Computational Intelligence in Electrical Engineering
Vol. 13, No. 4, 2023
Research Paper

Offline Identification of the Author using Heterogeneous Data based on Deep Learning

Seyed Nadi Mahamed Khosroshahi, Seyed Naser Razavi [*], Amin Babazadeh Sangar, Kambiz Majidzadeh

Ph.D. Student, Department of IT and Computer Engineering, Urmia Branch, Islamic Azad University, Urmia, Iran.

Abstract:

Handwriting recognition has always been a challenge; therefore, it has attracted the attention of many researchers. The present study presents an offline system for the automatic detection of human handwriting under different experimental conditions. This system includes input data, image processing unit, and output unit. In this study, a right-to-left dataset is designed based on the standards of the American Society for Experiments and Materials (ASTM). An improved deep convolution neural network (DCNN) model based on a pre-trained network is designed to extract features hierarchically from raw handwritten data. A significant advantage in this study is the use of heterogeneous data. Another significant aspect of the present study is that the proposed DCNN model is independent of any particular language and can be used for different languages. The results show that the proposed DCNN model has a very good performance for identifying the author based on heterogeneous data.

Keywords: Offline Identification of the Author, Heterogeneous Data, Feature Learning, Deep Neural Network.

https://doi.org/10.22108/isee.2021.127816.1459

مقاله پژوهشی

شناسایی آفلاین (غیر برخط) نویسنده با استفاده از داده‌های نامتجانس دستخط بر پایهٔ یادگیری عمیق

سید نادی محامد خسروشاهی[۱]، سید ناصر رضوی[۲*]، امین بابازاده سنگر[۳] و کامبیز مجیدزاده[٤]

۱- دانشجوی دکتری، دانشکده مهندسی کامپیوتر و فناوری اطلاعات – واحد ارومیه – دانشگاه آزاد اسلامی- ارومیه – ایران
n.khosroshahi@iaurmia.ac.ir

۲- استادیار، دانشکده مهندسی کامپیوتر و فناوری اطلاعات – واحد ارومیه – دانشگاه آزاد اسلامی- ارومیه – ایران
n.razavi@tabrizu.ac.ir

۳- استادیار، دانشکده مهندسی کامپیوتر و فناوری اطلاعات – واحد ارومیه – دانشگاه آزاد اسلامی- ارومیه – ایران
Bsamin2@liveutm.onmicrosoft.com

٤- استادیار، دانشکده مهندسی کامپیوتر و فناوری اطلاعات – واحد ارومیه – دانشگاه آزاد اسلامی- ارومیه – ایران
K.Majidzadeh@iaurmia.ac.ir

چکیده: تشخیص دستخط همواره مسئله چالش‌برانگیزی بوده است؛ ازاین‌رو، توجه محققان زیادی را به خود جلب کرده است. مطالعهٔ حاضر یک سیستم آفلاین (غیر برخط) تشخیص خودکار دست‌نوشته‌های انسان را در شرایط آزمایشی مختلف ارائه می‌دهد. این سیستم شامل داده‌های ورودی، واحد پردازش تصویر و واحد خروجی است. در این مطالعه، یک مجموعه داده راست به چپ بر پایهٔ استانداردهای آمریکایی (ASTM) طراحی شده است. یک مدل شبکهٔ عصبی کانولوشن عمیق (DCNN) بهبودیافته بر پایهٔ شبکهٔ از پیش آموزش‌دیده، برای استخراج ویژگی‌ها به‌صورت سلسله‌مراتبی از داده‌های خام دستخط طراحی شده است. یک مزیت درخور توجه در این مطالعه استفاده از داده‌های نامتجانس است. یکی دیگر از جنبه‌های شایان توجه مطالعهٔ حاضر این است که مدل پیشنهادی DCNN مستقل از هر زبان خاصی است و می‌تواند برای زبان‌های مختلف استفاده شود. نتایج نشان می‌دهند مدل پیشنهادی DCNN، عملکرد بسیار خوبی برای شناسایی نویسنده بر پایهٔ داده‌های نامتجانس دستخط دارد.

واژه‌های کلیدی: شناسایی آفلاین نویسنده، داده نامتجانس، یادگیری ویژگی، شبکهٔ عصبی عمیق.

۱- مقدمه

در طول زمان‌های گذشته، از دستخط برای ایجاد ارتباط میان مردم استفاده شده که شامل نمادهایی برای نمایش زبان‌های مختلف است [۱-۲]. دست‌نوشته‌ها ویژگی‌های قابل اندازه‌گیری دارند که می‌توانند نویسندگان را توصیف کنند [۳]. شناسایی دست‌نوشته‌ها یکی از فعال‌ترین حوزه‌های تحقیقاتی در زمینهٔ بینایی ماشین و پردازش الگو است. از کاربردهای تشخیص دست‌نوشته‌ها می‌توان به دسته‌بندی نامه‌های پستی، خواندن خودکار مبالغ چک‌های بانکی، شناسایی خودکار اطلاعات ثبت‌شده در فرم‌ها، شناسایی اسناد دست‌نویس جعلی و غیره اشاره کرد. در تمامی این کاربردها، همواره مسئله دقت و سرعت سیستم اهمیت بسیار زیادی داشته است. جرم جعل در اسناد

[۱] تاریخ ارسال مقاله: ۱۳۹۹/۱۲/۲٤
تاریخ پذیرش مقاله: ۱٤۰۰/۰۸/۱۷
نام نویسندهٔ مسئول: سید ناصر رضوی
نشانی نویسندهٔ مسئول: : ایران – ارومیه- دانشگاه آزاد اسلامی- دانشکده مهندسی کامپیوتر و فناوری اطلاعات

دست‌نویس به سرعت درحال‌رشد است؛ بنابراین، شناسایی اسناد دست‌نویس جعلی به‌عنوان یک مشکل بسیار حیاتی برای سازمان‌های دولتی (مانند دادگستری و دفاتر ثبت اسناد) در نظر گرفته می‌شود. به‌طور کلی، شناسایی اسناد جعلی پیچیدگی زیادی دارد. روش عملی در کشف اسناد جعلی شامل دو مرحله است؛ به دست آوردن نمونه‌های دستخط از مظنونین و استخراج تفاوت بین ویژگی‌های دستخط. تمامی روش‌ها بر پایهٔ استخراج ویژگی‌ها از سند اصلی و مقایسهٔ آنها با سند جعلی استوار است. ویژگی‌های دستخط به سبک‌های نوشتاری زبان بستگی دارد. این به‌عنوان یک نشانهٔ اساسی در تجزیه و تحلیل اسناد دست‌نویس در نظر گرفته می‌شود.

بیشتر محققان زبان انگلیسی را برای تجزیه و تحلیل اسناد دست‌نویس انتخاب کرده‌اند [٤-٥]. با این حال، تعداد کمی از محققان روی زبان‌های راست به چپ مانند عربی، فارسی، اردو و غیره تمرکز کرده‌اند؛ بنابراین، مجموعه داده‌های استاندارد بسیار کمی در زبان‌های راست به چپ وجود دارد. زبان‌های راست به چپ به‌طور گسترده در میان اقشار مختلفی از مردم جهان استفاده می‌شوند و ازنظر چگونگی پیوند چند حرف با یکدیگر سبک نوشتاری خاصی دارند. مشکل اصلی در تحلیل دست‌نوشته‌های راست به چپ، کمبود مجموعه داده‌های جامع است. مجموعه داده‌ها شامل انواع حروف، کلمات، جملات، اعداد و انواع مختلف اتصالات است. یک دلیل و توجیه خلأ تحقیق دربارهٔ زبان‌های راست به چپ این است که آنها از روش‌های پیچیده‌ای برای اتصال حروف استفاده می‌کنند؛ ازاین‌رو، آنها به‌عنوان سبک‌های نوشتاری ناشناخته و کم مطالعه باقی مانده‌اند.

به‌طور کلی، مطالعات تحقیقاتی در زمینهٔ اسناد دست‌نویس، به سه گروه عمده طبقه‌بندی می‌شوند: تشخیص دستخط[1]، شناسایی نویسنده[2] و تأیید نویسنده[3]. نیاز سازمان‌های دولتی (مانند دادگستری و دفاتر ثبت اسناد) عمدتاً مربوط به شناسایی و تأیید نویسنده است. شناسایی و تأیید نویسنده معمولاً توسط متخصصان براساس شناسایی بصری انجام می‌شوند که این کار زمان‌بر، خسته‌کننده و نادرست است. مطالعات تحقیقاتی بسیاری برای حل این مشکلات انجام شده‌اند. برخی از مطالعات بر پایهٔ رویکردهای سنتی یادگیری ماشین و برخی دیگر بر پایهٔ رویکردهای جدید یادگیری ماشین، مانند یادگیری عمیق [٦-٧] استوارند. بسیاری از محققان روش‌های طبقه‌بندی را برای مسئله شناسایی نویسنده پیشنهاد داده‌اند [٨-٩]. استخراج ویژگی در یادگیری ماشین نقشی اساسی دارد و یک زمینه مهم در ادبیات فن محسوب می‌شود. در رویکردهای سنتی یادگیری ماشین، مراحل استخراج و طبقه‌بندی ویژگی‌ها از یکدیگر جدا هستند؛ در حالی که در رویکردهای جدید یادگیری ماشین، این مراحل با هم ادغام شده‌اند. این رویکردها، برخلاف رویکردهای سنتی، نیازی به دانش قبلی از مسئله ندارند و می‌توانند به‌صورت سلسله‌مراتبی ویژگی‌های مطلوب را از داده‌های خام استخراج کنند. به‌طور کلی، تشخیص دست‌نوشته‌ها به‌صورت مرسوم به دو دسته آنلاین (برخط) و آفلاین (غیر برخط) تقسیم می‌شوند. در تشخیص آنلاین ترتیب زمانی از مختصات‌ها دریافت می‌شود که بیان‌کنندهٔ حرکات نوک قلم شخص است؛ در حالی که در روش‌های آفلاین تنها تصویر متن در دسترس است [١٠-١١].

آوایدا و همکاران [١٢] یک روش شناسایی نویسنده بر پایهٔ متون عربی با استفاده از ویژگی‌های آماری و ساختاری ارائه دادند. در این روش الگوریتم نزدیک‌ترین همسایه به همراه معیارهای فاصله اقلیدسی استفاده شد. همچنین، از الگوریتم‌های کاهش داده، برای کاهش ابعاد داده استفاده شد. دویست و پنجاه نویسنده، یک پایگاه دادهٔ ۵۰۰ پاراگرافی را به زبان عربی نوشته‌اند. نتایج نشان‌دهندهٔ عملکرد مطلوب روش بیان‌شده در شناسایی نویسنده بود. شهابی و همکاران [١٣] روشی را برای شناسایی آفلاین اسناد فارسی دست‌نویس با استفاده از فیلتر چندکاناله گابور[4] ارائه دادند. این روش می‌توانست ویژگی‌ها را استخراج کند و یک مجموعه دادهٔ محدود را با توجه به معیارهای فاصله اقلیدسی طبقه‌بندی کند. با روش بیان‌شده عملکرد خوبی دربارهٔ اسناد دست‌نویس فارسی حاصل شد. باغشاه و همکاران [١٤] روش جدیدی را برای شناسایی آنلاین (برخط) دست‌نوشته‌های فارسی ارائه دادند. در روش بیان‌شده، ویژگی‌ها با توجه به معیارهای فاصله اقلیدسی

