

خلاصه:

در این مقاله روش جدید برای تشخیص keypoint معرفی شده است که روش های سنتی و شبکه عصبی کانولوشنی با معماری چند لایه اما کم عمق ترکیب می شود. فیلترهای دست ساز ساختارهای anchor را برای فیلترهای آموخته شده فراهم می کنند که ویژگی های تکرار شونده را بومی سازی ، امتیازدهی و رتبه بندی می کند. نمایش فضای مقیاس برای استخراج نقاط کلیدی در سطوح مختلف در شبکه استفاده می شود. و یک تابع هزینه را برای شناسایی ویژگی های مناسب و قوی که در طیف وسیعی از مقیاس ها وجود دارد و به حداکثر رساندن نمره تکرارپذیری طراحی می کنیم.

مدل Key.Net معرفی شده در مقاله بر روی داده هایی که به صورت مصنوعی از ImageNet ایجاد شده و در معیار HPatches ارزیابی می شود ، آموزش دیده است. نتایج نشان می دهد که رویکرد این مقاله از نظر تکرارپذیری ، عملکرد و پیچیدگی بهتر از detectors پیشرفته موجود عمل می کند.

مقدمه:

پیشرفتهای تحقیقاتی در آشکارسازها و توصیفگرهای ویژگی محلی وجود داشته منجر به پیشرفتهای چشمگیری در زمینه هایی مانند تطبیق تصویر ، تشخیص object ، نوبری خود هدایت شده یا بازسازی سه بعدی شده است.

اگرچه جهت کلی روشهای تطبیق تصویر در حال حرکت به سمت سیستمهای مبتنی بر یادگیری و بخصوص شبکه عصبی کانولوشنی است ، اما مزیت روشهای یادگیری نسبت به روشهای دست ساز به روشنی در تشخیص کلیدواژه اثبات نشده است.

با وجود ناکارآمدی غیر عملی روشهای اولیه به طور خاص ، شبکه های عصبی (CNN) Convolutional توانستند به طور قابل توجهی خطای تطبیق را در توصیف کنندگان محلی کاهش دهند.

این کارها باعث تلاشهای تحقیقاتی بیشتر و در نتیجه بهبود کارایی توصیف کنندگان مبتنی بر CNN می شود.

با این حال ، محبوبیت روزافزون همدست های واقعیت افزوده و همچنین برنامه های گوشی های هوشمند ، توجه بیشتری را به detectors ویژگی محلی قابل اعتماد و کارآمد جلب کرده است که می تواند برای تخمین سطح ، بازسازی سه بعدی پراکنده ، کسب مدل سه بعدی یا ترازبندی اشیا و ... استفاده شود.

به طور سنتی ، آشکارسازهای ویژگی محلی بر اساس فیلترهای مهندسی شده به اصطلاح دست ساز ساخته می شدند.

به عنوان مثال ، رویکردهایی مانند Harris-Laplace، Difference of Gaussians یا Hessian-Affine از ترکیب مشتقات تصویر برای محاسبه نقشه های ویژگی استفاده می کنند ، که به طرز قابل توجهی مشابه عملیات در لایه های آموزش دیده CNN است.

فقط با چند لایه ، یک شبکه می تواند رفتار detectors های سنتی را با یادگیری مقادیر مناسب در فیلترهای کانولوشن تقلید کند.

با این حال، برخلاف موفقیت در توصیفات تصویر محلی مبتنی بر CNN، پیشرفت در ارائه شده توسط روشهای کاملاً مبتنی بر CNN که اخیراً پیشنهاد شده اند، از نظر معیارهای پذیرفته شده گسترده ای مانند تکرارپذیری محدود هستند.

یکی از دلایل دقت کم آنها هنگام تخمین پارامترهای ترکیبی مناطق مشخصه است. مقاومت به تغییرات مقیاس به ویژه مشکل ساز به نظر می رسد در حالی که پارامترهای دیگر مانند جهت گیری غالب را می توان به خوبی توسط CNN کنترل کرد

این باعث ایجاد انگیزه در معماری جدید در ایم مقاله می شود که Keynet نامیده می شود که از فیلترهای دست ساز و آموخته شده و همچنین نمایش چند مقیاس استفاده می کند.

معماری Key.Net در شکل ۱ نشان داده شده است. معماری پیشنهادی Key.Net ترکیبی از فیلترهای ساخته شده و آموخته شده برای استخراج ویژگی ها در مقیاس های مختلف است. نقشه های ویژگی نمونه برداری شده و بهم پیوسته اند فیلتر آخرین آموخته شده برای به دست آوردن نقشه پاسخ نهایی ، میزان فضای مقیاس را ترکیب می کند.

معرفی فیلترهای دست ساز ، که به عنوان لنگر نرم عمل می کنند ، امکان کاهش تعداد پارامترهای استفاده شده توسط detectors پیشرفته را فراهم می کند ، در حالی که عملکرد را از نظر تکرارپذیری حفظ می کنند.

