

## روشی نوین برای بهبود عملکرد توصیف گر نقاط کلیدی شبکه سریع با بهره‌گیری از کلاسه‌بند بیزین شبه ساده

ایمان شریفی راد<sup>۱</sup>، فرشاد تاجری پور<sup>۲</sup>، زهره عظیمی فر<sup>۳</sup>

۱- دانشگاه شیراز، Sharifirad@hotmail.com

۲- دانشگاه شیراز، Tajeri@shirazu.ac.ir

۳- دانشگاه شیراز، Azimifar@shirazu.ac.ir

### خلاصه

مسئله اصلی که در این مقاله مورد بررسی قرار گرفته است، ارائه روشی نوین برای تطبیق نقاط کلیدی در تصاویر دیجیتال است. تطبیق نقاط کلیدی از اساسی‌ترین بخش‌های ساخت تصاویر دوخته شده، دنبال کردن اشیاء متحرک، ساخت مدل سه‌بعدی از تصاویر دوبعدی و تشخیص اشیاء است. در این راه ما از توصیف گر ویژگی محلی نقاط کلیدی شبکه سریع بهره برده‌ایم که در آن از روش نزدیک‌ترین همسایه برای یافتن نقطه کلیدی متناظر استفاده شده است. همچنین ما با فرض شرط استقلال گروهی ویژگی‌ها، از کلاسه‌بند بیزین شبه ساده برای بهبود عملیات تطبیق نقاط کلیدی تصویر استفاده کرده‌ایم. نتیجه آزمایش‌های انجام شده بر روی مجموعه داده‌های استاندارد در زمینه تطبیق الگو، بهبود قابل توجهی در زمینه نرخ تشخیص صحیح و سرعت اجرا در مرحله تطبیق نقاط کلیدی را نشان می‌دهد.

کلمات کلیدی: پردازش تصویر، بینایی ماشین، توصیف گر ویژگی‌های محلی، تطبیق الگو.

### ۱. مقدمه

شناسایی وصله‌های بافت دار اطراف نقاط کلیدی در تصاویری که مورد تغییرات ظاهری و نوری مختلفی واقع شده‌اند، از اساسی‌ترین و چالش برانگیزترین مسائل بینایی ماشین محسوب می‌شود. نتیجه تطبیق نقاط کلیدی به دست آمده در تصاویر مختلف می‌تواند در بسیاری از کاربردها نظیر دوخت تصاویر، استخراج اطلاعات شکل‌های سه بعدی، رهگیری اشیاء در تصاویر ویدئویی استفاده شود. این موضوع همچنین نقش بسیار مهمی در تشخیص موضوعی اشیاء و بازیابی تصاویر مشابه دارد. با توجه به اهمیت فراوان تشخیص وصله‌ها در نحوه کارکرد الگوریتم‌های مذکور، سرعت و دقت از اصلی‌ترین معیارهای آن است. البته به دلیل تغییرات زیاد شکل ظاهری سطح اطراف نقاط کلیدی با تغییر زاویه عکس برداری، دست یافتن همزمان به هر دو معیار سرعت و دقت دشوار به نظر می‌رسد.

<sup>۱</sup> دانشجوی تحصیلات تکمیلی

<sup>۲</sup> استادیار دانشگاه شیراز

<sup>۳</sup> استادیار دانشگاه شیراز

<sup>۴</sup> Patch

راه حل معمول حل این مسئله، ساخت توصیف‌گرهایی از وصله‌های محلی تصویر با خاصیت affine invariant و مقایسه آن‌ها با توصیف‌گرهای استخراج شده از سایر تصاویر است؛ که معمولاً شامل بخش‌های انتخاب مقیاس، تصحیح چرخش و نرمال سازی شدت تصویر می‌باشد [۱]. نتیجه تحقیق‌های اخیر نشان می‌دهند که برای جفت‌سازی تکه‌های تصویر، مقایسه ساده شدت تعدادی جفت نقطه، دارای کارایی مناسبی است [۲]. از مهم‌ترین و جدیدترین روش‌های ارائه شده در این زمینه، می‌توان به الگوریتم نقاط کلیدی شبکه‌ی سریع<sup>۵</sup> اشاره نمود که یک توصیف‌گر سریع، مقاوم و بهینه می‌باشد و از سیستم بینایی انسان و نحوه عملکرد شبکه‌ی چشم الهام گرفته شده است. روش مذکور در عین حفظ سادگی توانسته است از نظر مصرف بهینه حافظه، سرعت و عملکرد، بهتر از سایر الگوریتم‌های شناخته شده در این زمینه عمل کند.

