

دانشگاه پیام نور
دانشکده فنی و مهندسی

پایان نامه
برای دریافت درجه کارشناسی ارشد
رشته مدیریت فناوری اطلاعات
گروه مهندسی کامپیوتر و فناوری ارتباطات

ارائه الگوریتمی برای جداسازی، تشخیص و تصدیق اتوماتیک امضا در چک‌های بانکی دستنویس فارسی

نویسنده: یونس اکبری

استاد راهنما:

دکتر جواد صدری

استاد مشاور:

مهندس ملیحه حبیبی

تیر ۱۳۹۰



دانشگاه پیام نور

دانشکده فنی و مهندسی

پایان نامه

برای دریافت درجه کارشناسی ارشد

رشته مدیریت فناوری اطلاعات

گروه مهندسی کامپیوتر و فناوری ارتباطات

ارائه الگوریتمی برای جداسازی، تشخیص و اتوماتیک امضا در چک‌های بانکی دستنویس فارسی

نویسنده: یونس اکبری

استاد راهنما:

دکتر جواد صدری

استاد مشاور:

مهندس ملیحه حبیبی

تیر ۱۳۹۰

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

تقدیم

به پدر و مادر عزیزم به پاس تعبیر عظیم و انسانی‌شان از کلمه ایثار و از خود گذشتگی و به پاس عاطفه سرشار و گرمای امید بخش وجودشان که در این سردترین روزگاران بهترین پشتیبان است.

تقدیر و تشکر

خدا را سپاس می‌گویم که به من توفیق داد تا بتوانم قدمی هر چند ناچیز در راه آموختن علم و کسب معرفت برداشته و ناهمواری‌های این راه دشوار را بر خود هموار سازم.

از استاد ارجمند و گرانقدرم، جناب آقای دکتر جواد صدری، که در راهنمایی من در این پایان‌نامه زحمات فراوان و طاقت‌فرسایی را متحمل شدند و در راه کمک به تدوین هر چه بهتر این مجموعه، از هیچ زحمتی فروگذار نکردند، کمال تشکر و قدردانی را می‌نمایم. راهنمایی‌های ارزنده ایشان همواره سختی‌های راه تحقیق و پژوهش را بر اینجانب آسان نمود.

همچنین از سرکار خانم مهندس ملیحه حبیبی به خاطر دلسوزی بی‌دریغشان در جهت راهنمایی و مشاوره من برای بهتر انجام یافتن این مجموعه صمیمانه سپاسگزارم.

در پایان از تمامی دوستان و عزیزانم از جمله دوستان عزیزم از جمله آقایان محمد جواد جلیلی و مسعود تقی‌زاده و سرکار خانم کندویی و سایر دوستان و همکاران عزیزی که هر یک به نوعی در تهیه مجموعه حاضر همواره امیدبخش من در طی تدوین آن بوده و مرا در این مهم یاری نمودند، تشکر و قدردانی می‌نمایم.

برای کلیه این عزیزان آرزوی توفیق، سعادت و سلامتی را از خداوند بزرگ خواستارم.

