



هوش مصنوعی

سیستم فرم خوان فارسی (Formreader)


نشانی: تهران - خیابان ستارخان، سه راه تهران ویلا، خیابان عنبر افشان، خیابان فرخ کیش، پلاک ۴، واحد ۲

نمابر: ۶۶۵۵۲۳۰۶ (۰۲۱)

شماره تماس: ۶۶۵۵۲۳۰۵ (۰۲۱)

وب سایت: [Http/ www.aisoft.ir](http://www.aisoft.ir)

پست الکترونیکی: [info@aisoft.ir](mailto:info@aisoft.ir)

صفحه ۲	شرکت هوش مصنوعی رایورز	
	سیستم فرم خوان فارسی (Formreader)	

## فهرست مطالب

۳	معرفی سیستم فرم خوان فارسی (Formreader)
۴	قابلیت های سیستم
۹	کاربردهای سیستم



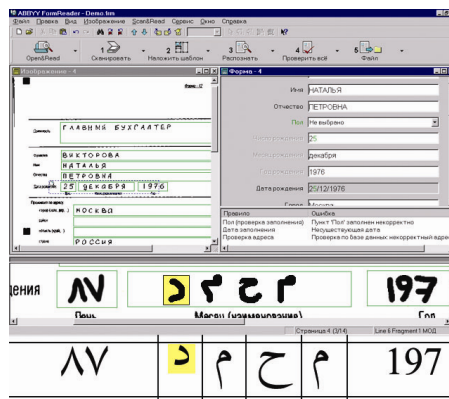
## ❖ معرفی سیستم فرم خوان فارسی (Formreader)


همگام با پیشرفت و مکانیزه شدن امور مختلف و کنار گذاشتن کاغذ (paperless) و تبدیل آنها به فایل‌های دیجیتالی، نرم افزارهای دیگری تولید می شوند تا روند ورود اطلاعات به کامپیوتر (Data entry) را بدون دخالت انسان انجام دهند.

این سیستمها بر اساس برنامه های مبتنی بر هوش مصنوعی عمل می کنند و می توانند برخی از کارها را یاد گرفته و با سرعت بالای ورود اطلاعات و خطای پایین خود، نیاز به اپراتور را بسیار کاهش دهند. نوعی از این نرم افزارها، سیستم فرم خوان (Formreader) می باشد که اخیرا در کشورهای پیشرفته دسته بزرگی از امور مربوط به ورود اطلاعات فرمهای مختلف به کامپیوتر را انجام می دهد.

فرم خوان (Formreader) یک نرم افزار قدرتمند جهت تشخیص انواع فرمها، پیدا کردن آیتمهای اطلاعاتی مورد نیاز و استخراج (Capture) آنها می باشد که چک کردن اطلاعات استخراج شده بوسیله کاربر و ورود آنها به پایگاه های داده (Database) توسط این نرم افزار انجام می شود.

این سیستم دارای تکنولوژی OCR تشخیص حروف چاپی، ICR تشخیص حروف و اعداد، دستنویس ناپیوسته، OMR (Optical Mark Recognition) به منظور خواندن شکلها، علامتها و نشانه ها و OBR (Optical Barcode Recognition) به منظور خواندن بارکد (Barcode) می باشد.



صفحه ۴	شرکت هوش مصنوعی رایورز	
	سیستم فرم خوان فارسی (Formreader)	


## ❖ قابلیت های سیستم

### قابلیت تعریف فرم های پیچیده

- امکان تعریف هر نوع پرسش نامه و فرم در محیط کاملا گرافیکی
- امکان تعریف انواع پاسخ نامه ، پرسش نامه و فرم با هر تعداد گزینه ، سوال و هر نوع قالب بندی گزینه ها
- امکان تعریف انواع کدینگ ( معمولی ، باینری ، عددی ، بارکد) عمودی و افقی برای فرم یا پرسش نامه
- امکان تعریف محدوده های تصویر برای تشخیص اعداد و حروف فارسی دست نویس و تایپ شده
- امکان تعریف محدوده های تصویر برای ذخیره تمام یا بخشی از تصویر فرم و پرسش نامه
- امکان تعریف رنگ فرم و حذف رنگ های اضافی در فرم
- امکان تعریف فرم ها یا پرسش نامه های بیش از یک صفحه با قالب بندی گوناگون برای هر صفحه
- امکان تعریف و طراحی پنجره ورود اطلاعات برای فیلدهای اطلاعاتی که تشخیص آنها توسط سیستم ممکن نباشد.