توصیف‌گرهای ارائه شده در سال‌های اخیر از روش نزدیک‌ترین همسایه به عنوان کلاسه‌بند، برای یافتن نقطه متناظر با یک نقطه خاص در تصویر ثانویه استفاده کرده‌اند. در [۴] توصیف‌گری ارائه شده است که با تبدیل مسئله یافتن تطابق به یک مسئله کلاسه‌بندی، نقاط متناظر در تصاویر را پیدا می‌کند. بردار ویژگی مورد استفاده در این روش با انتخاب تصادفی تعدادی جفت نقطه ساخته می‌شود. این روش استخراج ویژگی در مقایسه با توصیف‌گر نقاط کلیدی شبکه‌ی سریع که جهت انتخاب دقیق‌تر جفت نقاط از ساختاری الهام گرفته شده از شبکه‌ی چشم بهره می‌برد، ضعیف‌تر است [۵].

در پژوهش انجام شده از دیدگاه کلاسه‌بندی مورد استفاده در مرحله تطبیق الگو در [۴]، برای بهبود دقت و بار محاسباتی الگوریتم نقاط کلیدی شبکه‌ی سریع، استفاده کرده‌ایم. در ادامه، ابتدا مروری بر تحقیق‌های انجام شده در این زمینه خواهیم داشت. در بخش روش ارائه شده، ما با در نظر گرفتن روش‌های پیشین، الگوریتم پیشنهادی خود را به تفصیل بیان کرده‌ایم. همچنین ما از مجموعه داده استاندارد graffiti برای آزمایش الگوریتم پیشنهادی استفاده کرده‌ایم که نتایج آن در بخش تحلیل و ارزیابی آورده شده است.

## ۲. مرور پژوهش‌های پیشین

تشخیص تکه‌های تصویر که از اطراف نقاط کلیدی استخراج شده‌اند، به دلیل نیاز به پایداری مناسب نسبت به انسداد جزئی و بهینگی محاسباتی، از ضرورت‌های موجود در بسیاری از مسائل بینایی ماشین است. در این زمینه دو دسته کلی از روش‌ها ارائه شده است.

اولین دسته، متکی بر توصیف‌گرهای محلی است که تلاش دارند نسبت به انواع خاصی از تغییر شکل‌ها، نامتغییر یا حداقل مقاوم باشند [۱] [۶]. این روش‌ها عموماً به جهت چرخ و مقیاس تخمین زده شده توسط یابنده‌های نقاط کلیدی وابسته هستند. در میان این روش‌ها می‌توان به روش تبدیل ویژگی نامتغییر به مقیاس<sup>۱</sup> [۱] اشاره کرد که بردار ویژگی آن از هیستوگرام‌های شیب‌های محلی ساخته و نشان داده شده که بسیار خوب عمل می‌کند. در [۷] روشی ارائه شده است که نشان می‌دهد می‌توان بخش مربوط به کاهش بعد را با ساخت یک توصیف‌گر دودویی کوتاه خلاصه کرد؛ به صورتی که هر بیت به صورت مستقل محاسبه شده باشد. در این روش از طریق مقایسه شدت تصویر در ۵۱۲ جفت از پیکسل‌ها، بردار ویژگی ساخته می‌شود. موقعیت پیکسل‌ها به صورت تصادفی بر اساس یک توزیع نرمال به مرکزیت وصله انتخاب شده است. بردار ویژگی به دست آمده، نسبت به تغییر مقیاس و چرخش، نامتغییر نیست. همچنین به عنوان یک راه حل دیگر، در [۸] توصیف‌گری دودویی ارائه شده که بردار ویژگی آن نسبت چرخش و نوز نامتغییر است. به طور مشابه، در [۹] یک توصیف‌گر دودویی معرفی شده که نسبت به تغییر مقیاس و چرخش نامتغییر است. در [۵] یک توصیف‌گر ویژگی دودویی معرفی شده که نسبت به چرخش و تغییر مقیاس، نامتغییر و الگوی نمونه برداری آن مشابه با توزیع سلول‌های شبکه‌ی چشم انسان است.

