

دانشگاه یزد

گزارش کارآموزی

ارائه شده به:

دکتر ولی درهمی

توسط:

فرزانه گرامی ( ۹۳۳۸۷۶۳ )

مجتمع فنی و مهندسی

گروه مهندسی کامپیوتر

تابستان ۹۶

## چکیده:

تکنولوژی OCR عبارت است از تشخیص خودکار متون موجود در تصاویر و اسناد و تبدیل آن‌ها به متون قابل جستجو و ویرایش توسط کامپیوتر لذا برای استخراج متن و داده از اسناد و مدارک گزینه‌ی بسیار مناسبی است. این فرایند موجب صرفه‌جویی در زمان برای داده‌هایی است که احتیاج داریم چندباره آن‌ها را مورد پردازش‌های مختلف قرار دهیم و با افزایش روز افزون الکترونیکی شدن کارها موقعیت‌های جدیدی برای استفاده از آن پدید می‌آید.

در اینجا با نمونه‌ای از مدارک در قالب فرم‌های پزشکی سر و کار داریم که لازم است قسمت‌های مشخصی از داده‌های نوشته شده روی آن‌ها را استخراج کرده تا در ادامه اقداماتی از قبیل نگهداری سوابق پزشکی بیماران و استفاده در راستای الکترونیک کردن روند تشخیص مشکلات پزشکی، سهولت بپذیرد.

در همین راستا تلاش بر این بود تا با بررسی روش‌ها و ابزارهای مختلف موجود به استخراج داده‌های مورد نیاز از فرم‌ها بپردازیم و با تغییر پارامترهایی که این ابزارها در اختیار ما قرار می‌دهند نتایج تشخیص را بهبود ببخشیم.

## فهرست مطالب

۱. فصل اول: معرفی شرکت .....	۱
۱.۱. شرکت هوشپاد درخشان ایساتیس (هدا): .....	۱
۱.۱.۱. زمینه‌های فعالیت و نمونه پروژه‌های انجام شده: .....	۱
۱.۱.۲. مهم‌ترین پروژه‌های در دست اجرا: .....	۴
۱.۱.۳. ساختار سازمانی: .....	۴
۲. فصل دوم: استخراج داده از فرم‌های دست نویس به وسیله ی OCR .....	۵
۲.۱. معرفی طرح: .....	۵
۲.۱.۱. ساختار فرم‌ها: .....	۵
۲.۲. معرفی و بررسی نرم افزارهای OCR: .....	۹
۲.۲.۱. ABBYY FineReader .....	۹
۲.۲.۲. Nuance OmniPage .....	۱۰
۲.۲.۳. IRIS Readiris .....	۱۱
۲.۲.۴. Soda PDF .....	۱۱
۲.۳. راه اندازی کتابخانه پردازش تصویر در جاوا: .....	۱۲
۲.۳.۱. پردازش تصویر در جاوا با کمک فریم ورک marvin: .....	۱۲
۲.۳.۲. استفاده از کتابخانه ی ImageJ: .....	۱۳
۲.۴. پیاده‌سازی برنامه OCR با استفاده از کتابخانه‌ها و SDKهای آماده: .....	۱۴
۲.۴.۱. Aspose.OCR for Java .....	۱۴
۲.۴.۲. Asprise Java OCR SDK .....	۱۴
۲.۴.۳. Tess4J .....	۱۶

۳. فصل سوم: نتیجه گیری..... ۱۸

## ۱. فصل اول: معرفی شرکت

### ۱.۱. شرکت هوشپاد درخشان ایساتیس (هدا):

شرکت هوشپاد درخشان ایساتیس با همکاری آزمایشگاه هوش محاسباتی و رباتیک گروه مهندسی کامپیوتر با به کارگیری روش‌های هوشمند در زمینه‌های مدلسازی و پیش‌بینی بر اساس داده، استخراج دانش از داده، سیستم‌های مبتنی بر پردازش تصویر هوشمند، کنترل و اتوماسیون، هدایت ربات‌های متحرک و کنترل بازوهای رباتی فعال می‌باشد.

شرکت تحت مدیریت دکتر ولی درهمی و بهره‌گیری از دانش همکاران گروه (آقای دکتر رضاییان، آقای دکتر لطیف، آقای دکتر محمد رضا زارع، و آقای دکتر دوست حسینی)، هفت محقق ارشد (مهندس نادی، مهندس فتحی نژاد، مهندس نعمتی، مهندس محمد کریمی، مهندس کریمی، مهندس آبادیان زاده مهندس شایانفر)، شش دانشجوی دکترا و ده دانشجوی کارشناسی ارشد فعالیت می‌کند.

### ۱.۱.۱. زمینه‌های فعالیت و نمونه پروژه‌های انجام شده:

۱. مدل‌سازی با توجه به داده‌های سیستم:

یکی از موارد مورد نیاز در صنعت مدل‌سازی یک فرایند، می‌باشد. بدین منظور با توجه به داده‌هایی که به عنوان ورودی و خروجی از فرایند جمع می‌گردد و تعیین یک ساختار مناسب به عنوان تقریب زننده، مدل‌سازی رفتار آن صورت می‌گیرد.

نمونه پروژه‌ها:

- مدل‌سازی تغییرات دمای شمش در کوره پیشگرم شرکت سهامی ذوب آهن اصفهان با استفاده از شبکه‌های عصبی
- مدل‌سازی زمان و هزینه نصب انشعابات گاز و ارائه یک سیستم پشتیبان تصمیم برای تعیین تعداد پیمانکار

۲. پیش بینی رفتار سیستم با استفاده از داده‌ها:

پیش‌بینی رفتار یک فرایند یا سیستم اجتماعی در صنعت و جامعه کاربرد فراوانی دارد. در اینجا داده‌های مختلف جمع‌آوری شده، سپس ویژگی‌های مناسبی که می‌تواند از داده‌ها به عنوان ورودی تعیین گردد به دست می‌آید. در مرحله بعد یک ساختار مناسب ب عنوان تقریب زنده انتخاب و پارامترهای آن جهت پیش‌بینی تنظیم می‌گردد.

