

## Feature Extraction for Object Recognition Inspired by Human Visual System

H. Sufikarimi<sup>1</sup>, K. Mohammadi<sup>2\*</sup>

<sup>1</sup>Ph.D Student, Electronic School, Electrical Engineering Faculty, Iran University of Science and Technology, Tehran, Iran

<sup>2</sup> Professor, Electronic School, Electrical Engineering Faculty, Iran University of Science and Technology, Tehran, Iran

Receipt in the Online Submission System: 5 May 2018, Received in Revised Form: 6 June 2018, Accepted: 30 June 2018

---

### Abstract

In this paper, we tried to present a robust and reliable approach to object recognition by inspiring human visual system. A famous model, inspiring mammalian visual system, is HMAX (Hierarchical Model and X). It shows significant accuracy rates on object recognition tasks. However, there are some differences between this model and human visual system. Indeed cortex's functions are not properly modeled. Unrepeatability under fixed conditions, redundancy, high computing load and being slow are some drawbacks of HMAX. By modeling the secondary visual cortex and adding to the HMAX, we tried to introduce a more accurate model of the human visual system and cover the drawbacks of the previous models. The proposed approach functionally mimics the secondary visual cortex. Attending to high-level features, selecting discriminative and repeatable features, it has higher performance than standard HMAX. The added parts have negligible computation load. Therefore, it does not slow down this model. On the contrary, by selecting brief and useful features, the speed of the model is increased. The proposed approach is compared to the standard HMAX in terms of speed and accuracy rate. The results showed the advantage of proposed approach rather than the standard HMAX. In addition, the effect of the number of features and training images on their performance was shown. It is shown that the proposed approach has a better performance than the standard HMAX especially when the number of feature and training images is small.

**Keywords:** *Object Recognition, Feature Extraction, Biologically Inspired, Repeatability, HMAX*

---

#### Corresponding Author

**Address:** Electronic School, Electrical Engineering Faculty, Iran University of Science and Technology, P. O. Box: 168461-3114, Tehran, Iran  
**Tel:** +98-21-77240487  
**Fax:** +98-21-77240486  
**E-mail:** [mohammadi@iust.ac.ir](mailto:mohammadi@iust.ac.ir)

## استخراج ویژگی‌ها جهت بازشناسی اشیا با الهام از بینایی انسان

هیوا صوفی کریمی<sup>۱</sup>، کریم محمدی<sup>۲\*</sup>

<sup>۱</sup> دانشجوی دکترای مهندسی برق، گروه الکترونیک، دانشکده‌ی مهندسی برق، دانشگاه علم و صنعت ایران، تهران  
<sup>۲</sup> استاد، گروه الکترونیک، دانشکده‌ی مهندسی برق، دانشگاه علم و صنعت ایران، تهران

تاریخ ثبت در سامانه: ۱۳۹۷/۲/۱۵، بازنگری: ۱۳۹۷/۳/۱۶، پذیرش قطعی: ۱۳۹۷/۴/۹

### چکیده

در این مقاله سعی شده است تا با الگو برداری از سامانه‌ی بینایی انسان، یک روش مقاوم و تکرارپذیر برای بازشناسی اشیا ارائه شود. یکی از معروف‌ترین مدل‌های ارائه شده مبتنی بر بینایی انسان، مدل *HMAX* می‌باشد که عمل‌کرد مناسبی در بازشناسی اشیا از خود نشان داده است. اما تفاوت‌هایی نیز بین این مدل و بینایی انسان وجود دارد، به طوری که رویه‌ی مغز به طور کامل مدل نشده است. از جمله نواقص این مدل می‌توان به تکرارناپذیری (حتی در شرایط ثابت)، وجود افزونگی بسیار زیاد و در نتیجه حجم بالای محاسبات و کند بودن اشاره کرد. در این مقاله، سعی شده است تا با مدل‌کردن عمل‌کرد بخش ثانویه‌ی قشر بینایی و اضافه کردن آن به *HMAX*، مدل کامل‌تری از بینایی انسان ارائه گشته و نقاط ضعف مدل *HMAX*، پوشش داده شود. بخش اضافه شده، مانند بخش ثانویه‌ی قشر بینایی، با تمرکز روی ویژگی‌های سطح بالاتر و انتخاب ویژگی‌های متمایزکننده و البته تکرارپذیر، باعث بهبود یافتن عمل‌کرد مدل خواهد شد. بخش اضافه شده، بار محاسباتی بسیار اندکی داشته به طوری که نه تنها باعث کند شدن مدل نمی‌شود، بلکه با انتخاب ویژگی‌های مختصر و مفید، باعث افزایش سرعت نیز خواهد شد. روش پیشنهادی از لحاظ دقت و زمان پردازش با روش استاندارد مقایسه شده و برتری مدل پیشنهادی نشان داده شده است. علاوه بر آن، تاثیر تعداد ویژگی‌های استخراج شده و تعداد تصاویر مورد استفاده جهت آموزش، مورد بررسی قرار گرفته است تا برتری روش پیشنهادی، به ویژه در زمانی که تعداد تصاویر اندکی در دست می‌باشد، نشان داده شود.

کلیدواژه‌ها: بازشناسی اشیا، استخراج ویژگی، مدل زیست شناختی، تکرارپذیری، *HMAX*

\*نویسنده مسئول

نشانی: گروه الکترونیک، دانشکده‌ی مهندسی برق، دانشگاه علم و صنعت ایران، تهران، ایران، کد پستی: ۱۶۸۴۶۱-۳۱۱۴

تلفن: ۷۷۲۴۰۴۸۷ (۲۱) ۹۸+

دورنگار: ۷۷۲۴۰۴۸۶ (۲۱) ۹۸+

پست الکترونیکی: mohammadi@iust.ac.ir

## ۱- مقدمه

ساخته می‌شود. در انتهای این ساختار، یک دسته‌بند وجود دارد که با توجه به نتایج به دست آمده، وجود یا عدم وجود شی مورد نظر در تصویر ورودی را بررسی می‌کند.

در این مدل از قشر بینایی، ترکیبی متوالی از توابع فیلتری و یک تابع غیرخطی پیشینه‌گیری محلی وجود دارد. برای مدل کردن سلول‌های ساده، از توابع گبور و پایه‌ی شعاعی<sup>۹</sup>، و برای مدل کردن عمل کرد سلول‌های پیچیده، از عمل‌گر پیشینه‌گیری محلی استفاده شده است. در این مدل، سعی شده تا ساختار سلسله‌مراتبی بینایی انسان حفظ شود. روند پردازش در بینایی انسان به شیوه‌ی پیش‌خور می‌باشد و با جلو رفتن در مسیر پردازش، به تدریج میدان دریافت سلول‌ها افزایش می‌یابد. این خصوصیات باعث افزایش سرعت پردازش و مقاوم شدن نسبی آن در برابر تغییرات، از جمله جابه‌جایی و اندازه می‌شود.

در این مدل نیز، مانند دیگر فرایندهای<sup>۱۰</sup> یادگیری ماشین، داده‌ها و تصاویر خام ورودی، توسط تعدادی داده‌ی جدید که از داده‌های ورودی استخراج می‌شوند، بازتعریف می‌شوند. به این داده‌ی جدید، ویژگی<sup>۱۱</sup> گفته می‌شود. این بازتعریف به فرایند کمک می‌کند که با حذف داده‌ی غیرمفید و انتخاب اطلاعات اصلی، حجم محاسبات را کاهش داده و هم‌چنین دقت فرایند را نیز افزایش دهد. شیوه‌ی استخراج ویژگی‌ها، نقش بسیار مهمی در عمل‌کرد فرایند خواهد داشت. این ویژگی‌ها باید شامل اطلاعات مهم بوده و داده‌های ورودی را به بهترین نحو ممکن برای ماشین بازتعریف کنند. در HMAX بهترین قسمت‌های کوچکی از تصاویر ورودی به عنوان ویژگی، استخراج می‌شود، که این ویژگی‌ها تکه<sup>۱۲</sup> نامیده می‌شوند. در روش استاندارد، استخراج تکه‌ها به روش تصادفی انجام می‌شود که این روی‌کرد باعث ایجاد نقاط ضعفی خواهد شد که از آن جمله می‌توان به موارد زیر اشاره کرد:

- ۱- تکرارناپذیری، حتی در شرایط یک‌سان
- ۲- وجود افزونگی در الگوهای ذخیره شده
- ۳- حساس بودن به تغییر در چرخش
- ۴- بالا بودن بار محاسباتی و در نتیجه کند بودن فرایند برای بهبود و ارتقای مدل استاندارد تلاش‌های زیادی انجام گرفته است. برای مثال، در [۷] برای دسته‌بندی، به جای استفاده از بردارهای ماشین پشتیبان<sup>۱۳</sup>، از دستگاه یادگیری

