**گزارش مرحله به مرحله عملکرد**

: Step 1تهیه چند عکس به عنوان ورودی 🡪 complete mouse folder in pc /

: Step 2 uplod عکس ها که در content بارگذاری میشوند.

: Step 3 ایجاد پوشه all\_input و جابجایی عکس ها در آن پوشه

: Step 4 فکر کردن به چگونگی تهیه بک گراند

میتوانستم یک سری عکس ثابت تهیه کنم و از آنها استفاده کنم اما

* چالشی که در ذهنم ایجاد شد این بود که ممکنه عکس object و بک گراند بسیار از هم متفاوت باشه و این شرایط را برای اموزش بسیار راحت میکنه پس برای اینکه شرایط پیچیده تر شود از یک طیف رنگی نزدیک به تصویر به عنوان بک گراند استفاده کنم.
* میتوانستم شبکه ای train کنم با کمک cnn که وقتی تصویری به عنوان ورودی داده شد رنگ آن را predict کنه و سپس بر اساس رنگ predict شده بک گراند بسازه که این کار اصلااااا بهینه نیس چون این کار را با کمک خواندن مقادیر RGB هر پیکسل و سپس میانگین گیری انها انجام میدهم تا به رنگی نزدیک برسم .

: Step 5 سرچ کردن درباره طیف رنگی تصاویر تا اسم آن را پیدا کنم : Result: Gradient



: Step 6تهیه تصاویر با طیف رنگی با کمک پایتون و open cv search:

<https://stackoverflow.com/questions/25668828/how-to-create-colour-gradient-in-python>

<https://note.nkmk.me/en/python-numpy-generate-gradation-image/>

* **در حین یافتن اموزشی برای تهیه چنین تصاویری با ریپازیتوری زیر اشنا شدم و از آن استفاده کردم**

<https://github.com/HYOUG/color-gradient-generator>

: Step 7 میانگین گرفتن از هر پیکسل در هر تصویر در پوشه ورودی

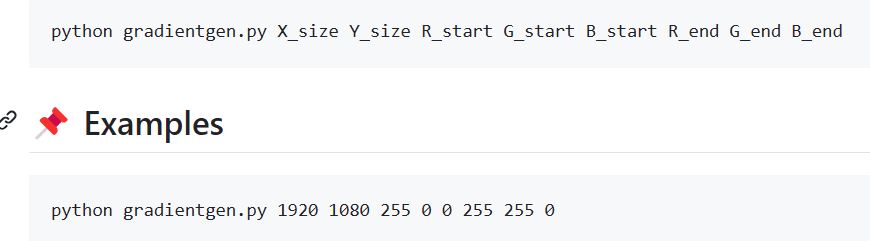
how get the average color by using the following:

<https://stackoverflow.com/questions/43111029/how-to-find-the-average-colour-of-an-image-in-python-with-opencv>

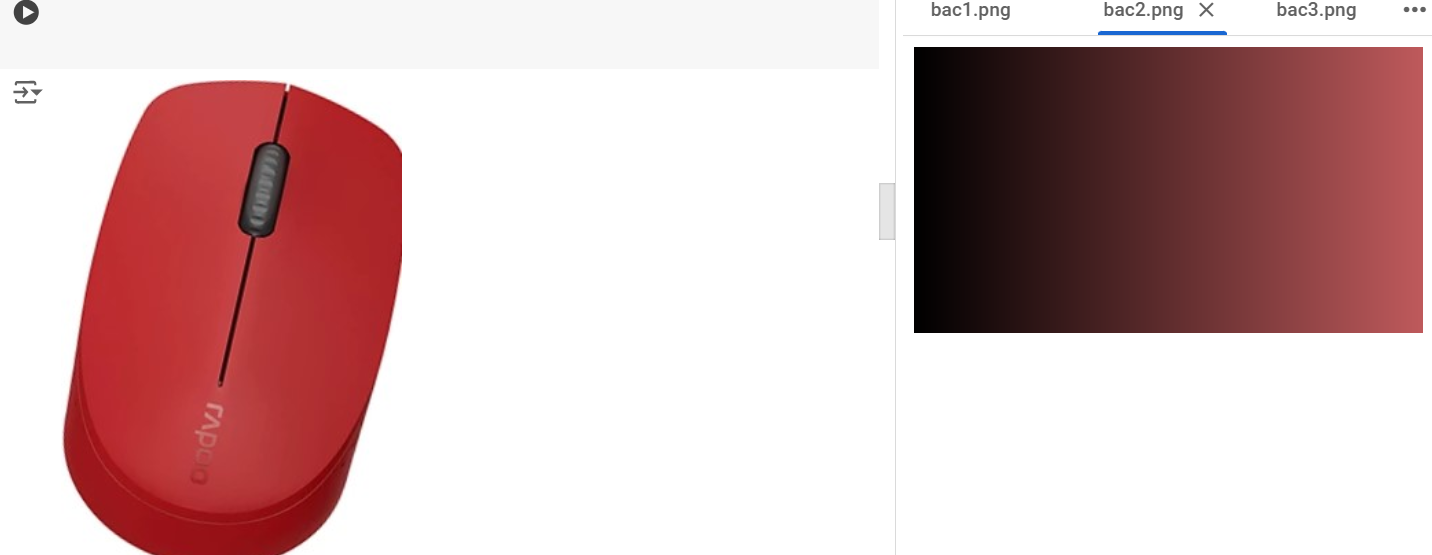


: Step 8 تهیه بک گراند از رنگ های فیکس به میانگین های رنگی

* پس از اجرای کد زیر یک تصویر در content جنریت میشه در یک فانکشن کدی مینویسم که از هر میانگین رنگی عکسی یک طیف از رنگ سفید به رنگ مد نظر جنریت میکنه و برای آن یک نام اختصاص میده و در پوشه background تصویر را انتقال میده   
  سپس بار دیگر از همان میانگین رنگی به رنگ مشکی و از رنگ مد نظر به زرد و قرمز و آبی هم این کار را انجام میدهیم ، پس برای هر تصویر 5 بک گراند تهیه میکنیم .