دسته دوم بر تکنیک‌های یادگیری آماری برای محاسبه یک مدل آماری از وصله، متکی است. روشی که در [۱۰] ارائه شده از مدل ترکیبی گاوسی<sup>۷</sup> و آنالیز مؤلفه اصلی<sup>۸</sup> استفاده کرده اما برای اعوجاج ژرفانمایی راه حلی ارائه نکرده است. در [۶] روشی ارائه شده است که برای تشخیص ویژگی نقاط به صورت مستقل از ژست، از یک کلاسه‌بند درخت تصادفی [۳] استفاده می‌کند. همچنین در [۴] روشی معرفی شده که از یک کلاسه‌بند سلسله مراتبی تصادفی<sup>۹</sup> به جای درخت تصادفی استفاده کرده است و اثبات می‌کند که ساختار درختی استفاده شده در [۳] با یک ساختار سلسله مراتبی قابل جایگزینی است. هر ساختار سلسله مراتبی شامل یک مجموعه کوچک از آزمایش‌های دودویی است و احتمال عضویت هر وصله را در یکی از کلاس‌های یاد گرفته شده را مشخص

<sup>۵</sup> Fast retina keypoints (FREAK)

<sup>۶</sup> Scale Invariant Feature Transform (SIFT)

<sup>۷</sup> Gaussian Mixture Model (GMM)

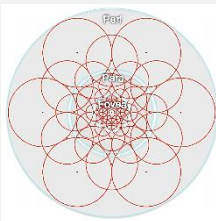
<sup>۸</sup> Principal Component Analysis (PCA)

<sup>۹</sup> Randomized Ferns

می‌کند. سپس از فرمول بیزین ساده<sup>۱۰</sup> برای ترکیب این احتمالات استفاده می‌شود. روش سلسله مراتبی تصادفی، از نظر قابلیت اطمینان با روش درخت تصادفی برابری می‌کند؛ اما روشی سریع‌تر بوده و برای پیاده‌سازی ساده‌تر است.

### ۳. روش پیشنهادی

با در نظر گرفتن تحقیقات انجام شده در این زمینه، راهکار ارائه شده ما مبتنی بر استفاده از کلاس‌بند سلسله مراتبی تصادفی [۴] برای بهبود دقت و سرعت توصیف گر نقاط کلیدی شبکه‌های سریع است. در این راه ابتدا بایستی بردار ویژگی مناسبی از هر نقطه کلیدی تهیه شود. بردار ویژگی تهیه شده نقش بسیار مهمی در دقت و عملکرد تطبیق گر نقاط کلیدی خواهد داشت؛ لذا ویژگی‌های استخراج شده از هر نقطه باید خصوصیات محلی آن را به خوبی مدل نماید. برای استخراج بردار ویژگی، ما از توصیف گر نقاط کلیدی شبکه‌های سریع استفاده کرده‌ایم که جهت چرخش نقاط کلیدی را برای ساخت بردار ویژگی در نظر می‌گیرد. به دلیل ماهیت دودویی بودن بردار ویژگی مذکور، این روش سریع و ساده است و به خوبی قابلیت استفاده در کاربردهای بلادرنگ را داراست. عملکرد مناسب بردارهای ویژگی دودویی در [۷] نشان داده شده است. همچنین روش مذکور از الگوی نمونه‌گیری الهام گرفته شده از چشم انسان با نرخ نمونه‌گیری بیشتر در مرکز استفاده می‌کند (شکل ۱). این موضوع باعث بهتر مدل شدن خصوصیات ظاهری نواحی اطراف نقاط کلیدی می‌شود. همچنین این روش در مقایسه با سایر توصیف گرهای محلی دودویی ارائه شده در این زمینه، بردارهای ویژگی با ابعاد کمتر و متمایزکنندگی بیشتر استخراج می‌کند.