بازشناسی اشیا، یکی از وظایف پرکاربرد در مبحث بینایی ماشین می‌باشد که توجه زیادی را به خود جلب کرده است. از کاربردهای آن می‌توان به سامانه‌های امنیتی، پردازش تصاویر پزشکی، سامانه‌های کمک‌رسان، کارهای صنعتی و غیره اشاره کرد. اما بازشناسی اشیا برای کامپیوتر و ماشین، وظیفه‌ای پرچالش و خطیر می‌باشد، زیرا شکل و ظاهر اشیا بر اثر عوامل محیطی تغییر کرده و وظیفه‌ی بازشناسی را برای ماشین دشوار می‌سازد. این تغییرات محیطی، شامل تغییر در روشنایی، چرخش، جابه‌جایی، اندازه و شکل ظاهری می‌باشد. تعدادی از روش‌های ارائه شده بر اساس محاسبات ریاضی می‌باشند که اگرچه نتایج خوبی از خود نشان داده‌اند، اما در برابر بسیاری از چالش‌ها ناتوان بوده [۱ - ۳] و عمل‌کرد آن‌ها با بینایی انسان فاصله‌ی زیادی دارد. سامانه‌ی بینایی انسان با وجود تمام تنوعی که از تغییر شرایط محیطی و ظاهر اشیا ناشی می‌شود، باز هم وظیفه‌ی بازشناسی اشیا را به راحتی و با دقت بالا انجام می‌دهد. به همین دلیل، انگیزه‌ی زیادی برای الگو برداری و ایده‌گیری از سامانه‌ی بینایی انسان وجود دارد. در این راستا، مقایسه‌ی خوبی بین مدل شناخته شده و پرکاربرد SIFT<sup>۱</sup>، به عنوان مدل مبتنی بر محاسبات ریاضی، و مدل مبتنی بر بینایی انسان، انجام شده است که به وضوح نشان می‌دهد مدل زیست شناختی نسبت به مدل ریاضی عمل‌کرد بهتری از خود نشان داده است [۴].

تحقیقات هابل<sup>۲</sup> و ویزل<sup>۳</sup> نشان داد که بینایی انسان یک ساختار سلسله‌مراتبی دارد که در آن با حرکت از لایه‌های اولیه‌ی بینایی به سمت لایه‌های بالاتر، میدان دریافت عصب‌ها<sup>۴</sup> افزایش می‌یابد، به گونه‌ای که عصب‌های IT تنها به الگوی خاصی پاسخ می‌دهند. معروف‌ترین مدل ارائه شده بر اساس بینایی انسان، توسط پوجیو<sup>۵</sup> و رایزنه‌بر<sup>۶</sup> معرفی شد [۵] و سپس بهبود یافته‌ی آن به عنوان HMAX<sup>۷</sup> استاندارد ارائه گردید [۶]. مدل HMAX دارای یک معماری سلسله‌مراتبی چهار لایه‌ای می‌باشد که از پشت سر هم قرار گرفتن تابع‌های فیلتری و تابع‌های غیرخطی، به نام پیشینه‌گیری محلی<sup>۸</sup>

<sup>۱</sup> Scale Invariant Feature Transform

<sup>۲</sup> Hubel

<sup>۳</sup> Wiesel

<sup>۴</sup> Neurons

<sup>۵</sup> Poggio

<sup>۶</sup> Riesenhuber

<sup>۷</sup> Hierarchical Model And X

<sup>۸</sup> Max-pooling

<sup>۹</sup> Radial Based Function(RBF)

<sup>۱۰</sup> Algorithm

<sup>۱۱</sup> Feature

<sup>۱۲</sup> Patch

<sup>۱۳</sup> Support Vector Machine(SVM)

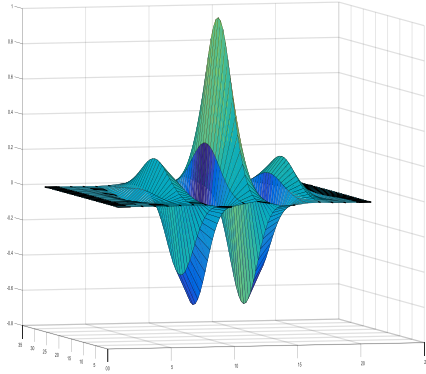
ویژگی انجام شده است. در این مدل، ابتدا چندین برابر تعداد ویژگی مورد نیاز، از تصاویر آموزشی استخراج می‌شود. سپس با استفاده از فرایند بیشینه‌کردن، اطلاعات متقابل آن‌ها را دسته‌بندی می‌کند و به تعداد مورد نیاز، بهترین ویژگی‌ها را انتخاب کرده و در هنگام یادگیری و بررسی تصاویر جدید، از آن‌ها استفاده می‌کند. پژوهش‌های دیگری نیز در زمینه‌ی استخراج ویژگی و تاثیر آن بر عمل‌کرد مدل استاندارد انجام شده است [۱۶-۱۸]. حجم کارهای انجام شده در زمینه‌ی استخراج ویژگی، نشان دهنده‌ی اهمیت و تاثیرگذاری روش استخراج ویژگی در عمل‌کرد مدل می‌باشد.

فرایندهای بازشناسی اشیاء، با استخراج اطلاعات تصویر و بازتعریف آن‌ها به عنوان ویژگی، کار بازشناسی شی را انجام می‌دهند. سعی بر آن است تا برای مقاوم‌سازی این فرایندها در برابر تغییرات شی از لحاظ هندسی (برای مثال تغییرات در مقیاس، چرخش، جابه‌جایی و روشنایی)، از ویژگی‌هایی استفاده شود که ذاتاً نسبت به تغییرات گفته شده مقاوم‌تر باشند. به این ویژگی‌ها، ویژگی‌های غیرتصادفی<sup>۲</sup> نیز می‌گویند. یکی از این ویژگی‌ها، گوشه‌ها و لبه‌های تصویر می‌باشد که نسبتاً مقاومت خوبی در برابر مقیاس، چرخش و روشنایی دارد. نتایج آزمایش‌های انجام گرفته روی انسان نیز نشان می‌دهد که گوشه‌ها و لبه‌ها نقش تعیین‌کننده‌ای در بازشناسی اشیاء در بینایی انسان دارد [۱۹]. در بینایی انسان، گوشه‌ها و لبه‌ها به عنوان ویژگی‌های اولیه، در قشر اولیه‌ی بینایی استخراج شده و در بخش ثانویه، محل تلاقی خطوط و گوشه‌ها استخراج می‌شوند. درحالی‌که این بخش از پردازش بخش بینایی انسان، در مدل HMAX در نظر گرفته نشده است. هدف این پژوهش، بهبود و ارتقای مدل HMAX، و ارائه‌ی یک مدل کامل‌تر از بینایی انسان می‌باشد. به همین منظور، یک لایه به مدل استاندارد اضافه شده، که عمل‌کرد قشر ثانویه‌ی بینایی انسان را مدل می‌کند. این بخش، نقش به‌سزایی در انتخاب ویژگی‌های مفید از تصاویر ورودی دارد. این هوشمندی در استخراج ویژگی باعث حذف افزونگی، تکرارپذیری فرایند و افزایش سرعت پردازش می‌شود. هم‌چنین تاثیر روش پیشنهادی بر دقت، تاثیر تعداد تصاویر مورد استفاده در آموزش و تعداد ویژگی‌های استفاده شده، مورد بررسی قرار داده می‌شود. یکی دیگر از ایرادهای عمده‌ی HMAX، تکرارناپذیری آن است که با استفاده از روش پیشنهادی، این ایراد نیز مرتفع شده است. علاوه بر مزایای ذکر شده، روش پیشنهادی با تعداد ویژگی

