  ،  $i=1, \dots, 5$  ،  $j=1, \dots, 7$  شماره گذاری شده اند. به ازای هر نقطه  $(x_0, y_0)$  از کانتور و مقدار  $\theta_j$  مقدار  $\rho_j$  از رابطه (۴) محاسبه می‌شود. در صورتیکه  $\rho_j$  در محدوده ناحیه  $i$  از محور  $\rho$  قرار گیرد، مقدار شمارنده سلول که با  $A(i, j)$  مربوط به ناحیه نشان داده می‌شود، یک واحد افزایش می یابد.

$$\rho_j = x_0 \cos(\theta_j) + y_0 \sin(\theta_j) \quad (4)$$

$$A(i, j) = A(i, j) + 1$$

#### ۴-۲ فازی کردن تبدیل هاف

یکی از مشکلاتی که در تبدیل هاف معمولی وجود دارد این است که اگر این تبدیل روی داده های نویزی اعمال شود، در فضای پارامترها بجای یک ماکزیمم مشخص چندین ماکزیمم محلی می‌خواهیم. دو راهکار برای این مسئله پیشنهاد شده است که هر دو آنها مفهوم فازی را در تبدیل هاف بکار می گیرند. در راهکار اول نقاطی که قرار است روی آنها تبدیل هاف صورت بگیرد، بصورت فازی در نظر گرفته می‌شوند. برای مثال در تبدیل یک نقطه، کلیه نقاط درون دایره ای حول آن نقطه را در نظر می‌گیرند. این روش باعث نرم شدن خروجی تبدیل هاف می‌شود، اما از لحاظ حجم محاسبات مزیت قابل ملاحظه‌ای نسبت به هموار سازی دو بعدی در حوزه تبدیل ندارد. تحقیقات انجام گرفته در این زمینه نشان میدهد که می‌توان حجم محاسبات در این روش را با تمهیداتی در حد یک هموار سازی یک بعدی کاهش داد [۳۲، ۳۳]. راهکار دیگری را که می توان برای فازی کردن تبدیل هاف بکار برد این است که سلولهای فضای پارامترها را بصورت فازی در نظر بگیریم. به عبارت دیگر وقتی قرار است محتوای یک سلول در فضای هاف افزایش یابد، محتوای سلولهای همسایه آن نیز از این افزایش متأثر شوند. به این روش تبدیل هاف با سلول فازی، FCHT<sup>۳</sup> گفته می‌شود. می توان نشان داد که حجم

بدر	م، ح، ک
میترا	م، ح، ک
نسیم	م، ح، ک
کامپیوست	م، ح، ک
رؤیا	م، ح، ک
زر	م، ح، ک

شکل ۶- قلمهای بکارگرفته شده در مجموعه داده‌ها

#### ۴- الگوریتم ۲: بازشناسی حروف با استفاده از تبدیل هاف فازی و شبکه عصبی

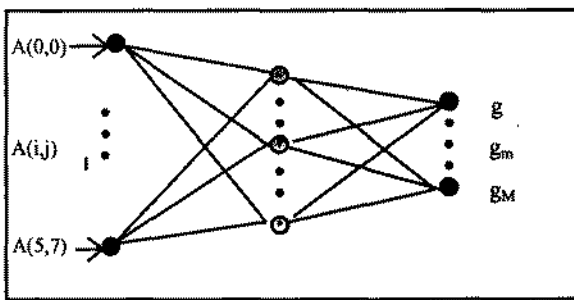
در این الگوریتم از تبدیل هاف فازی برای استخراج ویژگی و از یک شبکه عصبی پرسپترون برای بازشناسی استفاده شده است. نوآوری این روش در بکارگیری تبدیل هاف فازی برای استخراج ویژگی است. با استفاده از تبدیل هاف فازی، ویژگیهای مناسبی از کانتور حروف استخراج و به شبکه عصبی اعمال می‌شود. انواع شبکه های عصبی با ویژگیهای مختلف برای بازشناسی متون عربی و فارسی بکار گرفته شده اند. در زیر به برخی از مقالات ارائه شده در این زمینه اشاره می‌کنیم.

در مقاله [۱۲] یک رویکرد ساختاری برای بازشناسی حروف دستنویس عربی با استفاده از شبکه عصبی ارائه شده است. اسکلت حرف ورودی از سمت راست به چپ، با استفاده از یک پنجره ۳×۳ برای ایجاد یک گراف که نشان دهنده حرف باشد، دنبال می‌شود. آنگاه مجموعه‌ای از ویژگیهایی ساختاری مانند خطوط راست، انحناها و حلقه‌ها از این گراف استخراج می‌شوند. بنابراین هر حرف با مجموعه ای از این ویژگیها مانند نقطه، همزه، خط، انحنا و حلقه و رابطه بین آنها که در یک ماتریس نشان داده شده، توصیف می‌شود. طبقه بندی حروف با یک شبکه عصبی پنج لایه انجام می‌شود. لایه ورودی شبکه از ۱۵۰ گره تشکیل شده است. مجموعه داده‌ها شامل ۱۰۰۰ نمونه از حروف نوشته شده توسط ۱۰ نفر بوده است. نیمی از این نمونه‌ها برای آموزش و نیمی دیگر برای آزمایش الگوریتم بکار گرفته شده است. میزان بازشناسی ۹۲ درصد گزارش شده است.