یونس اکبری

تیر ماه ۹۰

چکیده

در زمینه‌ی تشخیص (بازشناسی) الگو، تحقیقات بسیاری در کشور عزیزمان انجام شده است. یکی از شاخه‌های تشخیص الگو، تشخیص دست‌خط است، که در این زمینه نیز تلاش‌های بسیاری صورت گرفته است. در سال‌های اخیر به زمینه‌های کاربردی تشخیص دست‌خط توجه بسیاری شده است. در زمینه‌های کاربردی، سازمان‌های دولتی و غیر دولتی می‌توانند به منظور کاهش حجم قابل توجهی از اسناد دست‌نویس و کاهش زمان پردازش دستی اسناد از سیستم‌های پردازش خودکار استفاده کنند. سیستم‌های تشخیص نوری کاراکتر زیادی در زبان‌های لاتین، چینی، ژاپنی و دیگر زبان‌ها به این منظور طراحی شده‌اند. در این پایان‌نامه، سیستمی کاربردی برای پردازش خودکار چک‌های بانکی دست‌نویس فارسی ارائه شده است. بانک‌ها و موسسات مالی اعتباری، از سازمان‌هایی می‌باشند که روزانه اوراق بهادار دست‌نویس فراوانی به صورت دستی توسط نیروی انسانی متخصص در آن‌ها پردازش می‌شوند. تا کنون تحقیقات اندکی برای طراحی سیستمی به منظور پردازش خودکار چک در ایران انجام شده است ولی هیچ‌کدام موفق به ارائه‌ی یک سیستم کاربردی نشده‌اند. در این پایان‌نامه، عدم موفقیت تحقیقات انجام شده بر روی پردازش چک و مشکلات خودکارسازی سیستم بانکداری کنونی ایران بررسی شده است. توجه اصلی در این پایان‌نامه تشخیص و بررسی صحت امضا در سیستم بانکی است و الگوریتمی بدین منظور ارائه شده است. امضا مهمترین فیلد چک است که در صورت بررسی نادرست، اطمینان مشتریان از پردازش اتوماتیک در سیستم بانکداری سلب می‌شود. الگوریتم ارائه شده برای کاهش خطای نیروی انسانی پیاده‌سازی شده است.

کلمات کلیدی: تشخیص رقم فارسی، بررسی صحت امضا، پردازش خودکار چک، بانک‌اطلاعات

نمونه‌های دست‌خط، تشخیص دست‌خط فارسی، تشخیص نوری کاراکتر.

فهرست مطالب

صفحه	عنوان
۱	فصل اول
۱-۱	تعریف مساله و بیان سوال‌های اصلی تحقیق
۲-۱	سابقه و ضرورت انجام تحقیق
۳-۱	هدف‌ها
۴-۱	کاربرد تحقیق و استفاده کننده گان
۵-۱	جنبه جدید بودن و نوآوری تحقیق
۶-۱	روش انجام تحقیق
۷-۱	روش تجزیه و تحلیل اطلاعات
۸-۱	نمایی کلی از پایان نامه
۹-۱	مفاهیم اصلی در تشخیص دست خط
۱-۹-۱	یادگیری ماشین
۲-۹-۱	تشخیص الگو
۳-۹-۱	روشهای تشخیص الگو
۴-۹-۱	بازشناسی نوری کاراکتر
۵-۹-۱	انواع سیستمهای OCR
۶-۹-۱	پردازش تصویر
۱۰-۱	ساختار این پایان نامه
۱۸	فصل دوم
۱۸	مروری بر کارهای گذشته
۱-۲	کارهای انجام شده برای جمع‌آوری بانک‌های اطلاعاتی
۲-۲	کارهای انجام شده برای تشخیص امضا

۳۷	فصل سوم	۳
۳۷	تعریف مسئله و ارائه راه حل	
۳۸	مشکلات چک های سنتی موجود در سیستم بانکی ایران	۱-۳
۴۳	ویژگی های طرح پیشنهادی چک	۲-۳
۴۶	فصل چهارم	۴
۴۶	ایجاد بانک اطلاعات	
۴۷	جمع آوری بانک اطلاعاتی	۱-۴
۵۴	مرحله پیش پردازش و جداسازی فیله های طرح پیشنهادی جدید چک	۲-۴
۵۸	نتایج آماری پس از جداسازی ارقام و کلمات	۳-۴
۶۱	فصل پنجم	۵
۶۱	بازشناسی (تشخیص) امضا	
۶۴	ایجاد بانک اطلاعاتی امضا با جداسازی امضا از چک های پیشنهادی و فرم ها	۱-۵
۶۶	پردازش اولیه	۲-۵
۶۹	استخراج ویژگی ها	۳-۵
۷۳	بررسی صحت امضا	۴-۵
۷۷	فصل ششم	۶
۷۸	نحوی پیاده سازی سیستم تشخیص امضا	۱-۶
۸۱	نتایج آزمایش ها	۲-۶
۸۲	نمونه ای از نتایج سیستم	۱-۲-۶
۸۵	فصل هفتم	۷
۸۷	پیشنهادات	۱-۷
۸۹	مراجع	۸
۹۲	واژه نامه	۹