شدید<sup>۱</sup> استفاده شده است. اگرچه این روش نیز از ساختار پیش‌خور استفاده می‌کند و از روش‌های دارای پس‌خورد سریع‌تر است، اما به علت تصادفی بودن ماهیت آن، هم‌چنان مشکل تکرارناپذیری و قابلیت اطمینان پایین را دارد. علاوه بر آن از لحاظ سرعت، مزیت جندانی نسبت به روش قبلی ندارد، زیرا بخش دسته‌بندی، سهم بسیار اندکی (کم‌تر از ۱ درصد) از زمان محاسبات را به خود اختصاص داده است. بنابراین، بهبودهای این روش، جزئی بوده و مشکلات تکرارناپذیری و قابلیت اطمینان پایین هم‌چنان وجود دارد. در [۸] بعد از استخراج ویژگی‌ها، آن‌ها را با استفاده از روش معرفی دودویی محلی بازتعریف کرده‌اند. به این دلیل که بررسی میزان تطبیق در داده‌های دودویی، که مبتنی بر روش‌های آماری می‌باشد، سریع‌تر از بررسی میزان همبستگی (استفاده شده در HMAX استاندارد) است. در نتیجه، سرعت این روش از روش HMAX استاندارد بیشتر می‌باشد. ایراد این کار، عدم استفاده از اطلاعات محلی است که می‌تواند در بالا بردن دقت بسیار مفید باشد. در [۹] برای افزایش سرعت مدل از یک مرحله‌ی خوشه‌بندی استفاده شده است. این تغییر، سرعت مدل را در کارهایی که حجم آن‌ها بالا است بهبود می‌بخشد. البته این خوشه‌بندی باعث حذف بسیاری از اطلاعات خواهد شد که تاثیر زیادی در دقت مدل خواهد داشت. از این‌رو، افزایش سرعت به قیمت کاهش دقت تمام می‌شود. در [۱۰] سعی شده است تا با استفاده از نقشه‌ی اهمیت، نواحی مهم‌تر تصاویر پیدا شده و سپس از خوشه‌بندی استفاده شود. پس از آن، ویژگی‌های استخراج شده با نمونه‌های از قبل ذخیره شده در حافظه، مقایسه می‌شوند. در این کار، تمرکز روی مدل کردن لایه‌ی ۶ از بخش V1 می‌باشد. در [۱۱-۱۲] با تغییر بانک فیلتری در اولین لایه‌ی HMAX، سعی شده است تا دقت این مدل افزایش داده شود. اما نتایج به‌دست آمده، بهبود اندکی را در دقت و عمل‌کرد نشان می‌دهد. در [۱۳] تاثیر تمرکز روی بخش‌های خاصی از تصویر، مورد بررسی قرار گرفته است و با اضافه کردن این راه‌کار به مدل استاندارد، عمل‌کرد آن بهبود داده شده است. در [۱۴] سعی شده است تا بعد از استخراج ویژگی، میزان تشابه بین ویژگی‌های استخراج شده بررسی شود و ویژگی‌هایی که تشابه زیادی با یک‌دیگر دارند، حذف شوند. بدین وسیله، برای افزایش عمل‌کرد فرایند و بالا بردن سرعت آن، بدون کاهش دقت، تلاش کرده‌اند. در [۱۵] تغییراتی در بخش استخراج و انتخاب

<sup>۲</sup> Non Accidental Properties(NAP)<sup>۱</sup> Extreme Learning Machine

در این رابطه،  $\theta$  معرف جهت فیلتر،  $\lambda$  طول موج،  $\sigma$  پهنای باند موثر و  $\gamma$  ضریب کروی بودن می‌باشد. این پارامترها به شیوه‌ی تجربی و به گونه‌ای تعیین شده‌اند که بیش‌ترین شباهت را به بینایی انسان داشته باشند [۶].



شکل (۱) - یک مثال از فیلتر گبور

### ۲-۱-۲- لایه‌ی C1

در قشر اولیه‌ی بینایی بعد از سلول‌های ساده، سلول‌های پیچیده قرار دارند. این سلول‌ها دارای میدان دریافت بزرگ‌تری نسبت به لایه‌ی قبل از خود بوده و نسبت به تغییرات جابه‌جایی و مقیاس، از خود مقاومت نشان می‌دهند. این عمل‌کرد، با استفاده از یک عمل‌گر غیرخطی مدل شده است. این عمل‌گر، یک بیشینه‌یاب محلی بوده، که به طور هم‌زمان مقدار بیشینه را در یک پنجره‌ی محلی در دو مقیاس کناری پیدا می‌کند و ابعاد تصویر را کاهش می‌دهد. این عمل‌گر علاوه بر افزایش مقاومت در برابر تغییرات جابه‌جایی و مقیاس، با کاهش ابعاد تصویر به کاهش بار محاسباتی نیز کمک می‌کند.

### ۲-۱-۳- انتخاب ویژگی

این مرحله تنها در زمان آموزش اجرا می‌شود و در زمان آزمون حذف می‌گردد. از تصاویر به دست آمده از مرحله‌ی C1 تعدادی ویژگی جهت معرفی تصاویر آموزشی انتخاب می‌شوند. ویژگی‌های انتخاب شده، شامل تکه‌هایی از تصاویر با ابعاد  $4 \times 4$ ،  $8 \times 8$ ،  $12 \times 12$  و  $16 \times 16$  می‌باشد. در HMAX استاندارد از هر کدام از ابعاد، ۲۵۰ تکه به‌عنوان ویژگی استخراج می‌شود. بنابراین، ۱۰۰۰ تکه برای بازتعریف شی در روند پردازش، استخراج می‌شود. این ویژگی‌ها هم در زمان آموزش و هم در زمان آزمون مورد استفاده قرار می‌گیرند. این ویژگی‌ها به روش تصادفی از تصاویر آموزشی استخراج می‌شوند. قابل ذکر است که ماهیت تصادفی این انتخاب باعث ایجاد افزونگی و انتخاب ویژگی‌های نامناسب خواهد شد.

بسیار کم‌تری قادر به بازشناسی تصاویر خواهد بود. در مدل پیشنهادی، سعی شده است تا کمبودهای مدل HMAX نسبت به بینایی انسان برطرف گردد.

در بخش ۲، به معرفی مدل HMAX استاندارد و روش پیشنهادی این مقاله می‌پردازیم. بخش ۳، به بررسی نتایج و تحلیل آن‌ها، و بخش ۴، نیز به نتیجه‌گیری اختصاص دارد.

## ۲- مواد و روش‌ها

در ابتدا به معرفی اجمالی مدل HMAX می‌پردازیم. سپس مدل پیشنهادی این مقاله را ارائه می‌دهیم.

### ۲-۱- معرفی HMAX استاندارد

مدل ارائه شده در [۶]، که به عنوان مدل استاندارد معرفی شده است، دارای چهار لایه‌ی متوالی و یک دسته‌بند می‌باشد. این ساختار، مانند بینایی انسان، یک ساختار سلسله‌مراتبی پیش‌خور<sup>۱</sup> است، تا بدین طریق سرعت پردازش و انتقال اطلاعات پردازش‌شده در هر لایه به لایه‌ی بعدی به درستی مدل شده باشد. طبق یافته‌های هابل و ویزل، قشر بینایی انسان از دو نوع سلول ساده و پیچیده تشکیل شده است. در مدل HMAX سلول‌های ساده با یک تابع فیلتری و سلول‌های پیچیده با استفاده از یک تابع غیرخطی، که وظیفه‌ی آن محاسبه‌ی بیشینه‌ی محلی می‌باشد، مدل شده است.

### ۲-۱-۱- لایه‌ی S1

تصویر دریافت شده در شبکیه‌ی چشم، بعد از عبور از بخش  $LGN^2$  وارد بخش اولیه‌ی بینایی می‌شود. در این بخش، اطلاعات اولیه‌ی تصویر، از قبیل خطوط و لبه‌ها، استخراج می‌شود. برای مدل کردن این توانایی قشر بینایی، یک لایه در نظر گرفته شده است. این بانک فیلتری، شامل ۶۴ فیلتر متفاوت می‌باشد. این فیلترها در ۱۶ مقیاس و ۴ جهت متفاوت دسته‌بندی می‌شوند. مقیاس فیلترها از  $7 \times 7$  تا  $37 \times 37$  بوده و زوایای فیلتری ۰، ۴۵، ۹۰ و ۱۳۵ درجه می‌باشند. فیلتر گبور از ترکیب یک تابع گوسی و یک تابع سینوسی به دست می‌آید (شکل ۱). روابط این فیلتر به صورت زیر تعریف می‌شود [۶]:

$$F(x, y) = \exp\left(-\frac{(x_o^2 + \gamma^2 y_o^2)}{2\sigma^2}\right) \times \cos\left(\frac{2\pi}{\lambda} x_o\right) \quad (1)$$

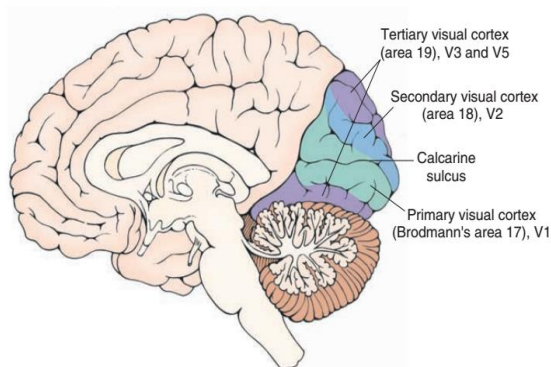
$$\begin{aligned} x_o &= x \cos \theta + y \sin \theta \\ y_o &= -x \sin \theta + y \cos \theta \end{aligned} \quad (2)$$