استخراج شدند و پس از آن یک طبقه‌بند فازی برای شناسایی اسناد دست‌نویس استفاده شد. این روش به دست‌نوشته‌های زبان فارسی اعمال شد و در مقایسه با سایر رویکردهای موجود به صحت بالاتری از شناسایی دست یافت. احمد و همکاران [٤] یک طبقه‌بند مبتنی بر فاصله را با مجموعه‌ای از ویژگی‌های استخراج‌شده با الگوریتم مور[٥] به‌منظور شناسایی آفلاین نویسنده بررسی کردند. این روش به چهار مجموعه دادهٔ جامع اعمال شد و به میزان دقت چشمگیری در شناسایی اسناد دست‌نویس دست یافت. وو و همکاران [۱۵] روش شناسایی آفلاین متون را بر پایهٔ طبقه‌بند مبتنی بر فاصله پیشنهاد کردند. در این مطالعه برای استخراج ویژگی‌های ساختاری از شش مجموعه دادهٔ مختلف، از فیلتر ایزوتروپیک استفاده شد. سه مورد از این مجموعه داده‌ها مربوط به زبان انگلیسی، یکی مربوط به زبان چینی و دو مورد دیگر مربوط به ترکیبی از این زبان‌ها بود. نتایج تجربی نشان دادند روش بیان‌شده در شناسایی متون بهتر از سایر روش‌های مقایسه‌ای عمل می‌کند. کومار و همکاران [١٦] یک روش شناسایی آفلاین دستخط را بررسی کردند. در این مطالعه، پنج نوع ویژگی استخراج‌شده از متن‌های دست‌نویس ارزیابی شدند. برای کاهش ابعاد ویژگی‌ها، از روش تجزیه و تحلیل خطی فیشر[٦] و آنالیز اجزای اصلی[٧] استفاده شد. با استفاده از ماشین بردار پشتیبان[٨] (SVM) و شبکهٔ عصبی کارایی روش بیان‌شده تأیید شد. روش‌های شناسایی نویسنده مبتنی بر یادگیری ماشین سنتی به‌طور گسترده استفاده شده‌اند؛ اما مشکلات متعددی دارند. برخی از این مشکلات عبارت‌اند از وابسته‌بودن به دانش تخصصی، حساسیت به تغییرات شرایط محیطی و محدودیت استخراج ویژگی‌های جدید. براساس مشکلات بیان‌شده، لازم است روش‌های خودکار شناسایی نویسنده براساس رویکردهای جدید یادگیری ماشین، مانند رویکردهای یادگیری عمیق بررسی شوند که ویژگی‌های مطلوب مربوط به هر مسئله را ازطریق داده‌های خام به‌صورت سلسله‌مراتبی می‌توانند بیاموزند [٦-٧].

منچالا و همکاران [۱۷] سیستمی را برای شناسایی دستخط با استفاده از یادگیری عمیق ارائه دادند. این سیستم برای یافتن شباهت‌ها و همچنین تفاوت‌ها در میان نمونه‌های مختلف دستخط آموزش دید. در این سیستم تصاویر دست‌نویس به متون دیجیتالی تبدیل شدند. نتایج نشان دادند این سیستم برای متن‌هایی که نویز کمتری دارند بهترین دقت را ارائه می‌دهد. همچنین دقت سیستم بیان‌شده کاملاً به مجموعه داده بستگی دارد و در صورت افزایش داده‌ها می‌توان با این سیستم به دقت بیشتری دست یافت. آداک و همکاران [۸] روش شناسایی و تأیید نویسنده را از دست‌نوشته‌های آفلاین بنگالی[٩] بررسی کردند. در این روش، برخی از ویژگی‌های مهندسی از این دست‌نوشته‌ها، استخراج و با استفاده از مدل ماشین بردار پشتیبان ارزیابی شدند. ویژگی‌های خودکار با استفاده از مدل شبکهٔ عصبی کانولوشنال عمیق[١٠] (DCNN) نیز از این دست‌نوشته‌ها استخراج شدند. در این مطالعه دو پایگاه داده از دو مجموعه مختلف با ۱۰۰ نویسنده برای آزمایش طراحی شدند. پس از آزمایش مشاهده شد مدل شبکهٔ عصبی کانولوشنال عمیق در مقایسه با سایر مدل‌ها نتایج بهتری ارائه می‌دهد. ژانگ و همکاران [۱۸] یک شبکهٔ عصبی بازگشتی[١١] (RNN) را برای شناسایی آنلاین نویسنده بررسی کردند. داده‌های دست‌نویس هر نویسنده با مجموعه‌ای از RHSها[١٢] نشان داده شدند. از یک مدل شبکهٔ عصبی با حافظه کوتاه‌مدت دو جهته برای رمزگذاری هر RHS (در یک بردار با طول ثابت) برای طبقه‌بندی استفاده شد. آزمایش‌های مربوطه روی مجموعه داده‌های انگلیسی (۱۳۳ نویسنده) و چینی (١٨٦ نویسنده) انجام و مزایای روش آنها در مقایسه با سایر روش‌های پیشرفته تأیید شدند. کاربون و همکاران [۱۹] یک سیستم تشخیص آنلاین دستخط را ارائه دادند که می‌تواند ۱۰۲ زبان مختلف را با استفاده از شبکهٔ عصبی عمیق[١٣] پشتیبانی کند. این سیستم روش‌های شناسایی متوالی را با رمزگذاری جدید ورودی با استفاده از منحنی‌های Bézier ترکیب کرده است. نتایج نشان دادند سیستم بیان‌شده در مقایسه با روش‌های دیگر نتایج بهتری ارائه می‌دهد. جاویدی و همکاران [۲۰] روشی را برای شناسایی نویسنده (مستقل از متن) بر پایهٔ یادگیری عمیق ارائه دادند. در این مطالعه یک نسخهٔ توسعه‌یافته از ResNet با ترکیب شبکه‌های residual عمیق و یک توصیف‌کنندهٔ دستخط سنتی برای تجزیه و تحلیل دستخط استفاده شده است.

توصیف‌کننده ضخامت دستخط را به‌عنوان یک ویژگی اولیه و ضروری دستخط تجزیه و تحلیل می‌کند. این روش می‌تواند هویت نویسنده مستقل از متن را ارائه دهد که برای یادگیری مدل خود نیازی به محتوای دست‌نویس یکسان ندارد. رویکرد پیشنهادی روی مجموعه داده‌های عمومی و مشهور ارزیابی شد. نتایج نشان دادند شبکهٔ ترکیبی پیشنهادی نسبت به روش‌های مقایسه‌ای بهتر عمل می‌کند و می‌تواند برای برنامه‌های کاربردی در دنیای واقعی استفاده شود. یانگ و همکاران [۲۱] یک روش یادگیری عمیق را برای شناسایی نویسنده بر پایهٔ زبان چینی با استفاده از ترکیب ویژگی‌ها ارائه دادند. در این مطالعه از ترکیب ویژگی‌های عمیق و ویژگی‌های دستی برای به دست آوردن ویژگی‌های دستخط از تصاویر دست‌نویس استفاده شد. نتایج نشان دادند این روش عملکرد بهتری در شناسایی حروف چینی نسبت به سایر روش‌های مقایسه‌ای دارد. وانگ و همکاران [۲۲] یک روش شناسایی خودکار نویسنده را بر پایهٔ یادگیری عمیق ارائه دادند. در این مطالعه، ترکیبی از شبکه‌های u-net و resnet به‌عنوان مدل پیشنهادی در نظر گرفته شد. روش پیشنهادی آنها روی مجموعه دادهٔ ICDAR17 ارزیابی شد و نتایج بهتری نسبت به مدل‌های مقایسه‌ای ارائه داد.

بررسی مطالعات شناسایی نویسنده نشان می‌دهد اگرچه تاکنون مطالعات زیادی در این زمینه انجام شده است، محدودیت‌هایی در این مطالعات وجود دارد. در بیشتر این مطالعات، از روش‌های سنتی مبتنی بر استخراج و انتخاب ویژگی‌ها برای شناسایی نویسنده استفاده شده است. علاوه بر این، در بیشتر این مطالعات شرایط محیطی مختلف در تهیه پایگاه داده‌های مختلف در نظر گرفته نشده است؛ در صورتی که برای ورود به حوزهٔ کاربردی، در نظر گرفتن کلیه شرایط محیطی ضروری است. براساس این، اولین هدف این مقاله ارائه یک سیستم شناسایی آفلاین نویسنده در شرایط آزمایشی مختلف است که مستقل از هر زبانی باشد. همچنین، با بررسی مطالعات پیشین مشاهده می‌شود مجموعه دادهٔ جامعی در رابطه با دست‌نوشته‌های راست به چپ وجود ندارد که بتواند به‌عنوان یک مجموعه پایگاه دادهٔ مرجع برای بررسی زبان‌های راست به چپ استفاده شود. درواقع دست‌نوشته‌های راست به چپ به‌طور عمده در مطالعات مرتبط نادیده گرفته شده‌اند. براساس این، در دومین هدف این مطالعه، تلاش شده است با تمرکز بر شناسایی نویسنده براساس دست‌نوشته‌های راست به چپ، این شکاف تحقیقاتی بررسی شود که به‌عنوان یک موضوع بسیار مهم و بحث‌برانگیز در سازمان‌های دولتی (مانند دادگستری و دفاتر ثبت اسناد) شناخته می‌شود. برای این منظور، یک مجموعه دادهٔ راست به چپ شامل کلمات، جملات و اعداد جمع‌آوری شده است. این مجموعه داده شامل ۸٦۳۰٤ نمونه از افراد مختلف با جنسیت، گروه سنی، شغل و سطح تحصیلات مختلف است. این مجموعه داده در فواصل زمانی مختلف در شرایط آزمایشی مختلف براساس استانداردهای آمریکایی (ASTM) جمع‌آوری شده است [۲۳]. همچنین، یادگیری عمیق به‌طور گسترده و با موفقیت زیادی در تجزیه و تحلیل تصاویر و سیگنال‌ها استفاده شده است. در سومین هدف این مقاله، یک مدل شبکهٔ عصبی کانولوشن عمیق بهبودیافته بر پایهٔ شبکه از پیش آموزش‌دیده[۱٤] طراحی شده است تا ویژگی‌ها را به‌صورت سلسله‌مراتبی از داده‌های خام دستخط یاد بگیرد. مهم‌ترین جنبهٔ مدل پیشنهادی قابلیت آن در طبقه‌بندی مجموعه داده‌های نامتجانس است؛ به این معنا که در هر دوره، اگرچه نمونه‌های تصادفی برای مراحل آموزش و ارزیابی، به یک شخص خاص تعلق داشته است، ممکن است لزوماً یکسان نباشند؛ حتی ممکن است هیچ شباهتی نداشته باشند. استفاده از نمونه‌های نامتجانس چهارمین هدف این مقاله است که به‌طور عمده در مطالعات قبلی نادیده گرفته شده است. درواقع، این نوآوری در روش شناسایی، برجسته‌ترین جنبهٔ مطالعهٔ حاضر است.

ادامهٔ مقاله به‌صورت زیر تدوین شده است؛ در بخش ۲، شبکه‌های عصبی کانولوشنال و بازگشتی بررسی می‌شوند. در بخش ۳، روش پیشنهادی برای شناسایی نویسنده ارائه می‌شود. در بخش ٤ نتایج شبیه‌سازی بررسی می‌شوند و درنهایت، بخش ۵ مربوط به نتیجه‌گیری است.

۲- مواد و روش‌ها

در این بخش ابتدا شبکه‌های عصبی کانولوشنال[۱۵] (CNN) و پس از آن، شبکه‌های حافظه طولانی کوتاه‌مدت[۱۶] (LSTM) بررسی می‌شوند که زیرمجموعه‌ای از شبکه‌های عصبی بازگشتی‌اند.

۲-۱- شبکه‌های عصبی کانولوشنال

شبکهٔ عصبی کانولوشنال، درواقع یک شبکهٔ عصبی بهبودیافته است. در این شبکه، چندین لایه با روشی قدرتمند در کنار هم آموزش می‌بینند [۲۴]. این روش، بسیار کارآمد بوده و یکی از رایج‌ترین روش‌ها در کاربردهای مختلف بینایی ماشین است. همانند شبکه‌های عصبی مصنوعی[۱۷]، تصمیم خروجی نهایی شبکهٔ عصبی کانولوشنال براساس وزن و بایاس لایه‌های قبلی در ساختار شبکه است.

در این شبکه، دو مرحله برای آموزش وجود دارد؛ مرحلهٔ انتشار پیش‌رو[۱۸] و مرحلهٔ پس‌انتشار[۱۹] (BP) [۲۵].

