

# بهینه سازی ویژگی‌های لایه‌ی میانی مدل شناسایی اشیاء HMAX

مینا غیرتمند<sup>۱</sup>، حمید خالوزاده<sup>۱</sup> و حمید سلطانیان زاده<sup>۲،۳</sup>

<sup>۱</sup> دانشکده برق، دانشگاه صنعتی خواجه نصیرالدین طوسی، آقطب علمی کنترل و پردازش هوشمند، دانشگاه تهران، <sup>۳</sup> بیمارستان هنری فورد آمریکا

m.gheiratmand@ee.kntu.ac.ir, h\_khaloozadeh@kntu.ac.ir, hszadeh@ut.ac.ir

## چکیده

مدل شناسایی اشیاء HMAX از گروه مدل‌های سلسله مراتبی بدون پسخوری است که ساختار و پارامترهای آن بر اساس خصوصیات زیستی کورتکس بینایی نخستینها انتخاب شده‌است. به خاطر همین ویژگی و نیز موفقیت آن در شناسایی اشیاء در صحنه‌های واقعی، این مدل هم در علوم اعصاب و هم در بینایی ماشینی بسیار مورد توجه قرار گرفته است. در این مقاله مدل HMAX اصلاح شده، که دارای قابلیت یادگیری در اتصالات لایه‌ی C1 به S2 می‌باشد، پیاده‌سازی شده و به منظور بهینه‌سازی ویژگی‌های لایه‌ی S2 که در مدل اصلی به طور تصادفی انتخاب می‌شوند، پیشنهاداتی ارائه شده‌است. دو روش خوشه‌بندی k-Means و انتخاب ویژگی ترتیبی پس‌رو بهبود یافته به این منظور به کار رفته است. همچنین با استفاده از روش‌های شبیه‌سازی اتفاقی Monte Carlo اثر نویز بر عملکرد مدل بررسی شده و با روش PCA مقایسه شده‌است. نتایج نشان می‌دهد که با استفاده از روش انتخاب ویژگی پس‌رو در مرحله‌ی یادگیری، می‌توان ضمن افزایش سرعت پردازش سیستم به عملکرد بهتری نیز دست یافت. همچنین مشاهده شد که مدل اصلی و پیشنهادی بر خلاف روش PCA نسبت به افزایش واریانس نویز ورودی مقاوم نیستند. این در حالی است که در شرایط بدون نویز عملکرد مدل پیشنهادی ۱۰٪ بیشتر از روش PCA است.

کلمات کلیدی- مدل HMAX، انتخاب ویژگی، عملکرد دسته‌بندی کننده، نویز گوسی، شبیه‌سازی مونت کارلو.

## ۱- مقدمه

اطلاعات بینایی توسط نورونها در مسیر پیش‌رو نواحی مختلف مسیر قدامی کورتکس بینایی طراحی شده است.

با پیش‌روی در لایه‌های این سیستم به ویژگی‌هایی از شیء دست می‌یابیم که نسبت به مقیاس و موقعیت تغییرناپذیرند و در عین حال نسبت به شیء مورد نظر انتخابگر هستند [2-13,5]. به منظور شناسایی اشیاء، این ویژگی‌های استخراج شده از تصویر ورودی به یک دسته‌بندی کننده فرستاده می‌شوند. ویژگی‌های استخراج شده از این روش در مقایسه با ویژگی‌های استخراجی

شناسایی اشیاء از اعمال بسیار مهم در کورتکس بینایی و سیستم‌های بینایی ماشینی است. روش‌های مختلف شناسایی اشیاء و کاستی‌های آنها در [2,4,11] بحث شده‌است. مدل HMAX [13] که در ۱۹۹۹ بر اساس بسط آرایش سلولهای ساده و پیچیده در ناحیه‌ی بینایی V1 پیشنهاد شد، به دسته‌ی مدل‌های سلسله‌مراتبی تعلق دارد. این مدل برپایه‌ی روند پردازش

شده و در نهایت مقاومت مدل بهبود یافته در مقابل نویز ورودی سنجیده و با روش PCA مقایسه شده است.

## ۲- پیاده سازی مدل

شکل ۱ دیگرام بلوکی مدل بهبودیافته‌ی HMAX را به همراه نواحی متناظر آن در کورتکس بینایی نشان می‌دهد.