<sup>۱</sup> FeedForward

<sup>۲</sup> Lateral geniculate nucleus

## ۲-۲- مدل پیشنهادی این مقاله

مدل HMAX الگوی نسبتاً خوبی از بینایی انسان ارائه داده است، اما تفاوت‌هایی نیز بین این مدل و سامانه‌ی بینایی انسان وجود دارد. در مدل HMAX استاندارد، از یک روش تصادفی برای استخراج ویژگی استفاده می‌شود، که با عمل کرد بینایی انسان متناقض می‌باشد. برای مثال، در سامانه‌ی بینایی انسان، قشر ثانویه مسئول استخراج اطلاعات مفید از گوشه‌ها و نقاط برجسته‌ی تصویر می‌باشد (شکل ۳). این نقاط از اهمیت بالایی برخوردار بوده و در بسیاری از موارد با حذف آن‌ها از تصویر، سامانه‌ی بینایی انسان قادر به بازشناسی آن شی نمی‌باشد. در حالی که در HMAX، این بخش مدل نشده و مکانیزمی برای استخراج این ویژگی‌های مهم در نظر گرفته نشده است. توجه به نقاط مهمی چون گوشه‌های تصویر، علاوه بر بالا بردن دقت بازشناسی، در بالا رفتن تغییرناپذیری فرایند نسبت به تغییرات ناشی از چرخش و مقیاس نیز موثر می‌باشد.



شکل (۳) - بخش اولیه و ثانویه بینایی در قشر بینایی [۲۴]

استفاده از روند تصادفی در استخراج ویژگی، نقاط ضعیفی را برای HMAX به همراه دارد، که تعدادی از آن‌ها عبارتند از:

- ۱- تکرارناپذیری، حتی در شرایط یک‌سان: از آنجایی که ویژگی‌ها به روش تصادفی انتخاب می‌شوند، احتمال انتخاب مجدد آن‌ها و به دست آوردن همان نتایج در اجرای مجدد برنامه، پایین می‌باشد.
- ۲- وجود افزونگی در الگوهای استخراج شده: وقتی انتخاب ویژگی حالت تصادفی دارد، احتمال انتخاب نقاط تکراری و یا انتخاب بخش‌هایی که هیچ اطلاعات مفیدی جهت بازشناسی شی ارائه نمی‌کنند، بسیار بالا می‌باشد.
- ۳- حساس بودن به تغییرات چرخشی: از آنجا که فرایند از روش فاصله‌ی اقلیدسی برای بررسی میزان تطبیق الگوها استفاده می‌کند، بنابراین نسبت به تغییرات چرخش، بسیار حساس می‌باشد.

## ۲-۱-۴- لایه‌ی S2

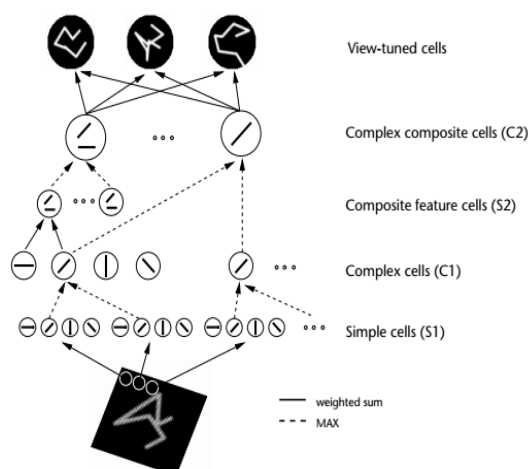
در بخش ثانویه بینایی، پردازش‌های سطح بالاتر از قبیل تطبیق الگو، انجام خواهد شد. در این لایه، از تابعی مشابه تابع پایه‌ی شعاعی، جهت بررسی انطباق الگوهای ذخیره شده با تصویر جدید استفاده می‌شود [۶]. این رابطه به صورت زیر تعریف می‌شود:

$$TM(x) = \exp(-\|x - P_i\|) \quad (۳)$$

که در آن  $x$  موقعیت در تصویر و  $P_i$  الگوی ذخیره شده می‌باشد. قابل ذکر است که بررسی تطبیق هر کدام از الگوهای ذخیره شده، با تمام تصاویر به دست آمده از C1 در کلیه‌ی زوایا و مقیاس‌ها انجام می‌شود. واضح است که حجم محاسباتی این لایه بسیار بالا می‌باشد، به گونه‌ای که تقریباً ۹۷ درصد از زمان پردازش، مربوط به این لایه می‌باشد [۲۰].

## ۲-۱-۵- لایه‌ی C2

در لایه‌ی S2، هر الگو با تمام ۳۲ تصویر حاصل از مرحله‌ی C1 مقایسه می‌شود. این مقایسه، به روش محاسبه‌ی فاصله‌ی اقلیدسی در هر نقطه از تصویر می‌باشد. نتایج حاصل از این مقایسه، تعدادی عدد است که نشان‌دهنده‌ی میزان تشابه الگو با تصویر در آن نقطه می‌باشد. از بین این نتایج، بهترین عدد به عنوان بیش‌ترین میزان تشابه الگو و تصویر انتخاب می‌شود. این عمل برای تمام الگوها تکرار می‌شود، بنابراین در خروجی این لایه، به تعداد الگوها، عدد اسکالر خواهیم داشت که معرف میزان تشابه تصویر جدید با الگوهای ذخیره شده می‌باشد. این اعداد برای بازشناسی شی به یک دسته‌بند ماشین بردار پشتیبان ارسال می‌شوند (شکل ۲).



شکل (۲) - ساختار سلسله‌مراتبی بینایی ساخته شده از سلول‌های

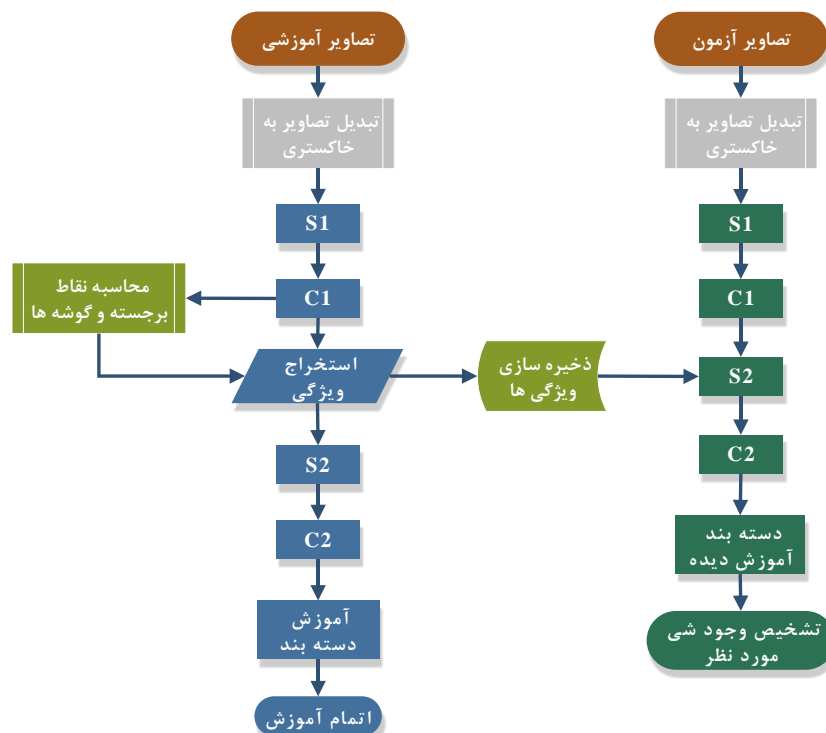
ساده و پیچیده [۲۳]

شوند، باز هم سامانه‌ی بینایی با احتمال زیادی قادر به بازشناسی آن خواهد بود. درحالی‌که اگر گوشه‌ها حذف و بقیه‌ی اجزا نگه داشته شوند، احتمال بازشناسی آن بسیار اندک خواهد بود.