BP روشی برای محاسبهٔ گرادیان تابع اتلاف نسبت به وزن‌ها است. BP سیگنال‌های خطا را در شبکه حین آموزش پس می‌زند و باعث به‌روزرسانی وزن‌ها می‌شود. در مرحلهٔ اول، داده‌های ورودی به شبکه اعمال می‌شوند و این عمل چیزی به‌جز ضرب نقطه‌ای بین ورودی و پارامترهای هر نورون و اعمال عملیات کانولوشن در هر لایه نیست و درنهایت، خروجی شبکه محاسبه می‌شود. به‌منظور تنظیم پارامترهای شبکه یا به عبارت دیگر آموزش شبکه، از نتیجهٔ خروجی برای محاسبهٔ میزان خطای شبکه استفاده می‌شود. برای این کار، خروجی شبکه با استفاده از یک تابع خطا[۲۰] با پاسخ صحیح، مقایسه و به این ترتیب، میزان خطا محاسبه می‌شود. در مرحلهٔ بعد، براساس میزان خطای محاسبه‌شده، مرحلهٔ پس‌انتشار آغاز می‌شود. در این مرحله، گرادیانت هر پارامتر با توجه به قاعدهٔ زنجیره‌ای محاسبه می‌شود و تمامی پارامترها، با توجه به تأثیرشان بر خطای ایجادشده در شبکه، به‌روزرسانی می‌شوند. بعد از به‌روزرسانی پارامترها، مرحلهٔ بعدی انتشار پیش‌رو آغاز خواهد شد. بعد از تکرار تعداد مناسبی از این مراحل، آموزش شبکه به پایان می‌رسد. در این شبکه، خروجی هر لایه همان ویژگی‌ها هستند که بعد کمتری نسبت به دادهٔ اصلی دارند.

به‌طور کلی، یک شبکه کانولوشنال از سه لایه اصلی تشکیل می‌شود که عبارت‌اند از لایه کانولوشنال، لایه ادغام[۲۱] و لایه تمام متصل[۲۲] (FC) [۲۴]. برای جلوگیری از فرایند بیش‌برازش[۲۳] و بهبود عملکرد شبکه از لایه‌های حذف تصادفی[۲۴] و نرمال‌ساز دسته‌ای[۲۵] نیز استفاده می‌شود. همچنین در شبکه‌های عصبی نیاز است پس از هر لایه از تابع فعال‌سازی استفاده شود که در ادامه، این لایه‌ها و توابع به‌طور خلاصه معرفی می‌شوند.

لایه کانولوشنال: شامل فیلترهایی (کرنل‌هایی) است که روی داده‌های ورودی می‌لغزند. یک کرنل، یک ماتریس است که با دادهٔ ورودی کانوالو می‌شود. این لایه عمل کانولوشن را روی داده‌های ورودی با استفاده از کرنل انجام می‌دهد. خروجی کانولوشن را نگاشت ویژگی می‌نامند. عملگر کانولوشن به شرح زیر است:

$$y_k = \sum_{n=0}^{N-1} x_n h_{k-n} \qquad (۱)$$

که x سیگنال، h فیلتر، N تعداد عناصر در x و y بردار خروجی است.

لایه ادغام: این لایه که به کاهش نمونه[۲۶] نیز معروف است، ابعاد نورون‌های خروجی از لایه کانولوشنال را کاهش می‌دهد و باعث کاهش محاسبات و همچنین جلوگیری از پدیدهٔ بیش‌برازش می‌شود. در این پژوهش از لایه ادغام بیشینه[۲۷] استفاده شده که فقط مقادیر بیشینه در هر نگاشت ویژگی را انتخاب کرده است و باعث کاهش تعداد نورون های خروجی می‌شود.

لایه FC: دارای اتصال کامل به تمامی فعال‌سازی‌ها در لایه قبلی است.

لایه حذف تصادفی: از این لایه به‌منظور جلوگیری از پدیدهٔ بیش‌برازش استفاده می‌شود [۲۵]. نحوهٔ کار آن به این صورت است که در هر مرحله از آموزش، هر نورون با احتمالی از شبکه بیرون انداخته شده است؛ به طوری که درنهایت، یک شبکه کاهش داده شده باقی می‌ماند.

لایه نرمال‌سازی دسته‌ای: این لایه به‌منظور نرمال‌سازی داده‌ها در داخل شبکه انجام می‌شود [۲۶]. زمانی که

محاسبات مختلف روی دادهٔ ورودی اعمال شود، توزیع داده‌ها تغییر خواهد کرد. این لایه با هدف کاهش تغییر کوواریانس داخلی، سرعت آموزش شبکه را افزایش می‌دهد و باعث تسریع در همگرایی می‌شود. تبدیل لایه نرمال‌سازی دسته‌ای به شرح زیر است:

$$\begin{aligned}\mu_B &= \frac{1}{n}\sum_{i=1}^{n} y_i^{(l-1)} \\ \sigma_B^2 &= \frac{1}{n}\sum_{i=1}^{n} (y_i^{(l-1)} - \mu_B)^2 \\ \hat{y}^{(l-1)} &= \frac{y^{(l-1)} - \mu_B}{\sqrt{(\sigma_B^2 + \varepsilon)}} \\ z^{(l)} &= \gamma^{(l)}\, \hat{y}^{(l-1)} + \beta^{(l)}\end{aligned} \qquad (۲)$$

که μ_B و σ_B^2 به‌ترتیب میانگین و واریانس دسته هستند. ε یک ثابت کوچک برای ثبات عددی، l شماره لایه، $y^{(l-1)}$ بردار ورودی به لایه نرمال‌ساز و $z^{(l)}$ بردار خروجی نرمال مربوط به یک نورون است و $\gamma^{(l)}$ و $\beta^{(l)}$ به ترتیب پارامترهای مربوط به مقیاس و تغییر نرخ یادگیری‌اند.

تابع فعال‌سازی: پس از هر لایه کانولوشن، یک تابع فعال‌سازی اعمال می‌شود. تابع فعال‌سازی یک عملگر است که خروجی را به مجموعه‌ای از ورودی‌ها نگاشت می‌کند و برای غیرخطی‌کردن ساختار شبکه استفاده می‌شود [۲۷]. تابع Relu یکی از پرکاربردترین توابع فعال‌سازی است و این ویژگی را دارد که غیرخطی‌بودن را به ساختار شبکه اعمال کند؛ بنابراین، در برابر تغییرات جزئی در ورودی مقاوم است. رابطه ۳ تابع Relu را نشان می‌دهد.

$$f(x) = \begin{cases} x & x > 0 \\ 0 & x \le 0 \end{cases} \qquad (۳)$$

تابع سافت‌مکس[۲۸]: این تابع توزیع احتمالی کلاس‌های خروجی را محاسبه می‌کند که رابطه آن به فرم زیر است:

$$p_i = \frac{e^{x_j}}{\sum_1^k e^{x_K}} \quad \text{for } j = 1, \ldots k \qquad (٤)$$

که در آن، x ورودی شبکه است و مقادیر خروجی p بین صفر و یک هستند که مجموع آنها برابر با یک است.

۲-۲- شبکه‌های عصبی بازگشتی

شبکه‌های عصبی بازگشتی شاخه‌ای مهم از شبکه‌های عصبی عمیق‌اند که به‌منظور تحلیل سیستم‌های پیچیده استفاده می‌شوند. این شبکه‌ها می‌توانند با کاهش ابعاد دادهٔ ورودی X_t، بار محاسباتی را کاهش دهند و همچنین باعث بهبود عملکرد آموزش شوند. علاوه بر این، این شبکه‌ها امکان تلفیق اطلاعات بین ورودی‌های مختلف را به‌منظور دستیابی به ویژگی‌هایی فراهم می‌کنند که نمی‌توان با استفاده از روش‌های سنتی استخراج کرد [۲۸-۳۰]. شبکه حافظه طولانی کوتاه‌مدت ازجمله شبکه‌های عصبی بازگشتی‌اند که به‌منظور رفع ضعف‌های شبکه‌های بازگشتی ازجمله حل مشکل پراکندگی گرادیان یا مشکلات انفجاری گرادیان به کار برده می‌شوند [۲۸-۳۰]. برخلاف شبکهٔ عصبی بازگشتی که صرفاً جمع متوازن سیگنال‌های ورودی را محاسبه می‌کند و سپس از یک تابع فعال‌ساز عبور می‌دهد، هر واحد LSTM از یک حافظه C_t در زمان t بهره می‌برد. یک سلول حافظه از چهار عنصر اصلی تشکیل شده است: یک دروازه ورودی یا دروازه به‌روزرسانی Γ_u، یک نورون با اتصال خودبازگشتی، یک دروازه فراموشی Γ_f و یک دروازه خروجی Γ_o. فعال‌سازی واحد LSTM به‌صورت رابطه زیر تعریف می‌شود [۲۸-۳۰]:

$$h_t = \Gamma_o . \tanh(C_t) \qquad (۵)$$

که در آن، Γ_o دروازه خروجی و کنترل‌کنندهٔ میزان محتوایی است که ازطریق حافظه ارائه می‌شود. دروازه خروجی با رابطه زیر محاسبه می‌شود [۲۸-۳۰]:

$$\Gamma_o = \sigma(W_o.[h_{t-1}, X_t] + b_0) \qquad (٦)$$

که در آن، σ تابع فعال‌سازی سافت‌مکس است و W_o و b_0 به‌ترتیب ماتریس وزن و بردار بایاس اولیه‌اند. سلول حافظه C_t نیز با فراموشی نسبی حافظه فعلی و اضافه‌کردن محتوای حافظه جدید به‌صورت $\tilde{C}_t$ از رابطه (۸) به‌روزرسانی می‌شود که در آن، محتوای حافظه جدید از رابطه (۷) به دست می‌آید [۲۸-۳۰]:

$$\tilde{C}_t = \tanh(W_C.[h_{t-1}, X_t] + b_C) \quad (۷)$$

$$C_t = \Gamma_f.C_{t-1} + \Gamma_u.\hat{C}_t \quad (۸)$$

میزان حافظه فعلی که باید فراموش شود، توسط دروازه فراموشی Γ_f کنترل می‌شود و مقدار حافظه جدید که باید به سلول حافظه اضافه شود، توسط دروازه به‌روزرسانی (دروازه ورودی) Γ_u انجام می‌گیرد. این عملیات در رابطه‌های (۹) و (۱۰) نشان داده شده است [۲۸-۳۰]:

$$\Gamma_f = \sigma(W_f.[h_{t-1}, X_t] + b_f) \quad (۹)$$

$$\Gamma_u = \sigma(W_u.[h_{t-1}, X_t] + bu) \quad (۱۰)$$

شکل ۱ ساختار یک شبکهٔ عصبی بازگشتی LSTM را نشان می‌دهد. در این شبکه که یک ورودی یا همان X_t دارد، دو خروجی تولید می‌شود: یک خروجی C_t و خروجی دیگر h_t؛ h_t به دو بخش تقسیم می‌شود؛ بخشی به گام زمانی بعد، منتقل و بخشی نیز در صورت نیاز به تولید خروجی در گام زمانی فعلی استفاده می‌شود. دروازه فراموشی Γ_f وظیفهٔ کنترل جریان اطلاعات از گام زمانی قبلی را دارد. این دروازه مشخص می‌کند آیا اطلاعات حافظه از گام زمانی قبل استفاده شود یا خیر و اگر باید از گام زمانی قبل چیزی وارد شود، به چه میزان باشد. دروازه به‌روزرسانی Γ_u وظیفهٔ کنترل جریان اطلاعات جدید را بر عهده دارد. این دروازه مشخص می‌کند آیا در گام زمانی فعلی باید از اطلاعات جدید استفاده شود یا خیر و اگر بلی به چه میزان. دروازه خروجی Γ_o نیز مشخص می‌کند چه میزان از اطلاعات گام زمانی قبل با اطلاعات گام زمانی فعلی به گام زمانی بعد منتقل شود.

۳- روش پیشنهادی

در این بخش، روش پیشنهادی مقاله ارائه می‌شود. شکل ۲ بلوک دیاگرام الگوریتم پیشنهادی را نشان می‌دهد.