در مقاله [۲۹] سیستمی برای بازشناسی متون چایی عربی با استفاده از ویژگیهای گشتاوری و شبکه عصبی ارائه شده است. ابتدا حروف کلمات با استفاده از هیستوگرام عمودی جدا می‌شوند. سپس شش ویژگی گشتاوری مختلف برای هر حرف استخراج می‌شود. برای بازشناسی از یک شبکه عصبی پرسپترون استفاده شده است.

در الگوریتم دیگری [۲۰] برای بازشناسی بدنه حروف از گشتاورهای شبه زرنیکه<sup>۱</sup> به عنوان ویژگی و از یک شبکه عصبی پرسپترون بعنوان طبقه بندی کننده

پس از اینکه از تصویر کانتور حرف ورودی تبدیل هاف گرفته شد، مقادیر ۲۵ سلول حوزه هاف به ورودی شبکه اعمال می‌شوند (شکل ۹). خروجی‌های گره‌های لایه انتهایی شبکه به عنوان توابع ممیز  $g_m(x)$  در نظر می‌گرفته می‌شوند و حرف ورودی به کلاسی منتسب می‌شود که خروجی آن بیشتر است. هرچه مقدار خروجی به یک نزدیکتر باشد، احتمال تعلق بدنه حرف ورودی به کلاس موردنظر بیشتر خواهد بود. برای شناسایی گروه‌های مختلف حروف (ابتدایی، میانی، انتهایی و مجزا) نیاز به چهار شبکه عصبی مختلف داریم. شبکه‌های عصبی مورد نظر از یک لایه ورودی، یک لایه پنهان و یک لایه خروجی تشکیل شده‌اند. تعداد گره‌های لایه ورودی ۳۵ و تعداد گره‌های لایه خروجی برحسب تعداد کلاس‌های هر گروه ۱۲ یا ۱۹ است. تعداد گره‌های لایه پنهان با سعی و خطا در مرحله آموزش تعیین شده اند و برای چهار گروه حروف ابتدایی، میانی، انتهایی و مجزا به ترتیب ۱۹، ۱۹، ۲۵ و ۲۵ است.



شکل ۹- شبکه عصبی بکار گرفته شده در الگوریتم ۲

۴-۴ آزمایش الگوریتم ۲

برای ارزیابی الگوریتم ۲ از همان مجموعه داده‌های الگوریتم ۱ استفاده شده است. میزان بازشناسی درست ۹۰/۵۲٪ است. میزان بازشناسی برای گروه‌های مختلف حروف در جدول ۴ آمده است.

۵-تعریف گزاره های شرط عضویت با استفاده از ویژگیهای ساختاری

در این مرحله گزاره های شرط عضویت مناسب، بر اساس ارتباط منطقی بین ویژگیهای ساختاری حروف و محل و نوع نقاط و حفره‌ها، تعریف می‌شوند. برای درک بهتر این گزاره ها ابتدا ویژگیهای بکار گرفته شده معرفی می‌شوند و سپس چگونگی تعریف گزاره‌های شرط عضویت بیان می‌شوند.

محل و نوع نقاط وحفره ها

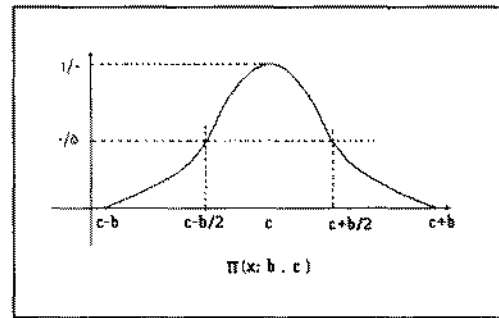
با استفاده از یک الگوریتم مناسب محل و نوع نقاط و حفره‌ها شناسایی می‌شوند [۲۴ و ۲۶]. محل نقاط می‌تواند بالا (U) یا پایین (D) باشد. نوع نقاط می‌تواند یک (1)، دو (2) یا سه (3) نقطه باشد. برای مثال حروف 'ب، پ، ت، ث' از لحاظ ویژگی نقاط بصورت گزاره  $PU3 | PU2 | PD3 | PD1$  توصیف می‌شوند. معنی این گزاره این است که این حروف یک یا سه نقطه در پایین یا دو یا سه نقطه در بالا دارند. حالت بدون نقطه بودن بدنه را نیز با (NP) نشان می‌دهیم. از لحاظ ویژگی حفره، بدنه حرف می‌تواند با یک حفره (H1) یا با دو حفره (H2) یا بدون حفره (NH) باشد.

