تشخیص آفلاین کاراکتر عربی :

آخرین پیشرفت

چکیده

شبیه سازی ماشین خواندن انسان، برای تقریبا سه دهه موضوع تحقیقات فشرده ای بوده است. تعداد زیادی از مقالات و گزارشات پژوهشی در مورد کاراکترهای لاتین، چینی و ژاپنی منتشر شده است. با این حال، کار کمی در مورد تشخیص خودکار کاراکترهای عربی به دلیل پیچیدگی متن چاپ شده و دست نوشته انجام شده است و این مشکل هنوز هم یک میدان تحقیقاتی باز است. هدف اصلی این مقاله، ارائه حالت عربی پژوهش تشخیص کاراکتر در سراسر دو دهه گذشته است.

کلمات کلیدی: تقسیم بندی، تشخیص دست خط، تشخیص آفلاین، کاراکترهای عربی، مدل های مخفی Markov، طبقه بندهای شبکه عصبی، استخراج ویژگی، تشخیص کاراکتر نوری

1. مقدمه

سیستم های تشخیص کاراکتر می توانند به طور فوق العاده به پیشرفت فرآیند اتوماسیون کمک نمایند و می توانند تعامل بین انسان و ماشین را در بسیاری از کاربردها، از جمله اتوماسیون اداری، تأیید بررسی و انواع زیادی از بانکداری، کاربردهای کسب و کار و ورود اطلاعات بهبود ببخشند.

روش های مختلف پوشش داده شده تحت اصطلاح کلی تشخیص کاراکتر در دو دسته آنلاین یا آفلاین قرار می گیرند که هر یک دارای الگوریتم های تشخیص و سخت افزار خاص خود هستند

در سیستم های تشخیص کاراکتر آنلاین، کامپیوتر، نمادها را زمانی که ترسیم می شوند، تشخیص می دهد. (1–4) رایج ترین سطح نوشتن، تبلت دیجیتالی است که از طریق یک قلم خاص در تماس با سطح تبلت عمل می کند و مختصات نقاط ترسیم شده را در یک فرکانس ثابت ساطع می کند. قطع تماس، انتقال یک کاراکتر خاص را بی درنگ برانگیخته می کند. بنابراین، ثبت بر روی تبلت، موجب تولید رشته های مختصات از هم جدا شده توسط علائم می شود، زمانی که قلم دیگر سطح تبلت را لمس نمی کند.

تشخیص آنلاین دارای چند ویژگی جالب است. اول، تشخیص به جای تصاویر دو بعدی، همانند مورد تشخیص آفلاین، بر روی داده های یک بعدی انجام می شود. خط نوشتن با دنباله ای از نقاط نمایش داده می شود که محل آن، یک تابع از زمان است. این کار دارای نتایج متعدد مهمی است:

- نظم نوشتن در دسترس است و می توان آن را توسط فرایند شناسایی استفاده نمود.
- خط نوشتن، هیچ عرضی ندارد.
- اطلاعات زمانی، مانند سرعت نیز می تواند در نظر گرفته شود.
- علاوه بر این، بلند کردن های قلم می تواند در فرآیند تشخیص مفید باشد.

در میان سیستم های آنلاین که کاراکترهای جدا شده عربی را تشخیص می دهد، چندین روش را می توان در مراجع یافت (5) - (11). (Amin) (12) سه روش را برای تشخیص آنلاین کلمات خط شکسته عربی دست نوشته معرفی نموده است. اولی یک روش ساختاری (13) بر اساس بخش بندی کلمه به کاراکترها است. سپس کاراکترها با استفاده از روش شبیه به روش کاراکترهای جدا شده تشخیص داده می شوند. (15) تشخیص کلمه با ساخت همه کلمات ممکن، با دنبال نمودن هر مسیر در گراف هم ارزی شبکه کار می کند. نمودارهای دودویی (14) نیز برای کنار گذاشتن ترکیب فاقد شرایط لازم از حروف استفاده می شوند. روش دوم، روشی نحوی (15) بر اساس تقسیم بندی کلمات به شکل های هندسی اولیه مانند منحنی ها و تکان ها است. یک اتوماسیون، شکلهای هندسی اولیه را به یک لیست از کاراکترهای تشکیل دهنده کلمه تبدیل می نماید: (16، 17) هر کلمه با توجه به بردار از برخی از پارامترهای از پیش تعیین شده مشخص می شود: در نهایت، روش سوم از یک رویکرد کلی استفاده می کند. علاوه بر

این، به منظور ارتقای نرخ تشخیص، یک تحلیلگر نحوی و معنایی که ساختار گرامری و معنای جمله عربی را تایید می نماید، است استفاده می شود. (18)

Al-Emmami و Usher (19)، یک سیستم را برای تشخیص آنلاین کلمات دست نویس عربی ارائه نمودند. کلمات به تکان هایی مبتنی بر روش پیشنهادی توسط Belaid تقسیم می شوند. (20) در فرایند یادگیری اولیه، مشخصات تکان ها از هر کاراکتر به سیستم تغذیه می شود، در حالی که در فرایند شناسایی، پارامترهای هر تکان یافت می شوند و قوانین خاص برای انتخاب مجموعه ای از تکان ها به کار برده می شوند که به بهترین شکل با ویژگی های یکی از کاراکترهای ذخیره شده مطابقت می یابند. با این حال، چند کلمه در فرآیند یادگیری و آزمایش مورد استفاده قرار می گیرند که عملکرد سیستم را سوال برانگیز می سازد.

تشخیص آفلاین، پس از اینکه نوشتن یا چاپ کامل می شود، انجام می شود. تشخیص کاراکتر نوری، OCR، (21-29) با تشخیص کاراکترهای نوری پردازش شده به جای تشخیص کاراکترهای مغناطیسی پردازش شده سرو کار دارد. در یک سیستم معمولی OCR، کاراکترختی ورودی خوانده می شوند و توسط یک اسکنر نوری دیجیتالی می شوند. سپس هر کاراکتر موقعیت یابی و بخش بندی می شود ماتریس حاصل به یک پیش پردازنده برای صاف کردن، کاهش نویز، و نرمالسازی اندازه تغذیه می شود. تشخیص آفلاین را می توان به عنوان کلی ترین مورد در نظر گرفت: هیچ دستگاه خاصی برای نوشتن لازم نیست و تفسیر سیگنال، همانند این مورد در تشخیص انسان، مستقل از تولید سیگنال است.

در طول سه دهه گذشته، بسیاری از روش های مختلف توسط تعداد زیادی از دانشمندان به منظور تشخیص کاراکترها بررسی شده است. انواع روش ها پیشنهاد شده است و توسط محققان در نقاط مختلف جهان آزمایش شده، از جمله روش های آماری، (30-32) روش ساختاری و نحوی، (33-35)، شبکه های عصبی، (36-38) سیستم های خبره (39- 41) و یادگیری ماشین. (42-44)

بسیاری از مقالات با تشخیص کاراکترهای لاتین، چینی و ژاپنی مرتبط هستند. با این حال، اگر چه تقریبا یک سوم از یک میلیارد نفر در سراسر جهان، در چند زبان مختلف، از کاراکترهای عربی برای نوشتن استفاده می نماید،

پیشرفت تحقیقاتی کمی، به صورت آنلاین و آفلاین برای تشخیص خودکار کاراکترهای عربی به دست آمده است. این یک نتیجه از عدم حمایت کافی از نظر بودجه، و کاربردهای دیگر مانند پایگاه داده های متن عربی، واژه نامه، و غیره و البته به دلیل ماهیت خط شکسته قوانین نوشتن آن است.

اگر چه نظرسنجی های دیگر با هر دوی کاراکترهای آنلاین و آفلاین عربی بررسی شده اند، (45-47) این مقاله تلاش می کند تا به طور خلاصه همه کارهای انجام شده در دو دهه گذشته را تنها در مورد سیستم های آنلاین در تلاش برای اشاره دقیق به حوزه های دقیق که باید با آنها مقابله شود، خلاصه نماید. ادامه این مقاله به شرح زیر سازماندهی شده است: بخش 2 به بررسی برخی از ویژگی های اساسی نوشتن عربی می پردازد. بخش 3 روش های مختلف برای تقسیم بندی و استخراج ویژگی را پوشش می دهد و روش های مختلف اتخاذ شده برای تشخیص را ارائه می دهد. در نهایت، نتایج در بخش 4 ارائه شده است.

2. مشخصات کلی نوشتن عربی

مقایسه ویژگی های مختلف از متون زبان های عربی، لاتین، عبری و هندی در جدول 1 بیان شده است. عربی از راست به چپ نوشته می شود. متن عربی (با دستگاه چاپ و یا دست نوشته) به صورت خط شکسته است و حروف عربی به طور معمول به خط پایه متصل می شوند. نشان داده خواهد شد که این قابلیت اتصال در فرآیند تقسیم بندی مهم است. برخی از متون چاپی ماشین و دست نوشته، خط شکسته نیستند، اما بسیاری از متون عربی اینگونه هستند، و در نتیجه تعجب آور نیست که میزان تشخیص کاراکتر عربی کمتر از کاراکترهای ناپیوسته مانند چاپ به زبان انگلیسی است.

مشخصات	عربی	لاتین	عبری	هندی
توجیه	R-to-L	L-to-R	R-to-L	L-to-R
خط شکسته	Yes	No	No	Yes
تفکیک کننده	Yes	No	No	Yes
تعداد واج ها	2	5	11	—
اشکال حروف	1–4	2	1	1
تعداد حروف	28	26	22	40
کاراکترهای مکمل	3	—	—	—

جدول 1. مقایسه متون مختلف

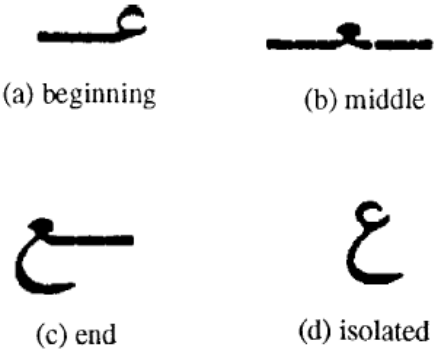
نوشتن عربی شبیه به زبان انگلیسی است که در آن از حروف (که از 29 حرف اساسی تشکیل شده است)، اعداد، علائم نقطه گذاری، و همچنین فضاها و نماد های خاص استفاده می شود. با این حال، با نشان دادن حروف صدادار از عربی با بهره گیری از نشانه گذاری های مختلف حروف، از انگلیسی متفاوت است. حضور و عدم حضور تفکیک واکه نشان دهنده معانی متفاوت است که در غیر این صورت همان کلمه را نشان می دهد. به عنوان مثال، **مدرسه** کلمه عربی برای هر دوی "مدرسه" و "معلم" است. اگر کلمه جدا شود، تفکیک کننده ها، برای تمایز بین دو معنی امکان پذیر، ضروری هستند. اگر این در یک جمله رخ می دهد، اطلاعات متنی ذاتی در جمله را می توان برای پی بردن به معنای مناسب استفاده نمود. در این مقاله، موضوع تفکیک واکه در نظر گرفته نشده است، زیرا برای نوشتن عربی، به کار نگرفتن این تفکیک ها، رایج تر است. تفکیک کننده تنها در نسخه های خطی قدیمی و یا در مناطق بسیار محدود یافت می شوند.

