



## کاربرد تطابق شکل در بازشناسی ارقام دستنویس فارسی

علیرضا درویش<sup>۱</sup>، احسان اله کبیر<sup>۲\*</sup>، حسین خسروی<sup>۳</sup>

۱- دانشجوی کارشناسی ارشد رشته مهندسی برق- الکترونیک، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس

۲- دانشیار بخش مهندسی برق، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس

۳- دانشجوی کارشناسی ارشد رشته مهندسی برق- الکترونیک، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس

\* تهران، صندوق پستی ۱۴۳-۱۴۱۱۵، [kabir@modares.ac.ir](mailto:kabir@modares.ac.ir)

[Downloaded from FarsiOCR.ir](http://Downloaded from FarsiOCR.ir)

**چکیده-** در این مقاله، از یک الگوریتم تطابق شکل برای بازشناسی ارقام دستنویس فارسی استفاده می شود. برای هر نقطه نمونه برداری شده روی کانتور شکل، توصیفگری براساس توزیع مکانی نقاط دیگر کانتور بدست می آید. برای تعیین میزان شباهت دو شکل، ابتدا بر اساس این توصیفگرها تناظری یک به یک بین نقاط نمونه برداری شده روی کانتور شکل اول با نقاط روی کانتور شکل دوم بدست می آید. جمع فواصل بین نقاط متناظر دو شکل معیاری برای عدم شباهت آنهاست. سپس تبدیلی تعریف می شود که نقاط کانتور شکل اول را بر روی نقاط متناظرشان در کانتور شکل دوم قرار دهد. میزان پیچیدگی این تبدیل معیاری دیگر برای عدم شباهت دو شکل است.

با بهینه کردن پارامترهای مختلف این الگوریتم، از آن برای بازشناسی ارقام دستنویس فارسی استفاده کرده ایم. از روش طبقه بندی نزدیکترین فاصله از نماینده یا نماینده های هر کلاس استفاده شده است. در یک آزمایش بر روی مجموعه ای شامل ۲۸۸ رقم که افراد مختلف نوشته اند، ۸۹/۹٪ بازشناسی درست حاصل شده است. این آزمایش بدون هیچگونه پس پردازشی انجام شده است.

**کلید واژگان:** تطابق شکل، کانتور، توصیفگر محلی، تبدیل هندسی، بازشناسی ارقام دستنویس.

### ۱- مقدمه

بازشناسی حروف و ارقام دستنویس همواره از موضوعات مورد علاقه برای تحقیق بوده است [۱، ۲، ۳، ۴]. در زمینه بازشناسی حروف و ارقام دستنویس عربی و فارسی نیز، کارهای زیادی صورت گرفته است [۵]. در یک تحقیق، از ویژگیهای گشتاوری و طبقه بند بیز برای بازشناسی حروف دستنویس فارسی استفاده شده است [۶]. در تحقیق دیگری از ویژگیهای مکانهای مشخصه و طبقه بند بیز و

زنجره مارکف برای بازشناسی ارقام استفاده شده است [۷]. از روش 3-tuple نیز برای استخراج ویژگی از ارقام دستنویس استفاده شده است [۸]. مسروری از DTW برای بازشناسی ارقام استفاده کرده است [۹]. از منطق فازی نیز برای تقسیم حروف به چند دسته مختلف استفاده شده است [۱۰]. در تحقیق دیگری بازشناسی ارقام دستنویس فارسی به کمک روشهای فازی انجام شده است [۱۱]. کاری که ما انجام داده ایم مبتنی بر توصیف شکل ارقام دستنویس است،

در ادامه در بخشهای دوم و سوم به معرفی الگوریتم تطابق شکل می پردازیم و در بخش چهارم چگونگی استفاده از این الگوریتم را برای بازشناسی ارقام دستنویس فارسی توضیح می دهیم. در بخش پنجم به مقایسه این روش با روشهای دیگر می پردازیم. بخش ششم به نتیجه گیری و ارائه پیشنهاد هایی برای بهبود بازشناسی اختصاص دارد.

## ۲-توصیف شکل ارقام

برای توصیف شکل یک رقم، ابتدا کانتور آن بدست می آید و سپس تعدادی نقطه در روی کانتور نمونه برداری می شوند. برای هر نقطه، براساس فاصله نسبی آن با سایر نقاط روی کانتور، توصیفی برای آن نقطه تعریف می شود. برای مقایسه دو رقم، ابتدا بر روی کانتور هر یک از آنها نقاط لازم نمونه برداری می شوند. سپس براساس معیار فاصله ای که تعریف می شود، تناظری یک به یک بین نقاط نمونه برداری شده دو رقم بدست می آید. در نتیجه فاصله بین دو رقم از هم به صورت جمع فاصله بین تک تک نقاط متناظر آنها قابل محاسبه خواهد بود. کانتور شکل ارقام به کمک لبه یاب کنی<sup>۱</sup> مشخص می شود. تعداد نقاط نمونه برداری شده روی کانتور بسته به دقت و کاربرد متفاوت است. البته مشخص است که هر چه تعداد این نقاط بیشتر باشد، توصیف دقیقتری از شکل بدست می آید. اما این نکته را هم باید در نظر داشت که با افزایش تعداد نقاط، حجم کار برنامه و در نتیجه زمان اجرای آن افزایش می یابد. بنابراین در این مورد باید انتخابی مناسب داشته باشیم.

به این صورت است که از یک الگوریتم تطابق الگو [۱۲] برای توصیف و مقایسه دو رقم دستنویس استفاده کرده ایم و سعی کرده ایم با بهینه کردن این الگوریتم، آن را برای بازشناسی ارقام دستنویس فارسی اصلاح کنیم. در این روش ابتدا باید توصیفی از شکل مورد نظر داشته باشیم و در مرحله بعد معیاری برای مقایسه توصیفگرها تعریف کنیم.

