

به نام خدا



دانشگاه صنعتی شریف

دانشکده مهندسی برق

یادگیری عمیق

دکتر فاطمی زاده

نیم سال ۱-۰۰

پروژه نهایی Deep Learning

Joint Object Detection and Depth Estimation in Image

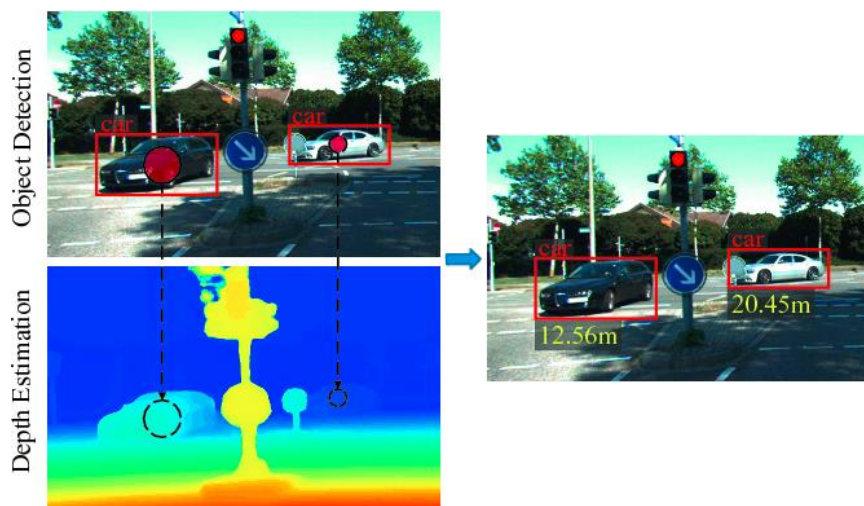
موعد تحویل: ۱۵ بهمن ۱۴۰۰

دستیاران پروژه:

رضا کریم زاده - بهراد احمدپور

- مقدمه:

در این پروژه قصد داریم تا با استفاده از ساختارهای تشخیص عمق، سیستمی را پیاده سازی کنیم که عملیات Object Detection انتخابی را بر سوژه هایی که از حد عمق مشخصی، نزدیک تر هستند، اعمال گرداند. بدین منظور دو شبکه مجزا را ابتدا آموزش و تست نموده و سپس به یکدیگر متصل نموده تا عملیات موردنظر تکمیل گردد.



شرح:

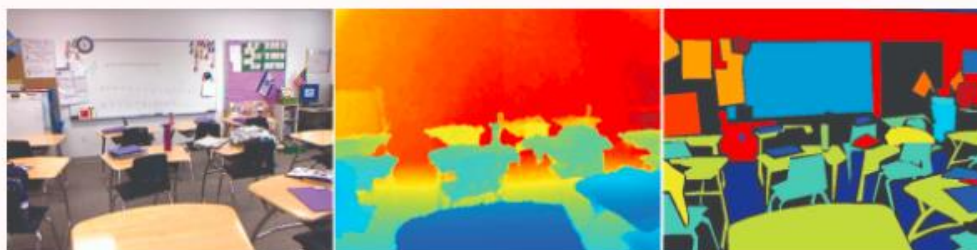
• کلیت پروژه:

با گسترش ماشین‌های هوشمند، تحقیقات حوزه پردازش عمق استریم ویدئویی و تصاویر، مورد توجه قرار گرفتند. یکی کاربردهای این حوزه، ترکیب شبکه‌های پردازش عمق با شبکه‌های تشخیص اشیا^۱ می‌باشد. بدین صورت که داده‌های ورودی، به جای آنکه که مستقیم وارد شبکه تشخیص اشیا شود - که تحقیقات گسترده‌ای در دهه گذشته در خصوص افزایش بازدهی و دقت آن‌ها صورت گرفته است - ابتدا وارد یک شبکه تشخیص عمق شده و عمق تقریبی هر بخش مشخص می‌گردد. سپس بسته به نظر آموزش دهنده شبکه، فیلتر بر تصاویر به صورتی اعمال می‌گردد که تمرکز عملیات پردازشی مشخص گردد.

