

Assignment NO.9 Solutions

Digital Image Processing | Fall 1400 | Dr.Mohammadi

Teacher Assistant(s) : Fatemeh Anvari – Fatemeh Hajizadeh

Student name : **Amin Fathi**

Student id : **400722102**

**Problem 1**

R-CNN از آنجا که با 2000 پروپوزال در شبکه cnn کار می کند بنابراین برای هر تصویر باید 2000 پروپوزال استخراج شود و تحویل شبکه کانولوشنی شود که سرعت بسیار پایینی دارد Fast R-CNN نسخه سرعت یافته قبلی است و به جای دادن خروجی پروپوزال ها به شبکه کانولوشنی ، خود تصویر را به cnn می دهد و در نتیجه خبری از 2000 هزار پروپوزالی که سرعت را کند کرده بودند نیست که البته این دلیل دقت عملکرد این شبکه را پایین می اورد . در این مدل بعد از مرحله کانولوشن پروپوزال ها را تشخیص داده و به fc پاس میدهیم . دو الگوريتم R-CNNو Fast R-CNNبراي پيدا كردن نواحي پروپوزال از الگوريتم selective searchاستفاده ميكنند كه الگوريتم زمانبري است بنابر این طراحان الگوريتم Faster R-CNNتصميم گرفتند كه اين الگوريتم را از پروسه تشخيص objectحذف كنند و به شبكه اجازه دهند كه به طور خودكار نواحي پروپوزال را بيابد. در اين الگوريتم مانند R-CNNتصوير به يك شبكه كانولوشني داده ميشود تا feature map ها استخراج شوند اما پس از استخراج feature mapها آنها را به selective searchپاس نميدهد و به جاي اينكار يك شبكه كانولوشني جداگانه طراحي كردند تا نواحي پروپوزال را پيشبيني كند و پس از reshape كردن خروجي ها آنها را به يك الگوريتم کلاسیفیکیشن پاس مي دهند. Yolo برخلاف دو الگوریتم قبلی که دو مرحله ای بودند و فقط قسمتی از تصویر را در نظر می گرفتند است .

در واقع الگوريتم YOLOتصوير را به يك شبكه S\*Sتبديل ميكند كه در هر خانه از اين شبكه تعداد mمحدوده وجود دارد. براي هر محدوده الگوريتم probabilityو offsetتعيين ميكند و جعبه هايي كه شانسي بيشتر از threshold مشخص شده داشته باشند به عنوان objectتشخيص داده ميشوند. الگوريتم YOLOچندين برابر از ساير الگوريتم ها سريع تر است ولي ضعف آن در كاركردن با اشياء كوچك داخل تصوير است .

Ssd در واقع تصویر را به چند بلوک تقسیم می کندو چک میکند هر بلوک که ایا در خودش شی ای دارد یا نه که اگه نباشد حکم background دارد

Retina net يك مدل تشخيص اشياء تك مرحله اي است كه از lossاي به نام focal lossبراي بررسي عدم تعادل در كلاس ها در حين training استفاده مي كند

منابع :

[Object Detection Explained: R-CNN | by Chingis Oinar | Towards Data Science](https://towardsdatascience.com/object-detection-explained-r-cnn-a6c813937a76)

[R-CNN, Fast R-CNN, Faster R-CNN, YOLO — Object Detection Algorithms | by Rohith Gandhi | Towards Data Science](https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e)

[RetinaNet Model for object detection explanation - TowardsMachineLearning](https://towardsmachinelearning.org/retinanet-model-for-object-detection-explanation/)

[How single-shot detector (SSD) works? | ArcGIS Developer](https://developers.arcgis.com/python/guide/how-ssd-works/)

[RetinaNet Explained | Papers With Code](https://paperswithcode.com/method/retinanet#:~:text=RetinaNet%20is%20a%20one%2Dstage,learning%20on%20hard%20negative%20examples.)