الفبای عربی به صورت عددی توسط یک کد تبادل ارتباطات استاندارد تایید شده توسط سازمان استاندارد و علم اوزان ومقادیر عرب (ASMO) ارائه شده است. شبیه به کد استاندارد آمریکایی برای تبادل اطلاعات (ASCII)، هر کاراکتر در کد ASMO توسط یک بایت نشان داده می شود. حرف انگلیسی دارای دو شکل ممکن است، بزرگ و کوچ. کد ASCII، نمایش جداگانه برای هر دوی این اشکال را را فراهم می کند، در حالی که حرف عربی تنها یک نمایش در جدول ASMO دارد. با این حال، نمی توان گفت که حرف عربی تنها یک شکل دارد. در مقابل، یک نامه عربی می تواند بسته به موقعیت نسبی آن در متن دارای چهار اشکال مختلف باشد. به عنوان مثال، حرف (ع A'in)

دارای چهار شکل مختلف است: در آغاز کلمه (قبل از یک فاصله)، در وسط کلمه (بدون فضای اطراف آن)، در پایان کلمه (به دنبال فاصله)، و به صورت جداگانه (قبل از حرف متصل نشده و بعد از آن فاصله). این چهار امکان در شکل 1 نشان داده شده است.

جدول 2، نشان دهنده اشکال مختلف کاراکترهای عربی در موقعیت های مختلف از کلمه است.

علاوه بر این، کاراکترهای مختلف عربی ممکن است دقیقا دارای یک شکل باشند، و تنها با اضافه نمودن یک کاراکتر مکمل از یکدیگر متمایز می شوند. (کاراکترهای مکمل: بخشی از یک کاراکتر که برای تکمیل یک کاراکتر عربی مورد نیاز است). اینها معمولا یک نقطه، یک گروه از نقطه ها و یا زیگزاگ (حمزه) هستند. اینها ممکن است در، بالا، و یا زیر خط پایه به نظر برسند وبه طور متفاوت، به عنوان مثال، بالا، پایین و یا در محدوده کاراکتر قرار گیرند. شکل 2، دو مجموعه از کاراکترها را نشان می دهد که اولین مجموعه دارای پنج کاراکتر و مجموعه دیگر سه کاراکتر دارد. واضح است که هر مجموعه شامل کاراکترهایی می شود که تنها به واسطه موقعیت و / یا تعداد نقاط مرتبط با آن متفاوت هستند. شایان ذکر است که هرگونه ساییدگی یا حذف این کاراکترهای مکمل منجر به ارائه اطلاعات نادرست کاراکتر می شود. از این رو، هر الگوریتم نازک شدن باید به طور موثر با این نقاط برای تغییر ندادن هویت کاراکتر مقابله نماید.



شکل. 1. اشکال مختلف حرف عربی 'A' در '.

نوشتن عربی خط شکسته است و به گونه ای است که کلمات با فاصله از هم جدا می شوند. با این حال، یک کلمه را می توان به واحدهای کوچکتر به نام زیرکلمه ها (بخشی از یک کلمه شامل یک یا چند کاراکتر متصل بیشتر) تقسیم نمود. برخی از کاراکترهای عربی با کاراکترهای پشت سر هم قابل اتصال نیستند. بنابراین، اگر یکی از این کاراکترها

در یک کلمه وجود داشته باشد، آن کلمه را به دو زیرکلمه تقسیم می کند. این کاراکترها تنها در دم زیرکلمه به نظر می رستد، و کاراکتر بعدی، سر زیرکلمه بعدی را تشکیل می دهد. شکل 3 سه کلمه عربی با یک، دو و سه زیرکلمه را نشان می دهد. اولین کلمه متشکل از یک زیرکلمه است که دارای نه حرف می باشد؛ دومی دارای دو زیرکلمه با سه و یک حرف است. آخرین کلمه شامل پنج زیرکلمه است که هر کدام شامل تنها یک حرف می باشد.

نوشتن عربی می تواند، به طور کلی، به سبک های تایپ شده (نسخ)، دست نوشته (Ruq'a) و هنری (و یا تزئینی خوشنویسی، Kufi، خط دیوانی، سلطنتی، و خط ثلث) طبقه بندی شود همانطور که در شکل 4 نشان داده شده است. سبک های دست نویس و تزئینی معمولا شامل ترکیبات عمودی از کاراکترها به نام دو یا چند حرف متصل بهم می شود. این ویژگی، تعیین مرزهای کاراکترها را دشوار می سازد. علاوه بر این، کاراکترها با یک فونت دارای اندازه های مختلف هستند (یعنی، مثال کاراکترها ممکن است دارای عرض های مختلف باشند، حتی اگر دو کاراکتر دارای فونت و نقطه اندازه یکسان باشند). از این رو، تقسیم بندی کلمه بر اساس عرض اندازه ثابت را نمی توان برای عربی استفاده نمود.

3. تشخیص کاراکترهای عربی

دو استراتژی وجود دارند که برای تشخیص کاراکتر عربی چاپی و دست نوشته استفاده می شوند. اینها را می توان به شرح زیر طبقه بندی نمود:

(1) استراتژی های جامع گرا که در آن تشخیص در سطح کلی بر اساس ارائه کلی کلمات است و در آن هیچ تلاشی برای شناسایی کاراکترهای جداگانه وجود ندارد. این استراتژی ها در اصل برای تشخیص گفتار معرفی می شوند و می توانند به دو رده تقسیم شوند:

(1. 1) روش های مبتنی بر اندازه گیری های فاصله با استفاده از برنامه نویسی پویا. (48، 49)

(1. 2) روش های مبتنی بر چارچوب احتمالاتی (مدل های مخفی Markov). (50-55)

(2) استراتژی های تحلیلی که در آن کلمات به عنوان یک کل در نظر گرفته نمی شوند، بلکه به عنوان توالی از واحدهای اندازه کوچک در نظر گرفته می شوند و تشخیص به طور مستقیم در سطح کلمه انجام نمی شود، بلکه در سطح میانی با این واحدها برخورد می شود، که می تواند حروف الفبا، بخش ها، شبه حروف، و غیره باشد (47، 56، 57)

3.1. تقسیم بندی کلمه

مرحله تقسیم بندی، یک گام ضروری در تشخیص متن عربی چاپ شده است. هر گونه خطا در بخش بندی شکل اولیه از کاراکترهای عربی، ارائه متفاوتی از جزء کاراکتر را تولید خواهد نمود. دو روش برای بخش بندی کلمات عربی چاپی و دست نوشته در کاراکترهای فردی استفاده می شود: بخش بندی های صریح و ضمنی:

(1) تقسیم بندی ضمنی (تقسیم بندی مستقیم): در این روش، کلمات به طور مستقیم به حروف تقسیم می شوند. این نوع از تقسیم بندی معمولا با قوانینی طراحی می شود که برای شناسایی تمام نقاط تقسیم بندی کاراکتر تلاش می نماید.

(2) تقسیم بندی صریح: در این مورد، کلمات از نظر خارجی به شبه حروف تقسیم بندی می شوند که پس از آن به صورت جداگانه تشخیص داده می شوند. این رویکرد معمولا به دلیل پیچیدگی فزاینده پیدا کردن فرضیه های کلمه مطلوب گران تر است.

در تمام کاراکترهای چاپی عربی، عرض در یک نقطه اتصال بسیار کمتر از عرض کاراکتر است. این ویژگی در استفاده از روش تقسیم بندی پایه ضروری است. (56، 57، 59) خط مبنا یک خط میانی در کلمه عربی است که در آن تمام ارتباطات بین کاراکترهای پی در پی رخ می دهند. اگر یک طرح عمودی برای پیکسل ها دوسطحی روی معادله کلمه انجام شود (1)]،

$$v(j) = \sum_{i} w(i, j) \qquad (1)$$

که در آن $w(i,j)$ یا صفر یا یک است و **i**، **j** به ترتیب شاخص سطر و ستون است، نقطه اتصال دارای مجموع کمتر از مقدار متوسط (AV) [معادله (2)] است

$$AV = (1/Nc) \sum_{j=1}^{Nc} Xj \qquad (2)$$

و در آن **Nc** تعداد ستون است و **Xj** تعداد پیکسل های سیاه ستون **j** ام است.

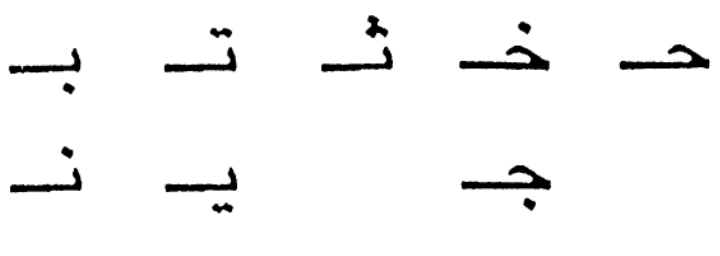
جدول 2. الفبای پایه برای کاراکترهای عربی و اشکال آنها در موقعیت های مختلف در کلمه

	isolated (i)	end (e)	middle (m)	beginning (b)
alif	ﺍ	ﺎ	ﺎ	ﺍ
ba	ﺏ	ﺐ	ﺒ	ﺑ
ta	ﺕ	ﺖ	ﺘ	ﺗ
tha	ﺙ	ﺚ	ﺜ	ﺛ
jim	ﺝ	ﺞ	ﺠ	ﺟ
ha	ﺡ	ﺢ	ﺤ	ﺣ
kha	ﺥ	ﺦ	ﺨ	ﺧ
dal	ﺩ	ﺪ	ﺪ	ﺩ
dhal	ﺫ	ﺬ	ﺬ	ﺫ
ra	ﺭ	ﺮ	ﺮ	ﺭ
zan	ﺯ	ﺰ	ﺰ	ﺯ
siin	ﺱ	ﺲ	ﺴ	ﺳ
shiin	ﺵ	ﺶ	ﺸ	ﺷ
sadd	ﺹ	ﺺ	ﺼ	ﺻ
dad	ﺽ	ﺾ	ﻀ	ﺿ
tahn	ﻁ	ﻂ	ﻄ	ﻃ
zah	ﻅ	ﻆ	ﻈ	ﻇ
ayn	ﻉ	ﻊ	ﻌ	ﻋ
ghayn	ﻍ	ﻎ	ﻐ	ﻏ
fa	ﻑ	ﻒ	ﻔ	ﻓ
qaf	ﻕ	ﻖ	ﻘ	ﻗ
kaf	ﻙ	ﻚ	ﻜ	ﻛ
lam	ﻝ	ﻞ	ﻠ	ﻟ
miim	ﻡ	ﻢ	ﻤ	ﻣ
noon	ﻥ	ﻦ	ﻨ	ﻧ
ha	ﻩ	ﻪ	ﻬ	ﻫ
waw	ﻭ	ﻮ	ﻮ	ﻭ
ya	ﻯ	ﻲ	ﻴ	ﻳ
lamalif	ﻻ	ﻼ	ﻼ	ﻻ
tamarbot	ﺓ	ﺔ		

از این رو، هر بخش با مقدار مجموع بسیار کمتر از AV باید یک مرز بین کاراکترهای مختلف باشد. با این حال اگر هیستوگرام تولید شده از طرح عمودی از شرط معادله (3) پیروی نکند، این کاراکتر، بخش بندی نشده باقی می ماند، همانطور که در شکل 5 نشان داده شده است.