فهرست شکل‌ها

صفحه	عنوان
۶	شکل ۱-۱ دیاگرام کلی سیستم
۷	شکل ۲-۱ نمونه‌ای از طرح جدید پیشنهادی چک
۹	شکل ۳-۱ بلوک دیاگرام طراحی یک سیستم شناسایی الگو
۱۳	شکل ۴-۱ cross pad
۲۰	شکل ۱-۲ نمونه‌ای از فرم‌های NTS
۲۱	شکل ۲-۲ نمونه‌ای از بانک اطلاعاتی CEDAR
۲۲	شکل ۳-۲ نمونه‌ای از فرم‌های CENPARMI
۲۳	شکل ۴-۲ نمونه‌ای از فرم‌های IFHCDB
۲۴	شکل ۵-۲ نمونه‌ای از چک‌های عربی CENPARMI
۲۴	شکل ۶-۲ نمونه‌ای از چک‌های بانکی انگلیسی CENPARMI
۲۵	شکل ۷-۲ نمونه‌ای از امضاهای
۲۶	شکل ۸-۲ ویژگی‌های استخراج شده در (جستینو و همکاران، ۲۰۰۵)
۲۸	شکل ۹-۲ تقسیمات عمودی تصویر امضا
۲۸	شکل ۱۰-۲ تقسیمات افقی تصویر امضا
۲۹	شکل ۱۱-۲ پراکندگی هریک از نقاط در مجموعه آموزشی
۳۲	شکل ۱۲-۲ دیاگرام کلی سیستم
۳۳	شکل ۱۳-۲ مراحل از پیش پردازش تا استخراج ویژگی
۳۴	شکل ۱۴-۲ تقسیم قاب تصویر امضا
۳۴	شکل ۱۵-۲ ویژگی ۳ (دارامولا ساموئل و همکاران، ۲۰۱۰)
۳۵	شکل ۱۶-۲ فاصله از بالا تا اولین نقطه امضا (سفید) و از پایین
۳۵	شکل ۱۷-۲ نقاط دارای بیش از دو همسایگی
۳۹	شکل ۱-۳ مرحله Binarization بر روی چک‌های سنتی
۴۰	شکل ۲-۳ تداخل امضا، شماره حساب و گیرنده حساب
۴۰	شکل ۳-۳ تاریخ و مبلغ عددی (a). مبلغ عددی (b). تاریخ عددی
۴۱	شکل ۴-۳ جابه جایی فیلد مبلغ عددی درد و نمونه چک بانک ملی ایران
۴۲	شکل ۵-۳ اسکن زاویه دار
۴۳	شکل ۶-۳ نمونه‌ای از طرح پیشنهادی جدید