برای مثال عکس بک گراند برای تصویر زیر، از مشکی به میانگین رنگی ، تصویر زیر را تولید می کند



: Step 9به این مرحله رسیدم که حال چگونه تصاویر را بر رویbackground ها بزارم ؟

* گفتم ابتدا عکس های ورودی را agumentation روش اجرا کنم تا object ها زیاد بشه سپس بر روی بک گراندا قرار بدم.

چون اگه قبلش از قرار دادن عکس بر رویbackground ها بزارم بعد agumentation اجرا کنم دیگه نمیتونم براش label تعیین کنم پیدا کردن مختصات ها سخت میشود و بهینه نیس.

: Step 10 سرچ درباره : agumentation میخواستم که عکس ها را با خود لایه های کراس agument کنم که گفتم سرچ کنم ببینم yolo خودش این کار را انجام میده یا نه ؟ که فهمیدم yolo در تنظیمات yolo انگار پارامتر هایی هنگام training وجود داره که دیتا رو agument میکنه. پس از متوجه شدن این که yolo این کار را میکنه به تهیه دیتاست برای yolo میپردازم.

: Step 11 در این مرحله به تهیه لیبل ها میپردازیم » برای این که لیبل ها را تهیه کنیم نیاز است که مختصات object را بدانیم پس یعنی عکس object باید بر روی background قرار گرفته شده باشد .

: Step 12 میخواستم بدانم عکسbackground که ورودیyolo است چه سایزی باید باشه که با سرچ در سایت

<https://docs.ultralytics.com/modes/train/>

به جواب زیر رسیدم



: Step 13 برای این که چگونه تصویر شی را بر روی بک گراند قرار دهم به سایت های زیر مراجعه کردم و در نهایت از سایتی که با رنگ زرد مشخص شده است استفاده کردم. بلند و فید کردن عکس object بر روی background تولید شده ، Adding Objects to Image in Python with blend؟

Pasting, Merging & Blending Images with Python (PIL):

<https://holypython.com/python-pil-tutorial/pasting-merging-blending-images-with-python-pil/>

Adding Objects to Image in Python using mask of objects:

<https://medium.com/@alexppppp/adding-objects-to-image-in-python-133f165b9a01>

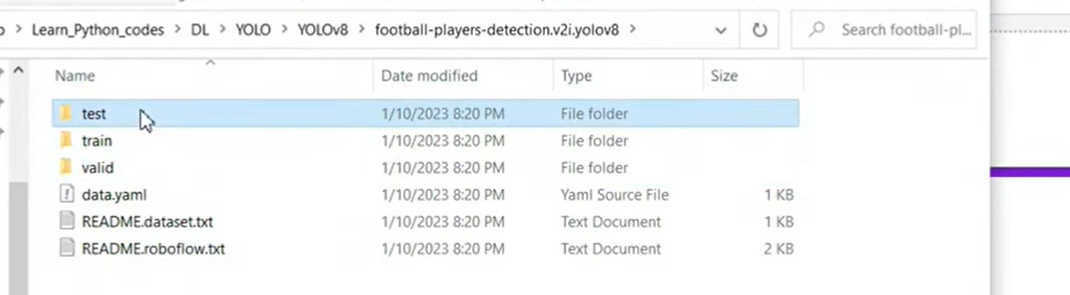
blend-overlapping-images-in-python:

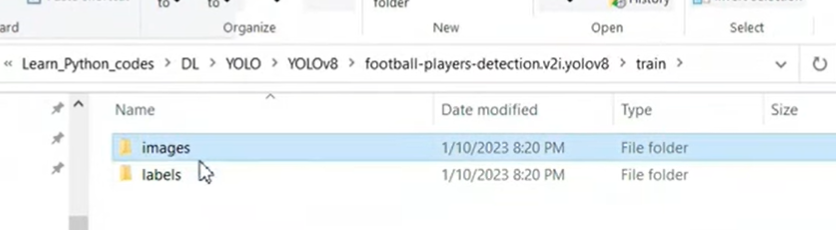
<https://stackoverflow.com/questions/29106702/blend-overlapping-images-in-python>

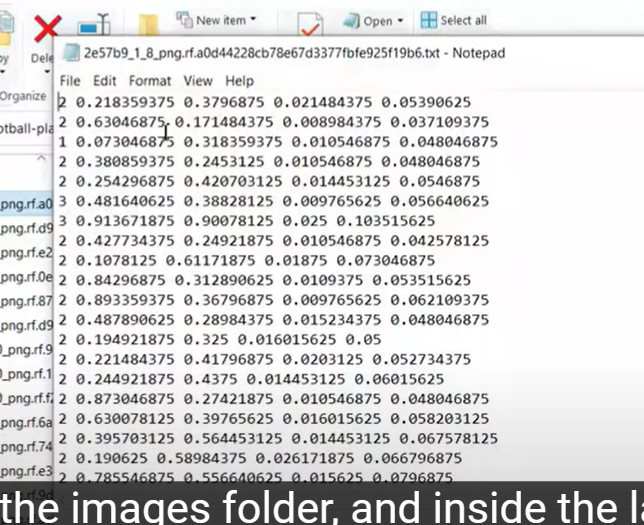
[Blending images using OpenCV](https://medium.com/@maximinusjoshus?source=post_page-----bfc9ab3697b7--------------------------------):

