پردیس علوم
دانشکده ریاضی، آمار، علوم کامپیوتر

بررسی شبکه سیامی و بکارگیری آن برای تشخیص برخط امضای دستنویس

نگارنده:

آناهیتا دوستی سنجانی

استاد راهنما:

دکتر باقر باباعلی

پروژه برای دریافت درجه کارشناسی
در رشته علوم کامپیوتر

بهمن ۹۶

سپاس‌گزاری

سپاس فراوان از زحمات فراوان استاد گرانقدرم جناب آقای دکتر باباعلی

آناهیتا دوستی سنجانی

بهمن ۱۳۹۶

فهرست مطالب

چکیده ۵

۱ مقدمه **۶**

۱.۱ بیومتریک ۶

۱.۱.۱ ارزیابی عملکرد سیستم های بیومتریک ۸

۲.۱ تشخیص و بازشناسی امضا ۸

۳.۱ شناسایی الگو ۱۰

۲ مفاهیم مورد نیاز **۱۲**

۱.۲ شبکه های عصبی ۱۲

۱.۱.۲ معماری شبکه های عصبی ۱۳

۲.۱.۲ فرآیند آموزش شبکه عصبی ۱۴

۲.۲ یادگیری One-shot ۱۵

۱.۲.۲ تفاوت طبقه بندی استاندارد و One-shot ۱۵

۳.۲ شبکه های سیامی ۱۷

۱.۳.۲ معماری شبکه های سیامی ۱۷

۲.۳.۲ کاربرد شبکه سیامی در طبقه بندی One-shot ۱۹

۴.۲ مقالات مرتبط ۲۰

۱.۴.۲ شبکه های سیامی برای تشخیص تصویر One-shot ۲۰
۲.۴.۲ معماری سیامی بازگشتی برای یادگیری شباهت جملات ۲۲
۳.۴.۲ کاربرد شبکه سیامی در تشخیص امضا ۲۴

۳ دادگان امضا **۲۶**
۱.۳ دادگان SVC۲۰۰۴ . . . . . ۲۶
۱.۱.۳ ویژگی های SVC۲۰۰۴ . . . . . ۲۷
۲.۱.۳ نتایج بدست آمده روی SVC۲۰۰۴ . . . . . ۲۷
۲.۳ دادگان BiosecurID ۲۸
۱.۲.۳ ویژگی های BiosecurID ۲۸
۲.۲.۳ نتایج بدست آمده روی BiosecurID ۲۸

۴ سیستم های تشخیص امضا موجود **۲۹**
۱.۴ تشخیص امضا با شبکه های عصبی بازگشتی ۳۲

۵ تشخیص امضا با شبکه سیامی **۳۵**
۱.۵ شبکه سیامی بدون رمزگذاری ۳۶
۱.۱.۵ فرآیند کلی ۳۶
۲.۱.۵ پیش پردازش ۳۷
۳.۱.۵ معماری ۳۷
۴.۱.۵ پروتکل ۳۷
۵.۱.۵ نتایج ۳۹
۲.۵ شبکه سیامی با رمزگذاری ۴۱
۳.۵ شبکه سه تایی Triplet ۴۲

۶ نتیجه گیری **۴۳**

چکیده

شبکه سیامي Siamese از معماری خاصی در بین شبکه های عصبی مصنوعی برخوردار است. این شبکه بر خلاف سایر شبکه های مرسوم، از دو شبکه مشابه تشکیل شده و در نتیجه دو ورودی دارد. دو شبکه مشابه؛ خروجی خود را در اختیار یک لایه مشترک قرار داده و در نهایت از این لایه مشترک یک خروجی گرفته میشود. وظیفه شبکه این است که تشخیص دهد آیا دو ورودی از یک کلاس هستند یا دو کلاس مجزا. به این ترتیب در مرحله آموزش، شبکه میزان شباهت بین دو الگوی ورودی را یاد میگیرد و در زمان تست، با ارائه یک خروجی در بازه صفر تا یک، میزان شباهت دو الگوی ورودی را مشخص مینماید. در این پروژه، هدف بررسی عملکرد دقیق این شبکه و بکارگیری آن برای کاربرد تصدیق برخط امضای دستنویس با فرض بازنمایی امضا به دو صورت مختلف میباشد.

فصل ۱

مقدمه

پیشرفت در تکنولوژی در دهه های اخیر همانطور که بی شک در تمامی زمینه ها به امور سهولت بخشیده است، ضرورت روش های حفظ امنیت و تعیین هویت را بیش از پیش افزایش داده است. در این بین تشخیص و بازشناسی امضا همواره یکی از مهمترین و پرکاربردترین ابزار بیومتریک برای تشخیص هویت و کنترل دسترسی می باشد. در حالی که پارامتر های بیومتریک دیگر مانند اثر انگشت و عنبیه چشم در سال های اخیر محبوبیت رو به افزایشی برای تعیین هویت داشته اند، نمی توان منکر این مسئله شد که بسیاری از سیستم های موجود، از مهمترین آن ها سیستم های بانکی و دولتی، سال ها بر مبنای امضای افراد به فعالیت پرداخته اند، بنابراین تشخیص امضا با کمترین خطای ممکن بدلیل همه گیری و قدمت از اهمیت فراوانی برخوردار است.

۱.۱ بیومتریک

بیومتریک مجموعه ای از محاسبات و اندازه ها مربوط به خصوصیت شخصی یک انسان است که از آنها میتوان در اهراز هویت و تشخیص افراد برای مصارفی چون کنترل دسترسی

استفاده کرد. بدیهی است که ویژگی های بیومتریک مورد استفاده باید تا حد ممکن غیرقابل تغییر و غیر قابل تقلید باشند. خصوصیات بیومتریک به دو دسته کلی تقسیم می شوند:

- **خصوصیات فیزیکی:** خصوصیاتی هستند که در انسان از بدو تولد وجود داشته و در طی زندگی تقریبا به کلی ثابت هستند اعم از اثر انگشت، عنیبه، DNA و غیره.
- **خصوصیات رفتاری:** این ویژگی ها مربوط به رفتار هستند و از بین آنها می توان به صدا، ریتم تایپ کردن، دست خط و امضا اشاره کرد.

این ویژگی ها هر یک در کاربرد های مختلف از مزیت های منحصر به فردی برخوردارند. به طور کلی از ۹ پارامتر زیر می توان برای ارزیابی هر ویژگی بیومتریک استفاده کرد.

۱. عمومیت: اینکه هر شخص دارای آن ویژگی باشد.

۲. یکتایی: میزان تفکیک دهندگی در بین جمعیت (اثر انگشت کاملا یکتا است.)

۳. دوام: استناد به ویژگی برای چه مدت ممکن است.

۴. قابلیت ارزیابی: سهولت ارزیابی ویژگی برای افراد.

۵. کارایی: دقت، سرعت و پایداری روش های برپایه ویژگی مورد بررسی.

۶. مقبولیت: میزان پذیرش تکنولوژی

۷. جایگزینی: سهولت و هزینه جایگزین کردن ویژگی انتخاب شده در سیستم های موجود

۸. امکان تصدیق هویت: در تصدیق هویت مشخصه یک فرد به پایگاه اطلاعات ارسال می‌شود و هدف بررسی آن به منظور تصدیق هویت آن فرد می‌باشد که پاسخ سیستم الزاماً مثبت یا منفی است.

۹. امکان تشخیص هویت: در سیستم‌های تشخیص هویت مشخصه بیومتریک فرد به سیستم ارائه می‌شود و سیستم با جستجوی پایگاه اطلاعات مشخصات فرد را در صورت موجود بودن استخراج می‌کند.

طبق معیارهای بالا ویژگی امضا بدلیل استفاده گسترده فعلی از امکان مقبولیت و جایگزینی بسیار بالایی برخوردار است.

۱.۱.۱ ارزیابی عملکرد سیستم های بیومتریک

معیار های مشخصی وجود داردند که با آنها میتوان سیستم مبتنی بر بیومتریک را ارزیابی کرد. تعدادی از این معیارها به شرح زیرند.

- **نرخ قبولی غلط (FAR)**[۱]

احتمال اینکه سیستم نمونه ای را به اشتباه به عنوان نمونه ی حقیقی قبول کند.

- **نرخ رد غلط (FRR)**[۲]

احتمال این که سیستم نمونه ای صحیح را به اشتباه رد کند.

- **نرخ خطای مساوی (EER)**[۳]

نرخی که در آن خطای قبول و رد یکسان است.

۲.۱ تشخیص و بازشناسی امضا

تشخیص امضا که یک بیومتریک رفتاری است به دو صورت شکل می گیرد.

[۱] False Accept Rate

[۲] False Reject Rate

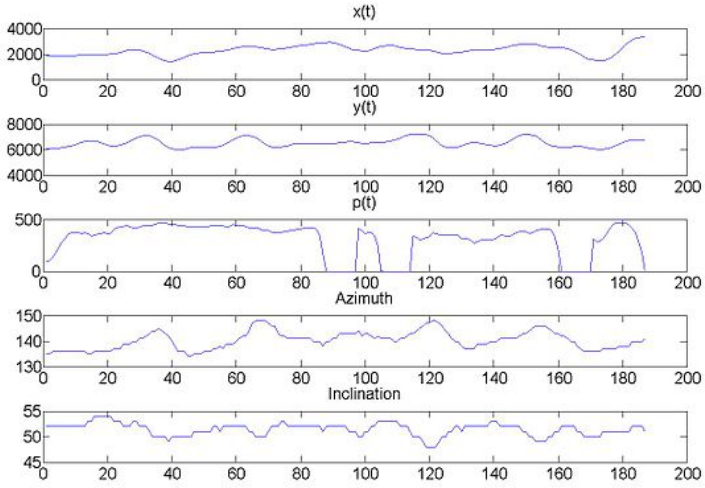
[۳] Equal Error Rate

- **استاتیک یا برون خط**

در این حالت، کاربر روی کاغذ امضا کرده، سپس این کاغذ توسط یک اسکنر نوری یا دوربین اسکن شده و وارد سیستم دیجیتال می شود و سیستم تشخیص امضا با بررسی و آنالیز شکل امضا، نتیجه خود را اعلام می کند.