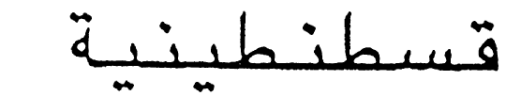
با بررسی کاراکترهای عربی، می توانیم دریابیم که فاصله بین قله های پی در پی، از یک سوم عرض کاراکتر عربی تجاوز نمی کند. یعنی

$$|d_k| < d_l/3 \qquad (3)$$

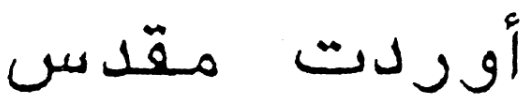


(a) (b)

شکل. 2. کاراکترهای عربی متفاوت تنها با توجه به موقعیت و تعداد نقاط مرتبط.



(a)



(b) (c)

شکل. 3. کلمات عربی با زیرکلمه های تشکیل دهنده.

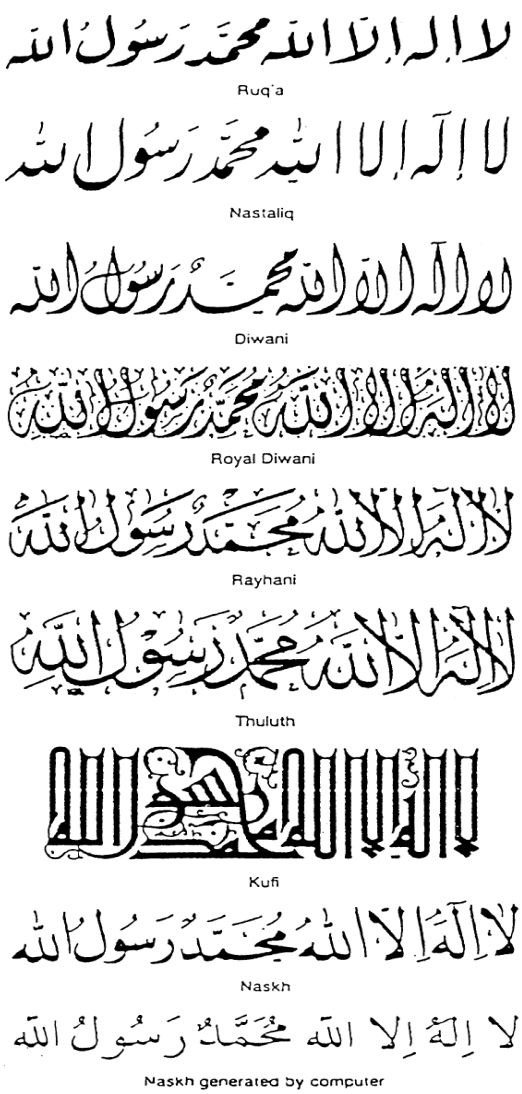
که در آن d_k فاصله بین قله k ام و قله K + 1، و d_l عرض کلی کاراکتر است.

علاوه بر این، در پایان یک کلمه یا یک زیرکلمه، معادله (4) نیز برقرار است.

$$L_{k+1} > 1.5L_k \qquad (4)$$

که در آن L_k قله k ام در هیستوگرام است. این قانون به دلیل درون اتصال کاراکترهای عربی و اشکال آنها در پایان یک کلمه است.

این رویکرد به شدت به یک مقدار آستانه از پیش تعریف شده مربوط به عرض کاراکتر بستگی دارد. علاوه بر این، این روش به طور موثر برای تصاویر اریب کارنخواهد کرد.

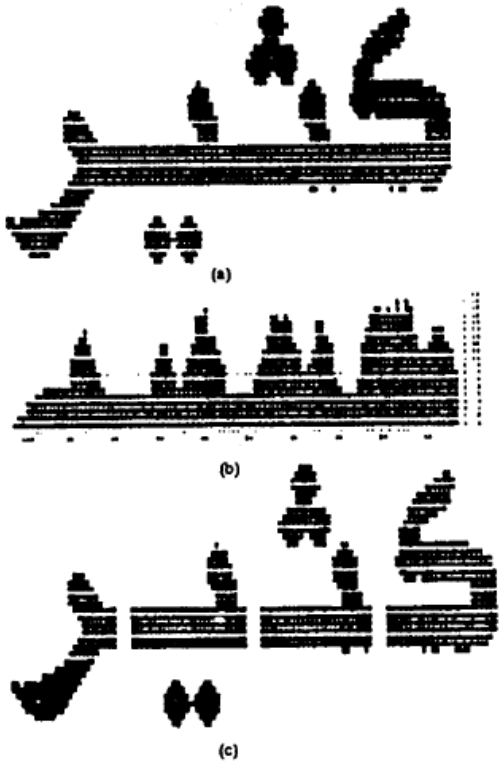


شکل. 4. سبک ها و فونت های مختلف برای نوشتن متن عربی.

Almuallim و Yamaguchi (57) یک تکنیک تشخیص ساختاری را برای کلمات دست نویس عربی ارائه نمودند. سیستم آنها شامل چهار مرحله می شود. اولی، پیش پردازش، که در آن کلمه نازک می شود و خط وسط کلمه تشخیص داده می شود. از آنجا که بخش بندی یک کلمه خط شکسته به حروف دشوار است، کلمات به تکان های

جداگانه تقسیم بندی می شوند و به عنوان کاراکترهای مکمل، تکان ها با یک حلقه و تکان ها بدون حلقه طبقه بندی می شوند. سپس این تکان ها با استفاده از خواص هندسی و توپولوژیک آنها طبقه بندی می شوند. در نهایت، موقعیت نسبی تکان های طبقه بندی شده مورد بررسی قرار می گیرند، و تکان ها در مراحل مختلف در رشته ای از کاراکترها که نشان دهنده کلمه تشخیص داده شده است ترکیب می شوند. خرابی های سیستم در اکثر موارد به دلیل تقسیم بندی نادرست کلمات رخ می دهد.

تقسیم بندی نیز با دنبال کردن خط واصل بیرونی (60) از یک کلمه معینو محاسبه فاصله بین نقاط کرانی از تقاطع خط واصل با یک خط عمودی به دست می آید. این تقسیم بندی بر اساس یک اسکن افقی از راست به چپ خط واصل بسته با استفاده از پنجره قابل تنظیم با عرض W می باشد. برای هر موقعیت از پنجره، فاصله عمودی میانگین h_{av} در سراسر این پنجره محاسبه می شود. در مرز بین دو کاراکتر، شرایط زیر باید رعایت شوند:



شکل. 5. یک مثال از تقسیم بندی کلمه عربی **کثیر** به کاراکترها: (الف) کلمه عربی، (ب) هیستوگرام، (ج) کلمه تقسیم بندی شده به کاراکترها.

(1) $h_{av} < T$. در این مورد، یک منطقه سکوت شناسایی می شود، به این معنی که فاصله عمودی متوسط روی پنجره باید کمتر از یک آستانه از پیش تعیین خاص T باشد.

(2) مرزهای شناسایی باید روی همان خط افقی (خط پایه) قرار گیرد.

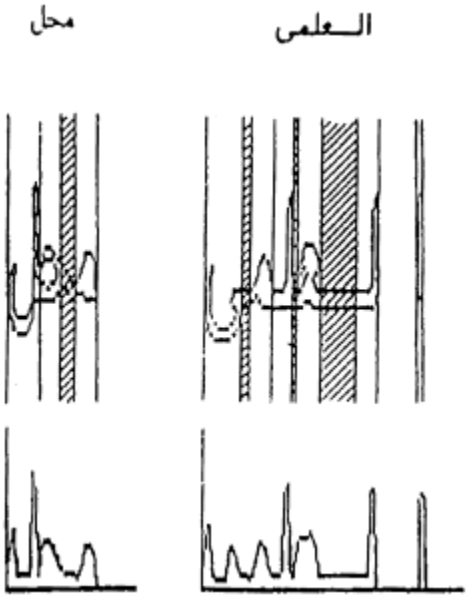
(3) هیچ کاراکتر مکمل نباید (در بالا یا زیر خط پایه) در یک منطقه سکوت واقع شوند.

تنظیم مجدد پارامترهای w و T و همچنین ردیابی به سمت عقب ممکن است رخ دهد، اگر تقسیم بندی به یک شکل کاراکتر مردود منجر شود. شکل 6 چند نمونه از این روش نشان می دهد.

El-Khaly و Sid-Ahmed (61)، یک کلمه نازک شده را به کاراکترها توسط خط مبنای زیر تقسیم بندی نمودند و زمانی که پیکسل ها شروع به بالاتر رفتن کردند یا پایین تر از آن آمدند، کار تشخیص را انجام دادند.

Abdelazim و Hashish (62) از تکنیک پیمودن یک منحنی انرژی (مشابه با مورد استفاده شده در تشخیص گفتار، برای متمایز نمودن سخن صحبت شده از پس زمینه سکوت) استفاده نمودند که نشان دهنده تعداد پیکسل های سیاه در هر ستون از کلمه دیجیتالی برای تقسیم بندی کلمه به کاراکترها است. این منحنی پیموده می شود و یک مقدار آستانه برای موارد اولیه از مناطق سکوت انتخاب می شود.

Shoukry (63) از یک الگوریتم ترتیبی بر اساس اصل ردیابی زمان-ورودی استفاده نمود که بستگی به خواص اتصال متن به دست آمده در حوزه تصویر باینری دارد. این الگوریتم برخی از شباهت ها را به یک الگوریتم ابداع شده توسط Wakayama (64) برای اسکلت بندی تصاویر دودویی دارد.