نمونه پروژه‌ها:

- پیش‌بینی خاموشی‌های ناخواسته در مناطق مختلف استان یزد با استفاده از شبکه‌های عصبی فازی
- طراحی سیستم فازی برای تشخیص بیماری‌های ریوی بر اساس تنفس‌سنجی
- پیش‌بینی تغییر روند سهام با توجه به مقاله‌های خبری

۳. استخراج دانش از داده‌ها:

یکی از کلیدهای طلایی در فناوری اطلاعات، استخراج دانش از داده‌هایی است که در یک سازمان وجود دارد. به طور معمول سازمان‌ها دارای داده‌های زیادی می‌باشند، با ترکیب این داده‌ها با داده‌های مرتبط دیگر می‌توان اطلاعات گرانمایی را برای سازمان یا صنعت تولید کرد.

نمونه پروژه‌ها:

- تشخیص خرابی (دستکاری) کنتور با استفاده از روش‌های داده کاوی و شبکه عصبی

۴. طراحی سیستم‌های هوشمند:

نمونه پروژه‌ها:

- سیستم توزین عمود حرکت هوشمند
- استخراج عوامل کدهای خرابی دستورالعمل‌های بازدید ترانسفورماتورها و ارائه راهکار مبتنی بر آن جهت هوشمندسازی زمان و موارد بازدید

۵. طراحی و پیاده سازی سیستم‌های نرم‌افزاری:

نمونه پروژه‌ها:

- آنکاریون: سیستم نرم‌افزاری مدیریت پروژه (PMS)
- سیستم خبره تغذیه

## ۶. کنترل ربات متحرک و بازوی ربات:

پس از عصر اطلاعات، هم اینک وارد عصر ربات‌ها شده‌ایم. ربات‌ها نه تنها باعث تسهیل امور، و سرعت بخشیدن به کارها می‌شوند بلکه چالش‌های انجام امور توسط انسان مانند انتقال بیماری، امنیت، آسیب‌دیدگی روحی و جسمی را کاهش می‌دهند. حرکت و کنترل ربات‌های متحرک در حضور موانع، و شرایط ناهموار از نیازهای صنعت می‌باشد. در این مورد می‌توان از ترکیب روش‌های یادگیری سنتی با روش‌های یادگیری نوین همچون یادگیری تقویتی بهره برد. یکی دیگر از نیازهای صنعت بازوهای رباتی است که وظیفه مشترک در بین آنها حرکت از یک نقطه دلخواه و رسیدن به یک هدف می‌باشد. با توجه به اینکه دوربین به عنوان یک سنسور ارزان قیمت و غنی از اطلاعات محسوب می‌شود، ما می‌توانیم از آن برای هدایت بازوی ربات استفاده کنیم.

نمونه پروژه‌ها:

- کنترل بینا مبنای بازوی ربات با بکارگیری روش‌های هوشمند در تقریب ماتریس ژاکوبین
- کنترل بینا مبنای بازوی ربات با استفاده از یادگیری تقویتی
- ترکیب یادگیری با ناظر و یادگیری تقویتی در حرکت ربات متحرک

## ۷. تشخیص اشیا و موارد مطرح با پردازش تصویر:

دوربین‌ها به عنوان سنسورهای ارزان قیمتی که حجم اطلاعات زیادی را منتقل می‌کنند شناخته شده‌اند. در بسیاری از بخش‌های صنعت به خصوص کنترل کیفیت از چشم برای تشخیص عیوب، ارزیابی و اندازه‌گیری استفاده می‌شوند. در اینگونه پروژه‌ها ابتدا تصویر گرفته شده و با انجام پردازش تصویر موارد خواسته شده بدست می‌آید.

نمونه پروژه‌ها:

- تشخیص ترک در پلوس ماشین پراید

## ۸. طراحی و پیاده سازی وب سایت:

امروزه با توجه به گسترش روزافزون استفاده از اینترنت، یکی از راه‌های موثر اطلاع‌رسانی، به کارگیری صفحات وب می‌باشد که یکی از مزایای مهم استفاده از صفحات وب امکان دسترسی به اطلاعات در ۲۴ ساعت روز و ۳۶۵ روز سال و در تمام نقاط جهان می‌باشد. در نتیجه با استفاده از سایت طراحی شده، امکان اطلاع‌رسانی جامعی از فعالیت‌های تجارت مربوطه وجود دارد.

نمونه پروژه‌ها:

- مطالعه، برنامه‌ریزی و راه‌اندازی سامانه مدیریت اطلاعات هیئت نظارت و ارزیابی آموزش عالی استان یزد به منظور ثبت و انعکاس آمار و اطلاعات دانشگاه‌های سطح استان (آدرس:

<http://nezarat.yazd.ac.ir>

- وب سایت آزمایشگاه هوش محاسباتی و رباتیک دانشگاه یزد (آدرس: <http://ce.yazd.ac.ir/lcir>)

• وب سایت زورخانه صاحب الزمان (آدرس: <http://www.zurkhanehsahebalzaman.ir>)

#### ۱.۱.۲. مهم ترین پروژه های در دست اجرا:

۱. یافتن موارد بازدید انجام نشده در پیمان های تعمیرات، و استخراج عوامل عدم انجام و ارائه راهکار