تعداد دندانها ها و وجود سرکش

تعداد دندانها ها و وجود سرکش از ویژگیهای ساختاری هستند که در مرحله جداسازی تعیین می‌شوند [۲۴ و ۲۶]. برای مثال بدنه حرف 'ب' از لحاظ دندانها بصورت D1 و بدنه 'ی' به صورت ND توصیف می‌شود، که نشانگر این است که

محاسبات در این روش در حد یک هموارسازی یک بعدی است [۲۵]. در این مقاله برای فازی کردن تبدیل هاف راهکار دوم در نظر گرفته شده است. برای این منظور تعلق یک نقطه مانند  $(\rho_j, \theta_j)$  به سلول  $(\rho_k, \theta_k)$  را بصورت فازی با یک تابع عضویت  $\Pi$  در نظر می‌گیریم. از این تابع در کاربردهای دیگر شناسایی الگو نیز استفاده شده است [۱۳۸]. نمودار کلی این تابع در شکل ۷ آمده است. همانطور که در این شکل نشان داده شده، در نقطه c مقدار تابع ۱ و در نقاط  $c+b/2$  و  $c-b/2$  مقدار این تابع ۰/۵ است. با فرض اینکه محور p به فاصله های مساوی  $\Delta\rho$  تقسیم شده باشد، یک نمونه تابع عضویت در سلول  $(\rho_k, \theta_j)$  که شرایط تابع  $\Pi$  را داشته باشد بصورت زیر تعریف می‌شود.

$$\mu_k(\rho\rho) = \pi(\rho; \Delta\rho, \rho_k) = 1 - \left| \frac{\rho - \rho_k}{\Delta\rho} \right| \quad (5)$$

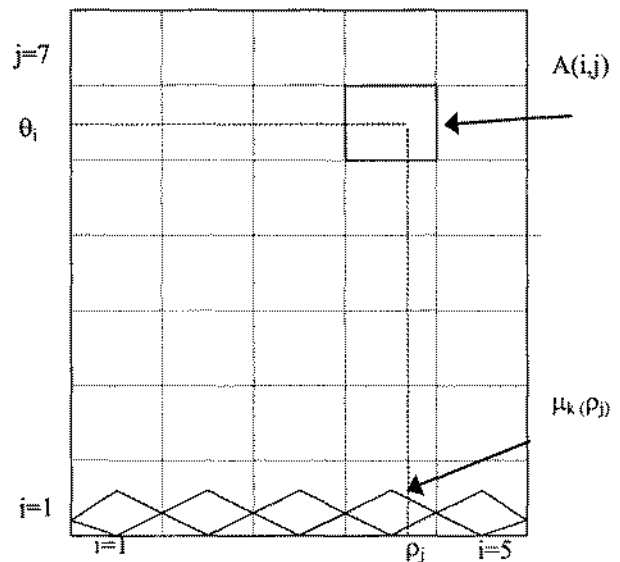


شکل ۷- تابع عضویت  $\Pi$

با تعریف تابع عضویت  $\mu_k(\rho_j)$  در معادله (۶)، اثرگذاری یک نقطه از کانتور ورودی بر مقدار شمارنده  $A(i,j)$  و دو همسایه اش در سطر j با رابطه (۶) نشان داده می‌شود. در این رابطه  $\mu_k(\rho_j)$  مقدار عضویت  $\rho_j$  در k امین سلول سطر j است (شکل ۸).

$$A(i, j) = A(i, j) + \mu_k(\rho_j) \quad (6)$$

$$k = i - 1, i, i + 1$$



شکل ۸- نمایشی از تبدیل هاف فازی

۳-۴ طبقه بندی

در الگوریتم ۲ از یک شبکه عصبی پرسپترون جلو سو با روش آموزش پس انتشار خطا استفاده شده است [۳۹].

**ارتفاع، پهنا و نسبت ارتفاع به پهناي حروف**

این ویژگیها که به ترتیب آنها را با  $W, H$  و  $AR$  نشان می‌دهیم، ویژگیهای ساده‌ای هستند که در بازشناسی برخی از حروف کمک مؤثری می‌کنند. ویژگیهای ارتفاع و پهنا معمولاً نسبت به پهناي قلم ( $pt$ ) نرمالیزه می‌شوند. برای مثال حرف  $T$  با استفاده از این ویژگی بصورت  $AR > 0.7 \ \& \ H/pt > 4$  توصیف می‌شود. کلیه ویژگیهای بکار گرفته شده در گزاره‌های شرط عضویت در جدول ۱ آورده شده‌اند.

جدول ۱- ویژگیهای بکاربرده شده در گزاره های شرط عضویت

نوع ویژگی	تعداد ویژگی	نماد ویژگی
محل و نوع نقاط	۷	$NP, PD1, PD2, PD3, PUI, PU2, PU3$
تعداد حرفه‌ها	۳	$NH, HI, H2$
تعداد دندانها و وجود سرکش	۶	$NG, G1, G2, ND, D1, D2$
ویژگیهای مکان مشخصه	۸۱	$Loc[0], Loc[1], \dots, Loc[80]$
هیستوگرام تعداد برخوردها در راستاهای افقی و عمودی	۶	$NC1, NC2, NC3, NR1, NR2, NR3$
ارتفاع، پهنا و نسبت ارتفاع به پهناي حروف	۳	$W, H, AR$