<https://medium.com/featurepreneur/blending-images-using-opencv-bfc9ab3697b7>

: Step 14 برای یاداوری این که پوشه ها و ساختار دیتا برای اموزش yolo چگونه باشد به یک پروژه که با v8 پیاده سازی شده بود مراجعه کردم :

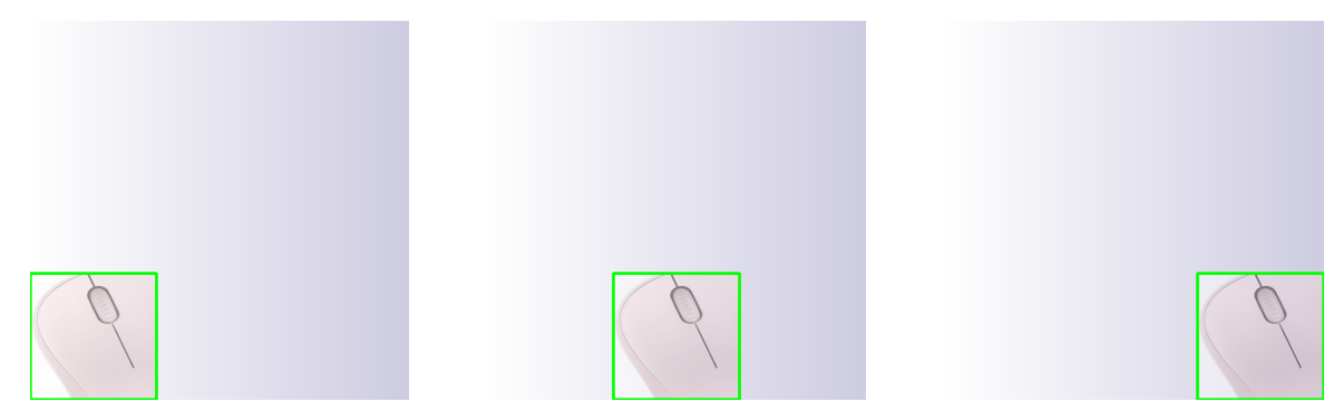
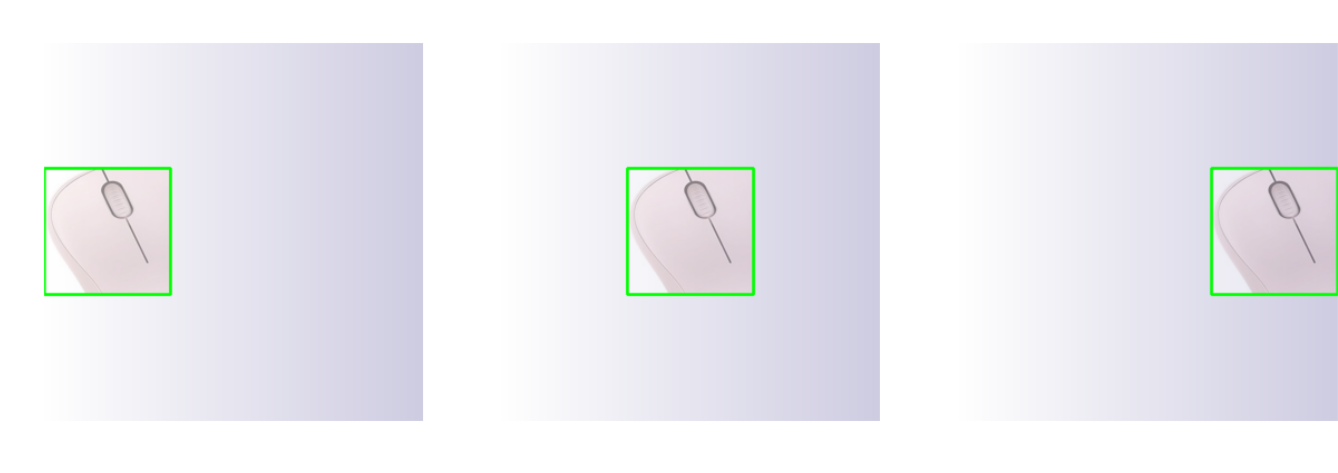
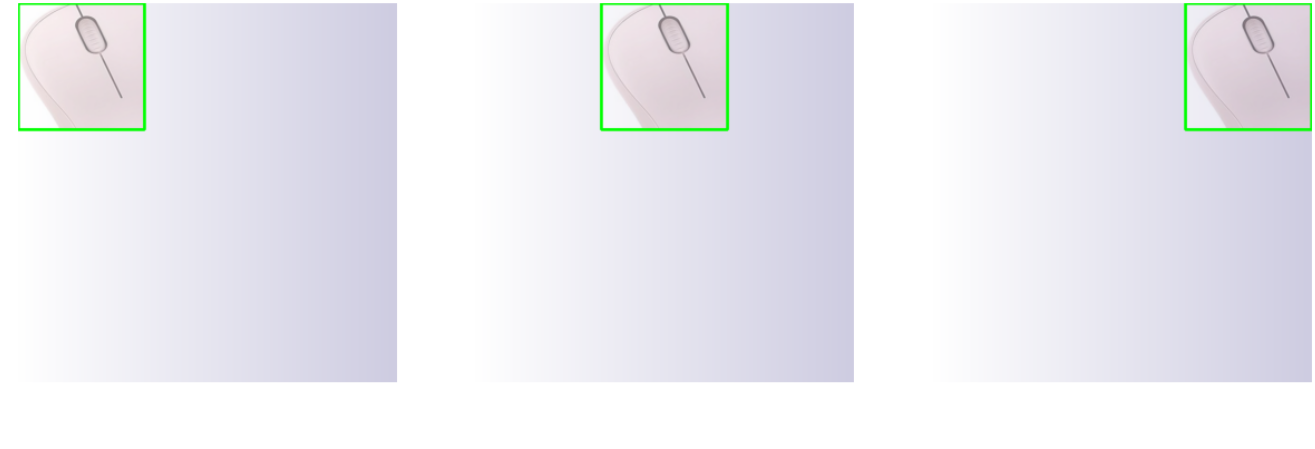