این مدل بر روی نمایش چند مقیاس از تصاویر در اندازه کامل عمل می کند و یک نقشه پاسخ را شامل نمره نقطه کلیدی برای هر پیکسل برمی گرداند.

ورودی چند مقیاسه به شبکه اجازه می دهد تا نقاط کلیدی پایدار را ارائه دهد ، بنابراین قدرت تغییرات مقیاس را فراهم می کند.

در حالت ایده آل ، یک detectors قوی قادر است ویژگی های یکسانی را برای تصاویری که تحت تغییرات هندسی یا فوتومتریک مختلفی قرار دارند ، ارائه دهد.

تعدادی از کارهای مرتبط عملکرد هدف خود را برای پرداختن به این موضوع متمرکز کرده اند ، اگرچه این کارها یا براساس تکه های محلی یا از دست دادن رگرسیون نقشه کلی یا به اصطلاح نقشه جهانی بود

ما یک اپراتور کاملاً متفاوت ، چند مقیاس را طراحی می کنیم که نکات کلیدی را در مناطق چند مقیاس پیشنهاد می کند.

ما به طور گسترده ای از معیار HPatches که اخیراً معرفی شده برای میزان دقت استفاده میکنیم.

به طور خلاصه ، کارهای انجام شده در این تحقیق به شرح زیر است:

- یک key point detectors که ویژگی های دست ساخته و آموخته CNN را با هم ترکیب می کند.
- یک تابع هزینه و اپراتور جدید برای شناسایی و رتبه بندی نقاط کلیدی پایدار در مقیاس ها
- تشخیص ویژگی چند مقیاس با معماری کم عمق

بقیه مقاله به شرح زیر است.

ما کارهای انجام شده مربوطه را در بخش ۲ مشاهده می کنیم. بخش ۳ معماری ترکیبی پیشنهادی Key.Net فیلترهای CNN ساخته و ساخته شده و آموخته شده را ارائه می دهد و بخش ۴ تابع هزینه را معرفی می کند. جزئیات اجرا در بخش ۵ آورده شده و نتایج در بخش ۶ ارائه شده است.

۲. کارهای مرتبط

تحقیقات بسیاری وجود دارد که به طور گسترده در مورد روش های تشخیص ویژگی بحث می کند. ما کارهای مرتبط را در دو دسته اصلی ارائه می دهیم: روش های سنتی و مبتنی بر یادگیری.

۲.۱ ردیاب های دست ساز یا ردیاب های مهندسی شده

ردیاب های ویژگی های سنتی ساختارهای هندسی را از طریق الگوریتم های مهندسی شده بومی سازی می کنند ، که اغلب از آنها به عنوان روش های دستی یاد می شود.

ردیاب های Harris و Hessian از مشتقات تصویر درجه اول و دوم برای یافتن گوشه ها یا لکه های تصاویر استفاده کردند. این آشکارسازها بیشتر برای مقابله با تحولات چند مقیاسی و تغییر شکل گسترش یافتند.

بعداً ، SURF با استفاده از تصاویر یکپارچه و تقریب ماتریس Hessian ، روند شناسایی را تسریع کرد. اگرچه آشکارسازهای گوشه قوی و کارآمد هستند ، اما روش های دیگر به دنبال ساختارهای جایگزین در تصاویر هستند. SIFT به دنبال لکه هایی در سطح چند مقیاس بود و MSER مناطق پایدار را به عنوان نقاط کلیدی تقسیم و انتخاب کرد.

۲.۲. detectors های یاد گرفته شده

موفقیت روشهای آموخته شده در تشخیص object و توصیف ویژگی ها ، محققین را بر آن داشت تا تکنیک های مشابه را برای ردیاب های ویژگی کشف کند. FAST یکی از اولین تلاشها برای استفاده از یادگیری ماشین برای استخراج یک آشکارساز گوشه بود.

کارهای بعدی با بهینه سازی FAST ، افزودن توصیفگر یا برآورد جهت گیری ، FAST را گسترش دادند. آخرین پیشرفت های CNN نیز در شناسایی ویژگی ها تأثیر داشته است.

TILDE چندین مدل رگرسیون خطی چند قطعه ای را آموزش داده است تا نقاط بهره ای را که تحت تغییرات شدید مقاوم هستند ، شناسایی کند.

در تحقیقی دیگر فرمول جدیدی را برای آموزش CNN براساس محدودیت های متغیر ویژگی معرفی شد.

در تحقیقی دیگر پیش بینی کرد که چه ویژگی ها و توصیفات با هم مطابقت دارند.

در تحقیقی دیگر برای یادگیری پیش بینی پارامترهای ترکیبی یک ویژگی محلی ، از ضریب توصیف کننده استفاده کرد.

اخیراً ، یک شبکه را معرفی کرد تا یاد بگیرد. علاوه بر این ، CNN های دیگر نیز برای انجام وظایفی فراتر از تشخیص یا تطبیق مطالعه شد.