روشی که ما از آن برای توصیف شکل استفاده می کنیم، بر اساس کانتور شکل است. در سالهای اخیر کارهای زیادی در زمینه توصیف شکل بر اساس کانتور انجام شده است. چنگ و یان [۱۳] ویژگیهایی را از کانتور ارقام استخراج کرده اند. این ویژگیها شامل توصیفگر فوریه کانتور، میانگین و واریانس فاصله پیکسلهای روی کانتور تا مرکز ثقل آن و همچنین طول و مساحت نرمالیزه شده کانتور است. ارقام بر حسب آنکه چند کانتور دارند به چند دسته تقسیم می شوند و بر اساس این ویژگیها شناسایی می شوند. همچنین در روش دیگری، چاکراواری و کمپلا بر اساس نقاط بحرانی توصیفی ریاضی از کانتور حروف و ارقام ارائه دادند [۱۴]. یک نقطه بحرانی، نقطه ایست که مشتق اول تابع کانتور در آن صفر است. در این روش کانتور هر شکل بر اساس نقاط بحرانش بوسیله گرافی نمایش داده می شود. روش ما که بر اساس کانتور شکل است شامل سه مرحله است [۱۲]:

۱. حل مسأله تناظر بین نقاط کانتور شکل اول با نقاط کانتور شکل دوم.
۲. استفاده از این تناظر برای تعریف تبدیل هندسی که دو شکل را روی هم قرار دهد.
۳. محاسبه فاصله بین دو شکل به صورت مجموع خطاهای تطابق نقاط متناظر دو شکل و جمله ای که میزان پیچیدگی این تبدیل هندسی را نشان می دهد.

---

1. Canny Edge Detector

## ۲-۱- توصیفگر محلی برای نقاط شکل

اگر مجموعه بردارهایی را که از یک نقطه روی کانتور به تمام نقاط دیگر روی آن کانتور رسم می‌شود در نظر بگیریم، بردارهایی داریم که توصیف تمام شکل را نسبت به نقطه مرجع نشان می‌دهند. براساس این بردارها، می‌توان توصیفی برای تک تک نقاط نمونه برداری شده بدست آورد [۱۲].

برای تعریف این توصیفگر شکل ۱ را در نظر بگیرید. برای هر نقطه  $p_i$  بر روی شکل اول، می‌توان هیستوگرام  $h_i$  را براساس فاصله و زاویه نسبی پاره خط واصل آن نقطه با  $n$ - نقطه دیگر روی کانتور تعریف کرد.

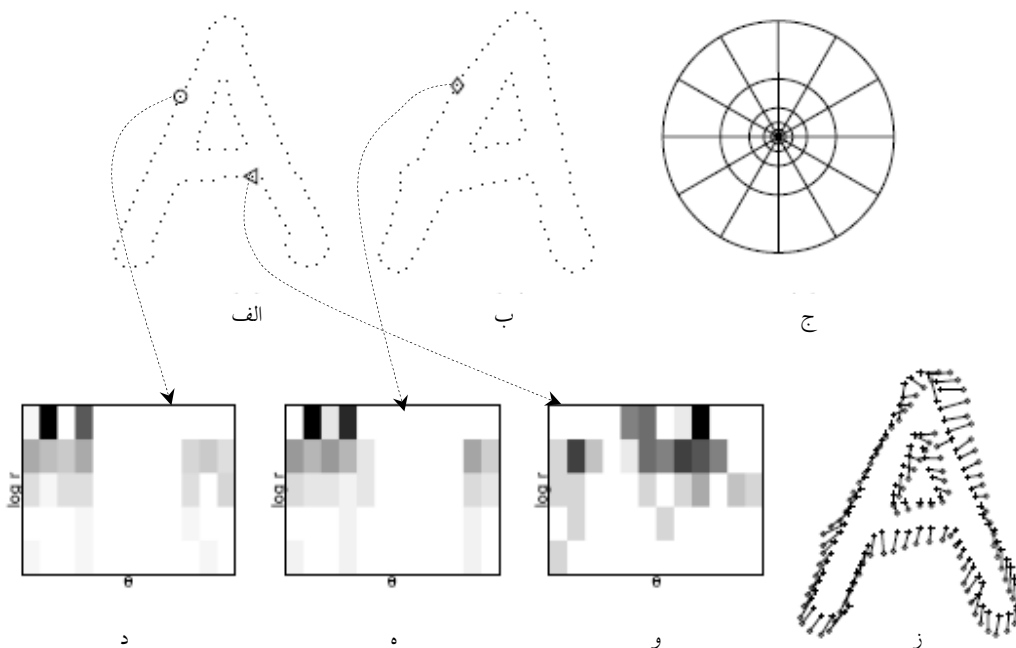
$$h_i(k) = \#\{q \neq p_i : (q - p_i) \in \text{bin}(k)\} \quad (1)$$

این هیستوگرام به عنوان توصیفگر اطلاعات جانبی نقطه  $p_i$  تعریف می‌شود. همانگونه که در شکل ۱ مشخص است، برای محاسبه این هیستوگرام از مختصات لگاریتمی قطبی استفاده می‌شود زیرا هدف ما این است که توصیفگر محاسبه شده در رابطه (۱) نسبت به نقاطی که در نزدیک آن قرار دارند حساستر باشد.

در این مختصات، عناصر از نظر زاویه‌ای به ۱۲ قسمت مساوی ۳۰ درجه تقسیم شده‌اند و از نظر شعاعی هم درجه بندی شعاع عناصر به صورت لگاریتمی است و از ۰/۱۲۵ تا ۲ تقسیم شده است. به کمک توصیفگرهای هر نقطه، تناظری یک به یک بین تمام نقاط روی کانتور شکل اول با تمام نقاط روی کانتور شکل دوم بدست می‌آوریم. برای محاسبه این تناظر باید از یک معیار فاصله بین این توصیفگرها استفاده کنیم، نقطه  $p_i$  را روی شکل اول و نقطه  $q_j$  را روی شکل دوم در نظر بگیریم. اگر  $(p_i, q_j)$  فاصله این دو نقطه باشد آنگاه خواهیم داشت:

$$c_{ij} = (p_i, q_j) = \frac{1}{2} \sum_{k=1}^K \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)} \quad (2)$$