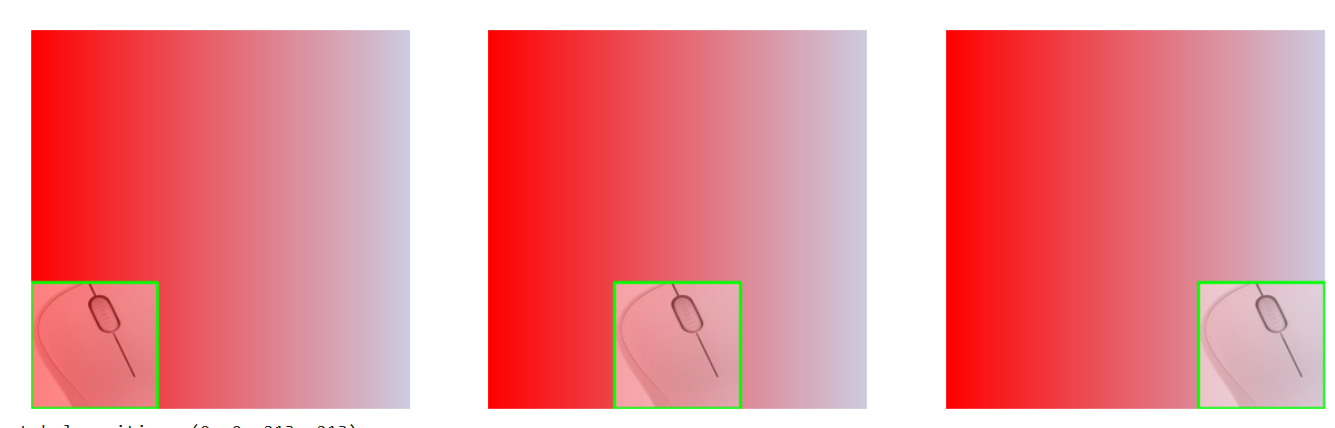
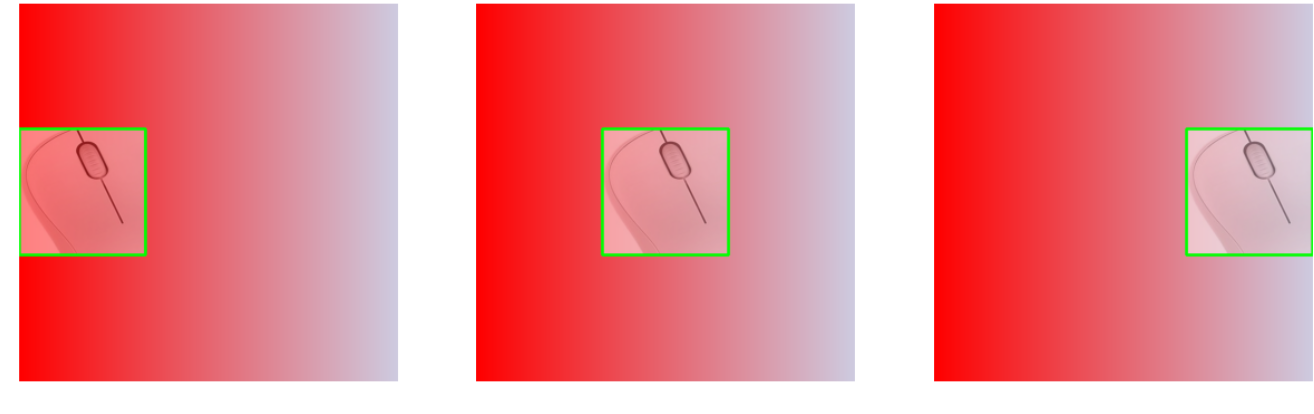
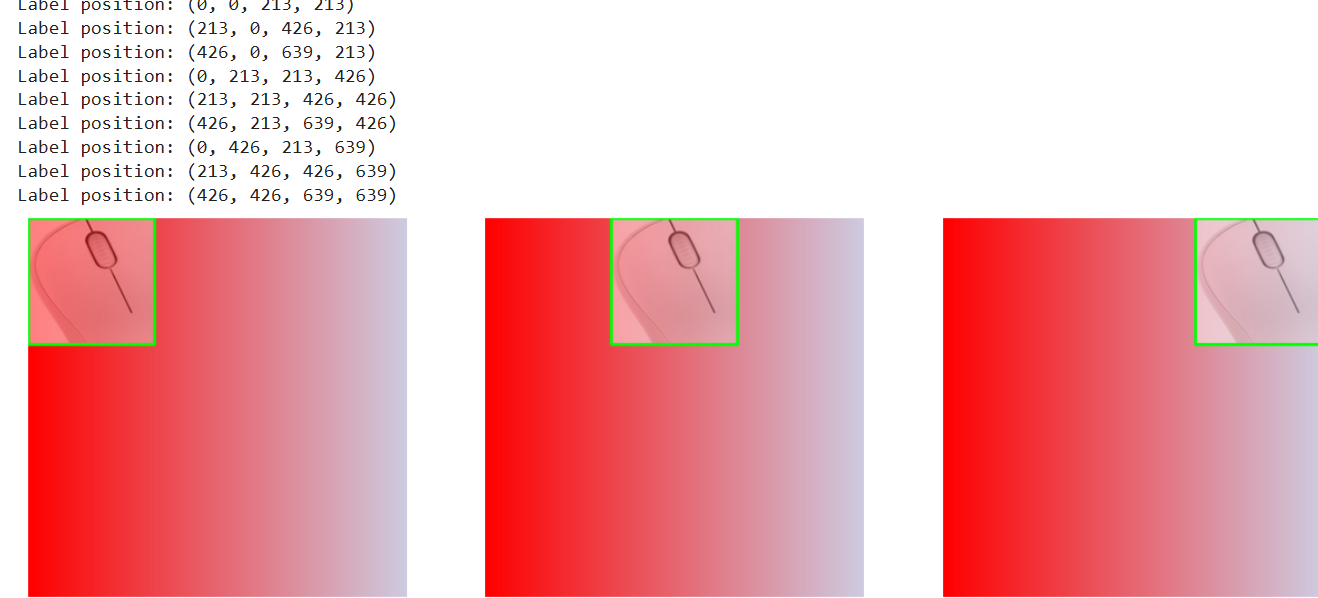
به ازای هر عکس یک فایل labels داریم هر labels یک فایل txt می باشد و هم نام با نام اسم عکس است و فایل تکست label برای یک image به صورت زیر می باشد که عدد اول شماره کلاس و چهار عدد بعدی مختصات object می باشد .



