

## یک مدل یادگیری عمیق بهینه برای تایید دست خط

پروانه افضل<sup>۱</sup>، عبدالرضا رضاپور<sup>۲\*</sup>، احمد رضایی جوردهی<sup>۳</sup>

۱- دانشجوی دکتری، گروه مهندسی کامپیوتر، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران

۲- استادیار، گروه مهندسی کامپیوتر، واحد آستانه اشرفیه، دانشگاه آزاد اسلامی، آستانه اشرفیه، ایران

۳- استادیار، گروه مهندسی برق، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران

رسید مقاله: ۳ فروردین ۱۴۰۲

پذیرش مقاله: ۲۸ مرداد ۱۴۰۲

### چکیده

دست نوشته، به عنوان راهی متداول برای ارتباط در میان جوامع انسانی و یک زیست سنج رفتاری به منظور احراز هویت، شناخته شده است. تایید هویت نویسنده یک متن مکتوب از کاربردهای قابل توجهی در حوزه سیستم های امنیتی، پزشکی قانونی و اسناد تاریخی و ادبی، برخوردار است. در این مقاله، برای نخستین بار، یک معماری عمیق بهینه شده با بهره مندی از شبکه سیامی، یادگیری انتقالی و الگوریتم بهینه ساز گرگ خاکستری، به منظور تایید هویت نویسنده به صورت برون خط از روی دست خط فارسی ارایه می گردد. به دلیل عدم دسترسی به مجموعه داده مناسب در زبان فارسی، یک مجموعه داده شامل کلمات، جملات و ارقام فارسی جمع آوری شده است. علاوه بر آن، از دو مجموعه داده شناخته شده IAM و KHATT نیز برای تجزیه و تحلیل روش پیشنهادی استفاده گردید. نتایج آزمایشات و ارزیابی ها بر روی این مجموعه داده ها، بر توانایی رویکرد پیشنهادی برای تایید هویت نویسنده از روی دست خط در زبان های مختلف و سبک های گوناگون نگارش، تاکید دارند.

**کلمات کلیدی:** تایید هویت نویسنده، دست خط، شبکه سیامی، یادگیری انتقالی، الگوریتم بهینه ساز گرگ خاکستری.

### ۱ مقدمه

تایید دست خط، یک روش زیست سنج رفتاری برای احراز هویت شخصی با کاربردهای مهم در سیستم های امنیتی و بررسی اسناد تاریخی و ادبی و امور مختلف مالی، اداری و حقوقی است [۱]. تصدیق برون خط هویت از روی دست نوشته، به فرآیند تجزیه و تحلیل تصویر متن مکتوب برای مطابقت دادن و یافتن میزان شباهت میان نمونه اصلی و نمونه مورد سوال برای تشخیص اصالت تألیف، اشاره دارد [۲]. این فرآیند بر این اساس که هر فرد از سبک دست خط خاص خود برخوردار است، امکان پذیر است [۳].

\* عهده دار مکاتبات

آدرس الکترونیکی: Abdoreza.Rezapour@iau.ac.ir

دو رهیافت در سیستم‌های تایید هویت با توجه به ماهیت روش‌های جمع‌آوری داده‌های دست‌خط وجود دارند: برخط<sup>۱</sup> و برون‌خط. در رهیافت برخط، نمونه‌های دست‌خط از طریق یک دستگاه دیجیتالی مانند قلم‌های الکترونیکی و صفحه‌های قابل لمس، جمع‌آوری می‌شوند. لذا، ویژگی‌های پویا و مختلفی مانند سرعت نوشتن و میزان فشار قلم در هنگام نوشتن قابل دسترس می‌باشند. اما در رهیافت برون‌خط، هیچ‌گونه اطلاعات زمانی و پویایی در خصوص دست‌نویسته وجود ندارد و تنها تصویر متن مکتوب به صورت اسکن شده پس از تکمیل فرآیند نوشتن، قابل استفاده است [۴]. علی‌رغم آنکه تصویر دست‌نویسته به عنوان تنها منبع موجود برای استخراج ویژگی می‌تواند چالش‌برانگیز باشد، رهیافت برون‌خط از اهمیت به‌سزایی در برنامه‌های کاربردی برخوردار است، زیرا دستگاه‌های ضبط دیجیتالی در همه‌جا در دسترس نیستند و در خصوص تجزیه و تحلیل و تایید هویت در اسناد تاریخی و یا در برخی از اسناد پزشکی قانونی، جمع‌آوری اطلاعات پویای نوشتار، غیرممکن است [۵].

سیستم‌های تایید هویت مبتنی بر دست‌خط از دید محتوای متنی نیز به دو دسته‌ی وابسته به متن<sup>۲</sup> و مستقل از متن<sup>۳</sup> تقسیم می‌شوند. در رویکرد وابسته به متن، متن داده‌های آموزشی و متن داده‌های آزمایشی یکسان می‌باشند و لذا سیستم در چنین حالتی دارای کاربردهای خاص و محدودتری است. اما در مقابل، در سیستم‌های مستقل از متن، محدودیتی برای یکسان بودن محتوای داده‌های آموزشی و داده‌های آزمایشی وجود نداشته و این سیستم‌ها کاربردی‌تر می‌باشند [۶].

در زمینه‌ی تایید هویت نویسنده، چالش‌های مهمی مطرح می‌باشند، از جمله این که ممکن است دو نویسنده‌ی مختلف برای نوشتن به سبک‌های مشابه، آموزش دیده باشند. جاعل نیز در نمونه‌های جعلی سعی می‌کند تا بیشترین شباهت با نمونه‌ی اصلی وجود داشته باشد. از طرف دیگر، تغییرات مختلفی ممکن است در سبک نوشتار نویسنده با توجه به حالت یا وضعیت نویسنده در زمان نوشتن یا با گذشت زمان، مشاهده شود که از آن به عنوان تغییرات درون کلاسی<sup>۴</sup> یاد می‌شود و لذا به طور کلی یادگیری سبک دست‌خط یک فرد از روی چندین نمونه دست‌نویسته برای در نظر گرفتن تغییرات درون کلاسی در نمونه‌های دست‌خط، مطلوب‌تر است. بررسی پژوهش‌های موجود، نشان می‌دهد که در زمینه‌ی تایید هویت نویسنده براساس دست‌خط و به خصوص دست‌نویسته‌ی فارسی، مطالعات محدودی انجام شده است و بیشتر پژوهش‌های انجام شده در زمینه‌ی تایید هویت نویسنده براساس امضا، تشخیص هویت یا شناسایی نویسنده<sup>۵</sup> براساس دست‌خط و یا تشخیص<sup>۶</sup> حروف و کلمات از روی دست‌نویسته می‌باشند. باید به این نکته نیز توجه شود که تایید هویت نویسنده براساس دست‌خط دشوارتر از تایید هویت نویسنده براساس امضا است، چرا که امضا معمولاً یک طرح ثابت انتخاب شده توسط فرد است اما در تایید هویت براساس دست‌خط، متن دست‌نویسته می‌تواند متغیر و شامل مجموعه‌ای متنوع از حروف، ارقام و نمادهای مختلف نگارشی باشد.