در مدل HMAX، برای حفظ سرعت فرایند و هم‌چنین الگوبرداری از بینایی انسان، از یک ساختار پیش‌خور استفاده شده است، که در آن نتایج هر لایه به مرحله‌ی بعدی ارسال گردیده و از پس‌خورد استفاده نشده است. در مدل پیشنهادی این ویژگی حفظ شده و مسیر ورودی و خروجی داده در لایه‌ی اضافه شده در راستای حرکت پردازشی داده‌ها می‌باشد. همان‌گونه که در شکل (۴) نشان داده شده است، مراحل آموزش و آزمون بسیار شبیه به هم بوده و تنها در دو بخش با هم تفاوت دارند. اولین تفاوت در لایه‌ی بعد از لایه‌ی C1 می‌باشد. در زمان آموزش، بعد از لایه‌ی C1 ویژگی‌ها از تصاویر استخراج می‌شوند، درحالی‌که در مرحله‌ی آزمون استخراج ویژگی نخواهیم داشت، بلکه از ویژگی‌های استخراج شده در مرحله‌ی آموزش استفاده می‌کنیم. دومین تفاوت در مرحله‌ی بعد از لایه‌ی C2 می‌باشد. این مرحله مربوط به دسته‌بند می‌باشد. در مرحله‌ی آموزش، دسته‌بند آموزش داده می‌شود، درحالی‌که در مرحله‌ی آزمون، از دسته‌بند آموزش‌دیده در مرحله‌ی آموزش استفاده می‌شود.

۴- بالا بودن بار محاسباتی: همان‌گونه که اشاره شد، فرایند در زمینه‌ی استخراج ویژگی دچار افزونگی خواهد شد که بار محاسباتی بسیار بالایی به آن تحمیل می‌کند و در نتیجه عمل‌کرد فرایند بسیار کند خواهد بود.

استخراج ویژگی، یکی از مهم‌ترین مراحل در بازشناسی می‌باشد. ویژگی‌ها باید به گونه‌ای انتخاب شوند که تصاویر را به بهترین نحو معرفی نمایند. به عبارت دیگر، ویژگی‌ها باید تمایزدهنده باشند. علاوه بر آن، روند استخراج ویژگی‌ها باید تکرارپذیر بوده، تا نتایج حاصله تکرارپذیر و قابل اعتماد باشند. برای گریز از این مشکل، پیشنهاد می‌شود که ویژگی‌ها توسط یک روش تکرارپذیر استخراج شوند. به همین منظور لازم است گوشه‌ها و نقاط برجسته‌ی تصویر به‌عنوان ویژگی استخراج شوند. این کار علاوه بر کمک به تکرارپذیر بودن فرایند، تا حدود زیادی از افزونگی نیز جلوگیری خواهد کرد. زیرا هر نقطه، تنها یک بار انتخاب می‌شود. درحالی‌که در حالت تصادفی ممکن است یک نقطه‌ی خاص به دفعات زیادی انتخاب شود. افزونگی باعث افزایش بار محاسباتی و هم‌چنین احتمال بیش‌برازش<sup>۱</sup> فرایند خواهد شد. تحقیقات نشان می‌دهد که سامانه‌ی بینایی انسان، حساسیت بیش‌تری نسبت به گوشه‌های تصویر از خود نشان می‌دهد، به گونه‌ای که اگر گوشه‌های یک شی موجود باشد و بقیه‌ی اجزای آن شی حذف



شکل (۴) - طرح مراحل آموزش و آزمایش در مدل پیشنهادی

<sup>۱</sup> Overfit

### ۳- یافته‌ها و بحث

برای بررسی و سنجش صحت روش، از پایگاه داده‌ی Caltech5، متشکل از پنج دسته از اشیا (هواپیما، برگ، خودرو، چهره و موتورسیکلت) و پایگاه GRAZ-02، متشکل از سه دسته از اشیا (دوچرخه، خودرو و انسان) استفاده شده است. در این پایگاه‌های داده، کلیه تصاویر رنگی می‌باشند، درحالی‌که تصاویر ورودی به مدل باید خاکستری باشند. به همین دلیل در اولین مرحله، تصاویر از حالت رنگی به خاکستری تبدیل می‌شوند. مساله‌ی دسته‌بندی به حالت دو دسته‌ای تعریف شده است. به همین دلیل تصاویر آموزشی و آزمون، خود به دو دسته از تصاویر مثبت و منفی تقسیم می‌شوند. دسته‌ی مثبت شامل تصاویری می‌باشد که شی مورد نظر در آن قرار دارد و دسته‌ی منفی به تصاویری گفته می‌شود که در آن تصویر شی مورد نظر قرار ندارد.

#### ۳-۱- نتایج

در این بخش، نتایج حاصل از بررسی روش پیشنهادی و روش استاندارد ارائه خواهد شد. مقایسه‌ها در رابطه با دقت دسته‌بندی و زمان محاسبات تحت شرایط متفاوت انجام گرفته است.

#### ۳-۱-۱- بررسی دقت دسته‌بندی

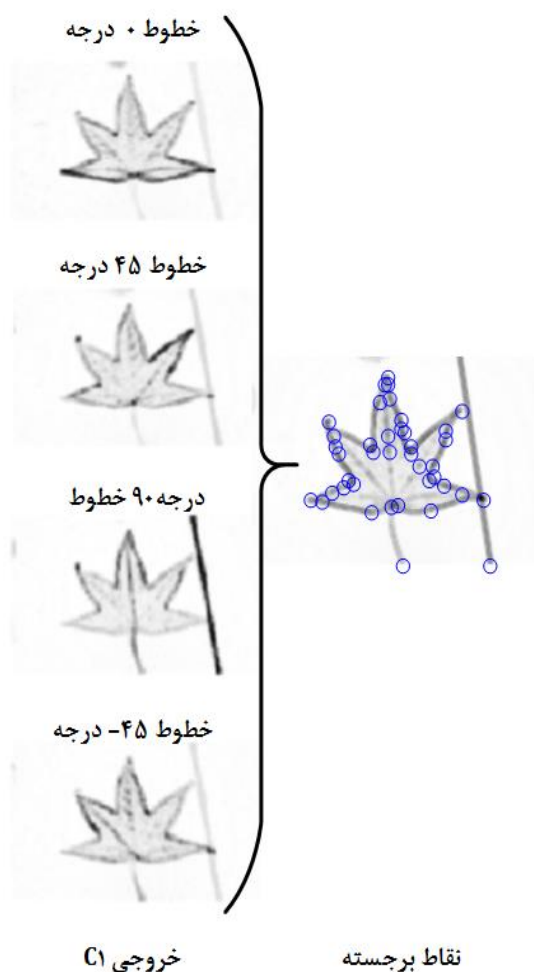
در ابتدا عمل‌کرد کلی دو روش، تحت شرایط یکسان مورد بررسی قرار گرفته است. در این بررسی از تعداد ۱۵ تصویر مثبت و ۱۵ تصویر منفی در زمان آموزش استفاده شده و برای آزمون، ۳۰ تصویر منفی و ۳۰ تصویر مثبت در نظر گرفته شده است. برای تصاویر منفی، از تصاویر پس‌زمینه<sup>۱</sup> استفاده شده است. به ازای هر تصویر آموزشی، یک تکه، و در مجموع تعداد ۱۵ تکه به عنوان ویژگی استخراج شده است. ابعاد این تکه‌ها ۸×۸، ۱۲×۱۲ و ۱۶×۱۶ می‌باشد. هرکدام از آزمون‌ها پنج بار تکرار شده و میانگین آن‌ها برای مقایسه در نظر گرفته شده است.

جدول (۱) - مقایسه‌ی دقت دسته‌بندی در روش استاندارد و روش پیشنهادی

روش پیشنهادی	روش استاندارد	دقت بازشناسی (درصد)
۹۱/۶۷	۸۶/۶۷	هواپیما
۹۳/۳۳	۹۱/۶۷	خودرو
۹۳/۳۳	۹۱/۶۷	موتورسیکلت
۹۸/۳۳	۹۸/۳۳	چهره
۹۶/۶۷	۹۳/۳۳	برگ

در روش پیشنهادی، یک لایه به مدل HMAX اضافه شده است که کار بخش ثانویه‌ی بینایی را انجام می‌دهد. این لایه نقش محاسبه‌ی نقاط برجسته، از جمله نقاط تلاقی خطوط را بر عهده دارد. اطلاعات به دست آمده به مرحله‌ی استخراج ویژگی ارسال می‌شود تا ویژگی‌های مفید از این نقاط انتخاب شوند.

لبه‌های تصویر به کمک تصاویر به دست آمده از مرحله‌ی S1 برجسته می‌شود. در مرحله‌ی S1، تعداد ۶۴ تصویر که معرف جهت خطوط تصویر می‌باشند، به دست می‌آید. از آنجا که گوشه‌های تصویر، محل تلاقی خطوط می‌باشند، در نتیجه در نقاط تلاقی، حداقل دو جهت متفاوت دارای مقدار قابل ملاحظه می‌باشند. با در نظر گرفتن این نکته، سعی در پیدا کردن گوشه‌های تصویر شده است. برای انجام این کار، تصاویر به دست آمده را با یکدیگر تلفیق نموده و مقادیر بیشینه را محاسبه می‌نمائیم (شکل ۵).



