

تشخیص جنسیت نویسنده مستقل از متن و زبان نوشتاری با استفاده از پالایش پویای نمادین مبتنی بر تبدیل رادن

کاظم نوری هفت چشمه، رضا خدادادی، یونس اکبری، سیدمحمد رضوی و حسن احمدی ترشیزی

از قبیل معلم‌ها و منشی‌ها می‌توانند تمایز بین دست‌خط زن و مرد را یاد بگیرند [۵] و [۶]. روانشناسان، تفاوت‌ها در دست‌خط مردان و زنان را ناشی از تفاوت در هماهنگی حرکتی یا انواع هورمون‌های متفاوت می‌دانند. محققان در حوزه دست‌خط نیز با ارتباط جنسیت و دست‌خط موافقت [۷] و [۸]. به طور معمول، روانشناسان پیشنهاد می‌کنند که صفاتی مانند تمیزبودن، به خوبی سازمان یافته بودن، کوچک و متقارن بودن، از ویژگی‌های دست‌خط زنان است. از طرف دیگر شتاب زده بودن، ناهمواری، کثیف بودن، تند و تیز و شیب‌دار بودن دست‌خط، بیشتر متعلق به نویسندگان مرد می‌باشد [۶] و [۹]. نمونه‌ای از دست‌خط زنان و مردان در شکل ۱ نشان داده شده است.

با پیشرفت در تحلیل تصویر و تکنیک‌های طبقه‌بندی الگو، تحلیل دستی دست‌خط با سیستم‌های خودکار جایگزین شد. یک سیستم برای طبقه‌بندی جنسیت از روی دست‌خط با استفاده از مجموعه ویژگی‌های هدفمند در محاسبه یک زیرمجموعه متمایز از خصوصیات ارائه گردید [۸] که در میان صفات مختلف متمایز از مردان و زنان نویسنده، روی کجی، دواری و انحنای پاکیزگی و خوانایی و بافت نوشتاری تمرکز شده است. در زمینه تشخیص جنسیت از روی دست‌خط تا کنون کارهای محدودی از قبیل [۸] و [۱۰] تا [۱۷] انجام پذیرفته است. در [۸] روش‌های فرکتال و کد زنجیره‌ای برای استخراج ویژگی و طبقه‌بندی شبکه‌های عصبی و ماشین بردار پشتیبان^۱ به منظور تشخیص جنسیت مورد استفاده قرار گرفته‌اند. نویسندگان [۱۰] به کارگیری روش‌های فرکتال^۲ و چندفرکتالی^۳ را برای استخراج ویژگی و طبقه‌بندی ماشین بردار پشتیبان را به منظور تشخیص جنسیت مورد مطالعه قرار داده‌اند. در [۱۱] بانک اطلاعاتی از دست‌خط‌های برون خط و برخط^۴ به منظور تشخیص جنسیت، مورد ارزیابی قرار گرفته و از مدل مخلوط گوسی^۵ به عنوان طبقه‌بند استفاده شده است. نویسندگان در [۱۲] از تکنیک استخراج ویژگی جهت اصلاح شده^۶ برای استخراج ویژگی‌ها بهره برده‌اند و از طبقه‌بند شبکه‌های عصبی چندلایه استفاده نموده‌اند. مقاله [۱۳] از ویژگی‌های جهت^۷ و ویژگی‌های انتقال^۸ برای استخراج ویژگی‌ها استفاده کرده و طبقه‌بند شبکه‌های عصبی را به کار برده است. مقاله [۱۴] دو ویژگی مبتنی بر شیب را مورد بررسی قرار داده است که عبارتند از نمودار شیب‌گرا^۹ که بر توزیع جهت شیب در تصاویر تأکید دارد و عملکرد بالایی را در برنامه‌های مختلف از قبیل

چکیده: در این مقاله یک سیستم خودکار بر مبنای روشی جدید در استخراج ویژگی برای تشخیص جنسیت افراد از روی تصاویر اسکن شده نمونه دست‌خط ارائه شده است. در روش پیشنهادی به منظور نشان دادن تمایز بین نمونه‌های دست‌خط زنان و مردان، ابتدا از تصویر دست‌نویسته، تبدیل رادن گرفته می‌شود و سپس با استفاده از یک ابزار تحلیلی در سیستم‌های دینامیکی با عنوان پالایش پویای نمادین، ویژگی‌های هر نمونه دست‌خط استخراج می‌گردد. آموزش و طبقه‌بندی ویژگی‌های استخراج شده از نمونه‌های دست‌خط با شبکه عصبی پرسپترون چندلایه انجام شده است. در پایان با هدف بررسی کارایی روش پیشنهادی، آزمایش‌هایی بر روی بانک اطلاعاتی MSHD صورت پذیرفت. علاوه بر آزمایش تشخیص جنسیت بر روی کل بانک اطلاعاتی، دو چالش جدید تشخیص جنسیت مستقل از متن و زبان نوشتاری نیز بررسی شده است. آزمایش‌های انجام شده نشان می‌دهد روش پیشنهادی میزان دقت تشخیص را نسبت به کارهای قبلی که از روش‌های جدیدی در تحلیل دست‌خط از قبیل فرکتال‌ها، کدهای زنجیره‌ای و بافت‌ها بهره می‌برند، بهبود داده است. بهترین نرخ دقت به دست آمده در آزمایش‌ها ۸۴٫۹ درصد گزارش شده است.

کلیدواژه: پالایش پویای نمادین، تبدیل رادن، تشخیص جنسیت نویسنده، دست‌خط برون خط.

۱- مقدمه

نوشتن یکی از قدیمی‌ترین روش‌های برقراری ارتباط در بین تمدن‌هاست که در طول زمان تکامل و توسعه یافته است. یک فرد نوشتن را با کپی کردن شکل‌ها از یک کتاب استاندارد که خود آن به عنوان تابعی از موقعیت جغرافیایی، زمانی، اجتماعی و تمدنی است یاد می‌گیرد. تحلیل دست‌خط موجب ایجاد فضای جستجوی جذاب برای روان‌شناسان، خط‌شناسان، تحلیل‌گران پزشکی قانونی، مؤسسات مالی-اعتباری، بانک‌ها و دستگاه قضایی شده است. ارتباط بین شخصیت و دست‌خط قابل بحث است و هنوز در زمینه‌های علمی، اعتبارسنجی می‌شود [۱] و [۲]. تحقیقات روانشناسی نشان داده‌اند که جنسیت می‌تواند از طریق دست‌خط با درجات مختلفی از موفقیت، پیش‌بینی شود [۳] و [۴]. برای مثال افرادی

این مقاله در تاریخ ۷ آبان ماه ۱۳۹۴ دریافت و در تاریخ ۱۳ مهر ماه ۱۳۹۵ بازنگری شد.

کاظم نوری هفت چشمه، گروه ریاضی، دانشکده ریاضی، آمار و علوم کامپیوتر، دانشگاه سمنان، سمنان، (email: knouri@semnan.ac.ir).

رضا خدادادی، دانشکده فنی و مهندسی، دانشگاه آزاد اسلامی واحد مشهد، مشهد، (email: re_khodadadi1362@yahoo.com).

یونس اکبری، گروه ریاضی، دانشکده ریاضی، آمار و علوم کامپیوتر، دانشگاه سمنان، سمنان، (email: akbari_younes@semnan.ac.ir).

سیدمحمد رضوی، گروه مهندسی الکترونیک، دانشکده مهندسی برق و کامپیوتر، دانشگاه بیرجند، بیرجند، (email: smrazavi@birjand.ac.ir).

حسن احمدی ترشیزی، گروه کامپیوتر، دانشکده فنی و مهندسی، دانشگاه آزاد اسلامی واحد مشهد، مشهد، (email: hahmadi@mshdiau.ac.ir).

1. Support Vector Machine
2. Fractal
3. Multi Fractal
4. Off-line and On-line
5. Gaussian Mixture Models
6. Modified Direction Feature
7. Direction Feature
8. Transition Feature
9. Histogram of Oriented Gradient

Internet est un système n
un ensemble standardisé de
réseau de réseaux, sans
aussi bien publics, privés,

Internet est un système
informatiques, utilisant
de transfert de donn
centre névralgique,

Internet est un système mo
standardisé de protocoles de ti
névralgique, composé de mill
et gouvernementaux. Internet tr

(الف)

Internet se est un m
informatique, utilis
de transfert de donn
centre névralgique, e

internet est un système mond
informatiques, utilisant un
de transfert de données
sans centre névralgique,

Internet est un système m
informatiques, utilisant un
de transfert de données, e
sans centre névralgique,

(ب)

شکل ۱: نمونه‌های از دستخط زنان و مردان، (الف) دستخط زنان و (ب) دستخط مردان [۸].