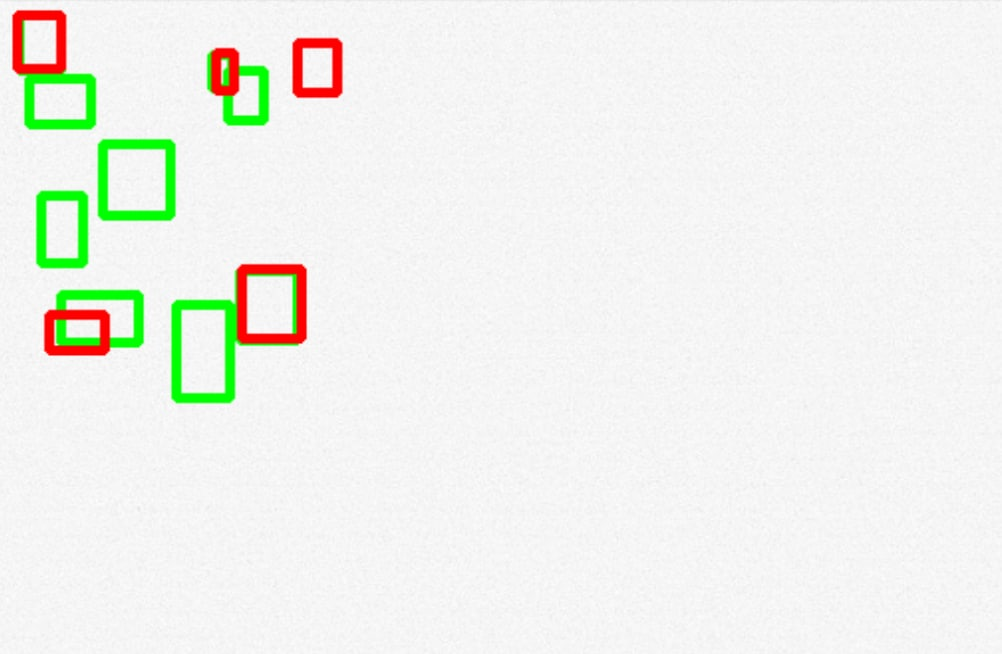
**Problem 2**

در این مقاله برای تولید object bounding box proposal ها از روش جدید به نام edges استفاده شده است . استفاده از edgeها مزایای محاسباتی زیادی را ارائه می‌دهد، زیرا ممکن است به طور موثر محاسبه شوند و نقشه‌های edge به‌دست‌آمده sparse هستند. در این مقاله، نحوه شناسایی مستقیم object proposal ها از نقشه های edge را بررسی می کنند. کار اصلی آن ها مشاهدات زیر است: تعداد خطوطی که به طور کامل توسط یک box محصور شده اند، نشان دهنده احتمال وجود یک شی در box است. اگر تمام پیکسل های edge متعلق به contour در داخل box قرار داشته باشند، می گویند یک contour به طور کامل توسط یک box محصور شده است. edge‌ها معمولاً با مرزهای شی مطابقت دارند و به همین دلیل جعبه‌هایی که مجموعه‌ای از لبه‌ها را قویاٌ محصور می‌کنند احتمالاً حاوی یک شی هستند. با این حال، برخی از edge‌هایی که در box مرزی یک شی قرار دارند، ممکن است بخشی از شی موجود نباشند. به طور خاص، پیکسل‌های لبه‌ای که متعلق به خطوطی هستند که مرزهای box را در بر می‌گیرند، احتمالاً با اشیاء یا ساختارهایی مطابقت دارند که در خارج از box قرار دارند.

بنابر نتایج این مقاله امتیاز دادن به یک box بر اساس تعداد خطوطی که به طور کامل در بر می‌گیرد، یک معیارproposal موثر ایجاد می‌کند. در مقابل، شمارش تعداد پیکسل های edge در box چندان موثر نیست. رویکرد این مقاله تا حدی شباهت هایی به معیارهای superpixels straddling دارد. با این حال، به جای اندازه گیری تعداد خطوط straddling ، چنین خطوطی را در نظر نمی گیرند. از آنجایی که تعداد کادرهای محدود کننده ممکن در یک تصویر زیاد است، باید بتوان به طور موثر به نامزدها امتیاز داد. آن ها از Structured Edge detector استفاده می کنند تا نقشه edge اولیه را بدست آورند. برای کمک به محاسبات بعدی، پیکسل های edge همسایه با جهت گیری مشابه در کنار هم قرار می گیرند تا گروه ها را تشکیل دهند.