- **دینامیک یا برخط**

در این حالت، کاربر امضای خود را توسط یک تبلت دیجیتال و با استفاده از قلم نوری ثبت می کند. این دستگاه دیجیتال امضا را به صورت بلادرنگ ثبت می کند. اطلاعات ثبت شده معمولا شامل سری زمانی های زیر می باشد.



شکل ۱.۱: نمونه ای از سری ها زمانی مربوط به امضای برخط

- مختصات در صفحه x(t) و y(t)
- فشار قلم p(t) [۴]
- زاویه قلم از جهت مثبت محور طول az(t) [۵]

[۴] pressure

[۵] azimuth

- زاویه قلم از جهت مثبت محور ارتفاع in(t) [۶]
- بالا یا پایین بودن سر قلم [۷]

برای تشخیص امضا از روش های شناسایی الگو استفاده می کنیم. پرکاربرد ترین این روش ها برای این کاربرد کشش زمانی پویا [۸] ، مدل مخفی مارکف[۹] و کوانتیزاسیون برداری[۱۰] هستند.

۳.۱ شناسایی الگو

شناسایی الگو، شاخه ای از هوش مصنوعی[۱۱] است که با طبقه‌بندی[۱۲] و توصیف مشاهدات سروکار دارد. شناسایی الگو به ما کمك می‌کند داده‌ها (الگوها) را با تکیه بر دانش قبلی یا اطلاعات آماری استخراج شده از الگوها، طبقه‌بندی نماییم. الگوهایی که می بایست طبقه‌بندی شوند، معمولا گروهی از سنجش‌ها یا مشاهدات هستند که مجموعه نقاطی را در یك فضای چند بعدی مناسب تعریف می‌نمایند. یك سیستم شناسایی الگوی کامل متشکل است از یك حسگر که مشاهداتی را که می‌بایست توصیف یا طبقه‌بندی شوند جمع‌آوری می‌نماید، یك سازوکار برای استخراج ویژگی‌ها که اطلاعات عددی یا نمادین را از مشاهدات، محاسبه می‌کند (این اطلاعات عددی را با یك بردار بنام بردار ویژگی‌ها نمایش می‌دهند) و یك نظام طبقه‌بندی یا توصیف که وظیفه اصلی طبقه‌بندی یا توصیف الگوها را با تکیه بر ویژگی‌های استخراج شده عهده‌دار است.

در سال های اخیر پیشرفت های زیادی در شاخه شبکه های عصبی از هوش مصنوعی صورت

[۶] inclination

[۷] pen up/down

[۸] Dynamic Time Warping (DTW)

[۹] Hidden Markov Model

[۱۰] Vector Quantization

[۱۱] Artificial Intelligence

[۱۲] Classification

گرفته است و روش های زیادی تحت عنوان یادگیری عمیق ارائه شده اند که از ابتدا در رابطه با شناسایی الگو مورد توجه بوده اند. با توجه به گستره ی وسیع این روش ها می کوشیم کارایی این روش ها را بر روی مسئله تشخیص امضا بسنجیم.

فصل ۲

مفاهیم مورد نیاز

در این بخش به معرفی مفاهیم تخصصی مورد استفاده در این پروژه، اعم از شبکه های عصبی، یادگیری One-shot و شبکه های سیامی می پردازیم.

۱.۲ شبکه های عصبی

[۱] یک شبکه عصبی که در لفظ اصلی شبکه عصبی مصنوعی نامیده (ANN)[۲] می شود را میتوان اینگونه تعریف کرد:

شبکه عصبی یک سیستم محاسباتی است که از تعدادی واحد محاسباتی ساده که به هم متصل اند تشکیل می شود. این واحد های محاسباتی یا نورون ها اطلاعات را وابسته به وضعیت پویای خود و پاسخ آن به ورودی های خارجی پردازش می کنند. [۱]

[۱] Neural Networks

[۲] Artificial Neural Network

۱.۱.۲ معماری شبکه های عصبی

شبکه های عصبی به طور کلی از اعضای زیر ساخته می شوند:

- **گره[۳] های ورودی (لایه ورودی)**

گره ها در واقع همان نرون های شبکه عصبی هستند. هیچ نوع محاسباتی در گره های ورودی انجام نمی شود. یک بلوک از گره ها را لایه می نامیم.

- **گره ها مخفی (لایه مخفی)[۴]**

این لایه جایی است که محاسبات میانی در آن صورت گرفته و اطلاعات را از لایه ورودی به لایه های بعدی انتقال میدهد. لایه مخفی می تواند از شبکه حذف شود که در آن صورت یک شبکه تک لایه خواهیم داشت.

- **اتصالات و وزن ها[۵]**

شبکه عصبی شامل اتصالاتی است که هر یک خروجی یک گره را به ورودی گره بعدی منتقل می کند. هر اتصال از گره i به گره j دارای یک وزن w_{ij} است.

- **تابع فعال سازی[۶]**

تابع فعال سازی یک گره، خروجی آن گره را بر اساس ورودی های آن مشخص می کند. این تابع بسته به نوع مسئله و موقعیت گره تعیین میشود و معمولاً برای تمامی گره های یک لایه یکسان است. برخی توابع فعال سازی خروجی صفر و یک دارند و برخی دیگر دارای خروجی پیوسته هستند. از انواع مختلف و پر کاربرد این توابع می توان به تابع پله ای، سیگموید، تانژانت هایپربولیک، Softsign و Rectified linear unit (ReLU) اشاره کرد. نکته مهم درمورد توابع فعال سازی ماهیت غیرخطی آنهاست. اگر چه استفاده از

[۳] node

[۴] Hidden layer

[۵] Connections and weights

[۶] Activation function

توابع خطی مشکلی در ساختار شبکه ایجاد نمی کند، اما آن را به یک تخمین زننده خطی محدود میکند. لذا در اکثر کاربردها همچون دسته بندی[۷] ، چنین شبکه ای قادر نخواهد بود در صورتی که داده ها خطی جدایی پذیر[۸] نباشند به نتیجه برسد. از آنجایی که تقریبا تمامی مسائل موجود دارای داده هایی با روابط غیرخطی هستند انتخاب توابع غیرخطی حداقل برای بخشی از شبکه ضروری است.

- **قاعده یادگیری[۹]**

قاعده یادگیری یک قاعده یا الگوریتم است که پارامترهای شبکه عصبی را تنظیم می کند تا شبکه بتواند از ورودی داده شده به خروجی مطلوب برسد. این تغییراتی که این الگوریتم در شبکه ایجاد می کند عملاً فرآیند یادگیری شبکه نام دارد و عملاً باعث تغییرات متوالی وزن های اتصالات بر طبق این الگوریتم میشود.

۲.۱.۲ فرآیند آموزش شبکه عصبی

به عنوان مثال یک مسئله طبقه بندی با ناظر [۱۰] را در نظر بگیرید. در داده های مربوط به مسئله، ما نمونه هایی به شکل $\left(x^{(i)}, t^{(i)}\right)$ داریم که در آن ها $x^{(i)}$ یک ورودی و $t^{(i)}$ طبقه یا دسته مربوط به آن است.

هدف شبکه عصبی در این مسئله این است که با دادن $x^{(i)}$ به شبکه به عنوان ورودی، $y^{(i)}$ را بطوری به عنوان خروجی شبکه محاسبه کند که مقدار آن تا جای ممکن به مقدار هدف یعنی $t^{(i)}$ نزدیک باشد تا خطا به کمترین حد ممکن برسد. در هر نرون رابطه بین ورودی و خروجی در صورتی که $x^{(i)} = \left(x_1^{(i)}, x_2^{(i)}, \cdots, x_n^{(i)}\right)$ و تابع فعال سازی آن نرون f باشد به صورت زیر است.

---

[۷] classification

[۸] Linearly separable

[۹] Learning rule

[۱۰] Supervised classification

$$y = f\left(\sum_{i=1}^{n} w_i x_i\right)$$
$$\epsilon = |y - t|$$

برای کمینه کردن خطا، با استفاده از یک قاعده یادگیری وزن های شبکه را در هر دور اجرا تغییر می دهیم. معروف ترین و پرکاربردترین قانون یادگیری مورد استفاده، قاعده ی کاهش گرادیان[۱۱] نام دارد.

۲.۲ یادگیری One-shot

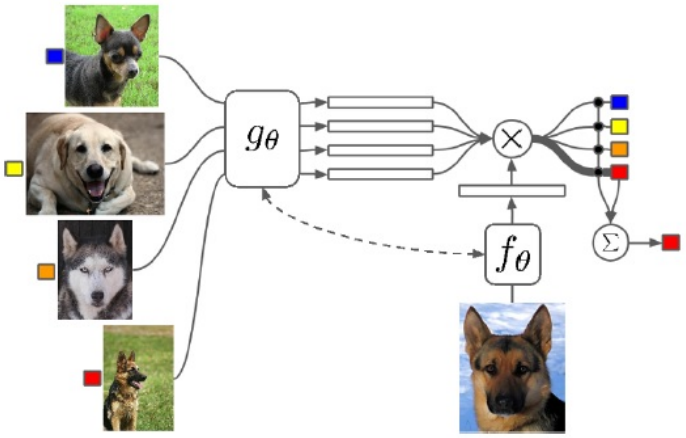
روش های طبقه بندی مبتنی بر داده، همواره وابستگی شدیدی به اندازه و توازن دادگان مورد استفاده اش دارد تا جامعیت طبقه بندی تضمین شود. این وابستگی در مواردی که چنین دسترسی ای موجود نیست و در بسیاری از کاربردهای واقعی، یک ضعف اینگونه روش هاست. به همین دلیل روش هایی که این وابستگی را به طرز مطلوبی کاهش دهند، همواره مورد توجه محققین بوده اند. از موفق ترین و جدیدترین این روش ها می توان از یادگیری one-shot نام برد.