که در این رابطه  $h_i(k)$  و  $h_j(k)$ ، معرف هیستوگرامهای نرمالیزه شده  $k$  عنصری مربوط به  $p_i$  و  $q_j$  هستند. پس از تعریف رابطه (۲) فاصله  $c_{ij}$  بین تمام نقاط روی شکل اول با تمام نقاط روی شکل دوم را به صورت ماتریس فاصله (۳) در اختیار خواهیم داشت:



شکل ۱. الف و ب) نقاط نمونه برداری شده از لبه دو شکل ج) دیاگرام مختصات لگاریتمی-قطبی (در این دیاگرام از ۵ عنصر شعاعی و ۱۲ عنصر زاویه ای استفاده شده است). د، ه، و) هیستوگرام بدست آمده برای سه نقطه نشان داده شده در شکل الف و ب. ز) نقاط متناظر دو شکل الف و ب [۱۲].

### ۳- تبدیل اسپلاین صفحه نازک<sup>۲</sup>

هنگامیکه مجموعه‌ای از نقاط متناظر را در روی کانتور دو رقم در اختیار داریم، می‌توانیم تبدیلی را طراحی بکنیم که بر اساس آن، این نقاط را روی هم قرار دهیم. با محاسبه میزان پیچیدگی این تبدیل، معیاری برای میزان اختلاف دو رقم بدست می‌آید. تبدیلی که ما در این قسمت استفاده کرده‌ایم، از خانواده تبدیلات اسپلاین صفحه نازک است که برای نمایش تبدیل در مختصات قابل انعطاف مناسب است [۱۲]. برای تعریف این تبدیل به صورت زیر عمل می‌کنیم.

فرض کنید بردار  $V$  شامل مقادیر خروجی تابعی است که بر مجموعه نقاط بردار  $P$  شامل نقاط  $P_i = (x_i, y_i)$  اعمال می‌شود. مجموعه نقاط بردار  $V$  را با  $(x'_i, y'_i)$  نشان می‌دهیم. تبدیل  $f(x, y)$  مجموعه نقاط  $(x_i, y_i)$  را به صورتی بر مجموعه نقاط  $(x'_i, y'_i)$  می‌نشانند که انرژی لازم برای انجام این تبدیل حداقل شود.

$$f(x, y) = a_1 + a_x x + a_y y + \sum_{i=1}^n w_i U(\|(x, y) - (x_i, y_i)\|) \quad (5)$$

که تابع کرنل  $u(r)$  به صورت زیر است.

$$U(r) = r^2 \log r^2 \quad (6)$$

برای اینکه تابع  $f(x, y)$  مشتقات مرتبه دوم پیوسته داشته باشد باید داشته باشیم

$$\sum_{i=1}^n w_i = 0 \quad \text{and} \quad \sum_{i=1}^n w_i x_i = \sum_{i=1}^n w_i y_i = 0$$

بر اساس تبدیل (۵) می‌توان یک سیستم خطی را به صورت زیر برای ضرایب TPS<sup>۱</sup> در نظر گرفت.

$$\begin{pmatrix} K & | & P \\ \hline P^T & | & 0 \end{pmatrix} \begin{pmatrix} w \\ a \end{pmatrix} = \begin{pmatrix} v \\ 0 \end{pmatrix} \quad (7)$$

که در آن ماتریسهای  $K$  و  $P$  به صورت زیر هستند:

$$Dis = \begin{bmatrix} C_{11} & \cdot & \cdot & C_{n1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ C_{n1} & \cdot & \cdot & C_{nn} \end{bmatrix} \quad (3)$$

حال هدف ما این است که برای نقطه  $p_i$  روی شکل اول نقطه  $q_{\pi(i)}$  را روی شکل دوم به گونه‌ای بدست آوریم که فاصله کل بین دو شکل حداقل شود.

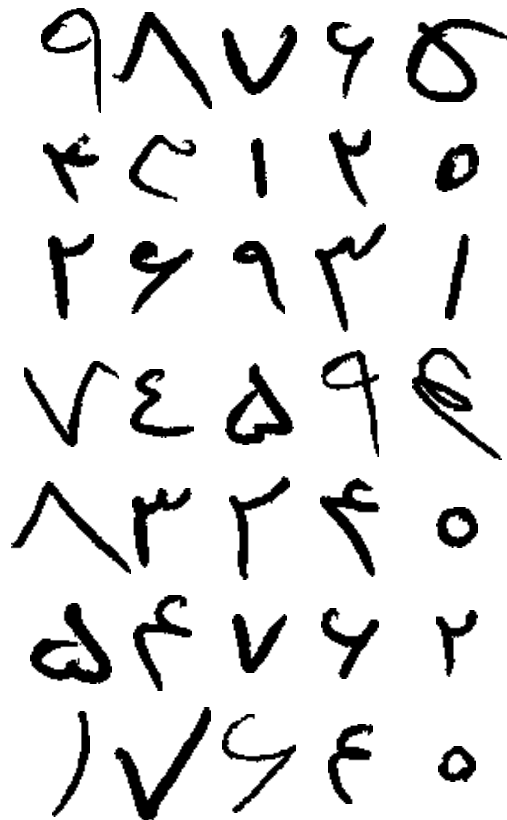
$$C_{context} = H(\pi) = \sum_i c(p_i, q_{\pi(i)}) \quad (4)$$

