پردیس علوم
دانشکده ریاضی، آمار و علوم کامپیوتر

تصدیق هویت با استفاده از دست خط به صورت بر خط (Online Signature Verification) با استفاده از روش‌های یادگیری ماشین

نگارنده: محمد حسام شهریاری

استاد راهنما: دکتر باقر باباعلی

پایان‌نامه برای دریافت درجه کارشناسی
در رشته علوم کامپیوتر

مرداد ماه ۱۴۰۱

چکیده

این روز ها برای تمام امور روزمره انسان مانند ورود به یک مکان امن نیاز به تصدیق هویت داریم. برای این منظور. سطوح بالاتر امنیتی با تاثیر متقابل ساده تر مورد نیاز است که میتوان با استفاده از تصدیق هویت بیومتریک [1] به آن دست یافت تصدیق هویت بیومتریک با استفاده از ویژگی های فیزیکی و رفتاری افراد به ما در تصدیق هویت آنها کمک میکند. این ویژگی ها باید خاصیت های مشخصی مانند خاص بودن. ماندگاری. مقبولیت. قابل جمع آوری بودن و ... را داشته باشند. تصدیق هویت با استفاده از دست خط [2] یکی از روش های تصدیق هویت بیومتریک است که مشخص میکند که آیا یک دستخط صحیح بوده و یا جعل شده است. شناسایی دستخط امضا و تصدیق آن یک حوضه فعال برای تحقیق در سال های اخیر بوده است. سیستم های شناسایی امضا برای شناسیی فرد از بین تمام افراد موجود در یک مجموعه مورد استفاده قرار میگیرند در حالی که سیستم های تصدیق امضا برای تصدیق هویت یک فرد با مقایسه یک امضای خاص با امضای وی که از قبل موجود است مورد استفاده قرار میگیرند. .

[1] Bio-Metric verification
[2] Handwriting verification

فهرست مطالب

[۳] Smoothing
[۴] Convert to Grayscale
[۵] Noise Reduction
[۶] Rotation
[۷] Binarization
[۸] Cropping
[۹] Thinning
[۱۰] Normalization
[۱۱] Extended shadow-code
[۱۲] Interest point matching
[۱۳] Reciever Operating Characteristics
[۱۴] Area Under ROC Curve

[۱۵] Template Matching
[۱۶] Neural Networks
[۱۷] Support Vector Machines
[۱۸] Hidden Markov Model
[۱۹] Gaussian Mixture Model
[۲۰] K-Nearest Neighbour
[۲۱] Deep Learning

فصل ۱

مفاهیم مقدماتی

امضا برای تصدیق هویت افراد مورد قبول ترین وسیله از نظر اجتماعی و قانونی بوده و در نتیجه در سطوح بالایی مورد حمله و جعل قرار میگیرد. تصدیق امضا نقش مهمی در شناسایی امضای جعلی و کاربرد بیومتریک بازی میکند. بیومتریک ویژگی های خاص فیزیکی و رفتاری افراد با هدف شناسایی یا تصدیق هویت را اندازه گیری میکند. ویژگی های فیزیکی در شاخصه بیومتریک شامل iris ، هندسه دست، اثر چهره و اثر انگشت میشود که در بین این ها iris در گذر زمان تغییر نکرده و در نتیجه تنوع بین کلاسیک کمتری دارند. این حالات نیازمند سخت افزار خاص و پر هزینه ای برای اندازه گیری تصوری بیومتریک دارند. ویژگی های رفتاری در شاخصه بیومتریک شامل امضا، صدا، الگوی ضربه زدن به کلید و gait میباشند که پیشرفته ترین این ویژگی ها تکنولوژی های امضا و صدا میباشند. امضای دست نویس شده یک شاخصه شناخته شده بیومتریکی است. که یک برتری مهم امضای دست نویس شده نسبت به دیگر روش های تشخیص و تصدیق هویت این است که تنها زمانی میتوان از آن استفاده کرد که شخص هوشیار بوده و مایل به نوشتن باشد. در حالی که در تکنولوژی اثر انگشت شخص میتواند غیر هوشیار نیز باشد. تصدیق امضای دست نویس شده به دو حالت برخط و آفلاین تقسیم بندی میشود.

۱.۱ تاریخچه امضا

امضا به شکل امروزی خود دستخوش تغییر و تحولات بسیاری شده. که از سالیان قبل شروع میشود. به طور خلاصه امضا عبارت است از اسم یا علامتی که در انتهای اسناد و نوشته ها گذاشته میشود و بدین معناست که مفاد آن نوشته مستند به فعل و تایید امضا کننده میباشد و با اراده او انجام شده است. اثر انگشت افراد بیسواد در حکم امضای آنها است. همچنین در قانون مدنی و سایر قوانین تعریفی از امضا نشده است. به علت حمله اسکندر مقدونی به ایران و آتش سوزی تخت جمشید و پس از آن ویرانی پایتخت ساسانی به دست اعراب، از ایران باستان به جز کتیبه های

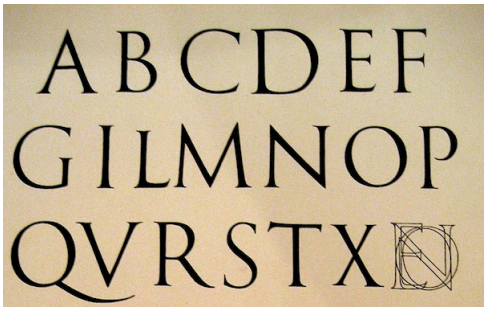
منتسب به شاهان هخامنشی و ساسانی اسناد و مدارک دیگری در دسترس نیست و کتیبه هایی که از شوش و همدان و تخت جمشید از داریوش اول، خشایارشا، اردشیر اول و داریوش دوم به دست آمده و نیز کتیبه های بیستون عموما فاقد اثر مهر و یا امضا هستند و ظاهرا علت مهمور نبودن کتیبه ها این است که پادشاهان صادر کننده فرمان سرلوحه کتیبه های سنگی و فلزی خود را به نام و سمت معرفی مینمودند و از طرفی نقش آثار مهر روی کتیبه های سنگی و فلزی ظاهرا خالی از اشکال است. به طور کلی میتوان گفت به مفهوم امروزی آثار و نمونه ای از امضا از ایران باستان در دسترس نیست اما مهر های مختلفی از زمان هخامنشی در موزه های بزرگ دنیا موجود است که به شکل استوانه ای میباشد و برای مهر کردن فرمان ها به کار میرفته است، این نوع مهر ها از سنگ های مختلف انتخاب و بر روی آن اسم و تصویر پادشاه حک گردیده است.

۱.۱.۱ امضا در فرهنگ های مختلف

مانند بسیاری از جنبه های مختلف جامعه مدرن امروزی ، امضای دست نویس شده نیز ریشه های خود را از روم باستان میگیرد. با این وجود که حروف چینی قدیمی ترین سیستم نوشتاری دنیا را تشکیل میدهند، که به ۲۰۰۰ سال قبل از میلاد مسیح بازمیگردد، امضا در چین بیشتر به شکل یک ضربه بوسیله یک ظرف جوهر و کمی خمیر قرمز انجام میشد. که افراد با حمل این ابزار سند های رسمی را امضا میکردند. در شکل ۱.۱ این ابزار نشان داده شده اند

شکل ۱.۱ : ابزار مورد نیاز برای امضا در چین[۳۶]

هم چنین مثال هایی از نسخه های خطی سومری که به ۳۱۰۰ سال قبل از میلاد مسیح بازمیگردند وجود دارند که از حروف و نماد های مختلف برای تصدیق هویت خود استفاده میکرده اند. با این حال این رومی ها بودند که پایه های دستخط و حروفی که امروزه از آن استفاده میکنیم را بنا نهادند.
دست نوشته های لاتین در حدود قرن هفتم قبل از میلاد مسیح تکامل پیدا کردند. هم چنین این دست نوشته ها با حروف بزرگ مرتب شده که به نام حروف بزرگ رومی شناخته میشوند نوشته میشدند. در شکل زیر ۲.۱ حروف بزرگ رومی که در کتیبه ها مورد استفاده قرار گرفته اند را میبینیم



شکل ۲.۱: حروف بزرگ رومی[۳۶]

همزمان با اینکه این الفبا در گرانیت و سنگ مرمر تراشیده میشد، این حروف در اسناد نوشتاری برای مکاتبه و اسناد دیگر نیز استفاده میشدند. تا حدود سال ۵۰۰ میلادی نسخه های اولیه حروف کوچک پدیدار شده و شروع به گسترش پیدا کردند.

در چند قرن بعدی دست نوشته ها تکامل پیدا کردند. سرانجام تبدیل به پایه ی دست نوشته ی ایتالیایی که در قرن پانزدهم با شکل خط شکسته ای که به آن ایتالیک[۱] گفته میشد شدند. که این موضوع منجر به شکل گیری سبک حروف ایتالیک شد که توسط طراحان در ونیز در اوایل قرن شانزدهم مورد استفاده قرار میگرفت که این اشکال امروزه نیز وجود دارند.