۳-۱- جمع‌آوری داده‌ها

در روش پیشنهادی یک پایگاه دادهٔ جامع طراحی و توسعه داده شده است. همچنین، در این مطالعه علاوه بر پایگاه دادهٔ طراحی‌شده از پایگاه داده‌های IAM، CVL، KHATT و IFN/ENIT نیز استفاده شده است تا روش پیشنهادی بهتر ارزیابی شود. پایگاه دادهٔ IAM [۳۱] شامل ٤٨٩٩ نمونه از ۱۵۰ نویسنده و پایگاه دادهٔ CVL [۳۲] شامل ۱۸۵٤ نمونه از ۳۰۹ نویسنده به زبان انگلیسی است. پایگاه دادهٔ KHATT [۳۳] شامل ۱۰۸۹۸ نمونه از ۸۲۸ نویسنده و پایگاه دادهٔ IFN/ENIT [۳٤] شامل ۲٦٤٥۹ نمونه از ٤۱۱ نویسنده به زبان عربی است. این پایگاه داده‌ها در مطالعات اخیر بسیار استفاده شده‌اند و در زمینهٔ شناسایی نویسنده از پایگاه داده‌های قابل اعتماد و پرکاربرد هستند.

در پایگاه دادهٔ پیشنهادی، بر پایهٔ استانداردهای ASTM، نمونه‌های دستخط از ٦۲ شرکت‌کننده در بازه‌های زمانی مختلف در شرایط محیطی مختلف جمع‌آوری شدند (شکل ۳ را ببینید). از ٦۲ شرکت‌کننده در این آزمایش، ۳٤ نفر مرد و ۲۸ نفر زن، با میانگین سنی ۲۲ تا ٥٤ سال‌اند. همچنین ٦ نفر از این شرکت‌کنندگان چپ‌دست و ٥٦ نفر راست‌دست بودند. درنهایت، براساس استانداردهای از پیش تعریف‌شده، ویژگی‌های بافتی و ساختاری دستخط به دست آمدند. ویژگی‌های بافتی مربوط به یک سند دست‌نویس به عواملی مانند کاغذ، جوهر، ابزارهای نوشتن و غیره بستگی دارد؛ در صورتی که ویژگی‌های ساختاری به سبک‌های نوشتاری، نحوهٔ پیوستن حروف و غیره اشاره دارد. گفتنی است زبان فارسی به‌عنوان زبان مقصد انتخاب شده که مجموعه دادهٔ مربوط به آن به دست آمده است. انتخاب این زبان به این علت است که محققان به نمونه‌های این زبان دسترسی آسان و کافی داشتند.

هر حرف از زبان فارسی چهار فرم نوشتاری مختلف دارد؛ فرم جداگانه[۲۹] (S)، فرم شروع کلمه[۳۰] (BOW)، فرم میانی کلمه[۳۱] (MOW) و فرم انتهایی کلمه[۳۲] (EOW). انتخاب فرم به موقعیت آن در یک کلمه بستگی دارد. از شرکت‌کنندگان چهار جمله مربوط به چهار فرم بیان‌شده گرفته شد و از آنها خواسته شد تا دستورالعمل‌های زیر را دنبال کنند: در مرحلهٔ اول، هر جمله دوازده بار براساس استاندارد ASTM روی یک برگ کاغذ جداگانه نوشته شد. در مرحلهٔ دوم، هر چهار جمله در یک برگ جداگانهٔ دیگر

نوشته شدند. از دو نوع کاغذ استاندارد متفاوت «PaperOne» و «Double-A» استفاده شد که مشخصات آنها در جدول ۱ نشان داده شده است [۳۵].

در این مطالعه از دو نوع مختلف از خودکارهای استاندارد، به مارک «Schneider» و «Faber-Castell» با رنگ‌های «آبی» و «سیاه» استفاده شد. مشخصات این خودکارها در جدول ۲ آورده شده است [۳۶-۳۷]. نمونه‌هایی از فرم‌های نوشتاری شرکت‌کنندگان در شکل ۳ نشان داده شده‌اند. گفتنی است شرکت‌کنندگان فرم اصلی را به زبان فارسی پر کردند و به‌ازای هر فرد، مجموعه کاملی از نمونه‌ها گرفته شد.

همه نمونه‌ها روی دو پد نوشتاری مختلف نوشته شدند که به‌عنوان پدهای «سخت» و «نرم» هستند. استفاده از این دو نوع پد، برای نشان‌دادن میزان فشار قلم در نظر گرفته شده است. نمونه‌های جمع‌آوری‌شده توسط RICOH Aficio MP 6001 با رزولوشن ۳۰۰ dpi در حالت رنگی اسکن شدند. ترتیب نمونه‌های جمع‌آوری‌شده، اطلاعات و جزئیات موردنیاز به‌عنوان کتاب کد در مجموعه داده‌ها ذخیره شده است. مجموعه دادهٔ جمع‌آوری‌شده از ۶۲ شرکت‌کننده شامل ٤٤١ صفحه و ٤٢٠٣ جمله از هم جدا شده است. همان‌طور که در شکل ٤ نشان داده شده است، ارتفاع نمونه‌های جمله‌ای ۲۳٦ پیکسل و عرض آنها متغیر است. اندازهٔ صفحات نمونه‌ها برابر با ١٦٥٦ × ۲۳۳۹ پیکسل است. از این پس، این مجموعه دادهٔ خاص «DANA_HW» نام‌گذاری می‌شود و در دسترس همه محققان در پلتفرم GitHub قرار می‌گیرد.

۳-۲- پیش‌پردازش داده‌ها

در این مطالعه، به‌منظور کاهش زمان اجرا و حجم محاسبات، پس از جداسازی ٤٢٠٣ جمله (مربوط به ۶۲ شرکت‌کننده با اندازهٔ ۲۳٦ پیکسل و عرض متغیر)، اندازهٔ جملات ابتدا به ۱۱۲ پیکسل و عرض متغیر تبدیل می‌شود. سپس با استفاده از روش تقسیم‌بندی[۳۳]، ٤٢٠٣ جمله به ۸٦۳۰٤ نمونه با اندازهٔ ۱۱۲ × ۱۱۲ پیکسل تقسیم می‌شوند؛ پس از آن، نمونه‌ها نرمال می‌شوند. عملیات تقسیم‌بندی یکی از جملات در شکل ۵ نشان داده شده است. مطابق شکل ۵، هر جمله به نمونه‌های ۱۱۲ × ۱۱۲ به‌صورت اتوماتیک تقسیم شده است. در برخی از نمونه‌های تقسیم‌شده، تصویر حاوی شکاف‌هایی بین کلمات است که با نماد (a) نمایش داده شده‌اند. در برخی از نمونه‌های دیگر تصویر یا حاوی هیچ دادهٔ ارزشمندی برای پردازش نیست که با نماد (b) یا حاوی داده‌های غیرقابل توجه یا کمی برای پردازش‌اند که با نماد (c) مشخص شده‌اند؛ حتی گاهی شرکت‌کنندگان در نمونه‌گیری با هدف تصحیح، بخشی از دست‌نوشته را مخدوش کرده‌اند که در تقسیم‌بندی با نماد (d) نمایش داده شده‌اند. چنین بخش‌هایی به کاهش دقت در شبکهٔ پیشنهادی منجر می‌شود؛ با وجود این، آنها از مجموعه داده حذف نمی‌شوند تا نمونه‌ها اصلی و دست‌نخورده باقی بمانند. چنین بخش‌هایی «بخش‌های فریبنده» نام‌گذاری شده‌اند. برخی از این بخش‌های فریبنده در شکل ٦ نشان داده شده‌اند.

۳-۳- شبکهٔ عمیق پیشنهادی

شبکهٔ عمیق پیشنهادی در این مطالعه از ترکیب یک شبکه از پیش آموزش‌دیده کانولوشنال Resnet-152 [۳۸] با شبکهٔ LSTM ایجاد شده است. با ترکیب شبکهٔ Resnet-152 با شبکهٔ LSTM می‌توان از مزایای هر دو شبکه به‌طور هم‌زمان استفاده کرد. در بسیاری از مطالعات، از ترکیب شبکه‌های LSTM با شبکه‌های کانولوشنال عمیق به‌منظور کاهش ابعاد ویژگی، افزایش پایداری، کاهش نوسانات، بهبود فرایند آموزش و افزایش صحت شناسایی استفاده شده است [۲۹-۳۰]. شبکهٔ پیشنهادی بر پایهٔ شبکه از پیش آموزش‌دیدهٔ Resnet-152 با یک بلوک پیشنهادی ترکیب می‌شود که شامل دو لایه LSTM، سه لایه نرمال‌سازی دسته‌ای، سه لایه حذف تصادفی و دو لایه FC است (شکل ۲ را ببینید). شبکه‌های از پیش آموزش‌دیده از چند لایه تشکیل شده‌اند که هر لایه ویژگی‌های خاصی را یاد می‌گیرد. لایه‌های اولیه ویژگی‌های پایه‌ای و سطح پایین و لایه‌های بعدی ویژگی‌های پیچیده و سطح بالا را یاد می‌گیرند. در این فرایند ماتریس وزن با روند آموزش تشکیل و تنظیم می‌شود. معماری بلوک پیشنهادی به‌صورت زیر انتخاب شده است: (۱) یک لایه FC با تابع خطی[۳٤] به

همراه یک لایه نرمال‌ساز دسته‌ای با تابع Relu که پس از آن، یک لایه حذف تصادفی قرار می‌گیرد. (۲) یک لایه LSTM با تابع Relu که پس از آن، لایه‌های نرمال‌ساز دسته‌ای و حذف تصادفی قرار می‌گیرند. (۳) معماری مرحلهٔ قبل، یک‌بار دیگر تکرار می‌شود. (٤) یک لایه FC با تابع غیرخطی سافت‌مکس برای دسترسی به لایه خروجی استفاده می‌شود. در شبکهٔ پیشنهادی، خروجی شبکه از پیش آموزش‌دیده یک بردار ویژگی با اندازهٔ ۵۱۲ × ۲۵٦ است. در اولین لایه بلوک پیشنهادی، یعنی FC، تابع خطی روی وزن‌های قابل یادگیری ویژگی‌های به‌دست‌آمده (w) اعمال می‌شود. مقادیر پیش‌بینی‌شدهٔ بایاس $u\left(-\sqrt{1/w},\sqrt{1/w}\right)$ در نظر گرفته می‌شود تا ابعاد بردار ویژگی را به ۱ × ۲۵٦ تغییر دهد. همان‌طور که ملاحظه می‌شود، کاهش ابعاد در لایه‌های پنهان از ۱۱۲ × ۱۱۲ (اندازهٔ ورودی) به ۱۲۸ (بردار ویژگی انتخاب‌شده) ادامه یافته است که درنهایت، بردار ویژگی انتخاب‌شده به یک لایه FC با تابع غیرخطی سافت‌مکس متصل می‌شود (شکل ۲ را ببینید).

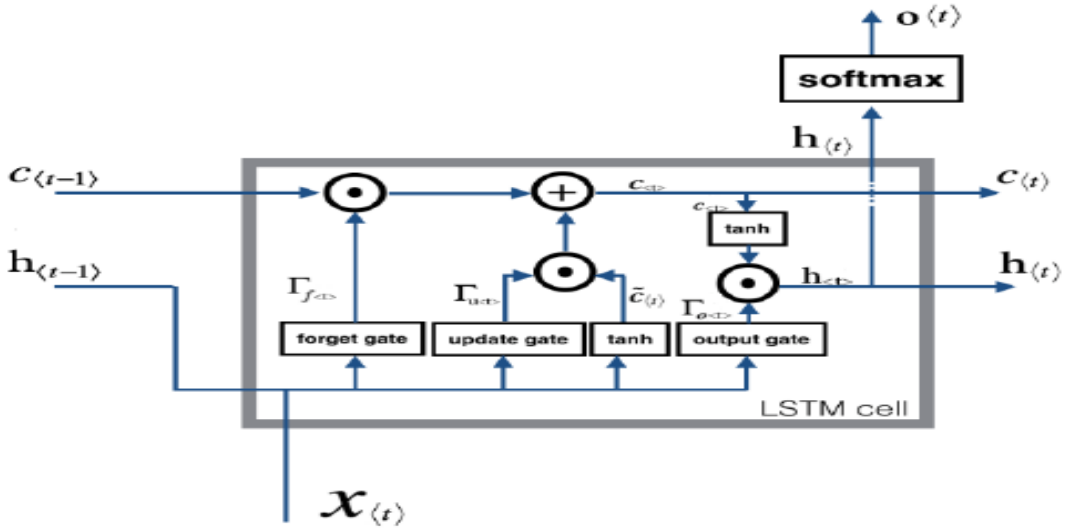
در این مطالعه همه ابر پارامترهای شبکهٔ پیشنهادی به دقت تنظیم شده‌اند تا بهترین نرخ همگرایی را به دست آورند و درنهایت، تابع خطای کراس آنتروپی[۳۵] و بهینه‌ساز SGD[۳٦] با نرخ یادگیری ۰/۰۱ انتخاب شده‌اند. روش مرسوم پس‌انتشار خطا با اندازه دستهٔ ۱۰۰ برای آموزش شبکه استفاده شده است. ابر پارامترهای بهینهٔ انتخاب‌شده برای مدل پیشنهادی در جدول ۳ نشان داده شده است.