در واقع  $\pi$  تابعی یک به یک است که به هر نقطه  $i$  روی شکل اول، نقطه  $\pi(i)$  را در روی شکل دوم به عنوان نزدیکترین نقطه از نظر تشابه نسبت می‌دهد. این مساله به عنوان یک مساله ریاضی در نظر گرفته می‌شود که ورودی آن ماتریس Dis است و خروجی آن، نقاط  $\pi(i)$  روی شکل دوم برای تک تک نقاط  $i$  روی شکل اول است. برای حل این مساله می‌توان از روشهای مجاری<sup>۱</sup> [۱۵] و یا یونکر [۱۶] استفاده کرد، که ما از هر دو روش استفاده کردیم. نتایج مربوطه در بخش ۶ آورده شده‌اند. پس از اجرای این الگوریتم، تناظری یک به یک بین نقاط شکل اول و شکل دوم بدست می‌آید و رابطه (۴) مقدار حداقل خود را دارا می‌شود. حال از این رابطه می‌توان به عنوان معیاری برای تعیین فاصله دو شکل از هم استفاده کرد. یعنی هر چه این مقدار بیشتر باشد، شباهت دو شکل با هم کمتر است. همانگونه که اشاره شد، رابطه (۴) یکی از دو معیاری است که در تعیین معیار کلی فاصله بین دو شکل به کار می‌روند. معیار دیگر از محاسبه میزان پیچیدگی انجام یک تبدیل هندسی که این دو شکل را روی هم قرار می‌دهد قابل محاسبه است و همانطور که بعدا اشاره می‌شود، مجموع وزنی این دو فاصله را می‌توان به عنوان معیار فاصله کلی بین دو شکل در نظر گرفت [۱۲].

تبدیل لازم (رابطه ۱۵)، می توان معیاری کلی از ترکیب وزنی این دو بدست آورد که از آن در بازشناسی ارقام استفاده می شود.

#### ۴- بازشناسی ارقام دستنویس فارسی

مجموعه داده ای که ما بازشناسی ارقام را بر روی آن انجام می دهیم، مجموعه ای شامل ۱۲۸۸ رقم دستنویس فارسی است که توسط افراد مختلف نوشته شده است و سعی در آن شده است که گوناگونی ارقام فارسی، که در بین عموم شایع است، در آن حفظ شود. نمونه هایی از این ارقام در شکل ۲ مشاهده می شود.



شکل ۲. نمونه هایی از ارقام مجموعه داده

$$K_{ij} = U(\|(x_i, y_i) - (x_j, y_j)\|) \quad (8)$$

$$ith \text{ row of } P = (1, x_i, y_i)$$

و 0 هم یک ماتریس صفر ۳×۳ است.

$a, w$  بردارهای ضرایب در رابطه (۵) هستند.

$$a = [a_l, a_x, a_y]^T \quad (9)$$

$$w = [w_l, w_2, \dots, w_n]^T, v = [v_l, v_2, \dots, v_n]^T \quad (10)$$

حال همانطور که بوکستین [۱۷] نشان داده است اگر  $v_i$  ها را که نقاط متناظر روی کانتور شکل دوم هستند، به صورت زیر در نظر بگیریم:

$$v_1 = (x'_1, y'_1), v_2 = (x'_2, y'_2), \dots, v_n = (x'_n, y'_n) \quad (11)$$

و بردار  $Y$  را به صورت زیر تعریف کنیم.

$$Y = (V | 000)^T \quad (12)$$

که بردار  $Y$  یک بردار ستونی به طول  $n+3$  است. آنگاه مقادیر  $W = (w_l, \dots, w_n)^T$  و  $A = (a_l, a_x, a_y)^T$  از رابطه زیر بدست می آیند.

$$L^{-1}Y = (w_l, a_l, a_x, a_y)^T \quad (13)$$

که در آن برای  $L$  رابطه زیر را داریم.

$$L = \begin{bmatrix} K & P \\ P^T & 0 \end{bmatrix} \quad (14)$$

حال برای محاسبه انرژی لازم برای این تبدیل از رابطه ۱۵ که بوسیله بوکستین نشان داده شده است استفاده می کنیم [۱۷].

$$C_{ips} = I_F = wkw^T \quad (15)$$

رابطه (۱۵) معیاری برای پیچیدگی تبدیل بدست می دهد. هر چه این مقدار بیشتر باشد، فاصله دو شکل از هم بیشتر است. در نهایت به کمک دو معیار فاصله ای که بدست آوردیم، یعنی معیار فاصله براساس توصیفگرهای نقاط متناظر (رابطه ۴) و معیار فاصله براساس پیچیدگی

#### ۴-۱- بازشناسی به کمک انتخاب نماینده

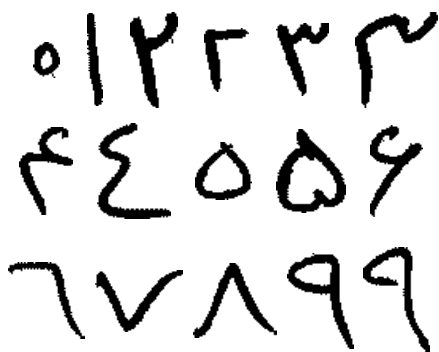
برای طبقه بندی ارقام دستنویس، ما از روش کمترین فاصله از نماینده‌های هر کلاس استفاده کرده ایم. در این روش بر حسب گوناگونی هر رقم، تعداد مناسبی نمونه را که بتوانند نماینده‌های خوبی برای توصیف آن رقم باشند، انتخاب می‌کنیم. هر ورودی از مجموعه آزمایش را با هر یک از این نمونه‌ها مقایسه می‌کنیم و آن را به کلاسی که کمترین فاصله را با آن دارد نسبت می‌دهیم. روابط (۴) و (۱۷) به عنوان معیارهای فاصله مورد استفاده قرار گرفته اند. بلونجی و همکارانش از روش سه همسایه نزدیکتر، 3-NN، به عنوان معیار فاصله استفاده کرده اند [۱۲]. این روش نسبت به روش ما نتایج بهتری دارد زیرا در این روش هر نمونه ورودی با تعداد زیادی از نمونه‌های ارقام مقایسه می‌شود، در نتیجه تصمیم دقیقتری گرفته می‌شود. از طرف دیگر در هر بار اجرای الگوریتم کلاسی انتخاب می‌شود که سه بار به عنوان انتخاب برتر، انتخاب شده است در نتیجه درجه اطمینان تصمیمی که گرفته می‌شود زیاد و خطا کم می‌شود. با این وجود، این روش زمان اجرای بسیار طولانی تری نسبت به روش کمترین فاصله از نماینده کلاس دارد. بنابراین چون تعداد کلاسهای ارقام مورد استفاده ما و همچنین گوناگونی ارقام در هر کلاس به گونه‌ای است که با انتخاب نماینده‌های مناسب به کمک روش کمترین فاصله از نماینده کلاس می‌توانیم به نتایج مطلوبی دست یابیم، از این روش استفاده کرده ایم.