دست نوشته به شکل ابتدایی خود از این الفبا و دست خط ها الهام گرفته است. اشکال ظریف به صورت جدی در قرن شانزدهم و هفدهم میلادی توسعه پیدا کرد و مدارس خطاطی شروع به شکل گیری کردند. در حالی که متخصصان این مهارت را به شکل شگفت انگیزی بالا بردند، خوش خطی ماهرانه جزو پیش شرط های هر فرد تحصیل کرده ای بود.

در سال ۱۶۷۷ پارلمان انگلستان قانون Statue of Frauds را تصویب کرد، که منجر میشد تمام اسناد رسمی شامل معاملات دارایی ها، وصیت ها، اجاره نامه ها و ... نوشته و امضا شوند تا جلوی کلاهبرداری گرفته شود.

تا پیش از این قانون امضا مختص اشراف و پادشاهان بود. یکی از اولین امضا های ثبت شده در سال ۱۰۹۸ به واقعه ای اشاره دارد که در آن فرمانده ارتش اسپانیا El Cid سند وقف او به کلیسای جامع والنسیا را امضا میکند. در شکل زیر این امضا نشان داده شده است.

شکل ۳.۱: یکی ار اولین امضا های ثبت شده در تاریخ[۳۶]

[۱] Italic

تا زمانی که اعلامیه استقلال ایالات متحده آمریکا در سال ۱۷۷۶ نوشته شود، امضای دست نویس شده استاندارد طلایی مقید کردن قرارداد ها به صورت قانونی به شمار می آمد. مقصود امضا خاص بودن آن برای هر شخص بود. John Hancock ، یک تاجر ثروتمند و رییس دومین مجلس کنگره، تا به امروز به دلیل امضای زیبا و طولانی خود بر روی یکی از مهمترین اسناد تاریخ آمریکا به یاد مانده است. او اولین فردی بود که اعلامیه استقلال ایالات متحده را امضا کرد. در شکل زیر این امضا نشان داده شده است.

شکل ۴.۱: اولین امضای نوشته شده بر روی اعلامیه استقلال آمریکا[۳۶]

در نهایت در سال ۱۸۳۷ Samuel B. Morse اولین نسخه های سیستمی را سرهم بندی کرد که یک آهنربای الکتریکی در یک سمت سیمی را کنترل کرده و دنباله ای از کلیک های کوتاه و بلند را در این روند تولید کند. هر دنباله نمایانگر حرفی از حروف الفبا طبق تصویر زیر است.

A ●−	J ●−−−	S ●●●
B −●●●	K −●−	T −
C −●−●	L ●−●●	U ●●−
D −●●	M−−	V ●●●−
E ●	N −●	W ●−−
F ●●−●	O −−−	X −●●−
G −−●	P ●−−●	Y−●−−
H ●●●●	Q−−●−	Z −−●●
I ●●	R ●−●	

شکل ۵.۱: کد مورس[۳۶]

که این عمل شروع کننده برقراری ارتباط به صورت الکترونیک بود و طولی نکشید تا اینکه تلگراف به صورت گسترده مورد پذیرش قرار بگیرد. در ۱۸۶۹ امضای دیجیتال ارسال شده از این طریق به عنوان یک الزام آور به صورت قانونی در ایالات متحده مورد پذیرش قرار گرفت.

۲.۱ تشخیص امضا

امضای دست نویس شده یک شاخصه شناخته شده بیومتریکی است. که یک برتری مهم امضای دست نویس شده نسبت به دیگر روش های تشخیص و تصدیق هویت این است که تنها زمانی میتوان از آن استفاده کرد که شخص هوشیار بوده و مایل به نوشتن باشد. در حالی که در تکنولوژی اثر انگشت شخص میتواند غیر هوشیار نیز باشد تشخیص امضا شامل دو بخش است

- تصدیق امضا[۲]
- شناسایی امضا[۳]

۱.۲.۱ شناسایی امضا

در این حالت کاربر نمونه ای بیومتریکی از امضای خود ارائه داده و وظیفه سیستم تشخیص هویت کاربر از بین تمام کاربران موجود در سیستم است.

۲.۲.۱ تصدیق امضا

در این حالت کاربر مدعی داشتن یک هویت میشود و این ادعا را با فراهم کردن نمونه ای بیومتریک از امضای خود انجام میدهد. وظیفه سیستم در این حالت بررسی درست بودن این ادعا با مقایسه نمونه داده شده و نمونه بیومتریکی هویت مورد ادعا میباشد. که تصدیق امضا خود شامل دو بخش است

- تصدیق امضای برخط[۴]
- تصدیق امضای برون خط[۵]

۱.۲.۲.۱ تصدیق امضای برون خط

تصدیق امضا به صورت برون خط شامل یک مدرک است که در آن امضا وجود دارد. این امضا اسکن میشود تا بتوانیم به نمایش دیجیتالی تصویر آن دست پیدا کنیم.

---

[۲] Signature Verification
[۳] Signature Identification
[۴] Online Signature Verification
[۵] Offline Signature Verification

۲.۲.۲.۱ تصدیق امضای برخط

تصدیق امضا به صورت برخط از یک سخت افزار خاص استفاده میکند. مانند تبلت یا خودکار دیجیتالی مخصوص. که در حالت برخط شکل و پویایی حرکت قلم اندازه گیری میشود.

۳.۲.۱ روش های مختلف تصدیق امضا

تصدیق امضا به صورت برخط(یا پویا) از امضایی استفاده میکند که توسط یک تبلت حساس به فشار تولید شده که در آن ویژگی های پویای یک امضا علاوه بر شکل آن به دست می آید. ویژگی های پویا شامل تعداد و ترتیب حرکت قلم، سرعت نهایی امضا ،فشار قلم در هر مرحله و... میشود. که باعث میشود جعل امضا دشوار تر بوده و امضا خاص تر باشد. در یک سیستم تصدیق امضا به صورت برخط اشخاص ابتدا نمونه هایی از امضا را به صورت مرجع ارائه میدهند. هنگامی که یک کاربر امضایی را با هدف تصدیق هویت ارائه دهد(امضای تست) ، این امضا با امضای مرجع آن شخص مقایسه میشود. اگر عدم شباهت از مقدار خاصی بیشتر باشد، کاربر تصدیق نمیشود. در حین تصدیق هویت امضای تست با تمام امضا های مرجع آن شخص مقایسه میشود، که باعث میشود خروجی شامل مقادیر فاصله متعددی باشد. سیستم باید روشی را انتخاب کند که در آن این مقادیر فاصله متعدد در یک مقدار خاص که بیانگر میزان تفاوت امضای تست و مجموعه امضای مرجع است ترکیب شده، و با مقدار خاص از پیش تایین شده مقایسه شود تا هویت فرد تصدیق گردد. مقدار تفاوت ترکیبی میتواند مینیمم، ماکسیمم و یا میانگین مقادیر فاصله مختلف باشد.به طور خاص، یک سیستم تصدیق یکی از این حالت ها را انتخاب کرده و بقیه روش ها را کنار میگذارد. برای محاسبه کارایی یک سیستم پویا دو عامل مهم وجود دارند: FRR[۶] یک امضای درست و FAR[۷] یک امضای جعلی. از آنجایی که این دو خطا به صورت معکوس با یکدیگر رابطه دارند معمولا EER[۸] که در آن FAR با FRR برابر است گزارش میشود.

۳.۱ جعل امضا (Signature Forgery)

تصدیق امضای دست نویس شده به صورت برخط پروسه بررسی اصلی یا جعلی بودن یک امضا است. یک امضا به سادگی قابل جعل شدن است که امضا های جعل شده به سه دسته ساده، رندوم و ماهرانه تقسیم میشوند

[۶] False Rejection Rate
[۷] False Acceptance Rate
[۸] Equal Error Rate

۱.۳.۱ جعل امضای تصادفی

در این نوع جعل امضا جاعل بدون دانستن اسم نویسنده یا امضای صحیح امضای جعلی تولید میکند. در این نوع جعل امضا جاعل معمولا از امضای خود به عنوان امضای جعلی استفاده کرده که تشخیص جعلی بودن آن در بیشتر مواقع بسیار ساده است چرا که این امضا معمولا معنا و تصویر کاملا متفاوتی با امضای اصلی دارد.

۲.۳.۱ جعل امضای ساده

که در آن جاعل هیچ ایده ای نسبت به شکل امضایی که قصد جعل کردن آن را دارد، ندارد اما مشخصات شخصی که قصد جعل امضای او را دارد در اختیار دارد. این روش نسبت به بقیه روش ها ساده تر تشخیص داده میشود چرا که معمولا امضای جعل شده شباهت زیادی به امضای صحیح ندارد. در برخی مواقع نیز این روش جعل امضا به بررسی کننده درستی امضا کمک میکند تا شخص جاعل را با توجه به نوع دست خط و رفتار های نوشتاری در امضای جعلی شناسایی کند.

۳.۳.۱ جعل امضای ماهرانه

در این روش جاعل نمونه ای از امضایی که قصد جعل آن را دارد در اختیار دارد. کیفیت شبیه سازی بستگی به میزان تمرین جاعل از قبل، توانایی های جاعل و میزان توجه جاعل به جزییات در هنگام شبیه سازی امضا دارد. امضای جعل شده به صورت ماهرانه شباهت بیشتری به امضای اصلی دارد. مشکل تصدیق امضا هنگامی سخت و سخت تر میشود که از جعل ساده به جعل ماهرانه برویم. در حال حاضر درخواست رو به رشدی برای این وجود دارد که پروسه تصدیق هویت افراد سریع تر و دقیق تر شود. در نتیجه طراحی یک سیستم تصدیق امضا تبدیل به چالش مهمی شده است.