<sup>1</sup> Online

<sup>2</sup> Text-dependent

<sup>3</sup> Text-independent

<sup>4</sup> Intra-class Variations

<sup>5</sup> Writer Identification

<sup>6</sup> Recognition

لذا در این مقاله، پس از جمع‌آوری یک مجموعه داده از نمونه‌های دست‌خط فارسی از ۵۰ نویسنده با در نظر گرفتن تغییرات درون کلاسی، یک رویکرد بهینه‌ی برون‌خط و مستقل از متن برای تایید هویت نویسنده از روی دست‌خط ارائه می‌گردد. در رویکرد پیشنهادی مبتنی بر شبکه سیامی<sup>۱</sup>، هر زوج تصویر (اصلی-اصلی یا اصلی-جعلی) به دو زیرشبکه‌ی یکسان وارد شده و بردارهای ویژگی به دست آمده از دو زیر شبکه‌ی یکسان برای تشخیص میزان شباهت میان نمونه‌ی اصلی و نمونه‌ی مورد سوال و تایید هویت نویسنده، مورد استفاده قرار می‌گیرند. در ساختار شبکه سیامی از مدل‌های از پیش آموزش دیده شده<sup>۲</sup> به کمک یادگیری انتقالی<sup>۳</sup> استفاده شده است و همچنین برای افزایش دقت<sup>۴</sup> در آموزش شبکه، الگوریتم بهینه‌ساز گرگ خاکستری<sup>۵</sup> (GWO) برای برای تنظیم ابرپارامترها به کار گرفته شده است.

ساختار مقاله در ادامه به این شرح است: بخش ۲ شامل مروری بر پیشینه تحقیق در زمینه‌ی تایید هویت نویسنده می‌باشد. شرح روش انجام تحقیق در بخش ۳ ارائه شده است. داده‌های مورد استفاده، آزمایشات انجام شده جهت ارزیابی عملکرد رویکرد پیشنهادی و بحث پیرامون نتایج در بخش ۴ شرح داده شده‌اند. نتیجه‌گیری و پیشنهادها نیز در بخش ۵ بیان شده‌اند.

## ۲ پیشینه تحقیق

تاکنون، روش‌های مختلفی برای تایید هویت نویسنده با استفاده از راهکارهای یادگیری ماشین و یادگیری عمیق پیشنهاد شده است. در هنگام استفاده از روش‌های رایج یادگیری ماشین [۷]، ویژگی‌های مورد استفاده برای بازنمایی و آموزش مدل طبقه‌بندی، باید از قبل توسط طراح تعریف شده و از تصویر استخراج شوند. این فرآیند بسیار وقت گیر است و در انتخاب توصیف‌کننده‌های مناسب ویژگی‌ها، به دقت زیادی نیاز دارد. با این حال، تلاش‌های تحقیقاتی گوناگونی برای تعریف و استخراج بهترین مجموعه ویژگی‌های متمایزکننده انجام شده است [۸]. در بخش عمده‌ای از تحقیقات اخیر، از روش‌های بازنمایی آماری برای استخراج ویژگی‌ها از تصویر امضا استفاده شده است که از جمله آن‌ها می‌توان به استفاده از توصیف‌کننده‌های مبتنی بر جهت مانند هیستوگرام گرادینان‌های جهت دار [۹] اشاره نمود.

بنسفی و همکاران [۱۰] از روشی مبتنی بر تجزیه و تحلیل یک نمونه‌ی منحصر به فرد از یک کلمه دست‌نویس برای تشخیص اصالت نویسنده‌ی دست‌خط استفاده کردند. در این روش از معیار فاصله ویرایش لوناشتاین<sup>۶</sup> استفاده شد و روش پیشنهادی بر روی بخشی از مجموعه داده IAM شامل ۱۰۰ نویسنده، آزمایش شد. نتایج نشان داد که دقت مدل پیشنهادی بر روی کل مجموعه داده (۱۰۰ نویسنده) ۸۷ درصد و بر روی ۴۰ نویسنده ۹۲ درصد است.

<sup>1</sup> Siamese Network

<sup>2</sup> Pre-trained Models

<sup>3</sup> Transfer Learning

<sup>4</sup> Accuracy

<sup>5</sup> Gray Wolf Optimizer

<sup>6</sup> Levenshtein Edit Distance

جانگ و همکاران [۱۱] روشی را برای تشخیص اصالت نویسنده بر پایه متون کره‌ای ارائه دادند. در روش پیشنهادی این پژوهش ابتدا تصاویر رنگی تبدیل به تصاویر سیاه و سفید شدند و تصاویر از پس زمینه، با استفاده از یک حد آستانه، جدا شده و پس از آن نویزها تشخیص داده شده و حذف شدند. در نهایت تصاویر با استفاده از الگوریتم باریک سازی ایجاد شدند. گوشه، انتها و نقاط تقاطع تشخیص داده شده و بخش‌های بین این نقاط حذف شدند. در نهایت معیار فاصله بدون نرمال سازی و با نرمال سازی تعیین شد. در این پژوهش میزان خطا در حالت بدون نرمال سازی ۳۷ درصد و در حالت نرمال شده ۳۶ درصد گزارش شده است.

در سال‌های اخیر، در حوزه تایید هویت نویسنده نیز به دلیل مشکلات استخراج ویژگی‌ها، علاقه زیادی به روش‌هایی وجود دارد که به ویژگی‌های دست‌ساز تکیه نمی‌کنند. روش‌های یادگیری عمیق می‌توانند ویژگی‌های یک تصویر را مستقیماً از خود تصویر یاد بگیرند و امکان ترکیب وظایف استخراج ویژگی‌ها و طبقه‌بندی را در یک چارچوب واحد فراهم کنند [۱].

هافمن و همکاران [۱۲] یک مدل استخراج ویژگی مستقل از نویسنده با استفاده از شبکه عصبی پیچشی را پیشنهاد کردند که با امضاهای اصلی و جعلی یک گروه از امضاکنندگان آموزش دیده بود. چالیک و همکاران [۵] از شبکه عصبی پیچشی عمیق برای تشخیص و تایید خودکار امضای دست‌نویس در حالت برون‌خط استفاده کردند. در معماری شبکه پیشنهادی، از ترکیب الگوریتم نزدیک‌ترین همسایه به منظور طبقه‌بندی بردار ویژگی‌ها استفاده شد. الگوریتم پیشنهادی در این مطالعه بر روی سه مجموعه داده‌ی مختلف از جمله GPDS، MCYT و CEDAR ارزیابی شد. همچنین این الگوریتم با شبکه‌های از پیش آموزش دیده مقایسه شد و به صحت‌های ۹۶ درصد و ۹۸ درصد دست یافت.