شکل. 6. کلمات بخش بندی شده عربی و ارتفاعات حد فاصل مربوطه.

سیستم SARAT (65)، حد فاصل های بیرونی را برای بخش بندی یک کلمه عربی به کاراکترها استفاده می نماید. این کلمه، با تعیین نقاط شروع و پایان کلمه، به یک سری از منحنی ها تقسیم می شود. هر گاه حد فاصل بیرونی، علامت را تغییر دهد (از انحنای مثبت به منفی) یک کاراکتر تقسیم بندی می شود.

Kurdy و Joukhadar (66) از تابع فاصله فوقانی زیرکلمه استفاده نمودند که مجموعه ای از بالاترین نقاط در هر ستون است. آنها به هر نقطه تابع، یک نشانه را با مقایسه ارتقاع نقطه با این ارتفاع و یک نام رمزی نقطه در سمت راست آن منسوب نمودند. با استفاده از دستور زبان، آنها دنباله ای از نشانه های یک زیرکلمه را برای پیدا کردن نقاط اتصال تجزیه نمودند.

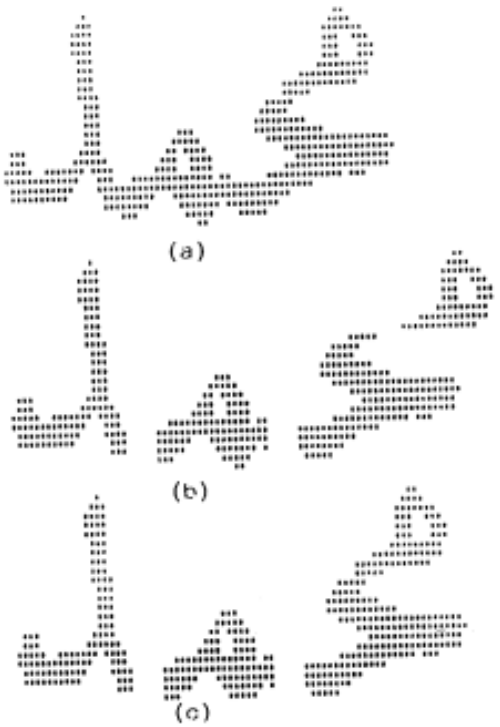
در نهایت، Amin و Al-Sadoun (67، 68) یک تکنیک جدید را برای بخش بندی متن عربی اتخذ نمودند. این الگوریتم را می توان برای هر فونت به کار برد و پوشش کاراکترها را میسر می سازد. دو مشکل عمده با روش تقسیم بندی سنتی که بستگی به خط مبنا دارد، وجود دارد:

(1) همپوشانی کاراکترهای عربی مجاور به طور طبیعی رخ می دهد، شکل 7a را ببینید. از این رو، هیچ خط مبنایی وجود ندارد. (پدیده او در هر دو متن عربی تایپ شده و دست نوشته رایج است.

(2) ارتباط بین دو کاراکتر اغلب کوتاه است. بنابراین، قرار دادن نقاط تقسیم بندی، یک کار دشوار است. در بسیاری از موارد، نقاط تقسیم بندی بالقوه در یک کاراکتر قرار خواهد گرفت، نه بین کاراکترها.

کلمه شکل 7a با استفاده از یک تکنیک خط مبنا تقسیم شده است. شکل 7b نشان دهنده تقسیم بندی مناسب است و نتیجه روش تقسیم بندی جدید در شکل 7c نشان داده شده است.

روش جدید را می توان به چهار مرحله اصلی تقسیم نمود. اولی، مرحله دیجیتالی نمودن است که در آن تصویر اصلی به یک تصویر باینری با استفاده از یک اسکنر (300 نقطه در اینچ) تبدیل می شود. دوم، یک مرحله پیش پردازش است که در آن کلمه عربی با استفاده از یک الگوریتم نازک شدن موازی نازک می شود. سوم، اسکلت تصویر از سمت راست به چپ با استفاده از یک پنجره 3x3 ترسیم می شود و یک درخت دودویی ساخته می شود. کد Freeman (69) برای توصیف شکل اسکلت استفاده می شود. در نهایت، درخت دودویی به زیردرخت ها تقسیم می شود به طوری که هر یک از زیر درخت ها، یک کاراکتر را در تصویر توصیف می نماید.



شکل. 7. مثال از یک کلمه عربی **محمد** و تکنیک های مختلف تقسیم بندی.

3.2 استخراج و تشخیص ویژگی

مشخص شده است که ویژگی ها، کوچکترین مجموعه ای را که می تواند برای مقاصد تمایز و برای شناسایی منحصر به فرد برای هر یک از کاراکترها استفاده شود نشان می دهند. ویژگی ها را می توان به دو دسته طبقه بندی نمود:

(1) ویژگی های موضعی که معمولاً هندسی هستند (به عنوان مثال بخش های مقعر / محدب، نوع اتصالات: اتصالات تقاطع / T- / نقطه پایانی، و غیره).

(2) ویژگی های کلی که معمولا توپولوژیک (اتصال، تعدادی از قطعات متصل شده، تعداد حفره، و غیره) و یا آماری (تبدیل فوریه، لحظات ثابت، و غیره) هستند.

Nouh و همکاران. (70) یک مجموعه کاراکتر عربی استاندارد را برای تسهیل پردازش کامپیوتری کاراکترهای عربی پیشنهاد نمودند. در این کار، سیزده ویژگی، و یا رادیکال، که نشان دهنده بخش هایی از کاراکترها است، توسط بازرسی انتخاب می شوند. تشخیص بر اساس یک درخت تصمیم گیری و یک اندازه گیری همبستگی قوی است. عیب سیستم پیشنهادی، این فرض است که کاراکترهای دریافتی با توجه به قوانین استاندارد مشخص شده تولید می شوند.

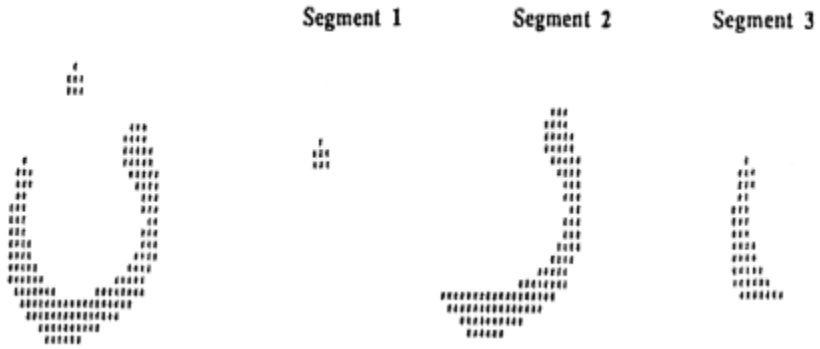
Parhami و Taraghi (71) یک تکنیک را برای خودکار شناسایی متن فارسی چاپی (که شبیه به متن عربی است) ارائه نمودند. نویسندگان، ابتدا با شناسایی یک سری از نقاط اتصال بالقوه روی خط مبنا که در آن ضخامت خط از یا به ضخامت خط مبنا تغییر می کند، زیرکلمه را به کاراکترها تقسیم بندی نمودند. اگر چه آنها برخی از قوانین را برای کاراکترها در پایان زیرکلمه دست نخورده نگهداشتند، برخی از کاراکترهای گسترده تر (به عنوان مثال **س**) به سه بخش تقسیم بندی نمودند. سپس آنها بیست ویژگی ها را مبتنی بر برخی از خواص هندسی از علامت فارسی برای ساخت یک بردار 24 بیتی انتخاب نمودند که با ورودی های یک جدول که در آن مطابقت دقیق در ابتدا بررسی می شود، مقایسه می شود. این سیستم به شدت وابسته به فونت است و انتظار می رود که فرایند تقسیم بندی نتایج نادرست را در برخی موارد ارائه دهد.

جدول مراجعه برای تشخیص کاراکترهای عربی دست نوشته جدا شده استفاده می شود. (72) در این روش، کاراکتر در یک فریم قرار داده می شود که به شش مستطیل تقسیم می شود و یک الگوریتم ردیابی حد فاصل برای کدنویسی حدفاصل به عنوان مجموعه ای از بردارهای جهتی با استفاده از یک کد Freeman استفاده می شود. با این حال، این اطلاعات برای تعیین کاراکترهای عربی کافی نمی باشد، بنابراین اطلاعات اضافی مربوط به تعداد نقاط و موقعیت آنها اضافه می شود. در صورتی که هیچ تطبیقی وجود نداشته باشد، این سیستم، بردار ویژگی را به جدول اضافه می کند و آن کاراکتر را به عنوان یک ورودی جدید در نظر می گیرد.

Amin و Masini (56) یک رویکرد ساختاری را برای تشخیص متن عربی چاپی اتخاذ نمودند. کلمات و زیرکلمه ها با استفاده از روش خط مبنا به کاراکترها تقسیم می شوند. ویژگی هایی از قبیل میله های عمودی و افقی از کاراکتر با استفاده از تصویرسازی های افقی و عمودی استخراج می شوند. چهار درخت تصمیم گیری انتخاب شده با توجه به موقعیت کاراکتر درون کلمه استفاده می شود که توسط فرآیند تقسیم بندی محاسبه شده است. ساختار چهار درخت تصمیم گیری، جستجوی سریع را برای کاراکتر مناسب میسر می سازد. علاوه بر این، درختان در کاراکتر متمایز که دارای یک شکل هستند اما در موقعیت های مختلف در یک کلمه به نظر می رسند استفاده می شوند.

Amin و Mari (57) یک تکنیک جدید را برای یک متن عربی چند فونتی که شامل تشخیص کاراکتر و کلمه می شود پیشنهاد نمودند. یک کاراکتر توسط یک فرآیند اسکن افقی (شکل 8) به بسیاری از بخش ها تقسیم می شود. به این ترتیب، بخش ها برای تشکیل یک شکل پایه از کاراکتر متصل می شوند. بخش های غیرمتصل با هر بخش دیگر، کاراکترهای مکمل در نظر گرفته می شوند. با استفاده از کد Freeman، (69) فرایند تشخیص حد فاصل برای این بخش ها اعمال می شود تا شکل پایه کاراکتر ترسیم شود و یک بردار جهتی را از طریق یک پنجره 2x2 تولید نماید. سپس یک درخت تصمیم گیری برای تشخیص کاراکتر استفاده می شود. در نهایت، یک الگوریتم Viterbi (73) برای تشخیص کلمه عربی به منظور ارتقای نرخ تشخیص استفاده می شود. مزیت اصلی این روش این است که استفاده از فرآیند یادگیری خودکار میسر می شود.