- شکل ۳-۷ چهارنقطه راهنما به منظور مکان‌یابی و تشخیص زوایا ۴۴
- شکل ۳-۸ استفاده از فیله‌های محدود شده برای تاریخ، مبلغ عددی و امضا ۴۴
- شکل ۴-۱ معادل اعداد فارسی و لاتین ۴۷
- شکل ۴-۲ ساختار سلسله مراتبی پایگاه داده ۵۰
- شکل ۴-۳ نمونه‌ای از فیله‌ها، از قام و کلمات بانک اطلاعاتی ۵۱
- شکل ۴-۴ فرم جمع‌آوری نمونه دست‌خط و امضا ۵۳
- شکل ۴-۵ مرحله Binarization ۵۴
- شکل ۴-۶ تشخیص زاویه به وجود آمده پس از اسکن ۵۵
- شکل ۴-۷ استفاده از بارکد برای تشخیص نوع چک ۵۶
- شکل ۴-۸ جداسازی فیله‌ها به کمک مرکز مستطیل‌های راهنما ۵۷
- شکل ۴-۹ جداسازی ارقام از فیله‌های تاریخ و مبلغ عددی ۵۸
- شکل ۵-۱ دیاگرام تشخیص امضای برخط (تانگ کیو و همکاران، ۲۰۰۴) ۶۲
- شکل ۵-۲ کلی سیستم تشخیص امضا برون خط ۶۳
- شکل ۵-۳ نمونه‌ای از فرم‌ها و چک‌های بانک اطلاعاتی امضا ۶۵
- شکل ۵-۴ مرحله Binarization ۶۶
- شکل ۵-۵ مرحله Noise removal ۶۷
- شکل ۵-۶ مرحله ROI ۶۸
- شکل ۵-۷ مرحله Morphology ۶۸
- شکل ۵-۸ دیاگرام مرحله استخراج ویژگی ۶۹
- شکل ۵-۹ نمونه‌ای از تقسیم قاب تصویر امضا ۷۰
- شکل ۵-۱۰ ویژگی نقاط در قسمتی از تصویر که بزرگنمایی شده است ۷۱
- شکل ۵-۱۱ ویژگی انحنای مسیر امضا ۷۲
- شکل ۶-۱ معادل کاراکترهای MICR ۷۸
- شکل ۶-۲ مرحله پیش‌پردازش و جداسازی بر روی فیله‌های چک ۷۹
- شکل ۶-۳ مرحله بررسی صحت امضا ۸۰
- شکل ۶-۴ امضاهای آموزشی نمونه امضا ۸۲
- شکل ۶-۵ امضاهای مجموعه آزمایشی که به درستی تشخیص داده شده است ۸۳
- شکل ۶-۶ امضاهای آموزشی نمونه امضا ۸۳
- شکل ۶-۷ نمونه امضایی از مجموعه آزمایشی که به اشتباه رد شده است ۸۴

فهرست جدول‌ها

صفحه	عنوان
۵۹	جدول ۱-۴ نتایج اولیه از جداسازی ارقام.....
۵۹	جدول ۲-۴ نتایج نهایی ارقام در مجموعه آموزشی و آزمایشی
۶۰	جدول ۳-۴ نتایج کلمات جدا شده.....
۸۱	۱-۶ نتایج بر حسب FAR.....

فهرست علائم اختصاری

ARAN	Aspect Ratio Adaptive Normalizat	نرمالسازی نسبت انطباق ظاهری
ATM	Automated Teller Machine	ماشین تحویلدار اتوماتیک
DIA	Document Image Analysis	تحلیل تصاویر اسناد
DPI	Dots Per Inch	نقطه در اینچ
FPGA	Field Programmable Gate Arrays	مدار مجتمع قابل برنامه ریزی
GUI	Graphical User Interface	واسط گرافیکی کاربر
KNN	K- Nearest Neighbor	kامین نزدیکترین همسایه
MATLAB	Matrix Laboratory	آزمایشگاه ماتریس (نام نرم افزار محاسباتی)
MLP	Multi Layer Perceptron	پرسپترون چند لایه
MSE	Median Square Errors	میانگین مربعات خطاها
NN	Neural Network	شبکه عصبی
OCR	Optical Character Recognition	تشخیص نوری کاراکترها
RGB	Red Green Blue	قرمز سبز آبی
RBF	Radial Basis Function	تابع اصلی پرتوی
SVM	Support Vector Machine	ماشین بردار پشتیبان
TIFF	Tagged Image File Format	فرمت فایل تصویری برچسب دار (ضمیمه دار)

فصل اول

مقدمه

شناسایی و پردازش خودکار دست‌نوشته‌ها یکی از چالش‌های دانشمندان علوم کامپیوتر می‌باشد. این چالش‌ها در تحقیقات، وقتی بیشتر به چشم می‌آید که ساختار نوشته بسیار قدیمی باشد یا این ساختار توسط محققین طراحی نشده باشد، مانند دست‌نوشته‌ها و کتب قدیمی که شناسایی و پردازش آن بسیار کار سخت و پیچیده‌ای می‌باشد. اگر ساختار فرمی، که توسط نویسنده‌ای (فردی) پر شده، استاندارد باشد و یا آن فرم توسط محققین که می‌خواهد دست‌خط را شناسایی کند طراحی شده باشد کار به مراتب ساده‌تر و دقیق‌تر خواهد بود. پس اگر سازمان‌های دولتی و غیر دولتی مانند بانک‌ها و موسسات مالی و اعتباری، تصمیم بگیرند که اسناد خود را به صورت خودکار پردازش کنند، بهترین راه‌حل، طراحی فرم‌های استاندارد می‌باشد که پردازش خودکار¹ اسناد را ساده می‌کند.