۱.۲.۲ تفاوت طبقه بندی استاندارد و One-shot

یادگیری استاندارد[۱۲] روشی است که تقریباً در تمامی مدل های طبقه بندی مورد استفاده قرار می گیرد. در شبکه های عصبی طبقه بندی، ورودی به شبکه داده شده و احتمال تعلق آن به هر دسته مشخص از داده ها به عنوان خروجی محاسبه می شود. برای این کار در روش استاندارد، شبکه با داده هایی از هر دسته بندی آموزش داده می شود. لذا برای اخذ دقت مناسب لازم است به داده های آموزشی، به تعداد زیاد و مشابه نمونه هایی که بعدا توسط شبکه مورد بررسی قرار می گیرند، در اختیار داشته باشیم.

[۱۱] Gradient Descent

[۱۲] Standard Classification



شکل ۱.۲: مثالی از طبقه بندی One-shot . مدل تنها به یک تصویر از هر یک از ۴ دسته دسترسی دارد.

در مقابل، **طبقه بندی One-shot** [۱۳] برای آموزش، تنها به یک داده نمونه از هر دسته طبقه بندی نیاز دارد. این قابلیت شباهت زیاید به روند یادگیر در انسان دارد. به عنوان مثال، برای شناخت انواع حیوانات، معمولاً کافی است که فرد از هر نوع حیوان یک تصویر مشاهده کند و علیرغم محدود بودن این داده ها قادر خواهد بود نمونه های جدید از این حیوانات را از هم تشخیص دهد. این موضوع از اهمیت زیادی در یادگیری ماشین برخوردار است. فرآیند استاندارد برای تشخیص به تعداد زیادی داده از هر دسته نیاز دارد که در موارد بسیاری دسترسی به حجم کافی از داده ممکن نیست. به علاوه، در طبقه بندی One-shot ، هزینه و مدت زمان آموزش مدل به طرز قابل توجهی کاهش می یابد.

به عنوان مثال، برای تشخیص چهره، مدل طبقه بندی One-shot را روی مجموعه ای از تصاویر که افراد کمی را در زوایا، نور و شرایط متفاوت نشان میدهند، آموزش می دهیم. سپس برای تشخیص چهره شخص X در یک تصویر جدید (تصاویر شخص X در نمونه های آموزشی وجود ندارد)، کافی است یک تصویر از آن شخص را به مدل ارائه دهیم. مدل با داشتن تنها همین یک نمونه خواهد توانست وجود یا عدم وجود چهره شخص X را در تصویر ورودی تشخیص دهد.

[۱۳] One-shot Classification

۳.۲ شبکه های سیامی

شبکه های سیامی [۱۴] نوع خاصی از شبکه های عصبی هستند. هدف این نوع شبکه تمیز دادن بین داده هاست، به اینصورت که با دریافت دو نمونه میزان مشابهت آنها را گزارش میدهد. این نوع شبکه اولین بار توسط Yann LeCun به هدف تشخیص امضای غیربرخط ارائه شد.

۱.۳.۲ معماری شبکه های سیامی

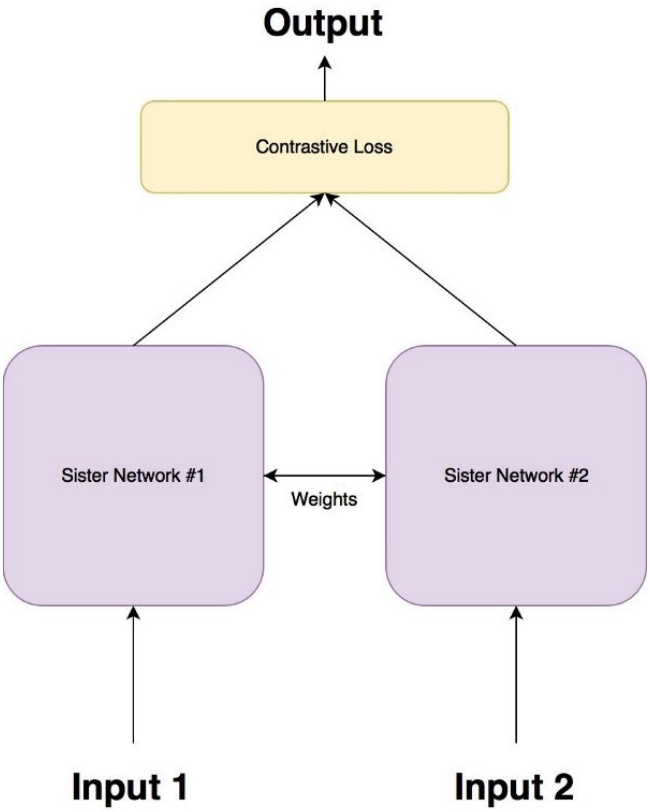
یک شبکه سیامی از دو شبکه عصبی کاملا یکسان (معماری این دو شبکه یک نوع خاص محدود نمی شود و در موارد مختلف از ساختارهایی مانند CNN LSTM، MLP، و غیره استفاده شده است) تشکیل شده که هر یک، یکی از دو نمونه ورودی را دریافت می کنند. سپس خروجی آنها به بخش نهایی داده میشود که وظیفه مقایسه دو نمونه را دارد. این لایه با استفاده از یک متریک مشخص، مقداری معرف شباهت دو نمونه محاسبه می کند. نکته مهم این است که دو زیرشبکه وزن های خود را به اشتراک می گذارند. این موضوع به سهولت و سرعت فرآیند آموزش کمک می کند.

آموزش این شبکه از طریق پس انتشار خطا [۱۵] انجام میشود.

از آنجایی که وزن های دو زیر شبکه باید در تمامی زمانها با یکدیگر برابر باشند، تابع خطا باید متقارن باشد تا در پس انتشار، تغییر وزن ها برای هر دو زیرشبکه یکسان باشند.

[۱۴] Siamese Networks

[۱۵] Error Backpropacation



شکل ۲.۲: شمای کلی شبکه سیامی

دو ویژگی مهم این شبکه به شرح زیر است:

- **انسجام پیش بینی**[۱۶] :

اشتراک وزن ها در زیر شبکه ها باعث می شود که نگاشت دو نمونه بسیار شبیه، به نقاط بسیار متفاوت در فضای ویژگی، [۱۷] توسط شبکه های نظیرشان ممکن نباشد زیرا دو زیرشبکه عملاً یک تابع را محاسبه می کنند.

- **تقارن شبکه**[۱۸]:

تقارن بدین معنی است که با ارائه دو نمونه متفاوت با هر ترتیبی، شبکه یک مقدار شباهت

[۱۶] Consistency

[۱۷] Feature Space

[۱۸] Symmetry

را محاسبه می کند. چنین ویژگی ای از ماهیت متقارن تابع متریک برگزیده برای شبکه و اشتراک وزن ها نتیجه می شود.

۲.۳.۲ کاربرد شبکه سیامی در طبقه بندی One-shot

توانایی شبکه های سیامی در تمیز دادن نمونه ها از دسته های متفاوت، آنها را به ابزار موثری در طبقه بندی بدل می کند. در طبقه بندی One-shot ، از هر دسته یک داده نمونه موجود است. اگر بتوان شبکه سیامی را به درستی آموزش داد میتوانیم با مقایسه ورودی با نماینده هر دسته تشخیص دهیم که ورودی به کدام دسته تعلق دارد.
در سالهای اخیر، این نوع استفاده از شبکه سیامی در طیف گسترده ای از مسائل در مقالات زیادی بررسی شده اند، که در ادامه به تعدادی از آنها خواهیم پرداخت.
بطور کلی در این مقالات فرآیند طبقه بندی به صورت زیر است:

۱. **Verification Tasks (training)**

این بخش که در واقع مرحله یادگیری شبکه است. در این بخش نمونه های آموزشی که از دسته آنها مطلعیم به صورت دو به دو به عنوان ورودی به شبکه داده می شوند و تمایز آنها محاسبه میشود. در فرآیند آموزش سعی داریم در صورتی که هر دو نمونه از یک دسته باشند مقدار خروجی را بیشینه و در غیر این صورت کمینه کنیم. لازم به ذکر است که بهتر است بخشی از داده های آموزشی به عنوان داده های validation مورد استفاده قرار گیرند.

۲. **One-shot Tasks (test)**

در این بخش، با فرض وجود N دسته $C_1, C_2, \cdots, C_N$ برای طبقه بندی هر نمونه آزمایشی احتیاج به N آزمایش داریم. به این صورت که مجموعه زیر از داده ها موجود

است:

$$S = \{(X_1, y_1), (X_2, y_2), \ldots, (X_N, y_N)\}$$

که X_i و y_i ها به ترتیب یک نمونه و نام دسته آن است. بطوریکه $\forall(X_k, y_k)\ X_k \in C_k$. برای طبقه بندی نمونه آزمایشی $\hat{x}$ آن را با تمام اعضای S به شبکه می دهیم. دسته $\hat{x}$ بر اساس این آزمایش ها دسته ای است که بیشترین شباهت را تولید می کند.

$$\hat{y} = argmax_C\ y_{out}(\hat{x})$$

به فرآیند فوق یادگیری n One-shot جهته [۱۹] می گوییم.