گش [۱۳] از شبکه‌های عصبی بازگشتی برای تایید و تشخیص امضای دست‌نویس در حالت برون‌خط استفاده نمود. در این پژوهش از ویژگی‌های مختلف ساختاری و جهت‌دار از هر نمونه امضا که به صورت محلی استخراج شده بود، استفاده شد و بردار ویژگی تولید شده با استفاده از دو مدل مختلف شبکه‌های عصبی بازگشتی شامل LSTM و BLSTM طبقه‌بندی شد. الگوریتم پیشنهادی بر روی شش پایگاه داده عمومی (GPDS، Synethic، GPDS-300، MCYT-75، CEDAR، BHsig260 Hindi و BHsig260 Bengali) که به طور گسترده در مطالعات مربوط به تشخیص خودکار نویسنده‌ی امضاء استفاده می‌شوند، ارزیابی شد. همچنین، الگوریتم پیشنهادی با شبکه‌های پیچشی مقایسه شد و به نتایج بهتری نسبت به این شبکه‌ها دست یافت.

یکی از مهم‌ترین محدودیت‌های مساله تایید امضای دست‌نویس، داده‌های ناکافی برای مراحل آموزش است که باعث ایجاد محدودیت‌هایی در استفاده از روش‌های یادگیری عمیق می‌گردد. یابیچی و همکاران [۱۴]، نحوه‌ی جدیدی از بهره‌مندی از شبکه‌های مولد تخصصی<sup>۱</sup> را به عنوان یک روش داده‌افزایی در نمونه‌های آموزشی، برای حل مشکل داده‌های ناکافی در مساله تایید امضا ارائه می‌کنند. محققان، روش داده‌افزایی پیشنهادی خود را با استفاده از دو مجموعه داده‌ی محبوب در ادبیات فن (GPDS و MCYT) و بر روی چهار مدل

<sup>1</sup> Generative Adversarial Network (GAN)

مختلف از شبکه عصبی پیچشی (VGG16، VGG19، ResNet50 و DenseNet121) مورد ارزیابی قرار دادند و به نتایج قابل قبولی دست یافتند.

دی و همکاران [۱۵] یک مدل تایید مستقل از نویسنده به نام SigNet با استفاده از یک شبکه سیامی متشکل از شبکه‌های عصبی پیچشی دوقلو با وزن مشترک ارایه کردند. این دو شبکه به طور موازی با جفت‌های مختلفی از تصاویر امضای مشابه و متفاوت برای یادگیری فضای ویژگی مشاهدات مشابه، آموزش دیدند. مارگر و همکارانش [۱۶] یک مدل تایید امضا را پیشنهاد نمودند که از معماری شبکه سه گانه با مدل‌های پیشرفته شبکه عصبی پیچشی استفاده کرده است. شبکه سه گانه یک فرمت از شبکه سیامی است که شامل سه شاخه از همان شبکه پیشرو با پارامترهای مشترک است. برای آموزش، یک سه گانه از امضا (امضای مرجع، نمونه مثبت و نمونه منفی) به شبکه داده می‌شود و تابع هزینه بر اساس فاصله اقلیدسی بین جفت‌های مثبت و منفی محاسبه می‌شود.

آداک و همکاران [۱۷] روشی را بر پایه یادگیری تجمعی<sup>۱</sup> با استفاده از شبکه سیامی برای تشخیص اصالت نویسنده ارایه دادند. در این روش ابتدا تابع توزیع احتمال برای استخراج ویژگی‌های مهندسی شامل ویژگی‌های متنی مانند روش گرفتن قلم و شیب نوشتاری و همچنین ویژگی‌های مربوط به شکل حروف و نویسه‌ها، استفاده شد و پس از آن این ویژگی‌ها با ویژگی‌هایی که به طور خودکار از شبکه‌های پیچشی به دست آمدند، ترکیب شدند و در نهایت به عنوان ورودی به شبکه سیامی اعمال شدند تا تشخیص اصالت نویسنده انجام شود. روش ارایه شده با استفاده از مجموعه داده‌های متشکل از ۱۰۰ نویسنده با ۳۰۰ صفحه دست‌نوشته مبتنی بر زبان بنگالی ارزیابی شد. نتایج نشان داد که ویژگی‌های متن با شبکه سیامی دقت ۹۷ درصد و ویژگی‌های مرتبط با شکل حروف و نویسه‌ها با شبکه سیامی دقت ۹۶ درصد را برای تشخیص ارایه می‌دهند. این در حالی است که شبکه عصبی تک لایه به جای شبکه سیامی برای دسته‌بندی با همین ویژگی‌ها به ترتیب دقت ۹۶ درصد و ۹۵ درصد را ارایه می‌دهد. شیخ و همکاران [۲] روشی را مبتنی بر یادگیری تجمعی بر پایه شبکه‌های عمیق پیشرفته سیامی خودرمزگذار<sup>۲</sup> برای تشخیص اصالت نویسنده ارایه دادند. در این روش ویژگی‌های مهندسی با استفاده از روش SIFT<sup>۳</sup> و GSC<sup>۴</sup> استخراج شدند و با ویژگی‌هایی که به صورت خودکار با استفاده از شبکه عمیق استخراج شدند، ترکیب شدند. برای ارزیابی مدل پیشنهادی تنها از کلمه پرتکرار "and" استخراج شده از مجموعه داده CEDAR که شامل ۱۵۵۱۸ کلمه از ۱۵۶۷ نویسنده است، استفاده شد. ارزیابی‌ها نشان داد که هرچند استخراج ویژگی‌ها با استفاده از SIFT و GSC مکملی برای روش‌های استخراج ویژگی توسط شبکه‌های عمیق هستند؛ اما فرآیند استخراج ویژگی به صورت دستی امری بسیار زمان‌بر است و ۷۰ درصد از زمان یادگیری برای این کار مصرف می‌شود. نتایج نشان داد که بهترین دقت مدل پیشنهادی با استفاده از ترکیب ویژگی‌های SIFT با شبکه عمیق سیامی است که برابر است با ۹۹ درصد برای مجموعه داده آموزش و ۶۳ درصد برای مجموعه داده ارزیابی، که این خود بیانگر وجود بیش‌برازش در سامانه است.

<sup>1</sup> Ensemble Learning

<sup>2</sup> Auto Encoder

<sup>3</sup> Scale Invariant Feature Transform

<sup>4</sup> Gradient Structural Concavity

پرچم و همکارانش [۱] یک مدل تایید امضای جدید را با هدف ارایه یک مدل دقیق با عملکردی بالا پیشنهاد نمودند. در این مدل، یک معماری ترکیبی به نام CBCapsNet شامل مدل‌های پیشنهادی شبکه عصبی پیچشی و شبکه عصبی کپسول<sup>۱</sup> ارایه گردید تا قابلیت شبکه اصلی در مرحله استخراج ویژگی بهبود یافته و دقت تشخیص افزایش یابد. این محققان تلاش نمودند تا با استفاده از مزایای شبکه‌های پیچشی در شناسایی و استخراج ویژگی‌ها، ضعف آنها در تشخیص و تمایز بین دگرگونی‌ها و تغییرات مکانی در مولفه‌های تصویر را با استفاده از قدرت شبکه‌های عصبی کپسول جبران نمایند. علاوه بر این، با طراحی یک سازوکار آموزشی جدید که در آن یک شبکه واحد به طور هم‌زمان توسط دو تصویر آموزش داده می‌شود، پارامترهای آموزشی را کاهش دادند. همان‌گونه که از بررسی پژوهش‌های پیشین مشخص می‌باشد، تحقیقات گسترده و متنوعی در خصوص تشخیص اصالت نویسنده بر اساس دست خط و به‌ویژه دست خط فارسی صورت پذیرفته و بیشتر پژوهش‌ها در این زمینه بر پایه‌ی امضا بوده است.