**بررسی برخی از گزاره های شرط عضویت**

بررسی خطا های موجود در مرحله بازشناسی بدنه حروف با الگوریتم‌های ۱ و ۲ نشان می‌دهد که اکثر خطاها، بین کلاسهای مشخصی مثل کلاسهای  $\bar{b}$ ,  $\bar{f}$ ,  $\bar{d}$ ,  $\bar{r}$ ,  $\bar{c}$ ,  $\bar{e}$  و  $\bar{s}$ , رخ می‌دهند. برای کاهش خطای بین این کلاسها، از گزاره‌های مناسبی برای توصیف آنها استفاده شده است. جدول ۲ برخی از این گزاره های شرط عضویت و کلاسهای مربوط به آنها را نشان می‌دهد.

جدول ۲- برخی از گزاره های شرط عضویت و کلاسهای مربوط به آنها

کلاسهای مربوطه	گزاره های شرط عضویت
$\bar{f}$	$H1 \ \& \ PUI$
$\bar{b}, \bar{p}, \bar{t}, \bar{d}$	$NH \ \& \ (PU2 \   \ PU3 \   \ PD1 \   \ PD3)$
$\bar{d}$	$NC2 > 0.5 \ \& \ NR2 > 0.8$
$\bar{r}$	$NC2 < 0.5$
$\bar{w}$	$NC3 > 0.06 \ \& \ NR2 > 0.2 \ \& \ H1$
$\bar{e}$	$Loc[66] > 0.07 \ \& \ Loc[58] < 0.02$
$\bar{c}$	$Loc[66] < 0.05 \ \& \ Loc[58] > 0.03$

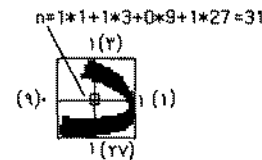
**۶- بازشناسی نهایی حروف**

برای بازشناسی حرف ورودی، ابتدا نقاط و علائم آن جدا شده و بدنه حرف استخراج می‌شود. آنگاه مطابق آنچه در بخش ۲ مقاله ارائه شد، کانتور مناسب آن تعیین می‌شود. در الگوریتم ۱ این کانتور بصورت کدهای فریمین به طبقه‌بندی کننده‌های کلاسهای مختلف که مجموعه‌ای از اتوماتون‌های آماری هستند ارائه می‌شود. هر کدام از طبقه‌بندی کننده‌ها مقدار تابع ممیز را برای کلاس مربوط به خود مشخص می‌کنند. در الگوریتم ۲ با استفاده از مختصات نقاط روی کانتور و تبدیل هاف فازی، ۲۵ ویژگی استخراج می‌شود و به ورودی شبکه عصبی اعمال می‌شود. هر کدام از گره‌های خروجی این شبکه عصبی مقدار تابع ممیز را برای کلاس مربوط به خود تعیین می‌کنند. در مرحله نهایی هر دو الگوریتم، بدنه‌های پیشنهادی به ترتیب مقدار تابع ممیز آنها با توجه به گزاره های شرط عضویت مربوط به هر کلاس بررسی می‌شوند. در نهایت کلاسی که بیشترین مقدار تابع ممیز را داشته و گزاره های شرط عضویت آن نیز برقرار باشند، با انتساب نقاط و

$\bar{b}$  یک دندان دارد و  $\bar{y}$  دندان ندارد. حرف  $\bar{g}$  از لحاظ وجود سرکش بصورت  $G1 \ | \ G2$  توصیف می‌شود، به این معنی که می‌تواند یک یا دو سرکش داشته باشد.

**ویژگیهای مکان مشخصه**

ویژگیهای مکان مشخصه معمولاً در راستاهای عمودی، افقی و یا زوایای ۴۵ و ۱۳۵ درجه تعیین می‌شوند. بردار ویژگیهای مکانهای مشخصه برای هر تصویر به اینصورت محاسبه می‌شود که به هر نقطه از زمینه تصویر بدنه حرف، یک عدد نسبت می‌دهیم که این عدد با توجه به اینکه خطوط عمودی و افقی رسم شده از آن نقطه در جهت های چهار گانه بالا، پایین، راست و چپ بدنه حرف را در چند نقطه قطع می‌کنند، محاسبه می‌شود. اگر تعداد قطع بدنه را به ۲ محدود کنیم، می‌توانیم از یک عدد چهار رقمی در مبنای ۳ برای نمایش نقاط مختلف نواحی زمینه حرف ورودی استفاده کنیم (شکل ۱۰). بردار مکانهای مشخصه در این حالت دارای ۸۱ عنصر است که هر کدام فراوانی نسبی عدد مربوط به خود را در تصویر ورودی نشان می‌دهند [۲۴]. در صورت لزوم می‌توانیم از هریک از عناصر بردار ویژگیهای مکان مشخصه برای توصیف و تمایز یک بدنه استفاده کنیم. برای مثال حرف  $\bar{d}$  را در شکل ۱۰ از لحاظ این ویژگی می‌توان بصورت  $Loc[31] > 0.2$  توصیف کرد.