فصل ۲

روش ها و چالش های تصدیق امضا

۱.۲ انواع روش های استفاده شده

مقالات متعددی بر روی سیستم های تصدیق امضای دست نویس شده و روش های مورد استفاده آن ها انجام شده. اخیرا روش های متعددی به همراه مطالعات متنوعی بر روی استخراج ویژگی ها و طبقه بندی با استفاده از HMM[۱]، SVM[۲]، FFT[۳]، MLP[۴] و NN[۵] انجام شده است. تعداد زیادی از استراتژی های تطابق استفاده شده در آنالیز امضا ها تطابق هایی کلی نگر اند، تطابق ناحیه ای و تطابق چند ناحیه ای. بعضی از پراکنده ترین تکنیک های استفاده شده شامل فاصله اقلیدسی، تطبیق مرتجع[۶]، همبستگی ناحیه ای[۷]، تطبیق درختی [۸]، تطبیق با استراحت [۹]، انشعاب و ترکیب [۱۰]، تطبیق رشته ای [۱۱]، ،HMM NN و SVM میشوند.

Govindaraju Venu Lei، Hansheng مطالعه مقایسه مانندی را از ویژگی هایی که به طور معمول استفاده میشوند انجام داده اند. که با تعمیم مقادیر مدل وابسته به ویژگی از قبل موجود مدل با ثباتی را توسعه میدهند که

---

[۱] Hidden Markov Model
[۲] Support Vector Machine
[۳] Fast Fourier Transform
[۴] Multi Layer Perceptron
[۵] Neural Networks
[۶] Elastic Matching
[۷] Regional Correlation
[۸] Tree Matching
[۹] Relaxation Matching
[۱۰] Split And Merge
[۱۱] String Matching

مقادیر وابسته به فاصله را اندازه گیری کند. نتایج آزمایش ها نشان میدهند که مشخصات ساده ای مانند مختصات XY، سرعت نوشتن و زاویه با محور X جزو با ثبات ترین ویژگی ها هستند، همچنین نمونه گیری مجدد از دنباله ها به طور یکسان لزوما کارایی تصدیق هویت را بهبود نمیدهند[۳۱].

M.M.Fahmy Dr.Maged یک سیستم تصدیق امضای برخط را بر پایه استخراج ویژگی های روش تبدیلات موجک های گسسته(DWT) و طبقه بندی بر پایه شیکه عصبی پیش خور با خطای بازگشتی[۱۲] ارائه داده است. برای بالا بردن تفاوت بین یک امضای اصلی و جعلی، امضا در دامنه ی DWT تصدیق میشود. یک چند تطبیق دهنده شامل شش شبکه ی عصبی میشود که از نمایش های متعددی استفاده کرده و تطابق برای سیگنال بیومتریک ورودی یکسانی جهت تصدیق امضا استفاده میشود.
نرخ شناسایی برای هر کدام از شناسایی کننده های هر شبکه عصبی مورد بحث قرار گرفته و این نرخ ها با یکدیگر مقایسه میشوند. آزمایشاتی بر روی دادگان برای پنج کاربر که هر کدام شامل ۲۰ امضای صحیح و ۲۰ امضای جعل شده به صورت ماهرانه هستند انجام میشود. که نرخ شناسایی موفق امضا های صحیح به دست آمده برابر ٪۹۵ است.[۳۲]

Krinninger Sebastian Gruber، Thiemo Gruber، Christian روش جدیدی را برای تصدیق امضا به صورت برخط با استفاده از SVM بر پایه تابع هسته ای LCSS ارائه داده اند. که در این روش شباهت های دو سری زمانی به وسیله طول LCSS با استفاده از تابع هسته ای محاسبه میشوند. این تکنیک جدید نشان میدهد که LCSS، SVM میتواند در صورتی که شش امضای صحیح جهت آموزش استفاده شده باشند افراد را به صورت قابل اعتمادی شناسایی کند. مشخص میشود که ارزیابی شباهت امضای برخط بر پایه LCSS حتی نسبت به روش مشابه بر پایه DTW برتری دارد.[۷]

Sharma Abhishek و Sundaram Suresh رویکرد جدیدی بر پایه مدل را ارائه داده اند، این روش GMM در چهارچوب DTW را برای تصدیق امضای برخط استفاده میکند.
ابتدا، ویژگی های آماری وابسته به نویسنده برای تطبیق امضا استخراج میشود. سپس مشخصات یک Warping Path با استفاده از یک مشتق در ویژگی بر پایه Path Warping تجزیه و تحلیل میشوند که برای تصدیق امضا کاربردی است.
در انتها ادغامی از ویژگی بر پایه Path Warping پیشنهاد شده با امتیازDTW نرمال سازی شده برای بهبود کارایی تصدیق هویت سیستم بر پایه DTW انجام میشود. این روش جدید بر پایه مدل به طور موفقی بر روی دادگان MCYT پیاده سازی شده است و اولین روشی است که ویژگی های استخراج شده از GMM را در یک الگوریتم تطبیق DTW

[۱۲] Feed Forward Back Error Neural Network

برای بهبود تصدیق امضای برخط استفاده میکند.[۳۳][۳۴]

Gabriel مشکل آموزش سیستم های تصدیق امضای برخط زمانی که تعداد نمونه های آموزشی کوچک هستند را بررسی کرد، که در آن تعداد امضا های موجود به ازای هر کاربر محدود است. نه استراتژی طبقه بندی مختلف بر پایه ،GMM و UBM[۱۳] محاسبه میشوند. این مدل ها طوری طراحی شده اند که تحت شرایطی که در آن داده های نمونه کوچک اند کار کنند همچنین توسط سه آزمایش مختلف بررسی میشوند. کارایی این روش ها هنگامی که مجموعه آموزشی کمتر از نصف نمونه ها را شامل میشود (حدودا ۱۲ امضا به ازای هر کاربر) با سرعت بیشتری تنزل پیدا میکند. تصمیم گیری با تخمین زدن نرخ شباهت و مقایسه آن با آستانه تصمیم EER انجام میشود. دقت به دست آمده در مدل های GMM-SVM به طور قابل توجهی بهتر از مدل های GMM-UBM هنگامی که زیر مجموعه آموزشی موجود شامل حداقل ٪۵۰ کل دادگان میشود است.[۳۵]

drott Beatrice و Hassan-Reza Thomas رویکردی را ارائه داده اند که در آن طبقه بندی امضا های جعلی و صحیح در ابتدا توسط طبقه بندی دو دویی به وسیله ویژگی های مهندسی شده ساده انجام میشود، سپس به وسیله تکنیک های یادگیری ماشین مانند Logistic ،MLP regression و در نهایت توسط یک رویکرد یادگیری عمیق[۱۴] به همراه یک شبکه عصبی حلقه ای[۱۵] عمل طبقه بندی کامل میشود. رویکرد یادگیری عمیق در مسئله تصدیق امضا نتایج امیدوار کننده ای را نشان میدهد ولی همچنان نیاز به بهبود دارد.[۶]

۲.۲ روش حل مسئله

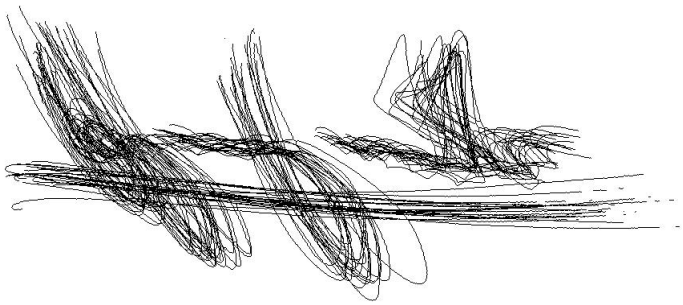
برای بررسی اصلی یا جعلی بودن یک امضا روش جدیدی به صورت تر کیبی از دو روش GMM[۱۶] و LCSS[۱۷] با استفاده از دیتابیس های امضای موجود به طور مثال (MCYT-۱۰۰) ارائه میشود. ابتذا امضا نرمال سازی شده و سپس پارامتر های GMM ب وسیله روش بیشترین شباهت[۱۸] تخمین زده میشوند. تکنیک روش تخمین بیشترین شباهت پارامتر هایی را پیدا میکند که شباهت ترکیبی از داده هایی که باید مستقل و یا به صورت یکسان توزیع شده باشند را بیشینه کنند. در ترکیب گاوسی ، این روش تغییر پذیری اساسی ویژگی های نقطه ای را به صورت آماری محاسبه میکند، که برای توصیف اثر امضا استفاده میشوند. سپس الگوریتم تشخیصی LCSS که شباهت

[۱۳] Universal Background Model
[۱۴] Deep Learning
[۱۵] Convolutional Neural Network
[۱۶] Gaussian Mixture Model
[۱۷] Longest Common Subsequence
[۱۸] Maximum Likelihood

سری زمانی امضا ها را محاسبه میکند استفاده میشود. پس از آن یک متغیز آستانه تنظیم میشود و با مقایسه مقادیر امضای تست با مقادیر امضا های دادگان تصمیم این که امضا اصلی و یا جعلی بوده گرفته میشود. برای محاسبه کارایی LCCS این روش با روش DTW[۱۹] که بیشیترین استفاده را در برابر دیگر روش ها دارد مقایسه میشود.