شکل ۱- الگوی نمونه‌گیری توصیف گر نقاط کلیدی شبکه‌های سریع [۵]

هر کدام از دایره‌ها در شکل ۱ یک ناحیه نمونه‌برداری است که توسط یک فیلتر گاوسی هموار می‌شود. اندازه هر فیلتر گاوسی هموارساز<sup>۱۱</sup> متناسب با قطر دایره متناظر در الگوی نمونه‌گیری است. در این روش ابتدا جهت چرخش هر نقطه کلیدی را به دست آورده سپس بعد از اعمال فیلتر هموارساز، اختلاف بین جفت نقاط را محاسبه کرده و به وسیله آستانه‌گیری<sup>۱۲</sup>، بردار ویژگی<sup>۱۳</sup> F را می‌سازیم [۵]. به عبارت دیگر بردار ویژگی F یک رشته دودویی است که به وسیله دنباله‌ای از تفاوت گاوسی‌های<sup>۱۳</sup> یک بیتی ساخته می‌شود.

$$F = \sum_{0 \leq a < N} 2^a T(P_a) \quad (1)$$

در (۱) عبارت  $P_a$  یک جفت نقطه و  $N$  اندازه بردار ویژگی مورد بحث است. همچنین عبارت  $T(P_a)$  از رابطه (۲) محاسبه می‌شود.

$$T(P_a) = \begin{cases} 1 & \text{if } (I(P_a^{r_1}) - I(P_a^{r_2}) > 0, \\ 0 & \text{Otherwise,} \end{cases} \quad (2)$$

<sup>۱۰</sup> Naïve Bayes

<sup>۱۱</sup> Gaussian smoothing filter

<sup>۱۲</sup> Thresholding

<sup>۱۳</sup> Difference of Gaussians (DOG)



در فرمول (۲) عبارت  $I(P_a^{r_1})$  ناحیه هموار شده مربوط به نقطه اول در جفت نقاط مربوط به  $P_a$  است. با توجه به اینکه تعداد انتخاب جفت نقاط در الگوی نمونه‌برداری شده می‌تواند خیلی زیاد باشد و بیشتر آن‌ها برای توصیف تصویر بهینه نباشند؛ ما برای یادگیری بهترین جفت‌ها توسط داده‌های آموزشی از الگوریتمی مشابه با روش ORB [۸] استفاده کرده‌ایم. به منظور انتخاب زیر مجموعه‌ای مناسب از جفت نقاط، مراحل زیر را دنبال می‌کنیم:

- (۱) از حدود ۵۰۰۰۰ نقطه کلیدی استخراج شده، ماتریسی به نام  $D$  می‌سازیم که هر ردیف آن نشان دهنده بردار ویژگی یک نقطه کلیدی است که به وسیله تمام جفت نقاط موجود در الگوی نمونه‌گیری شکل ۱ ساخته شده است.
- (۲) سپس میانگین هر ستون را محاسبه کرده و برای به دست آوردن یک ویژگی متمایز کننده، بالا بودن واریانس آن مورد توجه است. میانگین نزدیک به ۰/۵ بیشترین واریانس یک توزیع دودویی را نتیجه می‌دهد.
- (۳) در مرحله بعد ستون‌ها را با توجه به بیشترین واریانس مرتب می‌کنیم.
- (۴) در مرحله آخر بهترین ستون‌ها (با میانگین ۰/۵) را نگه می‌داریم و از بین ستون‌های باقیمانده، مواردی که دارای کمترین همبستگی<sup>۱۴</sup> با ستون‌های انتخاب شده هستند را انتخاب می‌کنیم.

با انجام مراحل مذکور زیر مجموعه‌ای از بهترین جفت نقاط برای توصیف بهینه تصویر انتخاب می‌شود. این مسئله باعث می‌شود بردار ویژگی بهتری از مجموعه نقاط کلیدی انتخاب شود. همچنین می‌توان مجموعه تکه‌های امکان‌پذیر اطراف یک نقطه کلیدی را تحت تاثیر تغییرات ژرفانمایی، به عنوان یک کلاس در نظر گرفت. در [۴] نشان داده شده است که با آموزش یک کلاس‌بند سلسله مراتبی تصادفی می‌توان بر مسئله تطبیق نقاط کلیدی فائق آمد. مجموعه داده‌های آموزشی با اعمال تبدیلات هوموگرافی تصادفی بر روی تصویرهای وصله‌های محلی هر نقطه کلیدی ساخته می‌شود که برای آموزش کلاس‌بند سلسله مراتبی تصادفی استفاده می‌شود. کلاس‌بند مورد استفاده باید قادر باشد مسائلی با تعداد کلاس‌های زیاد (بیش از ۲۰۰ کلاس) را بدون کاهش محسوس دقت و سرعت، به خوبی دسته‌بندی نماید. ما در این راه از اطلاعات آماری بردار ویژگی توصیف گر نقاط کلیدی شبکه سریع مربوط به وصله‌های دیده شده از چشم اندازه‌های مختلف استفاده می‌کنیم. سپس بردار ویژگی را به گروه‌هایی از ویژگی‌ها بخش‌بندی کرده و بین گروه‌ها، شرط استقلال را در نظر می‌گیریم [۴].

