

## گزارشکار تشخیص شناسه و مبلغ چک با yolov8l:

در این بخش از پروژه برای تشخیص شناسه چک و مبلغ آن بر روی دیتاست محدودی از تصاویر (28 تصویر) مدل را آموزش دادیم

### مراحل کار :

تسک object detection دوکلاسه ایجاد کردیم ، برای انجام این پروژه ابتدا تصاویر مربوط به چک های صیادی (جدید ) و چک های قدیمی را با استفاده از roboflow ، نشانه گذاری کردیم و خروجی api آن را در گوگل کولب آموزش دادیم

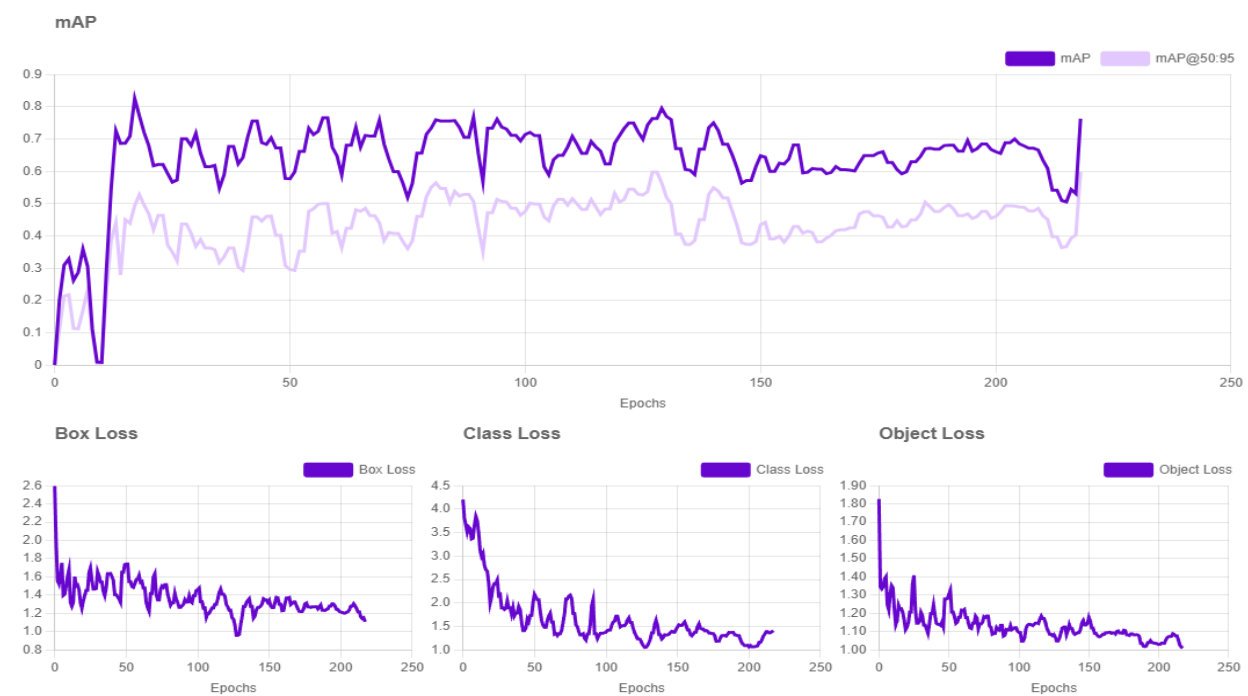
مدل استفاده شده yolov8l با 100 اپیک بود که 18 تصویر برای آموزش مدل و 5 تصویر برای ارزیابی و 5 تصویر برای تست نهایتا استفاده شده

عملکرد مدل ، روند آموزشی آن و نتایج پیش ها در لینک زیر بصورت پابلیک در دسترس است :

<https://drive.google.com/drive/folders/19SQ8xCgXLnzq-PzRL3rGAH8vBUOIGXsh>

بعضی از خروجی های مهم مدل:

Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
100/100	11.1G	0.4329	0.3738	0.8463	9	640:
100% 4/4 [00:02<00:00, 1.49it/s]						
	Class	Images	Instances	Box (P	R	mAP50
mAP50-95): 100% 1/1 [00:00<00:00, 7.45it/s]						
	all	5	12	0.822	0.443	
0.482	0.362					



بر روی تصاویر داده شده به مدل تغییر سایز اعمال نشده است همچنین از روش های زیر برای داده افزایش استفاده کردیم.