۱.۲.۲ چالش های مسئله

یکی از اصلی ترین چالش های مسئله تصدیق امضا تغییر پذیری درون کلاسی بالای آن است. مثلا در مقایسه با ویزگی های فیزیکی بیومتریکی، مانند اثر انگشت یا iris، امضای دست نویس شده یک شخص معمولا در بین نمونه های مختلف آن تفاوت و تغیر بالایی دارد که این موضوع در تصویر زیر نشان داده شده است.



شکل ۱.۲: تغییر پذیری امضای اشخاص در نمونه های مختلف[۵]

که این مشکل زمانی که جعل امضای ماهرانه را در نظر بگیریم با وجود تغییر پذیری کم بین کلاسی تشدید میشود. چرا که این نوع جعل امضا با هدف قرار دادن یک شخص خاص انجام میشوند، که در آن فردی با تمرین و تکرار سعی در تقلید امضای یک شخص دیگر را دارد. از این رو جعل امضای ماهرانه تا مقدار زیادی گرایش به نمایش امضای اصلی دارد.
چالش بزرگ دیگری در آموزش سیستم تصدیق امضا وجود دانش جزیی در حین آموزش است. در یک سناریوی واقع بینانه، در حین آموزش ما تنها به امضا های درست کاربران سیستم دسترسی داریم. اما در عمل ما نه تنها میخواهیم امضا های درست را شناسایی کنیم، بلکه امضا های جعلی را هم رد کنیم. که این موضوع چالش بر انگیزی است، چرا که در حین آموزش سیستم آگاهی ای از این که چگونه امضای اصلی را از جعلی متمایز کند ندارد.
در نهایت چالش دیگری که برای این مسئله وجود دارد این است که میزان داده ی موجود برای هر کاربر به نسبت کاربرد عملی بسیار محدود است. در مرحله ثبت نام معمولا از کاربران تعداد محدودی نمونه امضا درخواست

[۱۹] Dynamic Time Warping

میشود. به بیانی دیگر، حتی اگر تعداد زیادی کاربر در سیستم ثبت نام کرده باشند، طبقه بند باید برای یک کاربر جدید که تنها تعداد محدودی نمونه امضا از وی موجود است نیز بهینه عمل کند.

۲.۲.۲ اهداف مسئله

هدف این مقاله ارائه روشی برای تصدیق امضا های دست نویس به صورت برخط میباشد که شامل مراحل زیر است
۱.جمع آوری امضا های دست نویس به وسیله یک تبلت دیجیتال یا دستگاهی حساس به فشار به همراه یک قلم دیجیتال
۲.استخراج ویژگی ها از یک مرتب کننده[۲۰] وابسته به مدل، مثلا .GMM
۳.محاسبه شباهت امضا با استفاده از یکی از الگوریتم های تشخیص شباهت مثلا LCSS
۴.محاسبه ،EER FRR، FAR و منحنی های ROC[۲۱] برای اندازه گیری میزان کارایی تصدیق امضا
۵. محاسبه کارایی با مقایسه نتایج به دست أمده با روش های مختلف

[۲۰] Classifier
[۲۱] Receiver Operation Characteristics

فصل ۳

آزمایش ها و نتایج

در این بخش مروری بر روش های استفاده شده در مقالات مختلف و دادگان موجود برای مسئله داریم و در ادامه به مقایسه کارایی آنها با یکدیگر میپردازیم

۱.۳ دادگان

تعداد زیادی از مقالات و مطالعات انجام شده بر روی مسئله تصدیق امضا با دادگان شخصی انجام شده. که این موضوع مقایسه کار ها را کمی دشوار میکند،چرا که بهبود کارایی طبفه بندی ممکن است به روشی بهتر، و یا دادگانی ساده تر و مرتب تر نسبت داده شود. با این حال در دهه گذشته تعدادی از مجموعه دادگان برای کار تحقیقاتی و عموم ساخته و قرار داده شده اند. پروسه به دست آوردن تصویر امضا برای اکثر دادگان عمومی شامل گام های مشابهی میباشد. امضا های درست در یک یا چند مرحله جمع آوری میشوند، که از کاربران میخواهند چندین نمونه از امضای خود را ارائه دهند. کاربر فرمی را شامل چندین جای خالی دریافت کرده، و نمونه ای از امضای خود را در هر جای خالی قرار میدهد.
جمع آوری امضا های جعلی با این حال روند متفاوتی را دنبال میکنند: کاربران نمونه هایی از امضا های درست را در اختیار دارند و از آنها خواسته میشود این امضا ها را یک یا چند بار شبیه سازی کنند. همچنین توانایی کاربران در جعل امضا به صورت مبتدی یا حرفه ای کاملا بی اهمیت بوده. پس از این که فرم ها تکمیل شدند، اسکن شده و پیش پردازش روی آنها انجام میشود. در ادامه انواع دادگان موجود به صورت عمومی به همراه توضیح مختصری از أنها ارائه میدهیم.

- CEDAR Dataset

- MCYT Dataset
- GPDS signature
- UTSig
- SigComp۲۰۰۹
- SigComp۱۱
- BHSig۲۶۰ dataset
- SVC۲۰۰۴
- SUSIG
- ATVS Dataset

CEDAR Dataset: این دادگان شامل ۵۵ نویسنده میشود که به ازای هر نویسنده ۲۴ امضای صحیح و ۲۴ امضای جعلی به فرمت PNG و رنگ خاکستری وجود دازد.

MCYT: دو زیر مجموعه از دادگانMCYT وجود دارد که با نام هایMCYT-۱۰۰ (که شامل داده از ۱۰۰ نویسنده:۲۵ امضای صحیح و ۲۵ امضای جعلی به ازای هر نویسنده) و MCYT-۷۵ (که شامل داده از ۷۵ نویسنده: ۱۵ امضای صحیح و ۱۵ امضای جعلی به ازای هر نویسنده) شناخته میشوند.

GPDS: که یک دادگان برون خط اسپانیایی است. زیر مجموعه های زیادی از دادگان GPDS وجود دازد. به عنوان مثال GPDS-۱۰۰ شامل داده از ۱۰۰ کاربر است، یا GPDS-۱۵۰ شامل داده از ۱۵۰ کاربر است، و GPDS-۹۶۰ داده های ۹۶۰ نویسنده را در بر دارد (که دیگر برای عموم در دسترس نیست.)
GPDS-Synthetic شامل داده های ۴۰۰۰ نویسنده است . همچنین تمام این دادگان شامل ۲۴ امضای صحیح و ۳۰ امضای جعلی به ازای هر کاربر به صورت خاکستری و در فرمت PNG میباشند.

UTSig: (University of Tehran Persian signature Dataset) یک دادگان برون خط به زبان فارسی است که شامل داده از ۱۱۵ نویسنده میشود: ۲۷ امضای صحیح و ۴۵ امضای جعلی به ازای هر کاربر به صورت خاکستری و در فایل های TIF.

SigComp۲۰۰۹: یک مسابقه ی تصدیق امضا است که شامل دو دادگان به صورت برخط و برون خط است که هر یک از این دادگان مجموعه های آموزش و ارزیابی را در بر دارند. مجموعه آموزشی شامل داده از ۱۲ نویسنده است: ۵ امضای صحیح و ۵ امضای جعلی به ازای هر نویسنده. مجموعه ارزیابی شامل داده از ۱۰۰ نویسنده است: ۱۲ امضای صحیح و ۶ امضای جعلی به ازای هر نویسنده.

SigComp۱۱: شامل دو زیر مجموعه از دادگان میشود: مجموعه نمونه های امضا ی چینی و مجموعه نمونه های امضای آلمانی. شامل نمونه های برخط و برون خط به صورت تصویر رنگی RGB میباشد. همچنین تعداد امضا های برخط با تعداد نمونه های برون خط متفاوت است، علاوه بر آن تعداد نمونه های امضا به زبان چینی با تعداد نمونه ها به زبان آلمانی نیز متفاوت است.

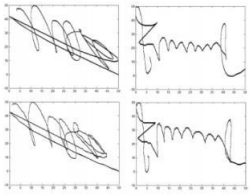
BHSig۲۶۰ signature Dataset: شامل نمونه امضای ۲۶۰ نویسنده میباشد ، ۱۰۰ تای آنها به صورت دادگان به زبان بنگلادشی و ۱۶۰ تای آنها به زبان هندی است. که هر کدام از آنها ۲۴ امضای صحیح و ۳۰ امضای جعلی به ازای هر نویسنده میباشند.

SVC۲۰۰۴: که اولین مسابقه بین المللی تصدیق امضا بود، شامل دو دادگان امضای دست نویس به صورت برخط میباشد. اولین دادگان تنها شامل اطلاعات به صورت مختصات و دومین دادگان شامل اطلاعات جانبی مانند جهت چرخش خودکار و فشار آن است. هر مجموعه دادگان شامل داده از ۱۰۰ نویسنده با ۲۰ امضای صحیح و ۲۰ امضای جعلی به ازای هر نویسنده است.