توجه کنید که اعضای کلاس ها در این بخش در مرحله verification حاضر نیستند.

۴.۲ مقالات مرتبط

۱.۴.۲ شبکه های سیامی برای تشخیص تصویر One-shot

[۲۰][۶] در این مقاله، کوش از شبکه های سیامی روی تصاویر استفاده می کند. داده های مورد بررسی مجموعه داده Omniglot است که شامل حروف دست نوشته ۵۰ الفبای مختلف است.

روش ارائه شده در این مقاله از این جهت مورد توجه است که میتواند با یک معماری سیامی ژرف و پیچیده قادر به تشخیص درست 92.0% است. این در حالی است که دقت انسان 95.5% و دقت روش HBPL 95.2% تنها ارقام بالاتر هستند. با توجه به اینکه HBPL یک روش ابداع شده برای داده های Omniglot است که از روش های استخراج ویژگی مشخص استفاده می کند، می توان گفت که شبکه سیامی بدلیل کلی بودن و عدم نیاز به مشخصات مانند HBPL

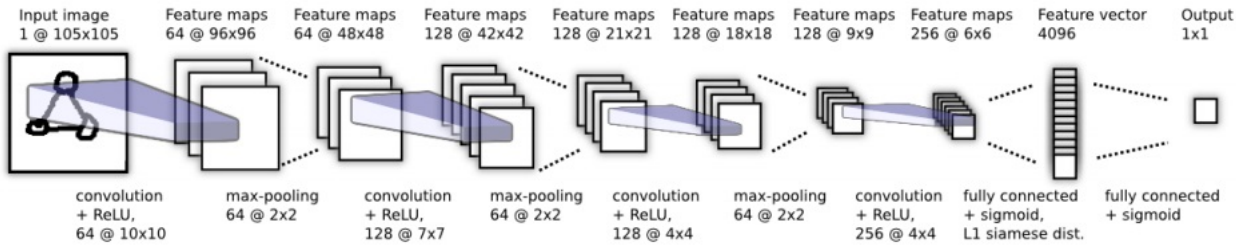
[۱۹] n-way one-shot learning

[۲۰] Siamses Networks for One-shot Image Recognition

دارای اهمیت بسیار است.

مقاله از فرآیند کلی ذکر شده در بالا پیروی می کند. فرض مهم این روش اینست که **اگر شبکه بتواند تفاوت دسته را در جفت تصاویر در مرحله verification تشخیص دهد، آنگاه خواهد توانست در مرحله One-shot نیز به درستی عمل کند.** از آنجایی که شبکه هیچ یک نمونه های مرحله One-shot را ندیده است، برای عملکرد بهتر توقع میرود که نمونه های مرحله verification دارای گوناگونی کافی باشند تا شبکه با آموزش بر روی آنها بتواند به معیارهایی کلی [۲۱] برای تشخیص دیگر انواع داده دست یابد.

مدل ارائه شده یک شبکه سیامی L لایه است که در آن دو زیر شبکه، شبکه عصبی پیچیده [۲۲] هستند. $h_{1,l}$ و $h_{2,l}$ بردار های مرتبط به لایه پنهان [۲۳] l-ام اولین و دومین زیر شبکه هستند. هر بخش پیچیده شامل دنباله ای از لایه های پیچیده است که از یک کانال [۲۴] با اندازه های



شکل ۳.۲: ساختار هر زیرشبکه پیچیده

متفاوت و گام ۱ استفاده می کنند. تعداد لایه ها به دلایل بهینه سازی از مضرب ۱۶ است. تابع فعال سازی استفاده شده برای این بخش ها ReLU است که گاهاً یک لایه max-pooling به گام ۲ به دنبال دارد.

لایه نهایی شبکه پیچیده مسطح میشود و به عنوان تک بردار به لایه بعد به صورت fully-

[۲۱] generalized

[۲۲] Convolutional Neural Networks (CNN)

[۲۳] Hidden Layer

[۲۴] channel

connected متصل می شود. در این لایه دو بردار بدست آمده از دو زیرشبکه تحت فاصله L_1 و تابع سیگموید قرار می گیرند. تابع سیگموید فاصله را به بازه $[0, 1]$ می برد. مقدار بدست آمده شباهت دو ورودی را مشخص می کند.

$$y_{out} = \sigma\left(\Sigma_j \alpha_j |h^{(j)}_{1,L-1} - h^{(j)}_{2,L-1}|\right)$$

آموزش و تابع هزینه

اگر فرض کنیم $y(X_1^{(i)}, X_2^{(i)})$ برچسب دسته نمونه i-ام از داده ها باشد، قرار میدهیم، $y(X_1^{(i)}, X_2^{(i)}) = 1$ اگر هر دو از یک دسته باشند. در غیر اینصورت قرار می دهیم $y(X_1^{(i)}, X_2^{(i)}) = 0$.

تابع خطا از نوع آنتروپی دوگانه منظم شده [۲۵] به شکل زیر است.

$$\mathcal{L}(X_1^{(i)}, X_2^{(i)}) = y(X_1^{(i)}, X_2^{(i)}) \log y_{out}(X_1^{(i)}, X_2^{(i)})$$
$$+(1 - y(X_1^{(i)}, X_2^{(i)})) \log(1 - y_{out}(X_1^{(i)}, X_2^{(i)}) + \lambda^T |w|^2$$

آموزش از طریق پس انتشار خطا با گرادیان کاهشی انجام می شود. از آنجایی وزن ها بین زیر شبکه ها مشترک هستند، گرادیان برای هردو یکسان محاسبه میشود. همچنین برای نتیجه بهتر از گشتاور استفاده میشود. بروزرسانی وزن ها به صورت زیر محاسبه می شود.

$$w^T_{kj}(X_1^{(i)}, X_2^{(i)}) = w^T_{kj} + \Delta w^T_{kj}(X_1^{(i)}, X_2^{(i)}) + 2\lambda_j |w_{kj}|$$
$$\Delta w^T_{kj}(X_1^{(i)}, X_2^{(i)}) = -\eta_j w^T_{kj} + \mu_j {}^{T-1}_{kj}$$

۲.۴.۲ معماری سیامی بازگشتی برای یادگیری شباهت جملات

[۲۶][۴] هدف این مقاله مقایسه جمله های انگلیسی از نظر معنایی است و با استفاده از یک معماری سیامی قادر به رسیدن به بهترین عملکرد در بین روش های موجود است.

[۲۵] regularized cross-entropy

[۲۶] Siamese Recurrent Architecture for Learning Sentence Similarity

داده های مربوطه در این پژوهش از نوع جمله هستند که ماهیتاً دنباله ای از کلمات با طول های متفاوت است. این کلمات و ترتیب آنها در معنا تاثیرگذار هستند لذا به نظر میرسد مدل مناسب برای آنها باید دارای حافظه باشد. در این پژوهش از شبکه بازگشتی LSTM استفاده شده است. مزیت LSTM به شبکه های بازگشتی ساده این است که بدلیل عدم وجود مشکل گرادیان محو شونده،میتواند وابستگی های بلند مدت [۲۷] در یک دنباله را به راحتی حفظ کند، که برای مسئله موجود ضروری است.

مدل

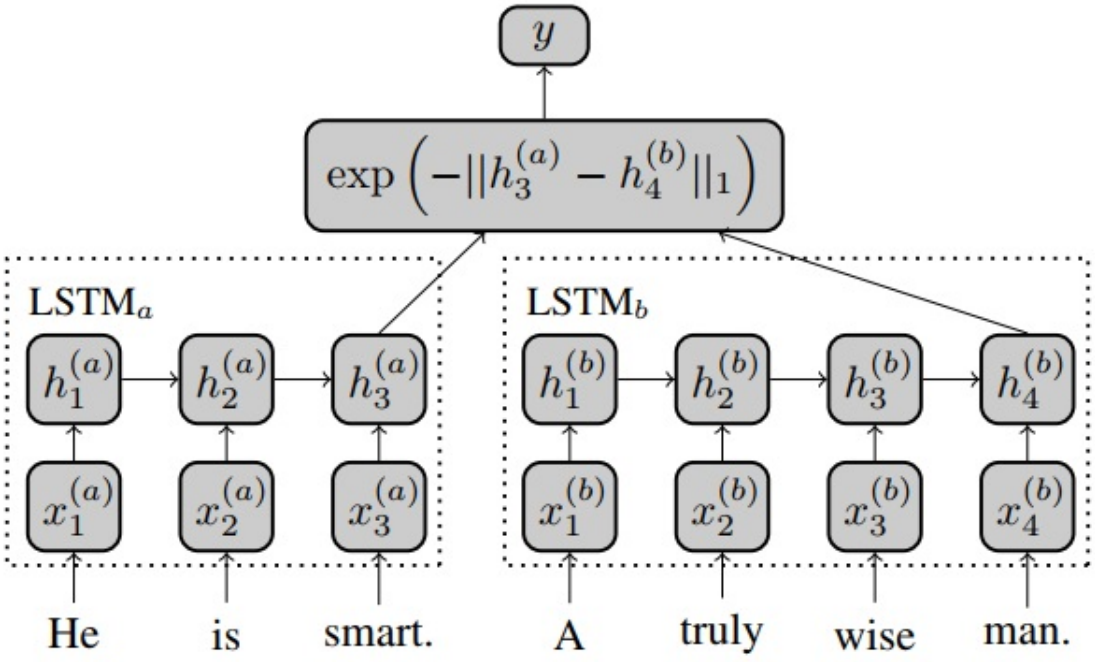
در شبکه سیامی ارائه شده، هر زیر شبکه یک LSTM است که این دو با هم مساوی و دارای وزن های مشترک هستند. هر جمله کلمه به کلمه به LSTM مربوط داده می شود و در نهایت آخرین خروجی LSTM که بار معنایی کل جمله را در بر دارد به لایه مقایسه کننده شبکه سیامی داده می شود.