برای بهبود وضعیت

۲. سیستم ارتباط جمعی کارمندان

۳. آشکارسازی هوشمند رخداد آتش به کمک تصاویر ویدئویی در انبار شرکت

۴. سیستم ارسال داده های پایش لحظه ای سیستم آنالایزر گازهای خروجی مولد کرافت و بلوک S1

نیروگاه سیکل ترکیبی یزد

#### ۱.۱.۳. ساختار سازمانی:

ساختار سازمانی شرکت به صورت زیر می باشد:

مدیر عامل: دکتر ولی درهمی

مدیر فنی و قائم مقام: مهندس فرزانه نادی

مدیر بخش نرم افزار: مهندس فاطمه آبادیان زاده

مدیر بخش سیستم های هوشمند: مهندس کریمی



## ۲. فصل دوم: استخراج داده از فرم‌های دست نویس به وسیله‌ی OCR

### ۲.۱. معرفی طرح:

در این طرح، ما با داشتن نمونه‌هایی اسکن شده از فرم‌های پزشکی مورد بررسی که حاوی اعداد و حروف می‌باشند تلاش می‌کنیم اطلاعات قسمت‌های تعریف شده‌ی مورد نیاز را به وسیله‌ی OCR از تصویر خوانده و برای استفاده‌های بعدی استخراج کرد.

برای این منظور در یک روند کلی ابتدا برنامه‌های تجاری و موجود فعلی برای انجام OCR بر روی اسناد را بررسی کرده سپس به راه اندازی کتابخانه پردازش تصویر ساده در جاوا برای خواندن تصاویر و نمایش آنها می‌پردازیم و در نهایت اقدام به راه اندازی یک کد OCR روی تصاویر برش داده شده از فرم‌ها می‌کنیم.

### ۲.۱.۱. ساختار فرم‌ها:

فرم‌هایی که برای بررسی در این طرح مد نظر می‌باشند شامل سه فرم می‌باشند. این فرم‌ها با کدهای F1، F2 و F3 در شکل ۱ تا شکل ۳ آورده شده‌اند.

جدول ۱. توضیحات فرم‌های مربوط به طرح

توضیحات کلی		
شماره فرم	کد توضیح*	توضیحات
F1	F1A	تاریخ بصورت فارسی مشابه ۹۵/۱۲/۱
	F1B	اعداد انگلیسی و بعضاً بصورت فارسی
	F1C	عبارت حداکثر سه حرفی انگلیسی با تعداد حالت‌های محدود
	F1D	گزینه مورد نظر علامت زده شده است
	F1E	عبارت حداکثر سه حرفی انگلیسی با تعداد حالت‌های محدود
F2	F2A	<ul style="list-style-type: none"><li>مقادیر بصورت انگلیسی هستند.</li><li>علامت ممیز بصورت "." و یا "/" وارد شده است.</li></ul>
F3	F3A	<ul style="list-style-type: none"><li>فیلدهای مشخص شده با فلش در فرم جهت تشخیص پردازش می‌شوند.</li><li>مقادیر بصورت انگلیس با تعداد حالت محدود می‌باشند.</li></ul>

\*کد توضیح: این کد به صورت یک مربع مشکی با متن سفید در هر فرم قابل مشاهده می‌باشد.

فرم چشم پزشکی (شماره ۱۲)

**F1A**

تاریخ :

شماره پرونده :

نام و نام خانوادگی :

VA	OD	<b>F1B</b>	OD	<b>F1B</b>
TA	OD	<b>F1B</b>	OS	<b>F1B</b>
Lens	OD	<b>F1C</b>	OS	<b>F1C</b>
Retinopathy OD	No <input type="checkbox"/> Yes <input type="checkbox"/>	Mild NPDR <input type="checkbox"/>	<b>F1D</b>	PDR <input type="checkbox"/>
		Modrae NPDR <input type="checkbox"/>		CSME <input type="checkbox"/>
		Sever NPDR <input type="checkbox"/>		TRD <input type="checkbox"/>
		Vit Hemorrhage <input type="checkbox"/>		
Retinopathy OS	No <input type="checkbox"/> Yes <input type="checkbox"/>	Mild NPDR <input type="checkbox"/>	<b>F1D</b>	PDR <input type="checkbox"/>
		Modrae NPDR <input type="checkbox"/>		CSME <input type="checkbox"/>
		Sever NPDR <input type="checkbox"/>		TRD <input type="checkbox"/>
		Vit Hemorrhage <input type="checkbox"/>		
Past history of treatment	MPC	OD <input type="checkbox"/>	<b>F1D</b>	OS <input type="checkbox"/>
	PRP	OD <input type="checkbox"/>		OS <input type="checkbox"/>
	Avastin	OD <input type="checkbox"/>		OS <input type="checkbox"/>
	Vitrectomy	OD <input type="checkbox"/>		OS <input type="checkbox"/>

OCT : .....

FA : .....

Comments : .....

Plan : **F1E** .....

Signature : .....

تلفن :

شماره پرونده :

نام و نام خانوادگی :