شکل ۱.۳: نمونه امضای جعلی و صحیح در دادگان SVC [۱]

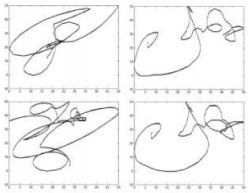
SUSIG: (Sabanci University Signature) به دو زیر مجموعه تقسیم میشود: بصری و کورکورانه. در زیر مجموعه بصری امضا ها بوسیله یک تبلت الکترونیکی مجهز به صفحه حساس به فشار که میتواند نویسنده امضا را در حین امضا زدن ببیند به دست می آیند. این مجموعه شامل داده از ۱۰۰ نویسنده با ۲۰ امضای صحیح و ۱۰ امضای جعلی(۵ امضای ماهرانه و ۵ امضای بسیار ماهرانه) به ازای هر نویسنده میباشد.

در زیر مجموعه کورکورانه امضا ها بوسیله یک تبلت و قلم الکترونیکی بدون مشاهده امضا کننده ها به دست می آیند. این دادگان شامل داده از ۱۰۰ نویسنده میباشد که ۱۰ امضای صحیح به ازای ۷۰ نویسنده، ۸ امضای صحیح به ازای ۳۰ نویسنده و امضا های جعلی مانند قسمت بصری میباشد. داده های ذخیره شده در هر دو دادگان شامل مختصات x-y میزان فشار و زمان هر امضا میباشد.



شکل ۲.۳: نمونه امضای صحیح و جعلی در دادگان SUISIG [۱]

ATVS Dataset: یک مجموعه دادگان امضا به صورت برخط است که شامل داده از ۳۵۰ کاربر با ۲۵ امضا به ازای هر کاربر میباشد.



شکل ۳.۳: نمونه امضای جعلی و اصلی در دادگان ATVS [۱]

۲.۳ اصول و روش ها

در این بخش به روش حل مسئله به صورت کلی پرداخته و سپس مرور کلی ای بر اجزای یک سیستم طبقه بند داریم

۱.۲.۳ پیش پردازش

پیش پردازش امضا های برخط معمولا برای حدف اختلافاتی صورت میگیرد که در فرآیند تصدیق بی اهمیت هستند. در مرحله پیش پردازش، امضا تحت فرآیند افزایشی ای جهت استخراج ویژگی ها قرار میگیرد. تصاویر

امضا نیازمند کمی دستکاری پیش از هرگونه عمل شناسایی میباشند. این فرآیند تصویر را آماده کرده و به جهت حذف اطلاعات بی اهمیت کیفیت را بالا میبرد همچنین انتخاب ویژگی های مهم تر را برای شناسایی و افزایش استحکام ویژگی هایی که باید انتخاب شوتد را بالا میبرد. علاوه بر این، مرحله پیش پردازش برای کاهش اختلال در تصاویر ورودی، و حذف تغیر در دست نوشته انجام میشود.
برای مقایسه فضایی یک امضا، وابستگی زمانی باید حذف شود. برخی نقاط یک امضا مانند نقطه شروع و پایان و نقاط تغییر مسیر، اطلاعات مهمی را در بر دارند. این نقاط، نقاط حساسی میباشند و در طی فرآیند استخراج شده و باقی می مانند.
در ادامه انواع روش های پیش پردازش تصدیق امضا و توضیح مختصری از آنها ارائه شده است.

۱.۱.۲.۳ هموار سازی[۱]

صفحات با کیفیت پایین مشکل گسسته بودن را دارند، که منجر به ناهمواری مسیر امضا میشود. استخراج ویژگی های موضعی از مسیر های ناهموار امضا و استفاده از آنها برای تصدیق امضا منجر به کارایی پایین امضا میشود. درنتیجه، هموار سازی برای تصاویر با کیفیت پایین در پیش پردازش ضروری است.

۲.۱.۲.۳ تغییر رنگ به خاکستری[۲]

اگر امضا ها به صورت رنگی باشند باید به طیف خاکستری رنگ تبدیل شوند. یک تصویر خاکستری تصویر ساده ایست که در آن تنها طیف رنگی موجود سایه های خاکستری رنگ می باشند. که در آن اطلاعات کمتری برای هر پیکسل تصویر مورد نیاز است. بیشتر منابعی که طیف رنگی را به خاکستری تبدیل کردند از فرمول زیر استفاده میکنند:

$$GrayColor = 0.299 * Red + 0.5876 * Green + 0.114 * Blue$$

---

[۱] Smoothing
[۲] Convert to Grayscale

شکل ۴.۳: تغییر رنگ امضا به خاکستری[۳۷]

۳.۱.۲.۳ کاهش اختلال[۳]

در بسیاری از موارد برخی از تصاویر ممکن است دچار اختلال باشند. روش های کاهش اختلال زیادی که بیشتر به صورت خطی اند وجود دارند(مانند میانگین گیری یا فیلتر های گاوسی[۴]) همچنین روش های غیر خطی ای (مانند میانه گیری یا فیلتر های درهم[۵]) نیز وجود دارند که در بسیاری از مطالعات از آنها استفاده میشود.

۴.۱.۲.۳ چرخش[۶]

برخی از امضا ها ممکن است در جهت حرکت عقربه های ساعت با زاویه خاصی چرخش داشته باشند بنابراین نیاز است که این تصاویر در خلاف آن جهت با همان زاویه چرخش داده شوند

---

[۳] Noise Reduction
[۴] Gaussian Filters
[۵] Fuzzy Filters
[۶] Rotation

۵.۱.۲.۳ دودویی سازی[۷]

یکی از گام های مهم در پردازش تصویر دودویی سازی آن است. این عمل تصویر را به سیاه و سفید(۰ یا ۱) با استفاده از الگوریتم هایی که به دو گروه تقسیم میشوند تبدیل میکند: دودویی سازی سراسری و دودویی سازی موضعی. چند مورد از روش های مهم دودویی سازی سراسری عبارتند از : Fixed Thresholding Method، Otsu Method، Kittler Method. که این روش ها از یک آستانه یکسان برای کل تصویر استفاده میکنند.
برخی از مهمترین روش های دودویی سازی موضعی عبارتند از: Niblack Method، Adaptive Method، Sauvola Method، Bernsen Method. این روش ها مقدار آستانه را به صورت موضعی و پیکسل به پیکسل محاسبه میکنند.

شکل ۵.۳: دودویی سازی امضا[۳۷]

۶.۱.۲.۳ برش[۸]

برش زدن عبارت است از حذف نواحی نامطلوب خارجی از یک تصویر که این عمل منجر به کاهش ابعاد آن تصویر میشود.

---

[۷] Binarization
[۸] Cropping

شکل ۶.۳: برش امضا[۳۷]

۷.۱.۲.۳ نازک کردن[۹]

نازک سازی به رویکرد عملیات ساختاری برای کاهش اندازه یک تصویر به کمترین مقدار ممکن گفته میشود. این عمل با استفاده از الگوریتم های متفاوتی همچون : Zhang Suen Thinning algorithm، Edge detection، Canny، Edge Based thinning algorithm، iterative algorithm using successive erosion، Optimized، Guo، Hall's parallel thinning algorithm انجام میشود.

۸.۱.۲.۳ عادی سازی[۱۰]

عادی یا نرمال سازی به فرآیندی گفته میشود که در آن تمام تصاویر دادگان به اندازه یکسان و ثابتی تبدیل میشوند. که این عمل در پیش پردازش عمل مهمی محسوب میشود.

۲.۲.۳ استخراج ویژگی ها

استخراج ویژگی ها شامل کاهش منابع مورد نیاز برای توصیف حجم زیادی از داده میشود. یکی از مسایل اساسی تحلیل داده پیچیده ریشه در تعداد متغیر های استفاده شده دارد. ویژگی های امضا بیانگر مقادیری است که می تواند از مسیر حرکت خودکار استخراج شوند، که این عمل با هدف توصیف امضا انجام میشود. ویژگی ها باید به ما اجازه دهند که بتوانیم بین امضا های کاربران مختلف تفاوت قایل شویم. ویژگی های امضا به دو گروه: موضعی و سراسری تقسیم میشوند.
ویژگی های سراسری به ویژگی هایی گفته میشود که از کل امضا استخراج می شوند. این ویژگی ها مقادیر متفاوتی از تصویر مانند شکل، رنگ و بافت را اندازه گیری میکنند. ویژگی های سراسری بیشتر در تصدیق امضا به صورت برون خط مورد استفاده قرار میگیرند. ویژگی های موضعی از قسمت های مختلف امضا استخراج میشوند. همچنین استفاده از ویژگی های موضعی در مقایسه با ویژگی های سراسری عملکرد بهتری دارند. چند مورد از ویژگی های

[۹] Thinning
[۱۰] Normalization

کلی ای که در برخی از سیستم های تصدیق امضای برخط مورد استفاده قرار میگیرند که معمولا به صورت پویا بوده و وابسته به عامل زمان اند عبارتند از: زمان امضا، تعداد دفعات برداشتن خودکار، مختصات امضا(xy) و مقادیر موقعیت خودکار، سرعت امضا و . ... همچنین در سال های اخیر علاقه به روش هایی که ویژگی ها را به صورت خودکار از روی تصویر امضا آموزش میبینند در مقابل طراحی روشی برای استخراج ویژگی ها مانند یادگیری عمیق رو به افزایش بوده است.
تلاش های زیادی برای یافتن ویژگی های بهینه در تصدیق امضا توسط محققان ارائه شده است که در ادامه به برخی از آنها میپردازیم.