براساس مطالعات صورت‌گرفته در ادبیات پیشین، از تکنیک‌های افزایش داده برای بهبود صحت، جلوگیری از پدیدهٔ بیش‌برازش و بهبود روند آموزش در شبکه‌های عصبی استفاده می‌شود [۳۹-٤۰]؛ با وجود این، تکنیک‌های افزایش داده باید با دقت انتخاب شوند و از هر تکنیک افزایش داده نمی‌توان برای داده‌های دست‌نویس استفاده کرد. در این مطالعه تکنیک‌های مقیاس خاکستری تصادفی[۳۷]، تغییر رنگ[۳۸] و چرخش تصادفی[۳۹] به‌عنوان تکنیک‌های افزایش داده انتخاب شده‌اند. نمونه‌هایی از داده‌های افزوده‌شده در شکل ۷ نشان داده شده‌اند. پس از استفاده از تکنیک‌های افزایش داده، مجموعه داده‌های آموزش ۵۰ درصد افزایش می‌یابند.

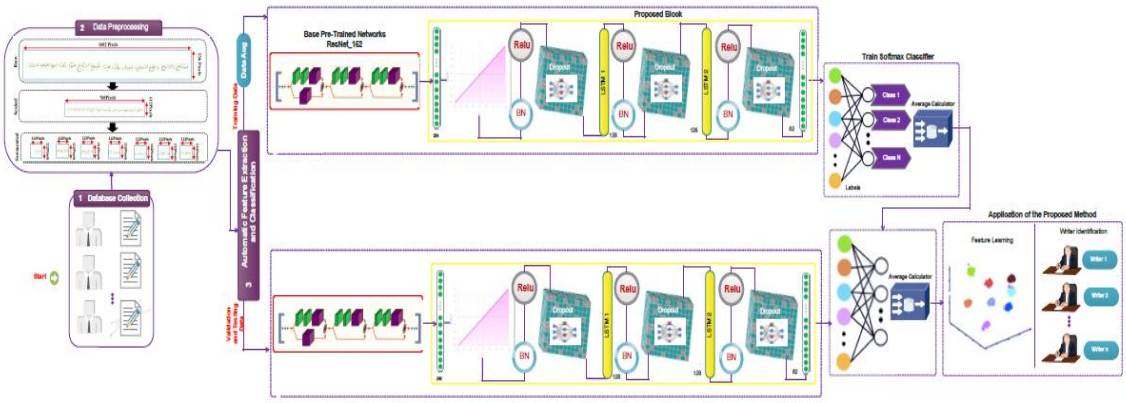
به‌عنوان آخرین مرحله برای افزایش و به حداکثر رساندن دقت پیش‌بینی، از تکنیک TTA[٤۰] استفاده می‌شود. همانند داده‌افزایی روی مجموعه دادهٔ آموزشی برای بهبود عملکرد مدل، هدف از TTA انجام تغییرات تصادفی روی مجموعه دادهٔ (تصاویر) آزمون است؛ بنابراین، به جای نشان‌دادن فقط یک‌بار تصاویر معمولی به مدل آموزش‌دیده، چندین‌بار تصاویر تقویت‌شده به آن نشان داده می‌شوند. پس از آن، میانگین پیش‌بینی‌های هر تصویر، محاسبه و به‌عنوان پیش‌بینی نهایی در نظر گرفته می‌شود. این روش، تکنیکی برای بهبود پیش‌بینی مدل است که به‌طور متداول برای محاسبهٔ پیش‌بینی‌های میانگین در بسیاری از مطالعات طبقه‌بندی استفاده شده است [٤۱-٤۲]. ساختار TTA در شکل ۸ نشان داده شده است. در این مطالعه از چرخش تصادفی به‌عنوان تکنیک افزایش داده در داده‌های آزمون استفاده شده است.

همان‌طور که گفته شد، در این مطالعه آموزش و ارزیابی مدل پیشنهادی با استفاده از داده‌های نامتجانس انجام می‌شود. شکل ۹ نمونه‌های دستخط نامتجانس را برای فرایند آموزش و ارزیابی نشان می‌دهد. با توجه به این واقعیت که تمام تصاویر ۱ تا ۲٤ متعلق به یک فرد است، تصاویر ۱ تا ۱۵ در مجموعه داده‌های آموزشی و تصاویر ۱٦ تا ۲٤ به مجموعه داده‌های اعتبارسنجی و ارزیابی تعلق دارند؛ برای مثال، تصاویر ۱ و ۳ بیشترین شباهت را با تصویر ۱۸ (a) دارند. همچنین، تصاویر ٦ و ۱٦ یکسان‌اند (b)؛ در مقابل، تصاویر ۱۹ و ۲۲ هیچ‌گونه همتایی در مجموعه داده‌های آموزشی ندارند (c).

از کل نمونه‌های جمع‌آوری‌شده (۸٦۳۰٤ نمونه)، ۵۱۷۸۲ نمونه برای داده‌های آموزش (٦۰ درصد)، ۲۵۸۹۰ نمونه (۳۰ درصد) برای داده‌های اعتبارسنجی و ۸٦۳۲ نمونه برای داده‌های آزمون (۱۰ درصد) استفاده می‌شود. علاوه بر این، تمام نمونه‌های اختصاص داده شده به مجموعه‌های آموزش و ارزیابی به‌طور تصادفی انتخاب می‌شوند.



شکل (۱): ساختار یک شبکهٔ عصبی بازگشتی LSTM [۲۸].



شکل (۲): بلوک دیاگرام الگوریتم پیشنهادی.

جدول (۱): جزئیات دو نوع کاغذ استاندارد استفاده‌شده در این مطالعه

خواص	واحد	PaperOne			DoubleA		
		هدف	تلرانس	روش	هدف	تلرانس	روش
وزن پایه	g/m^2	۸۰	±٤%	ISO ٥٣٦	۸۰	±٣%	ISO ٥٣٦
ضخامت	µm	۱۱۰	±٣	ISO ٥٣٤	۱۰٦/٥	±٣/٥	ISO ٥٣٤
زبری	ml/min	۱٤۰	±٤۰	ISO ۸۷۹۱-۲	۱٥۰	±٥۰	ISO۸۷۹۱
روشنایی ISO	%	۹۹	±۲	ISO ۲٤۷۰	۱۰۲/٥	±۱/٥	ISO ۲٤۷۰
تیرگی ISO	%	۹٥	±۲	ISO ۲٤۷۱	۹٤	-	ISO ۲٤۷۱
سفیدی CIE	#	۱٦۷	±۲	ISO ۱۱٤۷٥	۱٦۰	±۲	ISO ۱۱٤۷٥

جدول (۲): مشخصات خودکارهای استاندارد استفاده‌شده در این مطالعه؛

(الف) مشخصات خودکار Schneider، (ب) مشخصات خودکار Faber-Castell.

نوع	GTIN	مدل	رنگ	عرض خط	شماره مقاله
Ballpoint	400467 5004 567	Tops 505	Blue	Fine 0.4 mm	150503
Ballpoint	400467 5004 529	Tops 505	Black	Fine 0.4 mm	150501

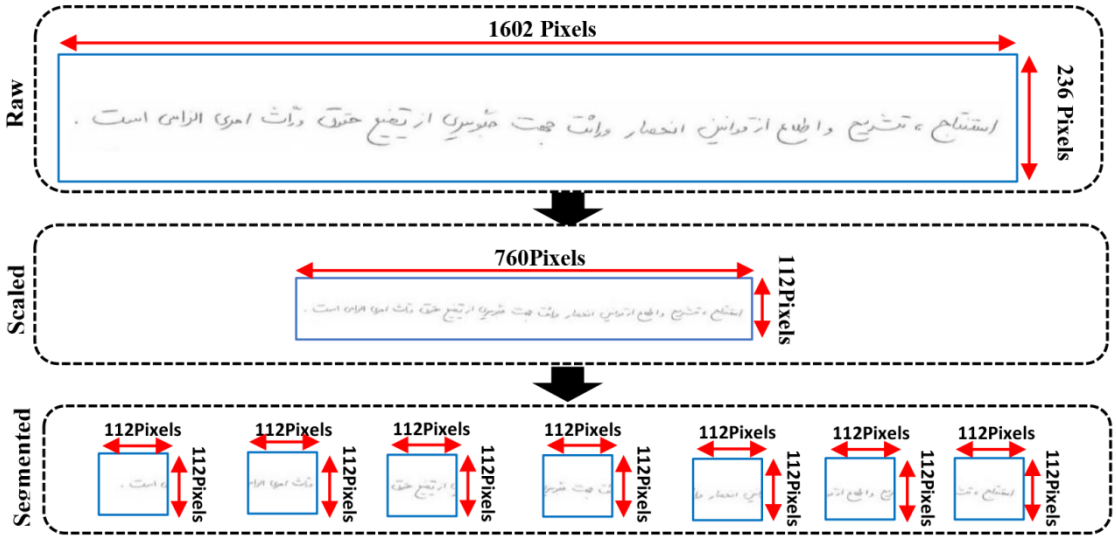
(ب) مشخصات خودکار **Faber-Castell**.

نوع	GTIN	مدل	رنگ	عرض خط	شماره مقاله
Ballpoint	8901180407516	TRI-Flow	[redacted]ue	Fine 0.7 mm	34 07 98
Ballpoint	8901180407998	TRI-Flow	[redacted]lack	Fine 0.7 mm	34 07 50

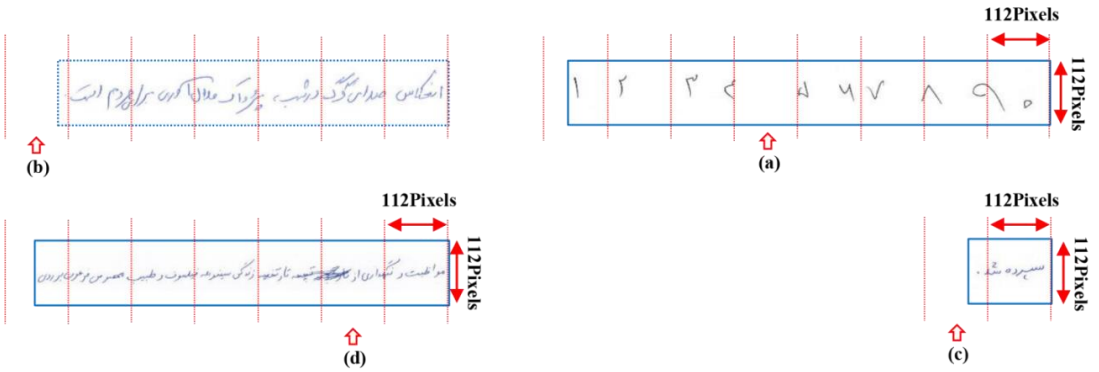
شکل (۳): نمونه‌هایی از فرم نوشتاری شرکت‌کنندگان

(الف) صفحه (ب) جمله

شکل (٤): اندازه نمونه‌ها؛ (الف) صفحه، (ب) جمله.



شکل (۵): عملیات تقسیم‌بندی یکی از جملات.