کلاس و مختصات bounding box

: Step 15 لازم به یاد اوری است که ما برای هر object 5 بک گراند ایجاد کردیم . الان تصویرمون که 640\*640 است را به 9 قسمت تقسیم می کنیم یعنی عکس object باید به ابعاد 213\*213 کاهش پیدا کند و هر بار در یکی از 9 مختصات قرار بگیرد . برای مثال عکس زیر نتیجه عملکردم را نشان میدهد:   
**نکته:** اگر در عکس زیر نگاه کنید متوجه خواهید شد که با خط سبز مختصات را با کمک opencv به عنوان bounding box نشون دادم ولی بعد متوجه شدم این کار درست نیس و در این حال نیازی نیس و این کار یادگیری را به شدت اسان میکند پس مدل ما خیلی ضعیف میشود و اصلا برای دیتای اموزش این کار صحیح نیس ، پس من این را حذف کردم لذا صرفا به عکس موجود برای چگونگی blend و add کردن و مختصات توجه کنید.





نکته : من فکر میکنم بهتر است این 9 تا عکس جای ثابت نباشه چون در این صورت شبکه به این 9 تا مختصات عادت میکنه بهتره کدی بنویسم که ب ازای یک عکس و یک بک گراند مثلا 9 تا مختصات رندوم ایجاد کنه و هر بار این تصاویر در جای جدید قرار بگیره ، فعلا قصد دارم به همین صورت با همیت دیتا شبکه را اموزش دهم سپس در انتها این کار را در یک فاز دیگر انجام دهم.

: Step 16تصاویر را به پوشه images انتقال دادیم و برای هر عکس label ان را با نام خود عکس تولید کردم و در پوشه labels قرار دادم

نکته: من فرض کردم یعنی عکس ها را طوری تولید کردم که ورودی اشیایی از یک نوع است، یعنی مثلا ورودی همه موس است و در این حالت predict میکنه به همین خاطر در هنگام ساخت labels اولین ورودی را در همه برچسب ها به طور دستی یعنی بدون اموزش شبکه ای که تشخیص دهد از چه نوعی است 1 قرار دادم . یعنی اگر من 2 تصویر نصغه از موس و 2 تصویر نصفه از هویج قرار بدم لیبل جفتشون یکیه و اگر بخوام فرق کنه نیاز داره شبکه ای باشه که روی همه تصاویر مثل imagenet اموزش دیده باشه و نمیدونم کارایی و عملکرد در این صورت به چه صورته ، نمیدانم برنامه نیاز دارد یا نه ؟ به هر حال برای توسعه پروژه میتوان این اقدام را در نظر گرفت.

پیدا کردن یک پروژه با yolo v10 : train-yolov10-object-detection-on-custom-dataset.ipynb

<https://colab.research.google.com/github/roboflow-ai/notebooks/blob/main/notebooks/train-yolov10-object-detection-on-custom-dataset.ipynb>

: Step 17 پوشه بندی فولدرهای دیتاست yolo v8 را درست میکنیم که به صورت زیر است :

data

Train

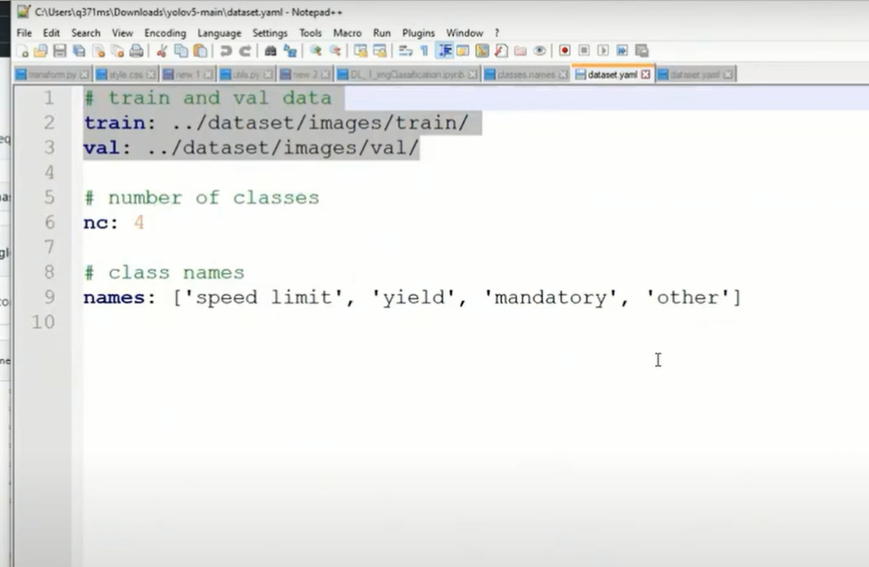
Images , labels

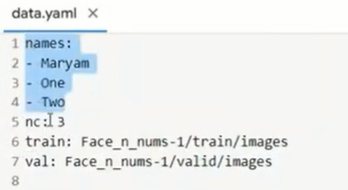
Vali

Images , labels

Dataset.yaml

: Step 18 با کمک و یاداوری از پروژه ها که فایل yaml انها به صورت زیر است yaml را ایجاد میکنیم:



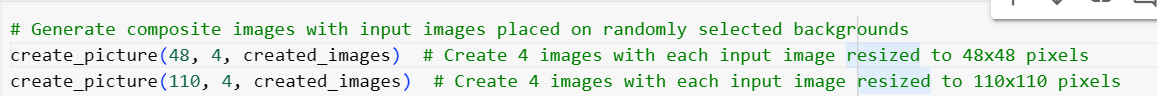


: Step 19 20 درصد از دیتاهایی که تهیه کردیم و همه را در پوشه های images , labels عه train جدا کردیم را به validation انتقال میدهیم .

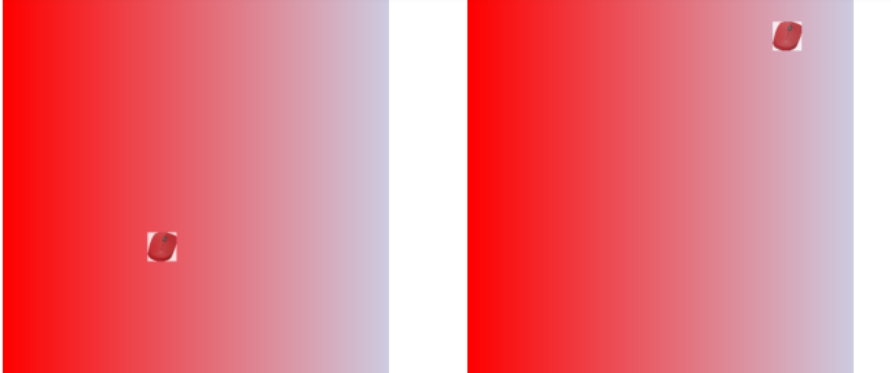
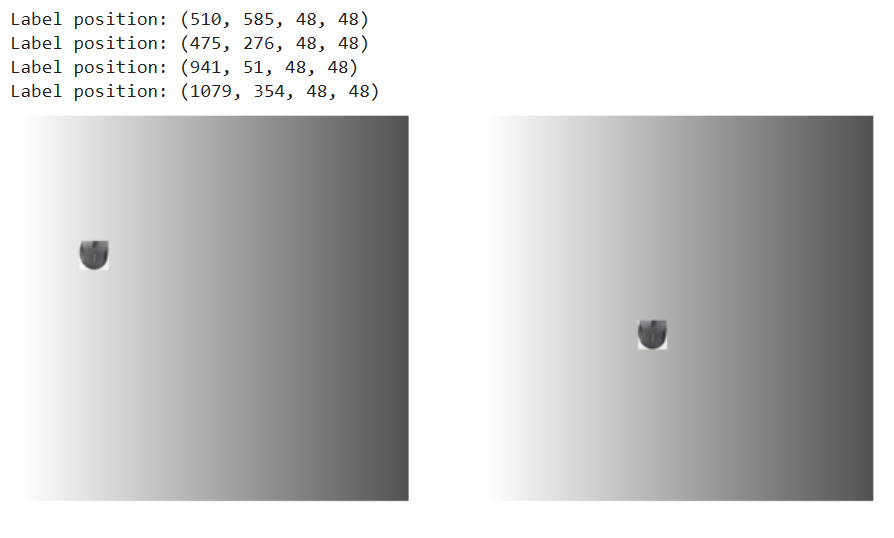
: Step 20 باید کدهای نوشته شده را طوری تغییر دهم که که این 9 جای ثابتی که انتخاب کردم صرفا عکس ها قرار نگیرند و به صورت رندوم و تصادفی بر روی بکگراند قرار بگیرد و این که اصلا 9 تا نباشه بتونم به تعدادی که مایلم انتخاب کنم و سایزشم 213 در 213 نباشه چون شبکه را به یک سایز ثابت عادت میدهم پس برای هر تصویر ورودی این موارد را در نظر میگیرم و تابع

def create\_picture(img\_size, num\_pic, created\_images):

را کد نویسی میکنم :

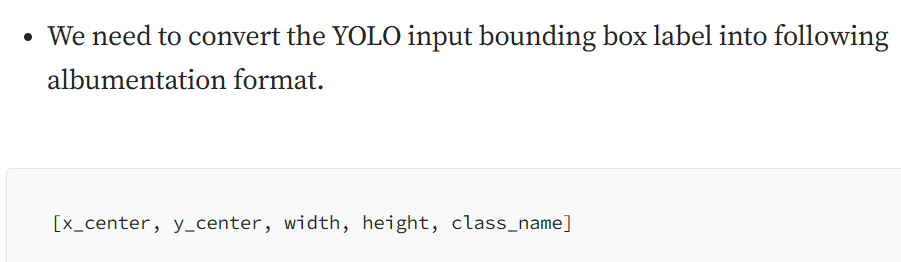


چند نمونه خروجی:



: Step 21 متوجه شدم که من coordinate را به صورت لیبل به yolo دادم و اموزش مشکل دارد با کمک سایت زیر که برای اگیومنیشن کمک کننده است فرمت صحیح label ها را برای YOLO پیدا کردم و تمامی label ها را به صورت استاندارد اصلاح کردم همچنین label ها باید به صورت نرمالیزیشن شده نیز باشند

<https://medium.com/red-buffer/apply-data-augmentation-on-yolov5-yolov8-dataset-958e89d4bc5d>