#### ۴-۲- انتخاب اولیه نماینده‌ها و تنظیم پارامترها

انتخاب نماینده برای ارقام دستنویس فارسی بایستی با توجه به گوناگونی آنها انجام شود. به عنوان مثال رقم ۴ نسبت به رقم ۱ از گوناگونی بیشتری برخوردار است، در

نتیجه نیاز به نماینده‌های بیشتری دارد. بنابراین با انتخاب نماینده‌های مناسب برای هر رقم می‌توان به نرخ شناسایی بالاتری برای آن رقم دست یافت. مساله دیگری که وجود دارد این است که گرچه با افزایش تعداد نماینده‌ها برای هر رقم احتمال آنکه اکثر شکلهای ورودی مربوط به آن کلاس درست شناسایی شوند بیشتر می‌شود، اما از طرف دیگر احتمال آنکه این نماینده‌ها شکلهای ورودی مربوط به ارقام دیگر را نیز به خود جذب کنند بالا می‌رود و در نتیجه باعث کاهش نرخ شناسایی می‌شوند. بنابراین در این زمینه باید انتخابی بهینه داشته باشیم.

برای اینکه بتوانیم پارامترهای مختلف الگوریتم تطابق شکل را به طرز مناسبی انتخاب کنیم و نیز نماینده‌های مناسبی برای بازشناسی ارقام پیدا کنیم، ابتدا ۳۲۰ رقم از مجموعه ارقام را به طور تصادفی انتخاب کردیم. هدف ما از انتخاب این نمونه‌ها این بود که ابتدا نماینده‌هایی برای هر رقم انتخاب کنیم و در مرحله بعد با انتخاب ترکیبات مختلفی از پارامترهای الگوریتم تطابق شکل، بازشناسی را بر روی مجموعه تمرین انجام دهیم. نمونه‌هایی که در این مرحله به عنوان نماینده انتخاب شدند در شکل ۳ نشان داده شده اند. در این مرحله برای ارقام ۰ و ۱ با توجه به این که نوع نگارشی کمتری دارند، تنها یک نماینده و برای سایر ارقام دو نماینده در نظر گرفته ایم.



شکل ۳. نماینده‌های انتخاب شده در کلاسهای ده گانه ارقام برای تعیین پارامترهای الگوریتم.

۱. فاصله‌ای که از جمع فاصله‌های بین تک تک نقاط متناظر دو شکل بدست می‌آید. این فاصله را با  $C_{context}$  نشان می‌دهیم (رابطه ۴).
۲. فاصله‌ای که بر حسب انرژی لازم برای تبدیل اسپلاین صفحه نازک بدست می‌آید. این فاصله را با  $C_{tps}$  نشان می‌دهیم (رابطه ۱۵).

$$C_{total} = W_{context}C_{content} + W_{tps}C_{tps} \quad (16)$$

$C_{total}$  فاصله‌ای است که از آن برای تعیین میزان اختلاف دو رقم مورد نظر استفاده می‌کنیم.  $W_{tps}$  و  $W_{context}$  ضرایب وزنی این دو فاصله هستند. بلونجی و همکارانش به ضرایب وزنی  $W_{context}=1.6$  و  $W_{tps}=0.3$  برای بازشناسی نهایی روی مجموعه ارقام دستنویس MNIST رسیده‌اند [۱۲]. به همین دلیل ما هم در ابتدا برای آنکه انتخابی تقریباً درست برای این ضرایب وزنی داشته باشیم همین مقادیر را انتخاب کردیم.

پس از تعیین پارامترهای فوق نماینده‌هایی را برای هر یک از ارقام انتخاب کردیم و با تغییر نمونه‌ها سعی کردیم نرخ بازشناسی را در مجموعه داده تمرین بیشتر کنیم. پس از آنکه این نمونه‌ها را انتخاب کردیم، در مرحله بعد سعی کردیم براساس این نمونه‌ها و تغییر پارامترهای الگوریتم به بهبود نرخ نمونه برداری در مجموعه تمرین پردازیم. آزمایشهای گوناگونی برای انتخاب بهترین پارامترهای الگوریتم فوق بر اساس نمونه‌هایی که در شکل ۳ نشان داده شده‌اند، انجام شد. تعداد عناصر زاویه‌ای مختصات لگاریتمی-قطبی که در این آزمایشها از آنها استفاده شد، ۸، ۱۰، ۱۲، ۱۴ و ۱۶ بود. برای تعداد عناصر شعاعی نیز از انتخابهای ۳ تا ۶ استفاده کردیم. برای تعداد نقاط نمونه برداری شده انتخابهای ۵۰، ۶۰، ۷۵ و ۱۰۰ را در نظر داشتیم. پارامترهایی که در نهایت انتخاب شدند، در جدول ۱ نشان داده شده‌اند. نرخ بازشناسی که با این پارامترها و به کمک

هدف ما در این قسمت این است که پارامترهایی را که منجر به بهترین بازشناسی بر روی مجموعه تمرین می‌شوند، به عنوان پارامترهای تثبیت شده برای بازشناسی بر روی کل مجموعه ارقام انتخاب کنیم. پارامترهای الگوریتم تطابق شکل که باید انتخاب شوند به صورت زیر هستند.

#### ۴-۲-۱- تعداد عناصر زاویه‌ای مختصات لگاریتمی-قطبی

منظور از تعداد عناصر زاویه‌ای مختصات لگاریتمی-قطبی، این است که در محاسبه هیستوگرام برای هر نقطه نمونه برداری شده روی کانتور، از چه تعداد عناصر زاویه‌ای استفاده می‌کنیم، به عنوان مثال همانگونه که در شکل ۱ ج. پیداست، تعداد عناصر زاویه‌ای ۱۲ است.