نشان دادن تمایز کارآمدتری بین دستخط زنان و مردان ارائه می‌دهد. هدف اصلی، مدل‌سازی با به کارگیری ابزاری جدید در سیستم‌های دینامیکی به منظور تشخیص جنسیت از روی دستخط می‌باشد. ابتدا با استفاده از تبدیل دوبعدی رادن^۴، تصویر دستخط به یک فرم موج تبدیل می‌شود. سپس با استفاده از پالایش پویای نمادین^۵ (SDF)، اطلاعات مهم این موج به عنوان ویژگی‌های متمایزکننده دستخط استخراج می‌شود. در آخر با استفاده از شبکه‌های پرسپترون چندلایه، طبقه‌بندی بر اساس جنسیت (مرد یا زن) صورت می‌پذیرد.

در بخش ۲ مفاهیم علمی استفاده‌شده در این تحقیق به صورت مختصر بیان می‌شود. در بخش ۳ ساختار بانک اطلاعاتی استفاده‌شده در آزمایش‌ها توضیح داده شده است. روش پیشنهادی و پیاده‌سازی آن در بخش ۴ بیان شده است. نتایج آزمایش‌های انجام‌شده در بخش ۵ گزارش می‌شود و بخش پایانی به نتیجه‌گیری و کارهای آینده از این مقاله اختصاص یافته است.

۲- مفاهیم علمی

خلاصه‌ای از مفاهیم علمی استفاده‌شده در این مقاله که شامل تبدیل رادن و پالایش پویای نمادین می‌باشد، در ادامه این بخش آمده است.

۲-۱ تبدیل رادن

تبدیل رادن، یک ابزار مهم در زمینه پردازش تصویر و تشخیص الگو با کاربردهای فراوانی در این حوزه‌ها می‌باشد [۱۸] و [۱۹]. به طور کلی، تبدیل رادن دوبعدی یک تصویر به صورت $RT_{\theta}[f(x, y)]$ نشان داده شده و در زیر تعریف می‌شود

$$RT_{\theta}[f(x, y)] = \iint f(x, y) \delta(r - x \cos \theta - y \sin \theta) dx dy \quad (1)$$

که $\delta(\cdot)$ تابع دیراک و $r \in (-\infty, +\infty)$ ، $\theta \in [0, \infty)$ است. خروجی تبدیل رادن که بر روی یک دستخط اعمال شده در شکل ۲ آمده است.

۲-۲ پالایش پویای نمادین

در این قسمت مفاهیم اساسی پالایش پویای نمادین برای استخراج ویژگی را ارائه می‌دهیم. گام‌های اساسی روش SDF برای استخراج ویژگی در زیر آورده شده است [۲۰]:

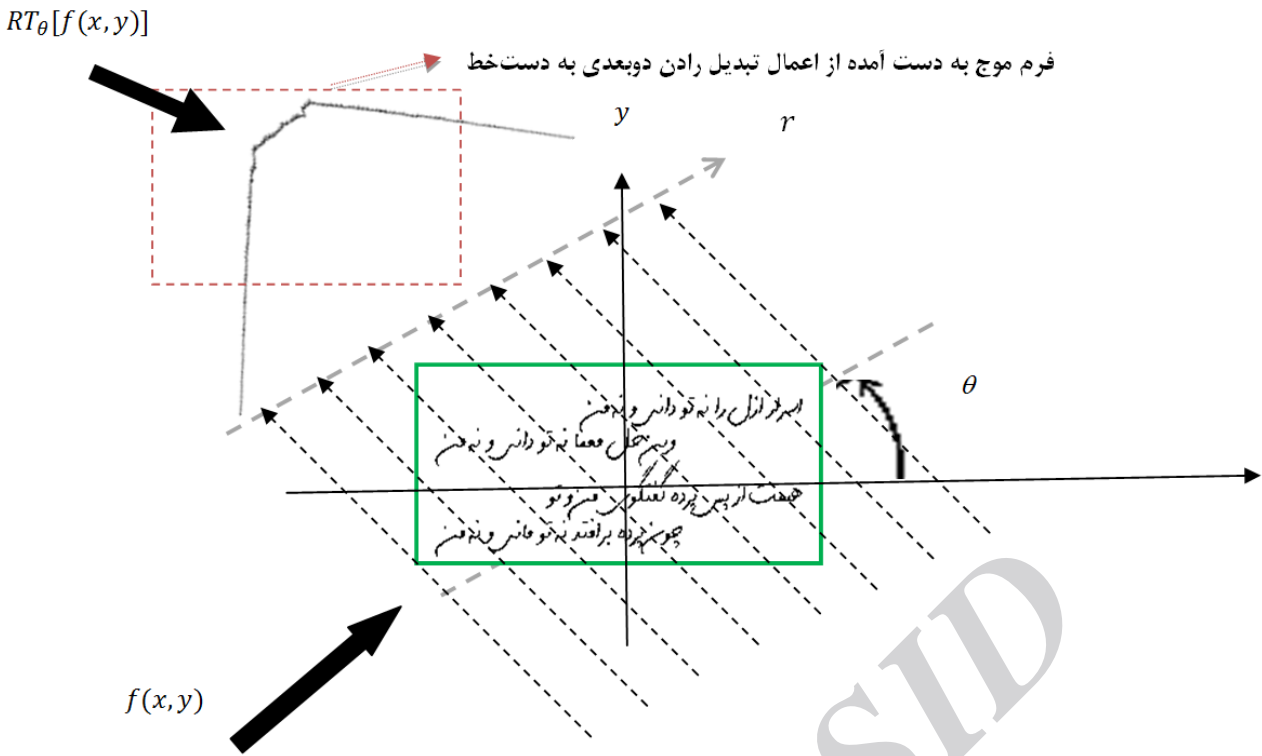
تشخیص چهره و تشخیص امضا و دستخط دارد. نمودار شیب‌گرا بر روی یک شبکه متراکم قرار گرفته بر روی تصویر محاسبه می‌شود و هر سلول شبکه نمودار مربوط به خود را دارا می‌باشد که توسط الگوریتم مربوطه محاسبه می‌گردد. ویژگی دوم الگوی باینری محلی شیب^۱ نامیده می‌شود که ویژگی جدید مبتنی بر شیب می‌باشد و اصل آن شامل بهره‌برداری از الگوهای باینری محلی یکنواخت به منظور محاسبه نمودار شیب‌گرا است. این دو روش بر روی بانک‌های اطلاعاتی به نام KHATT و IAM-۱ که به ترتیب شامل زبان‌های عربی و انگلیسی می‌باشد اعمال شده‌اند. در این مقاله از طبقه‌بند ماشین بردار پشتیبان استفاده شده است. در [۱۵] از تابع زاویه تانژانت، تابع انحنا و توصیف فوریه بر روی بانک اطلاعاتی برون‌خط که شامل ۳۷۶۶ دستخط از زنان و مردان می‌باشد، جهت استخراج ویژگی استفاده شده است. استفاده از توصیف فوریه نشان می‌دهد که فرکانس‌های بالاتر مربوط به ویژگی‌های دستخط مردان (تند و تیزبودن، آشفته) می‌باشد و اغلب ویژگی‌های دستخط زنان (از قبیل مرتب و تمیزبودن) دارای فرکانس بالا نمی‌باشند. همچنین نتایج به دست آمده از تابع زاویه تانژانت و تابع انحنا نشان می‌دهد که ویژگی تند و تیزبودن دستخط مردان تغییرات سریع‌تری در تابع زاویه تانژانت را دارا می‌باشد. مقاله [۱۶] از ویژگی‌های هندسی از قبیل ویژگی جهت، ویژگی انحنا، ویژگی خمیدگی، ویژگی مبتنی بر کد زنجیره‌ای و ویژگی مبتنی بر جهت لبه برای استخراج ویژگی بر روی بانک اطلاعاتی QUWI استفاده کرده است. این ویژگی‌ها به وسیله چند طبقه‌بند که شامل جنگل‌های تصادفی^۲ و تحلیل تفکیک‌کننده‌های هسته^۳ می‌باشند ترکیب می‌شوند. نوآوری این مقاله در ترکیب ویژگی‌ها به وسیله طبقه‌بندهای جنگل تصادفی و تحلیل تفکیک‌کننده‌های هسته می‌باشد.