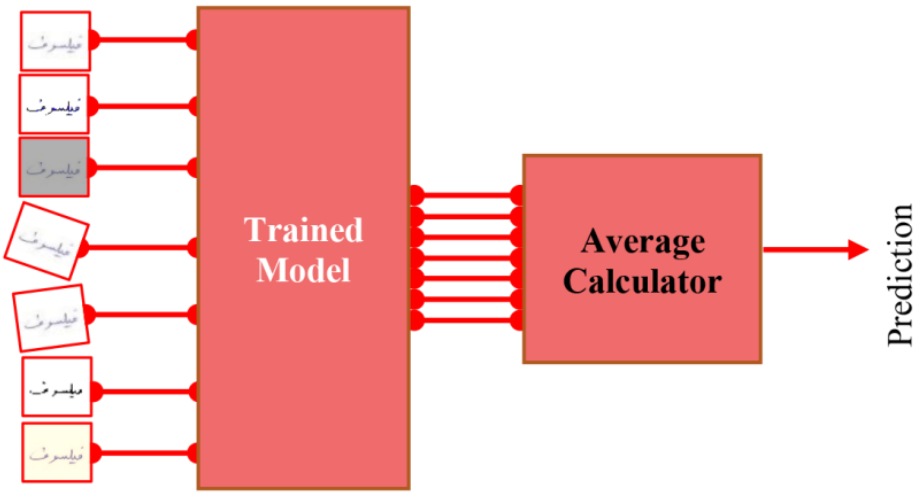


شکل (٦): بخش‌های فریبنده.

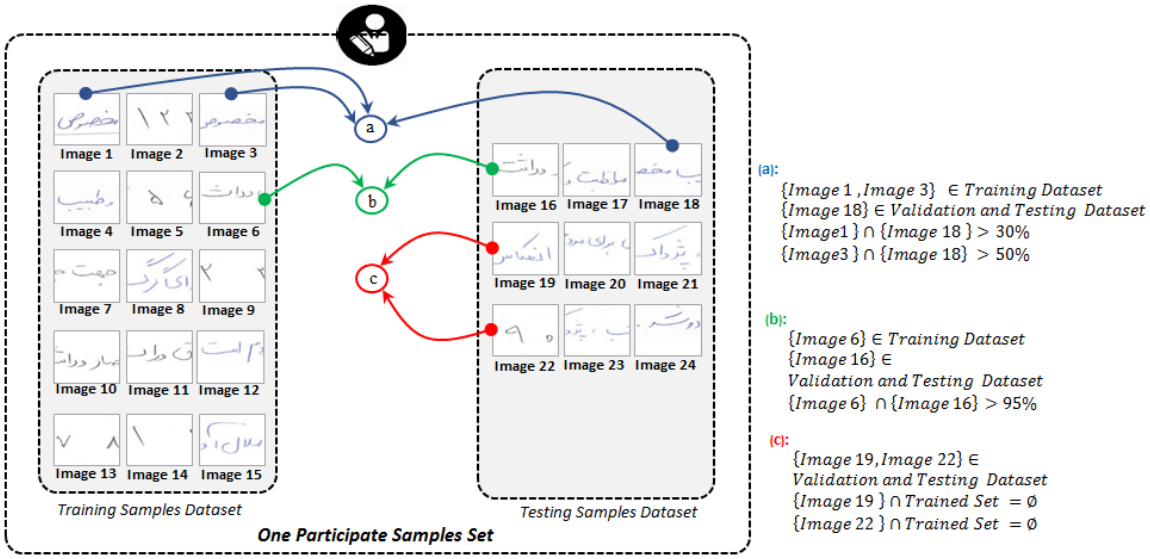
جدول (۳): ابر پارامترهای بهینهٔ استفاده‌شده در مدل پیشنهادی.

پارامتر	فضای جستجو	مقدار بهینه
بهینه‌ساز	RMSProp, Adam, Adamax, SGD, Adadelta	SGD
تابع خطا	MSE, Cross-entropy	Cross-entropy
نرخ لایه حذف تصادفی	۰، ۰/۲، ۰/۳، ۰/٤، ۰/۵	۰/۲
اندازه دسته	٤، ۸، ۱۰، ۱٦، ۳۲، ٦٤، ۱۰۰	۱۰۰
نرخ یادگیری	۰/۰۱، ۰/۰۰۱، ۰/۰۰۰۱	۰/۰۱
تابع فعال‌ساز در لایه FC اول بلوک پیشنهادی	Leaky-Relu, Sigmoid, Relu, Linear	Linear
تابع فعال‌ساز در لایه‌های مخفی بلوک پیشنهادی	Leaky-Relu, Sigmoid, Relu, Linear	Relu
تابع فعال‌ساز در لایه آخر بلوک پیشنهادی	Softmax, Sigmoid	Softmax

شکل (۷): نمونه‌هایی از داده‌های افزوده‌شده با استفاده از تکنیک داده‌افزایی.



شکل (۸): ساختار TTA.



شکل (۹): نمونه‌های دستخط نامتجانس.

٤- نتایج و بحث

روش پیشنهادی[٤١] (P-M) شناسایی نویسنده و کلیه نتایج و بررسی‌ها در پایتون با استفاده از کتابخانه‌های متنوعی انجام شده‌اند که مهم‌ترین آنها PyTorch و NumPy هستند. این بررسی‌ها روی یک سیستم رایانه‌ای با مشخصات زیر انجام شده‌اند: پردازندهٔ مرکزی Intel Core i7-6700K، پردازندهٔ گرافیکی GeForce GTX TITAN X 12 GB، رم ٦٤ گیگابایت DDR IV و هارددیسک ۱ ترابایت SSD. به‌منظور ارزیابی عملکرد روش پیشنهادی، از رابطهٔ مربوط به صحت استفاده می‌شود که به‌صورت زیر بیان می‌شود [٤٣]:

$$\text{Accuracy} = \frac{T_P + T_N}{T_P + F_P + T_N + F_N} \qquad (۱۱)$$

که در آن، TP موارد مثبتی است که به درستی مثبت تشخیص داده شده است. FP موارد منفی است که به اشتباه مثبت تشخیص داده شده است. TN موارد منفی است که به درستی منفی تشخیص داده شده است. FN موارد مثبتی است که به اشتباه منفی تشخیص داده شده است.

نتایج تجربی مدل پیشنهادی (شبکه از پیش آموزش‌دیدهٔ Resnet-152 همراه با بلوک پیشنهادی) و شبکه از پیش

آموزش‌دیدهٔ Resnet-152 بدون بلوک پیشنهادی در جدول ٤ آمده‌اند. مطابق این جدول، هر دو مدل هنگام استفاده از تکنیک TTA، عملکرد بهتری نسبت به استفاده از تکنیک TTA دارند. صحت ارزیابی مدل پیشنهادی (P-M) با تکنیک TTA، ۹۹/٦٦ درصد است؛ در حالی که صحت ارزیابی مدل پیشنهادی بدون استفاده از تکنیک TTA، ۹۵/۷۸ درصد است. همچنین صحت ارزیابی Resnet-152 با تکنیک TTA، ۹٦/۵۱ درصد است؛ در حالی که صحت ارزیابی Resnet-152 بدون استفاده از تکنیک TTA، ۹۳/٤۵ درصد است. همان‌طور که در جدول ٤ ملاحظه می‌شود، صحت و زمان اجرای مدل پیشنهادی از مدل Resnet-152 بیشتر است.

همچنین نمودار صحت و خطا برای داده‌های اعتبارسنجی مدل پیشنهادی (P-M) و مدل Resnet-152 با تکنیک TTA در شکل ۱۰ ارائه شده است. مطابق این شکل، خطای شبکه مدل پیشنهادی و مدل Resnet-152 با افزایش تکرار الگوریتم کاهش می‌یابد. همچنین مشاهده می‌شود مدل پیشنهادی و مدل Resnet-152 پس از ۲۰۰ تکرار تقریباً به صحت ۹۹ درصد و ۹٦ درصد می‌رسند.

برای ارزیابی مدل پیشنهادی، از چهار مجموعه دادهٔ بیان‌شده در بخش ۳-۱ استفاده شده است. نتایج شناسایی مدل پیشنهادی و مدل Resnet-152 با تکنیک TTA برای شناسایی نویسندگان با استفاده از هریک از چهار مجموعه داده در جدول ۵ آورده شده‌اند. مطابق جدول ۵، مدل پیشنهادی مبتنی بر شبکه بهبودیافتهٔ Resnet-152، از مدل Resnet-152 برای شناسایی نویسندگان با استفاده از هریک از چهار مجموعه داده بهتر عمل می‌کند.

صحت ارزیابی روش‌های مختلف برای شناسایی نویسنده در جدول ٦ ارائه و در شکل ۱۱ مقایسه شده‌اند که نتایج مدل پیشنهادی (P-M) در جدول ٦ به‌صورت برجسته مشخص شده‌اند. مقایسه‌ها بر پایهٔ چهار مجموعه دادهٔ بیان‌شده، یعنی IAM، CVL، KHATT و IFN/ENIT انجام می‌شوند. تعداد نویسندگان در هر بررسی نیز در جدول ٦ آمده است. گفتنی است تفاوت‌های ارائه‌شده در جدول ٦، دربارهٔ تعداد نمونه‌های بررسی‌شده در مجموعه داده‌های بیان‌شده، به‌دلیل در دسترس بودن آنها است.

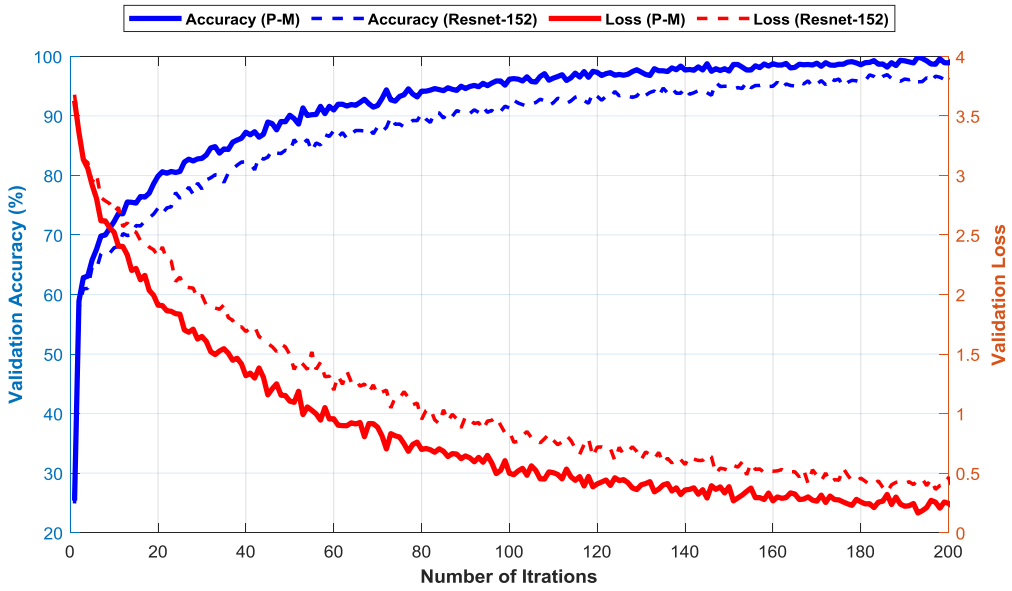
همان‌طور که در جدول ٦ و شکل ۱۱ نشان داده شده است، دربارهٔ تمام مجموعه داده‌های ارزیابی‌شده، مقادیر صحت طبقه‌بندی نشان‌دهندهٔ عملکرد بهتر مدل پیشنهادی در مقایسه با سایر روش‌هاست.