90° Rotate  
Upside Down

Grayscale  
Apply to 50% of images

Saturation  
Between -35% and +35%

Bounding Box: Brightness  
Between -35% and +35%

Bounding Box: Noise  
Up to 4.98% of pixels



تاریخ به حروف ستاره ام جعفری ۱۳۸۷/۱۱/۲۰

حساب ۲۱۹۳۲

محمدهادی عباسیان  
جاری ۲۱۹۳۲

به موجب این چک مبلغ هجده میلیون و شصت هزار و سیصد و پنجاه و پنج ریال صادر شد و به حساب جاری نزد بانک ملت واریز گردید.

در وجه حلال

یا به حواله کرد پرداختید

۲۷۵ ۵۲۷۲۱۰

۷,۶۰۰,۰۰۰/۰۰۰ ریال

تاریخ به حروف ستاره دوم ۱۳۸۷/۱۱/۲۰

حساب ۶۳۳/۱

شرکت کاوه مکانیک شیراز

به موجب این چک مبلغ هشتصد و بیست و یک هزار و سیصد و پنجاه و پنج ریال صادر شد و به حساب جاری نزد بانک ملت واریز گردید.

در وجه آقای امینی بهاری

یا به حواله کرد پرداختید

۲۱۸ ۸۲۴۸۶۵

۷,۰۰۰,۰۰۰/۰۰۰ ریال

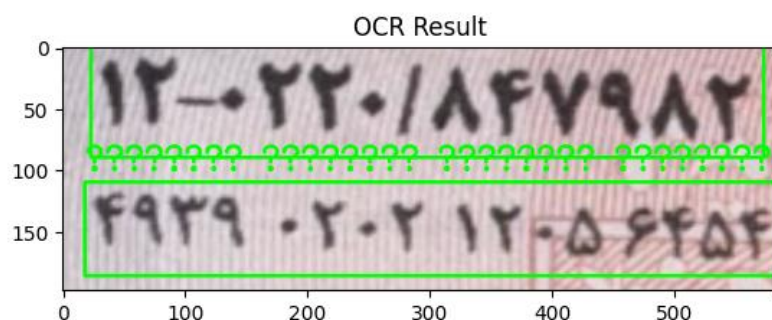
شرکت فنی  
کاوه مکانیک شیراز  
(سهامی خاص)  
KM.CC





## اعمال OCR فارسی بر روی شناسه و مبلغ چک :

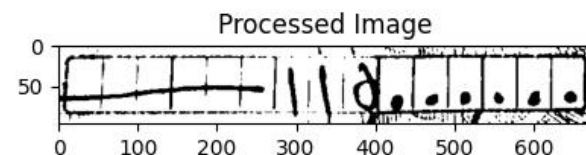
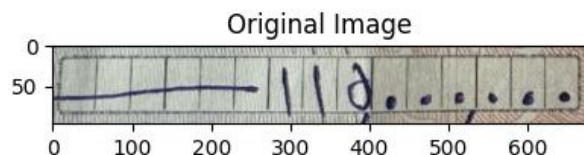
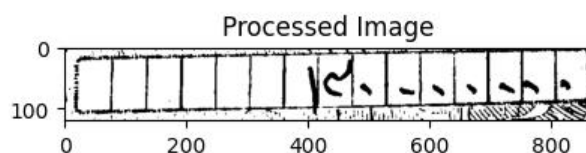
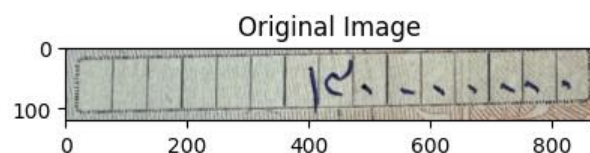
بعد از تشخیص و جدا سازی شناسه و مبلغ چک صیادی نوبت به اعمال OCR برای استخراج متن آنها است، برای تشخیص اعداد دیجیتال شناسه چک صیادی کتابخانه easyocr برای اسکن کامل و بدون خطا شناسه کافی بوده است ( مثال از این مورد در [این کد](#) قرار دارد )



CompleteDetected text: ۱۲-۰۲۲۰/۸۴۷۹۸۲ (Confidence: 0.79)

Detected text: ۴۹۳۹-۰۲۰۲۱۲-۵۶۴۵۴ (Confidence: 0.63)

همانطور که در خروجی مشخص است شناسه چک صیادی بصورت کامل استخراج شده است !