مقاله [۱۷] شامل مروری کامل بر بانک‌های اطلاعاتی مورد استفاده در تحلیل دستخط تا سال ۲۰۱۵ می‌باشد. این مقاله می‌تواند در آشنایی و انتخاب بانک‌های اطلاعاتی در زمینه‌های کاربردی مختلف دستخط از قبیل شناسایی نویسنده از روی دستخط و همچنین تشخیص جنسیت از روی دستخط به پژوهشگران کمک کند. تمامی کارهای بررسی‌شده در بالا به جز [۸] بر روی بانک‌های اطلاعاتی آزمایش شده‌اند که در اختیار عموم نمی‌باشد.

این مقاله، روش جدیدی در مرحله استخراج ویژگی به منظور

4. Radon Transform
5. Symbolic Dynamic Filtering

1. Gradient Local Binary Pattern
2. Random Forest
3. Kernel Discriminant Analysis



شکل ۲: خروجی تبدیل رادن اعمال شده بر روی دست خط.

سیگنال در یک ناحیه مشخص قرار گرفته شده باشد با نماد مرتبط با آن ناحیه کدگذاری خواهد شد. بنابراین یک نماد از الفبا به هر نقطه از سیگنال با توجه به ناحیه‌ای که در آن قرار گرفته است اختصاص می‌یابد و در نتیجه، آرایه‌ای از نمادها که دنباله نماد^۳ نامیده می‌شود ایجاد می‌گردد. در اینجا این نکته حائز اهمیت است که افراز می‌تواند با استفاده از افراز ماکسیمم آنترپی^۴ (MEP) یا افراز یکنواخت^۵ (UP) انجام پذیرد [۲۱]. ما در این مقاله از افراز ماکسیمم آنترپی استفاده نموده‌ایم.

۲-۲-۲ استخراج ویژگی‌ها توسط ساختار وضعیت ماشین‌های خودکار احتمالی متناهی

این قسمت، شامل ساختار وضعیت ماشین‌های خودکار احتمالی متناهی برای استخراج ویژگی‌ها می‌باشد. برای تولید یک ماشین خودکار احتمالی متناهی از یک دنباله نماد مشخص، می‌توان از دو روش \mathcal{E} -ماشین^۶ و زنجیره مارکوف با مرتبه D استفاده نمود. این مقاله از روش زنجیره مارکوف با مرتبه یک ($D=1$) استفاده نموده است. در ادامه روش زنجیره مارکوف با مرتبه D به صورت مختصر بیان شده است. یک پردازش نمادین تصادفی مانند $S \equiv \dots s_{-1} s_0 s_1 \dots$ ، زنجیره مارکوف با مرتبه D نامیده می‌شود اگر احتمال نماد بعدی فقط به D نماد قبلی وابسته باشد یعنی

$$P(S_i | S_{i-1} S_{i-2} \dots S_{i-D} \dots) = P(S_i | S_{i-1} S_{i-2} \dots S_{i-D}) \quad (2)$$

ساختار وضعیت ماشین خودکار احتمالی متناهی می‌تواند توسط زنجیره مارکوف و ماشین مارکوف با مرتبه D (که D عددی صحیح و مثبت است) تقریب زده شود. اساس ماشین مارکوف با مرتبه D ، جداسازی

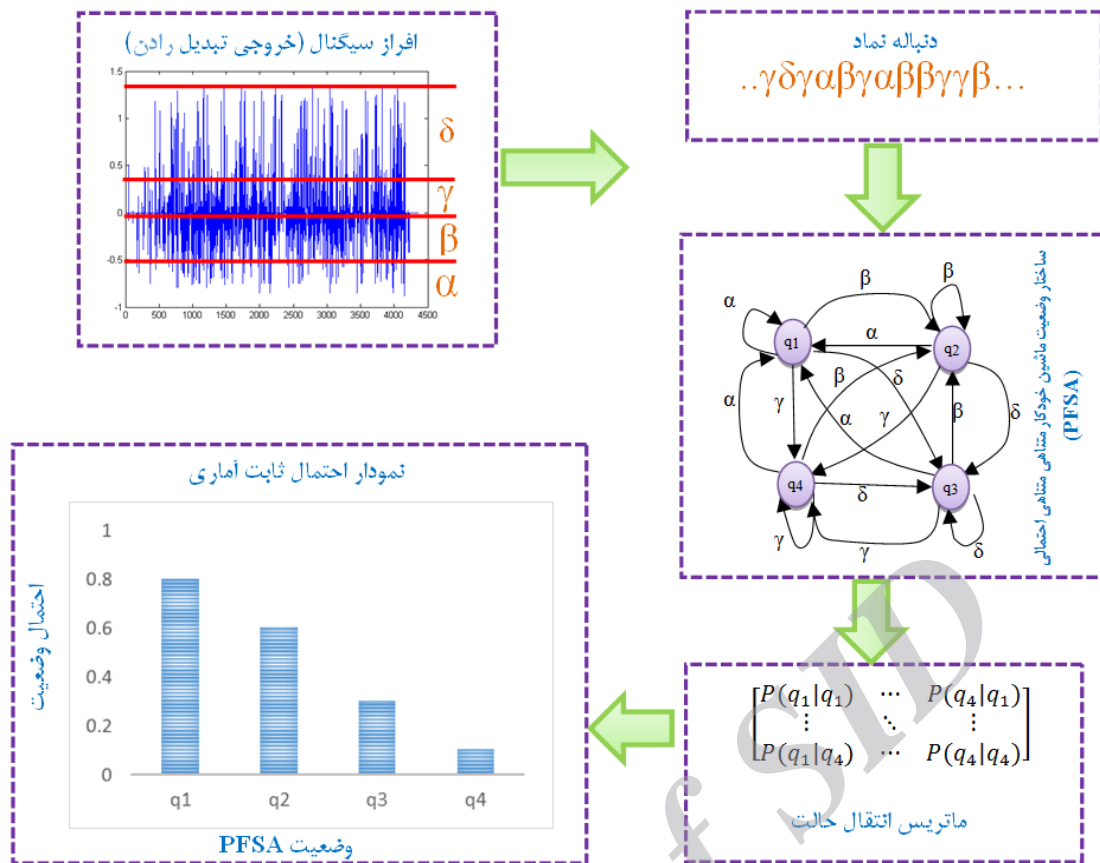
- نمادگذاری سیستم‌های پویا از داده‌های حسگر (برای مثال تصاویر و سری‌های زمانی) برای تولید دنباله‌های نمادین.
- فشرده‌سازی اطلاعات به طریق ساختار وضعیت ماشین‌های خودکار متناهی احتمالی^۱ (PFSA) از دنباله‌های نمادین تولیدشده در گام قبل برای تولید ویژگی.

۲-۲-۱ نمادگذاری سیستم‌های دینامیکی

فرض می‌کنیم Ω یک مجموعه فشرده (بسته و کراندار) در فضای فاز^۲ (شامل تمام حالات ممکن سیستم) سیستم دینامیکی سیگنال (در اینجا خروجی تبدیل رادن) باشد. مجموعه Ω به تعدادی سلول که تعداد آنها برابر با $|\Sigma|$ می‌باشد و به صورت مجموعه $Q = \{q_1, \dots, q_{|\Sigma|}\}$ نشان داده می‌شوند، افراز شده که دو به دو متمایز $(q_j \cap q_k = \emptyset, \forall j \neq k)$ و کامل می‌باشند $(\bigcup_{j=1}^{|\Sigma|} q_j = \Omega)$. در حقیقت Σ سلول‌های افراز شده را برچسب‌گذاری می‌کند. یک مسیر سیستم دینامیکی توسط یک مجموعه داده مانند $\{x_1, x_2, \dots\}$ که $x_i \in \Omega$ توصیف می‌شود. مسیر از طریق گذار از سلول‌های افراز شده با توجه به نماد مربوطه اختصاص یافته به هر نقطه x_i که توسط نگاشت $\mu: \Omega \rightarrow \Sigma$ تعریف شده است عبور می‌کند. این بدان معنی است که هر مسیر شامل دنباله‌ای از نمادها می‌باشد. فضای سیگنال (خروجی تبدیل رادن) به تعداد محدودی از سلول‌هایی که به عنوان نماد برچسب‌گذاری شده‌اند، افراز می‌شود که تعداد این سلول‌ها دقیقاً به اندازه کاردینالیته می‌باشد $(|\Sigma|)$. به عنوان مثال این روش بر روی یک نمونه فرم موج (خروجی تبدیل رادن) در شکل ۳ نشان داده شده است. در اینجا $\Sigma = \{\alpha, \beta, \gamma, \delta\}$ می‌باشد و در نتیجه $|\Sigma| = 4$ بوده و سه خط افراز که موازی با محور x ها می‌باشند، سیگنال (خروجی تبدیل رادن) را به چهار ناحیه کاملاً جدا افراز می‌کنند. هر یک از این مناطق جدا از هم با یک نماد از الفبا نام‌گذاری می‌شود. اگر هر نقطه از