۱.۲.۲.۳ ویژگی های هندسی

این ویژگی ها مشخصات هندسی تصویر امضا را توصیف میکنند.

۲.۲.۲.۳ تبدیلات ریاضی

مطالعات بسیاری به نوعی از یک یا چند تبدیل ریاضی برای استخراج ویژگی ها مانند Wavelet transform، Contourlet transform، Gabor wavelet transform، discrete Radon transform، transform Fourier استفاده میکنند

۳.۲.۲.۳ ویژگی های جهت دار

که تصویر را از نظر جهت ضربه امضا توصیف میکنند. این عمل با روش های مختلفی انجام میپذیرد

۴.۲.۲.۳ کد سایه قابل تعمیم[۱۱]

که این روش اطلاعات توزیع فضایی امضا را استخراج میکند

۵.۲.۲.۳ ویژگی های بافت امضا

که شامل نقاط پایانی، نقاط تقاطع و نقاط شاخه ای شدن امضا میشود. نقاط پایانی شامل نقطه شروع و پایان امضا میباشند. همچنین نقاط تقاطع به نقاطی گفته میشود که یک مسیر حرکت امضا با مسیر دیگر آن یکدیگر را منقطع کنند. نقاط شاخه ای نیز نقاطی هستند که یک مسیر حرکت امضا به دو شاخه تقسیم میشود.

[۱۱] Extended shadow-code

۶.۲.۲.۳ تطبیق نقطه مورد علاقه[۱۲]

مانند الگوریتم های SURF[۱۳] و SIFT[۱۴] که برای مسایل بینایی ماشین[۱۵] مورد استفاده قرار میگیرند

۳.۲.۳ آموزش مدل

در این بخش به اجزای یک سیستم تصدیق امضا، انواع روش های استفاده شده برای آموزش طبقه بند و برخی از مفاهیم مورد استفاده در مطالعات مختلف اشاره میکنیم. ابتدا مفاهیم ساده ای را تعریف میکنیم که در جداول و توضیحات نهایی مورد استفاده قرار میگیرند.

۱.۳.۲.۳ FAR (False Acceptance Rate)

به احتمالی که سیستم به صورت غلط یک مفدار ورودی را با یک الگوی موجود در دادگان تطابق دهد گفته میشود. این مقدار درصد ورودی های ناصحیحی را اندازه میگیرد که به طور غلط درست در نظر گرفته شده اند. یا به صورت خلاصه به درصد امضا های جعلی ای گفته میشود که صحیح در نظر گرفته شده اند. مقدار FAR از فرمول زیر محاسبه میشود:

FAR = ۱ـ کل امضا های صحیح / امضا های صحیح درست تشخیص داده شده

۲.۳.۲.۳ FRR (False Rejection Rate)

به احتمالی گفته میشود که سیستم نتواند یک ورودی را با الگوی آن در دادگان تطابق دهد. یعنی درصد ورودی های درستی را که به غلط رد شده اند محاسبه میکند. به طور خلاصه به درصدی از امضا های درست گفته میشود که جعلی در نظر گرفته شده اند. مقدار FRR از فرمول زیر محاسبه میگردد:

FRR = کل امضا های جعلی / امضا های جعلی که صحیح تشخیص داده شده اند

[۱۲] Interest point matching
[۱۳] Speeded Up Robust Features
[۱۴] Scale Invariant Feature Transform
[۱۵] Computer Vision

۳.۳.۲.۳ منحنی ROC[۱۶]

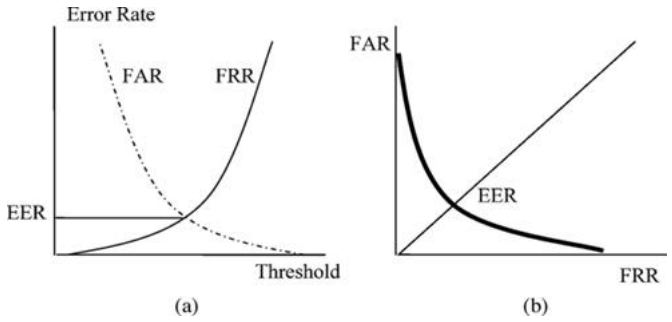
یک طبقه بند یک کلاسه میتواند توسط کسری از FRR و FAR محاسبه شود. منحنی ROC بیانگر تغییرات کسری از FAR با تغییر کسری از FRR است.

۴.۳.۲.۳ EER (Equal Error Rate)

بر مبنای منحنی ROC متغیر EER به گونه ای تعریف میشود که مقادیر FAR و FRR برابر باشند. به بیانی دیگر این متغیر تنها برای مقایسه دقت سیستم های مختلف مورد استفاده قرار میگیرد که دقت این سیستم ها برابر با عبارت زیر است:

دقت = ۱ - EER

همچنین در ادامه منحنی های مختلف مقادیر ذکر شده را در شکل های زیر نشان میدهیم که منحنی a بیانگر مقدار EER با استفاده از FAR و FRR بوده و منحنی b بیانگر مقدار EER با استفاده از منحنی ROC میباشد.



شکل ۷.۳: منحنی انواع خطا های استفاده شده در سیستم های تصدیق امضا[۴]

۵.۳.۲.۳ مساحت زیر منحنی ROC(AUC)[۱۷]

این متغیر تنها بیانگر عددی است که منحنی ROC را خلاصه کند.

[۱۶] Reciever Operating Characteristics
[۱۷] Area Under ROC Curve

۶.۳.۲.۳ ساختار یک سیستم تصدیق امضا

یک سیستم تصدیق امضا با توجه به مفاهیم ذکر شده از اجزای مختلفی تشکیل میشود که به این اجزا به ترتیب زیر اند:

- امضای ورودی
- پیش پردازش
- استخراج ویژگی ها با استفاده از تصویر امضا
- طبقه بندی و تصدیق
- نتیجه (اصلی یا جعلی بودن امضا)

که برای هر یک از بخش های فوق با توجه به نوع مسئله و دادگان موجود روش های متنوعی همان گونه که ذکر شد وجود دارد. در ادامه به برخی از آموزش دهنده های مورد استفاده برای تصدیق امضای برخط به همراه توضیح کوتاهی از آنها اشاره میکنیم.

۷.۳.۲.۳ تطبیق الگو[۱۸]

تطبیق الگو یک روش تشخیصی است که در آن نواحی ای از تصویر امضا که با الگوی موجود مشابهت دارند پیدا میشود. الگوریتم معمول استفاده شده برای تطبیق الگو DTW (Dynamic Type Wraping) است، با این حال الگوریتم های دیگری همچون mahalanobis distance و فاصله اقلیدسی نیز مورد استفاده قرار میگیرند.

۸.۳.۲.۳ شبکه عصبی[۱۹]

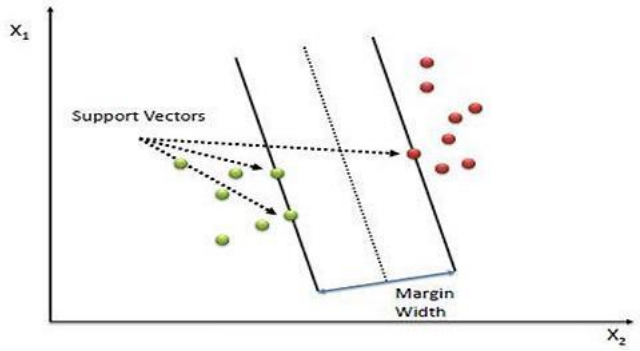
شبکه های عصبی در تشخیص الگو به صورت گسترده ای به دلیل قدرت پردازشی بالا، سادگی در استفاده و توانایی یادگیری آنها مورد استفاده قرار میگیرند. انواع معمول آن شامل پرسپترون چند لایه[۲۰] و شبکه عصبی معمولی میشوند. همچنین نوع دیگری از شبکه های عصبی وجود دارند که به آنها شبکه عصبی احتمالی[۲۱] گفته میشود که از عامل احتمال در طی فرآیند آموزش نیز استفاده میکنند. استفاده از انواع شبکه عصبی برای تصدیق امضا نیز بدین صورت است که ویژگی های استخراج شده امضا به شبکه عصبی برای آموزش داده شده و از نتیجه آموزش برای تشخیص الگو های امضا های مجموعه آموزشی استفاده میشود.

---

[۱۸] Template Matching
[۱۹] Neural Networks
[۲۰] Multi Layer Perceptron
[۲۱] Probablistic Neural Networks

۹.۳.۲.۳ SVM[۲۲]

SVM یک مجموعه روش یادگیری نظارتی[۲۳] است که در رگرسیون و طبقه بندی مورد استفاده قرار میگیرد. این روش از یک طبقه بند دودویی برای پیدا کردن مرز تصمیم گیری جهت جداسازی کلاس های مختلف استفاده میکند، با این حال در برخی موارد برای طبقه بندی چند کلاسه نیز مورد استفاده قرار میگیرد. در تصویر زیر ایده اصلی SVM به صورت ساده نمایش داده شده است.