مطالعه گزارش شده در مراجع (61)، (74) و (75) از توصیف کننده های غیرمتغیر با زمان برای تشخیص کاراکتر بهره گیری می نماید. روش های دیگر شامل مجموعه ای از توصیف کننده های فوریه از توالی های مختصات حدفاصل بیرونی که برای تشخیص استفاده می شود، می باشند. (60) همچنین، در مرجع (76) به هر کاراکتر، یک تابع منطقی تخصیص داده می شود که در آن کاراکترها، وابسته به وجود پیکسل های خاص در یک محل مشخص شده از تصویر، از قبل به چهار گروه طبقه بندی می شوند.



شکل. 8. بخش های اصلی کاراکتر **ن**.

در مرجع (45)، جدول مراجعه برای تشخیص کاراکترهای عربی جدا شده اتخاذ شده است. در این روش، کاراکتر در فریم پنجره قرار داده می شود و برای استخراج برخی از ویژگی ها به پنجره های کوچک تقسیم می شود. این ویژگیها عبارتند از نقاط انتهایی، نقاط تقاطع، گوشه ها، و رابطه بین طول و عرض فریم پنجره. کاراکترها توسط یک ارتباط بین نقاط ویژگی و مکان های آنها در درون فریم پنجره شناسایی می شوند. تشخیص با پیدا کردن یک تطبیق بین کاراکترهای ناشناخته و ورودی ها در یک جدول مراجعه به دست می آید.

برای افزایش نرخ تشخیص یک سیستم OCR، برخی از ویژگی های مورفولوژیکی زبان عربی را می توان مورد استفاده قرار داد. Amin و Al Fedaghi (58، 77)، یک روش را برای تصحیح املای کلمات عربی توصیف نمودند. آنها اشتباهات املایی و کلمات کاملی که دارای برخی از کاراکترهای ناشناخته با استفاده از یک الگوریتم هستند که به فراوانی های ریشه و الگوها به زبان عربی بستگی دارد را تصحیح می نمایند.

Haj Hassan (78، 79) یک روش نحوی-ساختاری را برای تشخیص متن عربی چاپ شده معرفی نمودند. کلمات به کاراکترها با استفاده از یک روش شبیه به روش پیشنهادی در مرجع (71) تقسیم می شوند. سپس شکلهای هندسی اولیه مانند افقی، عمودی، مورب با شیب های مثبت و منفی از کاراکتر استخراج می شوند. این شکلهای هندسی اولیه در منطقه از پیش تعریف شده در داخل کاراکتر تشخیص داده می شوند. در نهایت، زبان های توصیفی (کلمه باینری) برای توصیف کاراکتر استفاده می شوند.

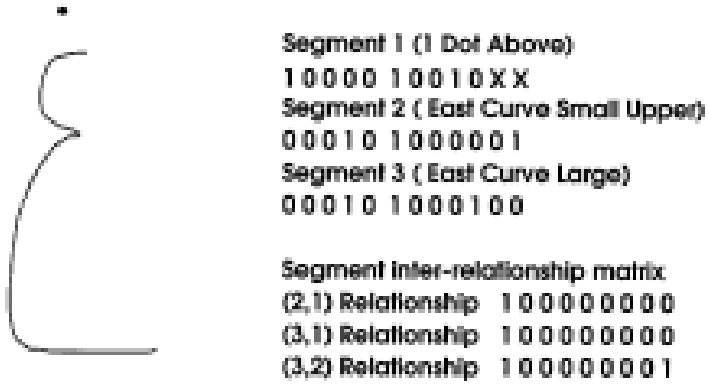
Al-Badr و Haralick (80) یک سیستم تشخیص کلمات عربی چاپی را بدون تقسیم بندی قبلی با استفاده از عملیات مورفولوژی ریاضی در تمام صفحه برای پیدا کردن مکان هایی که در آن شکلهای هندسی اولیه حاضر هستند ارائه می شوند، معرفی نمودند. سپس آنها آن شکلهای هندسی اولیه را در کاراکترها ترکیب نمودند و هویت های کاراکتر و موقعیت آنها بر روی صفحه را چاپ نمودند.

Sano و همکاران. (81) یک رویکرد ساختاری را با استفاده از روابط فازی برای تشخیص کاراکترهای عربی جداشده دست نوشته معرفی نمودند. هر الگوی ورودی توسط نقاط ویژگی به زیر الگوها (تکان ها) تقسیم می شود؛ نقاط پایانی، نقاط شاخه، تقاطعات و حداکثر نقطه انحنا، و غیره. تعداد زیرالگوها، بسته به کاراکتر ورودی از 1 تا 6 متغیر است. سپس این زیرالگو از نظر شباهت با عناصر اولیه (خط راست، دایره و نقطه حروف) ارائه می شوند. این الگوریتم بر روی تعداد کمی از نمونه های دست نوشته آزمایش می شود.

در نهایت، Bouslama (82) یک الگوریتم را بر اساس روش ساختاری و منطق فازی برای تشخیص کاراکترهای چاپی عربی جداسازی شده اتخاذ نمودند. روش ساختاری برای استخراج ویژگی ها از کاراکتر ورودی مانند تعداد تکان ها، قبل و بعد از تقسیم بندی، موقعیت مرکز ثقل هر زیر بخش، نسبت پیکسل سیاه زیر بخش با توجه به تعداد پیکسل های سیاه در اسکلت، کد زنجیره، نسبت طول فاصله بین نقاط نهایی و طول کلی هر زیر بخش، و غیره استفاده می شود. مفاهیم منطق فازی برای مدلسازی هرگونه تفاوت ها و یا عدم قطعیت ها در مقادیر استفاده می شوند تا ارائه بهتر و واقعی تر از این ویژگی ها میسر شود. علاوه بر این، قوانین فازی نیز برای طبقه بندی کاراکتر استفاده می شوند.

3. 3. طبقه بندهای شبکه عصبی

در میان کاربردهای بسیاری که برای شبکه های عصبی ارائه می شوند، تشخیص کاراکتر یکی از موفق ترین ها بوده است. در مقایسه با روش های دیگر مورد استفاده در تشخیص الگو، مزایای غالب بیان شده به نفع یک روش شبکه عصبی برای تشخیص الگوی این است که (1) نسبت به روش های دیگر، نیاز به ورودی کمتر از دانش در مورد مشکل دارد (2) قادر به پیاده سازی پارتیشن بندی پیچیده تر از فضای ویژگی است و (3) متمایل به پیاده سازی های پردازش موازی با عملکرد بالا است. با این حال، معایب راه حل های شبکه عصبی، در مقایسه با روش های آماری، شامل (1) مقدار گسترده ای از آموزش مورد نیاز، (2) عملیات کندتر زمانی که به عنوان یک شبیه سازی بر روی یک کامپیوتر معمولی پیاده سازی می شود و (3) در دسترس نبودن درک دقیق از فرایند تصمیم گیری (به عنوان مثال سطوح تصمیم گیری در فضاهای ویژگی) می شود. (83)



شکل. 9. نمایش کامل یک کاراکتر عربی برای لایه ورودی شبکه عصبی.

Amin and Al-Sadoun (84-86) یک رویکرد ساختاری را برای تشخیص کاراکترهای دست نوشته عربی پیشنهاد نمودند. در ابتدا تصویر باینری از کاراکتر با استفاده از یک الگوریتم نازک شدن موازی نازک می شود و سپس اسکلت تصویر از سمت راست ترسیم به چپ با استفاده از پنجره 3x3 به منظور ساخت یک نمودار برای نشان دادن کاراکتر ترسیم می شود. سپس ویژگی هایی مانند خطوط مستقیم، منحنی ها و حلقه ها از نمودار استخراج می شوند. در نهایت، یک شبکه عصبی مصنوعی پنج لایه برای طبقه بندی کاراکتر استفاده می شود. هر کاراکتر از نظر بخش های استفاده شده در سیستم مانند نقطه، حمزه، خط، منحنی و حلقه طبقه بندی می شود. روابط بین

بخش ها در ماتریس شی رابطه –میانی کد گذاری می شوند. طراحی کلی لایه ورودی از 150 نورون استفاده می کند. شکل 9 نمونه ای از ارائه کاراکتر با استفاده از این طرح لایه ورودی را نشان می دهد.

Altuwaijri و Bayoumi (87) یک سیستم را برای تشخیص کلمات عربی چاپی با استفاده از شبکه های عصبی مصنوعی (NN) معرفی نمودند. این سیستم را می توان در سه مرحله مختلف تشریح نمود: اول کلمه ورودی عربی به کاراکترها با استفاده از یک روش شبیه به آنچه در مرجع (56) آمده است تقسیم می شود. در مرحله بعد، شش لحظه برای استخراج ویژگی ها از تغذیه کاراکترهای بخش بندی شده به شبکه عصبی استفاده می شوند. در نهایت، یک شبکه پرسپترون چند لایه با آموزش انتشار بازگشتی با یک لایه پنهان برای طبقه بندی کاراکتر استفاده می شود.

در نهایت، Amin و Mansoor (88) از شبکه های عصبی مصنوعی برای تشخیص متن چاپی عربی استفاده نمودند. این روش را می توان به سه مرحله عمده خلاصه سازی نمود: مرحله اول، پیش پردازش است که در آن تصویر اصلی به یک تصویر باینری با استفاده از یک اسکنر 300 نقطه در اینچ و پس از آن تشکیل جزء متصل تبدیل می شود. دوم، سپس ویژگی های کلی کلمه عربی ورودی مانند تعداد زیرکلمه ها، تعداد قله در زیرکلمه، نوع و تعداد کاراکترهای مکمل، و غیره استخراج می شوند. در نهایت، یک شبکه عصبی مصنوعی سه لایه برای طبقه بندی کلمه استفاده می شود. طراحی کلی لایه ورودی در مجموع از 270 نورون استفاده می کند.

3. 4. روش های تصادفی

مدل های Markov پنهان (HMM) در حال حاضر تبدیل به الگوی رایج در تشخیص گفتار خودکارشده است. (89-91) به تازگی، چند محقق در تشخیص دست خط برای پس و پیش کردن فناوری HMM در زمینه آنها پس از تحقق این مورد که تصاویر کلمه را بتوان به توالی مشاهدات جذب نمود، تلاش کرده اند. (92-101) HMM یک خانواده از ابزار را برای مدل سازی فرآیندهای متوالی به شیوه ای آماری و مولد تشکیل داده اند. شهرت آنها ناشی از

نتایج به دست آمده در تشخیص گفتار است که عمدتا از وجود روش های آموزش اتوماتیک و مزایای استفاده از چارچوب احتمالاتی استنتاج می شوند.