### ۲-۱- لایه‌ی S1

این لایه معادل سلولهای ساده در ناحیه اول بینایی V1 می باشد که با استفاده از مجموعه‌ای از فیلترهای گابور به شکل زیر پیاده‌سازی می‌شوند. جزئیات این لایه و پارامترهای آن در [12] آمده است.

$$G_j^i(x,y) = \exp\left(-\frac{(x_r^2 + \gamma^2 y_r^2)}{2\sigma_i^2}\right) \times \cos\left(\frac{2\pi}{\lambda_i} x_r + \phi\right) \quad \forall x, y < s_i \quad (1)$$

$$i = 1, \dots, 16. \quad s_i = 7, 9, \dots, 35, 37.$$

$$j = 1, \dots, 4. \quad \theta_j = 0, 45^\circ, 90^\circ, 135^\circ.$$

### ۲-۲- لایه‌ی C1

این لایه که متناظر با سلولهای پیچیده‌ی ناحیه‌ی V1 است، تصاویر حاصل از بانک فیلتری (و یا عبارتی خروجی سلولهای ساده) را به گونه‌ای با هم ترکیب می‌کند که خصوصیات سلولهای پیچیده یعنی تغییرپذیری و حساسیت کمتر به موقعیت محرک در میدان گیرنده و گسترش باند فرکانسی (که در فیزیولوژی اثبات شده‌اند) محقق شود. خصوصیت اول با ماکزیمم‌گیری‌های محلی بر روی موقعیت‌های مختلف [9] و خصوصیت دوم با ماکزیمم‌گیری از مقیاس‌های مختلف و موقعیت‌های یکسان حاصل می‌شود.

### ۲-۳- لایه‌ی S2

این بخش از مدل شبکه‌ای از نورونهای RBF گوسی است و دارای قابلیت یادگیری می باشد. مراکز این نورونها از طریق نمونه برداری تصادفی با اندازه‌ی patch های مختلف (در اینجا ۴ اندازه) از ویژگی‌های C1 استخراج شده از مجموعه‌ای از تصاویر مثبت به دست می آیند. مرجع [2] اثر تعداد این الگوها بر عملکرد مدل را بررسی می کند. پس از مرحله‌ی یادگیری، برای هر تصویر ورودی جدید ویژگی‌های C1 آن با الگوهای ذخیره شده در S2 مقایسه شده و تابعی نمایشی از مجذور فاصله‌ی آنها تولید می‌شود.

سایر روش‌های مطرح در دسته‌بندی اشیاء قدرتمندتر عمل می‌کنند [11]. با این وجود ما در این مقاله نشان داده‌ایم که این ویژگی‌ها نسبت به نویز در تصاویر ورودی حساس‌اند.

مدل HMAX همچنین راهنمای بسیار خوبی برای مدلسازی نواحی بینایی مسیر قدامی در کورتکس است و با بهره‌گیری از آن می‌توان مدل‌های متکامل‌تری از سیستم بینایی ارائه داد: نظیر ترکیب مدل کمی پردازش حرکت در مسیر خلفی [8] با مدل کمی شناسایی اشیاء در مسیر قدامی. اما پایین بودن سرعت پردازش از محدودیت‌های این مدل در کاربردهای زمان-واقعی است [2]. همچنین برای تطابق بیشتر با سیستم‌های بسیار سریع شناسایی بیولوژیکی، بهبود مدل در این راستا حائز اهمیت می‌باشد. مدل HMAX مدلی مستقیم است در حالیکه در واقعیت سیستم بینایی انسان دارای مسیرهای پسخور فراوانی می‌باشد که احتمالاً به تصمیم‌گیری‌های درست در هر ناحیه‌ی قشری بینایی کمک می‌کند. در این راستا مکانیزم‌های توجهی برای بسط مدل HMAX بکار گرفته شده‌اند. مرجع [7] سیستمی برگرفته از HMAX را پیشنهاد می‌کند که تنها دارای یک لایه (V1) و یک مکانیزم توجهی است. با وجود اینکه این مدل توانسته از نظر زمانی (و در برخی کاربردها از نظر عملکرد) از HMAX پیشی بگیرد، اما منجر به حذف لایه‌هایی از مدل و در نتیجه دور شدن از سیستم بینایی بیولوژیکی شده‌است.