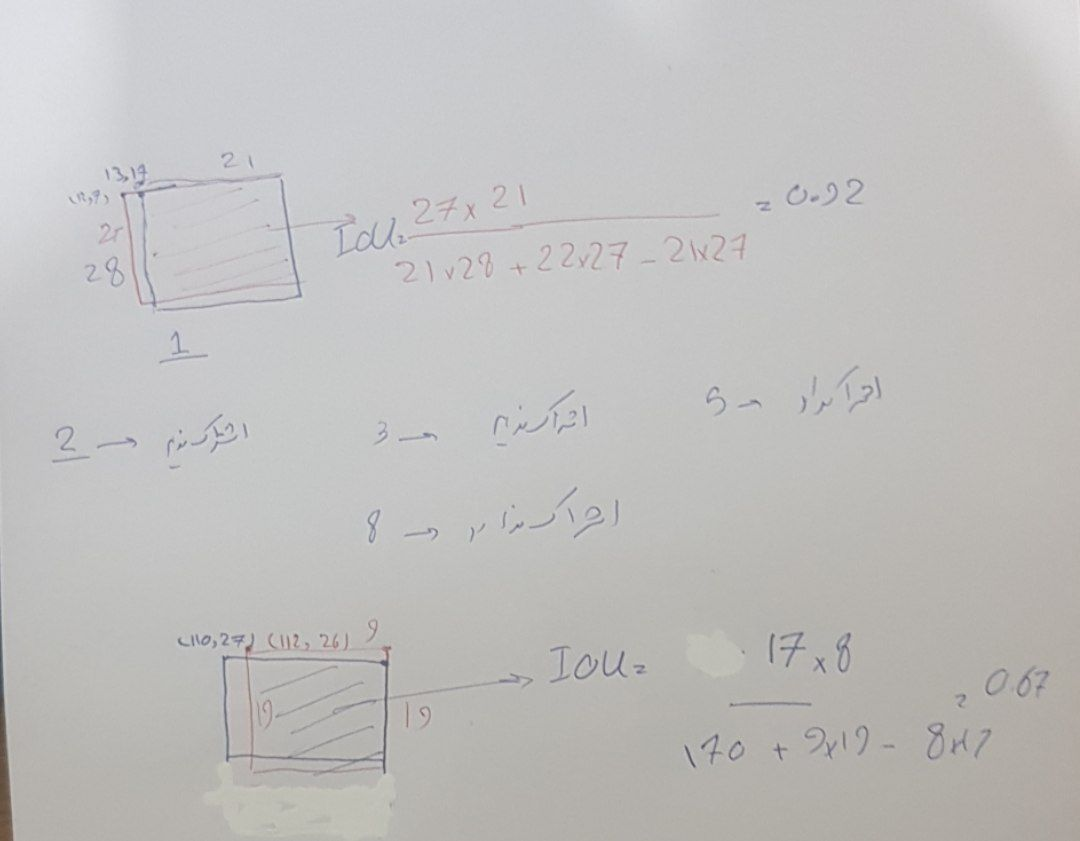
Problem 3

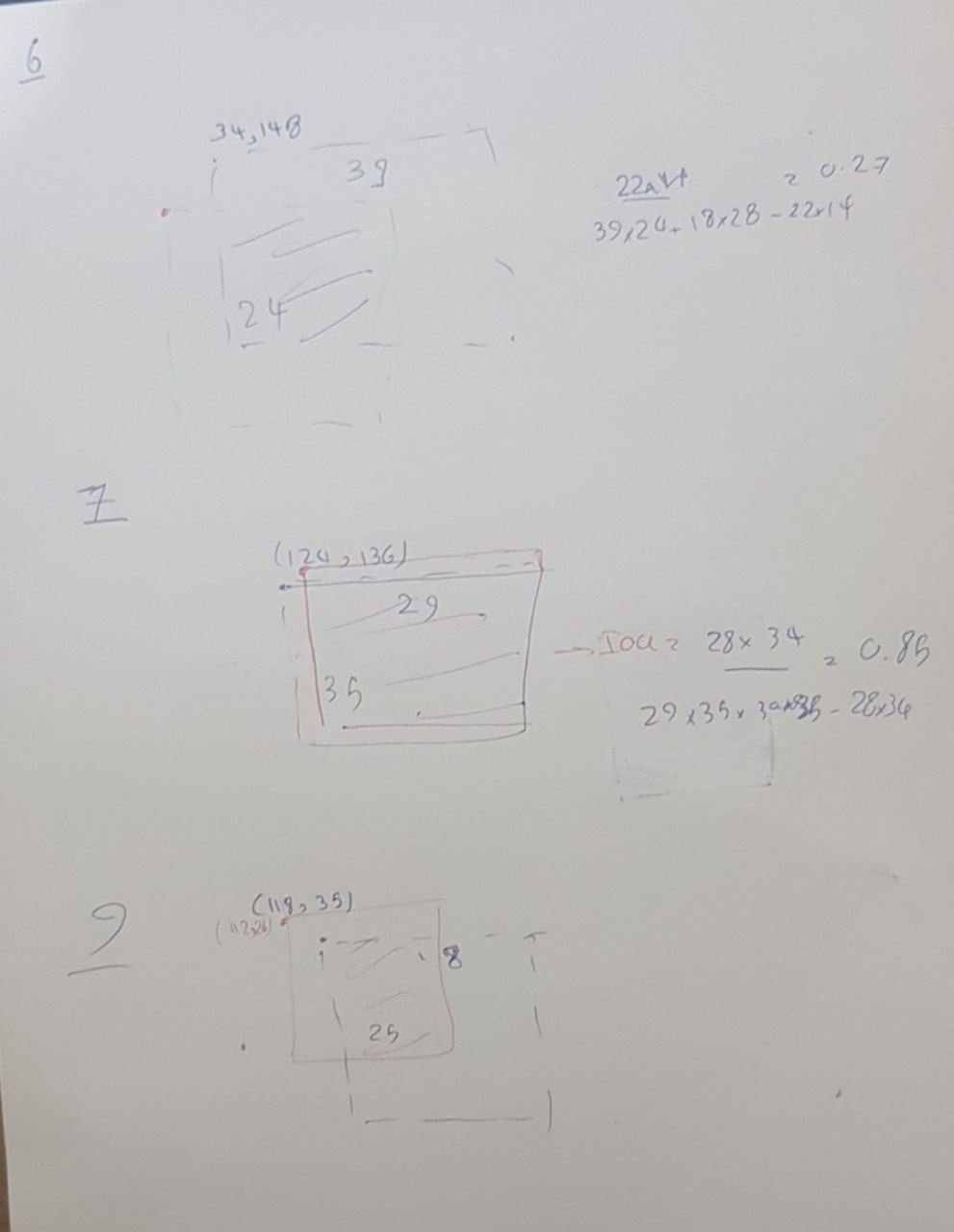
ابتدا دیتای داده شده را رسم میکنیم ، باکس های سبز ، باکس های پیش بینی هستند و باکس های قرمز باکس های garnd truth :



همانطور که مشاهده میشود 5 باکس قرمز وجود دارد ، 9 باکس سبز هم در تصویر هست که 3 باکس از ان ها با باکس های قرمز تطبیق حداکثری دارند و با کمی دقت قابل مشاهده هستند .