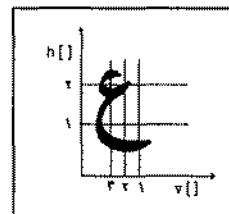
شکل ۱۰- نحوه محاسبه مکانهای مشخصه

**هیستوگرامهای تعداد برخوردها در راستاهای افقی و عمودی**

منظور از این دو هیستوگرام محاسبه تعداد دفعاتی است که یک خط عمودی یا افقی بدنه حرف ورودی را قطع می‌کند. نحوه محاسبه این هیستوگرامها در شکل ۱۱ نشان داده شده است. هر کدام از عناصر بردارهای هیستوگرام عمودی،  $V[j]$  و افقی،  $H[j]$  نشان دهنده این است که خط رسم شده در ستون یا ردیف متناظر آن، بدنه حرف را در چند نقطه قطع می‌کند. ویژگیهایی که از این دو هیستوگرام برای توصیف بدنه حروف بکار برده می‌شوند، چهار ویژگی از هیستوگرام عمودی ( $NC1, NC2, NC3, NC4$ ) و چهار ویژگی از هیستوگرام افقی ( $NR1, NR2, NR3, NR4$ ) هستند، که طبق معادلات (۷) محاسبه می‌شوند. هر کدام از این ویژگیها نشان دهنده این است که در تمام یا بخشی از بدنه حرف، فراوانی خطوط افقی یا عمودی که بدنه حرف را در ۱، ۲، ۳ یا ۴ نقطه قطع کرده‌اند چقدر است. برای مثال حرف  $\bar{d}$  از لحاظ این ویژگیها بصورت  $NC2 > 0.5 \ \& \ NR1 > 0.8$  توصیف می‌شود.

$$NC_i = \frac{\sum_{j, v[j]=i} v[j]}{\sum_j v[j]} \tag{7}$$

$$NR_i = \frac{\sum_{j, h[j]=i} h[j]}{\sum_j h[j]}$$



شکل ۱۱- نمونه ای از محاسبه هیستوگرام های تعداد برخوردها در راستاهای

عمودی و افقی

طبقه‌بندی استفاده شده است. بازشناسی نهایی حروف با توجه به تعداد نقاط آنها و برخی ویژگی‌های ساختاری انجام شده است. این الگوریتم‌ها روی مجموعه‌ای از حروف شامل ۱۱۵۰۰ نمونه از ده قلم مختلف آماده شدند و به ترتیب میزان شناسایی ۹۷/۱۳ و ۹۸/۲۲ درصد حاصل شد.

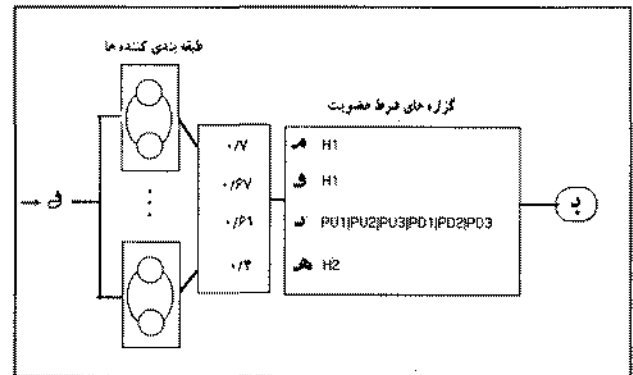
جدول ۳- نتایج آزمایش الگوریتم ۱

نام گروه	تعداد کلاسها	تعداد کلاسهای بدنه	میزان شناسایی بدنه (حروف (%)) (قبل از بازشناسی نهایی)	میزان شناسایی نهایی
ابتدایی	۲۵	۱۲	۹۳/۵	۹۷/۴۶
میانی	۲۵	۱۲	۹۲	۹۶/۳
انتهاایی	۴۴	۱۹	۸۷/۳	۹۶/۸
مجزا	۳۴	۱۹	۸۶/۹۵	۹۷/۷۹
نتایج کل	۱۱۵	۶۲	۸۹/۳	۹۷/۱۳

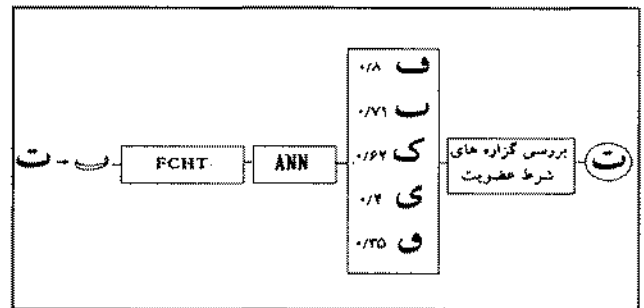
جدول ۴- نتایج آزمایش الگوریتم ۲

گروه حروف	مشخصات شبکه	تعداد کلاسهای بدنه	میزان شناسایی بدنه (حروف (%)) (قبل از بازشناسی نهایی)	میزان شناسایی نهایی
ابتدایی	۱۲ × ۱۹ × ۳۵	۱۲	۹۳/۵	۹۹/۱
میانی	۱۲ × ۱۹ × ۳۵	۱۲	۹۲	۹۸/۳
انتهاایی	۱۹ × ۲۵ × ۳۵	۱۹	۸۸/۵	۹۷/۴۵
مجزا	۱۹ × ۲۵ × ۳۵	۱۹	۸۹/۷۸	۹۸/۷۳
نتایج کل		۶۲	۹۰/۵۳	۹۸/۳۲

علائم به آن به عنوان کلاس حرف ورودی تعیین می‌شود. شکل های ۱۲ و ۱۳ به ترتیب نمونه ای از بازشناسی حروف با الگوریتم های ۱ و ۲ را نشان می‌دهند.