### قرائت حروف و اعداد

- امکان تشخیص حروف و اعداد دستنویس فارسی با دقت ۹۸ درصد با استفاده از ICR فارسی
- خواندن اعداد و متون چاپی فارسی ، عربی و ۱۶ زبان دیگر با دقت ۹۸ درصد با استفاده از OCR
- امکان تشخیص اعداد و حروف فارسی تایپ شده با دقت ۱۰۰ درصد
- امکان تشخیص ۱۵۰ کاراکتر تایپ شده و ۵۰ کاراکتر دست نویس در ثانیه

صفحه ۵	شرکت هوش مصنوعی رایورز	
	سیستم فرم خوان فارسی (Formreader)	

- عدم وابستگی به سایز و ابعاد ، شیوه نوشتن ، فونت و رنگ متن

### تشخیص ضربدر و چک مارک

- خواندن چک مارکها، شکلها و علامتها با استفاده از تکنولوژی OSR با دقت ۹۹/۵ درصد
- امکان تشخیص گزینه هایی که کم رنگ پر شده اند.
- امکان تشخیص گزینه هایی که کثیف پاک شده اند.
- امکان تشخیص گزینه هایی که تاخوردند ، پاره شده اند و یا کج اسکن شده اند.
- امکان تشخیص گزینه هایی که با خودکار یا ماژیک پر شده اند.

### تشخیص انواع بارکد

- خواندن بارکد (Barcode) با استفاده از تکنولوژی OBR با دقت ۹۹/۷ درصد
- امکان تعریف جهت بارکد ( راست به چپ ، چپ به راست ، بالا به پایین ، پایین به بالا )
- امکان تنظیم چیدمان بارکد ( افقی و عمودی )

### تنظیمات ذخیره سازی

- امکان طبقه بندی و بایگانی اطلاعات فرم های قرائت شده در جدول متفاوت
- امکان تعریف فرمت ذخیره سازی اطلاعات به صورت فایل اکسل ، اکسس یا متنی
- امکان ذخیره سازی تمام یا قسمتی از تصویر فرم به صورت تصویر واقعی یا تصویر پردازش شده
- امکان تعریف فرمت ذخیره سازی تصاویر ( gif ، jpg ، bmp و ... )



### امکان بازبینی (Verification) و تأیید (Validation) دسته ای اسناد

قابلیت یاد شده این امکان را می دهد تا بعد از تشخیص و خواندن آیتمهای مورد نیاز فرمها توسط سیستم ، کاربر بتواند اقدام به بازبینی و چک کردن (Verification) این اطلاعات با اصل سند نماید. همچنین بمنظور جلوگیری از اتلاف وقت و تسریع هر چه بیشتر این روند می توان فیلدهای مشابه هم را به صورت دسته ای به سیستم معرفی کرد تا سیستم بعد از چک کردن هر دسته با کاربر به صورت اتوماتیک به سراغ دسته بعدی برود.

بعد از مرحله باز بینی ، سیستم وارد مرحله تأیید (Validation) می شود . در این بخش سیستم با کمک گرفتن از قوانین تعریف شده برای آن (Validation rules) شروع به چک کردن اطلاعات خوانده شده با مراجعی (References) مثل پایگاه داده یا لغت نامه می کند. به این ترتیب صفحاتی که دارای اشتباهاتی هستند به صورت اتوماتیک علامت گذاری می گردند و اشتباه مورد نظر نیز مارک (Highlight) می شود تا کاربر بعدا با مراجعه به آنها اقدام به تصحیح این اشتباهات نماید ؛ در نهایت این اطلاعات صحیح که از دو فیلتر بالا رد شده اند به پایگاه داده (database) فرستاده می شوند.

### بایگانی الکترونیکی

- ایجاد بایگانی الکترونیکی بطور اتوماتیک
- امکان ذخیره سازی و بازیابی تصاویر فرم ها و پرسش نامه هایی که سیستم قادر به قرائت آنها نیست.
- امکان مدیریت سریع تصاویر در تعداد بالا
- امکان مشاهده اطلاعات جدول های متفاوت با یک کلیک ساده
- امکان شبیه سازی تصویر مجازی فرم از بانک اطلاعاتی



### قابلیت استخراج انعطاف پذیر اطلاعات (Flexicapture)

فرمها می توانند به دو گروه تقسیم شوند : فرمهای کاملا یکسان و همانند هم و فرمهای نیمه ساختار یافته و مشابه.