اهمیت این مدل همچنین در پایداری آن به خصوصیات پردازش قشری (نظیر عدم استفاده از اطلاعات هندسی) و در عین حال داشتن عملکردی دست کم در سطح بهترین سیستم‌های بینایی کامپیوتری است [11]. در مدل HMAX هیچ پارامتری به منظور دستیابی به عملکرد بهینه تنظیم نشده‌است، بلکه در عوض همه پارامترها به گونه‌ای تنظیم شده‌اند که با آنچه در مورد سیستم بینایی نخستین‌ها شناخته شده‌است منطبق باشد [11]. این امر ضمن تأکید بر قدرت این مدل سلسله مراتبی پیش‌رو در شناسایی اشیاء، راه را برای بهبود زمانی و عملکردی سیستم با اعمال نظارت بر بخش‌هایی از آن باز می‌گذارد. در این مقاله ما پس از پیاده‌سازی مدل، نقش عناصر مختلف آن را در عملکرد نهایی بررسی کرده‌ایم. این بررسی امکان حذف برخی افزونگی‌ها را نشان می‌دهد. با توجه به این مسأله، بدون تغییر ساختار اصلی سیستم دست به ساده‌سازی‌هایی زده‌ایم. سپس روش‌هایی را برای حذف ویژگی‌های زائد در لایه‌ی یادگیرنده‌ی مدل (S2) بکار برده‌ایم که از نظر زمان پردازش و نیز عملکرد بهبودهایی را به دست داده است.

در ادامه ابتدا ساختار مدل HMAX اصلاح شده (و نحوه‌ی استخراج ویژگی‌های C2) توضیح داده می‌شود. سپس روش‌های به کار رفته به منظور انتخاب ویژگی‌های بهتر در لایه‌ی S2 بررسی

## ۴-۲- لایه ی C2

این لایه بر روی خروجی‌های لایه‌ی S2 ماکزیمم‌گیری می‌کند و ویژگی‌هایی با ابعاد برابر با تعداد نورونهای S2 تولید می‌کند.

## ۵-۲- دسته بندی کننده

ویژگی‌های C2 از داده‌های آموزش و تست استخراج شده و به دسته‌بندی‌کننده فرستاده می‌شوند. نتایج بررسی عملکرد دسته‌بندی‌کننده‌های مختلف نشان داد که SVM خطی در مقایسه با SVM غیرخطی و KNN بهتر عمل می‌کند.

اساس یادگیری، بویژه در دوران کودکی، از نگهداری اطلاعات زائد و خطاساز پرهیز می‌کند.

به طور کلی عملکرد با افزایش تعداد ویژگی‌ها افزایش می‌یابد. این امر به دلیل افزایش قدرت تمایز مجموعه‌ی ویژگی‌هاست. اما با افزایش تعداد ویژگی‌ها احتمال وقوع بیش‌برازش نیز وجود دارد [5]. از آنجایی که خروجی سیستم از جنس ۰ و ۱ (عضویت یا عدم عضویت در دسته‌ای از اشیاء) است، آموزش شبکه RBF با توجه به ابعاد سیستم به هیچ عنوان کارآمد نمی‌باشد. با توجه به این مسأله ما دو روش دیگر را برای بهبود ویژگی‌های لایه ی S2 به کار برده ایم که در ادامه توضیح داده شده‌است.

## ۳- آزمایش ها و نتایج

### ۳-۱- مجموعه داده ها

در کلیه آزمایشات از تصاویر موجود در پایگاه داده‌ی CalTech5 استفاده شده‌است. مجموعه تصاویر موتورسیکلت (بیش از ۸۰۰ تصویر رنگی با اندازه‌های مختلف) و مجموعه تصاویر پس‌زمینه (حدود ۵۰۰ تصویر با روشنایی و اندازه‌های مختلف)، پس از انجام پیش‌پردازش‌هایی شامل تبدیل به مقیاس خاکستری و ارتفاع ۱۴۰ پیکسل (با حفظ نسبت ارتفاع به عرض تصاویر) به ترتیب به عنوان مجموعه‌های مثبت و منفی به کار رفته‌اند. هر مجموعه‌ی داده‌های آموزش شامل ۴۰ تصویر مثبت (شامل شیء مورد نظر یعنی موتورسیکلت) و ۵۰ تصویر منفی (پس‌زمینه) است. هر مجموعه‌ی تست نیز شامل ۵۰ تصویر مثبت و ۵۰ تصویر منفی می‌باشد. لازم به ذکر است که تصاویر، نرمالیزه یا قطعه‌بندی نمی‌شوند [2]. در هر آزمایش ویژگی‌های C2 کلیه تصاویر آموزش و تست استخراج شده و به دسته‌بندی‌کننده فرستاده می‌شود. برای بدست آوردن عددی قابل اطمینان برای عملکرد سیستم در هر آزمایش، به پیشنهاد [4] ۱۰ مجموعه تصویر آموزش و تست به طور تصادفی از میان پایگاه‌داده انتخاب شده و روی عملکردهای آنها میانگین‌گیری شده‌است.