شکل ۱۲- نمونه ای از شناسایی یک حرف با الگوریتم ۱



شکل ۱۳- نمونه ای از شناسایی یک حرف با الگوریتم ۲

### مراجع

- [1] S. Diehl and K. Eglowstein, "Tame The Paper Tiger," *BYTE*, Vol. 27, no 5, pp 220-238, 1991.
- [2] S. Mori, C. Y. Suen and K. Yamamoto, "Historical Review of OCR Research and Development," *Proc. of IEEE*, Vol. 80, No 7, 1992.
- [3] O. V. Trier, A. K. Jain and T. Taxt, "Feature Extraction Methods for Character Recognition- a Survey," *Pattern Recognition*, Vol. 29, No. 4, pp. 641-662, 1996.
- [4] K. Badie and M. Shimura, "Machine Recognition of Arabic Cursive Script," *Pattern Recognition in Practice*, pp. 315-323, 1980.
- [5] B. Parhami and M. Taraghi, "Automatic Recognition of Printed Farsi Text," *Pattern Recognition*, Vol. 14, No.1-6, pp. 395-403, 1981.
- [6] H. Al-Muallim and S. Yamaguchi, "A Method of Recognition of Arabic Cursive Handwriting," *IEEE Trans. PAMI*, Vol. 9, No.5, pp. 715-722, 1987.
- [7] T. S. El-Sheikh and R. M. Guindi, "Computer Recognition of Arabic Cursive Script," *Pattern Recognition*, Vol. 21, No.4, pp. 293-302, 1988.
- [8] A. Mahmoud, "Arabic Character Recognition using Fourier Descriptors and Character Contour Encoding," *Pattern Recognition*, Vol. 27, No.6, pp. 815-824, 1994.
- [9] I. S. I. Abuhaiba, S. A. Mahmoud and R. J. Green, "Recognition of Handwritten Cursive Arabic Characters," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 6, pp. 664-671, 1994.

همانطور که قبلاً ذکر شد، ۳۲ حرف الفبای فارسی با توجه به موقعیت آنها در کلمه (ابتدا، وسط، انتها و مجزا) اشکال مختلفی به خود می‌گیرند. مجموع کل این کلاسها ۱۱۵ است. چون هر دو الگوریتم ۱ و ۲ بر اساس شناسایی بدنه اصلی حروف کار می‌کنند، تعداد کلاسها در هر گروه کاهش پیدا می‌کند. مجموعه آزمایش شامل ۱۱۵۰۰ حرف است. نمونه‌ها مربوط به ۱۱۵ کلاس مختلف هستند که برای هر حرف ۱۰ قلم مختلف و از هر قلم ۱۰ نمونه در نظر گرفته شده است. از این ۱۰ نمونه، ۵ نمونه برای آموزش و ۵ نمونه برای آزمایش سیستم استفاده شده است. جدولهای ۳ و ۴ به ترتیب خلاصه نتایج بازشناسی را برای گروههای مختلف حروف نشان می‌دهند. همانطور که دیده می‌شود بکارگیری گزاره‌های شرط عضویت میزان بازشناسی را در حدود ۸ درصد افزایش داده است. میزان شناسایی برای الگوریتم ۲ حدود ۱/۲ درصد بیشتر از الگوریتم ۱ است. الگوریتم ۱ بعلت اینکه ویژگیهای ورودی خود را از تبدیل هاف فازی روی کانتور حروف می‌گیرد، نسبت به الگوریتم ۲ حساسیت کمتری به نویز دارد. زمان آموزش الگوریتم ۲ خیلی بیشتر از الگوریتم ۱ است، ولی سرعت شناسایی برای هر دو الگوریتم تقریباً یکسان است. اگر از یک کامپیوتر شخصی با ریزپردازنده ۴۸۶ با فرکانس ۱۰۰ مگاهرتز استفاده شود، سرعت بازشناسی حدود ۱۰۰ حرف در ثانیه است. در محاسبه این سرعت مراحل استخراج ویژگی، بازشناسی بدنه و بازشناسی نهایی در نظر گرفته شده‌اند.