همان‌طور که گفته شد مجموعه تمام ویژگی‌های ظاهری استخراج شده از تکه اطراف یک نقطه کلیدی را به عنوان یک کلاس در نظر می‌گیریم. بنابراین با داشتن تکه‌ای از اطراف یک نقطه کلیدی در یک تصویر، آن را به شبیه‌ترین کلاس انتساب می‌دهیم. اگر  $c_i, i = 1, \dots, H$  مجموعه تمام کلاس‌ها و  $f_j, j = 1, \dots, N$  بردار ویژگی نقاط کلیدی شبکه سریع مربوط به تکه مورد نظر جهت کلاس‌بندی باشد؛ آنگاه هدف ما یافتن کلاس بیشینه کننده عبارت (۳) است.

$$c_i^* = \operatorname{argmax}_{c_i} P(C = c_i | f_1, f_2, f_3, \dots, f_N) \quad (۳)$$

در فرمول (۳) عبارت  $C$  یک متغیر تصادفی است که نشان دهنده کلاس است. بر اساس فرمول بیز نتیجه می‌گیریم که:

$$P(C = c_i | f_1, f_2, f_3, \dots, f_N) = \frac{P(f_1, f_2, f_3, \dots, f_N | C = c_i)P(C = c_i)}{P(f_1, f_2, f_3, \dots, f_N)} \quad (۴)$$

در فرمول (۴) فرمول بیز استفاده شده را آورده‌ایم. فرض می‌کنیم که  $P(C)$  یک دانش مقدم یکنواخت است. از آنجایی که مخرج کسر، صرفاً یک ضریب مقیاسی و مستقل از کلاس است، مسئله ما به فرمول (۵) خلاصه می‌شود.

$$c_i^* = \operatorname{argmax}_{c_i} P(f_1, f_2, f_3, \dots, f_N | C = c_i) \quad (۵)$$

در فرمول (۵) هدف یافتن کلاس بیشینه کننده تابع  $P$  است.

طول بردار ویژگی نقاط کلیدی شبکه سریع ۵۱۲ است ( $N = 512$ ). بنابراین محاسبه احتمال توأم ممکن نیست؛ زیرا نیاز به تخمین زدن و ذخیره  $2^N$  مدخل برای هر کلاس دارد. یک راهکار خوب برای حل این مشکل بخش‌بندی بردار ویژگی به  $M$  گروه با اندازه  $S = N/M$  است. هر گروه به دست آمده یک Fern خوانده می‌شود و بایستی احتمال توأم به ازای هر Fern به صورت جداگانه محاسبه شود [۴]. بنابراین احتمال شرطی به صورت فرمول (۶) محاسبه می‌شود.

<sup>۱۴</sup> Correlation

$$P(f_1, f_2, f_3, \dots, f_N | C = c_i) = \prod_{k=1}^M P(F_k | C = c_i) \quad (6)$$

در فرمول (۶) عبارت  $F_k$  به Fern شماره  $k$  اشاره دارد و از فرمول (۷) محاسبه می‌شود.

$$F_k = \{f_{\sigma(k,1)}, f_{\sigma(k,2)}, \dots, f_{\sigma(k,S)}\}, \quad K = 1, 2, \dots, M \quad (7)$$

در فرمول (۷) عبارت  $\sigma(k, i)$  یک تابع جایگشت در بازه  $1 \dots N$  است. مراحل فوق نحوه ساخت بردار ویژگی نقاط کلیدی شبکه‌ی سریع مربوط به هر نقطه کلیدی و نحوه کار کلاسه‌بند سلسله مراتبی تصادفی را بیان می‌کنند.

