به نام خدا

تشخیص لبه‌های معنایی اجسام با استفاده از مدل‌های عمیق

گزارش روند پیشرفت

24 فروردین 1394

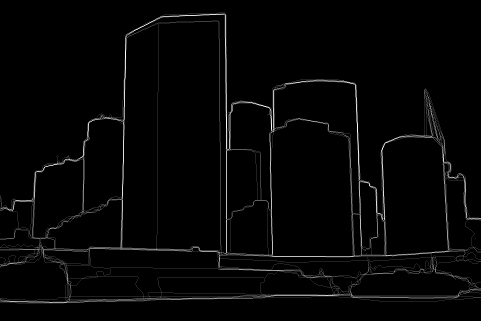
عرفان نوری

# تعریف مسئله

# مجموعه داده

برای این مسئله از مجموعه داده‌ی BSDS500 [1] استفاده خواهد شد. این مجموعه داده دارای 200 تصویر آموزش، 100 تصویر اعتبارسنجی و 200 تصویر آزمون می‌باشد. هر یک از این تصاویر از سه کانال رنگی تشکیل شده‌اند و اندازه‌ی آنها 321×481 پیکسل می‌باشد. برای همه‌ی این تصاویر تعدادی تصویر لبه توسط متخصصان تهیه شده است. تعداد این تصاویر لبه برای بیشتر تصاویر 5 عدد می‌باشد، هر چند این مقدار بین 4 و 8 متغییر است. به دلیل تفاوت میان تصاویر لبه‌ی به‌دست‌آمده برای تصاویر و همچنین تعدد آنها، از تصویر لبه‌ی میانگین به عنوان تصویر لبه‌ی اصلی هر تصویر استفاده می‌شود و مراحل یادگیری، اعتبارسنجی و آزمون خروجی نیز با استفاده از تصویر لبه‌ی میانگین[[1]](#footnote-1) انجام می‌گیرد. دیر زیر نمونه‌ای از یک تصویر و همچنین تصویر لبه‌ی میانگین نمایش داده شده است.





# معیارهای تشخیص کیفیت

مسئله‌ی تشخیص لبه‌ها را می‌توان به صورت یک مسئله‌ی کلاسه‌بندی فرمول‌بندی کرد که هدف در آن تشخیص پیکسل‌های لبه از پیکسل‌های غیرلبه می‌باشد. بنابراین می‌توان چارچوب معیار دقت-فراخوانی[[2]](#footnote-2) را با استفاده از لبه‌های نشان‌گذاری شده توسط انسان از مجموعه داده‌های BSDS500 به عنوان حقیقت مبنا[[3]](#footnote-3) استفاده کرد [2]. استفاده از این بخصوص زمانی اهمیت خود را نشان می‌دهد که کاربردهای تشخیص لبه‌ها را در مسائلی که از لبه‌های تصویر استفاده می‌کنند، مانند تشخیص عمق از دو تصویر stereo یا تشخیص اشیا، در نظر بگیریم. برای محاسبه‌ی این معیار باید مشخص کنیم که چه زمانی لبه‌های تشخیص داده شده صحیح هستند و چه زمانی خطا در تشخیص رخ داده است. هر نقطه بر روی منحنی دقت-فراخوانی با در نظر گرفتن خروجی تشخیص‌دهنده در یک آستانه‌ی مشخص است. تصویر خروجی باید آستانه‌گذاری[[4]](#footnote-4) شود تا یک تصویر دوگانی[[5]](#footnote-5) بوجود بیایید. زیرا تصاویر نشان‌گذاری شده توسط انسان نیز تصاویر دوگانی هستند. ابتدا محاسبه‌ی دقت و فراخوانی یک تصویر خروجی الگوریتم در مقابل یک تصویر نشان‌گذاری شده توسط انسان را بررسی می‌کنیم. یک راه‌حل ابتدایی در نظر گرفتن همه‌ی پیکسل‌هایی که در هر دو تصویر به عنوان لبه تشخیص داده شده‌اند و رد کردن بقیه پیکسل‌ها می‌باشد. مشکل این راه‌حل این است که الگوریتم‌هایی که لبه‌های مناسبی تولید می‌کنند ولی مکان آنها دقیق نیست، امتیاز کمی از این معیار کسب می‌کنند. پس لازم است به این نکته توجه شود که در انتخاب معیار مناسب به مسئله‌ی دقیق نبودن مکان لبه‌ها در تصویر خروجی الگوریتم توجه شود، زیرا حتی تصاویر نشان‌گذاری شده توسط انسان نیز این مشکل را نسبت به یکدیگر دارند. به عنوان راه‌حل بهتری برای محاسبه‌ی مقدار معیار، نقشه‌ی لبه‌ی خروجی الگوریتم را با همه‌ی نقشه‌های لبه‌ی نشان‌گذاری شده توسط انسان مقایسه می‌کنیم. اگر پیکسلی که به عنوان لبه تشخیص داده شده است در هیچ یک از نقشه‌های لبه‌ی نشان‌گذاری شده به عنوان لبه نباشد، در این صورت تشخیص آن پیکسل به عنوان لبه نادرست است و آن پیکسل به عنوان یک یقین کاذب[[6]](#footnote-6) شناخته می‌شود. تعداد اشتراکات در مورد تشخیص لبه‌بودن یک پیکسل نیز در مقایسه با همه‌ی تصاویر نشان‌گذاری شده شمارش می‌شود و میانگین گرفته می‌شود. بنابراین برای اینکه مقدار فراخوانی برای یک الگوریتم 1 باشد باید اطلاعات همه‌ی تصاویر نشان‌گذاری شده توسط انسان را تولید کند. پس بنابراین اگر خروجی یک الگوریتم باشد، یک منحنی دقت-فراخوانی از آن محاسبه می‌کنیم. هر نقطه بر روی این منحنی به طور مستقل با آستانه‌گذاری تصویر خروجی و در نتیجه تبدیل آن به یک تصویر دوگانی محاسبه شده و به دست می‌آید. منحنی دقت-فراخوانی اطلاعات بسیاری را در مورد الگوریتم نمایش می‌دهد. اگر لازم باشد فقط یک مقدار به عنوان معیار گزارش شود، می‌توان مقدار بیشینه‌ی F-measure را محاسبه کرد و آن را گزارش کرد.

# روش‌های مورد نظر

## استفاده از مدل‌های از پیش آموزش دیده شده

## طراحی شبکه کامل برای تشخیص لبه

## استفاده از شبکه‌ی شرطی خصمانه مولد

# مشکلات پیش رو

# برنامه‌ی پیش رو

1. mean Contour Map [↑](#footnote-ref-1)
2. Precision-Recall [↑](#footnote-ref-2)
3. Ground truth [↑](#footnote-ref-3)
4. Thresholding [↑](#footnote-ref-4)
5. binary [↑](#footnote-ref-5)
6. False Positive

   ترجمه‌های زیر برای این اصطلاحات پیشنهاد می‌شوند:

   یقین کاذب: False Positive شک کاذب: False Negative شک راستین: True Negative یقین راستین: True Positive [↑](#footnote-ref-6)