۱-۱ تعریف مساله و بیان سوال‌های اصلی تحقیق

علی‌رغم پیشرفت‌های شگرف در سرعت و پردازش کامپیوترها، هنوز تعداد زیادی از کاربردهای مهم در ادارات و سازمان‌ها به صورت دستی انجام می‌شود. از جمله این کاربردها پردازش چک‌های بانکی دست‌نویس در کشورهای مختلف از جمله کشور عزیزمان ایران می‌باشد. هرروزه صدها هزار نسخه چک دست‌نویس به بانک‌های کشورمان جهت پردازش ارجاع می‌شود و تمامی آنها در حال حاضر به صورت دستی پردازش شده و بایگانی می‌گردد. پردازش دستی چک‌ها به کمک اپراتورهای انسانی، عملی وقت‌گیر، خسته‌کننده و هزینه‌بر می‌باشد و همچنین بایگانی چک‌ها به صورت کاغذی، نیاز به فضای زیاد داشته و هزینه‌های نگهداری و امنیتی زیادی را به بانک‌ها تحمیل می‌کند. در این پایان‌نامه سیستمی بر اساس پردازش تصویر و یادگیری الگوها برای جداسازی و تشخیص اتوماتیک امضا در چک‌های بانکی فارسی ارائه می‌گردد و همچنین دستورالعملی برای طراحی و پیاده‌سازی چک‌های بانکی با ساختاری جدید که کاملاً منطبق بر سیستم‌های جداسازی اتوماتیک دست خط باشد ارائه می‌گردد. در این تحقیق به سوالات زیر پاسخ می‌دهیم:

- ۱- از آنجایی که سرعت در امور بانکی از اهمیت زیاد برخوردار است؛ آیا برای بهبود سرعت پردازش کامپیوتری و خودکار اوراق بهادار کاری صورت گرفته است؟

¹Automatic Processing

۲- چگونه می‌توان اوراق بهادار را با ساختار بهتری طراحی نمود؟

۳- آیا از لحاظ آماری می‌توان دقت در تشخیص صحیح دست خط‌ها را اندازه‌گیری نمود؟

۴- آیا سیستم نرم‌افزاری ارائه شده بر روی بسترهای سخت افزاری و نرم افزاری مرتبط با بانک‌ها نیز قابل اجرا می‌باشد؟

۵- برای بررسی صحت امضا از چه روشی استفاده شده است و کدام روش متناسب با ساختار چک‌های بانکی و دقت مورد نیاز بانک‌ها می‌باشد؟

۲-۱ سابقه و ضرورت انجام تحقیق

در مورد پردازش اتوماتیک چک‌ها باید به سابقه شاخه‌های کلی‌تر آن یعنی تشخیص دست خط، تشخیص کاراکترهای نوری (OCR) و تشخیص الگوها نگاهی انداخت. از زمانی که سیستم‌های خبره برای جاگزین کردن سیستم‌های کامپیوتری برای برعهده گرفتن فعالیت‌های فکری انسان‌ها وارد این عرصه شده کارهای زیادی برای اتوماتیک کردن پردازش انواع فرم‌ها انجام شده که این خود مستلزم پردازش انواع تصاویر و تشخیص الگوها بود. یکی از شاخه‌های پردازش تصویر پردازش دست خط‌ها است که با توجه به تفاوت‌ها و پیچیدگی‌های الفبای کشورهای گوناگون به یکی از بحث برانگیزترین شاخه‌های پردازش تصویر مبدل شد. طبیعتاً برای انجام این امر توسط نرم افزارها باید راه‌حلی الگوریتمی و مرحله به مرحله ارائه شود. که مراحل کلی و عمومی به شرح ذیل است:

الف- پردازش اولیه و کاهش نویز ب- جداسازی ج- استخراج ویژگی‌ها د- تشخیص نهایی
همچنین با توجه به تفاوت‌های زیاد بین دست خط افراد مختلف باید پایگاه داده‌های جامعی برای بهتر شدن تشخیص جمع‌آوری شود. همه موارد فوق در زمینه پردازش چک‌ها که خود نوعی از فرم‌ها می‌باشد صادق است.