آزمایش	تاریخ	آزمایش	تاریخ	آزمایش	تاریخ
FBS	۱۲۳	FBS	۱۲۳	FBS	۱۲۳
$T_{HPP}$	۱۶۵	$T_{HPP}$	۱۶۵	$T_{HPP}$	۱۶۵
BS * PM	۱۱۵	BS * PM	۱۱۵	BS * PM	۱۱۵
HbA <sup>1</sup> C	۱۰۱	HbA <sup>1</sup> C	۱۰۱	HbA <sup>1</sup> C	۱۰۱
MicroAlb	۵۲	MicroAlb	۵۲	MicroAlb	۵۲
BUN	۲۹	BUN	۲۹	BUN	۲۹
Cr	۱٫۲	Cr	۱٫۲	Cr	۱٫۲
TG	۲۳۳	TG	۲۳۳	TG	۲۳۳
Chol	۱۴۵	Chol	۱۴۵	Chol	۱۴۵
LDL	۵۸	LDL	۵۸	LDL	۵۸
HDL	۴۰	HDL	۴۰	HDL	۴۰
Uire Acid	۴۱۵	Uire Acid	۴۱۵	Uire Acid	۴۱۵
AST	۲۶	AST	۲۶	AST	۲۶
ALT	۲۷	ALT	۲۷	ALT	۲۷
BiL-T		BiL-T		BiL-T	
BiL-D		BiL-D		BiL-D	
Alp		Alp		Alp	
T <sup>r</sup>		T <sup>r</sup>		T <sup>r</sup>	
T <sup>r</sup>		T <sup>r</sup>		T <sup>r</sup>	
THS		THS		THS	
T <sup>r</sup> up		T <sup>r</sup> up		T <sup>r</sup> up	
Na	۱۳۶	Na	۱۳۶	Na	۱۳۶
K	۴۱۲	K	۴۱۲	K	۴۱۲
Ca	۲۱۷	Ca	۲۱۷	Ca	۲۱۷
Ph	۳۱۸	Ph	۳۱۸	Ph	۳۱۸
Vit D		Vit D		Vit D	
PSA		PSA		PSA	
Ferritin	۹۳۱۶	Ferritin	۹۳۱۶	Ferritin	۹۳۱۶
TIBC		TIBC		TIBC	
Fe		Fe		Fe	
GTT <sup>1</sup> h		GTT <sup>1</sup> h		GTT <sup>1</sup> h	
GTT <sup>r</sup> h		GTT <sup>r</sup> h		GTT <sup>r</sup> h	
GTT <sup>r</sup> h		GTT <sup>r</sup> h		GTT <sup>r</sup> h	
Urin vol <sup>1</sup> /h		Urin vol <sup>1</sup> /h		Urin vol <sup>1</sup> /h	
Urin pro <sup>1</sup> /h		Urin pro <sup>1</sup> /h		Urin pro <sup>1</sup> /h	
Urin Cr <sup>1</sup> /h		Urin Cr <sup>1</sup> /h		Urin Cr <sup>1</sup> /h	

شکل ۲. فرم شماره F2



## ۲.۲. معرفی و بررسی نرم افزارهای OCR:

در زمینه‌ی OCR تعداد زیادی نرم افزارهای متنوع که به صورت رایگان یا تجاری ارائه شده‌اند وجود دارد که با در نظر گرفتن عواملی مثل دقت و کیفیت تشخیص، تعداد زبان‌های قابل پشتیبانی، رابط کاربری، سرعت و پشتیبانی از پردازنده‌های چند هسته‌ای، توانایی ذخیره سازی فایل‌های خروجی با فرمت‌های متنوع و غیره نرم افزارهای تجاری نتایج قوی تر و بهتری ارائه می‌دهند که از بین این نرم افزارها، نرم افزارهای ABBYY FineReader، Nuance OmniPage و ReadIRIS در اکثر بررسی‌های انجام گرفته روی نرم افزارهای این زمینه با نتایج خوبی در رتبه‌های اولیه قرار گرفته‌اند.

### ۲.۲.۱. ABBYY FineReader:

سایت رسمی: [www.abbyy.com](http://www.abbyy.com)

این نرم افزار OCR هوشمند و با قابلیت استفاده ساده و نیز یک مبدل PDF می‌باشد. این برنامه عملکردی قدرتمند برای تبدیل تصاویر دریافت شده از اسکنر، دوربین عکاسی دیجیتال و یا فکس را همانند فایل‌های PDF به فرمت‌های متنی قابل ویرایش و با قابلیت جستجو تبدیل می‌نماید. این نرم افزار فرمت بندی و صفحه بندی اسناد را به صورت دقیق نگهداری می‌کند و نیز از طیف وسیعی از زبان‌ها برای شناسایی و فرمت فایل‌های خروجی پشتیبانی به عمل می‌آورد. نسخه 12 آن از طرف PC Magazine امتیاز عالی گرفته است.

آخرین نسخه‌ی منتشر شده: 14

سال انتشار: 2017

نسخه‌ی بررسی شده: ABBYY FineReader 12 Corporate

سال انتشار: 2013

این نرم افزار علاوه بر اینکه دقت خوبی دارد از تعداد زیادی زبان (۱۹۲ زبان در آخرین نسخه) پشتیبانی می‌کند که زبان عربی نیز جزء آنها هست ولی از زبان فارسی پشتیبانی نمی‌کند همچنین دارای ۳۹ لغت نامه برای زبان‌های مختلف است. ورودی آن می‌تواند انواع فایل‌های تصویری با فرمت‌های مختلف، فایل‌های اسکن شده، فایل‌های PDF و اسناد الکترونیکی باشد و برای ذخیره کردن فایل‌های خروجی فرمت‌های متنوعی از جمله Word، Excel، Powrepoint، Rich Text Format، HTML، txt، CSV، Epub، Djvu را دارد. این نرم افزار برای سیستم عامل‌های Windows، Mac OS X و Linux ارائه می‌شود همچنین این نرم افزار منبع باز نیست ولی SDKهایی مجزا مختص توسعه دهندگان برای استفاده در برنامه‌هایشان برای فروش ارائه می‌دهد.

در بررسی عملکرد آن روی فرم‌های نمونه قسمت‌های چاپی انگلیسی را با دقت خوبی تشخیص داد ولی در تشخیص دست نوشته ها نسبتا ضعیف عمل کرد و فقط برخی حروف واضح تر را درست تشخیص داد و با استفاده از زبان عربی برای قسمت‌های فارسی برخی کلمات را نسبتا خوب تشخیص می‌داد. در تشخیص جداول عملکرد متوسطی داشت (بسته به کیفیت تصویر در بعضی ساختار جدول را تشخیص نمی‌داد).