شکل (۵) - نمونه‌ای از تکه‌های استخراج شده به عنوان ویژگی

<sup>۱</sup> Background



و ۱۵ تصویر منفی در مرحله‌ی آموزش مورد استفاده قرار گرفت. در مرحله‌ی آزمایش از ۳۰ تصویر مثبت و ۳۰ تصویر منفی استفاده شد. از لحاظ زمان محاسبات، روش پیشنهادی به طور میانگین ۲۰ درصد سریع‌تر از روش استاندارد بود. دقت دسته‌بندی در جدول (۳) آورده شده است. در این آزمون نیز دقت و سرعت روش پیشنهادی از روش استاندارد بالاتر بوده است. قابل ذکر است که این پایگاه داده، به علت وجود عوامل پس‌زمینه و عوامل ناخواسته، دارای چالش بزرگ‌تری بوده و دقت مدل در آن پایین‌تر از پایگاه داده‌ی Caltech5 می‌باشد.

#### جدول (۳) - مقایسه‌ی دقت دسته‌بندی در روش پیشنهادی با

روش‌های انتخاب ویژگی در پایگاه داده‌ی GRAZ-02

دقت بازشناسی	دوچرخه	خودرو	انسان
روش استاندارد	۷۳/۳۳	۷۶/۶۷	۸۰
روش پیشنهادی	۸۳/۳۳	۷۸/۶۸	۸۳/۳۳

#### ۳-۱-۲- بررسی تاثیر تعداد تکه‌ها بر دقت و زمان محاسبات

در این آزمون سعی شده است تا اثر تعداد ویژگی‌های مورد استفاده جهت بازتعریف شی، روی دقت و زمان محاسبات، مورد ارزیابی قرار گیرد. روش استاندارد به شدت وابسته به تعداد ویژگی‌های اشیا می‌باشد و هنگامی که تعداد ویژگی‌های مورد استفاده کم باشند، عمل‌کرد آن دچار افت زیادی خواهد شد. علت این امر، تصادفی بودن روش استخراج این ویژگی‌ها می‌باشد. در استخراج تصادفی ویژگی، ممکن است قسمت‌هایی از تصویر به عنوان ویژگی استخراج شوند که هیچ ارزش اطلاعاتی برای فرایند نداشته باشند، و یا این‌که تعداد زیادی از آن‌ها تکراری بوده و دچار افزونگی شده باشد. واضح است که با افزایش تعداد ویژگی‌ها، این نقص تا حدودی جبران شده و دقت دسته‌بندی افزایش خواهد یافت. اما این افزایش تعداد ویژگی، بار محاسباتی بسیار زیادی را به فرایند تحمیل می‌کند و باعث کندی آن می‌شود. بنابراین روشی که بتواند با تعداد ویژگی کم، تصاویر را به بهترین نحو بازتعریف کند، از امتیاز ویژه‌ای برخوردار خواهد بود.

در مقابل، روش پیشنهادی در هنگامی که تعداد ویژگی‌های کمی وجود داشته باشد، عمل‌کرد به مراتب بهتری از خود نشان می‌دهد، که دلیل آن انتخاب هوشمندانه‌ی ویژگی‌ها است. روش پیشنهادی سعی می‌کند تا نقاط مهم تصویر را پیدا کرده و به ترتیب اهمیت، از آن‌ها انتخاب کند. بنابراین حتی با تعداد ویژگی‌های کم، دقت بسیار خوبی از خود نشان

همان‌گونه که در جدول (۱) نشان داده شده است، غیر از دسته‌ی "چهره" که دقت در هر دو روش برابر می‌باشد، در تمام موارد دیگر، دقت روش پیشنهادی بالاتر از روش استاندارد بوده است.

زمان محاسبات به طور میانگین در روش پیشنهادی ۲۱۰۰ ثانیه و در روش استاندارد ۲۵۰۰ ثانیه بوده است، که نشان می‌دهد روش پیشنهادی ۱۶ درصد از روش استاندارد سریع‌تر بوده است. مزیت روش پیشنهادی نسبت به روش استاندارد از لحاظ سرعت، هنگامی هویدا می‌شود که تعداد تصاویر و تعداد ویژگی‌های استخراج شده از آن‌ها، زیاد باشد. در این حالت روش پیشنهادی به وضوح سریع‌تر خواهد بود. اما هنگامی که تعداد ویژگی‌ها کم باشد، روش پیشنهادی کمی از روش استاندارد سریع‌تر خواهد بود. در بخش بعدی تاثیر تعداد ویژگی‌ها بر زمان محاسبات و دقت، به تفصیل بررسی شده است.

روش پیشنهادی هم‌چنین با روش‌های انتخاب ویژگی [۱۴]، GMAX [16] و B-MAX [8] نیز مقایسه شده است که نتایج به دست آمده در جدول (۲) نشان داده شده است. در دو روش انتخاب ویژگی و GMAX، با تمرکز بر روند استخراج ویژگی، سعی شده است تا دقت و سرعت فرایند بهبود داده شود. روش پیشنهادی نتایج بهتری نسبت به این روش‌ها نشان داده است. اما در روش B-MAX، نتایج مربوط به دو دسته‌ی "هواپیما" و "خودرو" از روش پیشنهادی بهتر بوده، اما در باقی دسته‌ها روش پیشنهادی نتایج بسیار بهتری از خود نشان داده است. البته قابل ذکر است که در هر سه روش، تمرکز روی دقت و سرعت بوده و تکرارپذیری و قابلیت اطمینان، مورد بررسی قرار نگرفته است.

#### جدول (۲) - مقایسه‌ی دقت دسته‌بندی در روش پیشنهادی با

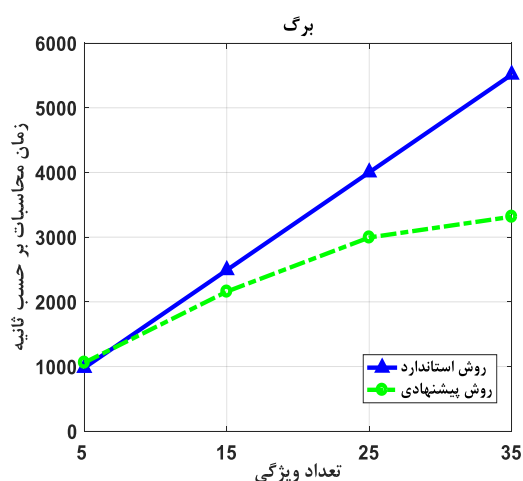
روش‌های انتخاب ویژگی، GMAX و B-MAX

در پایگاه داده‌ی Caltech5

دقت بازشناسی	روش پیشنهادی	GMAX	انتخاب ویژگی	B-MAX
هواپیما	۹۱/۶۷	۸۷	۹۰	۹۵
خودرو	۹۳/۳۳	۹۰	۹۰/۵	۹۸
موتورسیکلت	۹۳/۳۳	۹۳	۹۱/۶۷	۸۸
چهره	۹۸/۳۳	۹۲	۹۳/۳۳	۹۱
برگ	۹۶/۶۷	۸۵	۹۱/۶۷	۸۲

عمل‌کرد روش پیشنهادی روی پایگاه داده‌ی GRAZ-02 نیز مورد بررسی قرار گرفت. در این آزمون، تعداد ۱۵ تصویر مثبت

شده باشند. همان‌گونه که در شکل (۷) مشاهده می‌شود، این مرحله برای آزمون انجام شده روی "برگ"، در حدود ۳۵ ویژگی به ازای هر تصویر به همراه دارد.