#### ۴-۲-۲- تعداد عناصر شعاعی در مختصات لگاریتمی-قطبی

منظور از تعداد عناصر شعاعی مختصات لگاریتمی-قطبی این است که در محاسبه هیستوگرام برای هر نقطه نمونه برداری شده روی کانتور از چه تعداد عناصر شعاعی استفاده می‌کنیم. به عنوان مثال در شکل ۱ ج. پنج عنصر شعاعی استفاده شده است.

#### ۴-۲-۳- روش پیدا کردن نقاط متناظر بین دو شکل

برای پیدا کردن تناظری یک به یک بین نقاط نمونه برداری شده شکل اول و شکل دوم از دو روش ریاضی می‌توان استفاده کرد [۱۲]. روشهایی که به این منظور پیشنهاد شده‌اند روش مجاری [۱۵] و روش یونکر [۱۶] هستند.

#### ۴-۲-۴- انتخاب ضرایب وزنی فواصل

همانگونه که در بخش قبل اشاره شد، برای مقایسه دو شکل با هم از دو نوع فاصله استفاده می‌کنیم.

نمونه های انتخاب شده در شکل ۳ بر روی مجموعه تمرین بدست آمد ۸۸/۴٪ است.

جدول ۱. پارامترهای نهایی الگوریتم تطابق شکل برای بازشناسی ارقام فارسی

۱۲	تعداد عناصر زاویه ای مختصات لگاریتمی - قطبی
۵	تعداد عناصر شعاعی مختصات لگاریتمی - قطبی
۶۰	تعداد نقاط نمونه برداری شده روی هر کانتور
یونکر	روش پیدا کردن نقاط متناظر دو شکل
$W_{context} = 1.8$ $C_{TPS} = 0.2$	ضرایب وزنی فواصل

در مرحله آخر با افزودن نماینده هایی به نماینده های ۱ و ۹، که در شکل ۵ نشان داده شده اند، به نرخ بازشناسی ۸۹/۹٪ رسیدیم. ماتریس کارایی این آزمایش در جدول ۲ نشان داده شده است. همان طور که در این جدول دیده می شود، ارقام ۷ و ۸ به دلیل تنوع کمتری که دارند بیشترین نرخ شناسایی را داشته (۹۵٪ و ۹۷٪) و رقم ۳ به دلیل تنوع زیاد و نزدیکی نحوه نگارش آن با رقم ۲ دارای کمترین میزان بازشناسی است (۸۱٪).

۴ ۵ ۶ ۷ ۸ ۹

شکل ۴. نماینده های جدیدی که به نماینده های قبلی ۴، ۵، ۶ و ۷ اضافه شده اند.

۹ ۱

شکل ۵. نماینده های جدیدی که به نماینده های قبلی ۱ و ۹ اضافه شده اند.

دلیل افزودن نماینده ۴ به دیگر نماینده ها این است، که این نوع از رقم ۴ بوسیله سایر نمونه هایی که قبلا برای رقم ۴ داشتیم پوشش داده نمی شد. دو نمونه رقم ۵ هم به این دلیل اضافه شدند که چون به طور کامل بسته نشده اند، دارای کانتوری کاملا متفاوت با نمونه های قبلی رقم ۵ هستند. در مورد ارقام ۷ و ۸ هم، این دو نمونه به این دلیل اضافه شدند که نمونه های نامتقارن این ارقام را پوشش دهند.

### ۳-۴- بازشناسی بر روی کل مجموعه ارقام

پس از آنکه پارامترهای الگوریتم تطابق شکل تثبیت شدند، به کمک همان نماینده هایی که در بخش تمرین انتخاب کردیم، بازشناسی را بر روی کل مجموعه ارقام انجام دادیم. نرخ بازشناسی این آزمایش ۸۴.۷٪ شد. در مراحل بعد برای بهبود نرخ بازشناسی به اصلاح نماینده های انتخاب شده پرداختیم. به این منظور نماینده های ۳ و ۶ را اصلاح کردیم و یک نماینده به نماینده های قبلی ارقام ۴، ۷ و ۸ اضافه کردیم. همچنین دو نماینده به نماینده های رقم ۵ اضافه نمودیم. این نماینده های جدید در شکل ۴ نمایش داده شده اند. در این مرحله به نرخ بازشناسی ۸۸/۴٪ رسیدیم.

جدول ۲. ماتریس کارایی آزمایش نهایی

	0	1	1	2	2	3	3	4	4	4	5	5	5	5	6	6	7	7	8	8	9	9	9	Sum	%
0	120	.	.	.	.	.	.	.	.	2	2	7	1	.	.	.	.	.	.	.	.	.	.	132	90.91
1	3	61	59	6	1	.	.	1	1	.	.	.	.	2	.	.	.	.	.	.	2	.	.	136	88.24
2	.	.	.	85	24	3	.	.	.	3	.	.	.	.	1	1	.	.	.	.	.	2	1	120	90.83
3	.	.	.	5	5	52	48	.	1	11	.	.	.	.	.	.	.	.	.	.	.	.	.	122	81.97
4	.	.	2	7	.	4	2	40	45	20	.	.	.	.	.	.	.	.	.	.	.	.	.	120	87.50
5	12	.	.	.	.	.	1	.	.	1	39	28	24	11	.	.	.	.	3	.	.	.	.	119	85.71
6	.	.	1	5	.	.	.	.	.	3	.	.	.	.	95	16	.	3	.	.	2	.	.	125	88.80
7	.	.	.	3	.	.	.	.	.	.	.	.	.	.	3	.	81	56	.	.	.	.	.	143	95.80
8	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	119	15	3	.	.	138	97.10
9	.	.	3	2	.	.	.	2	.	1	.	.	.	.	1	.	.	.	.	.	38	49	29	125	92.80

## ۵- مقایسه با روشهای دیگران

در این بخش مقایسه ای بین نتایج روشی که ما برای بازشناسی ارقام دستنویس فارسی از آن استفاده کرده ایم با سایر روشهایی که در سالهای اخیر انجام شده اند، صورت می دهیم. این مقایسه در جدول ۳ نمایش داده شده است. همانگونه که مشخص است، نرخ بازشناسی ارقام به کمک تطابق شکل تنها از دو روش دیگر ([۷] و [۱۸]) کمتر است. اما باید توجه داشت که در اکثر این روشها از مراحل تکمیلی پس پردازش و یا پیش پردازش استفاده شده است که طبیعتاً نرخ بازشناسی را بالا خواهد برد. به عنوان مثال در پیش پردازش می توان کیفیت تصویری ارقام را بهبود داد [۱۸]، یا ارقامی مثل ۰ و ۱ را قبل از بازشناسی اصلی از بقیه ارقام جدا کرد [۸].