## ۲.۲.۲: Nuance OmniPage

سایت رسمی: [www.nuance.com](http://www.nuance.com)

OmniPage به عنوان یکی از بهترین نرم افزارهای OCR عمل می‌کند و دقت خوبی در حفظ ساختار، رنگ و فونت سند دارد. کمپانی ارائه دهنده‌ی این نرم افزار آن را در دو نسخه‌ی Standard و Professional (در نسخه جدید این مدل با عنوان Ultimate ارائه شده است) عرضه می‌کند. آزمایشگاه مستقل مجله معتبر PC Magazine طی تست‌هایی که از نرم افزار گرفته، کارایی و دقت بسیار بالای این نرم افزار را تایید کرده است. این نرم افزار همچنین از ۱۲۰ زبان پشتیبانی می‌کند (فاقد زبان فارسی) و پس از تشخیص امکان ویرایش و چرکنویسی را نیز فراهم می‌کند.

آخرین نسخه‌ی منتشر شده: 19.2

سال انتشار: 2015

نسخه‌ی بررسی شده: Nuance OmniPage Professional 17.1

سال انتشار: 2009

این نرم افزار به طور کلی از دقت خوبی برخوردار است. انواع مختلفی از فرمت‌های فایل‌های تصویری و اسکن شده و اسناد مانند فایل‌های PDF، Word، Powerpoint، Excel، txt و غیره را به عنوان ورودی می‌پذیرد و قابلیت ذخیره خروجی با فرمت‌های گفته شده و همچنین فرمت‌های HTML، XML، CVS و غیره را دارد. این نرم افزار برای Windows و Mac OS ارائه می‌شود و منبع باز نیست. همچنین قیمت نسخه‌ی اصلی این نرم افزار از بقیه بسیار گران تر است و همانند ABBYY برای استفاده‌ی توسعه دهندگان SDKهایی برای فروش ارائه داده است.

در بررسی فرم‌ها با این نرم افزار، تشخیص قسمت‌های چاپی انگلیسی دقت خوبی داشت ولی در تشخیص دست نوشته ها مثل FineReader ضعیف بود. این نرم افزار فاقد زبان فارسی و عربی است و قسمت‌های فارسی

را تشخیص نمی‌داد. در تشخیص ساختار جداول در صورت خوب بودن کیفیت تصویر کمی بهتر از FineReader عمل می‌کرد.

### ۳.۲.۲. IRIS Readiris

سایت رسمی: [www.irislink.com](http://www.irislink.com)

این نرم افزار یکی از قوی‌ترین نرم افزارهای OCR است، این نرم افزار معمولاً به صورت رایگان بر روی دستگاه های اسکنر ارائه می‌شود و بررسی‌ها دقت و کارایی خوبی را برای آن ذکر کرده اند. این نرم افزار از ۱۳۰ زبان پشتیبانی می‌کند و دارای ۱۵ لغت نامه‌ی داخلی برای زبان‌های پر استفاده‌تر است. نسخه جدید این نرم افزار که مخصوص کشورهای خاورمیانه عرضه شده علاوه بر زبان های لاتین از زبان های ابری، عربی و فارسی نیز پشتیبانی می‌کند.

آخرین نسخه‌ی منتشر شده: 16

سال انتشار: 2016

نسخه‌ی بررسی شده: Readiris Corporate 16

سال انتشار: 2016

این نرم افزار به طور کلی دقت و امکانات مناسبی برای پردازش اسناد و تصاویر دارد. انواع متنوعی از فایل‌ها مانند فرمت‌های مختلف تصاویر و اسکن‌ها، فایل‌های PDF، Word، Powerpoint، Excel و Djvu را به عنوان ورودی می‌گیرد و برای ذخیره‌ی خروجی علاوه بر فرمت‌های گفته شده فرمت‌های CSV، XML، Epub، txt و غیره را ارائه می‌دهد. این نرم افزار برای Windows و Mac OS ارائه می‌شود و منبع باز نیست.

نتیجه‌ی بررسی نمونه‌ها با این نرم افزار ضعیف تر از دو نرم افزار دیگر بود و حتی در قسمت‌های چاپی انگلیسی برای نمونه‌ها خوب عمل نمی‌کرد. با اینکه از زبان فارسی پشتیبانی می‌کند تشخیص قسمت‌های فارسی دقت بالایی نداشت و به طور کلی با بررسی چند نمونه‌ی فارسی دیگر می‌توان گفت در تشخیص زبان فارسی متوسط است مثلاً در فایل PDF فارسی کلمات را درست با درصدی خطا می‌دهد که در عکس‌ها این تشخیص بسیار ضعیف‌تر و با خطای بیشتر می‌شود.

### ۴.۲.۲. Soda PDF

سایت رسمی: [www.sodapdf.com](http://www.sodapdf.com)

این نرم افزار یک ابزار حرفه ای مدیریت فایل های PDF است که برای ایجاد، ویرایش و تبدیل فایل های PDF طراحی شده است همچنین سازگار با تکنولوژی OCR می باشد. در بررسی ها نتایج نسبتاً خوبی دارد و رتبه پایین تری نسبت به ۳ نرم افزار دیگر دارد و در تشخیص تطاویر با کیفیت پایین ضعیف تر عمل می کند. قابلیت پشتیبانی از ۸ زبان را دارد.

آخرین نسخه ی منتشر شده: 9.2

سال انتشار: 2017

نسخه ی بررسی شده: Soda PDF Professional 5 + OCR

سال انتشار: 2013

فایل های PDF و فایل های تصویری را به عنوان ورودی دریافت می کند و قابلیت ذخیره ی خروجی را به صورت فایل های PDF، Word، Excel و txt را دارد. فقط برای سیستم عامل Windows عرضه شده و منبع باز نیست.