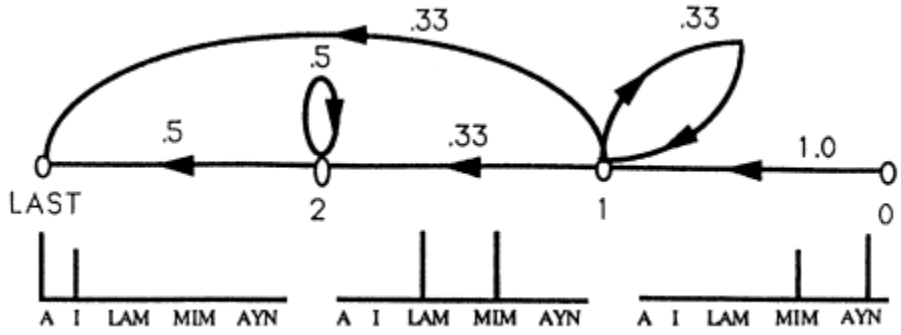
یک HMM را می توان بدین صورت تعریف نمود: (1) مجموعه ای از حالات {S} ، با یک S_I حالت اولیه و یک S_F در حالت نهایی. (2) ماتریس احتمال انتقال، $A = \{a_{ij}\}$ ، که در آن a_{ij} احتمال گذار برای گذار از حالت i به حالت j است. (3) ماتریس احتمال خروجی B. برای HMM گسسته، $B = \{b_j(O_k)\}$ ، که در آن O_k نشان دهنده یک نماد مشاهده گسسته است. برای یک HMM مداوم، $\text{HMM, B} = \{b_j(x)\}$ ، که در آن x مشاهدات مستمر بردارهای تصادفی K-بعدی را نشان می دهد. اگر توزیع حالت اولیه $\pi = \{\pi_i\}$ ، مجموعه پارامتری کامل از HMM را می توان به طور فشرده به صورت زیر بیان نمود:

$$\lambda = (A, B, \pi) \qquad (5)$$

یک HMM می تواند بر اساس توزیع های احتمال مشاهده گسسته و یا تابع چگالی احتمال ترکیبی پیوسته باشد. در HMM گسسته، توزیع های احتمال گسسته به اندازه کافی برای توصیف هر رویداد تصادفی با تعداد مناسب از پارامترها قدرتمند هستند. مزیت اصلی HMM پیوسته، توانایی مدلسازی مستقیم پارامترهای سیگنال پیوسته است. یک HMM نیمه پیوسته، یک چارچوب را برای متحد نمودن HMM گسسته و پیوسته را فراهم می کند.

با توجه به تعریف HMM، سه مشکل اساسی مورد نظر وجود دارند که باید برای کاربردهای دنیای واقعی حل شوند: مشکل ارزیابی، مشکل رمزگشایی، و مشکل یادگیری. راه حل های این سه مشکل الگوریتم رو به جلو-رو به عقب، الگوریتم Viterbi، و الگوریتم Baum-Walch هستند. برای جزئیات دقیق تر در تئوری و محاسبه HMM، خوانندگان می توانند به مرجع (102) مراجعه نمایند.

چالش عمده در سیستم های تشخیص نوشتن عربی از ماهیت خط شکسته داده ها ناشی می شود. بنابراین، هیچ یک از سیستم های HMM موجود، در صورت وجود، قادر نیستند که به دقت دست خط خط شکسته دست نوشته عربی نامحدود را تشخیص دهند. با این حال، چند محقق، HMM را روی متن چاپ شده عربی استفاده نمودند.



شکل. 10. HMM مرتبط با واژه: /ala / ayn lam a.

به عنوان مثال، Amin و Mari (57) الگوریتم Viterbi را به منظور ارتقای تشخیص متن چاپی چند فونتی عربی مورد استفاده قرار دادند. پس از وقوع خرابی در تشخیص یک کلمه معین، شبکه فرضیه های کاراکتر با احتمالات آن با استفاده از مدل های مختلف تولید می شود. شکل 10، HMM استفاده شده برای کلمه "ala" را نشان می دهد. تابع چگالی احتمال گسسته (PDF) در هر حالت از مدل کلمه نشان داده شده است. هر قسمت از PDF نشان دهنده احتمال مشاهده یک کاراکتر معین است. این موردتوسط ترکیب فراوانی های هجاهای بی معنی سه حرفی استخراج شده از فرهنگ لغت و امتیاز اعتماد OCR برای هر کاراکتر تعیین می شود. برای هر دنباله از شبکه، این سیستم، دنباله حالت احتمال حداکثر را با استفاده از الگوریتم Viterbi تعیین می کند که احتمال مشترک مشاهده و دنباله حالت را حداکثر می نماید.

Schwartz و همکاران، (53)، یک روش را برای استفاده مجدد از بسته تشخیص گفتار پیوسته موجود (BBN BYBLOS) در متن عربی چاپی معرفی نمودند. این ایده اینست که الگوریتم های HMM برای آموزش و تشخیص، از زبان مستقل هستند؛ تنها داده های آموزش واژگان ممکن است در کاربردهای خود تفاوت ایجاد نمایند. مشابه با گفتاری که در آن داده ها، یک گفته تک پیوسته است، تصویر خط متن به عنوان ورودی برای سیستم استفاده می شود. دنباله مشاهده از بردارهای ویژگی محاسبه شده به عنوان تابعی از موقعیت افقی در این خط تشکیل می شود (شکل 11 را ببینید). یک فریم تعریف شده به عنوان یک نوار عمودی باریک، با یک عرض که کسر کوچکی (به طور معمول در مورد 15/1) از ارتفاع خط است، به 20 سلول با هم متداخل برابر تقسیم می شود. اجزای بردار، مستقل از

زبان هستند و ویژگی های ساده مانند شدت محلی، مشتق عمودی و افقی شدت، شیب محلی، و غیره را برجسته می نمایند.

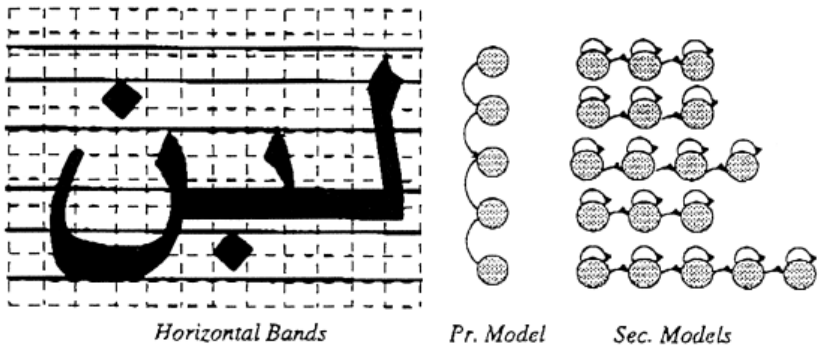
این سیستم از یک HMM راست به چپ برای هر کاراکتر استفاده می کند. یک مدل برای یک کلمه با الحاق مدل های کاراکتر آن به دست می آید. الگوریتم آموزش رو به جلو-رو به عقب برای استخراج برآوردهای احتمال حداکثر پارامترهای مدل استفاده می شود. این الگوریتم با تضمین توزیعات برای خوشه بندی حالات مشابه با آموزش ناکافی و بازآموزی وزن های حالات خوشه تقویت می شود. نویسندگان همچنین دو مدل را مورد استفاده قرار دادند: واژگان (به دست آمده با استفاده از یک مجموعه متن بزرگ) و مدل زبانی. مدل زبانی می تواند یک هجای بی معنی دو حرفی یا هجای بی معنی سه حرفی باشد که حاوی احتمالات کلمات در واژگان می شود. در مرحله تشخیص، یک الگوریتم جستجوی چند پاس (103) به جای الگوریتم Viterbi به دلیل وسعت فضای حالت استفاده می شود. این سیستم برای برای متون چند فونتی با یک پایگاه داده بزرگ عربی منجر به یک نرخ خطای کاراکتر میانگین 1.9٪ آزمایش شده است.

شکل. 11. استخراج ویژگی برای یک خط از متن.

Mahjoub (53) از یک طرح مرسوم و متداول برای تشخیص آنلاین کاراکترهای جداشده عربی پیروی نمود. هر کاراکتر توسط یک توالی مشاهده، مشخص شده از یک لیست از فواصل شعاعی نرمال نشان داده می شود. در طول آموزش، الگوریتم رو به جلو-رو به عقب برای استخراج حداکثر برآورد چگالی احتمال اعمال می شود.

مشابه با سیستم های تشخیص کاراکتر لاتین، استفاده از 1D-HMM، محدود به مشاهده توالی خطی، برای کلمات عربی، بهترین روش نیست. یک روش جایگزین، بزرگنمایی HMM با تعریف مدل های "دوبعدی" (106-104)

است. با این حال، این رویکرد منجر به پیچیدگی نمایی فرآیند تشخیص شد. Kuo و Aggazi (104) یک راه حل را با تقسیم تصویر به باندها (افقی یا عمودی) و مرتبط نمودن D-HMM1 برای هر باند برای این مسئله پیشنهاد دادند. علاوه بر این، مدل های چپ به راست (به نام ثانویه) در جهت افقی به دلیل ثبات کاراکترهای لاتین در این جهت ارائه می شود. علاوه بر این، مدل دیگری در جهت عمودی (به نام حالات فوق العاده اصلی) مشاهده تولید شده توسط مدل افقی را مرتبط می نماید. این معماری جذاب است، با این حال، در شمار مسئله استقلال خطی از حالت فوق العاده قرار می گیرد. چند راه حل در مراجع [104] و [105] پیشنهاد شده است. اولین راه حل، با اضافه کردن یک ماژول یک ماژول پس پردازش با در نظر گرفتن مدت زمان حل و فصل می شود، در حالی که راه حل دوم، محاسبه کلاس های معادل حالت فوق العاده است.



شکل. 12. PAW جدا شده به شش حالات فوق العاده.

در نهایت، Ben Amara و Belaid (55) از یک مدل سازی توزیع طول مدت حالت فوق العاده کارآمد استفاده نمودند. این کاربرد با زنجیره های متصل از کلمات چاپ شده عربی (PAW) سرو کار دارد. باندها با توجه به توپولوژی هر PAW تعیین می شوند. این سیستم از ویژگی های مختلف از جمله: ساقه، نقاط بالا، حلقه، نقطه پایین تر و descender (قسمتی از حروف کوچک که پایین تر از خط شروع می شوند)، و غیره استفاده می کند. (شکل 12). مشاهده برای مدل های ثانویه از بخش های پی در پی تشکیل شده است که در خطوط یافت می شوند و مدت زمان (طول) و محل آنها را کدنویسی می نمایند. در طول آموزش، تصویر به باندهای عمودی با توجه به مدت زمان و خطوط متعلق به حالات فوق العاده تقسیم می شود. توزیع احتمال مدت زمان توسط برآورد فراوانی ارتفاع باند برای حالت فوق العاده معین محاسبه می شود.