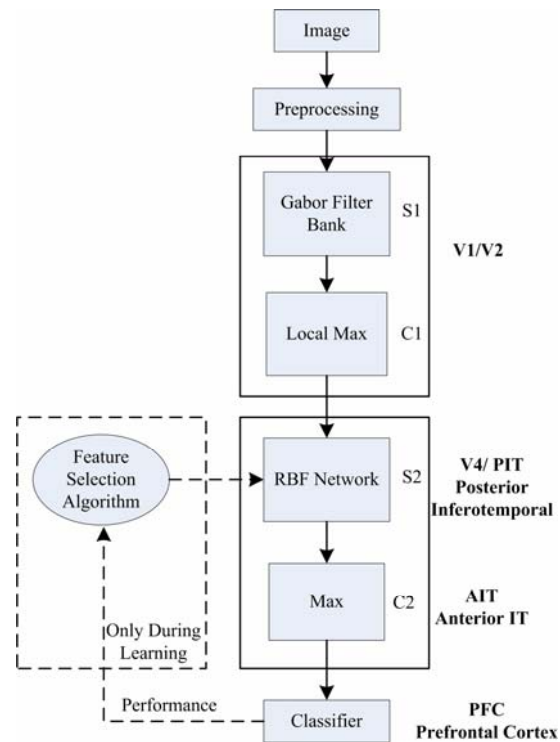
### ۳-۲- مدل اصلی با ویژگی های بهبود نیافته

عملکرد مدل اصلی با بکارگیری هر ۸ باند C1 و ۴ اندازه‌ی patch‌های S2 به منظور مقایسه با روشهای PCA و مدل بهبود یافته در لایه ی S2 با روش فوق محاسبه شده است. این مقدار برابر ۹۶٪ می‌باشد.

### ۳-۳- خوشه‌بندی مراکز RBF لایه S2 با

#### الگوریتم k-Means

روش خوشه‌بندی k- میانگین برای جدا کردن مراکز نورون‌های RBF در لایه S2 به دو مجموعه‌ی مؤثر و غیرمؤثر (یا مضر)



شکل ۱: دیاگرام بلوکی مدل بهبود یافته به همراه نواحی متناظر آن در کورتکس

## ۶-۲- بهینه‌سازی ویژگی‌های S2

ویژگی‌های لایه‌ی S2 به طور تصادفی از میان مجموعه‌ای از تصاویر مثبت یاد گرفته شده‌اند [2,4] و بنابراین ممکن است که برخی از آنها در تولید ویژگی‌های خاص شیء مورد نظر مفید نباشند. از آنجایی که این ویژگی‌ها بار محاسباتی بسیار زیادی را به سیستم تحمیل کرده و سیستم را کند می‌کنند، بهینه کردن آنها (در لایه S2) امری ضروری به نظر می‌رسد. این کار، مدل را به سیستم بینایی مغز نخستینها نظیر انسان نزدیک تر می‌سازد. چرا که این سیستم عملکردی بسیار سریع و بهینه دارد و بر

وحذف مجموعه‌ی بد برای کاهش محاسبات و افزایش عملکرد سیستم به کار رفته‌است. با توجه به نتایج حاصل از آزمایشات صورت گرفته بر روی مجموعه‌های مختلف داده‌ها، این الگوریتم روش مناسبی برای بهبود ویژگی‌های در لایه‌ی S2 ارزیابی نشده است. چرا که تفاوت چندانی در مجموعه نوروهای هر خوشه با خوشه‌ی دیگر دیده نمی‌شود، به این معنی که عملکرد سیستم با قرار دادن عناصر هر یک از خوشه‌ها در مراکز نوروهای لایه‌ی S2 تغییر چندانی نمی‌کند. با توجه به اینکه عناصر این لایه به طور تصادفی (از میان مجموعه‌ای از تصاویر مثبت) انتخاب شده‌اند، متمرکز نشدن آنها حول دو مرکز به صورت معنی‌دار چندین عیب به نظر نمی‌آید. با توجه به الگوریتم خوشه‌بندی k-Means مبنی بر کمینه کردن مجموع فواصل نقاط دسته‌ها از مراکز آنها [10] این روش برای جداسازی patch های انتخاب شده از تصاویر ایده‌آل با پس زمینه‌ی یکنواخت و یا هر دو مجموعه‌ی تصاویر مثبت و منفی، بهتر عمل می‌کند.