بدین منظور در این پروژه شما دو شبکه پردازش عمق و تشخیص اشیا را ابتدا به صورت جداگانه آموزش و تست کرده و سپس آن دو را با سازگار مناسب به یکدیگر متصل می‌نمایید.

• داده‌های پروژه:

در این پروژه دادگان^۲ مورد نیاز از طریق دستیاران آموزشی در اختیار شما قرار می‌گیرد. این دادگان شامل جفت تصاویر RGB Depth frames هستند که شما باید از آن‌ها برای آموزش و تست شبکه خود استفاده کنید. جزئیات و نحوه استفاده از دادگان در صفحه مربوطه شرح داده شده است. همچنین توصیه می‌شود از آنجایی که بسیاری از تحقیقات تشخیص عمق بر داده‌های مذکور صورت گرفته است، پیش از شروع به کار، مقاله^۳ مربوطه به دیتاست را مطالعه نموده تا دید مناسبی در خصوص نحوه کار با دیتاست و پروژه کسب کنید.



Output from the RGB camera (left), preprocessed depth (center) and a set of labels (right) for the image.

¹ [Object detection - Wikipedia](#)

² [NYU Depth Dataset V2](#)

³ [Indoor Segmentation and Support Inference from RGBD Images](#)

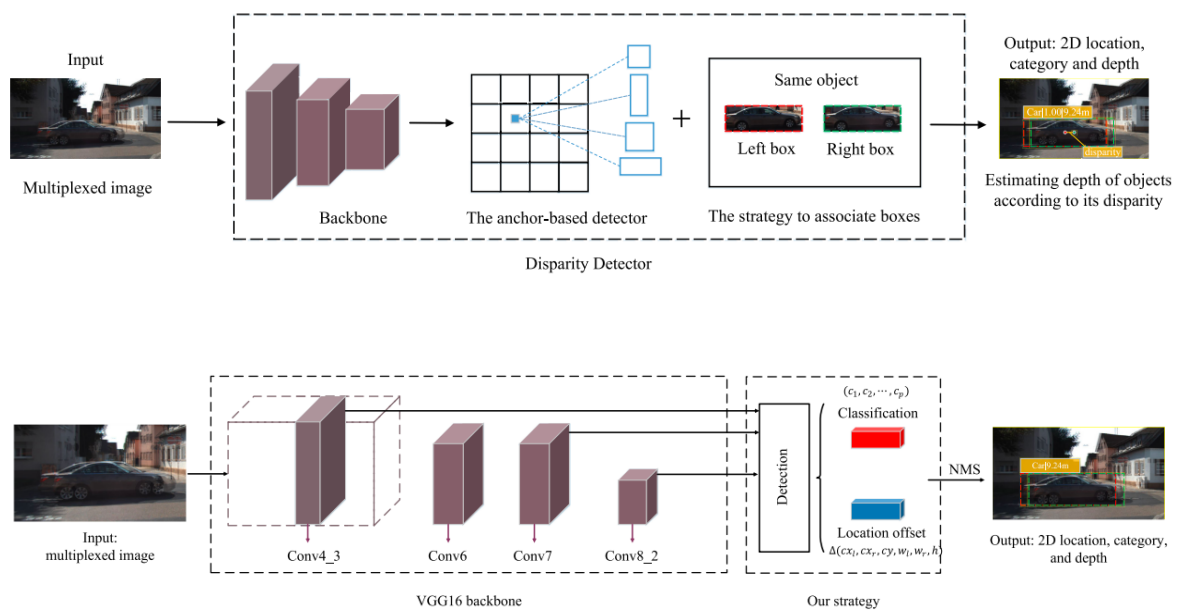
✓ لازم به ذکر است، در صورتی که در پیاده سازی و یا اجرای بخشی از پروژه، از یافته های مقاله یا پروژه دیگری استفاده نمودید، به منبع مورد نظر ارجاع داده و به صورت مختصر روش حل چالش از دید آن مقاله یا پروژه را شرح دهید.