جدول ۳. مقایسه بین نتایج روشهای مختلفی که در زمینه بازشناسی ارقام انجام شده اند.

x. روشهایی را نشان می دهد که از پس پردازش یا پیش پردازش استفاده کرده اند.

همچنین در پس پردازش نیز با توجه به اشتباهات متداولی که طبقه بندی اصلی می کند، با انتخاب ویژگیهایی متفاوت از ویژگیهای اصلی، می توان بازشناسی را اصلاح کرد [۷][۸].

همچنین باید توجه داشت که در روش این مقاله، تنها از یک تا ۴ نمونه به عنوان نماینده هر کلاس استفاده شده است، اگر همانطور که در بخش ۴-۱ ذکر شده است، از روش طبقه بندی نزدیکترین همسایه استفاده شود، با هزینه بیشتر، نتایج بهتری به دست خواهد آمد.

ردیف	مرجع	استخراج ویژگی	روش طبقه بندی	تعداد ارقام	نرخ بازشناسی
۱	[۷]	روش مکانهای مشخصه مورب	پنج همسایه نزدیکتر	۱۲۷۷۸	٪۸۵/۰۷ تا ٪۹۵/۵
۲	[۹]	DTW	کمترین فاصله از نماینده های هر کلاس	-----	٪۷۵
۳	[۱۰]	گشتاورهای هندسی و تبدیل فوریه	شبکه عصبی	-----	٪۶۰
۴	[۱۸]	ویژگیهای ساختاری	درخت تصمیم باینری	۳۲۰۰	٪۹۳
۵	[۸]	3-tuple	کمترین فاصله از نماینده های هر کلاس	۵۰۰	٪۸۱
۶	[۱۱]	zoning	روش فازی	۵۰۰	٪۸۳
۷	[۱۹]	انتخاب ویژگی با الگوریتمهای وراثتی	روش فازی	۵۰۰	٪۸۶/۲
۸	روش تطابق شکل	به کمک اطلاعات جانبی حول نقاط کانتور	کمترین فاصله از نماینده های هر کلاس	۱۲۸۸	٪۸۹/۹

## ۶- نتیجه گیری و پیشنهاد

در این مقاله از یک الگوریتم تطابق شکل [۱۲] برای بازشناسی ارقام دستنویس فارسی استفاده شده است. برای هر نقطه نمونه برداری شده روی کانتور شکل، توصیفگری بر اساس توزیع مکانی نقاط دیگر کانتور بدست آوردیم. برای تعیین میزان شباهت دو شکل، ابتدا بر اساس این توصیفگرها تناظری یک به یک بین نقاط نمونه برداری شده روی کانتور شکل اول با نقاط روی کانتور شکل دوم به دست آورده و جمع فواصل بین نقاط متناظر دو شکل را به عنوان معیاری برای عدم شباهت دو شکل انتخاب کردیم. سپس تبدیلی تعریف کردیم که نقاط کانتور شکل اول را بر روی نقاط متناظرشان در کانتور شکل دوم قرار دهد. میزان پیچیدگی این تبدیل معیار دیگری برای عدم شباهت دو شکل است. در نهایت با بهینه کردن پارامترهای مختلف این الگوریتم از آن برای شناسایی ارقام دستنویس فارسی استفاده کردیم.

در یک آزمایش بر روی ۱۲۸۸ رقم که افراد مختلف نوشته اند ۸۹.۹٪ بازشناسی درست حاصل شده است که این نتیجه بدون هیچ پس پردازشی بدست آمده است. البته مشخص است که تغییر هر یک از پارامترهایی که در جدول ۱ آمده است، در کارایی الگوریتم تطابق شکل تاثیر خواهد داشت، اما آنچه مهم است ترکیب مناسب این پارامترها است. مقدار هر یک از این پارامترها بر اساس ویژگیهایی که مجموعه ارقام دستنویس مورد استفاده ما دارند، دارای محدودیتهایی خواهند بود. برای مثال چون ارقام دستنویس بکار رفته از نظر اندازه نرمالیزه نیستند، تعداد نقاط روی مرز آنها برای ارقام مختلف متفاوت است.

برای بالا بردن نرخ بازشناسی می توان ارقام صفر را برحسب تعداد نقاط روی مرز آنها بازشناسی کرد و در نتیجه الگوریتم تطابق شکل را برای ارقام غیرصفر به کار برد. در

این صورت می توان از تعداد نقاط نمونه برداری شده بیشتری برای توصیف ارقام استفاده کرد و در نتیجه توصیف قوی تری از ارقام به دست آورد.

مشکل دیگری که در این نوع روش بازشناسی وجود دارد این است که با وجود آنکه ما در این مقاله برای هر رقم، سعی کردیم از تعداد مناسبی نماینده جهت معرفی آن استفاده کنیم و همچنین برای انتخاب نماینده های بهینه هم، آزمایشهای گوناگونی را پس از تثبیت پارامترهای الگوریتم انجام دادیم و تا حد امکان نیز به نماینده های بهینه برای بازشناسی کلی دست یافتیم، اما با این حال به دلیل گوناگونی نحوه نگارش ارقام دستنویس فارسی، این تعداد ۲۳ نماینده نمی تواند تمام حالات موجود را پوشش دهد. نحوه انتخاب نماینده و تعداد نماینده های هر کلاس هنوز جای کار دارد.