### ۷- نتیجه گیری

در این مقاله دو الگوریتم مختلف برای شناسایی حروف با قلم های مختلف ارائه شد. در الگوریتم ۱ از کد های زنجیره ای کانتور حروف بعنوان ویژگی و از اتوماتون‌های آماری برای طبقه بندی استفاده شده است. در الگوریتم ۲ از تبدیل هاف فازی برای استخراج ویژگی و از شبکه عصبی سه لایه پرسپترون برای

- [۲۶] ر. عزمی و ا. کبیر، 'معرفی روش جدیدی برای جداسازی حروف درمتون چایی فارسی بدون توجه به نوع قلم'، استقلال، نشریه مهندسی دانشگاه صنعتی اصفهان، سال ۱۸، شماره ۲، ص ۱-۱۰، ۱۳۷۸.
- [27] R. Azmi and E. Kabir, "A New Segmentation Technique for Omnifont Farsi Text" *Pattern Recognition Letters*, Vol. 22, No. 2, pp. 97-104, Jan. 2001.
- [28] K. S. Fu, "Syntactic Methods in Pattern Recognition," Academic Press, 1974.
- [29] M. Al-Tuwaijir and M. Bayoumi, "Arabic Text Recognition Using Neural Network," *Proc. Int. Symp. on Circuits and Systems (ISCAS'94)*, pp. 415-418, 1994.
- [۳۰] ک. فائز، ع. ختن زاد، و م. ح. شیرعلی شهرضا، 'تشخیص حروف و ارقام دستنویس فارسی با استفاده از گشتاورهای شبه زرنیکه و بکمک شبکه های عصبی'، مجموعه مقالات سومین کنفرانس مهندسی برق ایران، ص ۲۳۱-۲۴۰، دانشگاه علم و صنعت ایران، ۱۳۷۴.
- [31] A. Amin, and W. Mansoor, "Recognition of Arabic Printed Text Using Neural Networks," *Proc. 4th Int. Conf. on Document Analysis and Recognition*, Ulm, 1997.
- [32] J. H. Han, L. T. Koczy and T. Poston, "Fuzzy Hough Transform", *2nd IEEE International Conference on Fuzzy Systems*, Vol. 2, p. 803-808, 1993.
- [33] J.H. Han, L.T. Koczy and T. Poston, "Fuzzy Hough Transform," *Pattern Recognition Letters*, Vol. 15, No. 7, pp. 649-658, July 1994.
- [34] K. P. Philip, E. L. Dove, D.D. McPherson, N.L. Gotteiner, W. Stanford and K.B. Chandran. "The Fuzzy Hough Transform- Feature Extraction in Medical Images," *IEEE Transactions on Medical Imaging*, Vol. 13, No. 2, pp. 235-40, 1994.
- [35] V. Chatzis and I. Pitas, "Fuzzy Cell Hough Transform for Curve Detection," *Pattern Recognition*, Vol. 30, No.12, pp. 2031-2042, 1997.
- [36] S. Sural, P. K. Das, "Fuzzy Hough Transform and an MLP with Fuzzy Input/Output for Character Recognition," *Fuzzy Sets and Systems*, Vol. 105, No. 3, pp. 489-497, 1999.
- [37] A. Amin, "Recognition of Hand-Printed Latin Characters Based on Generalized Hough Transform and Decision Tree Learning Techniques," *Int. J. Pattern Recog. and Artificial Intelligence*, Vol. 14, No. 3, pp. 369-387, 2000.
- [38] S. K. Pal, "Fuzzy Mathematical Approach to Pattern Recognition," *Wiley Eastern Limited*, 1986.
- [39] R. P. Lippmann, "An Introduction to Computing with Neural Nets," *IEEE ASSP Magazine*, pp. 4-22, 1987.
- [10] K. Masruri and E. Kabir, "Recognition of Hand-Printed Farsi Characters by a Fuzzy Classifier," *Proc. 2nd Asian Conf. Computer Vision, ACCV'95*, pp. II.607-II610, 1995.
- [11] R. Azmi and E. Kabir, "A Recognition Algorithm for Hand-Printed Farsi Characters," *Proc. Int. Conf. on Telecommunication, ICT96*, Vol. 2, pp. 852-855, 1996.
- [12] A. Amin, H. Al-Sadoun, and S. Fischer, "Hand-Printed Arabic Character Recognition System Using an Artificial Network," *Pattern Recognition*, Vol. 29, No. 4, pp. 663-675, 1996.
- [13] A. Amin, "Off-Line Arabic Character Recognition: the State of the Art", *Pattern Recognition*, Vol. 31, No. 5, pp. 517-530, 1998.
- [14] F. Bouslama, "Neural Networks in the Recognition of Machine Printed Arabic Characters," *Int. J. Pattern Recognition and Artificial Intelligence*, Vol. 13, No. 3, pp. 395-414, 1999.
- [۱۵] م. احمدزاده و ا. کبیر، 'تشکستن کلمات تایپ شده فارسی به حروف'، مجموعه مقالات اولین کنفرانس بین المللی کاربرد کامپیوتر در علوم، فنون و پزشکی، ص ۱-۶، ۱۳۷۰.
- [۱۶] ا. کبیر، ک. بهاری و م. احمدزاده، 'بازشناسی متون تایپ شده فارسی'، مجموعه مقالات اولین کنفرانس مهندسی برق ایران، جلد دوم، ص ۲۸۵-۲۹۴، ۱۳۷۲.
- [۱۷] ح. فهیمی و ب. تیمساری، 'بازشناسی حروف در کلمات تایپ شده فارسی با استفاده از روش مورفولوژی'، مجموعه مقالات اولین کنفرانس مهندسی برق ایران، جلد ۲، ص ۲۷۷-۲۸۴، ۱۳۷۲.
- [۱۸] م. فهیمی و ر. ثانی، 'تشخیص حروف دستنویس فارسی'، مجموعه مقالات اولین کنفرانس انجمن کامپیوتر ایران، ص ۱۰۳-۱۱۰، ۱۳۷۴.
- [۱۹] م. فتحی و ع. برومند نیا، 'شناسایی ارقام و حروف مجزای فارسی در شرایط نور غیر یکنواخت'، مجموعه مقالات اولین کنفرانس انجمن کامپیوتر ایران، ص ۹۶-۱۰۲، ۱۳۷۴.
- [۲۰] م. ح. شیرعلی شهرضا، 'تشخیص کلمات و ارقام دستنویس فارسی بوسیله شبکه های عصبی'، پایان نامه دکتری، دانشکده برق، دانشگاه صنعتی امیرکبیر، ۱۳۷۴.
- [۲۱] ح. میرزا خانی و ک. فائز، 'روش نوین در شناسایی متون فارسی به کمک شبکه های عصبی'، مجموعه مقالات سومین کنفرانس الکترونیک، ص ۱۴۹-۱۵۴، ۱۳۷۴.
- [۲۲] ر. عزمی و ا. کبیر، 'ارائه دو الگوریتم برای شناسایی حروف چایی فارسی'، مجموعه مقالات سومین کنفرانس انجمن کامپیوتر ایران، ص ۱۹۱-۱۹۷، ۱۳۷۴.
- [۲۳] م. رحمتی و ن. زارعی، 'شناسایی حروف تاییبی فارسی با استفاده از تبدیل خود همبستگی مورفولوژی'، مجموعه مقالات سومین کنفرانس انجمن کامپیوتر ایران، ص ۱۸۰-۱۸۵، ۱۳۷۴.
- [۲۴] ر. عزمی، 'بازشناسی متون چایی فارسی'، رساله دکتری، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، ۱۳۷۸.
- [۲۵] ر. عزمی، و ا. کبیر، 'جداسازی حروف درمتون چایی فارسی'، مجموعه مقالات پنجمین کنفرانس برق، دانشگاه صنعتی شریف، صفحات ۳۰۲-۵ تا ۳۰۹، ۱۳۷۴.