### ۳ روش انجام تحقیق

در این بخش، ابتدا به مرور شبکه عصبی پیچشی<sup>۲</sup> (CNN)، یادگیری انتقالی و مدل‌های از پیش آموزش دیده شده بر مبنای آن، شبکه سیامی و الگوریتم بهینه‌ساز گرگ خاکستری و سپس به تشریح معماری روش پیشنهادی پرداخته می‌شود.

### ۳-۱ شبکه عصبی پیچشی، یادگیری انتقالی و مدل‌های از پیش آموزش دیده شده

شبکه‌های عصبی پیچشی، نوعی خاصی از شبکه‌های عصبی عمیق هستند که شامل چندین لایه پیچشی با اندازه کرنل‌های گوناگون، لایه‌های ادغام و لایه‌های کاملاً متصل می‌باشند. لایه‌های پیچشی پایینی در این شبکه‌ها به استخراج ویژگی‌های سطح پایین مانند لبه و رنگ پرداخته و در لایه‌های بالایی، ویژگی‌های سطح بالا نظیر خط و شیء استخراج می‌شوند. این شبکه‌ها در انواع وظایف بینایی ماشین بسیار پر کاربرد هستند [۱۸]. با این وجود، آموزش مدل‌های شبکه‌های عصبی پیچشی عمیق از مشکلات مربوط به عدم تنوع و کمیت داده رنج می‌برد. در این میان، استفاده مجدد از یک مدل از پیش آموزش دیده شده بر روی تعداد زیادی داده آموزشی، با هدف انتقال دانش کسب شده به یک مساله جدید می‌تواند راهگشا باشد و به عنوان مفهوم یادگیری انتقالی شناخته می‌شود. تاکنون مدل‌های از پیش آموزش داده شده بسیار متنوعی از شبکه عصبی پیچشی عمیق ارایه شده‌اند.

### ۳-۲ شبکه سیامی

شبکه عصبی سیامی نوع خاصی از معماری شبکه‌های عصبی شامل دو زیرشبکه با پیکربندی یکسان است. به‌روزرسانی پارامترها در هر دو زیرشبکه منعکس می‌شود و زیرشبکه‌ها از طریق یک تابع زیان<sup>۳</sup> به هم متصل می‌شوند. این تابع یک معیار شباهت را بر مبنای فاصله اقلیدسی بین بردارهای ویژگی حاصل از دو زیرشبکه،

<sup>1</sup> Capsule Neural Network

<sup>2</sup> Convolutional Neural Network

<sup>3</sup> Loss Function

محاسبه می کند [۱۹]. شبکه سیامی بر اساس مفهوم مقایسه الگوها عمل می کند و با استفاده از آن می توان به طور موثر الگوهای مشابه یا متفاوت را تشخیص داد. این شبکه به عنوان یکی از روش های پرکاربرد در حوزه تشخیص الگو، در مسایل مختلفی در حوزه بینایی ماشین از جمله تایید چهره [۲۰] و تایید امضا [۱۵] به نتایج قابل قبولی دست یافته است.

شبکه سیامی به دلیل ویژگی های خاص خود برای مسایل تایید دست خط مناسب است. دست خط شامل الگوها و جزییاتی است که به صورت یکتا برای هر فرد متفاوت است، و تشخیص و تایید صحت دست خط یکی از چالش های مهم در حوزه تایید هویت فردی است. شبکه سیامی قادر است ویژگی های مشترک بین دو دست خط را استخراج کند و بر اساس آنها تصمیم گیری کند.

در ادامه، به صورت موردی به علل انتخاب شبکه سیامی در تایید دست خط اشاره می کنیم:

- عدم نیاز به تعداد زیادی دست خط برجسب دار: یکی از مزایای اصلی استفاده از شبکه سیامی در تایید دست خط، عدم نیاز به تعداد زیادی دست خط برجسب دار برای آموزش است. شبکه سیامی با استفاده از معماری خاص خود، امکان استخراج ویژگی های مشترک بین نمونه های دست خط را فراهم می کند. به این ترتیب، با داشتن تعداد محدودی دست خط برجسب دار، می توان شبکه را آموزش داد و برای تایید دست خط های جدید استفاده کرد. این قابلیت بسیار مفید است زیرا در صورت محدود بودن داده های برجسب دار یا عدم دسترسی به تعداد زیادی نمونه دست خط، شبکه سیامی همچنان قادر به تشخیص و تایید دست خط خواهد بود.
- تشخیص الگوهای مشابه و متفاوت: شبکه سیامی به عنوان یک روش قدرتمند می تواند الگوهای مشابه و متفاوت در نمونه های دست خط را تشخیص دهد. با استخراج ویژگی های مشترک بین دست خط ها، می توان الگوهایی که در دست خط های یک فرد تکرار می شوند را تشخیص داد. همچنین، با تشخیص تفاوت ها و انحرافات بین دست خط ها، می توان به تایید یا رد دست خط پرداخت.
- قابلیت تشخیص تغییرات در دست خط: شبکه سیامی قادر است به تشخیص تغییرات در دست خط کمک کند. زمانی که یک فرد تغییراتی در دست خط خود ایجاد می کند، می توان با استفاده از شبکه سیامی این تغییرات را تشخیص داد.