اگر h_{T_1} و h_{T_2} آخرین لایه پنهان باشند، تشابه دو نمونه به صورت زیر محاسبه میشود:

$$y_{out}(h_{T_1}, h_{T_2}) = \exp(-\|h_{T_1} - h_{T_2}\|_1) \in [0, 1]$$

این پژوهش نیز از فاصله L_1 استفاده می کند. دلیل ترجیح این متریک به فاصله L_2 اینست که همان طور LeCun در مقاله خود مطرح می کند [۵]، L_2 میتواند در تابع هدف پلاتو های نامطلوبی ایجاد کند و باعث گرادیان محو شونده شود.

[۲۷] Long-term dependencies



شکل ۴.۲: شبکه ارائه شده. LSTM ها بصورت باز شده[۲۸] در زمان نمایش داده شده اند.

آموزش و تابع هزینه

تابع هزینه از نوع میانگین مربع خطا (MSE) [۲۹] است و فرآیند یادگیری از طریق پس انتشار خطا در زمان صورت می گیرد.

۳.۴.۲ کاربرد شبکه سیامی در تشخیص امضا

از مقالات ذکر شده نکات مهمی می توان دریافت کرد که در مسئله ما مورد استفاده قرار می گیرند.

اول از همه این که به نظر میرسد در عمل شبکه های سیامی برای یادگیری One-shot از کارایی لازم برخوردار هستند. از آنجایی که در مسئله ما نمونه امضای هر فرد تعداد محدودی دارد میتوان امیدوار بود یادگیری k-shot (در حالت ایده آل k=۱) توسط شبکه سیامی موثر

[۲۹] Mean-squared error

واقع شود.

دوم اینکه از آنجایی که ماهیت داده های امضای برخط از جنس دنباله هستند، به نظر میرسد روش مقاله دوم به این مسئله قابل تعمیم باشد. از آنجا که طول بردار امضا در داده های ما ممکن است بسیار بلندتر از جملات باشند، چالش اصلی اینست که آیا شبکه های LSTM قادر خواهند بود وابستگی های بسیار بلند را به نحو موثری تشخیص دهند.

فصل ۳

دادگان امضا

اهمیت دادگان[۱] مناسب در هر پروژه شناسایی الگو بدیهی است اما در استفاده از شبکه های عمیق، بدلیل وابستگی شدید سیستم به داده ها و ماهیت جعبه سیاه این روش، اهمیتی دو چندان پیدا می کند. برای دست یابی به نتایج بهتر، دادگان انتخاب شده باید دارای نمونه های متعدد و متفاوت باشند و دارای توزیعی مناسب روی داده ها باشد تا بایاس در فرآیند طبقه بندی اتفاق نیافتند و سیستم بدست آمده بتواند در عمل روی نمونه های جدید که به دادگان تعلق نداشته اند نیز عملکردی قابل قبول داشته باشد. در اصطلاح، این بدین معناست که سیستم دارای عمومیت[۲] است. در این بخش به معرفی دادگان به کار رفته در این پروژه می پردازیم که از نوع دادگان امضاهای برخط هستند.

۱.۳ دادگان SVC۲۰۰۴

[۲] دادگان SVC۲۰۰۴ در جریان *اولین دوره مسابقات بین المللی تصدیق امضا SVC۲۰۰۴* اولین بار ارائه شد. این دادگان در جریان هر دو قسمت این رقابت ها دادگان مورد استفاده تیم

[۱] dataset

[۲] generality

ها و معیار عملکرد سیستم آنها بود. این دادگان شامل دو قسمت ۱۰۰ تایی مجموعه امضا بوده، که هر مجموعه امضا شامل ۲۰ امضای اصل از یک شخص و ۲۰ امضای جعلی حرفه ای از روی امضای همان شخص است. البته مجموعه امضای ۱۰۰ نفر اول که متعلف به بخش اول مسابقات است اطلاعات مربوط به فشار و زاویه قلم را به همراه ندارد. نکته مهم دیگر اینست که امضاها اکثراً انگلیسی یا چینی هستند.

۱.۱.۳ ویژگی های SVC۲۰۰۴

- تعداد افراد[۳]: ۲۰۰ نفر
- تعداد امضاها اصل هر فرد: ۲۰ عدد
- تعداد امضاهای جعلی هر فرد: ۲۰ عدد
- ویژگی ها ذخیره شده برای هر نقطه از امضا:

X-coordinate	موقعیت اسکیل شده قلم روی محور x
Y-coordinate	موقعیت اسکیل شده قلم روی محور y
Time stamp	زمان سیستم در هنگام ثبت این پیشامد
Button status	حالت فعلی قلم (۰ برای بالا بودن و ۱ برای پایین بودن قلم)
Azimuth	زاویه ساعتگرد چرخش قلم حول محور x
Altitude (Inclination)	زاویه بین قلم و جهت مثبت محور z
Pressure	مقدار تنظیم شده فشار قلم

۲.۱.۳ نتایج بدست آمده روی SVC۲۰۰۴

در مسابقات بهترین خطا به صورت میانگین ۲.۸۴% بود.

[۳] users

۲.۳ دادگان BiosecurID

[۳] BiosecuID یک دادگان چند منظوره است که در سال ۲۰۱۰ توسط تیمی از محققان اسپانیایی قابل دسترسی قرار گرفت. این دادگان دارای ۷ خصوصیت بیومتریک است: گفتار، عنبیه چشم، صورت (شامل عکس و صورت های در حال صحبت)، امضا و دست خط (برخط و برون خط)، اثر انگشت، کف دست و .keystroking این دادگان با روندی مناسب تهیه شده است و دارای توزیع جمعیتی متعادلی بین زن و مرد است. همچنین اطلاعات جمعیت شناسی درباره اشخاص دادگان برای تحلیل های بیشتر موجود است.

۱.۲.۳ ویژگی های BiosecurID

ویژگی های بخش امضای برخط این دادگان به شرح زیر است. توجه کنید امضاها در ۴ جلسه با فواصل ۴ ماهه از افراد گرفته شده اند. بنابراین حتی امکان بررسی تغییرات زمانی در تحقیقات وجود دارد.

- تعداد افراد: ۴۰۰ نفر
- تعداد جلسات: ۴ جلسه
- تعداد امضاها اصل هر فرد: ۱۶ عدد (در هر جلسه ۴ عدد)
- تعداد امضاهای جعلی هر فرد: ۱۲ عدد (در هر جلسه ۳ عدد)

۲.۲.۳ نتایج بدست آمده روی BiosecurID

بهترین نتیجه بدست آمده رو این دادگان خطای ۵.۵۸% دارد.

فصل ۴

سیستم های تشخیص امضا موجود

در این بخش تعدادی از سیستم های تشخیص امضای موجود را که عملکردهای خوبی در این زمینه داشته اند به اختصار شرح میدهیم. در ادامه نیز تعدادی از تحقیقات جدید که در انجام این پروژه مورد بررسی قرار می گیرند را شرح خواهیم داد.

- **تشخیص امضای برخط بر اساس GA-SVM[۸]**

هوانگ و همکارانش روشی با استفاده از هر دو روش توصیف داده بردار پشتیبانی [۱] (با استفاده از (SVM) و الگوریتم ها ژنتیک (GA)) برای تشخیص امضای برخط ارائه دادند. در این روش یک مجموعه ویژگی ۲۷ پارامتری شامل شکل و ویژگی های دینامیکی از داده امضای برخط استخراج می شود. SVM میتواند به عنوان یک روش طبقه بندی تک رده ای مبتنی بر هسته[۲] به دقت توزیع حالت امضاهای اصل را پیشبینی کرده و امضاهای جعلی را به این صورت تشخیص دهد. داده های امضا از دادگان SVC۲۰۱۳ در این تحقیق برای آزمایش تشخیص استفاده شده است. این روش پیشنهادی EER ای ۴.۹۳ درصدی در دادگان امضای جعل بدست می آورد. این روش از پارامترهای

[۱] Support Vector Data Description

[۲] kernel-based

ویژگی سراسری [۳] استفاده می کند اما با این حال از مزایای ضد دخالت بودن قوی، محاسبات ریاضی راحت برخوردارست. این سیستم در قابلیت تشخیص جزئیات محلی کمی ضعیف عمل می کند.

- **تصدیق هویت با استفاده از تصدیق امضای برخط بهبود یافته [۹]**

کولماتف و همکارانش آزمایشی برای اطمینان صحت امضا با تطبیق دادن امضا با تمام امضاهای مرجع از شخص ادعا شده، با روش کشش زمانی پویا (DTW) ارائه کردند. مقدار فاصله امضای مورد آزمایش با نزدیک ترین و دورترین امضاهای مرجع با استفاده از میانگین متناظر با مجموعه مرجع نرمال سازی می شوند تا یک بردار ویژگی سه بعدی را بسازند. سپس این بردار ویژگی به دو رده ی اصلی و جعلی طبقه بندی می شود. با استفاده از تحلیل مولفه های اصلی (PCA) [۴] نرخ خطای ۱.۴% برای مجموعه داده ای از ۹۴ شخص و ۶۱۹ امضای آزمایش (شامل اصل و جعل) بدست آمده است. در این مقاله، گروه محققان نشان داده اند که به نظر میرسد در شناسایی الگوی دو رده ای در این مسئله اطلاعات مربوط به فشار، پارامتر متمایزکننده موثری نیست.