<sup>1</sup> Pseudo Zernike moment<sup>2</sup> Connected components<sup>3</sup> Fuzzy Cell Hough Transform



علوم و مهندسی کامپیوتر، نشریه علمی پژوهشی انجمن کامپیوتر ایران، مجلد ۱، شماره ۱ (الف)، بهار ۱۳۸۲

رضا عزمی کارشناسی و کارشناسی ارشد خود را در مهندسی الکترونیک به ترتیب از دانشگاه صنعتی امیر کبیر و دانشگاه تربیت مدرس در سالهای ۱۳۶۹ و ۱۳۷۲ دریافت کرد.



او سپس دوره دکترای خود را در بخش مهندسی برق دانشگاه تربیت مدرس آغاز کرد و در سال ۱۳۷۸ از رساله خود با عنوان "باز شناسی متون چاپی فارسی" دفاع کرد. او اکنون استادیار دانشکده فنی و مهندسی دانشگاه الزهراء است. زمینه های پژوهشی مورد علاقه او بازشناسی الگو و پردازش تصویر است.

احسان ا... کبیر کارشناسی ارشد پیوسته خود را در مهندسی برق و الکترونیک از دانشکده فنی دانشگاه تهران و دکترای خود را در مهندسی سیستمهای الکترونیک از دانشگاه اسکس در انگلستان، به ترتیب در سالهای ۱۳۶۴ و ۱۳۶۹ دریافت کرد.



او اکنون دانشیار بخش مهندسی برق، دانشگاه تربیت مدرس است. زمینه های پژوهشی مورد علاقه او بازشناسی الگو به ویژه بازشناسی متون چاپی و دستنویس و بینایی ماشین است.

کامبیز بدیع فارغ التحصیل دبیرستان البرز تهران بوده و درجات کارشناسی، کارشناسی ارشد و دکترای خود را در زمینه مهندسی الکترونیک و برق از انستیتو تکنولوژی توکیو دریافت نموده است. اهم فعالیت های تحقیقاتی نامبرده در حال حاضر در زمینه یادگیری ماشین و مدلسازی شناختی به طور اعم و مدلسازی فیزیکی، مدلسازی تفسیر، یادگیری از تجربه و زایش خلاق ایده ها به طور اخص می باشد. ایشان در حال حاضر عضو هیأت علمی و مدیر گروه تخصصی جامعه اطلاعاتی در مرکز تحقیقات مخابرات ایران و عضو هیأت مدیره انجمن کامپیوتر ایران میباشند.