با تحقیقات انجام شده، کشورهای آمریکا، کانادا، برزیل، ژاپن، از پیشگامان اتوماتیک کردن پردازش اوراق بهادار و چک‌ها می‌باشند که پیشرفت‌های زیادی در این زمینه حاصل شده است. اما در زبان فارسی عموماً کارهای انجام گرفته، ناقص و به صورت تحقیقات باقی مانده که یکی از دلایل

آن پیچیدگی‌های بالای زبان فارسی در امر پردازش و همچنین مشکلات ناشی از ساختارهای کنونی چک‌ها می‌باشد.

در حال حاضر میلیون‌ها چک بانکی دست‌نویس به صورت دستی و بدون استفاده از کامپیوتر در کشور عزیزمان ایران پردازش و ذخیره می‌گردد. هر ساله هزینه‌های زیادی را به مشتریان و بانک‌های کشور تحمیل می‌نماید. امید است که نتایج حاصل از این تحقیق که به صورت کاربردی نیز پیاده‌سازی شده راه‌گشای مشکلات پردازش اوراق بهادار در کشور عزیزمان باشد.

۳-۱ هدف‌ها

- ۱- ایجاد و توسعه بانک اطلاعاتی بر مبنای چک‌های بانکی دست‌نوشته فارسی
- ۲- ارائه یک طرح جدید برای چک‌های بانکی در کلیه بانک‌های کشور عزیزمان برای ساده‌سازی و بالا بردن دقت تشخیص و جداسازی چک‌های بانکی فارسی
- ۳- جداسازی و تشخیص امضا در چک‌های دست‌نوشته فارسی

۴-۱ کاربرد تحقیق و استفاده‌کنندگان

این سیستم در کلیه بانک‌های کشور می‌تواند برای پردازش اتوماتیک بخش زیادی از چک‌ها (حدوداً ۸۰ درصد) و ذخیره‌سازی آن‌ها به کمک کامپیوتر مورد استفاده قرار گیرد همچنین می‌توان پردازش اتوماتیک چک را در ساعات غیر اداری در سیستم‌های ATM انجام داد.

- کلیه موسسات بانکی و مالی در کشور
- کلیه موسسات پژوهشی و آموزش عالی در کشور

۵-۱ جنبه جدید بودن و نوآوری تحقیق

طرح پردازش (شامل: جداسازی، تشخیص و ذخیره‌سازی اتوماتیک) چک‌های فارسی در کشور ما کاملاً جدید بوده و تا کنون به دلیل پیچیدگی‌های زیاد فنی آن به اجرا در نیامده است؛ و حتی در سطح بین‌المللی به عنوان یک مسئله تحقیقاتی و کاملاً حل نشده مطرح می‌باشد و سیستم‌های معدودی برای این کار تا به حال معرفی یا پیاده‌سازی شده‌اند.