3. Symbol Sequence
 4. Maximum Entropy Partitioning
 5. Uniform Partitioning
 6. \mathcal{E} -Machine
 7. D-Markov Chain

1. Probabilistic Finite State Automata
 2. Phase Space



شکل ۳: مفهوم پالایش پویای نمادین (SDF) به عنوان استخراج کننده ویژگی.

$$P(q_k | q_l) = \frac{N(q_l, q_k)}{\sum_{i=1, \dots, |Q|} N(q_l, q_i)}, \quad \forall q_l, q_k \in Q \quad (3)$$

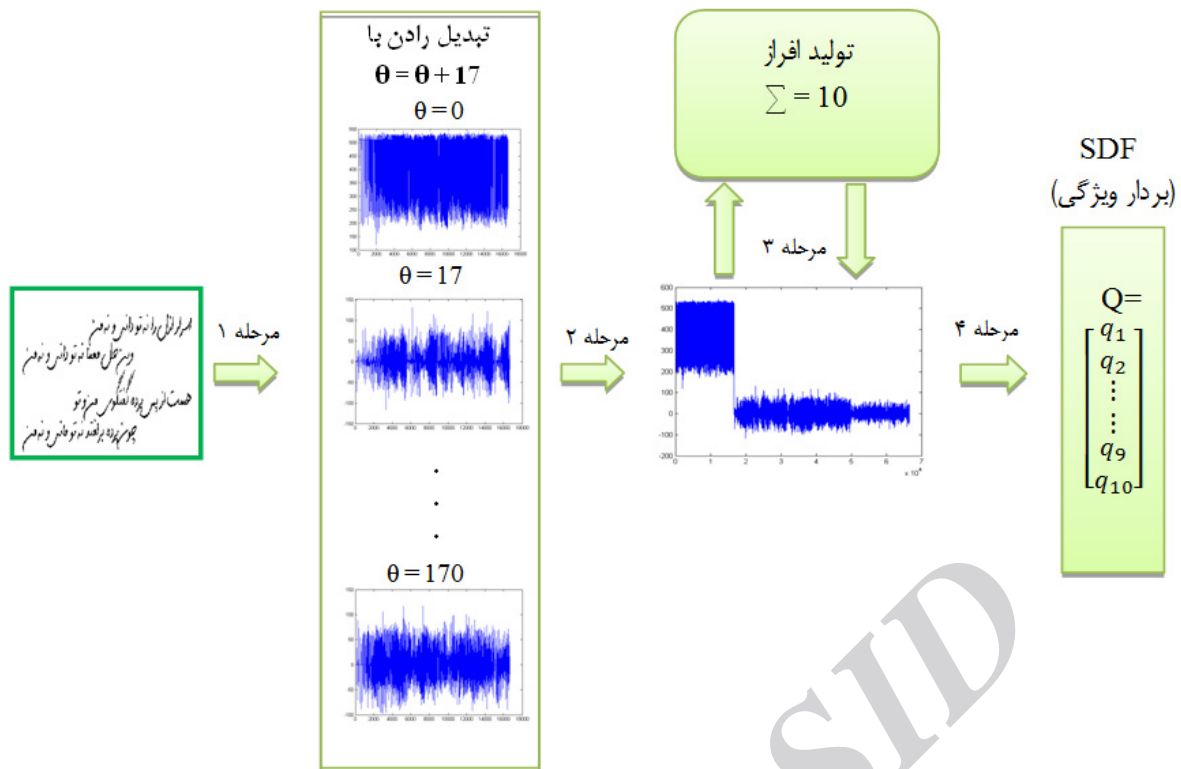
که $N(q_l, q_k)$ مقادیر کلی رویدادها می باشد هنگامی که q_k در مجاورت q_l در مسیر حرکت مثبت و به سمت جلوی دنباله نمادین قرار داشته باشد. رابطه فوق احتمال انتقال از وضعیت q_l به q_k است اگر q_k در مجاورت q_l ، در مسیری مثبت و به سمت جلو قرار گرفته باشد. ماتریس احتمالات انتقال $P(q_k | q_l) \forall q_k, q_l \in Q$ با استفاده از (۳) و به صورت زیر محاسبه می شود

$$\Pi = \begin{bmatrix} P(q_1 | q_1) & \dots & P(q_{|Q|} | q_1) \\ \vdots & \ddots & \vdots \\ P(q_1 | q_{|Q|}) & \dots & P(q_{|Q|} | q_{|Q|}) \end{bmatrix} \quad (4)$$

با انتخاب الگوریتم افراز مناسب می توان مطمئن بود که شرایط مورد نظر برآورده شده است. برای یک سیگنال مشخص (در اینجا خروجی تبدیل رادن)، بردار ویژه ρ مربوط به مقادیر ویژه ماتریس Π ساخته می شود. سپس بردار ρ که بردار احتمال وضعیت ثابت می باشد برای استخراج ویژگی از فرم موج، همان طور که در شکل ۳ نشان داده شده است به کار گرفته می شود. در نهایت، بردار ویژگی ρ در طبقه بند مورد استفاده قرار می گیرد [۲۰] و [۲۲]. در پایان به این نکته باید توجه شود که بر عکس مدل مخفی مارکوف که از یک ساختار نامعلوم قبلی استفاده می کند که در نتیجه به منظور به دست آوردن الگوی بهینه از داده های آموزشی نیز استفاده می کند، مدل D -مارکوف استفاده شده در این مقاله از یک ساختار معلوم که همان دنباله نمادهای به دست آمده از افراز مجموعه داده می باشد استفاده می کند.

وضعیت ها که باعث تولید کلمات (زیرمجموعه ای از دنباله نمادین) با طول مختلف با توجه به اهمیت نسبی آنها می شود، روی الفبای محدود با کاردینالیته $|\Sigma|$ (تمامی وضعیت های ممکن کمتر یا مساوی $|\Sigma|^D$ می باشد. با افزایش D ، حافظه مورد نیاز برای نگهداری حالات PFSA نیز افزایش می یابد. به هر حال مزیت داشتن حافظه اضافی با افزایش بار محاسباتی همراه خواهد بود. وضعیت های PFSA ترکیبات مختلف کلمات را روی دنباله نماد نشان می دهند. همان طور که در شکل ۳ نشان داده شده است در گراف مربوط به PFSA، لبه های جهت دار^۱ یک وضعیت را به وضعیت های دیگر متصل می کنند که احتمال انتقال بین وضعیت ها را نشان می دهد. بنابراین وضعیت ها تمام کلمات با طول خاص را نشان می دهند. از آنجایی که با افزایش D محاسبات انجام شده برای ساختن ماتریس انتقال حالت به طور چشم گیری افزایش می یابد. همچنین این افزایش محاسبات به افزایش حافظه مصرفی جهت اجرای الگوریتم و افزایش زمان منجر می شود. به علاوه در بررسی های اولیه این الگوریتم با بالابردن این پارامتر و در نظر گرفتن زمان اجرای الگوریتم بهبود قابل قبولی در نتایج مشاهده نشده و به دلیل این که نسبت هزینه دقت به زمان با افزایش این پارامتر افزایش می یابد در این مقاله $D=1$ در نظر گرفته شده است. در نتیجه تعداد وضعیت ها در PFSA مساوی تعداد کاردینالیته نمادها $|\Sigma| = |Q|$ می باشد یعنی تعداد کل حالات ممکن به صورت $Q = \{q_1, \dots, q_{|\Sigma|}\}$ خواهد بود. احتمال انتقال بین وضعیت ها در سراسر دنباله نمادین به صورت زیر محاسبه می شود

1. Directional Edge



شکل ۴: مراحل کلی استخراج ویژگی روش پیشنهادی.

جدول ۱: توزیع کلی بانک اطلاعاتی MSHD.