- **تشخیص امضای دست نویس برخط با استفاده از طبقه بندی توسط شبکه عصبی مبتنی بر تحلیل مولفه های اصلی[۱۰]**

ایرانمنش و همکارانش روشی سیستماتیک برای تشخیص امضا براساس استفاده از شبکه چندلایه پرسپترون (MLP) روی زیرمجموعه ای از ویژگی های بدست آمده از تحلیل مولفه های اصلی (PCA) معرفی کردند. روش پیشنهادی تعمیم انتخاب ویژگی ای از اطلاعات معمولاً بلااستفاده PCA نشان میدهد که نقش مهمی در رسیدن به نرخ خطای پایینتر دارد. آزمایش روی ۴۰۰۰ نمونه امضا از دادگان SIGMA انجام شده و به نرخ قبولی غلط (FAR) ۷.۴% و نرخ رد غلط (FAR) ۶.۴% دست یافته است. با استفاده ۸۰۰۰ نمونه امضا دقت ۹۳.۱% بدست آمده است. این مقاله نشان میدهد

[۳] Global feature parameters

[۴] Principal Component Analysis

نه تنها PCA بلکه عوامل دیگری مثل مقادیر استفاده شده در شبکه و مقادیر امتیاز به امضا میتوان تاثیر زیادی در افزایش دقت داشته باشد.

- **تشخیص امضا**[۱۱]

جین و همکارنش از یک تبلت دیجیتال برای ثبت اطلاعات دینامیک و فضایی امضا استفاده کردند. تشابه بین یک امضای ورودی و مجموعه مرجع با استفاده از تکنیک های تطبیق رشته محاسبه می شود و سپس با یک مقدار آستانه مقایسه می شود. به علاوه در این مقاله تعداد حرکت ها قلم [۵] به عنوان یک ویژگی سراسری استفاده می شود.

- **تشخیص امضای برخط و برون خط: یک روش ترکیبی**[۱۲]

رادیکا و همکارانش روی ویژگی های برخط و برون خط در کنار هم کار کردند تا بتوانند نتایجشان را در تصدیق و تشخیص امضا ترکیب کنند. روش های برخط و برون خط به صورت جداگانه امضا را تصدیق می کنند و در نهایت نتایج آنها در کنار هم توسط یک SVM تصدیق می گردد. این تحقیق همچنین روش های برخط، برون خط و ترکیبی را با یکدیگر مقایسه می کند. در روش برخط بردار ویژگی و در روش برون خط از گرادیان و ویژگی های تصویری برای تشخیص امضا استفاده می شود.

- **تشخیص امضا با استفاده از ویژگی های استاتیک و دینامیک** [۱۳]

وستا و همکاران او توضیح می دهند که بافت و ویژگی های توپولوژیکی، ویژگی های استاتیک یک تصویر امضا هستند در حالی که تبلت دیجیتال ویژگی هایی دینامیکی از قبیل فشار، نقاط قطع[۶] و زمان نوشتن امضا را به صورت بلادرنگ ثبت می کند. از 1D - log Gabor wavelet و اعداد اویلر برای تحلیل ویژگی های بافتی و توپولوژیکی به ترتیب استفاده شده است. یک الگوریتم طبقه بندی چند رده ای نتایج را ترکیب می کند و از این سه مجموعه ویژگی (دینامیک، بافت، توپولوژی) به دقت ۸۹.۱۸% میرسد.

[۵] stroke

[۶] breakpoints

- **تشخیص امضای برخط برای تصدیق هویت چند منظوره با استفاده از گوشی های هوشمند** [۱۴]

فورهاد و همکارانش یک سیستم تصدیق هویت چند عاملی بیومتریک پیاده سازی کردند که از پلتفرم موبایل استفاده می کند. این مدل به راحتی با استفاده از یک مدل تصدیق هویت تک عاملی یا چند عاملی قابل پیاده سازی است و باعث می شود دسترسی به یک سیستم تصدیق هویت پیچیده تر یا قابل اطمینان تر برای مصارف روزانه ممکن شود.

- **DCT بر اساس استخراج ویژگی برای تشخیص امضای دینامیک**[۱۵]

رشیدی و همکاران او روشی ساده و به صرفه را برای تشخیص امضا بر اساس یک تبدیل کسینوس گسسته که به ۴۴ سیگنال زمانی مانند موقعیت، سرعت، شتاب و زاویه قلم اعمال می شود، ارائه دادند. از الگوریتم انتخاب ویژگی جلورونده [۷] برای انتخاب زیرمجموعه ای ویژگی ها که دارای بهترین عملکرد است، استفاده شده است. سیستم ارائه شده با امضاهای جعلی حرفه ای و روش های طبقه بندی متفاوت امتحان شده است.

مقالات بعدی در ۱۴-امین کنفرانس بین المللی تحلیل و تشخیص نوشتار (IAPR) در سال ۲۰۱۷ ارائه شده اند و ایده های نویی با استفاده از شبکه های عصبی عمیق و سیامی ارائه میدهند.

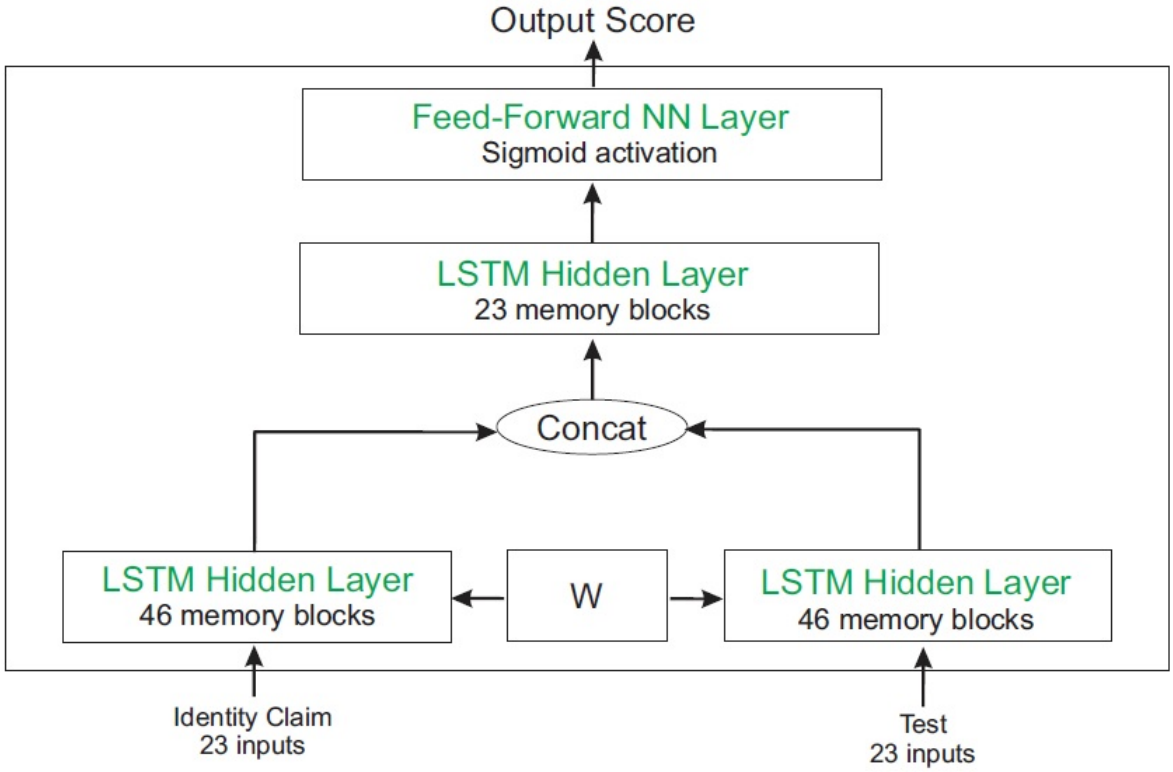
۱.۴ تشخیص امضا با شبکه های عصبی بازگشتی

[۱۶] در این تحقیق تولوسانا و همکاران او از دادگان BiosecurID که در فصل ۳ آمده است استفاده می کنند. آنها تلاش دارند به امضاها به چشم دنباله نگاه کنند و از طریق شبکه های بازگشتی این دنباله ها را با یکدیگر مقایسه کنند. به دلیل طول نسبتاً زیاد امضاها از شبکه LSTM استفاده می شود که قدرت به مراتب بیشتری در استخراج و حفظ ارتباطات طولانی

[۷] Forward feature selection

دارد. (همانند مثال موجود در فصل دو، در رابطه به معنای جملات. LSTM در واقع همانطور که از اسمش پیداست یک واحد حافظه است و خروجی نهایی آن درواقع یک فشرده سازی با طول ثابت از کل دنباله است.)

در این مقاله، نویسندگان از یک معماری سیامی استفاده می کنند، بطوریکه نه تنها دو زیر شبکه ی آن از جنس LSTM هستند بلکه لایه LSTM دیگری در بخش مقایسه وجود دارد. خروجی این لایه، سپس به یک لایه معمولی کاملاً متصل وصل می شود. ورودی شبکه سیامی دو امضا است و خروجی ۰ و ۱ به ترتیب به معنای جعلی و یا اصل بودن آن است.



شکل ۱.۴: معماری سیستم تشخیص امضا با استفاده از حافظه LSTM

به دلیل یکسان بودن دادگان، پروتکل استفاده شده در پروژه ما بگرفته از پروتکل این مقاله است که در فصل ۵ توضیح داده شده است. به طور کلی نتایج در دو حالت مقایسه یک امضا با تمام امضا ۱vs۱ و مقایسه امضا با ۴ امضای مرجع ۴vs۱ بررسی شده اند که بیشرفت قابل

ملاحظه ای را نسبت به نتایج مرجع میبینیم.