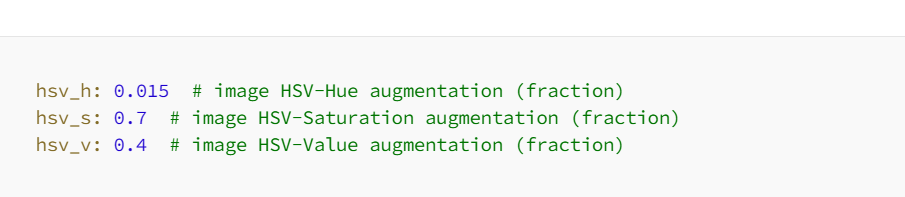
: Step 22 سایت های مفید و خوبی که برای YOLO Data Augmentation بررسی کردم :

از سایت خود ultralytics هایپر پارامترهای مختلف را مشاهده کردم

<https://docs.ultralytics.com/usage/cfg/#augmentation-settings>

و

<https://rumn.medium.com/yolo-data-augmentation-explained-turbocharge-your-object-detection-model-94c33278303a>

مثلا 

که من هم در کدم اعمال کردم و با کمک سایت بالا تمام هایپرپارامترهای اگیومنتیش های yolo را متوجه شدم.

و

<https://github.com/ultralytics/ultralytics/issues/8308>

<https://github.com/ultralytics/yolov5/issues/8021>

working with yolo

<https://docs.ultralytics.com/modes/train/#usage-examples>

<https://docs.ultralytics.com/tasks/detect/>

<https://docs.ultralytics.com/modes/#modes-at-a-glance>

<https://github.com/ultralytics/ultralytics>

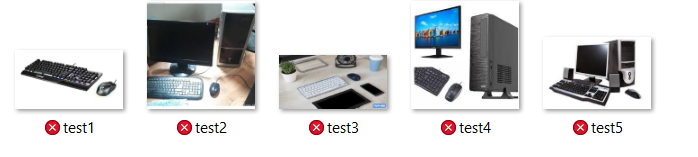
<https://docs.ultralytics.com/models/yolov8/#supported-tasks-and-modes>

yolo v10

<https://docs.ultralytics.com/models/yolov10/>

<https://docs.ultralytics.com/models/yolov10/#usage-examples>

: Step 23بعد از اموزش مدل با yolov 8 به پیش بینی تصاویر زیر میپردازیم :



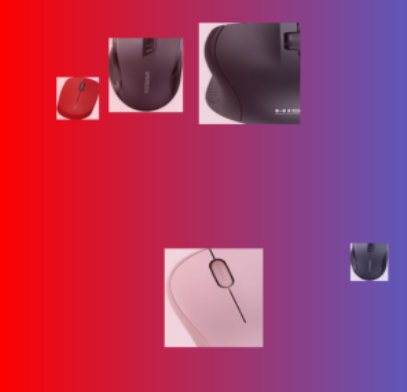
: Step 24تعداد تصاویر و سایز و پارامترها را بررسی میکنم تا نتیجه گیری کنم :

* تعداد تصاویر سایز کوچک را بیشتر کنم تعداد تصاویر 48 در 48 وقتی بیشتره نتیجه بهتره
* وجود بک گراند غیر رنگی پر چالش ممکنه مفید باشه
* وقتی تصاویر با سایز کوچک کمه اصلا هیچی در تصویر نشون نمیده
* وقتی تصاویر سایز بزرگ زیاده مانیتور و کیس هم موس تشخیص میده
* در هر عکسی تنها یک object هس باید تعداد object ها بیشتر باشه
* تمام عکس ها صاف هستند هیچ گونه چرخشی یا حالتی ندارند بهتر است روی تصاویر ورودی با کمک open cv اگیومنتیشن اعمال کنم.
* تعداد تصاویرو بالا ببرم
* وقتی تصاویر کوچک زیادم بشه حتی مارک موس هم یه ابجکت در نظر میگیره پس اعتدال
* بررسی با yolo8 s , yolo8 n

در هر عکسی تنها یک object هس باید تعداد object ها بیشتر باشه :

: Step 24 فانشکنی درست میکنم که برای هر بک گراند هر بار از تصاویر ورودی به طور رندوم حتی با قابلیت تکرار عددی بین 2 تا 5 انتخاب میکنه و تصاویر جدید را تولید میکنه

خروجی چند نمونه:



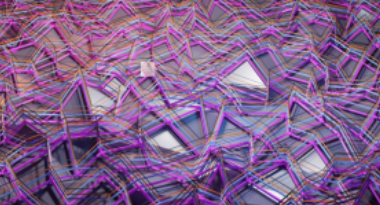
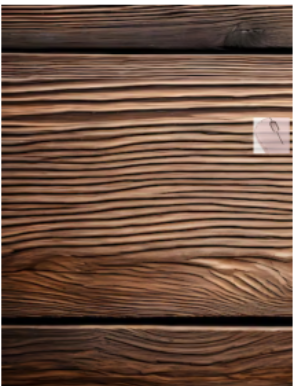
وجود بک گراند غیر رنگی پر چالش ممکنه مفید باشه : مارک دار – مربع های پنهان - اشکال رنگی مشابه موس و ....