4. نتایج

در این مقاله، مسائل مربوط به کاراکترهای عربی چاپ شده و دست نوشته ارائه شد و بسیاری از پژوهش های مهم به طور خلاصه در تلاش برای ارائه وضعیت فعلی پژوهش تشخیص کاراکتر عربی شرح داده شدند. این هنوز هم یک حوزه تحقیقاتی باز است و هنوز هیچ سیستم تجاری عربی OCR در دسترس وجود دارد. این به دلیل مشکل تقسیم بندی است که در واقع شبیه به تقسیم بندی دست خط شکسته در بسیاری از زبان ها، و به دلیل پیچیدگی کاراکترهای عربی است. علاوه بر این، تمام الگوریتم های ارائه شده در این مقاله با متن بی صدا سرو کار دارند و تشخیص تفکیک واکه ها، یک حوزه پژوهش بسیار مهم در زبان عربی است.

با این حال، ارائه نتایج مقایسه ای برای روش پیشنهادی تا کنون بسیار دشوار است. بسیاری از روش های مورد استفاده برای دست خط عربی روی پایگاه های داده کوچک و مختلف ایجاد شده توسط تعداد کمی از مردم مورد آزمایش قرار گرفتند، زمانی که هیچ پایگاه داده عربی رایج در دسترس وجود ندارد. این برای همه روش های مورد استفاده برای کاراکترهای چاپ عربی هم درست است. به همین دلیل است که نتایج در بخش های قبلی که به شرح تکنیک ها اختصاص داده نشده اند، گنجانده نشدند.

همانطور که قبلا گفته شد، هیچ تکنیک محاسباتی حیاتی در این زمینه هنوز به طور کامل بررسی نشده است. به این ترتیب، این زمینه برای تحقیقات آینده از اهمیت زیادی برخوردار است.

REFERENCES

1. J. Kim and C. C. Tappert, Handwriting recognition accuracy versus tablet resolution and sampling rate, Proc. 7th Int. Conf. on Pattern Recognition, Montreal, pp. 917—918 (1984).
2. J. R. Ward and T. Kuklinski, A model for variability effects in handprinted with implication for the design of handwritten character recognition system, IEEE ¹rans. Man Cybernet. 18, 438—451 (1988).
3. F. Nouboud and R. Plamondon, On-line recognition of handprinted characters: Survey and beta tests, Pattern Recognition 25, 1031—1044 (1990).
4. C. Tappert, C. Y. Suen and T. Wakahara, The state of the art in one-line handwriting recognition, IEEE ¹rans. Pattern Anal Mach. Intell. PAMI-12, 787—808 (1990).
5. A. Amin and A. Kaced, Reconnaisances des caracters Arabes Manuscrits, Actes Congres AFCET Informatique Logiciel et materiel applications et implications, pp. 35—44 (1979).
6. A. Amin, A. Kaced, J. P. Haton and R. Mohr, Handwritten Arabic characters recognition by the IRAC system, Proc. 5th Int. Conf. on Pattern Recognition, Miami, pp. 729—731 (1980).
7. A. Amin and A. Shoukry, Topological and statistical analysis of line drawing, Pattern Recognition, ¸ett. 1, 365—374 (1983).
8. M. S. El-Wakil, On-line recognition of handwritten isolated arabic characters, Pattern Recognition, 22(2), 97—105 (1989).
9. T. S. El-Sheikh and S. G. El-Taweel, Real-time Arabic handwritten character recognition, Pattern Recognition 23, 1323—1332 (1990).
10. A. Amin, Issue on Arabic character recognition, Arabian J. Sci. Engng 18, 319—341 (1993).
11. A. Alimi and O. Ghorbe, The analysis in an on-line recognition system of Arabic handwritten characters, Proc. 3rd Int. Conf. on Document Analysis and Recognition, Canada, pp. 890—893 (1995). Off-line Arabic character recognition 527
12. A. Amin, IRAC: Recognition and understanding systems, in Applied Arabic ¸inguistic and Signal and Information Processing, ed., pp. 159—170. R. Descout, Hemisphere, New York (1987).
13. A. Amin, Machine recognition of handwritten Arabic word by the IRAC II system, Proc. 6th Int. Conf. on Pattern Recognition, Munich, Germany, pp. 34—36 (1982).
14. E. M. Riseman and R. W. Ehrich, Contextual word recognition using binary diagrams, IEEE ¹rans. Comput. c-20, 397—403 (1971).
15. A. Amin, G. Masini and J. P. Haton, Recognition of handwritten Arabic words and sentences, Proc. 7th Int. Conf. on Pattern Recognition, Montreal, pp. 1055—1057 (1984).
16. A. Amin and G. Masini, Machine recognition of cursive Arabic words, Application of Digital image Processing Vol. IV G. Tescher, ed., pp. 1127—1135 (1982).
17. A. Amin and G. Masini, Deux Methodes de Reconnaissance de Mots pour l'Ecriture Arabe Manuscrite, Act. 5'eme Congre's AFCET, Reconnaissance des Formes et Intelligence Artificielle, pp. 837—848, Grenoble 1985).
18. A. Amin, Arabic handwritten recognition and understanding, Proc. Computer processing of the Arabic ¸anguage, Kuwait, pp. 1—40 (1985).
19. S. Al-Emami and M. Usher, On-line recognition of handwritten Arabic characters, IEEE ¹rans. Pattern Anal Mach. Intell. PAMI-12, 704—710 (1990).
20. A. Belaid and G. Masini, Segmentation of line drawings for recognition and interpretation, ¹echnol. Sci. Informatics 1(2), 121—134 (1983).
21. C. Y. Suen, M. Berthod and S. Mori, Automatic recognition of handprinted characters, the state of the art, Proc. IEEE 68, 469—483 (1980).
22. J. Schuermann, Reading machines, Proc. 6th Int. Conf on Pattern Recognition, Munich, Germany, pp. 1031—1044 (1982).
23. J. R. Ullmann, Advance in character recognition, in Application of Pattern Recognition, K. S. Fu, ed., pp. 197—236 (1982).

24. M. K. Brown and S. Ganapathy, Preprocessing technique for cursive script word recognition, Pattern Recognition 19(1), 1—12 (1983).
25. R. H. Davis and J. Lyall, Recognition of handwritten characters a review, Image »ision Comput. 4, 208—218 (1986).
26. V. K. Govindan and A. P. Shivaprasad, Character recognition—A review, Pattern Recogn. 23, 671—683 (1990).
27. S. Srihari, From pixel to paragraphs: the use of models in text recognition, Proc. 2nd Ann. Symp. on Document Analysis and Information Retrieval, Las Vegas, USA, pp. 47—64 (1993).
28. E. Lecolinet and O. Baret, Cursive ¼ord Recognition: Methods and Strategies, Fundamentals in Handwriting Recognition, S. Impedovo,ed., pp. 235—263 (1994).
29. O. D. Trier, A. Jain and T. Taxt, Feature extraction methods for character recognition, Pattern Recognition 29, 641—662 (1996).
30. J. W. The and R. T. Chin, On image analysis by the methods of moments, IEEE [1]rans. Pattern Anal Mach. Intell. PAMI-10, 496—508 (1988).
31. S.J. Raudys and A. Jain, Small sample size effect in statistical pattern recognition, IEEE [1]rans. Pattern Anal Mach. Intell. PAMI-1, 252—264 (1991).
32. T. Matsunage and H. Kida, An experimental study of learning curves for statistical pattern classifiers, Proc. 3rd Int. Conf. on Document Analysis and Recognition, Canada, pp. 1103—1106 (1995).
33. M. Berthod and J. P. Maroy, Learning in syntactic recognition of symbols drawn on a graphic tablet, Compu. Graphiics Image Process. 9, 166—182 (1979).
34. A. Belaid and J. P. Haton, A syntactic approach for handwritten mathematical formula recognition, IEEE [1]rans. Pattern Anal. Mach. Intell. PAMI-6, 105—111 (1984).
35. R. Freund, Syntatic analysis of handwritten characters by quasi-regular programmed array grammars, in Advances in Structural and Syntactic Pattern Recognition, H. Bunke, ed., pp. 310—319 (1992).
36. Y. Lecun, Backpropagation applied to handwritten zip code, Neural Comput. 1, 541—551 (1989).
37. I. Guyon, Application of neural network to character recognition, in Character and Handwriting Recognition in Expanding Frontiers, P. S. P. Wang, ed., pp. 353—382 (1991).
38. S-W. Lee and Y-J. Kim, A new type of recurrent neural network for handwritten character recognition, Proc. 3rd Int. Conf. On Document Ananlysis and Recognition, Canada, pp. 38—41 (1995).
39. C. Y. Suen and C. L. Yu, Performance Accessmant of a character recognition Expert System, Int. Expert System application EXPERSYS 90, pp. 195—200 (1990).
40. L. Likfooman—Solem, H. Maiutre and C. Sirait, An expert and vision system for analysis of Hebrew characters and autheutication of manuscript, Pattern Recognition 24, 121—137 (1991).
41. A. Amin, M. Bemford and A. Hoffman, A knowledge acquisition technique for recognizing Hand-printed Chinese characters, Proc. 13th Int. Conf. on Pattern Recognition, Austria, pp. 254—258 (1996).
42. A. Amin, D. Ziino and C. Sammut, Recognition of hand printed Latin characters using machine learning, Proc. 3rd Int. Conf. on Document analysis and Recognition, Canada, pp. 1098—1102 (1995).
43. A. Amin and P. Compton, Hand-Printed characters recognition using machine learning, Proc. 5th Int. workshop on Frontiers in Handwriting Recognition, Essex, England, pp. 247—250 (1996).
44. A. Amin, C. Sammut and K. C. Sum, Learning to recognize Hand-printed Chinese characters using Inductive Logic Programming, Int. J. Pattern Recognition Artificial Intell. 10, 829—847 (1996).
45. K. Jambi, Arabic charcter recognition: Many approaches and one decade, Arab J. Sc. Engng 16, 499—509 (1991).
46. B. Al-Badr, and S. Mahmoud, Survey and bibliography of Arabic optical text recognition, Signal Process. 41, 49—77 (1995).
47. A. Amin, Arabic character recognition, in Handbook of Character Recognition and Document Image Analysis, H. Bunke and P. S. P. Wang, eds, pp. 349—420, world scientific (May 1997).
48. M. Khemakhem, Reconnaissance de caracters imprimes par comparaison dynamique, These de Doctorate de 3 e'me cycle, University of Paris XI (1987).