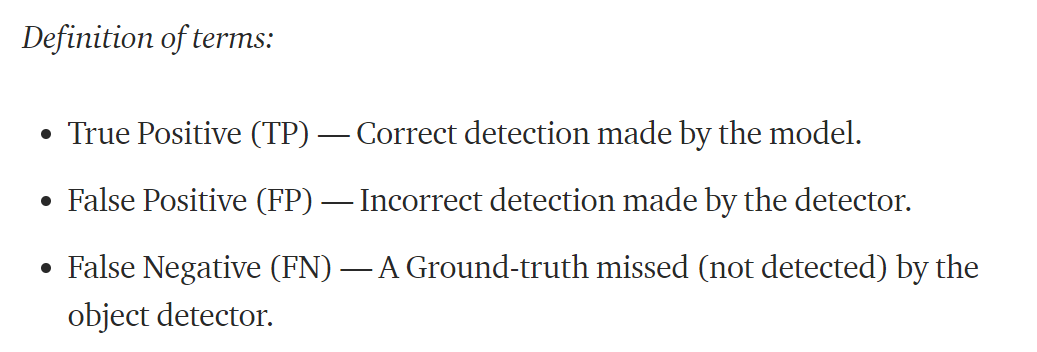
مقدار iou را برای سمپل هایی که داریم حساب میکنیم که برابر است بار مساحت مشترک تقسیم بر مساحت کل .

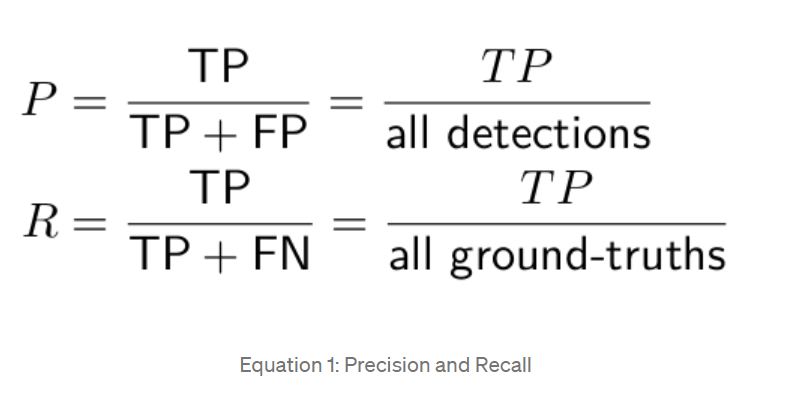




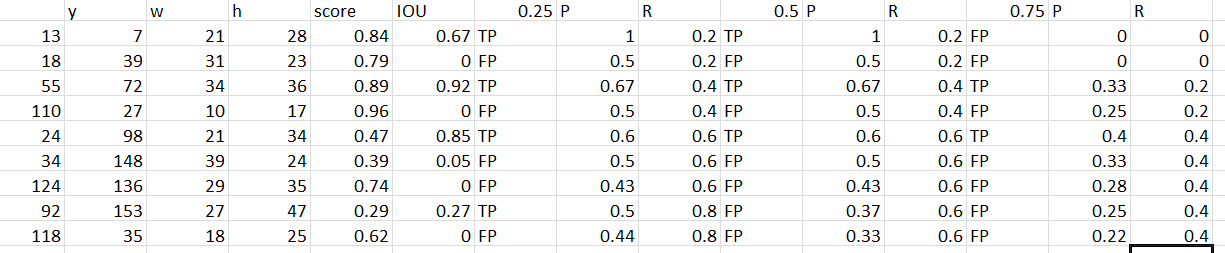
به ترتیب ترشولد iou را برابر 0.25 0.5 0.75 قرار میدهیم و چنانچه iou نمونه ای از این مقدار بیشتر بود tp حساب شده و در غیر این صورت fp است .

سپس با توجه به فرمول زیر :





مقادیر precision و ریکال را برای هر نمونه به ترتیب از score بالا حساب می کنیم .

که در نتیجه محسابات به شکل زیر میشود. 

اگر که نمودار ها را بخواهیم رسم کنیم به شکل زیر می شود.