۳۶ + ۴۸ = ۸۴	تعداد نویسندگان (زن و مرد)
۴۲ + ۴۲ = ۸۴	تعداد نویسندگان (مجموعه آموزش و آزمایش)
۶ + ۶ = ۱۲	تعداد متون برای هر نویسنده به زبان (عربی و فرانسه)
۱۲ × ۸۴ = ۱۰۰۸	مجموع نمونه‌ها

شده است.

در ادامه روش پیشنهادی این مقاله توضیح داده شده است.

۴- روش پیشنهادی و پیاده‌سازی

به طور کلی روش پیشنهادی در سه مرحله پیش‌پردازش، استخراج ویژگی و طبقه‌بندی در ادامه بیان شده است.

۴-۱ پیش‌پردازش

از تکنیک‌های پردازش تصویر به منظور آماده‌سازی تصویر برای مراحل بعد استفاده می‌شود. ابتدا تصاویر به وسیله الگوریتم Otsu [۲۴] به سیاه و سفید تبدیل می‌شوند و بعد از آن الگوریتم حذف نویز توضیح داده شده در [۲۵] اعمال می‌گردد.

۴-۲ استخراج ویژگی

استخراج ویژگی مرحله به مرحله در ادامه بیان می‌شود. همچنین این مراحل در شکل ۴ نشان داده شده است.

مرحله اول: تصویر دست‌خط با استفاده از تبدیل رادن دوبعدی به یک فرم موج تبدیل می‌شود. در تبدیل رادن به منظور نشان‌دادن جزئیات بیشتر دست‌خط، زوایای مختلفی را می‌توان در نظر گرفت. بهترین نتایج گرفته‌شده در این مقاله با فاصله زاویه‌ای ۱۷ درجه به دست آمده است به گونه‌ای که تصویر دست‌خط ابتدا با زاویه صفر درجه در نظر گرفته شده و در هر تکرار، ۱۷ درجه به آن اضافه شده است $(\theta = \theta + 17)$. این عمل ۱۱ مرتبه تکرار گردیده و ۱۱ فرم موج تولید می‌شود.

مرحله دوم: تمامی تبدیلات رادن ذکر شده در بالا به شکل یک بردار در کنار هم جمع می‌شوند.

۳- بانک اطلاعاتی

از آنجایی که در زبان فارسی بانک اطلاعاتی به منظور تشخیص جنسیت تا کنون ارائه نشده است، آزمایش‌ها بر روی بانک اطلاعاتی MSHD [۲۳] صورت گرفته است. ۸۴ نویسنده (زن و مرد) در تهیه این بانک شرکت داشته‌اند و نمونه دست‌خط هر نویسنده بر روی ۱۲ صفحه گرفته شده است (۶ صفحه اول به زبان فرانسه و ۶ صفحه بعدی به زبان عربی). هر ۱۲ متن در زبان‌های فرانسه و عربی، متن‌هایی با محتوای متفاوت می‌باشند به گونه‌ای که هر صفحه شامل متن متفاوت نسبت به صفحه دیگر در این دو زبان است. مراحل ذکر شده برای تمامی نویسندگان بانک اطلاعاتی در نظر گرفته شده است. همچنین این بانک اطلاعاتی به منظور پیاده‌سازی دو چالش جدید استقلال^۲ و وابستگی متن^۳ و زبان‌های نوشتاری مستقل^۴ و وابسته^۵ طراحی شده است. چالش مستقل بودن متن به معنای عدم مشاهده دست‌خط‌های مشابه در دو مجموعه آموزشی و آزمایشی است. همچنین چالش مستقل از زبان نوشتاری بدین معنی می‌باشد که زبان نوشتاری در دو مجموعه آموزش و آزمایش متفاوت باشند. در جدول ۱ توزیع کلی این بانک اطلاعاتی آورده

1. Multiscript Handwriting Database
2. Text Independent
3. Text Dependent
4. Script Independent
5. Script Dependent

6. Concatenating

جدول ۲: نتیجه آزمایش بر روی تمام داده‌های موجود در پایگاه داده MSHD.

روش استخراج ویژگی	اندازه بردار ویژگی	درصد تشخیص (شبکه عصبی)
کجی و خمیدگی	۸۸	۶۹٫۲۵
بافت	۲۶۷	۶۴٫۸۸
فرکتال	۶	۶۱٫۹۰
روش پیشنهادی	۱۰	۷۴٫۲۰

جدول ۳: ارزیابی متون مستقل روی پایگاه داده MSHD.

زبان نوشتاری	روش استخراج ویژگی	درصد تشخیص (شبکه عصبی)
کجی و خمیدگی		۶۶٫۲۷
فرانسه	بافت	۶۵٫۴۸
	فرکتال	۶۵٫۸۷
	کجی و خمیدگی	۶۹٫۴۴
عربی	بافت	۷۱٫۴۳
	فرکتال	۶۵٫۰۸
فرانسه	روش پیشنهادی	۸۴٫۱۰
عربی	روش پیشنهادی	۸۴٫۹۰

جدول ۴: ارزیابی متون وابسته روی پایگاه داده MSHD.

زبان نوشتاری	روش استخراج ویژگی	درصد تشخیص (شبکه عصبی)
کجی و خمیدگی		۶۷٫۰۶
فرانسه	بافت	۶۶٫۲۷
	فرکتال	۶۶٫۲۷
	کجی و خمیدگی	۷۲٫۶۲
عربی	بافت	۷۲٫۲۲
	فرکتال	۶۵٫۸۷
فرانسه	روش پیشنهادی	۷۳
عربی	روش پیشنهادی	۸۰٫۷۵

در جدول ۲ مقایسه‌ای از نتایج روش پیشنهادی ما و ویژگی‌های پرکاربردی از قبیل کجی و خمیدگی در متن، بافت و فرکتال‌ها که در [۸] مورد آزمایش قرار گرفته‌اند، صورت پذیرفته است. همان طور که ملاحظه می‌شود، نرخ طبقه‌بندی در روش پیشنهادی، ۷۴٫۲ درصد می‌باشد که نسبت به [۸] بهبود حاصل شده است. در ادامه دو چالش مهم (وابستگی و استقلال متون و زبان نوشتاری) که در این حوزه بسیار مورد توجه قرار گرفته است بررسی می‌گردند.

۱-۵ ارزیابی متون مستقل و وابسته

این آزمایش‌ها با هدف مطالعه چگونگی عملکرد ویژگی‌های متفاوت، روی نمونه‌های موجود در مجموعه‌های آموزشی و آزمایشی شامل محتوی متنی مستقل (وابسته) است. برای ارزیابی متون مستقل، ۳ متن اول به زبان فرانسه یا عربی ۴۲ نویسنده ابتدایی، برای مجموعه آموزشی و ۳ متن آخر ۴۲ نویسنده دیگر به عنوان مجموعه آزمایشی در نظر گرفته می‌شوند. سپس ۳ متن آخر به زبان فرانسه یا عربی ۴۲ نویسنده ابتدایی، برای مجموعه آموزشی و ۳ متن ابتدایی ۴۲ نویسنده دیگر به عنوان مجموعه آزمایشی در نظر گرفته می‌شوند. با این عمل اطمینان حاصل می‌شود که متنی با محتوای مشابه در مجموعه آموزش و مجموعه آزمایش وجود نخواهد داشت. میانگین نتایج برای دو حالت در زبان فرانسه و دو حالت در زبان عربی در جدول ۳ همراه با مقایسه ارزیابی مقاله [۸] نشان داده شده است. برای ارزیابی متون وابسته، ۶ متن به زبان فرانسه ۴۲ نویسنده ابتدایی، برای مجموعه آموزشی و ۶ متن به زبان فرانسه ۴۲ نویسنده دیگر، برای مجموعه آزمایشی در نظر گرفته می‌شوند. سپس ۶ متن به زبان عربی ۴۲ نویسنده ابتدایی به عنوان مجموعه آموزشی و ۶ متن به زبان عربی ۴۲ نویسنده دیگر به عنوان مجموعه آزمایشی در نظر گرفته می‌شوند. نتایج همراه با مقایسه ارزیابی مقاله [۸] در جدول ۴ آمده است.

نتایج نشان می‌دهد روش ارائه‌شده در این مقاله (SDF) از کارایی بالاتری نسبت به روش‌های ارائه‌شده در [۸] برخوردار است.