نتیجه ی بررسی نمونه ها با این نرم افزار اینگونه بود که در فایل نهایی عکس یا فرم اسکن شده را به عنوان پس زمینه ی سند و مواردی که تشخیص داده بود را به شکل متن قابل انتخاب و جستجو ذخیره می کرد. قدرت و دقت تشخیص آن نیز در فرم های اولیه نمونه ها بیشتر از قدرت تشخیص عبارات داخل ساختارهای جدولی بود.

## ۲.۳. راه اندازی کتابخانه پردازش تصویر در جاوا:

### ۲.۳.۱. پردازش تصویر در جاوا با کمک فریم ورک marvin :

برای استفاده از امکانات جالب این فریم ورک کافی است کارهای زیر را انجام دهیم. ابتدا نسخه ای از این فریم ورک را دانلود می کنیم. آن را از حالت فشرده خارج کرده و در مسیر پروژه تان قرار می دهیم. پوشه marvin شامل دو پوشه است یکی فریم ورک ها و دیگری پلاگین ها. جدا بودن این دو پوشه اجازه می دهد افزودن پلاگین جدید بدون هیچ تغییری در پروژه انجام شود چون همه ی پلاگین ها توسط بازتاب لود می شوند. سپس فایل jar فریم ورک را در لیست کتابخانه ها اضافه می کنیم. سپس طبق آموزش های ذکر شده در سایت جاواکاپ<sup>۲۱</sup> می توان

---

<sup>۱</sup> پردازش تصویر در جاوا با کمک فریم ورک marvin (قسمت اول)

<sup>۲</sup> پردازش تصویر در جاوا با کمک فریم ورک marvin (قسمت دوم)



[illegible]

شكل ٨ نتيجة الگوریتم  
*Sobel*

## ۴.۲. پیاده‌سازی برنامه OCR با استفاده از کتابخانه‌ها و SDKهای آماده:

پس از جستجو درباره‌ی OCR در جاوا و SDKها و کتابخانه‌های موجود در این زمینه برای پیاده‌سازی از Aspose.OCR for Java، Asprise Java OCR SDK و Tess4J استفاده کردم. ورودی برنامه‌های نوشته شده با هر کدام از این SDKها تصاویر برش داده شده از فرم F2 (شکل) که شامل اعداد و ممیز هستند و قطعه تصاویر برش داده شده از فرم F1 که شامل حروف تاییپی و دست نوشته بود، می‌باشد.

### ۴.۲.۱. Aspose.OCR for Java:

سایت اصلی: [www.aspose.com](http://www.aspose.com)

Aspose.OCR یک API تشخیص حروف و علامت‌ها (OCR & OMR) برای جاواست که امکان اضافه کردن این قابلیت‌ها به برنامه را می‌دهد. چند زبان و ویژگی‌های مختلف مثل مشخص کردن قسمتی از متن برای تشخیص را دارا می‌باشد. قیمت انواع مختلف این SDK از \$399 تا \$12572 متغیر بوده ولی امکان دانلود و استفاده به صورت Free Trial از نسخه‌ی ۱۷,۶ API را فراهم کرده است.

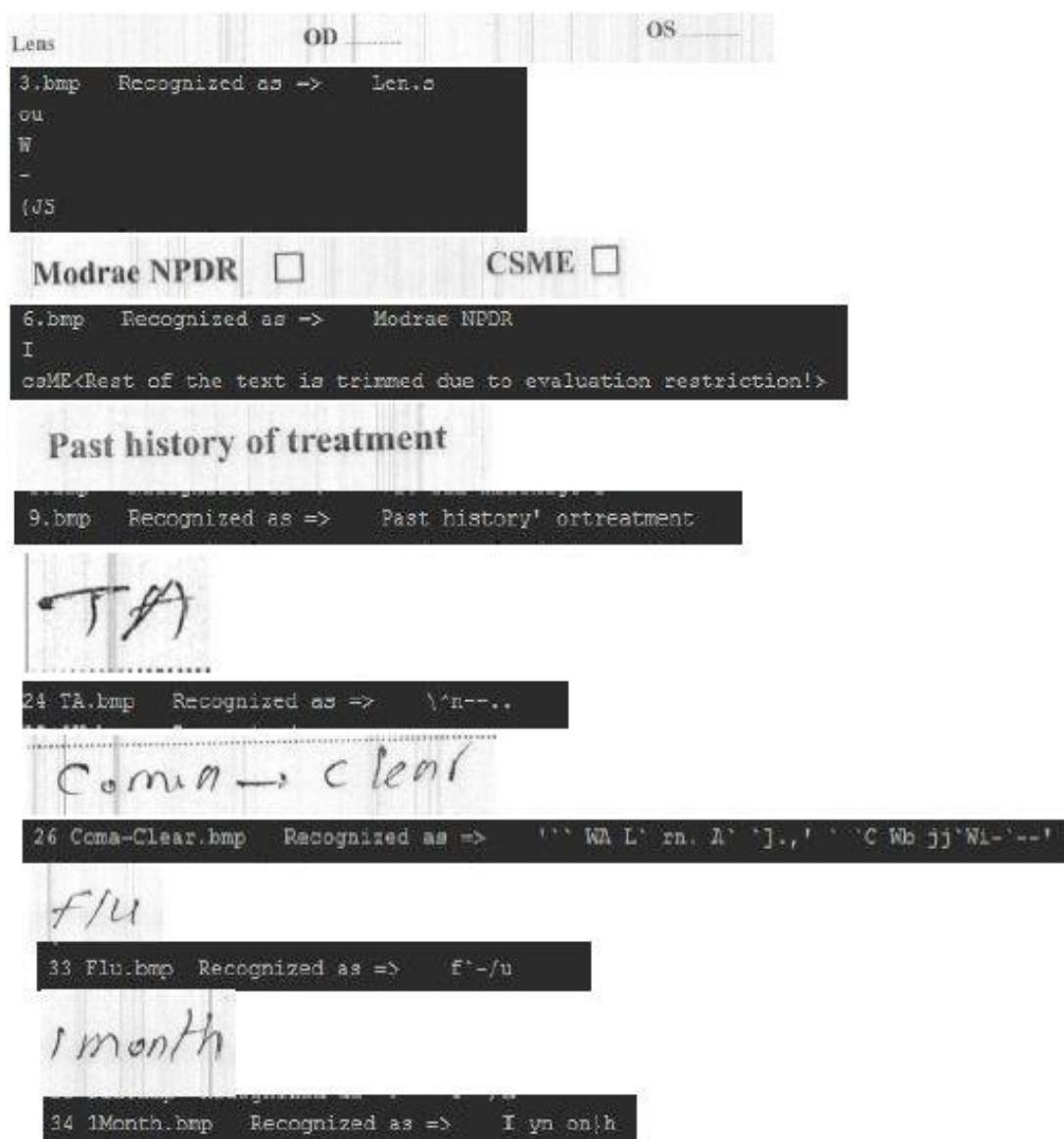
نتیجه‌ی تشخیص اعداد با استفاده از Aspose ابتدا دارای خطای بالا و دارای حروف نیز بود ولی بعد از تنظیم پارامتر Whitelist آن برای تشخیص اعداد ۰ تا ۹ و / درصد خطای تشخیص آن کمتر شد و با اینکه همچنان در تشخیص اعداد و گاهی تعداد کاراکترهای موجود در تصویر خطا دارد، خطای آن در تشخیص اعداد از دو SDK دیگر کمتر است. با این حال استفاده از آن برای تشخیص عکس‌های حاوی عبارات تاییپی و دست نوشته نمونه که از فرم F1 استخراج شده بودند نتایج ضعیف‌تری نسبت به تشخیص اعداد داشت. در تشخیص عبارات تاییپی در مقایسه با تشخیص دست نوشته‌ها بسیار بهتر عمل کرد زیرا تنها قادر به تشخیص برخی حروف کم از دست نوشته‌ها بود (شکل ۹).