۶-۱ روش انجام تحقیق

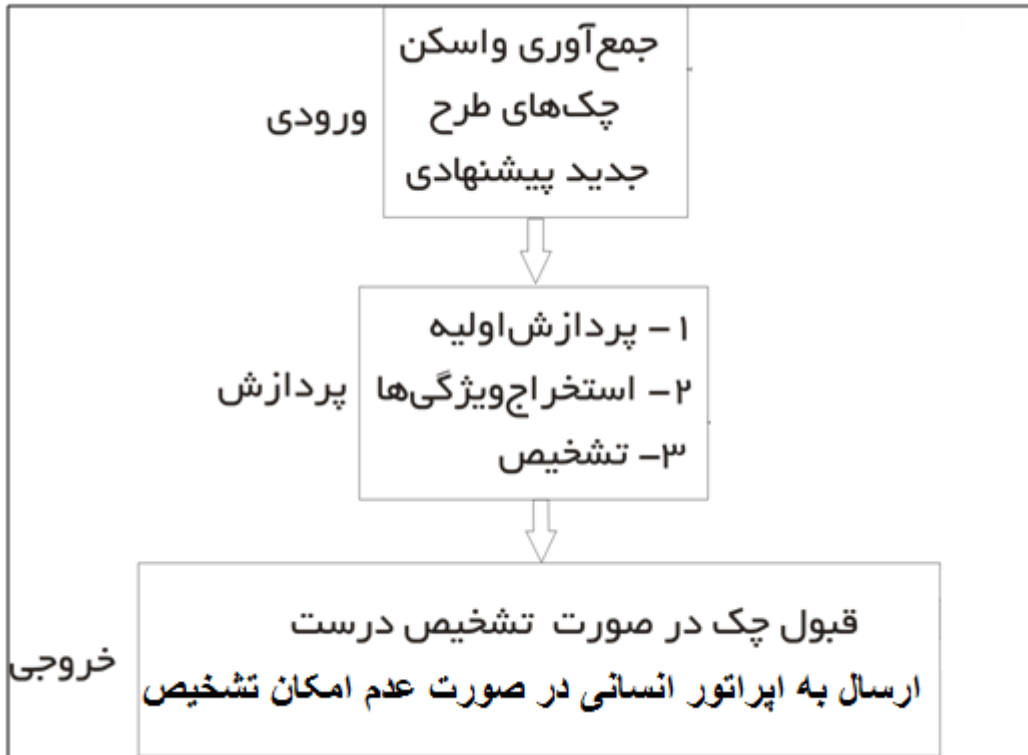
- ۱- جمع‌آوری نمونه‌های واقعی چک‌های فارسی در بانک‌ها و اسکن کردن حدوداً ۲۰۰۰ نسخه از آنها
- ۲- راه‌اندازی نرم‌افزار بر روی بستر لینوکس
- ۳- ایجاد و توسعه یک بانک اطلاعاتی برای ارقام، حروف، کلمات و تاریخ‌های استفاده در چک‌ها بر اساس مورد ۱ و ۲
- ۴- مطالعه کارهای انجام شده در زمینه تشخیص دست‌خط فارسی و عددی
- ۵- طراحی و پیاده‌سازی یک سیستم برای جداسازی و تشخیص امضای چک‌های بانکی
- ۶- آموزش، آزمایش و ارزیابی سیستم طراحی شده بر اساس موارد ۱ و ۲ و ۳

۷-۱ روش تجزیه و تحلیل اطلاعات

پس از اسکن کردن چک‌های جمع‌آوری شده آن‌ها به ۲ گروه آموزش (Training) و تست (Testing) با نسبت‌های ۸۰ درصد و ۲۰ درصد به صورت کاملاً تصادفی تقسیم می‌شوند و از هر گروه کلیه ارقام، تاریخ و مبالغ عددی و همچنین کلیه کلمات مقادیر حرفی چک استخراج شده و بانک‌های اصلی جدید برای ارقام حروف و کلمات تشکیل خواهد گردید. در پایان نیز بانک اطلاعاتی جهت تشخیص امضا ایجاد می‌شود.

۸-۱ نمایی کلی از پایان‌نامه

این پایان‌نامه به بررسی مشکلات چک‌های موجود در سیستم بانک‌داری ایران می‌پردازد و راه‌حلی سیستمی و کامل را برای مشکلات پردازش خودکار چک‌های بانکی فارسی پیشنهاد می‌کند. سیستمی نیز ارائه می‌شود که دیاگرام آن در شکل ۱-۱ نشان داده شده است.



شکل ۱-۱-۱ دیگرام کلی سیستم

در این پایان‌نامه طرح جدیدی برای چک‌های بانکی دست‌نویس فارسی ارائه می‌شود. به دلیل ساختار استاندارد طرح جدید، این طرح را می‌توان برای تمامی سیستم‌های بانکی دنیا به منظور استفاده برای پردازش خودکار چک‌های بانکی و دیگر اوراق بهادار سفارشی کرد. نمونه‌ای از طرح جدید در شکل ۲-۱ نشان داده شده است.