شکل ۸.۳: SVM [۴]

۱۰.۳.۲.۳ HMM[۲۴] و GMM[۲۵]

HMM و GMM هر دو روش های آماری ای هستند که به صورت گسترده در شناسایی و تصدیق امضا مورد استفاده قرار میگیرند. در HMM ایده اصلی به صورت این است که تطبیق امضا با مدل رخ دهد. این عمل با استفاده از محاسبه احتمال توزیع ویژگی های امضا انجام میشود. GMM از چند توزیع احتمال گاوسی چند بعدی استفاده کرده و داده های با ابعاد کم را طبقه بندی میکند. به عبارتی میتوان آن را HMM با یک حالت در نظر گرفت.

۱۱.۳.۲.۳ KNN[۲۶]

الگوریتم KNN روشی برای طبقه بندی اشیا با استفاده از نزدیک ترین نمونه های آموزشی در فضای ویژگی ها است. این الگوریتم بیشتر برای شناسایی الگو به کار میرود و در مسئله تصدیق امضا با یافتن نزدیک ترین امضا ها به یک

---

[۲۲] Support Vector Machines
[۲۳] Supervised Learning
[۲۴] Hidden Markov Model
[۲۵] Gaussian Mixture Model
[۲۶] K-Nearest Neighbour

امضا در مرحله آموزش این کار را انجام میدهد.

۱۲.۳.۲.۳ روش های ساختاری

در این روش ها از الگو هایی همچون درخت، گراف یا رشته هایی برای نمایش الگوی امضا استفاده میشود. برای تصدیق امضا در این روش ساختار امضای ورودی با ساختار تمام امضا های موجود در دادگان مقایسه شده و اگر تفاوت ساختاری امضایی با امضای ورودی از حدی کمتر بود امضا تصدیق میشود.

۱۳.۳.۲.۳ یادگیری عمیق[۲۷]

در سال های اخیر تمایل به روش هایی که بتوانند اطلاعات را از داده های خام یادگرفته(پیکسل های تصویر) و بر اساس آن ویژگی های داده را مرتب کرده و یادگیری کنند رو به افزایش است. که این موضوع خود تفاوت اصلی روش های یادگیری عمیق با روش های معمول یادگیری ماشین است. روش های معمول یادگیری عمیق شامل DNN[۲۸]، CNN[۲۹]، RNN[۳۰]، GAN[۳۱]، AE[۳۲]، RBM[۳۳]، DBN[۳۴] و غیره میشوند.

۳.۳ نتایج

در این بخش به مقایسه انواع روش های استفاده تاکنون برای تصدیق امضای برخط پرداخته و به برخی از سوالات پاسخ میدهیم. در ادامه به جمع بندی و پیشنهاداتی برای ادامه کار اشاره میکنیم

[۲۷] Deep Learning
[۲۸] Deep Neural Networks
[۲۹] Convolutional Neural Networks
[۳۰] Recurrent Neural Networks
[۳۱] Generative Adversarial Networks
[۳۲] Auto Encoders
[۳۳] Restricted Boltzmann Machines
[۳۴] Deep Belief Networks

۱.۳.۳ جداول به دست آمده برای روش های مختلف

جدول ۱.۳: روش های مختلف استفاده شده در تصدیق امضای برخط

Features and Classifiers	Dataset User/Sig	Accuracy
BPNN, Probabilistic model and Fusion [8]	SVC2004	FRR=0.3 FAR=0.5
Dynamic Time Warping(DTW) [11]	SVC2004	FRR=5.5 FAR=4.13
Support vecctor Machines(SVM) based on LCSS kernel function [7]	SVC2004	ERR 6.84
Neural Networks Classifiersand Fuzzy Inference [16]	20/60	FRR=21.5 FAR=3.5, FRR=3.5 FAR=0.0
HMM/ANN [18]	MCYT	ERR 0.12
PWC, HMM [21]	MCYT	EER=6.67 EER=2.12
Bp ANN	150	FRR=1.8 FAR=2
Bayes Classifiers [23]	94/1247	FRR=2.19 FAR=3.5
HMM [25]	2/120 40/1440 2/1100	FRR=6.67 FAR=0.0 FRR=9.94 FAR=0.5 FRR=11.3 FAR=2.0
String Matching [21]	102/1232	FRR=2.8 FAR=1.6
PDF Classifiers [27]	5/25	FAR=5

جدول ۲.۳: روش های مختلف استفاده شده بر روی دادگان SVC و SUSIG

Method	SVC2004 EER(%)	SUSIG EER(%)
Gruber, et al. [7]	6,84	
Mohammadi and Faez [10]	6.33	
Barkoula, et al. [13]	5.33	
Yahyatabar, et al. [15]	5.33	
Khalil, et al. [30]		3.06
Napa and Memon [20]		2.91
Yeung, et al. [22]	2.89	
Fayyaz,et al. [24]	2.15	
Kholmatov and Yanikoglu [26]		2.10
Ansari, et al. [28]	1.65	1.23
Ibrahim, et al. [29]		1.59

جدول ۳.۳: انواع روش ها و ویژگی های مختلف تصدیق امضای برخط

Dataset	Extracted Features	Classifier	Evaluation
Dataset consists of 10 writers with 10 genuine signatures and 10 forged signatures per each user [9]	Extracted some features as(coordinates, pressure, altitude and azimuth) which are function of time t	DTW algorithm was used to calculate warping distance to differentiate between a genuine signature and its forgery and make a decision about verification result	The system can detect fake signatures with an accuracy of 90.4%
MCYT-100 [3]	Extract the local features	Gaussian Mixture Model(GMM) and Longest Common Subsequences(LCSS)	0.4% Equal Error Rate for GMM-LCSS model
Chinese dataset from SigComp2011 and MCYT-100 dataset [12]	Extracted 15 function features	DTW was used to compute distance values between query signature and all reference signatures in the template dataset	Achieved 1.69% and 1.77% EER for Chinese and MCYT-100 datasets respectively
Japanese online dataset from ICDAR2013 which contains data from 11 writers for training set and 20 writers for evaluation set [14]	extracted using a combination of: Fourier Transform based features, Wavelet Transform based features and Global features	correlation	27.48% FAR, 25.54% FRR and 73.49% accuracy

Dataset	Extracted Features	Classifier	Evaluation
ATVS, SVC, SUSIG datasets [1]	The features have been learned from ATVS dataset by using a sparse autoencoder with one hidden layer	One-class classifier	0.83% EER for SVC2004 dataset and 0.77% EER for SUSIG dataset
SVC [17]	Wavelet Transform	Neural Network(NN)	3.5% EER
MCYT-100 dataset and SVC2004 dataset [19]	Basic functions, Geometric normalization, Extended functions, Time derivatives, Signal normalization	Hidden Markov Models(HMM)	0.74% and 0.05% EER to skilled and random forgeries, respectively for MCYT dataset, 6.9% and 3.02% EER to skilled and random forgeries for SVC2004 dataset, respectively

۲.۳.۳ جمع بندی و نتایج

با توجه به جدول های ارائه شده شبکه عصبی در مقایسه با HMM برای دادگان ثابت بهتر عمل میکند. همچنین شبکه های عصبی احتمالی نیز در مقایسه با پرسپترون چند لایه بهتر عمل میکنند. در این مقاله مطالع ای بر روش های معمول استفاده شده برای تصدیق امضا به صورت برخط ارائه شد، همچنین روش های مختلف استخراج ویژگی ها و طبقه بندی نیز مورد بررسی قرار گرفت. به مشکلات و معضلات تصدیق امضا و همچنین تاریحچه امضا نیز اشاره شد. همچنین از آنجایی که امضای اشخاص در گذر زمان به دلایل مختلفی همچون سن، بیماری و شرایط روحی اشخاص بستگی دارد، توجه به تمام موارد و روش های مختلف یادگیری اهمیت بالایی دارد. در فرآیند یادگیری و آموزش مطالعات متنوع و زیادی انجام شده. با این حال با ایجاد تغییر در بخش های مختلف یک سیستم تصدیق امضا هنوز هم جای پیشرفت وجود دارد که شامل(و نه محدود به) موارد زیر میشوند:

- استفاده از ویژگی های بهتر
- بهبود عملکرد طبقه بند با کاهش نمونه های آموزشی
- تقویت کردن دادگان و استفاده از دادگان بهتر
- ساختن و استفاده از مدل های گروهی(ترکیبی)

۳.۳.۳ پیشنهادات برای ادامه کار

همان طور که عنوان شد، مطالعات زیادی در دهه ی اخیر بر روی روش های مختلف تصدیق امضای برخط و بهبود آنها انجام شده. با این حال تمام این روش ها جای بهبود و پیشرفت نیز دارند. به عنوان مثال روش های استخراج ویژگی ها میتوانند به صورت ترکیبی برای استخراج ویژگی های بهتر و قویتر مورد استفاده قرار گیرند. همچنین علاقه به حیطه روش های یادگیری عمیق نیز رو به گسترش است که این حوضه جای پیشرفت بالایی دارد. علاوه بر آن مدل های آموزشی نیز در مراحل نسبتا خامی قرار دارند و ترکیب و مقایسه آنها باعث بهبود عملکرد میشود.