### ۳-۴- انتخاب ویژگی ترتیبی پس‌رو

#### (Sequential Backward Selection)

در این روش ابتدا یک معیار خطا در نظر گرفته، به ازاء بردار ویژگی اولیه آن را محاسبه می‌کنیم. سپس ویژگی‌ها را یکی یکی حذف کرده، اثر حذف هر ویژگی را بر معیار خطا می‌سنجیم. در پایان این مرحله مجموعه ویژگی‌ای را که کمترین خطا را داشته نگه می‌داریم و بدترین ویژگی را حذف می‌کنیم. در مرحله‌ی بعد همین کار را بر روی بردار جدید انجام می‌دهیم تا به بردار با تعداد ویژگی‌های مطلوب برسیم [10].

در اینجا عملکرد سیستم در دسته‌بندی را به عنوان معیار در نظر گرفته و با رویکردی مشابه رویکرد بالا در هر مرحله به حذف ویژگی‌های بدتر (که باعث کاهش عملکرد یا ثابت ماندن آن می‌شوند) پرداخته‌ایم. از آنجایی که روش انتخاب ویژگی پس‌رو برای حذف ویژگی‌های زائد در این مدل الگوریتم کندی است، ابتدا با توجه به ساختار مدل و بررسی تأثیر اجزاء مختلف آن بر عملکرد سیستم دست به ساده‌سازی‌هایی زده‌ایم تا بتوانیم آزمایشات مختلف مورد نظر از جمله افزودن نویز را بر روی مدل انجام دهیم.

جدول (۱): عملکرد سیستم به ازاء اندازه‌های مختلف ویژگی در

لایه‌ی S2 برای یک مجموعه داده

عملکرد	اندازه‌ی patch
٪۸۸	۴×۴
٪۹۲	۸×۸
٪۹۲	۱۲×۱۲
٪۹۳	۱۶×۱۶
٪۹۴	همه اندازه‌های patch

جدول (۱) عملکرد سیستم را به ازاء بهره‌گیری از اندازه‌های مختلف patch های تصادفی در لایه‌ی S2 برای یک مجموعه داده، نشان می‌دهد. با توجه به جدول فوق و نتایج نسبتاً مشابه برای سایر داده‌ها، و با توجه به اینکه به طور کلی ویژگی‌های با اندازه‌ی متوسط عملکرد بهتری دارند، اندازه patch ۸×۸ را برای انجام آزمایشات بعدی انتخاب می‌کنیم. ویژگی‌های متوسط نسبت به ویژگی‌های کوچکتر انتخابگری بیشتر و نسبت به ویژگی‌های بزرگ انعطاف بیشتری در انطباق با تعداد بیشتری از تصاویر مثبت دارند [5].

در مرحله‌ی بعد با جایگذاری patch های تصادفی با اندازه‌ی ۸×۸ در مراکز نوروهای RBF تأثیر باندهای مختلف مقیاس را بررسی می‌کنیم. جدول (۲) عملکرد سیستم را به ازاء شرکت باند مقیاس‌های مختلف C1 در استخراج ویژگی‌های نهایی را نشان می‌دهد.

جدول (۲): باندهای مقیاس تصاویر C1 که برای استخراج

ویژگی‌های C2 استفاده شده

عملکرد	باندهای مقیاس تصاویر C1
٪۹۲	باند ۱
٪۹۴	باند ۱-۲
٪۹۵	باند ۱-۳
٪۹۷	باند ۱-۴
٪۹۴	باند ۱-۵
٪۹۵	باند ۱-۶
٪۹۲	باند ۱-۷
٪۹۴	باند ۱-۸

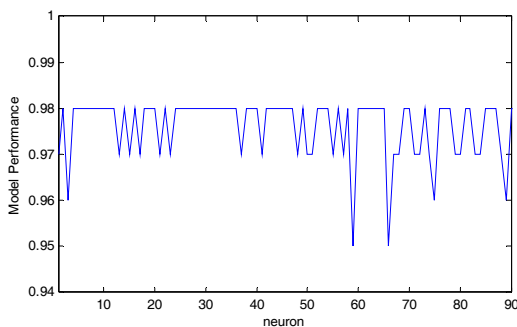
با توجه به جدول (۲) می‌توان مشاهده کرد که الزاماً وجود کلیه باندهای مقیاس در تصاویر C1 برای رسیدن به عملکرد بهتر ضروری نمی‌باشد. مشاهده می‌شود که حضور ۴ باند اول مقیاس، که معادل با ۸ باند اول مقیاس در سلولهای گابور لایه S1 می‌باشند، در مقایسه با حضور کلیه باندها منجر به عملکرد بهتری در شناسایی شده است. در این حالت، بهبود ویژگی‌های C2 استخراج شده از یک سو باعث بالا رفتن عملکرد دسته بندی کننده شده، و از سوی دیگر کاهش قابل ملاحظه‌ای در زمان لازم برای اجرای برنامه‌ها و نزدیک شدن سیستم به شرایط real-time سیستم بینایی انسان بوجود آورده است.