## فهرست منابع

[1] Ø.D.Trier; A.K.Jain; T.Taxt; "Feature extraction methods for character recognition – a survey"; Pattern Recogniton; vol. 29; no. 4; 1996; pp. 641–662

[2] R.Plamondon; S. Srihari; "On-line and off-line handwritten recognition: a comprehensive survey"; IEEE Trans. On Pattern Analysis and Machine Intelligence; vol. 22; no. 1; January 2000; pp. 63-83

[3] N. Arica; F.T. Yarman-Vural; "An overview of caracter recognition focused on off-line handwriting"; IEEE Trans. on Systems; Man and Cybernetics-Part C: Application and Reviews; vol. 31; no. 2; May 2001; pp. 216-233

[4] C. Suen; M. Berthod; "Automatic Recognition of Handprinted Characters- The

[13] D. Cheng; H. Yan; "Recognition of Handwritten Digits Based on Contour Information". Pattern Recognition; vol. 31; no. 3; 1999; pp. 235-255

[14] S. Chakravarthy; B. Kompella; "The Shape of handwritten characters"; Pattern Recognition Letters; vol. 24; 2003; pp. 1901-1913

[15] C. Papadimitriou; k. Steiglitz; "Combinatorial Optimization: Algorithm and Complexity"; Prentice Hall 1982.

[16] R. Jonker; T. Volgenant; "A shortest augmenting path algorithm for Dense and Sparse linear assignment problems"; Computing; vol. 38; 1987; pp. 325-240.

[17] Bookstein; F.L.; "Principal warps: Thin Plate Splines and Decompositions of deformations"; IEEE Trans. Pattern Analysis and Machine Intelligence; vol. 11; no. 6; 1989; pp. 567-585

[۱۸] ح. کتابدار؛ "شناسایی ارقام دستنویس فارسی به روش ساختاری"؛ ششمین کنفرانس مهندسی برق ایران؛ ۱۳۷۷؛ ص ۱۹\_۴ تا ۲۴\_۴

[۱۹] س. م. رضوی؛ ه. صدوقی یزدی؛ ا. کبیر؛ "انتخاب ویژگی برای بازشناسی ارقام دستنویس فارسی به کمک الگوریتمهای وراثتی"؛ هفتمین کنفرانس سالانه انجمن کامپیوتر ایران؛ ۱۳۸۰؛ ص ۲۸۵ تا ۲۹۲

state of the Art"; Proceedings of the IEEE vol 68; no 4; April 1980; pp. 469-487

[5] A. Amin; "Off-Line Arabic Character Recognition: The State of the Art"; Pattern Recognition; vol. 31; no. 5; 1998; pp. 517-530

[۶] ر. عزمی؛ ا. کبیر؛ "شناسایی آماری حروف دستنویس فارسی"؛ دومین کنفرانس مهندسی برق ایران؛ ۱۳۷۳؛ ص ۲۷۷ تا ۲۸۵

[۷] ح. ر. نفیسی؛ ا. کبیر؛ "شناسایی ارقام دستنویس فارسی"؛ دومین کنفرانس مهندسی برق ایران؛ ۱۳۷۳؛ ص ۲۹۵ تا ۳۰۴

[۸] س. م. رضوی؛ ا. کبیر؛ "خواندن اتوماتیک فرمهای انتخاب درس"؛ سومین کنفرانس بین المللی سالانه انجمن کامپیوتر ایران؛ ۱۳۷۶؛ ص ۶۴ تا ۶۹

[۹] ک. مسروری؛ و. پور محسنی خامنه؛ "شناسایی ارقام دستنویس با استفاده از الگوریتم DTW"؛ چهارمین کنفرانس بین المللی انجمن کامپیوتر ایران؛ ۱۳۷۷؛ ص ۱ تا ۷

[10] S. Mansoori; H. Hassibi; F. Rajabi; "A Heuristic Persian Handwritten Digit Recognition with Neural Network"; The 6th Iranian Conference on Electrical Engineering; vol. 3; 1998; pp. 131-135

[۱۱] و. جوهری مجد؛ س. م. رضوی؛ "بازشناسی فازی ارقام دستنویس فارسی"؛ اولین کنفرانس ماشین بینایی و پردازش تصویر ایران؛ ۱۳۷۹؛ ص ۱۴۴ تا ۱۵۱

[12] S. Belongie; J. Malik; J. Puzicha; "Shape Matching and Object Recognition using Shape context." IEEE transaction on Pattern Analysis and Machine Intelligence; vol. 24; no. 4; 2002; pp. 509-522

## Recognition of Handwritten Farsi Digits by Shape Matching

Alireza Darvish<sup>1</sup>, Ehsanollah Kabir<sup>2\*</sup>, Hossein Khosravi<sup>3</sup>

1. Msc. Student of Electrical Engineering, Department of Engineering, Tarbiat Modarres University

2. Associate Professor of Electrical Engineering, Department of Engineering, Tarbiat Modarres University

3. Msc. Student of Electrical Engineering, Department of Engineering, Tarbiat Modarres University

\* P.O Box 14115-143, kabir@modares.ac.ir

In this paper, we used a shape matching algorithm to recognize Farsi digits. For each sampled point on the contour of a shape, we obtain a descriptor showing the distribution of the other points of the contour, with respect to this point. Based on these descriptors, we find the corresponding points of the two contours and take the sum of their distances as a dissimilarity measure between two shapes. Then we define a geometric transformation that maps the sampled points of the one shape to the corresponding points of the other shape. The bending energy of this transform is taken as the second dissimilarity measure between two shapes. We optimized the parameters of the matching algorithm for the recognition of Farsi digits and used the method of minimum distance from the class prototypes for the recognition. In a test on a set of 1288 digits, we obtained a recognition rate of 89.9%. This result was obtained without any post processing.

**Keywords:** Shape Matching, Contour, Local Descriptor, Geometrical Transform, Handwritten Recognition, Farsi Digits.