۲-۵ ارزیابی زبان نوشتاری مستقل و وابسته

در آزمایش‌های زبان نوشتاری مستقل، مجموعه داده‌های آموزشی و مجموعه داده‌های آزمایشی شامل نمونه‌هایی با زبان نوشتاری متفاوت می‌باشند. آزمایش‌های زبان نوشتاری مستقل بسیار چالش برانگیزتر از آزمایش‌های زبان نوشتاری وابسته می‌باشد و نمونه‌ها و زبان‌های مورد استفاده در مجموعه‌های آموزشی با نمونه‌ها و زبان‌های مورد استفاده در مجموعه‌های آزمایشی متفاوت می‌باشند. برای انجام آزمایش‌های زبان نوشتاری مستقل روی بانک اطلاعاتی MSHD از ۶ نمونه زبان فرانسوی ۴۲ نویسنده ابتدایی در مجموعه‌های آموزشی استفاده می‌شود در حالی که

مرحله سوم: داده‌های آموزشی و آزمایشی که از مرحله قبل به دست آمده‌اند توسط الگوریتم افراز سیگنال که در اینجا الگوریتم ماکسیمم آنتروپی می‌باشد افراز می‌شوند. بهترین اندازه جهت افراز در آزمایش‌ها $Q=10$ است.

مرحله چهارم: استخراج بردار ویژگی از فرم‌های موج افرازشده از مرحله قبل با استفاده از روش SDF که در بخش دوم توضیح داده شد صورت می‌پذیرد.

۳-۴ طبقه‌بندی

در این مرحله بردارهای ویژگی به دو دسته آموزشی و آزمایشی تقسیم می‌شوند. داده‌های آموزشی به منظور ایجاد یک مدل در الگوریتم‌های یادگیری ماشین استفاده می‌شوند. سپس به کمک این مدل، الگوریتم‌های طبقه‌بندی، داده‌های آزمایشی را به طبقه‌های مورد نظر اختصاص می‌دهند. در این مقاله از طبقه‌بند پرکاربرد شبکه‌های عصبی پرسپترون چندلایه^۱ که دارای سه لایه شامل لایه‌های ورودی، مخفی و خروجی می‌باشد استفاده شده است.

نورون‌های لایه ورودی برابر ۱۰ (به اندازه بردار ویژگی)، نورون‌های لایه مخفی برابر با ۴ و تعداد نورون‌های لایه خروجی برابر با تعداد طبقه‌ها که در این مقاله ۲ می‌باشد (زن و مرد) در نظر گرفته شده است.

۵- نتایج آزمایش‌ها

جهت انجام آزمایش‌ها در ابتدا، نرخ تشخیص جنسیت را بر روی تمام مجموعه داده‌ها با استفاده از طبقه‌بند شبکه‌های پرسپترون چندلایه ارائه نموده، سپس ارزیابی متون مستقل و وابسته، زبان نوشتاری مستقل و وابسته و آزمایش تأثیر میزان داده در مجموعه آموزش مورد آزمایش قرار می‌گیرند. در چالش‌های مورد آزمایش از این موضوع که هیچ دست‌خط نویسنده‌ای در هر دو مجموعه آموزش و آزمایش قرار نداشته باشد، اطمینان حاصل نموده‌ایم. جدول ۲ نتیجه آزمایش را برای پایگاه داده MSHD که در بخش ۳ توضیح داده شد، مشتمل بر نمونه‌های ۴۲ نویسنده در هر یک از مجموعه‌های آزمایشی یا آموزشی را نشان می‌دهد.

1. Multi Layer Perceptron

جدول ۵: ارزیابی زبان نوشتاری مستقل روی پایگاه داده MSHD.

داده‌های آموزشی	داده‌های آزمایشی	روش استخراج ویژگی	درصد تشخیص (شبکه‌های عصبی)
عربی	عربی	کجی و خمیدگی	۶۹,۴۴
عربی	عربی	بافت	۶۸,۶۵
عربی	عربی	فرکتال	۶۹,۸۴
عربی	عربی	روش پیشنهادی	۷۴,۶۰
عربی	عربی	کجی و خمیدگی	۶۱,۹۰
عربی	عربی	بافت	۶۰,۷۱
عربی	عربی	فرکتال	۶۱,۱۱
عربی	عربی	روش پیشنهادی	۷۰,۲۰

جدول ۶: ارزیابی زبان نوشتاری وابسته روی پایگاه داده MSHD.

داده‌های آموزشی	داده‌های آزمایشی	روش استخراج ویژگی	درصد تشخیص (شبکه‌های عصبی)
عربی	عربی	کجی و خمیدگی	۶۹,۴۴
عربی	عربی	بافت	۶۸,۲۵
عربی	عربی	فرکتال	۶۱,۵۱
عربی	عربی	روش پیشنهادی	۷۳
عربی	عربی	کجی و خمیدگی	۷۳,۴۱
عربی	عربی	بافت	۷۱,۴۱
عربی	عربی	فرکتال	۶۲,۳۰
عربی	عربی	روش پیشنهادی	۷۶,۶۰

به منظور بررسی هزینه‌های محاسباتی الگوریتم پیشنهادی (زمان اجرا)، زمان اجرای الگوریتم پیشنهادی با بهترین الگوریتم [۸] که همان کجی و خمیدگی می‌باشد مقایسه شده است. این آزمایش بر روی یک سیستم با پردازنده ۳ هسته‌ای و حافظه موقت ۴ مگابایت انجام شده است. زمان اجرای الگوریتم‌ها برای ایجاد بردار ویژگی از یک نمونه آزمایشی (تصویر با فرمت tif) با اندازه ۱۲۳۹×۲۵۵۰ پیکسل محاسبه شده است. الگوریتم ایجاد بردار ویژگی بر اساس کجی و خمیدگی $۵/۱$ ثانیه و در الگوریتم پیشنهادی $۶/۲$ ثانیه زمان سپری شد. با وجود اختلاف $۱/۱$ ثانیه‌ای این دو الگوریتم به دلیل این که در این مسایل نسبت هزینه دقت به زمان از نسبت هزینه زمان به دقت از اهمیت بالاتری برخوردار می‌باشد، الگوریتم پیشنهادی با در نظر گرفتن اختلاف زمانی ناچیز می‌تواند به عنوان یک انتخاب مناسب در نظر گرفته شود.

در پایان، بررسی تأثیر تعداد نمونه‌های موجود در مجموعه‌های آزمایشی و آموزشی جهت اندازه‌گیری دقت تشخیص جنسیت بر روی نمونه‌ها، آزمایش‌هایی صورت پذیرفت. در ابتدا ۹۰% تعداد کل نمونه‌ها را در مجموعه آموزش و ۱۰% تعداد کل نمونه‌ها را در مجموعه آزمایش قرار می‌دهیم. سپس نمونه‌های موجود در مجموعه آموزشی را ۸۰% تعداد کل نمونه‌ها و نمونه‌های موجود در مجموعه آزمایش را ۲۰% تعداد کل نمونه‌ها در نظر می‌گیریم. همین روند را ادامه می‌دهیم تا ۵۰% تعداد کل نمونه‌ها در مجموعه آموزشی و ۵۰% تعداد کل نمونه‌ها در مجموعه آزمایشی قرار گیرند. نمودار نتایج این آزمایش‌ها در شکل ۵ نشان داده شده‌اند. نتایج نشان می‌دهد با افزایش تعداد داده‌ها در مجموعه آموزشی میزان درصد تشخیص به صورت چشمگیری افزایش یافته است.

