



دانشگاه صنعتی شریف
دانشکده مهندسی کامپیوتر
سمینار کارشناسی ارشد گرایش هوش مصنوعی

عنوان:
دسته‌بندی ریزدانه‌ای تصاویر
Fine-grained Image Classification

نگارش:
یاسر سوری
۹۲۲۰۴۷۴۴

استاد راهنما:
دکتر شهره کسایی

استاد ممتحن داخلی:
دکتر محمد تقی منظوری شلمانی

شهریور ۹۳

چکیده: دسته‌بندی تصاویر ریزدانه‌ای عبارت است از دسته‌بندی تصاویر در حالتی که دسته‌های مورد نظر همگی زیر دسته‌ی یک دسته‌ی کلی‌تر هستند. برای مثال وقتی دسته‌های ما همگی گونه‌های مختلف پرندگان هستند، همگی زیر دسته‌ی پرند هستند. در این حالت خاص مسئله دسته‌ها از نظر ظاهری معمولاً از نظر ظاهری بسیار به یکدیگر شبیه هستند به گونه‌ای که افراد غیر متخصص نمی‌توانند دسته‌ها را از یکدیگر تمایز دهند. در چنین شرایطی راه حل‌های ارائه شده برای مسئله دسته‌بندی تصاویر معمولی اکثراً نتایج خوبی کسب نمی‌کنند. لذا ارائه روش‌هایی جدید برای حل این مسئله لازم است. در این گزارش ابتدا به مرور روش‌های مهم در دسته‌بندی تصاویر معمولی و سپس به مرور روش‌های ارائه شده برای دسته‌بندی تصاویر ریزدانه‌ای می‌پردازیم. سپس روش انتخاب شده و دلایل انتخاب آن را بررسی می‌کنیم.

واژه‌های کلیدی: بینایی کامپیوتری، بازشناسی شیء، دسته‌بندی تصاویر، بازشناسی ریزدانه‌ای، دسته‌بندی تصاویر ریزدانه‌ای.

۱ مقدمه

در دسته‌بندی تصویر^۱ هر تصویر با توجه به محتوایش دسته‌بندی می‌شود. برای مثال آیا تصویر شامل شیء خودرو هست یا خیر. معمولاً در بینایی کامپیوتری مسئله بدین صورت است که تعدادی دسته مشخص را در نظر می‌گیریم (مثلاً انسان، خودرو، ساختمان، تلویزیون، صندلی، اسب و ...) سپس طبق چارچوب معمول یادگیری ماشین، توسط تعدادی تصویر شامل یکی از دسته‌ها (نمونه‌های مثبت) و تعدادی تصویر بدون شیء از آن دسته (نمونه‌های منفی) یادگیری برای آن دسته انجام می‌شود. در نهایت پس از یادگیری تمام دسته‌ها در مواجهه با تصویر جدید لازم است تشخیص دهیم که آیا شیء از هر کدام از آن دسته‌های مورد بررسی در تصویر وجود دارد یا خیر [۱].

مسئله مطرح بوده است.

مسئله در حالت کلی بسیار مشکل است.

یک مسئله معکوس است.

برای مقایسه کارایی باشد.

۲ کارهای پیشین

به صورت یک مسئله تخمین حالت سیستم پویا^۲ در نظر گرفت.

$$p(X_t|I_t) \propto \underbrace{p(I_t|X_t)}_{\text{درست‌نمایی}} \int \underbrace{p(X_t|X_{t-1})}_{\text{مدل جابجایی}} p(X_{t-1}|I_{t-1}) dX_{t-1} \quad (1)$$

References

- [1] L. Fei-Fei, R. Fergus, and P. Perona, "Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories," *Computer Vision and Image Understanding*, vol. 106, no. 1, pp. 59–70, 2007.

واژه‌نامه

Dynamic System^۲

Image classification^۱