اما easy ocr برای اسکن متن دست نویس مبلغ چک خوب عمل نکرده و هیچ چیز را در تصویر فوق و دو نمونه تست دیگر تشخیص نداده است. راه حل های چک شده برای حل این موضوع تا الان این موارد بوده است :

**بررسی و اعمال مدل های مبتنی بر YOLO برای تشخیص پلاک برای OCR مبلغ چک :** دو مدل که دقت بالایی بر روی تصاویر پلاک داشتند بر روی این بخش از چک اعمال شد که جوابگو نبودند :

<https://universe.roboflow.com/azsdf/persian-plates-number/>

<https://universe.roboflow.com/shahab-jafari-1vorv/persian-plate-characters-mvinj/model/>

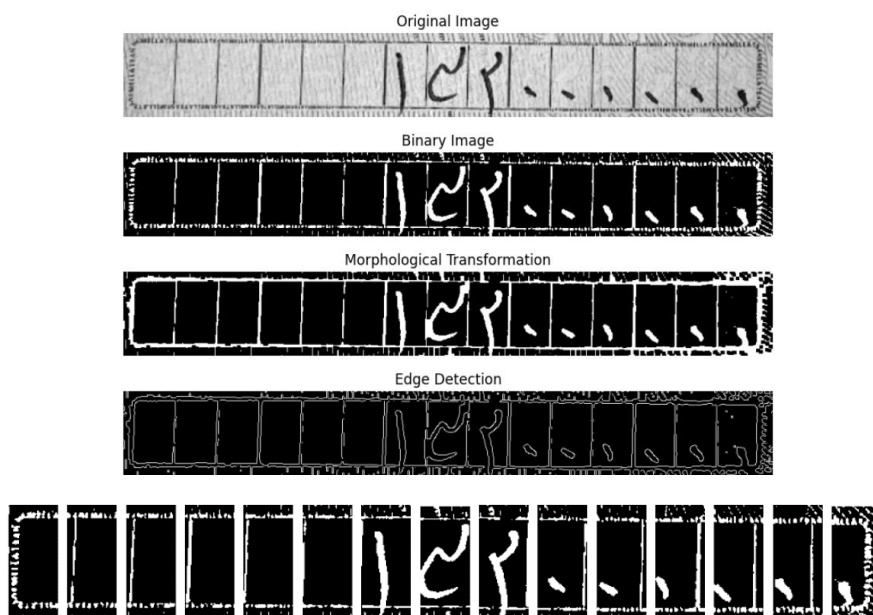
این دو مدل تنها زمانی عدد 0 را در شناسه چک را اسکن میکردند که مقدار confidence را کمتر از 10 % قرار دادم، (بنظرم دلیلش متفاوت بودن شکل نوشتاری صفر در پلاک خودرو با صفر در چک است که با فاین تیون کردن میتوان عملکردش را بهبود بخشید)

بعد از بررسی عملکرد پولو از کتابخانه های مبتنی بر شبکه عصبی برای تشخیص اعداد دستویس فارسی استفاده کردم ، نمونه با عملکرد مناسب ان مورد زیر است :

[https://github.com/aztofighi/Persian\\_Handwritten\\_Recognition/tree/main](https://github.com/aztofighi/Persian_Handwritten_Recognition/tree/main)

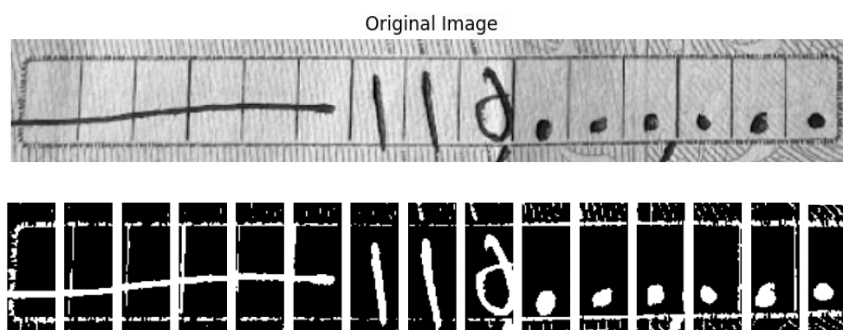
<https://replicate.com/aztofighi/persian-digit-detector> لینک آنلاین برای تست مدل

برای اینکه عملکرد تشخیص بهبود پیدا کند نیاز است تا کاراکتر های چک جدا شوند که در اینجا ابتدا ترشهود بر روی تصویر می زنیم تا تصویر باینری شود و بعد عملگر های مورفولوژیک اعمال می کنیم تا حاشیه های از بین رفته باز سازی شوند و در نهایت با الگوریتم نوشته شده 15 باکس مربوط به جایگاه های اعداد جدا و ذخیره می شوند...