بنابراین الگوریتم انتخاب ویژگی پس‌رو را بر روی ویژگی‌های اندازه‌ی ۸×۸ و باند اول تا چهارم C1 پیاده می‌کنیم. شکل ۲ نتیجه الگوریتم انتخاب ویژگی پس‌رو را پس از حذف ۱۰ ویژگی بدتر نشان می‌دهد. مشخص است که حذف برخی از ویژگی‌ها باعث بهبود عملکرد سیستم شده‌است (از ٪۹۷ به ٪۹۸). و همچنان تعداد زیادی از ویژگی‌ها کاندید حذف شدن هستند. افزایش عملکرد سیستم در ازاء حذف برخی ویژگی‌ها، مؤید این

واقعیت است که بسیاری از ویژگی‌ها زاید و مضر هستند. از طرفی وجود برخی ویژگی‌ها (نظیر ویژگی ۵۹ و ۶۶) به تصمیم‌گیری درست سیستم بسیار کمک می‌کند.

PCA ابعاد فضا تا حدود ۹۹٪ کاهش می‌یابد که از نظر محاسباتی بسیار به صرفه است. اما با توجه به جدول (۳) مشاهده می‌شود که عملکرد روش PCA در مقایسه با مدل اولیه و نیز مدل بهبود یافته بسیار ضعیف‌تر است.

همچنین عملکرد مدل و روش PCA برای تصاویر نویزی نیز مقایسه شده است. برای محاسبه عملکرد سیستم در دسته‌بندی تصاویر نویزی با استفاده از PCA، به صورت زیر عمل کرده‌ایم: ابتدا نویز گوسی (با میانگین صفر و واریانس‌های ۰,۱ و ۰,۱ و ۰,۵) به تصاویر آموزش و تست افزوده شده و برای هر مجموعه از داده‌ها میانگین عملکرد دسته‌بندی کننده بر روی ۱۰۰ تکرار حساب شده است. این کار بر روی هر ده مجموعه داده صورت گرفته و در نهایت از عملکردهای نهایی هر ۱۰ مجموعه میانگین گرفته شده است. نتایج در جدول (۳) آمده است. نسبت سیگنال به نویز در این آزمایشات به ترتیب برابر ۵، ۰,۵ و ۰,۱ است که از رابطه  $SNR = \frac{P_{Signal}}{P_{Noise}}$  محاسبه شده است. در این رابطه  $P$  معرف توان است.



شکل ۲: تأثیر حذف هر یک از نورونها در مرحله تکرار ۱۰ام

بررسی نتایج نشان می‌دهد که حضور نویز با واریانس ۰,۱ تأثیر چندانی در عملکرد روش PCA نمی‌گذارد ولی برای مدل ۱۰٪ کاهش عملکرد ایجاد می‌کند. این میزان معادل کاهش عملکرد روش PCA به ازاء افزودن نویز با واریانس ۰,۵ است. به طور کلی از جدول (۳) می‌توان نتیجه گرفت که مدل پیاده شده نسبت به نویز گوسی در تصاویر ورودی حساس‌تر از روش PCA است. بنا براین می‌توان انتظار داشت که بکارگیری PCA در بخش‌هایی از مدل، بتواند علاوه بر سریع‌تر شدن سیستم، حساسیت آن را نسبت به نویز خارجی کاهش دهد. در مقابل PCA نسبت به تغییرات مقیاس و موقعیت شیء مقاوم نمی‌باشد [۱]، در حالیکه مدل پیاده شده اصولاً با معیار تغییرناپذیری به مقیاس و موقعیت طراحی شده [2-4] و بر این اساس در مقابل این گونه تغییرات مقاوم است. ضمن اینکه تصاویر مثبت بکار رفته برای آموزش و تست دارای پس زمینه ی خالی و یکنواخت نبوده و پس زمینه‌های متفاوت و گوناگونی دارند، خود به عنوان نویز و عامل خطا عمل می‌کنند. موفقیت مدل در رابطه با تصاویر واقعی [11]، که در بینایی ماشینی نیز از اهمیت ویژه‌ای برخوردار است، مؤید توانایی این دسته از مدل‌های سلسله مراتبی بدون پس‌خور در شناسایی اشیاء است.