۶- نتیجه‌گیری و کارهای آینده

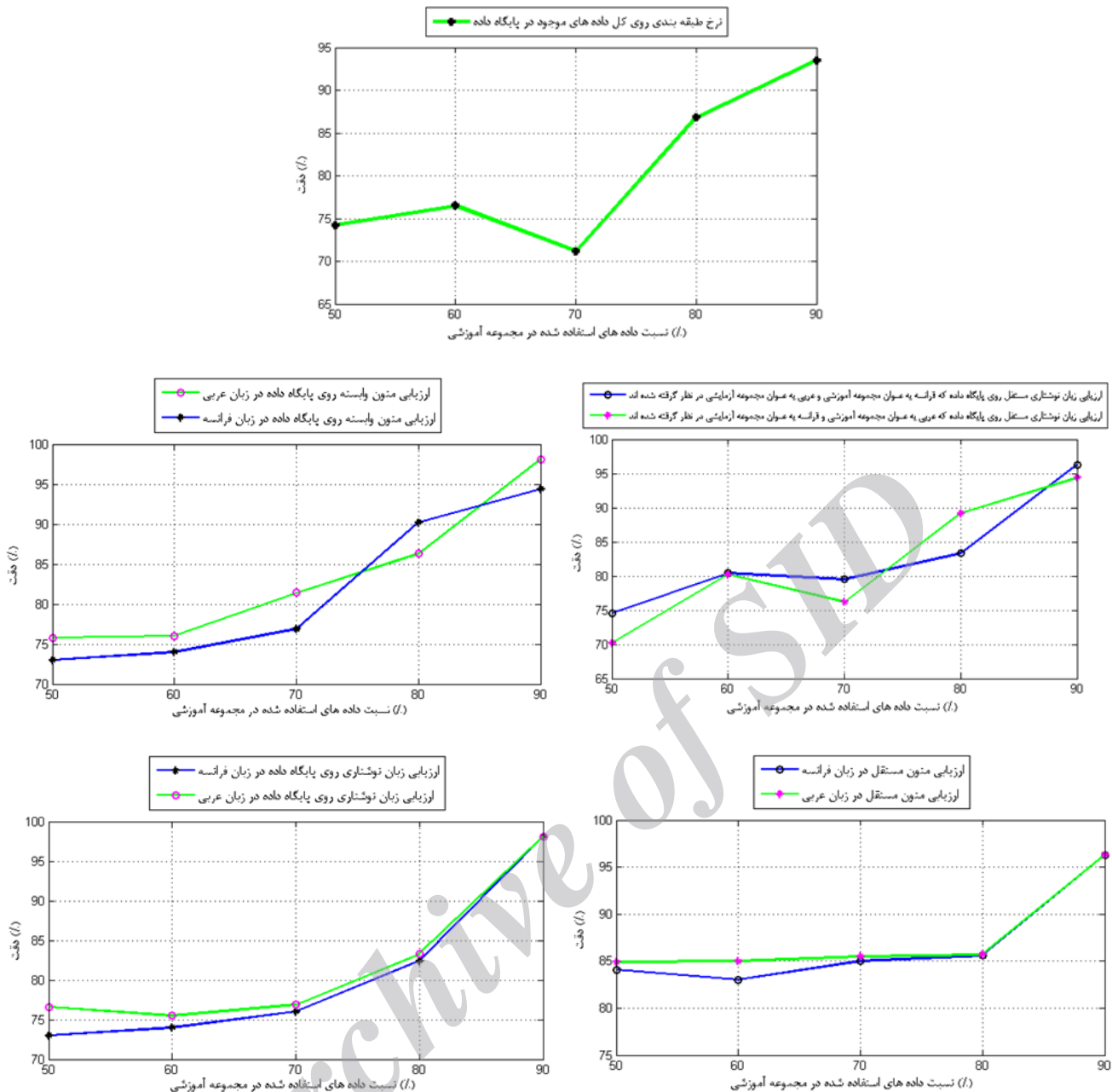
در این مقاله، تشخیص جنسیت از روی تصاویر اسکن شده دست‌خط بررسی گردید. اگرچه این موضوع یک محیط تحقیقی عمومی در مطالعات روانشناسان در دهه‌های متعددی است اما این مسأله کمتر در تحقیقات کامپیوتری مورد مطالعه قرار گرفته است. در این تحقیق، روش جدیدی جهت استخراج ویژگی به نام پالایش پویای نمادین (SDF) معرفی گردید. SDF، یک روش استخراج ویژگی برای طبقه‌بندی الگو در سیستم‌های دینامیکی پیچیده را ارائه می‌دهد و همچنین یک روش غیر خطی می‌باشد که اطلاعات مناسب را به عنوان بردار ویژگی استخراج می‌کند. این بردارهای ویژگی با استفاده از وضعیت ماشین‌های خودکار منتهای احتمالی که موجب تولید الگوهای آماری با ابعاد کم می‌شود، مدل شده است. برای ارزیابی روش پیشنهادی از نمونه‌های موجود در پایگاه داده MSHD استفاده نموده‌ایم. در چالش‌های مستقل و وابسته از متن و

از ۶ نمونه زبان عربی ۴۲ نویسنده دیگر، در مجموعه‌های آزمایشی استفاده می‌گردد. در روشی مشابه از ۶ نمونه زبان عربی ۴۲ نویسنده ابتدایی در مجموعه‌های آموزشی استفاده می‌شود و ۶ نمونه زبان فرانسوی ۴۲ نویسنده دیگر در مجموعه‌های آزمایشی مورد استفاده قرار می‌گیرند. جدول ۵ نتایج دقت به دست آمده روش پیشنهادی را در مقایسه با شیوه‌های دیگر استخراج ویژگی شامل کجی و خمیدگی در متن، بافت و فرکتال‌ها که در [۸] استفاده شده است نشان می‌دهد. نتایج نشان می‌دهد روش ارائه‌شده در این مقاله، کارایی بالاتری نسبت به کارهای قبلی گزارش شده دارد.

در آزمایش‌های زبان نوشتاری وابسته، نمونه‌های نوشتاری با رسم‌الخط مشابه (فرانسوی-عربی) هم در مجموعه‌های آموزشی و هم در آزمایشی استفاده می‌شوند. در بانک اطلاعاتی MSHD برای ارزیابی زبان نوشتاری وابسته در زبان فرانسوی، ۶ نمونه اول ۴۲ نویسنده ابتدایی در مجموعه‌های آموزشی و ۶ نمونه اول ۴۲ نویسنده دیگر به عنوان مجموعه‌های آزمایشی در نظر گرفته می‌شوند. برای ارزیابی زبان نوشتاری وابسته در زبان عربی، ۶ نمونه دوم ۴۲ نویسنده ابتدایی در مجموعه‌های آموزشی و ۶ نمونه دوم ۴۲ نویسنده دیگر به عنوان مجموعه‌های آزمایشی در نظر گرفته می‌شوند. جدول ۶ نتایج دقت به دست آمده روش پیشنهادی را در مقایسه با شیوه‌های دیگر که در [۸] استفاده شده است نشان می‌دهد. نتایج نشان می‌دهد روش ارائه‌شده در این مقاله نیز کارایی و دقت بالاتری به دست آورده است.

۳-۵ بررسی کارایی روش پیشنهادی

ابتدا ویژگی‌های روش پیشنهادی و سپس هزینه محاسباتی این روش ارزیابی می‌شود. در پایان نیز تأثیر تعداد نمونه‌های موجود در مجموعه آموزش و آزمایش با استفاده از روش پیشنهادی بررسی می‌شود. از آنجایی که روش پالایش پویای نمادین مبتنی بر تحلیل فرم موج می‌باشد و این روش ویژگی‌های متمایزکننده‌ای در سایر کاربردهای بازشناسی الگو از قبیل تحلیل سری‌های زمانی گرفته‌شده از رفتار حرکت ربات‌ها و بررسی صدمات سطحی در نمونه‌های آلیاژهای چندبلوری [۲۰] و [۲۱] از فرم موج به دست آورده است، در این مقاله با تبدیل دست‌خط به یک فرم موج با استفاده از تبدیل رادن به تحلیل این فرم موج با استفاده از پالایش پویای نمادین پرداخته شد. همان طور که در نتایج آزمایش‌ها مشاهده شده است در مسئله تشخیص جنسیت نیز ویژگی‌های متمایزکننده‌ای نسبت به روش‌های قبلی که بر روی این بانک اطلاعاتی گزارش شده‌اند استخراج شده است.



شکل ۵: نتایج تأثیر تعداد نمونه‌ها در مجموعه‌های آموزشی و آزمایشی.

مراجع

- [1] F. Adrian and G. Barrie, "Graphology and personality: another failure to validate graphological analysis," *Personality and Individual Differences*, vol. 8, no. 3, pp. 433-435, 1987.
- [2] R. J. Klimoshk and A. Rafaeli, "Inferring personal qualities through handwriting analysis," *J. of Occupational and Organizational Psychology*, vol. 56, no. 3, pp. 191-202, 1987.
- [3] E. Neter and G. Ben-Shakhar, "The predictive validity of graphological inferences: a meta-analytic approach," *Personality and Individual Differences*, vol. 10, no. 12, pp. 735-745, 1989.
- [4] R. P. Tett and C. A. Palmer, "The validity of handwriting elements in relation to self-report personality trait measures," *Personality and Individual Differences*, vol. 22, no. 1, pp. 11-18, 1997.
- [5] S. Hamid and K. M. Loewenthal, "Inferring gender from handwriting in Urdu and English," *The J. of Social Psychology*, vol. 136, no. 6, pp. 778-782, 1996.
- [6] H. James, "Sex differences in handwriting: a comment on spear," *British Educational Research J.*, vol. 17, no. 2, pp. 141-145, 1991.
- [7] W. N. Hayes, *Identifying Sex from Handwriting*, Percept Motor Skills, pp. 91-800, 1996.
- [8] I. Siddiqi, C. Djeddi, A. Raza, and L. Souici-Meslati, "Automatic analysis of handwriting for gender classification," *Pattern Analysis and Applications*, vol. 18, no. 8, pp. 887-899, 2014.