به‌منظور نشان‌دادن عملکرد مدل شبکهٔ عصبی کانولوشن عمیق (DCNN) با مجموعه داده‌های DANA_HW به‌عنوان ورودی، صحت ارزیابی با استفاده از مدل‌های دیگر نیز به دست آمده است؛ براساس این، داده‌های خام DANA_HW و چندین ویژگی مهندسی از مجموعه داده‌های DANA_HW همراه با شبکه پس‌انتشار خطا[٤٢] (BPNN) و ماشین بردار پشتیبان (SVM) به‌عنوان مدل‌های مقایسه‌ای انتخاب شده‌اند [۵۳-۵۵]. تابع پایه شعاعی گوسین[٤٣] به‌عنوان تابع کرنل ماشین بردار پشتیبان بوده و از روش جستجوی شبکه[٤٤] برای بهینه‌سازی پارامترهای کرنل استفاده شده است. معماری شبکهٔ BPNN از یک لایه مخفی تشکیل شده که در آن از تابع فعال‌ساز سیگموئید استفاده شده است. به‌منظور دستیابی به نتایج بهتری از مدل‌های BPNN و DCNN، ابر پارامترهای آنها با توجه به داده‌های مختلف تنظیم می‌شوند. پنج ویژگی استاندارد که نمی‌توانند از تغییرات زمانی تأثیر بگیرند به‌عنوان ویژگی‌های مهندسی انتخاب شده‌اند: مساحت[٤۵]، مختصات مرکزی[٤٦]، گریزازمرکز[٤۷]، کشیدگی[٤۸] و چولگی[٤۹] [۵٦]. صحت ارزیابی روش‌های مختلف بر پایهٔ یادگیری ویژگی از داده‌های خام و ویژگی‌های مهندسی در جدول ۷ ارائه شده است که نتایج مدل DCNN پیشنهادی با داده‌های خام به‌عنوان ورودی، یعنی روش پیشنهادی، در جدول ۷ برجسته شده‌اند. در روش پیشنهادی از معماری ارائه‌شده در بخش ۳-۳ استفاده شده است. مقایسهٔ عملکرد یادگیری ویژگی‌ها و ویژگی‌های مهندسی ارائه‌شده در جدول ۷ نشان می‌دهد یادگیری ویژگی از داده‌های خام با مدل DCNN پیشنهادی، نتایج بهتری نسبت به ویژگی‌های مهندسی ارائه می‌دهد (با افزایش صحت در حدود ۱۳ درصد). این نتیجه کاملاً به معماری منحصربه‌فرد DCNN پیشنهادی مربوط می‌شود که می‌تواند به‌صورت خودکار ویژگی‌های مفید را از داده‌های خام استخراج کند. علاوه بر این، استخراج ویژگی‌های مهندسی به دانش و تخصص قبلی نیاز دارد؛ در حالی که

یادگیری ویژگی از داده‌های خام وابستگی کمتری به دانش تخصصی دارد. مدل DCNN پیشنهادی با یادگیری ویژگی از داده‌های خام نتایج بهتری ارائه می‌دهد؛ در حالی که همه مدل‌های بررسی‌شده، یعنی DCNN، BPNN و SVM نتایج مشابهی را برای ویژگی‌های مهندسی ارائه می‌دهند. این نشان می‌دهد DCNN بدون توانایی یادگیری ویژگی نمی‌تواند نتایج بهتری در شناسایی نویسنده نسبت به روش‌های سنتی ارائه دهد.

جدول (٤): نتایج ارزیابی مدل Resnet-152 همراه با بلوک پیشنهادی (مدل پیشنهادی) در مقایسه با Resnet-152 بدون بلوک پیشنهادی.

نام شبکه		ارزیابی بدون TTA		ارزیابی با TTA		زمان ارزیابی با TTA (ms)	
		صحت آزمون		صحت آزمون		زمان آموزش	زمان آزمون
Resnet-152 همراه با بلوک پیشنهادی (P-M)		۹۵/۷۸		۹۹/٦٦		۱۷/٥٤	۱/۸٥
Resnet-152		۹۳/٤٥		۹٦/٥۱		۱٤/۳۲	۱/۱٥



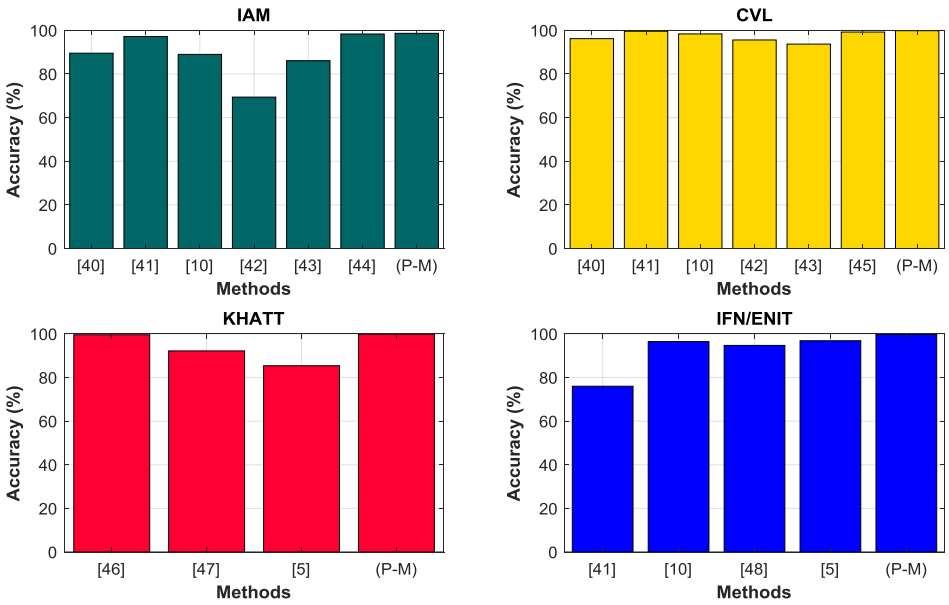
شکل (۱۰): نمودار صحت و خطا برای داده‌های اعتبارسنجی

جدول (٥): نتایج ارزیابی مدل پیشنهادی و مدل Resnet-152 برای چهار مجموعه دادهٔ جامع.

نام شبکه		پایگاه داده				ارزیابی با TTA
		نام داده	زبان	تعداد نویسندگان	تعداد نمونه‌ها	صحت آزمون (٪)
Resnet-152 همراه با بلوک پیشنهادی (P-M)		IAM	انگلیسی	۱٥۰	٤۸۹۹	۹۸/٦٤
		CVL	انگلیسی	۳۰۹	۱۸٥٤	۹۹/۸٥
		KHATT	عربی	۸۲۸	۱۰۸۹۸	۹۹/۸۸
		IFN/ENIT	عربی	٤۱۱	۲٦٤٥۹	۹۹/۷٦
Resnet-152		IAM	انگلیسی	۱٥۰	٤۸۹۹	۹۷/۹۲
		CVL	انگلیسی	۳۰۹	۱۸٥٤	۹۹/۳٤
		KHATT	عربی	۸۲۸	۱۰۸۹۸	۹۹/۳۸
		IFN/ENIT	عربی	٤۱۱	۲٦٤٥۹	۹۹/۱۱

جدول (٦): صحت روش پیشنهادی در مقایسه با سایر روش‌ها.

پایگاه داده			ارزیابی	
نام	زبان	تعداد نویسندگان	روش	صحت (٪)
IAM	English	٦٥٠	هاناد و همکاران [٤٤]	٨٩/٥٤
		٦٥٧	خان و همکاران [٤٥]	٩٧/٢
		٦٥٧	چاهی و همکاران [١٠]	٨٨/٩٩
		٦٥٧	والبرگ و همکاران [٤٦]	٦٩/٤
		٦٥٧	هی و همکاران [٤٧]	٨٦/١
		٢٢٠	کوتزنر و همکاران [٤٨]	٩٨/٣٤
		١٥٠	مدل پیشنهادی (P-M)	٩٨/٦٤
CVL	English	٣١٠	هاناد و همکاران [٤٤]	٩٦/٢
		٣١٠	خان و همکاران [٤٥]	٩٩/٦
		٣٠٩	چاهی و همکاران [١٠]	٩٨/٣٨
		٣١٠	والبرگ و همکاران [٤٦]	٩٥/٦
		٣١٠	هی و همکاران [٤٧]	٩٣/٧
		٣١١	بننور و همکاران [٤٩]	٩٩/٢٣
		٣٠٩	مدل پیشنهادی (P-M)	٩٩/٨٥
KHATT	Arabic	١٠٠٠	کریستلین و همکاران [٥٠]	٩٩/٦
		١٠٠٠	رحمان و همکاران [٥١]	٩٢/٢
		١٠٠٠	هاناد و همکاران [٥]	٨٥/٤
		٨٢٨	مدل پیشنهادی (P-M)	٩٩/٨٨
IFN/ENIT	Arabic	٤١١	خان و همکاران [٤٥]	٧٦
		٤١١	چاهی و همکاران [١٠]	٩٦/٤٧
		٤١١	صبا و همکاران [٥٢]	٩٤/٧٢
		٤١١	هاناد و همکاران [٥]	٩٦/٨٦
		٤١١	مدل پیشنهادی (P-M)	٩٩/٧٦



شکل (١١): مقایسهٔ صحت آزمون روش‌های مختلف.

جدول (۷): صحت آزمون روش پیشنهادی در مقایسه با سایر مدل‌ها

مدل‌ها		یادگیری ویژگی از دادهٔ خام		ویژگی‌های مهندسی
SVM		۸۷/۲۶		۸۶/۴۲
BPNN		۸۶/۱۸		۸۴/۲۳
DCNN		۹۹/۶۶		۸۵/۲۷

۵- نتیجه‌گیری

با توجه به پیچیدگی‌های سبک‌های نوشتاری و نیاز سازمان‌های دولتی (مانند دادگستری و دفاتر ثبت اسناد) به شناسایی دستخط نویسندگان، هدف این مطالعه ارائه یک روش جدید برای شناسایی آفلاین نویسنده با استفاده از نمونه‌های دستخط در شرایط آزمایشی مختلف است. دو ویژگی درخور توجه و مهم مطالعهٔ حاضر استفاده از داده‌های نامتجانس و استقلال روش پیشنهادی برای هر زبان خاص است. در این مطالعه یک مجموعه دادهٔ جامع بر پایهٔ استانداردهای ASTM طراحی شده است. یک مدل DCNN مبتنی بر شبکه از پیش آموزش‌دیده برای استخراج ویژگی‌ها به‌صورت سلسله‌مراتبی از دست‌نوشته‌های خام طراحی و توسعه یافته است.

مطالعهٔ حاضر نشان داد روش پیشنهادی می‌تواند ویژگی‌ها را از روی داده‌های خام دستخط بیاموزد و به صحت قابل قبولی برای شناسایی نویسنده دست یابد. مدل پیشنهادی بر پایهٔ شبکه از پیش آموزش‌دیده به همراه مجموعه دادهٔ طراحی‌شده و چهار نوع مجموعه دادهٔ جامع بررسی شد. نتایج نشان دادند مدل پیشنهادی (شبکه از پیش آموزش‌دیده همراه با بلوک پیشنهادی) از شبکه از پیش آموزش‌دیده بدون بلوک پیشنهادی در شناسایی نویسنده برای هریک از پنج مجموعه دادهٔ بیان‌شده، بهتر عمل می‌کند. همچنین، صحت روش‌های مختلف برای چهار نوع مجموعه دادهٔ جامع با مدل پیشنهادی مقایسه شد. نتایج نشان‌دهندهٔ صحت بالاتر مدل پیشنهادی در مقایسه با سایر روش‌ها برای همه مجموعه داده‌ها بود. علاوه بر این، مجموعه دادهٔ طراحی‌شده همراه با DCNN بررسی و با ویژگی‌های مهندسی و دو روش هوشمند BPNN و SVM مقایسه شد. نتایج نشان دادند روش پیشنهادی قادر به یادگیری ویژگی‌ها و به دست آوردن نتایج شناسایی قانع‌کننده‌ای است. در مقایسه با ویژگی‌های مهندسی، روش پیشنهادی صحت شناسایی را حدوداً ۱۳ درصد افزایش می‌دهد و همچنین وابستگی کمتری به دانش تخصص دارد. با توجه به نتایج ارائه‌شده، می‌توان بیان کرد روش پیشنهادی برای شناسایی خودکار نویسنده بسیار رضایت‌بخش و مناسب است و می‌تواند با ورود به حوزهٔ کاربردی، دستیار خوبی برای متخصصان شناسایی دستخط باشد.

مراجع

[1] S. N. Srihari, S.-H. Cha, H. Arora, and S. Lee, "Individuality of handwriting," *Journal of forensic science*, vol. 47, no. 4, pp. 1-17, 2002.

[2] N. Pokhriyal, K. Tayal, I. Nwogu, and V. Govindaraju, "Cognitive-biometric recognition from language usage: A feasibility study," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 1, pp. 134-143, 2016.

[3] H. E. Said, T. N. Tan, and K. D. Baker, "Personal identification based on handwriting," *Pattern Recognition*, vol. 33, no. 1, pp. 149-160, 2000.