۴.۳.۳ واژه نامه

Signature ——— امضا
Handwriting ——— دست نوشته
Verification Identity ——— تصدیق هویت
Identification Identity ——— شناسایی هویت
Classifier ——— طبقه بند

Biometric ——— بیومتریک
Ĭris ——— عنبیه چشم
gait ——— راه رفتن
Italic ——— خط شکسته ی ایتالیایی
Code Morse ——— کد مورس
Verification Signature ——— تصدیق امضا
Identification Signature ——— شناسایی امضا
Dataset ——— دادگان
Online ——— برخط
Offline ——— برون خط
Forgery Signature ——— جعل امضا
Preprocessing ——— پیش پردازش
Extraction Feature ——— استخراج ویژگی ها
Model Markov Hidden ——— مدل مخفی مارکوو
Machine Vector Support ——— ماشین بردار پشتیبانی
Transform Fourier ——— تبدیل فوریه
Transform Wavelet ——— تبدیل موجکی
Characteristics Operating Reciever ——— ویژگی های گیرنده عملیاتی
Perceptron Layer Multi ——— پرسپترون چند لایه
Networks Neural ——— شبکه عصبی
Matching Elastic ——— تطابق کشسان
Correlation Regional ——— همبستگی موضعی
Matching Tree ——— تطابق درختی
Merge and Split ——— انشعاب و ادغام
Matching String ——— تطابق رشته ای
Learning Deep ——— یادگیری عمیق
Network Neural Convolutional ——— شبکه عصبی حلقه ای

Network Neural Recrrent ——————————————————— شبکه عصبی تکرار شونده

Model Mixture Gaussian ——————————————————— مدل ترکیبی گاوسی

Sub-Sequence Common Longest ——————————————— طولانی ترین زیر دنباله مشترک

Likelihood Maximum ——————————————————— حداکثر شباهت

Warping Time Dynamic ——————————————————— پیچیدگی زمانی پویا

منابع

[1] Mohammad Hajizade Saffar, Mohsen Fayyaz, Mohammad Sabokru, Mahmmod Fathy, *Online Signature Verification using Deep Representation: A new Descriptor*, Arxiv, 2018

[2] VG Yogesh, *Online Signature Verification: A Survey*, NCRIET, 2015

[3] Shashidhar Sanda, Sravya Amirisetti, *Online Handwritten Verification System using Gaussian Mixture Model and Longest Common Sub-Sequenes*, Diva, 2017

[4] Nehal Hamdy Al-banhawy, Heba Mohsen, Neveen Ghali, *Signature Identification and Verification Systems: A Comparative Study on the Online and Offline Techniques*, Future Computing and Informatics Journal, 2020

[5] Luiz G. Hafemann, Robert Sabourin, Luiz S. Oliveira, *Offline Handwritten Signature Verification-Literature Review*, Arxiv, 2017

[6] Beatrice Drott, Thomas Hassan-Reza, *Online Handwritten Signature Verification using Machine Learning Techniques with a Deep Learning Approach*, 2015

[7] C. Gruber, T. Gruber, S. Krinniger, B. Sick, *Online Signature Verification With Support Vector Machines based on LCSS Kernel Functions*, IEEE, 2010

[8] Dr.Mohammad J. Alhaddad, *Multiple Classifiers to verify the Online Signature*, WCSIT, 2012

[9] Patil,B. V., Patil,P. R., *An Efficient DTW Algorithm for Online Signature Verification*, ICACCT, 2018

[10] M. H. Mohammadi, K. Faez, *Matching Between Important Points using Dynamic Time Warping for Online Signature Verification*, JBIO, 2012

[11] Ahmed Galib Reza, Hayotaek Lim, Md Jahangir Alam, *An Efficient Online Signature Verification Scheme Using Dynamic Programming of String Matching*, ICHIT, 2011

[12] Chen, Z., Xia, X., Luan, F., *Automatic Online Signature Verification based on Dynamic Function Features*, ICSESS, 2016

[13] K. Barkoula, G. Economou, S. Fotopoulos, *Online Signature verification based on Signatures Turning Angle Representation using Longest Common Subsequence Matching*, IJDAR, 2013

[14] Tahir M., Akram M. U., *Online Signature verification using Hybrid Features*, Infosec, 2015

[15] M. E. Yahyatabar, Y.Baleghi, M. R. Karimi, *Online Signature Verification: A Persian-language Specific Approach*, ICEE, 2013

[16] M. Khalid, Hamam Mokayed, R. Yusof, Osamu Ono, *Online Signature Verification with Neural Network Classifier and Fuzzy Inference*, AMS09, 2009

[17] M. R. Nilchiyan, R. B. Yusof, *Improved Wavelet-Based Online Signature Verification Scheme Considering Pen Scenario Information*, IEEE, 2013

[18] Zhong-Hua Quan, De-Shuang Huang, Kun-Hong Liu, Kwok-Wing Chau, *A Hybrid HMM/ANN based Approach for Online Signature Verification*, IJCNN, 2007

[19] J. Fierrez, J. Ortega-Garcia, D. Ramos, Gonzalez-Rodriguez, *HMM-based Online Signature Verification: Feature Extraction and Signature Modeling*, 2007

[20] S. B. Napa, N. Memon, *Online Signature Verification on Mobile Devices*, IEEE, 2014

[21] A.K. Jain, F.D Griess, S.D. Connell, *Online Signature verification*

[22] D-Y. Yeung, H. Chang, Y. Xiong, S. George, R. Kashi, T. Matsumoto, *SVC2004: First International Signature Verification Competition*, 2004

[23] Alisher Kholamatov, Berrin Yanikoglu, *Biometric Authentication Using Online Signatures*, ISCIS, 2004

[24] M. Fayyaz, M. H. Saffar, M. Sabokrou, M. Hoseini, M. Fathy, *Online Signature Verification based on Feature Represntation*, International Symposium on Artificial Intelligence and Signal Processing, 2015

[25] Mingfu Zou, Jianjun Tong, Changping Liu, Zhengliang Lou, *Online Signature Verification Using Local Shape Analysis*, Document Analysis and Recognition, 2003

[26] A. Kholamatov, B. Yanikoglu, *SUSIG: An Online Signature Database, Associated Protocols and Benchmark Results*, Pattern Analysis and Applications, 2009

[27] G. V. Kiran, R.S.R. Kunte, S. Samuel, *Online Signature Verification System using Probabilistic Feature Modeling*, Signal Processing and its Applications, 2001

[28] A. Q. Ansari, M. Hanmandlu, J. Kour, A. K. Singh, *Online Signature Verification using Segment-level Fuzzy Modelling*, Biometrics, IET, 2014

[29] M. T. Ibrahim, M. Kyan, G. Ling, *Online Signature Verification using Global Features*, CCECE, 2009

[30] M. I. Khalil, M. Moustafa, H. M. Abbas, *Enhanced DTW based Online Signature Verification*, IEEE, 2009

[31] H. Lei, V. Govindaraju, *A Comparative Strudy on the Consistency of Features in Online Signature Verification*, ICFHR, 2012

[32] M. M. Fahmy, *Online handwritten Signature Verification System based on DTW Features Extraction and Neural Network Classification*, Ain Shams engineering journal, 2010

[33] M. Faundez-Zanuy, *Online Signature Recognition based on VQ-DTW*, Pattern Recognition, 2007

[34] B. Kar, P. K. Dutta, T. K. Basu, C. VielHauer, J. Dittmann,*DTW based Verification Scheme of Biometric Signatures*, IEEE, 2006

[35] G. Zapata, J. D. Arias-Londono, J. Vargas-Bonilla, J. R. Orozco, *Online Signature Verification using Gaussian Mixture Models and Small-Sample learning Strategies*, Revista Facultad de Ingenieria, 2016

[36] Steve Moretti, *The Rise and Fall of Handwritten Signatures*, `https://www.stevemoretti.ca/post/the-rise-and-fall-of-hand-written-signatures`

[37] Prarthana Parmar, Jahnvi Mehta, Sakshi Sharma, Krupa Patel, Parth Singh, *A Survey of Handwritten Signature Verification System Methodologies*, JETIR, 2019

Abstract

Nowadays, human identifications are necessary for our routine activities such as entering any secure locations besides many other applications. To that end, higher security levels need with easier user interaction which can be achieved using bio-metric verification. Bio-metric verification helps us identify people based on their extracted physical or behavioural features. These features should have certain properties such as uniqueness, permanence, acceptability, collectability, and the cost to employ any bio-metric. Handwritten Signature Verification is one of the bio-metric verification which authenticates whether the signature is genuine or forged. Signature identification and its verification has been an active research topic in recent years. While identification systems are used to identify a person among all the users in a system, verification systems are used to verify an identity by comparing a special signature with the persons signature already available in the database.

College of Science
School of Mathematics, Statistics, and Computer Science

Online Signature Verification Using Machine Learning Approaches

Mohammad Hesam Shahriary

Supervisor: Bagher BabaAli

A thesis submitted in partial fulfillment of the requirements for
the degree of B.Sc. in Computer Science

August 2022