فرمهای نوع اول که همگی دارای یک شکل مشخص و دارای فیلدهای ثابتی هستند براحتی و با یک تعریف اولیه توسط سیستم عادی فرم خوان (Formreader) خوانده می شوند ؛ در حالیکه فرمهای نیمه ساختار یافته که تنها دارای اطلاعات و شکلهای مشابه هم هستند را نمی توان براحتی و مستقیم توسط سیستم فرمخوان عادی (Formreader) خواند و برای اینگونه فرمها یک ابزار اضافه ای با نام Flexicapture نیاز هست.

Flexicapture ابزار قدرتمندی است که در آن با تعریف یکسری الگوریتمها و قوانین با توجه به تنوع فرمهای نیمه ساختار یافته خود که Flexilayouts نامیده می شوند ، می توان اقدام به شناسایی هوشمند مکان فیلدهای روی این فرمهای مشابه که جایگاه ثابتی ندارند ولی نوع اطلاعات آنها یکسان است ، کرد. انواع نامه ها، فرمهای ثبتی، اسناد حقوقی، فاکتورها و صورتحسابهای مالی نمونه هایی از این گونه فرمها هستند.

بعد از تعریف این قوانین و الگوریتمها ( Flexilayouts ) باید آنها را به داخل سیستم فرم خوان بفرستیم (Export) تا فرم خوان با استفاده از آنها اقدام به شناسایی این فرمها و تشخیص و استخراج اطلاعات فیلدهای مورد نظر ما کند.

### از دیگر قابلیت های سیستم می توان به موارد زیر اشاره کرد :

در صورتیکه در فرم ، فیلدهای اطلاعاتی وجود داشته باشد که نرم افزار قادر به تشخیص مقدار آن ها نباشد ( مانند آدرس ) ، در این نرم افزار امکان تعریف پنجره ورود اطلاعات وجود دارد تا بتوان مقادیر این فیلدها را نیز در سیستم ذخیره کرد . پس از تعریف پنجره ورود اطلاعات برای پرسش نامه ، این


صفحه ۸	شرکت هوش مصنوعی رایورز	
	سیستم فرم خوان فارسی (Formreader)	

پنجره هنگام اسکن کردن هر پرسش نامه باز شده و این امکان را به اپراتور سیستم می دهد تا در همان لحظه با دیدن تصویر فیلد اطلاعاتی بر روی مانیتور، مقدار فیلد را در سیستم ذخیره نماید .  
در این نرم افزار عملیات پیش پردازش تصویر قبل از شروع تشخیص وجود دارد مانند اصلاح انحراف زاویه ، تشخیص جهت فرم ، تمیز کردن کثیفی های روی فرم و ...

موارد تکراری در بانکهای اطلاعاتی اعلام می شود .

رابط کاربری آسانی دارد و با تمام اسکنرها سازگار است .



صفحه ۹	شرکت هوش مصنوعی رایورز	
	سیستم فرم خوان فارسی (Formreader)	

## ❖ کاربردهای سیستم

این سیستم می تواند انواع فرمهای کاغذی که اطلاعات در آنها به صورت دستی یا ماشینی وارد شده است را با دقت تشخیص داده و اطلاعات مربوط به آنها را بخواند. همچنین می توان با استفاده از سیستم طراح فرم (Form designer) ، فرمهای ماشین خوان ( Machine Readable ) را تولید کرد. این فرم های کاغذی مواردی مانند زیر هستند :

- انواع بلیط مانند بلیط هواپیما ، قطار و ...
- چک های بانکی
- امتحانات و پرسشنامه های چند گزینه ای
- جداول مالی ، پرینتهای بانکی و ...
- برگه های جمع آوری اطلاعات مانند فرمهای آمارگیری و جداول آزمایشگاهی
- لیستهای بیمه ، درخواست خرید ، میزان مصرف و ...
- فرمهای ثبت نام ( ثبت نام کنکور ، حساب بانکی و ... )
- صورتحسابهای مالی ( حقوق و دستمزد و ... )
- و ...

این سیستم قادر است اطلاعات فایل های دسته ای تصاویر فرمها در حد چند هزار تایی را به انواع بانکهای اطلاعاتی ( Database ) وارد کند.