چالش های این بخش: الگوریتم نوشته شده برای جداسازی اعداد با چالش هایی روبرو است :

بعضا کاربر اعداد را خارج از باکس مربوطه می کشد که باعث اشتباه مدل می شود ، همچنین بعضا افراد مابعد یا ماقبل مبلغ خود خطوط هاشوری ایجاد می کنند که به نوعی از تغییر در چک جلوگیری کنند اما این موارد باعث اشتباه الگوریتم نوشته شده برای جداسازی می شود ، در این حالت با حذف عملگر مورفولوژیک که باعث چسبیدن اعداد به حاشیه کادر مبلغ شده بود موضوع حل و فصل شد ...



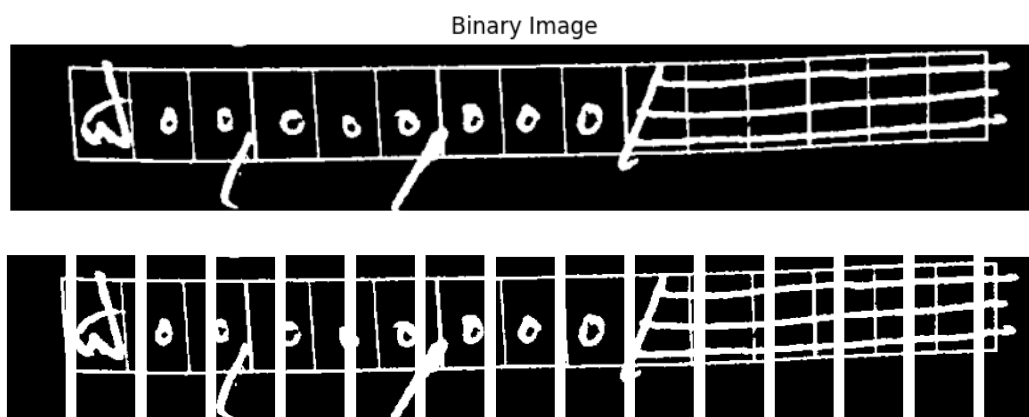
علاوه بر اینها اعداد جدا شده با این روش دارای نقاط اضافه ای هستند که مدل که بالاتر لینک داده شده است را به اشتباه می اندازد بطوری که بعضا اعداد مختلف را متعلق به کلاس 1 تشخیص می دهد !



مثال: مدل ذکر شده این کادر را کلاس 1 تشخیص داده است ، اما تصویر زیر را کلاس 5 تشخیص داده است :



بایاس مدل بالا بخاطر اینورت بودن عکس نیست و این موضع هم چک شده است بخاطر خطوط یا نقاط اضافه ای است که در اسکن تصویر یا جداسازی تصویر باقی مانده است ...



چالش دیگه برش اشتباه برای سلول اول است اگه فضای اضافی در تصویر وجود داشته باشد سلول اول به اشتباه انتخاب می شود و این خطا در اشتباه سلول اول در انتخاب سلول های بعدی نیز منتشر شده و تاثیر بد می گذارد ! هر چند این مدل های CNN به دقت های خیلی بالا رسیده اند اما برای کار با تصاویر نویزی مانند این مناسب نیستند و احتمالا overfit شده اند.

لیست چند مخزن دیگه برای تست آنها :

<https://github.com/Erfaniaa/persian-digits-recognition>

<https://github.com/msnmkh/Persian-Handwritten-Digits-Image-Recognition>

[https://github.com/sadafnazari/Persian\\_handwriting\\_recognition](https://github.com/sadafnazari/Persian_handwriting_recognition)

<https://github.com/MiladNooraei/Persian-HandWritten-Digits-Recognition>

<https://github.com/eslamisepehr/Persian-Handwritten-Digit-Recognition-CNN>

<https://github.com/mhsharifi96/Hand-Digits-Recognition>