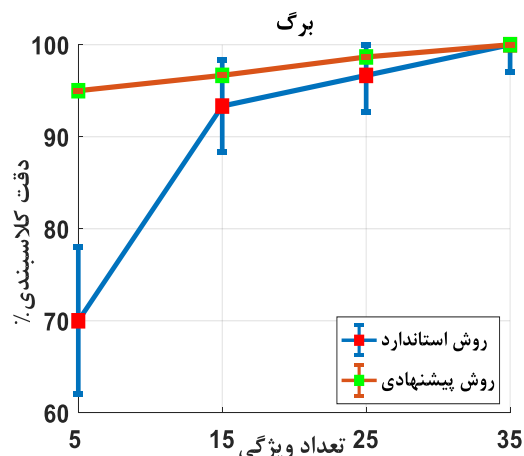
شکل (۷) - مقایسه‌ی تاثیر تعداد ویژگی‌ها بر زمان محاسبات در روش استاندارد و روش پیشنهادی

### ۳-۱-۳- تاثیر افزایش تصاویر آموزشی بر دقت

یکی از پارامترهای مهم در آموزش فرایند، تعداد تصاویر مورد استفاده در مرحله‌ی آموزش می‌باشد. این امر که فرایند بتواند با تعداد تصاویر کم‌تری آموزش ببیند، یک مزیت به حساب می‌آید [۲۱]. نتایج نشان می‌دهد که روش استاندارد نسبت به کم بودن تعداد تصاویر آموزشی بسیار حساس بوده و عمل کرد آن به شدت افت می‌کند. علت این امر نیز مربوط به تصادفی بودن استخراج ویژگی‌ها است. اگر تعداد تصاویر کم باشد و فرایند از روش تصادفی برای استخراج ویژگی استفاده کند، احتمال استخراج ویژگی‌های مربوط به پس‌زمینه و ویژگی‌های مربوط به هدف، برابر خواهد بود. بنابراین بازشناسی صحیح شی کاملاً تصادفی خواهد بود. همان‌گونه که در شکل (۸) و (۹) قابل مشاهده می‌باشد، هنگامی که تصاویر آموزشی کمی در اختیار داریم، عمل کرد روش پیشنهادی نسبت به روش استاندارد، بسیار بهتر می‌باشد. با افزایش تعداد تصاویر، عمل کرد دو روش به یک‌دیگر نزدیک می‌شود. در شکل (۸) نتایج این آزمون روی تصاویر "چهره" نشان داده شده است. تعداد تصاویر آزمون هم‌چنان ۳۰ تصویر مثبت و ۳۰ تصویر منفی می‌باشد. تعداد ویژگی‌های استخراج شده به ازای هر تصویر، ۱۵ عدد بوده است.

آزمون مذکور روی تصاویر "موتورسیکلت" نیز تکرار شد تا صحت عمل کرد، مورد ارزیابی بیش‌تری قرار گیرد (شکل ۹).

می‌دهد. با افزایش تعداد ویژگی‌ها، دقت این روش نیز روند افزایشی خواهد داشت، اما نرخ رشد افزایش دقت نسبت به افزایش تعداد ویژگی‌ها، بسیار ملایم می‌باشد، زیرا ویژگی‌هایی که دارای ارزش اطلاعاتی بیش‌تری هستند در همان ابتدای کار انتخاب شده‌اند و با افزایش تعداد ویژگی، آن‌هایی که دارای ارزش اطلاعاتی کم‌تری هستند، اضافه خواهند شد (شکل ۶).



شکل (۶) - مقایسه‌ی تاثیر تعداد ویژگی‌ها بر دقت در روش پیشنهادی و روش استاندارد

تاثیر دیگری که تعداد ویژگی‌ها روی این دو روش دارد، مربوط به زمان محاسبات می‌باشد. همان‌گونه که ذکر شد، در روش استاندارد با افزایش تعداد ویژگی‌ها، زمان محاسبات افزایش می‌یابد. این افزایش رابطه‌ای خطی با افزایش تعداد ویژگی‌ها داشته و با دو برابر شدن تعداد ویژگی‌ها، زمان پردازش این روش نیز دو برابر می‌شود. اما این رابطه در روش پیشنهادی به گونه‌ی دیگری می‌باشد. در روش پیشنهادی تمرکز روی نقاط برجسته و گوشه‌های تصویر می‌باشد. اگر در تصویر تعداد کمی نقطه‌ی برجسته وجود داشته باشد، تعداد ویژگی‌های استخراج شده، کم‌تر از تعداد تعیین شده در فرایند خواهد بود. برای مثال اگر تصویر آموزشی، یک مربع ساده باشد، تنها چهار نقطه به عنوان ویژگی وجود دارد که فرایند آن‌ها را استخراج کند. از طرفی، مکان هر نقطه‌ی برجسته در شکل، منحصر به فرد می‌باشد، بنابراین فقط یک بار در مرحله‌ی استخراج ویژگی از آن استفاده می‌شود. به عبارتی، به هیچ عنوان افزونگی نخواهیم داشت. به همین دلیل با افزایش تعداد ویژگی‌ها، زمان محاسبات تا حدودی افزایش خواهد یافت و بعد از یک مرحله به بعد، دیگر هیچ افزایشی در زمان محاسبات نخواهیم داشت. این مرحله، زمانی است که تمام نقاط برجسته‌ی تصویر انتخاب

تغییرات و قابلیت اطمینان پایین، هنگامی که تعداد تصاویر آموزشی و یا تعداد ویژگی‌ها کم باشد، بسیار بیش‌تر خواهد بود.

### ۲-۲-۳- کاهش بار محاسباتی

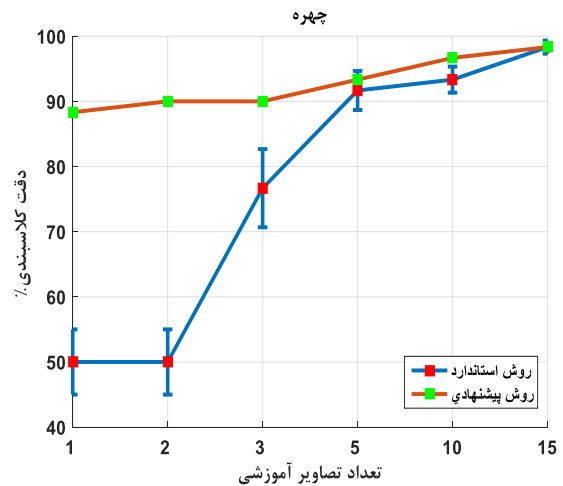
همان‌گونه که نمودارها نشان می‌دهند، روش پیشنهادی هیچ گونه افزونگی ندارد. هم‌چنین با انتخاب بهینه‌ی ویژگی‌ها، با کم‌ترین تعداد ویژگی، بهترین عمل‌کرد را خواهد داشت. اما بار محاسباتی روش استاندارد تحت تاثیر افزونگی به شدت افزایش خواهد یافت. چون در این روش هیچ رویه‌ای برای بررسی ویژگی‌ها وجود ندارد و احتمال انتخاب تکراری یک نقطه زیاد خواهد بود.

### ۳-۲-۳- افزایش دقت

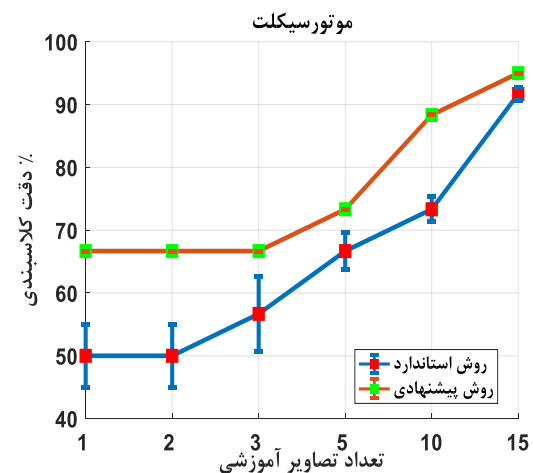
دقت یک فرایند بازشناسی، به شدت به ویژگی‌های استخراج شده‌ای وابسته می‌باشد که برای بازتعریف تصاویر مورد استفاده قرار می‌گیرند. این ویژگی‌ها باید به گونه‌ای باشند که بیش‌ترین تمایز را بین تصاویر مثبت و تصاویر منفی ایجاد کنند، تا دسته‌بند بتواند به درستی شی را بازشناسی کند. طبق بررسی‌های زیست شناختی، نقاط مربوط به مرزها، گوشه‌ها و نقلی که نسبت به اطراف خود برجسته‌تر باشند، بیش‌ترین تاثیر را در بازشناسی شی دارند. در روش پیشنهادی با تکیه بر این روی‌کرد بینایی انسان، سعی در استخراج نقاط مذکور شده است.