: Step 25 10 تا تصویر برای بک گراند دانلود کردم و انها را در پوشه بک گراند گذاشتم تا بک گراندهای چالشی هم داشته باشیم و این عکس ها برای هر برنامه ای ثابت در نظر بگیریم  
سوال؟

* میتونیم استفاده کنیم؟
* بر روی تعداد محدودیت داریم؟

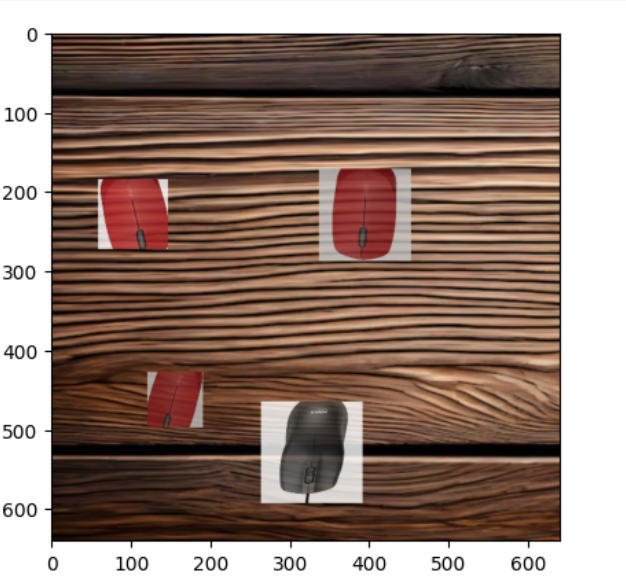
خروجی چند نمونه که هم بر روی تصاویر تکی هم چندتایی اعمال شد



: Step 26اعمال تغییرات تصادفی به تصاویر ورودی با استفاده از کتابخانه Augmentor، تغییرات تصادفی به تصاویر ورودی اعمال می‌شود تا تنوع بیشتری در مجموعه داده‌ها ایجاد شود.

این تغییرات شامل چرخش، وارونه‌سازی افقی و عمودی، تغییر کنتراست، روشنایی، زوم و برش تصادفی می‌شود. ابتدا یک (Pipeline) برای اعمال تغییرات تصادفی به تصاویر ورودی ایجاد می‌شود. سپس تغییرات مختلف با احتمالات معین به آن اضافه می‌شوند. چند تصویر تغییر داده شده را در بر روی بک گراند زیر میبینیم:



: Step 27 جلسه با دکتر محمدرضا محمدی برای ارائه گزارش

بعد از جلسه اقدامات زیر انجام شد »

* تصاویر ورودی و بکگراند و تست را در گوگل درایو بارگذاری کرده و آن ها را از طریق لینک دانلود میکنیم.
* کد مخصوص عکس گرفتن با وب کم را اعمال کردم.
* Yolo v8s را اعمال کردم و نتایج مانند شبیه yolo v8n شد ولی از انجایی که مدت زمان هر ایپاک زیاد است ران تایم GPU پاسخ گو نیست .

و

: Step 28

به جای addweighted میتوان از

### **[Seamless Cloning using OpenCV](https://learnopencv.com/seamless-cloning-using-opencv-python-cpp/" \t "_blank)**

<https://amroamroamro.github.io/mexopencv/opencv/cloning_demo.html>

<https://www.scaler.com/topics/seamless-cloning-in-image-processing/>

<https://learnopencv.com/seamless-cloning-using-opencv-python-cpp/>

<https://docs.opencv.org/4.x/df/da0/group__photo__clone.html>

استفاده کرد. ولی برای استفاده از seamless clone نیاز داریم ماسک هر تصویر را داشته باشیم»

.1 استفاده از تکنیک‌های آستانه‌گذاری (Thresholding):

با استفاده از آستانه‌گذاری می‌توان ناحیه شیء را از پس‌زمینه جدا کرد. این روش برای تصاویر با کنتراست بالا بین شیء و پس‌زمینه مناسب است. که من نمیتوانم بدلیل همخوانی تصاویرم با بک گراند های رنگی از این تکنیک استفاده کنم.

.2 استفاده از روش‌های تقسیم‌بندی مبتنی بر رنگ (Color-Based Segmentation):

اگر شیء رنگ مشخصی داشته باشد، می‌توان از تقسیم‌بندی مبتنی بر رنگ استفاده کرد. که این مورد هم از انجایی که رنگ ورودی مشخص نیست، امکان پذیر نیست.

3.استفاده از الگوریتم‌های تقسیم‌بندی پیشرفته‌تر:

استفاده از الگوریتم‌هایی مانند GrabCut برای جدا کردن شیء از پس‌زمینه. این روش نیازمند مقداردهی اولیه دستی ناحیه مستطیلی حول شیء است. این کار به ظاهر از انجایی که ما مختصات مکان object را در background داریم امکان پذیر است اما من این روش را اعمال کردم و نتایج تولید شده از object با mask تولید شده هیچ مطابقتی نداشت و اصلا مورد قبول نبود.

برای ارتقای پروژه و افزودن قابلیت‌های جدید :

1. استفاده از ImageNet و تصاویر دیگر به عنوان بک‌گراند :

- استفاده از ImageNet برای بررسی تصاویر و استخراج بک‌گراندهای مناسب برای ترکیب با تصاویر ورودی بعد از Predict تصویر ورودی.

- پیاده‌سازی یک سیستم خودکار برای انتخاب بک‌گراندهای مناسب بر اساس دسته‌بندی‌های موجود در ImageNet.

2 . استفاده از image captioning برای تولید تصاویر بک گراند

3. پیاده‌سازی YOLOv10:

- تست و اعمال YOLOv10 به جای YOLOv8 به منظور ارزیابی بهتر عملکرد و دقت شبکه.

- امکانات جدید و بهبودهای احتمالی در معماری YOLOv10 مانند دقت بیشتر در تشخیص اشیاء، سرعت بالاتر در پردازش و آموزش شبکه.

4. بلند کردن تصاویر با استفاده از Seamless Cloning:

- پیاده‌سازی روش‌های پیشرفته‌تر برای بلند کردن تصاویر با استفاده از Seamless Cloning در OpenCV.

- بهبود دقت و کیفیت تصاویر حاصل از این روش با بهینه‌سازی فرایندها و متدهای استفاده شده.