در ادامه با بهره‌گیری از کلاسه‌بند سلسله مراتبی تصادفی عملیات تطبیق نقاط کلیدی را به انجام می‌رسانیم؛ بنابراین نیاز به آموزش کلاسه‌بند مذکور داریم. برای این منظور به یک تصویر از موضوع مورد نظر نیاز داریم که به آن، یک تصویر مدل گفته می‌شود. روال آموزش با انتخاب تصادفی زیر مجموعه‌ای از نقاط کلیدی تصویر مدل آغاز می‌شود و به هر کدام از این نقاط کلیدی انتخاب شده یک شماره کلاس منحصر به فرد انتساب داده می‌شود. برای تهیه مجموعه داده آموزشی ابتدا تعدادی کافی از تصاویر را با اعمال تابع Affine Deformation ایجاد می‌کنیم. پارامترهای تابع مذکور از یک توزیع یکنواخت به صورت تصادفی مقدار گرفته‌اند. سپس به تصاویر ساخته شده نویز گاوسی اضافه می‌کنیم و آن‌ها را با یک فیلتر گاوسی به ابعاد  $7 \times 7$  هموار می‌کنیم. نتایج نشان می‌دهند، این کار باعث افزایش پایداری کلاسه‌بند نسبت به نویز در زمان اجرا می‌شود. مجموعه به دست آمده را به عنوان داده آموزشی به کار می‌گیریم. سپس از نقاط کلیدی موجود در هر تصویر مجموعه داده آموزشی، بردار ویژگی نقاط کلیدی شبکه‌ی سریع را مطابق دستورالعمل ذکر شده، استخراج می‌کنیم. همان‌طور که در (۶) آمد، برای آموزش کلاسه‌بند بایستی احتمال شرطی هر کلاس  $P(F_m | C = c_i)$  را به ازای هر Fern که همان  $F_m$  است و هر کلاس  $c_i$  تخمین بزنیم. به ازای هر Fern ما عبارت‌های مذکور را به صورت (۸) بیان می‌کنیم.

$$P_{k,c_i} = P(F_m = k | C = c_i) \quad (8)$$

بنابراین هر Fern می‌تواند  $K = 2^S$  مقدار بگیرد و ما باید برای هر کدام از آن‌ها با توجه به قید (۹) مقدار  $P_{k,c_i}$  را تخمین بزنیم.

$$\sum_{k=1}^K P_{k,c_i} = 1 \quad (9)$$

در (۹) باید مجموع  $P_{k,c_i}$  برابر با ۱ باشد. ساده‌ترین راه برای تخمین زدن پارامترها به وسیله داده‌های آموزشی، استفاده از روش تخمین بیشترین احتمال<sup>۱۰</sup> است. برای پارامترهای  $P_{k,c_i}$  این تخمین به صورت (۱۰) محاسبه می‌شود.

$$P_{k,c_i} = \frac{N_{k,c_i}}{N_{c_i}} \quad (10)$$

در فرمول (۱۰) عبارت  $N_{k,c_i}$  برابر با تعداد نمونه‌های آموزشی برای Fern شماره  $k$  در کلاس  $c_i$  است و عبارت  $N_{c_i}$  تعداد کل نمونه‌های آموزشی برای کلاس  $c_i$  است. بنابراین این پارامترها می‌توانند به صورت مستقل برای هر Fern محاسبه شود.

#### ۴. تحلیل و ارزیابی

در این بخش الگوریتم ارائه شده را با استفاده از مجموعه داده استاندارد graffiti مورد ارزیابی قرار داده‌ایم. در شکل ۲، نمونه‌هایی از این مجموعه داده استاندارد که به عنوان تصویر مدل، جهت آموزش و آزمایش روش ارائه شده استفاده کرده‌ایم، آورده شده است. ما برای ساخت مجموعه داده آموزشی تعداد

<sup>۱۰</sup> Maximum likelihood estimation

دفعات مشخصی (۱۰۰ الی ۲۵۰۰۰)، affine deformation، را با پارامترهای تصادفی بر روی تصاویر مدل‌های مذکور اعمال کرده‌ایم و بدین ترتیب به ازای هر تصویر مدل تعداد مشخصی نمونه آموزشی ساخته‌ایم. سپس الگوریتم گوشه‌یاب Harris را بر روی تصاویر ساخته شده اعمال و ۴۰۰ نقطه کلیدی را به صورت تصادفی از هر تصویر انتخاب کرده‌ایم و یک شماره کلاس منحصر به فرد، به آنها اختصاص داده‌ایم. همچنین مرحله کلاسه‌بندی با استفاده از وصله‌هایی با ابعاد ۴۸×۴۸ پیکسل با تعداد ۴۲ Fern که هر کدام شامل ۱۲ ویژگی است انجام شده است.