### ۳-۳- کارهای پیشنهادی برای آینده

اگرچه کارهای زیادی برای بهبود مدل مبتنی بر بینایی انسان ارائه شده است، اما هم‌چنان نقاط ضعفی وجود داشته که برای رفع آن‌ها به پژوهش و تحقیقات بیش‌تری نیاز است. یکی از این نقاط ضعف، استفاده نکردن از اطلاعات تکمیلی، مانند رنگ می‌باشد. مدل HMAX تنها از تصاویر خاکستری استفاده می‌کند. اضافه کردن مشخصات رنگ، کمک زیادی به افزایش دقت مدل می‌کند. البته در مقالاتی از جمله [۲۲]، ویژگی رنگ اضافه شده است، اما بار محاسباتی سه برابر شده که عملاً این بار محاسباتی نسبت به دقت اضافه شده توجیه‌پذیر نمی‌باشد. بنابراین استفاده از اطلاعات تکمیلی از جمله رنگ، بدون افزایش زیاد بار محاسباتی، یک چالش بزرگ می‌باشد. یک نقطه ضعف دیگر روش استاندارد، کاهش شدید عمل‌کرد در برابر چرخش می‌باشد. این مدل در برابر تغییرات مقیاس، جابه‌جایی و روشنایی بسیار عمل‌کرد خوبی از خود نشان



شکل (۸) - مقایسه‌ی تاثیر تعداد تصاویر آموزشی بر دقت بین روش استاندارد و روش پیشنهادی در بازشناسی چهره



شکل (۹) - مقایسه‌ی تاثیر تعداد تصاویر آموزشی بر دقت بین روش استاندارد و روش پیشنهادی در بازشناسی موتورسیکلت

### ۲-۳- بحث

روش پیشنهادی از جهات مختلفی نسبت به روش استاندارد برتری دارد:

#### ۱-۲-۳- تکرارپذیری و قابلیت اطمینان

یکی از مهم‌ترین مزیت‌های روش پیشنهادی نسبت به روش استاندارد، تکرارپذیری و قابلیت اطمینان آن است. در کلیه‌ی آزمون‌های انجام شده، نتایج مربوط به روش پیشنهادی در هر بار تکرار آزمون، یک‌سان بوده و هیچ تغییری در آن وجود نداشت. این خاصیت از تکرارپذیر بودن روش استخراج ویژگی ناشی می‌شود. در حالی که روش استاندارد در این زمینه ضعف دارد و حتی در شرایط یک‌سان نیز نتایج تغییر می‌کند. به عبارت دیگر تکرارپذیری روش استاندارد پایین می‌باشد. این

- [6] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, "Robust object recognition with cortex-like mechanisms. TL - 29," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29 VN-r, no. 3, pp. 411–426, 2007.
- [7] S. Seifzadeh, M. Rezaei, and O. Farahbakhsh, "A Computational Visual Neuroscience Model for Object Recognition," *J. Adv. Med. Sci. Appl. Technol.*, vol. 2, no. 4, p. 315, Jan. 2017.
- [8] H.-Z. Zhang, Y.-F. Lu, T.-K. Kang, and M.-T. Lim, "B-HMAX: A fast binary biologically inspired model for object recognition," *Neurocomputing*, vol. 218, pp. 242–250, Dec. 2016.
- [9] Yulong Wang, Qingtian Zhang, and Xiaolin Hu, "Distributed sparse HMAX model," in 2015 Chinese Automation Congress (CAC), 2015, no. 1, pp. 740–745.
- [10] Y. Li, W. Wu, B. Zhang, and F. Li, "Enhanced HMAX model with feedforward feature learning for multiclass categorization," *Front. Comput. Neurosci.*, vol. 9, no. October, pp. 1–14, Oct. 2015.
- [11] C. Thériault, N. Thome, and M. Cord, "Extended Coding and Pooling in the HMAX Model," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 764–777, Feb. 2013.
- [12] P. Mishra and B. K. Jenkins, "Hierarchical model for object recognition based on natural-stimuli adapted filters," in 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, 2010, pp. 950–953.
- [13] D. B. Walther and C. Koch, "Attention in hierarchical models of object recognition," in *Progress in Brain Research*, vol. 165, no. 06, 2007, pp. 57–78.
- [14] H. Sufikarimi and K. Mohammadi, "Speed up biological inspired object recognition, HMAX," in 2017 3rd Iranian Conference on Intelligent Systems and Signal Processing (ICSPIS), 2017, pp. 183–187.
- [15] M. Jazlaeiyan and H. S. Shahhoseini, "Optimal Feature Selection in Biologically Inspired Model for Object Recognition Using Mutual Information Maximisation," *Iran. J. Biomed. Eng.*, vol. 8, no. 4, pp. 371–383, 2015.
- [16] M. Ghodrati, S.-M. Khaligh-Razavi, R. Ebrahimpour, K. Rajaei, and M. Pooyan, "How Can Selection of Biologically Inspired Features Improve the Performance of a Robust Object Recognition Model?," *PLoS One*, vol. 7, no. 2, p. e32357, Feb. 2012.
- [17] J. Mutch and D. G. Lowe, "Object Class Recognition and Localization Using Sparse Features with Limited Receptive Fields," *Int. J. Comput. Vis.*, vol. 80, no. 1, pp. 45–57, Oct. 2008.
- [18] Y. Lu, M. Lim, H. Zhang, and T. Kang, "Enhanced hierarchical model of object recognition based on a novel patch selection method in salient regions," *IET Comput. Vis.*, vol. 9, no. 5, pp. 663–672, Oct. 2015.

می‌دهد، اما در برابر چرخش بسیار آسیب‌پذیر بوده و عمل‌کرد خود را از دست می‌دهد. بنابراین اصلاح این ضعف، یکی از ملزومات مدل استاندارد می‌باشد. در آخر باید به کند بودن نسبی فرایند اشاره کرد. اگرچه کارهایی برای افزایش سرعت انجام شده است، اما این روش هنوز برای بسیاری از کارهای بلادرنگ قابل استفاده نمی‌باشد. کند بودن فرایند به علت بار محاسباتی زیادی می‌باشد که در مرحله‌ی تطبیق الگو در S2 وجود دارد. ارائه‌ی روشی سریع‌تر در تطبیق الگو، می‌تواند این روش را تا حد زیادی بهینه کند.

#### ۴- نتیجه‌گیری

در این پژوهش سعی شد تا با مدل کردن بخش ثانویه‌ی قشر بینایی و ترکیب آن با مدل HMAX استاندارد، مدل کامل‌تری از بینایی انسان ارائه شود. با اضافه کردن بخشی که نقش آن استخراج نقاط برجسته و گوشه‌ها، به عنوان ویژگی‌های غیرتصادفی بوده است، یکی از مهم‌ترین معایب مدل استاندارد، یعنی تکرارناپذیری آن، مرتفع شده است. روش پیشنهادی با به کارگیری یک روش تکرارپذیر در استخراج ویژگی، بر این معضل فائق آمده است. علاوه بر آن روش پیشنهادی عمل‌کرد مدل استاندارد را تحت شرایط کمبود تصاویر آموزشی به صورت چشم‌گیری بهبود داده است. روش پیشنهادی با تکیه بر ویژگی‌های پرمحتوا، با تعداد ویژگی‌های کم‌تری می‌تواند بازشناسی شیء را انجام دهد، که این امر باعث افزایش سرعت مدل نیز شده است.

#### ۵- مراجع

- [1] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-Up Robust Features (SURF)," *Comput. Vis. Image Underst.*, vol. 110, no. 3, pp. 346–359, Jun. 2008.
- [2] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
- [3] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, vol. 1, pp. 886–893.
- [4] P. Moreno, M. J. Marín-Jiménez, A. Bernardino, J. Santos-Victor, and N. P. de la Blanca, "A Comparative Study of Local Descriptors for Object Category Recognition: SIFT vs HMAX," in *Pattern Recognition and Image Analysis*, no. June, 2007, pp. 515–522.
- [5] P. T. Riesenhuber M, "Hierarchical models of object recognition in cortex," *Nat. Neurosci.*, pp. 1019–1025, 1999.

- [22] B. Yang, L. Zhou, and Z. Deng, "C-HMAX: Artificial cognitive model inspired by the color vision mechanism of the human brain," *Tsinghua Sci. Technol.*, vol. 18, no. 1, pp. 51–56, 2013.
- [23] T. Serre, L. Wolf, and T. Poggio, "Object Recognition with Features Inspired by Visual Cortex," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2007, vol. 2, pp. 994–1000.
- [24] K. D. Flemming, "Essential Neuroscience," *Mayo Clin. Proc.*, vol. 81, no. 10, p. 1409, Oct. 2006.
- [19] I. Biederman, "Recognition-by-Components: A Theory of Human Image Understanding," *Psychol. Rev.*, vol. 94, no. 2, pp. 115–147, 1987.
- [20] A. Al Maashri, M. DeBole, C.-L. Yu, V. Narayanan, and C. Chakrabarti, "A hardware architecture for accelerating neuromorphic vision algorithms," in *2011 IEEE Workshop on Signal Processing Systems (SiPS)*, 2011, pp. 355–360.
- [21] Z. Guo and Z. J. Wang, "An Unsupervised Hierarchical Feature Learning Framework for One-Shot Image Recognition," *IEEE Trans. Multimed.*, vol. 15, no. 3, pp. 621–632, Apr. 2013.