#### ۴- نتیجه‌گیری

در این مقاله ابتدا با بررسی اجزاء مختلف مدل HMAX اصلاح شده و تأثیر آنها بر عملکرد سیستم در عمل دسته‌بندی، اقدام به حذف برخی افزونگی‌ها نظیر باندهای مقیاس C1 و اندازه‌ی patch‌های S2 کرده‌ایم. این کار بدون حذف یا تغییر

در مورد این مثال کاهش ویژگی‌ها تا حذف ۲۶ ویژگی (و عملکرد ۹۸٪) انجام شده است. با ادامه اجرای الگوریتم می‌توان ویژگی‌های زائد بیشتری را حذف کرد. قابل ذکر است که عملکرد ۹۷٪ که با ۱۰۰ ویژگی اولیه حاصل شده بود، با کاهش بیش از ۳۰ درصدی ویژگی‌ها نیز قابل دسترسی است. از آنجایی که این الگوریتم کند می‌باشد، برای افزایش سرعت الگوریتم تغییراتی در آن داده شده تا بتواند بدون نقض هدف اولیه یعنی کاهش تابع خطا (در اینجا افزایش عملکرد دسته‌بندی کننده) در صورت امکان در هر تکرار بیش از یک ویژگی بد را حذف کند. با وجود کند بودن الگوریتم انتخاب ویژگی پس‌رو، دقت این روش در مقایسه با الگوریتم k-Means بیشتر است و بکارگیری آن منجر به افزایش عملکرد سیستم می‌شود.

#### ۳-۵- بررسی حساسیت نسبت به نویز

در این بخش به بررسی اثر نویز در عملکرد سیستم و مقایسه‌ی آن با روش PCA می‌پردازیم. به منظور کاهش زمان محاسبات، از ویژگی‌های با اندازه‌ی ۸×۸ و باند اول تصاویر C1 در مدل استفاده کرده‌ایم. نتایج این بررسی در جدول (۳) آمده است.

#### ۳-۵-۱- PCA

در اینجا روش تحلیل مؤلفه‌های اصلی به منظور کاهش ابعاد داده‌ها پیش از فرستادن آنها به کلاسیفایر (SVM) به کار رفته است. PCA از روش‌های معمول در شناسایی اشیاء و به ویژه شناسایی چهره است [6,14].

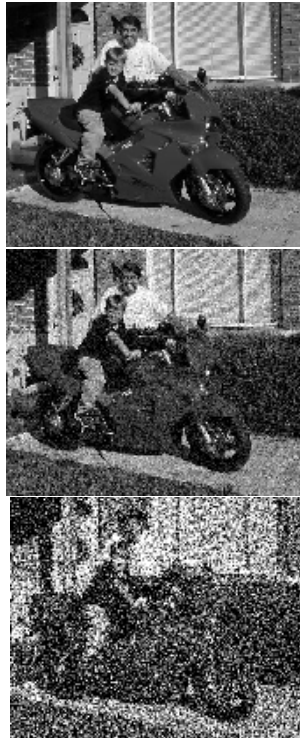
برای دسته‌بندی با استفاده از روش تحلیل مؤلفه‌های اصلی، ابتدا تصاویر درون مجموعه‌های ۱ تا ۱۰ از لحاظ ابعاد با هم مساوی شده‌اند (۲۱۰×۱۴۰). سپس مؤلفه‌های اصلی عمود بر هم (غیرصفر) تصاویر آموزش پیدا شده و به کمک آنها داده‌های آموزش و تست در فضای جدید تصویر می‌شوند. با استفاده از

جدول (۳): عملکرد مدل و روش PCA بر روی داده های بدون نویز و نویزی

درصد کاهش عملکرد			عملکرد به درصد (میانگین عملکرد ۱۰ مجموعه داده)				روش استخراج ویژگی
نویز با واریانس ۰.۵	نویز با واریانس ۰.۱ (SNR=0.5)	نویز با واریانس ۰.۰۱ (SNR=5)	نویز با واریانس ۰.۵	نویز با واریانس ۰.۱ (SNR=0.5)	نویز با واریانس ۰.۰۱ (SNR=5)	بدون نویز	
-	25.35	10.5	-	70.91	85	95	مدل ساده شده با ویژگی های بهبود یافته
13.75	1.15	0	73.57	84.32	85.6	85.3	PCA



(الف)



(ب)

شکل الف و ب: دو نمونه تصاویر موتورسیکلت، به ترتیب در سه حالت بدون نویز و نویزدار شده با SNR=5 و SNR=0.5

ساختار سیستم از یک سو و حفظ عملکرد دسته بندی از سوی دیگر، صورت گرفته است. پس از آن، دو روش خوشه بندی  $k$ -Means و انتخاب ویژگی ترتیبی پس رو به منظور حذف ویژگی های بد در لایه ی S2 مدل ساده شده به کار گرفته شده است. با توجه به تصادفی بودن ویژگی های این لایه روش خوشه بندی نتایج قابل قبولی به دست نمی دهد. اما با استفاده از روش انتخاب ویژگی توانستیم با کاهش ۳۰ درصدی ویژگی ها، ۱٪ افزایش عملکرد نیز داشته باشیم. این کاهش ویژگی در کاهش بار محاسباتی سیستم و سریع تر شدن آن نقش بسیار مؤثری داشته است.