ما برای هر دو فاز آموزش و آزمایش، تصاویر دگرپس شده را با استفاده از انتخاب کردن تصادفی پارامتر میزان چرخش تصویر در بازه  $[0: 2\pi]$  و پارامتر میزان تغییر مقیاس را در بازه  $[0.5: 1.5]$  ایجاد کرده‌ایم. در شکل ۳ وصله‌های اطراف برخی از نقاط کلیدی در تصاویر دگرپس شده را نمایش داده‌ایم. جهت افزایش اعتبار نتایج، به تصاویر دگرپس شده نویز گاوسی با میانگین صفر و واریانس زیاد (۲۵ برای شدت‌های خاکستری در محدوده ۰ تا ۲۵۵) افزوده‌ایم. سپس یک فیلتر هموارساز گاوسی با ابعاد ۷×۷ بر روی تصاویر آموزشی و آزمایشی اعمال کرده‌ایم. ما ۱۰۰۰ نمونه آزمایشی را با همان روش ساخت نمونه‌های آموزشی تهیه کرده‌ایم. بنابراین در بخش آزمایش، ما مکان نقاط کلیدی اصلی در تصویر مدل را بر اساس پارامترهای تابع دگرپس کننده منتقل کرده‌ایم و تکرارپذیری یابنده‌های نقاط کلیدی را در بخش آزمایش نادیده گرفته‌ایم و نرخ تشخیص صحیح را اندازه‌گیری کرده‌ایم.

برای ارزیابی الگوریتم ارائه شده، دقت روش ارائه شده را به ازای تعداد ۱۰۰ الی ۲۵۰۰۰ نمونه آموزشی (با ضریب گام ۱/۵) و همچنین به ازای تعداد ۱۰۰ الی ۲۵۰۰ کلاس (با گام ۱۰۰) در مقایسه با روش‌های نقاط کلیدی شبکه‌ی سریع و سلسله‌مراتبی تصادفی که به ترتیب در [۵] و [۴] ارائه شده‌اند، مورد ارزیابی قرار داده‌ایم. نتایج به دست آمده از این آزمایش‌ها به ترتیب در شکل ۴ و شکل ۵ آورده شده است. به دلیل عدم نیاز به بخش آموزش در روش نقاط کلیدی شبکه‌ی سریع، نمودار آن در شکل ۴ به صورت یک خط افقی دیده می‌شود.

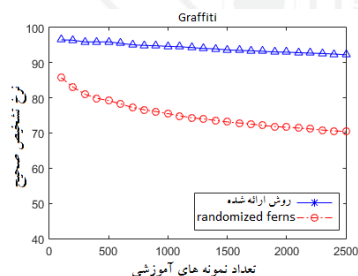


شکل ۳- نمونه‌هایی از وصله‌های دگرپس شده (هر ردیف مربوط به یک نقطه کلیدی)

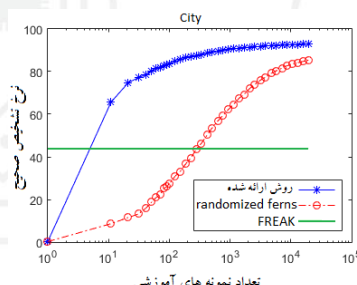


شکل ۲- تصاویر مدل استفاده شده در آموزش و آزمایش

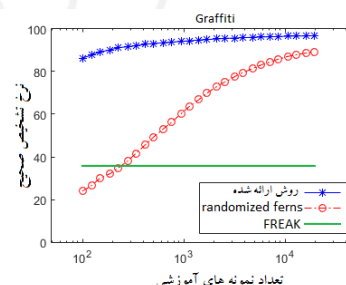
همان‌طور که در شکل ۴ مشاهده می‌شود نرخ تشخیص صحیح روش ارائه شده همواره بهتر از روش نقاط کلیدی شبکه‌ی سریع و روش سلسله‌مراتبی تصادفی بوده است؛ خصوصاً زمانی که تعداد نمونه‌های آموزشی کم باشد. دلیل این بهبود، استفاده از بردار ویژگی نقاط کلیدی شبکه‌ی سریع است که در مقایسه با بردار ویژگی پیش فرض الگوریتم سلسله‌مراتبی تصادفی، ویژگی ظاهری نقاط کلیدی را بهتر توصیف می‌کند. دلیل دیگر آن است که بردار ویژگی نقاط کلیدی شبکه‌ی سریع، پس از عملیات نرمال‌سازی جهت چرخش نقطه کلیدی، ساخته می‌شود؛ در حالی که در الگوریتم سلسله‌مراتبی تصادفی، جهت چرخش نقاط کلیدی در نظر گرفته نمی‌شود؛ لذا به تعداد نمونه‌های آموزشی بیشتری نیاز دارد. از جهتی دیگر ما از کلاسه‌بند Fern استفاده کرده‌ایم که در مقایسه با روش نزدیک‌ترین همسایه، که الگوریتم تطبیق پیش فرض در روش نقاط کلیدی شبکه‌ی سریع است دارای عملکرد بهتری است.