### ۳-۳ الگوریتم بهینه ساز گرگ خاکستری

الگوریتم بهینه ساز گرگ خاکستری [۲۱]، یکی از انواع الگوریتم های فرا ابتکاری [۲۲] است که با الهام از ساختار سلسله مراتبی رهبری و سازوکار شکار گرگ های خاکستری در طبیعت، ارائه گردید. در هر گله، ۴ رتبه اصلی از گرگ ها وجود دارند. گله توسط گرگ های گروه آلفا هدایت می شود و توسط گرگ های بتا، گرگ های زیردستی که به آلفا در فرآیند تصمیم گیری و حفظ پویایی دسته کمک می کنند، دنبال می شوند. در این میان، پایین ترین رتبه در میان گرگ های خاکستری، امگا است که همیشه باید مطیع دیگر گرگ های غالب باشند و آخرین گرگ هایی هستند که مجاز به خوردن از طعمه می باشند. اگر گرگ، آلفا، بتا یا امگا نباشد، به عنوان دلتا

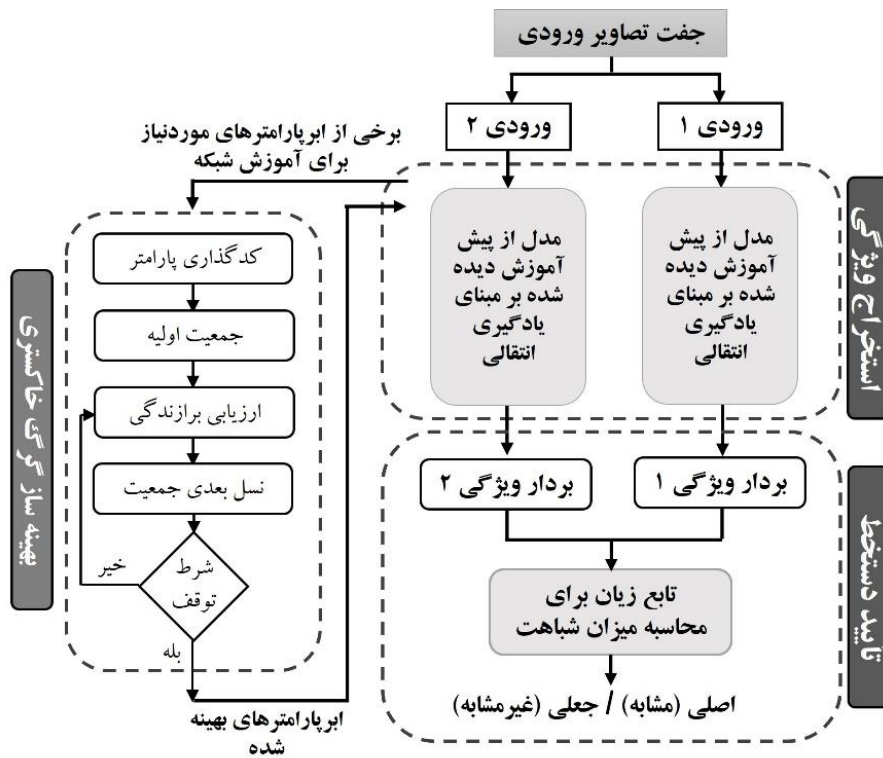
شناخته می شود. گرگ های دلتا باید به آلفا و بتا گزارش دهند. آن ها بر گرگ های امگا تسلط دارند و به عنوان پیشاهنگ، نگهبان، سالمند، شکارچی و نگهبان عمل می کنند. در الگوریتم بهینه ساز گرگ خاکستری مناسب ترین راه حل به عنوان آلفا شناخته می شود.

علت انتخاب الگوریتم گرگ خاکستری به منظور بهینه سازی مدل عمیق پیشنهادی برای تایید دست خط عبارتند از:

- سادگی و سرعت اجرا: الگوریتم گرگ خاکستری به دلیل ساختار ساده خود، به عنوان یک الگوریتم بهینه سازی مستقیم، سرعت اجرای بالایی دارد. این ویژگی برای بهینه سازی در مدل های عمیق که معمولاً دارای تعداد زیادی پارامتر هستند، بسیار مهم است.
- قابلیت جستجوی گسترده: الگوریتم گرگ خاکستری قادر به جستجو در فضای پارامترها به صورت گسترده است. این به معنی این است که الگوریتم می تواند در محدوده های گسترده تری از فضای جستجو حرکت کند و از این طریق به دست یابی به نقاط بهینه بهتر کمک کند.
- تعامل گروهی: الگوریتم گرگ خاکستری بر اساس رفتار گروهی گرگ ها طراحی شده است. این ویژگی به این الگوریتم کمک می کند تا بهینه سازی را به صورت هم زمان در بخش های مختلف فضای جستجو انجام دهد و از این طریق به جستجوی موثرتری برای پیدا کردن نقاط بهینه برسد.
- قابلیت همگرایی به نقاط بهینه محلی و سرعت خروج از مینیمم محلی: الگوریتم گرگ خاکستری قادر است به نقاط بهینه محلی همگرا شود و در عین حال از مینیمم محلی خارج شود. این به معنی این است که الگوریتم قادر است در محدوده های نزدیک بهینه های محلی گیر نکند و به جستجوی مناطق بهینه جدید بپردازد. این قابلیت سبب می شود تا این الگوریتم از اجتناب بالا از بهینه محلی سود ببرد که منجر به اجتناب از همپوشانی ویژگی ها در مساله انتخاب ویژگی می گردد.

### ۳-۴ معماری روش پیشنهادی

آنچنان که در شکل ۱ قابل مشاهده است، در معماری روش پیشنهادی، یک شبکه عصبی پیچشی سیامی ارایه گردید به گونه ای که در شاخه های شبکه سیامی به جای استفاده از یک شبکه عصبی پیچشی معمول، با هدف دستیابی به دقت بالاتر، از مدل های پیچشی از پیش آموزش دیده شده مبتنی بر یادگیری انتقالی استفاده شد. در این ساختار، ابتدا دو زیر شبکه با وزن های یکسان با یک جفت تصویر از دست نوشته (اصلی-اصلی یا اصلی-جعلی) به عنوان ورودی، تغذیه می شوند. سپس بردارهای ویژگی به دست آمده از این دو زیر شبکه ی یکسان، بر اساس تابع زیان برمبنای فاصله اقلیدسی برای تشخیص میزان شباهت میان نمونه ی اصلی و نمونه ی مورد سوال و در نهایت تایید دست خط مورد استفاده قرار می گیرند. همچنین از آنجایی که ابر پارامترها نقش مهمی در دقت و همگرایی شبکه های عصبی پیچشی ایفا می کنند، در آموزش شبکه، الگوریتم بهینه ساز گرگ خاکستری برای تنظیم برخی از ابر پارامترها مورد استفاده قرار گرفت.



شکل ۱. ساختار شبکه سیامی بهینه پیشنهادی برای تایید دستخط

#### ۴ بحث پیرامون نتایج

در این بخش، ابتدا مجموعه داده‌ی جمع‌آوری شده و دو مجموعه داده‌ی مورد استفاده شده جهت ارزیابی عملکرد روش پیشنهادی، نحوه ایجاد جفت تصاویر در مجموعه داده‌ها و پیش پردازش‌های لازم، معرفی شده و سپس نتایج حاصل از آزمایشات مورد بحث و بررسی قرار می‌گیرند.