همچنین اثر افزایش نویز با واریانس های مختلف در تصاویر ورودی بر روی عملکرد مدل بررسی و با ویژگی های استخراجی از روش PCA مقایسه شد. نتایج نشان می دهد که ویژگی های C2 استخراجی نسبت به نویز ورودی بسیار حساس اند. در حالیکه روش PCA نسبت به این نوع نویز مقاوم بوده و در مقابل تغییرات مقیاس و موقعیت حساس است.

به طور کلی کاهش زمان پردازش سیستم به لحاظ نزدیکی به سیستم های بیولوژیکی و نیز در کاربردهای بینایی ماشینی از اهمیت ویژه ای برخوردار است. بر پایه ی نتایج به دست آمده از آزمایشات انجام شده با روش های دقیق آماری در این مقاله، می توان از روش های ترکیبی برای بهینه سازی الگوهای S2 و لایه های دیگر مدل استفاده کرده و ضمن حفظ عملکرد بالای سیستم، زمان پردازش و حساسیت ویژگی های استخراجی آن به نویز را کاهش داد.

#### ضمیمه

دو نمونه تصویر مثبت از کلاس موتورسیکلت در پایگاه تصاویر CalTech5 به همراه تصاویر نویزی آنها با SNR 5 و 0.5 در این بخش آورده شده است.



## مراجع

- [ ] م. یعقوبی، ک. فائز، و س. معتمدی، "شناسایی گوش با استفاده از روش استخراج مشخصه HMAX و طبقه بندی کننده SVM"، سیزدهمین کنفرانس ملی انجمن کامپیوتر ایران، جزیره کیش، خلیج فارس، ایران، ۱۳۸۶.
- [2] T. Serre, L. Wolf, S. Bileschi, M. Reisenuber, and T. Poggio "Robust Object Recognition with Cortex-Like Mechanisms," *IEEE Trans. On Pattern Analysis and Machine Intelligence* Vol. 29, No. 3, March 2007
- [3] T. Serre, and M. Kouh "A Theory of Object Recognition: Computations and Circuits in the Feedforward Path of the Ventral Stream in Primate Visual Cortex," *AI Memo 2005-036, CBCL Memo 259*, December 2005
- [4] T. Serre, and L. Wolf "Object Recognition with Features Inspired by Visual Cortex," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2005
- [5] T. Serre, J. Louie, M. Reisenhuber, and T. Poggio "On the Role of Object-specific Features for Real World Object Recognition in Biological Vision," *Proc. Workshop Biologically Motivated Computer Vision*, 2002
- [6] M. Turk, and A. Pentland "Eigenfaces for Recognition," *J. Cognitive Neuroscience* Vol. 3, No. 1, 1991
- [7] E. Bermudez-Contreras, H. Buxton, and E. Spier "Attention can improve a simple model for Object Recognition," *J. Image and Vision Computing* Vol. 26, 2008
- [8] E.P. Simoncelli, and D.J. Heeger, "A Model of Neuronal Response in Visual Area MT," *Vision Research* Vol. 38, No. 5, 1998
- [9] I.M. Finn, and D. Ferster "Computational Diversity in Complex Cells of Cat Primary Visual Cortex," *J. of Neuroscience*, Vol. 27 No. 36, 2007
- [10] S. Theodoridis, and K. Koutroumbas "Pattern Recognition," Academic Press, 2<sup>nd</sup> Ed., 2003
- [11] T. Serre, L. Wolf, and T. Poggio "A New Biologically Motivated Framework for Robust Object Recognition," Technical Report CBCL Paper 243 / AI Memo 2004-26, Massachusetts Institute of Technology, Cambridge, MA, November 2004
- [12] T. Serre, "Learning a Dictionary of Shape-Components in Visual Cortex: Comparison with Neurons, Humans and Machines," Ph.D Thesis at MIT, 2006
- [13] M. Reisenuber, and T. Poggio "Hierarchical Models of Object Recognition in Cortex," *Nature Neuroscience*, Vol. 2, No. 11, November 1999
- [14] [www.face-rec.org](http://www.face-rec.org)