شکل ۵- مقایسه نرخ تشخیص صحیح به ازای تعداد کلاس‌های مختلف



شکل ۴- مقایسه نرخ تشخیص صحیح به ازای تعداد نمونه‌های آموزشی مختلف





همان‌طور که در شکل ۵ آورده شده است، با افزایش تعداد کلاس‌ها، نرخ تشخیص صحیح روش ارائه شده با شیب کمتری نسبت به روش سلسله مراتبی تصادفی کاهش پیدا می‌کند.

جدول ۱- مقایسه زمان اجرای الگوریتم ارائه شده با سایر روش‌ها

روش نقاط کلیدی شبکه سریع	روش سلسله مراتبی تصادفی	روش ارائه شده
۲۱۵۳۷۰	۵۳	۴۱
زمان تطبیق (نانوثانیه)		

در جدول ۱ مقایسه زمان اجرای مرحله تطبیق نقاط کلیدی روش ارائه شده با روش نقاط کلیدی شبکه سریع و سلسله مراتبی تصادفی آورده شده است.

### ۵. نتیجه‌گیری

در مقاله ارائه شده، روشی برای بهبود کارایی و دقت الگوریتم‌های نقاط کلیدی شبکه سریع و سلسله مراتبی تصادفی ارائه داده‌ایم. در این راه ما از بردار ویژگی نقاط کلیدی شبکه سریع برای استخراج ویژگی‌های مناسب از هر نقطه کلیدی بهره برده‌ایم. همچنین برای افزایش دقت تطبیق نقاط کلیدی از الگوریتم کلاس بندی Ferns استفاده کرده‌ایم. روش ارائه شده با مجموعه داده استاندارد مورد ارزیابی قرار داده و نتایج به دست آمده بهبود قابل توجهی را هم در زمینه دقت و هم در زمینه سرعت تطبیق نقاط نشان می‌دهد.

### ۶. مراجع

- [1] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
- [2] A. Böcker, S. Derksen, E. Schmidt, A. Teckentrup, and G. Schneider, "A Hierarchical Clustering Approach for Large Compound Libraries," *J. Chem. Inf. Model.*, vol. 45, no. 4, pp. 807–815, Jul. 2005.
- [3] Y. Amit, Geman. Donald "Shape Quantization and Recognition with Randomized Trees," *Neural Comput.*, vol. 9, no. 7, pp. 155–161, 1997.
- [4] M. Ozuysal, M. Calonder, V. Lepetit, and P. Fua, "Fast Keypoint Recognition Using Random Ferns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. Xx, pp. 1–14, 2010.
- [5] A. Alahi, R. Ortiz, and P. Vandergheynst, "FREAK: Fast Retina Keypoint," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 510–517.
- [6] V. Lepetit and P. Fua, "Keypoint recognition using randomized trees," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 9, pp. 1465–1479, Sep. 2006.
- [7] M. Calonder, V. Lepetit, M. Özuysal, T. Trzcinski, C. Strecha, and P. Fua, "BRIEF: Computing a local binary descriptor very fast," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 7, pp. 1281–1298, 2012.
- [8] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 2564–2571, 2011.
- [9] B. Brisk and O. R. B. Card, "Computer Vision Reading BRISK : Binary Robust Invariant Scalable Keypoints," pp. 1–8, 2011.
- [10] Li Fei-Fei, R. Fergus, and P. Perona, "One-shot learning of object categories," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 4, pp. 594–611, Apr. 2006.