بانک تحقیقات دانشگاهی (فاقد ارزش قانونی)

شماره سریال

تاریخ سال ماه روز

به موجب این چک مبلغ ریال

در وجه یا به حواله کرد پردازید.

محل امضاء (کاملاً داخل مستطیل)

شماره حساب

ریال هزار میلیون میلیارد

شکل ۲-۱ نمونه‌ای از طرح جدید پیشنهادی چک

۹-۱ مفاهیم اصلی در تشخیص دست خط^۱

در این قسمت مفاهیم اصلی، که در پردازش خودکار اسناد نقش دارند، بررسی می‌شوند. در پردازش خودکار اسناد مهمترین قدم، تشخیص دست خط فرد نویسنده می‌باشد. قبل از کاربرد و پیاده‌سازی هر علمی، باید نسبت به آن علم شناخت کامل پیدا شود. علم تشخیص دست خط نیز از این امر مستثنی نمی‌باشد. تعریف یادگیری ماشین، تشخیص الگو، تشخیص نوری کاراکتر و پردازش تصویر در شناخت مفهوم تشخیص دست خط بسیار موثر می‌باشند، زیرا از علومی می‌باشند که باعث پیدایش تشخیص دست خط شده‌اند.

¹Handwritten Recognition

۱-۹-۱ یادگیری ماشین^۱

تعریف فرهنگ لغات از یادگیری: یادگیری عبارت است از بدست آوردن دانش و یا فهم آن از طریق مطالعه، آموزش و یا تجربه، همچنین گفته شده است که یادگیری از بهبود عملکرد از طریق تجربه به دست می‌آید.

تعریف یادگیری ماشین: یادگیری ماشین عبارت است از اینکه چگونه می‌توان برنامه‌ای نوشت که از طریق تجربه یادگیری کرده و عملکرد خود را بهتر کند. یادگیری ممکن است باعث تغییر در ساختار برنامه و یا داده‌ها شود. برخی کاربردهای یادگیری ماشین از قبیل داده‌کاوی^۲، تشخیص گفتار^۳، شناسایی متن^۴، پردازش داده‌های اینترنتی^۵ و بیوانفورماتیک^۶ می‌باشند.

۱-۹-۲ تشخیص الگو^۷

وقتی که یک چهره را تشخیص می‌دهیم، کلمات را متوجه می‌شویم، حروف دست‌نوشته را می‌خوانیم، کلید ماشین داخل جیبمان را با لمس کردن شناسایی می‌کنیم و از روی بوی سیب متوجه می‌شویم که رسیده است یا نه، سبب شده است که به پیچیدگی عملیاتی که تحت عنوان تشخیص الگو در این فرایندها صورت می‌پذیرد کمتر توجه کنیم. تشخیص الگو، روشی است علمی که هدف آن دسته‌بندی اشیا به تعدادی موضوع یا کلاس می‌باشد. تشخیص الگو، شاخه‌ای از هوش مصنوعی^۸ و مبحث یادگیری ماشین است. می‌توان گفت تشخیص الگو، دریافت داده‌های خام و تصمیم‌گیری براساس طبقه‌بندی^۹ داده‌ها است. بیشتر تحقیقات در زمینه تشخیص الگو در رابطه با یادگیری نظارت شده یا یادگیری بدون نظارت است. نظام کلاسه‌بندی یا توصیف معمولاً مبتنی بر وجود یک مجموعه از الگوهایی است که قبلاً کلاسه‌بندی یا توصیف شده‌اند. این مجموعه الگوها را آموزشی و قانون یادگیری منتج شده را قانون یادگیری با نظارت می‌نامند. همچنین یادگیری می‌تواند بدون نظارت باشد

¹Machine learning

²Data Mining

³Speech Recognition

⁴Text Recognition

⁵Internet Data processing

⁶Bioinformatics

⁷Pattern Recognition

⁸Artificial Intelligence

⁹Classification