⚠ **تذکر:** بهره گیری از کدهای از پیش نوشته شده و یا ساختارهای مشابه موردی ندارد اما دانشجویان باید مسلط به شبکه و ساختار کد باشند و در گزارش پروژه نهایی، به صورت مشروح، یافته ها و چالش ها را شرح دهند و صرفاً کد و ساختار شبکه را کپی نکرده باشند، که در این صورت نمره ای از پروژه کسب نمی کنند.

• ساختار شبکه:

جهت پیاده سازی روند مدنظر، پیشنهاد می شود دو شبکه به صورت جداگانه مورد آموزش قرار گرفته و سپس به به یکدیگر متصل گردند. در واقع سامانه شما مجموعه ای از دادگان دریافت نموده و به صورت جداگانه عملکرد ساختار تشخیص عمق و تشخیص اشیا را سنجیده و در نهایت با اتصال این دو به یکدیگر سامانه را تکمیل می کنید.

همچنین شما باید روشی برای ارزیابی دقت سامانه متشکل از دو شبکه ارائه دهید که بر مبنای دقت جداگانه هر کدام از شبکه های تشخیص عمق و تشخیص اشیا است.



برای ساختار شبکه کانولوشنی^۴ و ساختار لایه‌های انکودر^۵ و بازگشتی^۶، میتوانید ساختارهای متفاوتی^۷ را مطالعه و بررسی کنید و ساختاری به صلاحدید خودتان پیاده‌سازی کنید. همچنین میتوانید از ساختاری مشابه ساختار ارائه شده در این مقاله^۸ استفاده کنید.

⚠ **تذکر:** نظر به تفاوت بنیادی در هدف نهایی پروژه و مقاله ارجاع داده‌شده، از ساختار مذکور صرفاً الهام گرفته و آن را به صورت کامل پیاده‌سازی نکنید زیرا بعدتر در اتصال آن به ساختار تشخیص اشیا به مشکل برمیخورید.

• ارزیابی:

انجام پروژه در غالب **تیم‌های دونفری** صورت می‌گیرد، بدین منظور توصیه می‌شود برای مدیریت پروژه و جهت سهول در ارزیابی نهایی، از ابزارهای کنترل پروژه مانند Jira, Trello استفاده کرده و برای مدیریت روند پیشبرد کدها و گزارشات از git استفاده کنید تا عملکرد هر یک از اعضا به صورت کامل مشخص گردد. عملکرد جداگانه هر شبکه (شبکه تشخیص عمق و تشخیص اشیا) با دادگان تست به صورت رقابتی بررسی و نمره‌دهی می‌شود و همانطور که قبل تر ذکر شد، برای ارزیابی ترکیب دو شبکه، از معیار پیشنهادی شما استفاده می‌شود. بدین منظور لازم است در این خصوص تحقیق کرده و در گزارش خود جزئیات آن را شرح دهید؛ توجه کنید که روش ارزیابی دقت عملکرد سامانه کلی، علاوه بر دوری از پیچیدگی، باید با دقت مناسب و بدون نیاز به تست جداگانه کل سامانه، نتیجه مورد نظر را از روی دقت هر شبکه، بدست آورد. پیشنهاد می‌شود در خصوص برخی از متدهای ارزیابی شبکه مانند معیار BLEU^۹ مطالعه فرمایید.

⚠ **توجه:** نمره پروژه تنها وابسته به گزارش نهایی می‌باشد و **پروژه شامل تحویل حضوری نمی‌باشد**، بدین منظور گزارشی تهیه نموده که در آن مدل طراحی شده، نتایج به دست آمده، نوآوری‌ها و سایر

⁴ [Convolutional neural network - Wikipedia](#)

⁵ [Autoencoder - Wikipedia](#)

⁶ [Recurrent neural network - Wikipedia](#)

⁷ [How do neural networks see depth in single images?](#)

⁸ [Joint Object Detection and Depth Estimation in Multiplexed Image](#)

⁹ معیار BLEU در اصل برای سنجش کیفیت ترجمه ماشین (Translation Machine) استفاده می‌شود. منظور از کیفیت در این تعریف، نزدیکی خروجی ماشین با خروجی درست (که توسط انسان بدست آمده) است. برای آشنایی بیشتر با این معیار می‌توانید به منابع زیر مراجعه کنید:

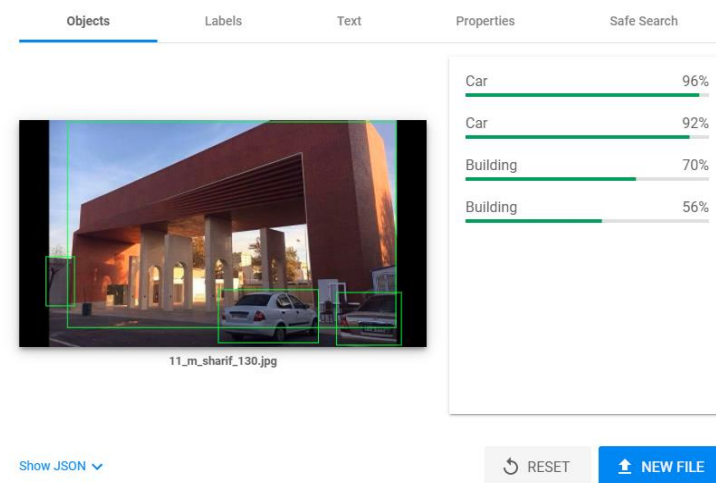
[BLEU - Wikipedia](#)

<https://www.aclweb.org/anthology/P02-1040.pdf>

[C5W3L06 Bleu Score \(Optional\) - YouTube](#)

مسائل مانند انتخاب بهینه ساز و هایپرپارامترهای آن و نحوه حل چالش‌های گوناگون را توضیح داده باشید. در نتیجه سعی کنید از همان مراحل اول به فکر تهیه و تدوین آن باشید.

لازم به ذکر چنانچه برای کاربردی شدن سامانه خود ابتکاری به خرج دادید که نحوه کار و تست آن را به بهینه می‌سازد (برای مثال نحوه دریافت و خروجی سامانه را به صورت وب‌اپلیکیشن درآورده‌اید) و یا در مواردی اعم از ساختار شبکه، نحوه اتصال، معیار ارزیابی پیشنهادی و ... روشی خلاقانه و خارج از چارچوب در اختیار گیرید **که مورد تائید دستیاران آموزشی می‌باشد**، تا سقف ۳۰ درصد نمره امتیازی کسب می‌کنید.



[Try it!](#) / [Cloud Vision API](#) / [Google Cloud](#)

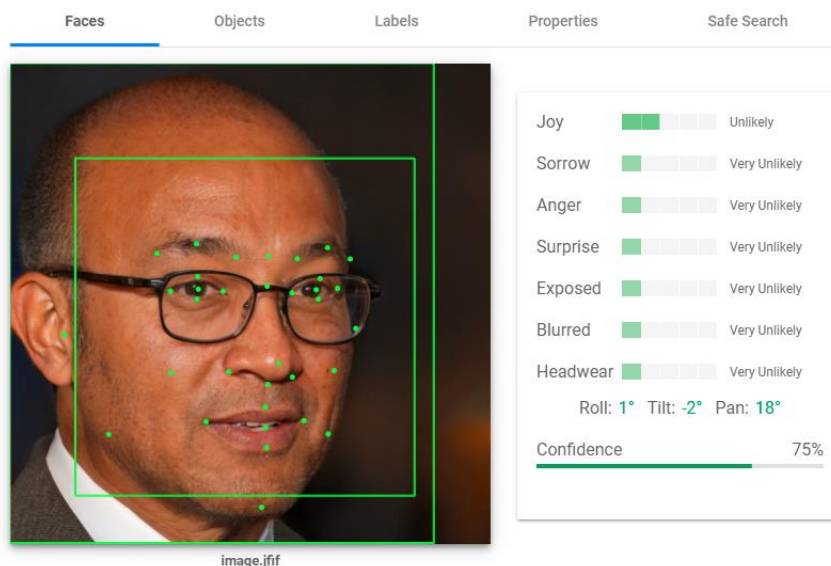


Photo from [This Person Does Not Exist](#)

موفق باشید