### ۴.۲.۲. Asprise Java OCR SDK:

سایت اصلی: [www.asprise.com](http://www.asprise.com)

Asprise Java OCR and barcode recognition SDK یک کتابخانه‌ی API با کارایی بالا برای استخراج متن و بارکد از اسناد اسکن شده ارائه می‌دهد که علاوه بر ویژگی‌های دیگر از زبان‌ها و فرمت‌های مختلف بارکدها پشتیبانی می‌کند. این SDK در سه نوع متفاوت با قیمت \$4998 تا \$12998 به فروش می‌رسد ولی اجازه‌ی دانلود Free Trial آن را می‌دهد که پس از گذشت مدتی منقضی می‌شود.

نتیجه‌ی تشخیص اعداد با استفاده از این کتابخانه دارای خطای بیشتری نسبت به Aspose بود و تعریف template برای آن برای معرفی حالت‌های مختلف اعداد موجود در تصاویر چندان مؤثر نبود. تنظیم پارامتر PROP\_LIMIT\_TO\_CHARSET آن ابتدا با خطا مواجه می‌شد در نتیجه از دوتای دیگر بیشتر خطا داشت و در تشخیص آن حروف و کاراکترهای دیگر نیز موجود بود ولی در تلاش‌های بعدی به درستی تنظیم شد و حروف دیگر در نتایج تشخیص وجود نداشت ولی همچنان در تشخیص اعداد از Aspose.OCR ضعیف تر بود. در تشخیص عبارات تایی عملکردی اندکی بهتر از Aspose.OCR داشت ولی در تشخیص دست نوشته‌ها همچنان تشخیص خوبی نداشت.