49. M. Khemakhem and M. C. Fehri, Recognition of Printed Arabic charcters by comparaison dynamique, Proc. 1st Kuwait Comput. Conf. pp. 448—462 (1989).
50. H. Y. Abdelazim and M. A. Hashish, Interactive font learning for Arabic OCR, Proc 1st. Kuwait Comput. Conf., pp. 464—482 (1989).
51. H. Y. Abdelazim and M. A. Hashish, Automatic recognition of handwritten Hindi numerals, Proc. 11th National Comput. Conf., Dhahran, pp. 287—299 (1989).
52. Z. Emam and M. A. Hashish, Application of Hidden Markov Model to the recognition of isolated Arabic word, Proc. 11th National Comput. Conf., Dhahran, pp. 761—774 (1989).
53. R. Schwartz, C. LaPre, J. Makhoul, C. Raphael and Y. Zhao, Language independent ocr using a continuous speech recognition system. Proc. 13th Int. Conf., on Pattern Recognition, Vol. C, pp 99—103, Vienna, Austria (1996). 528 A. AMIN
54. M. A. Mahjoub, Choix des parametres lies a l''apprentissage dans la reconnaissance en ligne des caractercs arabes par les chaines de markov cachees, in Forum de la Recherche en Informatique, Tunis, (Juillet 1996).
55. N. BenAmara and A. Belaid. Printed PAW recognition based on planar hidden Markov models. In 13th Int. Conf. on Pattern Recognition, Vol. B, Vienna, Austria, (1996).
56. A. Amin and G. Masini, Machine recognition of mutifonts printed Arabic texts, Proc. 8th Int. Conf. on Pattern Recognition, Paris, 392—395 (1986).
57. H. Almuallim and S. Yamaguchi, A method of recognition of Arabic cursive handwriting, IEEE, [1]rans. Pattern Anal. Mach. Intell. PAMI-9, 715—722 (1987).
58. A. Amin and J. F. Mari, Machine recognition and correction of printed Arabic text, IEEE [1]rans. Man Cybernet. 9, 1300—1306 (1989).
59. A. Amin and S. Al-Fedaghi, Machine recognition of printed Arabic text utilising a natural language morphology, Int. J. Man—Machine Stud. 35, 769—788 (1991).
60. T. El-Sheikh and R. Guindi, Computer recognition of Arabic cursive script, Pattern Recognition 21, 293—302 (1988).
61. F. El-Khaly and M. Sid-Ahmed, Machine recognition of optically captured machine printed Arabic text, Pattern Recognition 23, 1207—1214 (1990).
62. H. Abdelazim and M. Hashish, Arabic reading machine, Proc. 10th National Computer Conf. Riyadh, Saudi Arabia, pp. 733—740 (1988).
63. A. Shoukry, A sequential algorithm for the segmentation of typewritten Arabic digitized text, Arabian J. Sci. Engng 16, 543—556 (1991).
64. T. Wakayama, A core-line tracing algorithm based on maximal square moving, IEEE [1]rans. Pattern Anal. Mach. Intell. PAMI-4, 68—74 (1982).
65. V. Margner, SARAT—A system for the recognition of Arabic printed text, Proc. 11th Int. Conf. on Pattern Recognition, 561—564 (1992).
66. B. M. Kurdy and A. Joukhadar, Multifont recognition system for Arabic characters, Proc. 3rd Int. Conf. Exhibition on Multi-lingual Computing (Arabic and Roman Script), U.K. pp. 731—739 (1992).
67. A. Amin and H. Al-Sadoun, A segmentation technique of Arabic text, Proc. 11th Int. Conf. on Pattern Recognition, The Netherlands, pp. 441—445 (1992).
68. H. B. Al-Sadoun and A. Amin, A new structural technique for recognizing printed Arabic text, Int. J. Pattern Recognition Artificial Intell. 9, 101—125, (1995).
69. H. Freeman, On the encoding of arbitrary geometric configuration, IEEE. [1]rans. Electronic Comp. EC-10, 260—268 (1968).
70. A. Nouh, A. Sultan and R. Tulba, An approach for Arabic character recognition, J. Engng Sc. 6(2), 185—191 (1980).
71. B. Parhami and M. Taraghi, Automatic recognition of printed Farsi texts, Pattern Recognition 14, 395—403 (1981).

72. S. Saadallah and S. Yacu, Design of an Arabic character reading machine, Proc. Computer Process. Arabic ¸anguage, Kuwait (1985).
73. D. Forney, The Viterbi algorithm, Proc. IEEE 61, 268—278 (1973).
74. S. El-Dabi, R. Ramsis and A. Kamel, Arabic charcter recognition system: statistical approach for recognizing cursive typewritten text, Pattern Recognition 23, 485—495 (1990).
75. H. Al-Yousefi and S. S. Udpa, Recognition of Arabic characters, IEEE ¹rans. Pattern Anal. Mach. Intell. PAMI-14, 853—857 (1992).
76. A. Nouh, A. Ula and A. Sharaf-Edin, Boolean recognition technique for typewritten Arabic character set, Proc. 1st King Saud ºniv. Symp. on Computer Arabization, Riyadh, pp. 90—97 (1987).
77. S. Al-Fedagi and A. Amin, Automatic correction of spelling errors in Arabic, J. ºniv. Kuwait 19, 175—194 (1992).
78. F. Haj Hassan, Arabic character recognition, Arab School of Science and Technology. pp. 23—30 (1985).
79. F. Haj Hassan and W. Haj Ali, Printed Arabic text recognition, Arab. J. Sci. Engng 16, 511—518 (1991).
80. B. Al-Badr and R. Haralick, Segmentation-free word recognition with application to Arabic, Proc. 3rd Int. Conf. on Document Analysis and Recognition, Montreal, pp. 355—359 (1995).
81. M. Sano, T. Kosaki and F. Bouslama, Fuzzy structural approach for recognition of handwritten Arabic characters, Proc. Int. Conf. on Robotics »ision and Parallel Processing for Industrial Automation, Ipon, Malaysia, pp. 252—257 (1996).
82. F. Bouslama, Arabic character recognition by Fuzzy techniques, Proc. 5th European Congress on Intelligent Techniques and Soft Computing, Aachen, Germany (1997).
83. R. P. W. Duin, Superlearning and neural network magic, Pattern Recognition ¸ett. 15, 215—217 (1994).
84. A. Amin and H. Al-Sadoun, Arabic character recognition system using artificial neural network, Int. Workshop on Appl. of Neural Networks to Telecommunications, USA, pp. 99—105 (1993).
85. A. Amin and H. Al-Sadoun, Handprinted Arabic character recognition system, Proc. 12th Int. Conf. on Pattern Recognition, pp. 536—539 (1994).
86. A. Amin and H. Al-Sadoun, Handprinted Arabic character recognition system using an arficial neural network, Pattern Recognition 29, 663—675 (1996).
87. M. Altuwaijri, and M. Bayoumi, Arabic text recognition using neural networks, Proc. Int. Symp. on Circuits and Systems — ISCAS'94. pp. 415—418 (1994).
88. A. Amin and W. Mansoon, Recognition of printed Arabic text using Neural networks, Proc. 4th Int. Conf. on Document Analysis Recognition, Ulm, Germany, (August 1997).
89. L. R. Rabiner and B. H. Juang, An introduction to hidden Markov models, IEEE ASSP Mag. 3(1), 4—16 (1986).
90. A. B. Poritz, Hidden Markov models: a guided tour, Proc. IEEE, Int. Conf. Acoust. Speech, Signal Process. pp. 7—13 (1988).
91. L. R. Rabiner, A tutorial on Hidden Markov Models and selected application in speech recognition, Proc. IEEE 77(2) 257—286 (1989).
92. A. Kundu, Y. He and P. Bahl, Recognition of handwritten word: first and second order Hidden Markov models based approach, Pattern Recognition 22, 283—297 (1989).
93. M. Y. Chen, A. Kundu and S. N. Srihari, Unconstrained handwritten word recognition using continuous density variable duration hidden Markov models, Proc. IEEE, Int. Conf. on Acoust. Speech Signal Process. (ICASSP'93) (1993).
94. M. Y. Chen and A. Kundu, An alternative to variable duration HMM in handwritten word recognition, Proc. 3rd Int. Workshop on Frontiers in Handwritten Recognition, pp. 82—91 (1993).
95. M. Gilloux,J. M. Bertille and M. Leroux, Recognition of handwritten words in a limited dynamic vocabulary, Proc. 3rd Int. Workshop on Frontiers in Handwritten Recognition, pp. 417—422 (1993).
96. H. S. Park and S. W. Lee, Off-line recognition of large set handwritten Hangul (Korean script) with HMM, Proc. 3rd Int. Workshop on Frontiers in Handwritten Recognition, pp. 51—61 (1993). Off-line Arabic character recognition 529

97. M. Gilloux, Hidden Markov models in handwritten recognition, Fundamentals in Handwriting Recognition, S. Impedovo, ed., Springer-Verlag, pp. 264—288 (1994).
98. C. B. Bosc and S. Kuo. Connected and degraded text recognition using hidden Markov model, Pattern Recognition, 27, 1345—1363 (1994).
99. A. El-Yacoubi, J-M Bertille and M. Gilloux, Conjoined location and recognition of street names within a postal address delivery line, Proc. 3rd Int. Conf. on Document Analysis and Recognition, ICDAR'95, pp. 1124—1127. Montreal, Canada, (1995).
100. M.Y. Chen, A. Kundu, J Zhou, and S. N. Srigari, Off-line handwritten word recognition using hidden Markov mode, in USPS'92 (1992).
101. M. Chen, A. Kundu and J. Zhou, Off-line handwritten word recognition using HMM type stochastic network, IEEE ¹rans. Pattern Analysis Mach. Intell. PAMI 16, 481—496 (1994).
102. L. Rabiner and B-H. Juang, Fundamentals of Speech Recognition, Prentice-Hall, Englewood Cliffs NJ.
103. L. Nguyen et al., The 1994 bbn/byblos speech recognition system. Proc. ARPA Spoken ¸anguage ¹echnology ¼orkshop, pp. 77—81. Morgan Kaufmann, Austin, TX (1995).
104. S. Kuo and O. E. Agazzi, Keyword spotting in poorly printed documents using pseudo 2—D hidden Markov moddels, IEEE ¹rans. Pattern Analysis and Machine Intell. PAMI 16, 842—848 (1994).
105. M. Gilloux, Reconnaissance de chiffres manscrits par modele de Markov pseudo 2-D, in act du 3eme Colloque National sur l'Erit et le Document, 11—17, Rouen, France (1994).
106. G.Saon and A. Belaid, Recognition of unconstrained handwriten word using Markov random fields and HMM, Proc. 5th Int. Workshop on Frontiers in Handwritten Recognition (1996).