#### ۴-۱-۱ داده

#### ۴-۱-۱-۱ مجموعه داده راست به چپ جمع‌آوری شده

زبان فارسی از جمله زبان‌های راست به چپ با سبک‌های نوشتاری متفاوت بر اساس نحوه اتصال حروف به یکدیگر است. از آنجایی که تعداد کمی از مجموعه داده‌های راست به چپ در دسترس عموم وجود دارند، لذا ما یک مجموعه داده از داده‌های دست‌خط راست به چپ شامل کلمات، جملات و اعداد فارسی را با توجه به وجود تنوع در نمونه‌های دست‌خط یک نویسنده، طراحی کردیم. در جمع‌آوری این مجموعه داده، ۵۰ نویسنده شامل ۲۹ مرد و ۲۱ زن از مشاغل گوناگون و در سنین مختلف مابین ۱۶ تا ۶۳ سال شرکت داشتند. اکثر مشارکت‌کنندگان راست دست بودند و برای نوشتن از انواع مختلف قلم‌ها در دو رنگ آبی و مشکی استفاده کردند. به منظور مشاهده تغییرات در نمونه‌های دست‌خط هر نویسنده‌ی خاص، به هر شرکت‌کننده هفت برگ کاغذ داده شد و از آنها خواسته شد که محتوای متنی هر برگ را ده بار بنویسند. چند نمونه از تصاویر دست‌خط از مجموعه داده‌ی جمع‌آوری شده، در شکل ۲ نشان داده شده است.



دست‌نویس که تقریباً شامل یک یا دو کلمه بودند، برش داده و جدا نمودیم. همچنین برای افزایش داده‌ها از راهکارهای افزایش داده شامل "روشنایی"، "زوم" و "چرخش تصادفی" از کتابخانه‌ی Keras بهره بردیم. کل مجموعه داده‌های موجود نیز به منظور آموزش و بررسی میزان یادگیری سبک‌های دست‌خط نویسندگان، توسط مدل به سه زیر مجموعه‌ی آموزش (۷۰ درصد)، اعتبارسنجی (۱۰ درصد) و آزمایش (۲۰ درصد) تقسیم شدند.

#### ۴-۲ آزمایشات و نتایج

یکی از چالش‌های موجود در تنظیم ابرپارامترها جهت آموزش یک مدل عمیق، انتخاب دقیق نرخ یادگیری<sup>۱</sup> است که اندازه گام را در هر تکرار که وزن‌های مدل به‌روزرسانی می‌شوند، مشخص می‌کند. در نظر گرفتن یک مقدار بسیار کوچک برای نرخ یادگیری باعث ایجاد تغییرات کوچک‌تری در وزن‌ها در هر به‌روزرسانی و در نتیجه سبب افزایش تعداد دوره‌ها و تاخیر در فرآیند یادگیری مدل می‌شود. یک مقدار بیش از حد بزرگ نیز منتج به همگرایی سریع مدل به یک راه حل غیربهبهینه می‌گردد.

تعداد دوره<sup>۲</sup> بیانگر تعداد دفعاتی است که تمامی داده‌های آموزشی در چرخه‌ی آموزش مدل عمیق، مورد استفاده قرار می‌گیرند. اگر تعداد دوره‌های مورد استفاده برای آموزش مدل، بیش از حد لازم باشد، مدل الگوهایی را می‌آموزد که تا حد زیادی مختص نمونه داده‌های آموزشی هستند و این امر باعث می‌شود که مدل نتواند عملکرد خوبی بر روی مجموعه داده‌های آزمایش داشته باشد. به عبارت دیگر، قابلیت تعمیم مدل به دلیل بیش برآزش بر روی داده‌های آموزش از دست می‌رود. نظارت بر مقدار زیان و دقت بر روی مجموعه داده آموزش و نیز مجموعه داده اعتبارسنجی، برای بررسی و تعیین تعداد دوره‌ها حایز اهمیت است.

در این مطالعه، برای استفاده از الگوریتم بهینه‌ساز گرگ خاکستری به منظور تعیین نرخ بهینه یادگیری و نیز تعداد بهینه دوره، ابتدا تابع هدفی براساس دقت مدل در مجموعه اعتبارسنجی تعریف می‌گردد. این بدان معناست که به دنبال یافتن نرخ یادگیری یا تعداد دوره‌هایی هستیم که مدل را به دقت بالایی روی داده‌های اعتبارسنجی برساند. شبه کد استفاده از الگوریتم بهینه‌ساز گرگ خاکستری برای بهینه نمودن نرخ یادگیری / تعداد دوره‌ها در شکل ۴ به طور خلاصه نشان داده شده است.

<sup>1</sup> Learning Rate

<sup>2</sup> Epoch

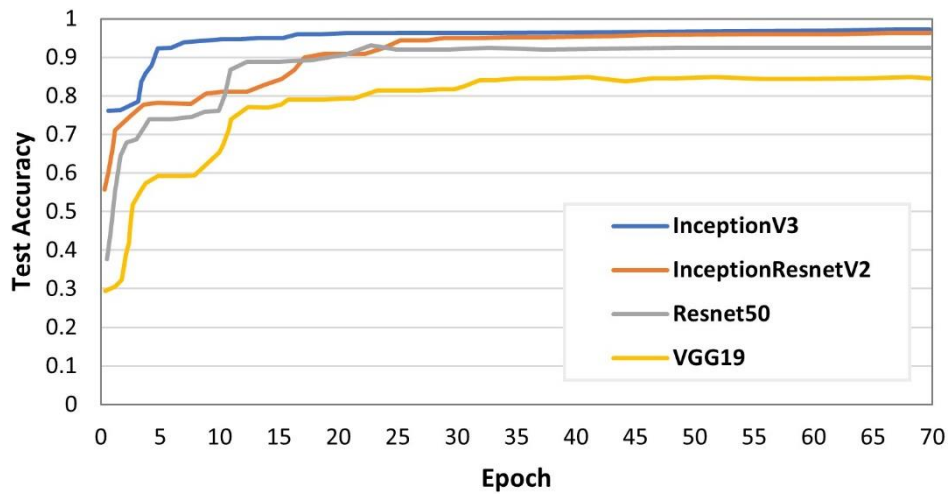
- ۱) تعریف تابع هدف با در نظر گرفتن دقت مدل در مجموعه اعتبارسنجی برای بررسی عملکرد شبکه در یافتن بهینه.
- ۲) تعریف مجموعه پارامترها با مشخص کردن پارامترهایی که قصد تنظیم آن‌ها را داریم (نرخ یادگیری/تعداد دوره‌ها) و مشخص کردن محدوده‌ای از این پارامترها (به عنوان مثال بازه ۰.۱ تا ۰.۰۰۰۱ برای نرخ یادگیری و بازه ۱ تا ۱۰۰ برای تعداد دوره‌ها)
- ۳) قراردادن تعدادی گرگ خاکستری به صورت تصادفی در محدوده‌ی مشخص شده، به گونه‌ای که هر گرگ خاکستری بیانگر یک موقعیت در محدوده‌ی نرخ یادگیری/تعداد دوره‌ها باشد.
- ۴) آموزش شبکه پیشنهادی برای هر گرگ خاکستری بر اساس نرخ یادگیری/تعداد دوره‌های مشخص شده و ارزیابی عملکرد شبکه بر اساس تابع هدف در نظر گرفته شده.
- ۵) به روزرسانی مکان‌های گرگ خاکستری بر اساس رفتار طبیعی گرگ‌ها.
- ۶) به روزرسانی پارامترهای همگرایی و کاوش بر اساس مقدار عملکرد گرگ‌ها.
- ۷) انجام عملیات‌های آمیزش (crossover) و جهش (mutation) به منظور حفظ تنوع جمعیت و کاوش در نقاط بهینه‌ی جدید
- ۸) تکرار مراحل ۴ تا ۷ و به روزرسانی جمعیت در هر نسل تا رسیدن به نرخ بهینه یادگیری/تعداد بهینه دوره‌ها در آموزش شبکه پیشنهادی