[4] A. A. Ahmed, H. R. Hasan, F. A. Hameed, and O. I. Al-Sanjary, "Writer identification on multi-script handwritten using optimum features," *Kurdistan Journal of Applied Research*, vol. 2, no. 3, pp. 178-185, 2017.

[5] Y. Hannad, I. Siddiqi, C. Djeddi, and M. E.-Y. El-Kettani, "Improving Arabic writer identification using score-level fusion of textural descriptors," *IET Biometrics*, vol. 8, no. 3, pp. 221-229, 2019.

[6] Z. Mousavi, S. Varahram, M. M. Ettefagh, M. H. Sadeghi, and S. N. Razavi, "Deep neural networks–based damage detection using vibration signals of finite element model and real intact state: An evaluation via a lab-scale offshore jacket structure," *Structural Health Monitoring*, p. 1475921720932614, 2020.

[7] Z. Mousavi, M. M. Ettefagh, M. H. Sadeghi, and S. N. Razavi, "Developing deep neural network for damage detection of beam-like structures using dynamic response based on FE model and real healthy state," *Applied Acoustics*, vol. 168, p. 107402, 2020.

[8] C. Adak, B. B. Chaudhuri, and M. Blumenstein, "An empirical study on writer identification and verification from intra-variable individual handwriting," *IEEE Access*, vol. 7, pp. 24738-24758, 2019.

[9] L. G. Hafemann, R. Sabourin, and L. S. Oliveira, "Characterizing and evaluating adversarial examples for Offline Handwritten Signature Verification," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 8, pp. 2153-2166, 2019.
[10] A. Chahi, Y. Ruichek, and R. Touahni, "Block wise local binary count for off-line text-independent writer identification," *Expert Systems with Applications*, vol. 93, pp. 1-14, 2018.
[11] X.-Y. Zhang, G.-S. Xie, C.-L. Liu, and Y. Bengio, "End-to-end online writer identification with recurrent neural network," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 2, pp. 285-292, 2016.
[12] S. M. Awaida and S. A. Mahmoud, "Writer identification of arabic text using statistical and structural features," *Cybernetics and Systems*, vol. 44, no. 1, pp. 57-76, 2013.
[13] F. Shahabi and M. Rahmati, "A new method for writer identification of handwritten Farsi documents," in *2009 10th International Conference on Document Analysis and Recognition*, 2009: IEEE, pp. 426-430.
[14] M. S. Baghshah, S. B. Shouraki, and S. Kasaei, "A novel fuzzy classifier using fuzzy LVQ to recognize online Persian handwriting," in *2006 2nd International Conference on Information & Communication Technologies*, 2006, vol. 1: IEEE, pp. 1878-1883.
[15] X. Wu, Y. Tang, and W. Bu, "Offline text-independent writer identification based on scale invariant feature transform," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 3, pp. 526-536, 2014.
[16] R. Kumar and M. Kaur, "A character based handwritten identification using neural network and SVM," *International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET)*, 2017.
[17] S. Y. Manchala, J. Kinthali, K. Kotha, J. Kumar, and J. Jayalaxmi, "Handwritten text recognition using deep learning with Tensorflow," *International Journal of Engineering and Technical Research*, vol. 9, no. 5, 2020.
[18] X.-Y. Zhang, G.-S. Xie, C.-L. Liu, and Y. Bengio, "End-to-end online writer identification with recurrent neural network," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 2, pp. 285-292, 2016.
[19] V. Carbune *et al.*, "Fast multi-language LSTM-based online handwriting recognition," *International Journal on Document Analysis and Recognition (IJDAR)*, pp. 1-14, 2020.
[20] M. Javidi and M. Jampour, "A deep learning framework for text-independent writer identification," *Engineering Applications of Artificial Intelligence*, vol. 95, p. 103912, 2020.
[21] Y. Xu, Y. Chen, Y. Cao, and Y. Zhao, "A Deep Learning Method for Chinese writer Identification with Feature Fusion," in *Journal of Physics: Conference Series*, 2021, vol. 1883, no. 1, p. 012142: IOP Publishing.
[22] Z. Wang, A. Maier, and V. Christlein, "Towards End-to-End Deep Learning-based Writer Identification," *INFORMATIK 2020*, 2021.
[23] E2290-07a, A., Standard Guide for Examination of Handwritten Items, in *ASTM International*. 2007: West Conshohocken.
[24] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning* (no. 2). MIT press Cambridge, 2016.
[25] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *arXiv preprint arXiv:1207.0580*, 2012.
[26] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International conference on machine learning*, 2015: PMLR, pp. 448-456.
[27] N. Siddique and H. Adeli, *Computational intelligence: synergies of fuzzy logic, neural networks and evolutionary computing*. John Wiley & Sons, 2013.
[28] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735-1780, 1997.
[29] S. Sheykhivand, Z. Mousavi, T. Y. Rezaii, and A. Farzamnia, "Recognizing Emotions Evoked by Music Using CNN-LSTM Networks on EEG Signals," *IEEE Access*, vol. 8, pp. 139332-139345, 2020.
[30] Z.-K. Gao, Y.-L. Li, Y.-X. Yang, and C. Ma, "A recurrence network-based convolutional neural network for fatigue driving detection from EEG," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 11, p. 113126, 2019.
[31] U.-V. Marti and H. Bunke, "The IAM-database: an English sentence database for offline handwriting recognition," *International Journal on Document Analysis and Recognition*, vol. 5, no. 1, pp. 39-46, 2002.
[32] F. Kleber, S. Fiel, M. Diem, and R. Sablatnig, "Cvl-database: An off-line database for writer retrieval, writer identification and word spotting," in *2013 12th international conference on document analysis and recognition*, 2013: IEEE, pp. 560-564.
[33] S. A. Mahmoud, H. Luqman, B. M. Al-Helali, G. BinMakhashen, and M. T. Parvez, "Online-KHATT: An Open-Vocabulary Database for Arabic Online-Text Processing," *The Open Cybernetics & Systemics Journal*, vol. 12, no. 1, 2018.
[34] M. Pechwitz, S. S. Maddouri, V. Märgner, N. Ellouze, and H. Amiri, "IFN/ENIT-database of handwritten Arabic words," in *Proc. of CIFED*, 2002, vol. 2: Citeseer, pp. 127-136.
[35] DoubleApaper. 2017; Available from: http://igepa-allcart.com/myuploads/WDAcjSnxTpKf2iBf1415718830.pdf.
[36] Schneiderpen. 2017; Available from: https://schneiderpen.com/en_us/office/tops-505-black-f-4004675004529.pdf/.
[37] Faber-Castell. 2017; Available from:

http://www.faber-castell.in/40526/Products/Exports/default_news.aspx.
[38] Z. Wu, C. Shen, and A. Van Den Hengel, "Wider or deeper: Revisiting the resnet model for visual recognition," *Pattern Recognition*, vol. 90, pp. 119-133, 2019.
[39] P. Domingos, "Bayesian averaging of classifiers and the overfitting problem," in *ICML*, 2000, vol. 747, pp. 223-230.
[40] D. M. Hawkins, "The problem of overfitting," *Journal of chemical information and computer sciences*, vol. 44, no. 1, pp. 1-12, 2004.
[41] D. Shanmugam, D. Blalock, G. Balakrishnan, and J. Guttag, "When and why test-time augmentation works," *arXiv preprint arXiv:2011.11156*, 2020.
[42] D. Jha *et al.*, "A comprehensive study on colorectal polyp segmentation with ResUNet++, conditional random field and test-time augmentation," *IEEE journal of biomedical and health informatics*, vol. 25, no. 6, pp. 2029-2040, 2021.
[43] A. R. Hassan and M. I. H. Bhuiyan, "Computer-aided sleep staging using complete ensemble empirical mode decomposition with adaptive noise and bootstrap aggregating," *Biomedical Signal Processing and Control*, vol. 24, pp. 1-10, 2016.
[44] Y. Hannad, I. Siddiqi, and M. E. Y. El Kettani, "Writer identification using texture descriptors of handwritten fragments," *Expert Systems with Applications*, vol. 47, pp. 14-22, 2016.
[45] F. A. Khan, M. A. Tahir, F. Khelifi, A. Bouridane, and R. Almotaeryi, "Robust off-line text independent writer identification using bagged discrete cosine transform features," *Expert Systems with Applications*, vol. 71, pp. 404-415, 2017.
[46] F. Wahlberg, "Gaussian process classification as metric learning for forensic writer identification," in *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*, 2018: IEEE, pp. 175-180
[47] S. He and L. Schomaker, "Deep adaptive learning for writer identification based on single handwritten word images," *Pattern Recognition*, vol. 88, pp. 64-74, 2019.
[48] T. Kutzner, C. F. Pazmiño-Zapatier, M. Gebhard, I. Bönninger, W.-D. Plath, and C. M. Travieso, "Writer identification using handwritten cursive texts and single character words," *Electronics*, vol. 8, no. 4, p. 391, 2019.
[49] A. Bennour, C. Djeddi, A. Gattal, I. Siddiqi, and T. Mekhaznia, "Handwriting based writer recognition using implicit shape codebook," *Forensic science international*, vol. 301, pp. 91-100, 2019.
[50] V. Christlein and A. Maier, "Encoding CNN activations for writer recognition," in *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*, 2018: IEEE, pp. 169-174.
[51] A. Rehman, S. Naz, M. I. Razzak, and I. A. Hameed, "Automatic visual features for writer identification: a deep learning approach," *IEEE access*, vol. 7, pp. 17149-17157, 2019.
[52] T. Saba, "Fuzzy ARTMAP Approach for Arabic Writer Identification using Novel Features Fusion," *J. Comput. Sci.*, vol. 14, no. 2, pp. 210-220, 2018.
[53] P. Santos, L. F. Villa, A. Reñones, A. Bustillo, and J. Maudes, "An SVM-based solution for fault detection in wind turbines," *Sensors*, vol. 15, no. 3, pp. 5627-5648, 2015.
[54] M. Hagan, H. Demuth, and M. Beale, "Neural Network Design (PWS, Boston, MA)," *Google Scholar Google Scholar Digital Library Digital Library*, 1996.
[55] S. Sheykhivand *et al.*, "Developing an efficient deep neural network for automatic detection of COVID-19 using chest X-ray images," *Alexandria Engineering Journal*, vol. 60, no. 3, pp. 2885-2903, 2021.
[56] A. Karouni, B. Daya, and S. Bahlak, "Offline signature recognition using neural networks approach," *Procedia Computer Science*, vol. 3, pp. 155-161, 2011.

[1] Handwriting Recognition
[2] Writer Identification
[3] Writer Verification
[4] Multi-Channel Gabor Filter (MGF)
[5] Moor
[6] Fisher's Linear Discriminant Analysis (LDA)
[7] Principal Component Analysis (PCA)
[8] Support Vector Machine (SVM)
[9] Bengali
[10] Deep Convolutional Neural Network (DCNN)
[11] Recurrent Neural Network (RNN)
[12] Random Hybrid Strokes (RHSs)
[13] Deep Neural Network (DNN)
[14] pre-trained network
[15] Convolutional Neural Network (CNN)
[16] Long Short Term Memory (LSTM)
[17] Artificial neural networks (ANNs)
[18] Feedforward

[19] Back Propagation (BP)
[20] Loss Function
[21] Pooling Layer
[22] Fully Connected
[23] Over Fitting
[24] Dropout
[25] Batch Normalization (BN)
[26] Down Sampling
[27] Max-Pooling
[28] Softmax
[29] Separated form (S)
[30] Beginning Of Word form (BOW)
[31] Middle Of Word form (MOW)
[32] End Of Word form (EOW)
[33] Segmentation
[34] Linear
[35] Cross-Entropy
[36] Stochastic Gradian Descend (SGD)
[37] RandomGrayscale
[38] ColorJitter
[39] RandomRotation
[40] Test Time Augmentation (TTA)
[41] Proposed Method (P-M)
[42] Back-Propagation Neural Network (BPNN)
[43] Gaussian Radial Basis Function (RBF)
[44] Grid search method
[45] Area
[46] Centroid Coordinates
[47] Eccentricity
[48] Kurtosis
[49] Skewness