۴vs۱	۱vs۱	
۷.۷۵	۱۰.۱۷	Basedline System
۵.۵۸	۶.۴۴	Proposed System

نتایج مقاله بخش ۴.۱

فصل ۵

تشخیص امضا با شبکه سیامی

همانگونه که در فصل ۲ اشاره کردیم، شبکه های سیامی ابزار قدرتمندی برای مقایسه نمونه های مختلف و محاسبه تفاوت آن هاست. در این پروژه تلاش می کنیم از این شبکه برای هدف تشخیص امضا بهره گیریم. ایده کلی به این صورت است که اگر بتوانیم شبکه سیامی ای آموزش دهیم که بتواند تشخیص دهد دو امضای داده شده متعلق به یک کلاس هستند یا خیر (یعنی آیا هر دو امضای حقیقی یک شخص هستند با اینکه یک امضا جعلی است؟) ، آنگاه با دسترسی به تعداد کمی امضای مرجع از هر شخص، پس از ارائه یک امضای جدید،با دادن مقایسه آن با امضاهای مرجع می توان تشخیص داد که آیا این امضا به شخص تعلق دارد یا خیر. (اگر تفاوت بین دو امضا از یک آستانه تعیین شده توسط شبکه بیشتر باشد آنگاه امضای جدید یا متعلق به شخصی دیگر و یا جعلی است.)

پیاده سازی این سیستم با استفاده از زبان برنامه نویسی python و با استفاده از بسته تخصصی keras و tensorflow انجام شده است.

۱.۵ شبکه سیامی بدون رمزگذاری

[۱] هر امضا دنباله ای از نقاط است و همانطور که امضاهای مختلف اشکال و اندازه های مختلفی دارند، بدیهی است که این دنباله ها از طول های متفاوتی برخوردار باشند. سرعت امضا کردن شخص نیز به این تفاوت طول کمک می کند زیر قلم نوری نقاط را در فاصله های زمانی ثابت ثبت می کند. به همین دلیل یک ایده این است که از طریقی همه دنباله ها را با روشی رمزگذاری و فشرده کرده تا همگی تبدیل به بلوک های هم اندازه شوند. کارکردن با داده ها هم طول بسیار ساده تر است زیرا چنین داده هایی را می توان با انواع مختلف شبکه ها پردازش کرد. با این حال در ابتدا تصمیم گرفتیم بدون رمزگذاری به تشخیص امضا بپردازیم. این کار با حذف مرحله رمزگذاری باعث افزایش سرعت سیستم می شود. همچنین با فشرده کردن دنباله های با طول های متغیر به طول ثابت با روش های مد نظر ما، بخشی از اطلاعات از دست می رود، بنابراین اگر بتوانیم از رمزگذاری صرف نظر کنیم ممکن است بتوانیم نتایج دقیق تری بدست آوریم.

به این منظور تمام امضاها را به طولی ثابت که برابر با بیشینه طول دنباله امضاهاست، با اضافه کردن صفر توسعه می دهیم. (در ابتدا امضاهایی که طول آنها انحراف از معیار زیادی دارند را از داده ها حذف می کنیم.)

۱.۱.۵ فرآیند کلی

ابتدا داده ها را به دو قسمت یادگیری و آزمایشی تقسیم می کنیم.

سپس از داده های یادگیری زوج هایی از امضاهای هر شخص می سازیم. این زوج ها نمونه های آموزشی ما هستند. اگر یک زوج امضا هر دو امضاهای حقیقی یک شخص باشند برچسب نمونه ۱ و در غیر این صورت ۰ است.

شبکه سیامی عمیق و کاملاً متصل [۲] می سازیم که به اندازه طول بزرگترین امضا، گره ورودی

[۱] encoding

[۲] fully connected

دارد. سپس این شبکه را با زوج امضاهای یادگیری آموزش می دهیم.
در آخر شبکه را با زوج امضاهایی از داده های آزمایشی، ارزیابی می کنیم.

۲.۱.۵ پیش پردازش

در این مرحله داده ها را نرمالیزه می کنیم و با اضافه کردن ۰ به انتهای دنباله امضا طول هر نمونه را به ماکزیمم طول موجود میرسانیم.

۳.۱.۵ معماری

معماری شبکه سیامی کاملاً متصل است و در آن از Drop out نیز استفاده شده است. تا به اینجا تنها از x و y امضا استفاده می کنیم ولی از آنجایی که شبکه عصبی ما تنها یک بردار را قبول می کند، امضای به شکل $((x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n))$ به شکل $(x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_n)$ به شبکه داده می شود. بدیهی است که ترتیب گذاشتن x و y به دلیل ماهیت شبکه های عصبی و اتصالات آن ها تفاوتی ایجاد نمی کند.
ضمناً خروجی های دو زیر شبکه با حالت تفریق (به صورت درایه به درایه) به هم می پیوندند و وارد بخش مقایسه گر شبکه می شوند.
تابع هدف در تمامی آزمایشات Binary cross-entropy است.
بهینه ساز adam مورد استفاده قرار گرفته است.
تابع فعال سازی برای زیر شبکه ها reLU است.
تابع فعال سازی لایه های مقایسه گر sigmoid است.

۴.۱.۵ پروتکل

پروتکل را برای دادگان BiosecurID به طور خاص شرح میدهیم. پروتکل SVC۲۰۰۴ به جز در اعداد منحصر به BiosecurID مشابه است.

در این دادگان، امضاهای با طول بیشتر از ۲۴۰۰ را حذف می کنیم و سپس تمام امضا ها را با افزایش ۰ [۳] به طول ۲۴۰۰ توسعه می دهیم.

برای هر فرد در دادگان، ۴ امضای حقیقی جلسه اول به عنوان امضاهای مرجع درنظر گرفته می شوند. این بدان معنی است که در آینده سیستم امضاهای جدید را در مقابل این ۴ امضا برای این شخص می سنجد.

امضاهای ۳۰۰ فرد اول از دادگان به عنوان داده های یادگیری و ۱۰۰ تای باقیمانده به عنوان داده های آزمایشی مورد استفاده قرار می گیرند. زوج امضاهای یادگیری به صورت زیر تشکیل می شوند:

$4 \times 12 \times 300 = 14400$ زوج از امضاهای مرجع و بقیه امضاهای حقیقی آن فرد برای هر ۳۰۰ نفر

$4 \times 12 \times 300 = 14400$ زوج از امضاهای مرجع و تمام امضاهای جعلی آن فرد برای هر ۳۰۰ نفر

بدیهی است که گروه اول برچسب ۱ و گروه دوم برچسب ۰ خواهند داشت. با این داده ها شبکه را آموزش می دهیم. برای ارزیابی شبکه روی داده های آموزشی به دو روش عمل می کنیم. درصد خطا را برای تمام زوج های موجود بدست می آوریم (۱vs۱) و یا میانگین خطا برای تمام زوج های یک امضا با ۴ امضای مرجع[۴] را بدست می آوریم. برای فرآیند آزمایش، امضای جدید ارائه شده برای یک شخص را باز به صورت ۱vs۱ و یا ۴vs۱ ارزیابی می کنیم.

[۳] zero-padding

[۴] 4vs1

۵.۱.۵ نتایج

نتایج مربوط به SVC۲۰۰۴

- آزمایش ۱-۱۰۰۰-۶۴-۶۴:

تعداد دور[۵]: ۱۰۰۰

	Precision	Recall	F۱
۱	۱.۰	۱.۰	۱.۰
۲	۰.۷۶۴۹	۰.۷۱۴۳	۰.۷۱۱۵
۳	۰.۶۶۳۷	۰.۶۳۲۱	۰.۶۳۱۱
۴	۰.۷۹۵۷	۰.۷۶۷۹	۰.۷۶۸۲
۵	۰.۹۹۶۵	۰.۹۹۶۴	۰.۹۹۶۴

- آزمایش ۲-۱۰۰۰-۶۴-۶۴:

تعداد دور[۶]: ۱۰۰۰

	Precision	Recall	F۱
۱	۰.۹۷۹۳	۰.۹۷۸۶	۰.۹۷۸۵
۲	۰.۷۹۷۷	۰.۷۵۷۱	۰.۷۵۶۳
۳	۰.۶۷۸۶	۰.۶۷۸۶	۰.۶۷۸۶
۴	۰.۷۷۷۸	۰.۷۶۴۳	۰.۷۶۵۴
۵	۰.۹۹۲۹	۰.۹۹۲۹	۰.۹۹۲۸

نتایج مربوط به BiosecurID

ابتدا از عمق کم شروع و عمق شبکه را زیاد کرده ایم. در کم عمق ترین حالت، شبکه علاوه بر یک لایه ورودی دارای یم لایه در هر زیرشبکه و یک لایه طبقه بندی کننده در آخر است که

[۵] epoch

[۶] epoch

خروجی هر دو زیرشبکه را به عنوان ورودی در کنار هم دریافت می کند.