**شکل ۴.** شبه کد الگوریتم بهینه‌ساز گرگ خاکستری برای نرخ یادگیری و تعداد دوره‌ها

در روش پیشنهادی، نرخ یادگیری در آموزش مدل پیشنهادی، به عنوان مقدار بهترین راه‌حل به دست آمده توسط الگوریتم بهینه‌ساز گرگ خاکستری، در نظر گرفته شده و نرخ یادگیری به دست آمده به عنوان ابر پارامتر بهینه در این روش، برابر با  $10^{-4}$  می‌باشد.

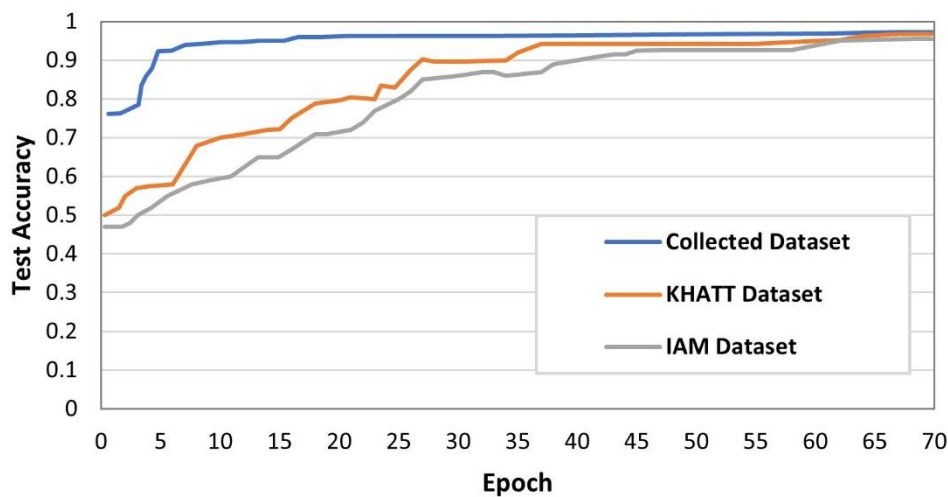
تعداد دوره‌های بهینه در نظر گرفته شده توسط الگوریتم بهینه‌ساز گرگ خاکستری نیز برای آموزش رویکرد پیشنهادی ۷۰ می‌باشد.

مدل پیشنهادی با استفاده از کتابخانه Kersa، بر روی NVIDIA GeForce GTX 1070 8GB اجرا گردید. برای دستیابی به بالاترین عملکرد، از چهار مدل پیشگی از پیش آموزش دیده شامل VGG19، Resnet50، InceptionV3 و InceptionResnetV2 در ساختار پیشنهادی خود به عنوان زیرشبکه‌های یکسان در شبکه سیامی استفاده نمودیم. مقایسه دقت بر روی داده‌های آزمایش از مجموعه داده‌ی جمع‌آوری شده که در شکل ۵ نشان داده شده است، بیانگر پیشرو بودن نتایج InceptionV3 با دقت ۹۷/۲۱ درصد است.



شکل ۵. مقایسه عملکرد مدل پیشنهادی در صورت استفاده از چهار مدل از پیش آموزش دیده شده‌ی مختلف در ساختار شبکه سیامی

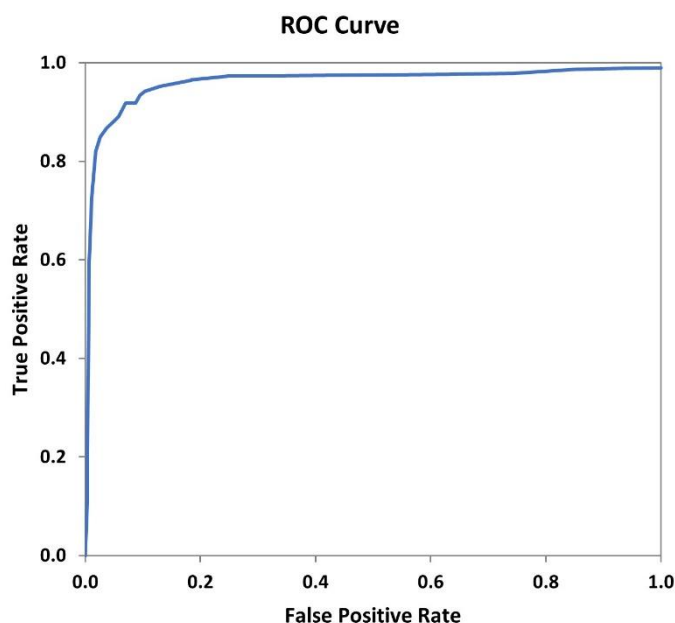
شکل ۶، نتایج ارزیابی مدل پیشنهادی بر روی هر سه مجموعه داده‌ی فارسی جمع‌آوری شده، IAM و KHATT به ترتیب با دقت ۹۷/۲۱ درصد، ۹۵/۴۷ درصد و ۹۶/۸۲ درصد را نشان می‌دهد.



شکل ۶. عملکرد مدل پیشنهادی بر روی سه مجموعه داده‌ی فارسی جمع‌آوری شده، IAM و KHATT

روشی دیگر برای ارزیابی کارایی و قابل اعتماد بودن رویکرد پیشنهادی، ترسیم نمودار مشخصه عملکرد سیستم (ROC<sup>۱</sup>) بر روی داده‌های آزمایش می‌باشد که نرخ مثبت صحیح را در مقابل نرخ مثبت کاذب ترسیم می‌کند. همان‌طور که از شکل ۷ قابل استنباط است، نرخ مثبت واقعی، قابل قبول می‌باشد.