زبان نوشتاری مستقل و وابسته جهت استخراج ویژگی از SDF استفاده کرده‌ایم. از طبقه‌بند شبکه‌های پرسپترون چندلایه، جهت تشخیص جنسیت از روی دست‌خط بهره برده‌ایم. نتایج به دست آمده با استفاده از روش پیشنهادی را با نتایج موجود در [۸] مقایسه نموده و نتایج به مراتب بهتری حاصل گردیده است.

با وجود این که مسئله تشخیص جنسیت از روی دست‌خط یک مسئله دوطبقه‌ای می‌باشد با این وجود تغییرات بالای درون طبقه‌ای در این مسئله باعث شده است که تشخیص جنسیت خود به تنهایی یک موضوع مهم در حوزه تحلیل اسناد دست‌نویس باشد اما در کارهای آینده سایر اطلاعات مربوط به افراد جامعه (Demographic) از قبیل سن، دست‌چپی/ راستی و سایر موارد بررسی خواهد شد. همچنین تأثیر این اطلاعات (جنسیت، سن و دست‌چپی/ راستی) در مسئله شناسایی نویسنده از روی دست‌خط مورد بررسی قرار خواهد گرفت.

- [24] N. Otsu, "A threshold selection method from gray-level," *IEEE Trans. on Systems*, vol. 9, no. 1, pp. 62-66, Jan. 1979.
- [25] M. Cheriet, N. Khama, C. L. Liu, and C. Suen, *Character Recognition Systems: A Guide for Students and Practitioners Anomaly Detection*, John Wiley and Sons Inc, 2007.
- کاظم نوری هفت‌چشمه** تحصیلات خود را در مقاطع کارشناسی، کارشناسی ارشد و دکتری ریاضی کاربردی به‌ترتیب در سال‌های ۱۳۸۳، ۱۳۸۶ و ۱۳۸۹ از دانشگاه علم و صنعت ایران به پایان رسانده است و هم‌اکنون دانشیار دانشکده ریاضی، آمار و علوم کامپیوتر دانشگاه سمنان می‌باشد. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: حل عددی معادلات دیفرانسیل معمولی و تصادفی، معادلات انتگرال، معادلات کسری، موجک‌ها و پایه‌های متعامد، ریاضیات مالی.
- رضا خدادادی** در سال ۱۳۸۶ مدرک کارشناسی مهندسی کامپیوتر گرایش نرم‌افزار خود را از دانشگاه پیام نور واحد بیرجند و در سال ۱۳۹۴ مدرک کارشناسی ارشد مهندسی کامپیوتر گرایش نرم‌افزار خود را از دانشگاه آزاد واحد مشهد دریافت نمود. از سال ۱۳۹۱ نام‌برده به عنوان کارشناس تحلیل‌گر سیستم در اداره کل راه و شهرسازی خراسان جنوبی مشغول به خدمت می‌باشد. زمینه‌های علمی مورد علاقه نام‌برده شامل موضوعاتی مانند پردازش تصویر، مدیریت شبکه‌های کامپیوتری، می‌باشد.
- یونس اکبری** در سال ۱۳۸۴ مدرک کارشناسی خود را در رشته مهندسی کامپیوتر اخذ نموده است و در سال ۱۳۹۰ در مقطع کارشناسی ارشد مدیریت فناوری اطلاعات از دانشگاه پیام‌نور واحد تهران دانش آموخته شده است. وی در حال حاضر دانشجوی دکتری ریاضی کاربردی (آنالیز عددی) دانشگاه سمنان می‌باشد. زمینه‌های تحقیقاتی مورد علاقه وی بازشناسی الگو، پردازش تصویر و سیستم‌های هوشمند است.
- سیدمحمد رضوی** در سال ۱۳۷۳ مدرک کارشناسی خود را در رشته مهندسی برق- الکترونیک دانشگاه صنعتی امیر کبیر (واحد تفرش) اخذ نموده است و در سال‌های ۱۳۷۵ و ۱۳۸۵ در مقاطع کارشناسی ارشد و دکتری در دانشگاه تربیت مدرس دانش آموخته شده است. وی از سال ۱۳۷۶ به عنوان عضو هیأت علمی در گروه الکترونیک و مخابرات دانشگاه بیرجند مشغول شده و اینک دانشیار گروه مهندسی الکترونیک دانشگاه بیرجند است. زمینه‌های تحقیقاتی مورد علاقه وی بازشناسی الگو، پردازش تصویر و سیستم‌های هوشمند است.
- حسین احمدی ترشیزی** شرح حال علمی و اجرایی ایشان در زمان انتشار مقاله در دسترس نبود.
- [9] V. Burr, "Judging gender from sampels of adult handwriting: accuracy and use of cues," *the J. of Social Psychology*, vol. 142, no. 6, pp. 691-700, 2002.
- [10] A. Chaabouni, H. Boubaker, M. Kherallah, and A. Alimi, "Fractal and multi-fractal for Arabic offline writer identification," in *Proc. of Int. Conf. on Pattern Recognition, ICPR'10*, pp. 3793-3796, 23-26 Aug. 2010.
- [11] M. Liwicki, A. Schlappbach, and H. Bunke, "Automatic gender detection using on-line and off-line information," *Pattern Analysis and Applications*, vol. 14, no. 1, pp. 87-92, 2011.
- [12] M. Blumenstein, X. Liu, and B. Verma, "An investigation of the modified direction feature for cursive character recognition," *Pattern Recognition*, vol. 40, no. 2, pp. 376-388, 2007.
- [13] M. Blumenstein, B. Verma, and H. Baseli, "A novel feature extraction technique for the recognition of segmented handwritten characters," in *Proc. of Int. Conf. on Document Analysis and Recognition, ICDAR'03*, pp. 137-141, 6 Aug. 2003.
- [14] N. Bouadjeneq, H. Nemmour, and Y. Chibani, "Age, gender and handedness prediction from handwriting using gradient features," in *Proc. of 13th Int. Conf. on Document Analysis and Recognition, ICDAR'15*, pp. 1116-1120, 23-26 Aug. 2015.
- [15] E. Sokic, A. Salihbegovic, and M. Ahic-Djokic, "Analysis of off-line handwritten text samples of different gender using shape descriptors," in *Proc. of Int. Symp. on Telecommunications*, pp. 114-120, 25-27 Oct. 2012.
- [16] S. Al Maadeed and A. Hassaine, "Automatic prediction of age, gender, and nationality in offline handwriting," *EURASIP J. on Image and Video Processing*, vol. 10, no. 1, pp. 19-29, 2015.
- [17] R. Hussain, A. Reza, I. Siddiqi, K. Khurshid, and C. Djeddi, "A comprehensive survey of handwritten document benchmarks: structure, usage and evaluation," *EURASIP J. on Image and Video Processing*, vol. 56, no. 6, pp. 1-24, Dec. 2015.
- [18] R. N. Bracewell, *Two-Dimensional Imaging*, Prentice Hall, 1995.
- [19] J. S. Lim, *Two-Dimensional Signal and Image Processing*, Prentice Hall, 1990.
- [20] S. Bahrapour, A. Ray, S. Sarkar, T. Damarla, and N. Nasrabadi, "Performance comparison of feature extraction algorithms for target detection and classification," *Pattern Recognition Letter*, vol. 34, no. 16, pp. 2126-2134, 1 Dec. 2013.
- [21] A. Ray, "Symbolic dynamic analysis of complex systems for anomaly detection," *Signal Process*, vol. 84, no. 7, pp. 1115-1130, Jul. 2004.
- [22] A. Subbu, A. Srivastav, A. Ray, and E. Keller, "Symbolic dynamic filtering for image analysis: theory and experimental validation," *Signal Image and Video Processing*, vol. 4, no. 3, pp. 319-329, Sept. 2010.
- [23] C. Djeddi, et al., "LAMIS-MSHD: a multi-script offline handwriting database," in *Proc. of the 14th Int. Conf. on Frontiers in Handwriting Recognition, ICFHR'14*, pp. 93-97, 2014.