- آزمایش ۱۰۰۰-۱۰۲۴-۱۰۲۴:

تعداد دور[۷]: ۱۰۰۰

	Precision	Recall	F۱	Accuracy	FAR	FRR
۱	۰.۵۳۵	۰.۵۱۸۹	۰.۴۲۱۷	۰.۵۱۸۸	۰.۰۸	۰.۸۹۶۴
۲	۰.۵۰۳	۰.۵۰۴	۰.۴۹۰۴	۰.۵۰۳۹	۰.۳۳۴۱	۰.۶۶۰۴
۳	۰.۴۹۲۸	۰.۵۰۴۱	۰.۴۹۱۲	۰.۵۰۴۰	۰.۳۴۵۸	۰.۶۷۴۳
۴	۰.۵۳۲۴	۰.۵۱۹۳	۰.۴۳۰۵	۰.۵۱۹۲	۰.۰۹۴۱	۰.۸۸۰۹

- آزمایش ۱۰۰۰-۲۰۴۸-۲۰۴۸:

تعداد دور[۸]: ۱۰۰۰

	Precision	Recall	F۱	Accuracy	FAR	FRR
۱	۰.۵۰۰۸	۰.۵۰۴۵	۰.۴۲۱۷	۰.۴۷۷۷	۰.۲۷۳۳	۰.۷۲۵۶
۲	۰.۴۸۵۵	۰.۴۸۵۵	۰.۴۸۵۵	۰.۴۸۵۵	۰.۵۱۳۳	۰.۵۱۵۶
۳	۰.۴۸۹۲	۰.۴۷۶۸	۰.۴۷۱۲	۰.۴۷۶۷	۰.۶۲۳۹	۰.۴۰۳۳
۴	۰.۵۳۱۳	۰.۵۲۹۹	۰.۵۱۰۸	۰.۵۲۹۸	۰.۲۷۶۶	۰.۶۷۰۴

- آزمایش ۱۰۰۰-۳۰۷۲-۳۰۷۲:

تعداد دور[۹]: ۱۰۰۰

[۷] epoch

[۸] epoch

[۹] epoch

FRR	FAR	Accuracy	F۱	Recall	Precision	
۰.۶۰۰۵	۰.۳۸۹۱	۰.۵۰۶۹	۰.۵۰۱۴	۰.۵۰۷	۰.۵۰۵۵	۱
۰.۳۱۴۹	۰.۶۴۹۱	۰.۵۱۶۵	۰.۵۰۲۸	۰.۵۱۶۶	۰.۵۲۰۳	۲
۰.۱۹۰۲	۰.۸۲۴۹	۰.۴۶۴۹	۰.۴۰۷۵	۰.۴۶۵	۰.۴۹۰۴	۳
۰.۱۹۸۴	۰.۸۳۵	۰.۴۷۷۷	۰.۴۱۹۲	۰.۴۷۷۷	۰.۴۷۱۷	۴

- آزمایش :۱۰۰۰-۱۲۸-۱۲۸x۵۱۲x۲۰۴۷

تعداد دور[۱۰]: ۱۰۰۰

FRR	FAR	Accuracy	F۱	Recall	Precision	
۰.۶۱۳۴	۰.۳۹۴۱	۰.۴۹۸۰	۰.۴۹۲	۰.۴۹۸۱	۰.۴۹۶۲	۱
۱.۰	۰.۰	۰.۵۰۳۹	۰.۳۳۷۸	۰.۵۰۴	۰.۲۵۴	۲
۱.۰	۰.۰	۰.۵۴۳۲	۰.۳۸۲۴	۰.۵۴۳۲	۰.۲۹۵۱	۳
۱.۰	۰.۰	۰.۵۰۸۶	۰.۳۴۳	۰.۵۰۸۷	۰.۲۵۸۸	۴

۲.۵ شبکه سیامی با رمزگذاری

از آنجایی که نتایج حالت بدون رمزگذاری رضایتبخش نیستند، تلاش داریم که این بار از یک شبکه autoencoder استفاده کرده و دنباله های امضا را به بردارهایی با طول ثابت تبدیل کنیم.

autoencoder مورد استفاده در این بخش از نوع ساده خواهد بود و برای ورودی آن همچنان نیاز به توسعه دنباله ها به طول بیشینه خواهیم داشت.

[۱۰] epoch

۳.۵ شبکه سه تایی Triplet

شبکه سه تایی از ایده های جدید در زمینه تشخیص تمایز بین نمونه ها است. این شبکه به نوعی تعمیم شبکه سیامی است و از ۳ زیرشبکه بهره می برد. نحوه دادن ورودی نیز کمی متمایز است. از آنجا که در برخی مسائل این نوع شبکه از عملکرد بهتری نسبت به شبکه سیامی برخوردار بوده است، یکی از انتخاب ها برای آزمایش های بعدی به حساب می آید.

فصل ۶

نتیجه گیری

نتایج آزمایشات ما نشان میدهد که روش شبکه سیامی بدون رمزنگاری، به مراتب عملکرد بهتری روی دادگان SVC۲۰۰۴ نسبت به BiosecurID دارد. این نتیجه را میتوان بدلیل کوتاهتر بودن امضاهای SVC۲۰۰۴ دانست. از آنجایی که در روش ما تمام امضاها با ۰ توسعه پیدا می کنند می توان توقع داشت که این سیستم با امضاهای بسیار طولانی مقدار زیادی گره اضافه به شبکه اضافه می کند. این گره های اضافی که خود به نوعی معرف متغیرهای جدید در تخمین شبکه هستند، باعث ناکارایی و یادگیری سخت تر شبکه می شود.
همچنین محدود کردن تعداد ویژگی ها به تنها x و y می تواند عامل زیانبار دیگری باشد. در آزمایش های آینده تلاش خواهیم کرد از ویژگی های دیگر امضا و از ویژگی های استخراجی دیگر مانند مشتق و زوایا در شبکه خود استفاده نماییم.

در کل مدل های تشخیص امضا امروزه با شتاب بسیاری در حال پیشرفت و معرفی هستند اما این سیستم ها همچنان نتوانسته اند در کاربرد جایگاه خود را اهراز نمایند. این به دلیل حساسیت تشخیص امضا و نیاز به پیشرفت بیشتر شبکه های حاضر بوده است.

کتاب‌نامه

[۱] Maureen Caudill، *Neural Network Primer: Part I*، .۱۹۸۹

[۲] First Internation Signature Verification Competition (SVC۲۰۰۴)، http://www.cse.ust.hk/svc2004/instructions.pdf

[۳] Fierrez، J.، Galbally، J.، Ortega-Garcia، J. et al. ، *BiosecurID: a multimodal biometric database*، Pattern Anal Applic (۲۰۱۰) :۱۳ .۲۳۵ https://doi.org/10.1007/s10044-009-0151-4

[۴] Mueller، J.، Thyagarajan، A.، Siamese Recurrent Architectures for Learning Sentence Similarity. In AAAI(pp. ،(۲۷۸۶-۲۷۹۲ ،۲۰۱۶) February).

[۵] Chopra، S.، Hadsell، R.، LeCun، Y. ، Learning a similarity metric discriminatively، with application to face verification. In Computer Vision and Pattern Recognition، .۲۰۰۵ CVPR .۲۰۰۵ IEEE Computer Society Conference on (Vol. ،۱ pp. .(۵۳۹-۵۴۶ IEEE، ،۲۰۰۵) June).

[۶] Koch، G.، Zemel، R.، Salakhutdinov، R. ، Siamese neural networks

for one-shot image recognition. In ICML Deep Learning Workshop (Vol. ،(۲ .(۲۰۱۵)

[۷] Lai، S.، Jin، L.، Yang، W. ، Online Signature Verification using Recurrent Neural Network and Length-normalized Path Signature. arXiv preprint arXiv:۱۷۰۵.۰۶۸۴۹، .(۲۰۱۵)

[۸] D. Huang and G. Jian، "On-line Signature Verification Based on GA-SVM،" Int. J. Online Eng.، vol. ،۱۱ no. ،۶ pp. ،۵۳—۴۹ .۲۰۱۵

[۹] A. Kholmatov and B. Yanikoglu، "Identity Authentication using Improved Online Signature Verification Method،" Pattern Recognit. Lett.، vol. ،۲۶ no. ،۱۵ pp. ،۲۴۰۸—۲۴۰۰ .۲۰۰۵

[۱۰] V. Iranmanesh، S. Mumtazah، S. Ahmad، W. Azizun، W. Adnan، S. Yussof، O. A. Arigbabu، and F. L. Malallah، "Online Handwritten Signature Verification Using Neural Network Classifier Based on Principal Component Analysis،" Sci. World J.، vol. ،۲۰۱۴ pp. ،۸—۱ .۲۰۱۴

[۱۱] A. K. Jain، F. D. Griess، and S. D. Connell، "On-Line Signature Verification،" Pattern Recognit. ،۳۵ vol. ،۳۵ pp. ،۲۹۷۲—۲۹۶۳ .۲۰۰۲

[۱۲] K. S. Radhika and S. Gopika، "Online and Offline Signature Verification: A Combined Approach،" Procedia - Procedia Comput. Sci.، vol. ،۴۶ pp. ،۱۶۰۰—۱۵۹۳ .۲۰۱۵

[۱۳] M. Vatsa، R. Singh، P. Mitra، and A. Noore، "Signature Verification

Using Static and Dynamic Features،" NEURAL Inf. Process.، pp. ۳۵۰–۳۵۵، ۲۰۰۴.

[۱۴] N. Forhad، B. Poon، M. A. Amin، and H. Yan، "Online Signature Verification for Multi-modal Authentication using Smart Phone،" Proc. Int. MultiConference Eng. Comput. Sci.، pp. ۱۸–۲۱، ۲۰۱۵.

[۱۵] S. Rashidi، A. Fallah، and F. Towhidkhah، "Feature Extraction Based DCT on Dynamic Signature Verification،" Sci. Iran.، vol. ۱۹، no. ۶، pp. ۱۸۱۰–۱۸۱۹، ۲۰۱۲.

[۱۶] R. Tolsana، et. al، Biometric Signature Verification Using Recurrent Neural Networks، ۱۴th IAPR International Conference on Document Analysis and Recognition، ۲۰۱۷.

[۱۷] M. Singhal، M. Dutta، Online Signature Verification: Present State of Technology، International Journal on Recent and Innovation Trends in Computing and Communication، Vol. ۴، no. ۹، p۶۶-۶۸، ۲۰۱۶.

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

A Survey of Siamese Networks and Their Application in Online Signature Verification

Anahita Doosti Sanjani

Supervisor: Dr. Bagher BabaAli

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

February 2018