<sup>۱</sup> Receiver Operating Characteristic



شکل ۷. نمودار ROC برای مدل پیشنهادی

## ۵ نتیجه گیری و پیشنهادات

در این مقاله، برای نخستین بار، از یک معماری عمیق بهینه شده برای تایید هویت نویسنده از روی تصاویر داده‌های دست‌نویس فارسی استفاده گردید. با وجود آن‌که دسترسی به اطلاعات و ویژگی‌های بیشتری از سبک دست‌خط نویسندگان از روی تصاویر کامل صفحات یا خطوط در متون دست‌نویس، بدیهی است، اما تایید هویت نویسنده از روی مقدار کمی از متن شامل یک یا دو کلمه می‌تواند در برنامه‌های کاربردی در دنیای واقعی از اهمیت ویژه‌ای برخوردار باشد. لذا در معماری پیشنهادی، مدل‌های گوناگون از شبکه‌های پیچشی از پیش آموزش داده‌شده با بهره‌مندی از یادگیری انتقالی، برای استخراج ویژگی‌های سبک دست‌خط از روی جفت تصاویر کلمات ورودی، مورد استفاده و بررسی قرار گرفتند و تلاش شد تا با استفاده از الگوریتم بهینه‌ساز گرگ خاکستری، برخی از ابرپارامترهای لازم برای آموزش مدل به منظور دستیابی به دقت بالاتر، تنظیم گردند. آزمایشات و ارزیابی‌های صورت گرفته بر روی مجموعه داده‌ی جمع‌آوری شده به زبان فارسی و نیز دو مجموعه‌ی شناخته‌شده به زبان‌های انگلیسی و عربی، بیانگر کارآمدی رویکرد پیشنهادی در تایید هویت نویسنده از روی تصاویر کلمات متون دست‌نویس در زبان‌های گوناگون در حوزه‌ی تجزیه و تحلیل اسناد پزشکی قانونی و اسناد تاریخی و اداری است. در پژوهش‌های آینده، تلاش خواهد شد تا از الگوریتم‌های بهینه‌سازی مختلف در تنظیم پارامترهای شبکه علاوه بر تنظیم ابرپارامترها برای آموزش بهتر مدل و دستیابی به عملکرد بالاتر استفاده گردد.

## منابع

- [1] Parcham, E., Ilbeygi, M., Amini, M., (2021). CBCapsNet: A novel writer-independent offline signature verification model using a CNN-based architecture and capsule neural networks, *Expert Systems with Applications*, 185.
- [2] Shaikh, M. A., Chauhan, M., Chu, J., Srihari, S., (2018). Hybrid feature learning for handwriting verification. In 2018 IEEE 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), 187-192.
- [3] Khan, F. A., Khelifi, F., Tahir, M. A., Bouridane, A., (2018). Dissimilarity gaussian mixture models for efficient offline handwritten text-independent identification using sift and rootsift descriptors. *IEEE Transactions on Information Forensics and Security*, 14, 289–303.
- [4] Chahi, A., Ruichek, Y., Touahni, R., (2018). Block wise local binary count for off-line text-independent writer identification. *Expert Systems with Applications*, 93, 1-14.
- [5] Çalik, N., Kurban, O. C., Yilmaz, A. R., Yildirim, T., Durak Ata, L., (2019). Large-scale offline signature recognition via deep neural networks and feature embedding. *Neurocomputing*, 359, 1–14.
- [6] Xiong, Y. J., Lu, Y., Wang, P.S.P., (2017). Off-line text-independent writer recognition: A survey. *International Journal of Pattern Recognition and Artificial Intelligence*, 31 (5), 1756008.
- [7] Falah Rad, M., Shakeri, M., Khoshhal Roudposhti, K., Shakerinia, I., (2022). Elderly Daily Activity-Based Mood Quality Estimation Using Decision-Making Methods and Smart Facilities (Smart Home, Smart Wristband, and Smartphone). *Journal of Operational Research and Its Applications (Applied Mathematics)-Lahijan Azad University*, 19(2), 1-21. In Persian.
- [8] Diaz, M., Ferrer, M. A., Impedovo, D., Malik, M. I., Pirlo, G., Plamondon, R., (2019). A perspective analysis of handwritten signature technology. *ACM Computing Surveys (CSUR)*, 51(6), 1–39.
- [9] Yilmaz, M. B., Yanikoglu, B., (2016). Score level fusion of classifiers in off-line signature verification. *Information Fusion*, 32, 109–119.
- [10] Bensefia, A., Paquet, T., (2016). Writer verification based on a single handwriting word samples. *EURASIP Journal on Image and Video Processing*, 1, 1-9.
- [11] Jang, W., Kim, S., Kim, Y., Lee, E. C., (2018). Automated Verification method of Korean word handwriting using geometric feature. In *Advances in Computer Science and Ubiquitous Computing*, 1340-1345.
- [12] Hafemann, L. G., Sabourin, R., Oliveira, L. S., (2017). Learning features for offline handwritten signature verification using deep convolutional neural networks. *Pattern Recognition*, 70, 163–176.
- [13] Ghosh, R., (2021). A Recurrent Neural Network based deep learning model for offline signature verification and recognition system. *Expert Systems with Applications*, 168, 114249.
- [14] Yapıcı, M. M., Tekerek, A., Topaloğlu, N., (2021). Deep learning-based data augmentation method and signature verification system for offline handwritten signature. *Pattern Analysis and Applications*, 24(1), 165-179.
- [15] Dey, S., Dutta, A., Toledo, J. I., Ghosh, S. K., Lladós, J., Pal, U., (2017). Signet: Convolutional siamese network for writer independent offline signature verification. *ArXiv Preprint*, ArXiv:1707.02131.
- [16] Maergner, P., Pondenkandath, V., Alberti, M., Liwicki, M., Riesen, K., Ingold, R., Fischer, A., (2019). Combining graph edit distance and triplet networks for offline signature verification. *Pattern Recognition Letters*, 125, 527–533.
- [17] Adak, C., Marinai, S., Chaudhuri, B. B., Blumenstein, M., (2018). Offline bengali writer verification by PDF-CNN and Siamese net. In 2018 IEEE 13th IAPR International Workshop on Document Analysis Systems (DAS), 381-386.
- [18] Javidi, M., Jampour, M., (2020). A deep learning framework for text-independent writer identification. *Engineering Applications of Artificial Intelligence*, 95.

- [19] Chakladar, D. D., Kumar, P., Roy, P. P., Dogra, D. P., Scheme, E., Chang, V., (2021). A multimodal-Siamese Neural Network (mSNN) for person verification using signatures and EEG, *Information Fusion*, 71, 17-27.
- [20] Pei, M., Yan, B., Hao, H., Zhao, M., (2023). Person-Specific Face Spoofing Detection Based on a Siamese Network, *Pattern Recognition*, 135, 109148.
- [21] Mirjalili, S., Mirjalili, S.M. & Lewis, A., (2014). Grey wolf optimizer. *Advances in Engineering software*, 69, 46–61.
- [22] Seyedi, I., Hamed, M., Tavakkoli-Moghaddam, R., (2022). Enhancing the Search Capability of the Imperialist Competitive Algorithm for Truck Scheduling Problem in the Cross-Docking System. *Journal of Operational Research and Its Applications (Applied Mathematics)-Lahijan Azad University*, 19(4), 37-61. In Persian.
- [23] Marti, U. V., Bunke, H., (2002). The IAM-database: an English sentence database for offline handwriting recognition, *International Journal on Document Analysis and Recognition*, 5(1), 39–46.
- [24] Mahmoud, S. A., Ahmad, I., Alshayeb, M., Al-Khatib, W. G., Parvez, M. T., Fink, G. A., Margner, V., Abed, H., (2012). KHATT: Arabic offline handwritten text database. in: *13th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, 447–452.