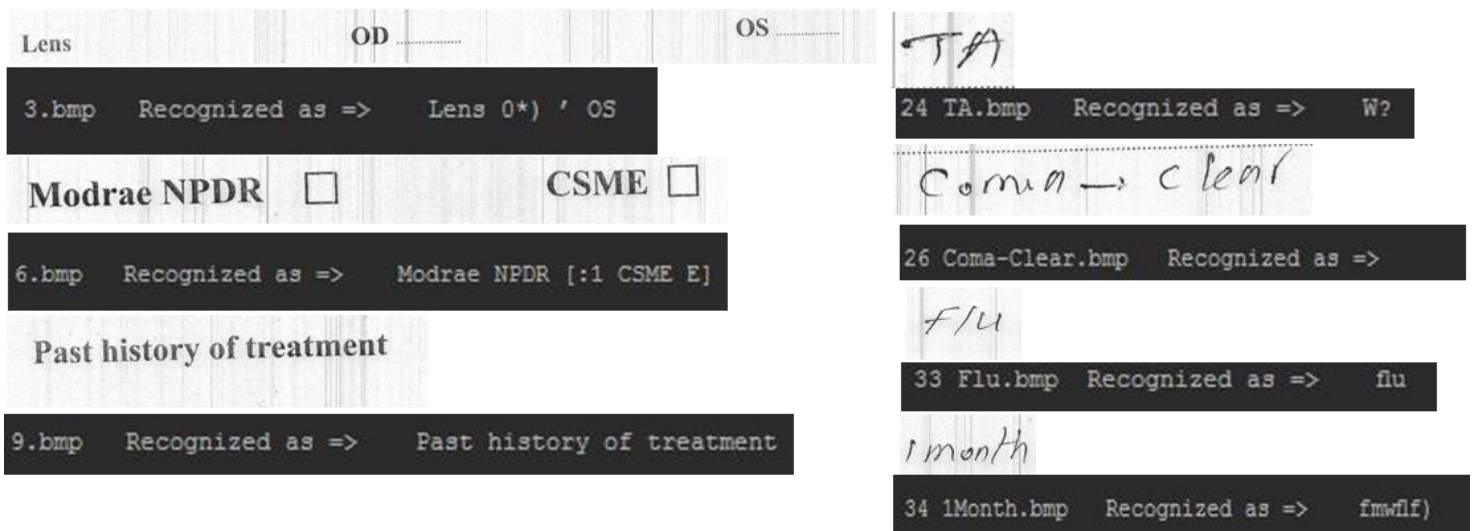


شکل ۹ نمونه‌هایی از تشخیص عبارات تایی و دست نوشته توسط Aspose.OCR

## ۲.۴.۳. Tess4J:

Tess4J یک JNA Wrapper از Tesseract OCR API برای جاواست. این کتابخانه نیز از زبان‌های مختلف پشتیبانی می‌کند ولی برای استفاده از آن چند کتابخانه‌ی دیگر هم نیاز است.

نتیجه‌ی تشخیص اعداد با استفاده از این کتابخانه دارای خطای بیشتری نسبت به Aspose بود ولی پس از تنظیم پارامتر Char\_whitelist آن برای اعداد و / خطای آن از Asprise کمتر شد گرچه در برخی تصاویر در تعداد اعداد را بیشتر از آنچه که هست تشخیص می‌دهد. در تشخیص حروف تاییی بهتر و با فرمتی تمیزتر از Aspose.OCR عمل کرد ولی در تشخیص دست نوشته مانند دو SDK قبل عملکرد خوبی نداشت (شکل ۱۰).



شکل ۱۰ نمونه‌هایی از تشخیص عبارات تاییی و دست نوشته توسط Tess4J

برای مشاهده‌ی نمونه‌ای از تشخیص‌ها و مقایسه در جدول ۲ اعداد فرم نمونه‌ی F2 (شکل ۱۱) و تشخیص آنها توسط Aspose.OCR و Tess4J آورده شده است.

جدول ۲ اعداد موجود در فرم F2 و نتیجه‌ی تشخیص آنها توسط Aspose.OCR و Tess4J

تعداد در فرم F2	تشخیص Aspose	تشخیص Tess4J
1.2	9/4	12
1.6	1/9	/5
1.6	14/63	7 / 4
3.8	3/446455	345277

[illegible]

4.2	4/2	411
4.2	4/2	4/2
4.5	4/5	6 /5
4.6	215 4	7 11
7.6	7/6	1/ 5
9.7	2//9	2/ 1
10.4	/914	7 9/ 4
26	26	25
27	14	21
27	524	2
29	49	9
40	44	140
42	426	41
42	42	7 1
52	52	52
53	653	53
58	0208	278
69.5	4725	67/5
93.6	2	9 314
115	116	1 5
123	123	523
136	136	66
139	401000/7	13 7
143	/43/46 006	12737 7
145	/45	1 45
146	/46	145
147	144	1
159	159	57
165	1695800	10
181	1424	1
199	/92	9/
201	28//00	3 4
233	37	3133

شکل ۱۱ اعداد انگلیسی، موجود در فرم F2



### ۳. فصل سوم: نتیجه گیری

همان طور که از ابتدا تا اینجا با ماهیت OCR، مسئله‌ی مورد بررسی و برخی راهکارهای موجود برای آن به عنوان مثالی از کاربرد آن آشنا شدیم دانستیم که در عصر امروز که تمام فرایندها و رویکردها به سمت الکترونیکی شدن می‌روند تکنولوژی‌ها و سیستم‌های هوشمند روز به روز مطرح تر شده و جای خود را هر چه بیشتر حتی در زندگی روزمره افراد باز می‌کنند. به کار گیری سامانه‌های هوشمند در انواع فرایندهای فعلی می‌تواند علاوه بر صرفه جویی در برخی فاکتورها مثل زمان و انرژی؛ باعث کاهش خطا، افزایش راندمان و بهبود کارآیی، ایجاد سهولت در کارهای مختلف و حتی پیشرفت هر چه بیشتر در خود سامانه‌های هوشمند و کاربردهای آنها در جنبه‌های مختلف زندگی انسان‌ها شود. با این حال چون این سامانه‌ها با محیط پویای انسان‌ها درگیر هستند مانند سایر سیستم‌های پویای دیگر همواره در حال تغییر و تکمیل هستند و نمی‌توان ادعا کرد که در هر لحظه همواره بی‌عیب هستند.

برای مثال در همین مسئله‌ی ما، با وجود پیشرفت خوبی که تا کنون در زمینه OCR صورت گرفته ولی نمی‌توان گفت که کارکرد ابزار موجود روی نمونه‌های ما کاملاً مورد قبول بوده. در واقع با خاص‌تر شدن پارامترها در مواردی مثل مسئله‌ی ما به نظر می‌رسد استفاده از این ابزار در راهکاری که خود آن را پیاده سازی کنیم عملکرد مطلوب‌تری داشته باشد. این طرح نمونه‌ی گویایی از کاربرد و نحوه‌ی بهره‌گیری از امکانات سیستم‌های هوشمند و زیرشاخه‌های هوش مصنوعی (پردازش تصویر) را در سطح ساده‌ای به خوبی بیان کرد و می‌توان مطمئن بود که با سرعتی که علم، به خصوص علوم کامپیوتر در حال پیشرفت است می‌توان شاهد ظهور تکنولوژی‌های خلاقانه و جالب برای استفاده‌ی مفید و راحت در جنبه‌های مختلف زندگی انسان بود.