Iranian Journal of Biomedical Engineering
8 (2015) 371-383, www.ijbme.org
DOI: 10.22041/ijbme.2014.14705

Optimal Feature Selection in Biologically Inspired Model for Object Recognition Using Mutual Information Maximisation

M. Jazlaeiyan[1], H. S. Shahhoseini[2*]

[1]M.Sc Graduate, Electronic Department, Faculty of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran
[2]Associate Professor, Electronic Department, Faculty of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran

Receipt in the online submission system 1 July 2015, received in revised form 7 July 2015, accepted 30 August 2015

Abstract

Human visual system operates superior than best machine vision systems in object recognition. So, researchers in machine vision and neuroscience try to model human visual system in order to employ it in machine. HMAX is one of the best operating models in this area. It is based on the function of brain cells in the ventral stream of visual cortex and contains four computational layers. In the learning stage, many image partitions called image patches are extracted randomly with different sizes from training images. This random selection of image patches is one of the drawbacks of HMAX which decreases the performance and increases the computational complexity of the algorithm. In this paper, a novel patch selection from the set of random patches is proposed. In this method, using a recursive approach, optimal patches are selected from optimal features of training images by mutual information maximization feature selection. The performance of proposed algorithm in binary classification (existence or non-existence of objects in the images) is compared with HMAX and the superiority is proved.

Keywords: ***HMAX, object recognition, visual system, feature selection***

[*] **Corresponding author**
Address: Electronic Department, Faculty of Electrical Engineering, Iran University of Science and Technology, Postal Code: 1684613114, Tehran, Iran
Tel: *+98-21-77240487*
Fax: *+98-21-77240486*
E-mail: *hshsh@iust.ac.ir*

© Copyright 2016 ISBME, http://www.ijbme.org

مجله مهندسی پزشکی زیستی
دوره ۸، شماره ٤، زمستان ۱۳۹۳، ۳۸۳-۳۷۱
شناسه دیجیتال: *10.22041/ijbme.2014.14705*

گزینش ویژگی‌های بهینه در مدل بیولوژیکی بازشناسی اشیا با حداکثرکردن اطلاعات متقابل

محمد جزلائیان[۱]، هادی‌شهریار شاه‌حسینی[۲*]

[۱] کارشناس ارشد مهندسی برق، گروه الکترونیک، دانشکده مهندسی برق، دانشگاه علم و صنعت ایران، تهران
[۲] دانشیار، گروه الکترونیک، دانشکده مهندسی برق، دانشگاه علم و صنعت ایران، تهران

تاریخ ثبت در سامانه: ۱۳۹٤/٤/۱۰، نسخه اصلاح شده: ۱۳۹٤/٤/۱٦، پذیرش قطعی: ۱۳۹٤/٦/۸

چکیده

عملکرد سیستم بینایی انسان در بازشناسی اشیای مختلف، از عملکرد بهترین سیستم‌های بینایی ماشین بهتر است. به همین دلیل محققان حوزه‌ی بینایی ماشین و علوم اعصاب همواره به دنبال مدل‌کردن عملکرد سیستم بینایی انسان جهت استفاده در سیستم‌های بینایی ماشین هستند. یکی‌از قوی‌ترین مدل‌های محاسباتی که در این زمینه توسعه یافته مدل ***HMAX*** است. این مدل، برپایه‌ی عملکرد سلول‌های مغز انسان در گذرگاه بطنی قشر بینایی طراحی شده و چهار لایه محاسباتی دارد. در مرحله‌ی یادگیری این مدل، تعداد زیادی از قسمت‌های کوچک تصاویر در اندازه‌های مختلف، که تکّه نامیده می‌شوند، در موقعیت‌های کاملاً تصادفی از مجموعه‌ی تصاویر آموزشی استخراج می‌شوند. استخراج تصادفی و بی‌هدف تکّه‌ها یکی‌از ضعف‌های اصلی مدل ***HMAX*** است که باعث کاهش کارایی و افزایش بار محاسباتی آن می‌شود. در این مقاله، یک مدل جدید برای گزینش تکّه‌های مرتبط‌تر و حذف تکّه‌های زائد از مجموعه‌ی تکّه‌های تصادفی پیشنهاد شده است. در این مدل، با یک روند بازگشتی، تکّه‌های بهینه از روی ویژگی‌های بهینه‌ای انتخاب شدند که با الگوریتم گزینش ویژگی حداکثرکردن اطلاعات متقابل از بین مجموعه‌ی ویژگی‌های تصاویر آموزشی گزینش شدند. عملکرد مدل پیشنهادی در مسایل دوکلاسی تشخیص حضور یا عدم حضور یک شئ در تصویر با مدل اصلی ***HMAX*** مقایسه گردید و برتری آن به اثبات رسید.

کلیدواژه‌ها: ***بازشناسی اشیا، سیستم بینایی، گزینش ویژگی، مدل*** *HMAX*

* **نویسنده مسئول**
نشانی: تهران، میدان رسالت، خیابان هنگام، خیابان دانشگاه، دانشگاه علم و صنعت ایران، دانشکده مهندسی برق، گروه الکترونیک، کد پستی: ۱٦۸٤٦۱۳۱۱٤
تلفن: ۷۷۲٤۰٤۸۷(۲۱)۹۸+،
دورنگار: ۷۷۲٤۰٤۸٦(۲۱)۹۸+،
پست الکترونیکی: *hshsh@iust.ac.ir*

© Copyright 2016 ISBME, http://www.ijbme.org

۱- مقدمه

امروزه سیستم‌های بینایی ماشین کاربردهای گسترده‌ای در زندگی انسان پیدا کرده‌اند. سیستم‌های احراز هویت، بازرسی تولیدات صنعتی، کنترل خودکار وسایل نقلیه، بازرسی و نظارت زنده در محیط‌های باز، کنترل ترافیک، پردازش اسناد و کنترل ربات‌ها نمونه‌هایی از این کاربردها هستند. اگرچه این سیستم‌ها پیشرفت‌های زیادی کرده‌اند، اما هنوز سیستم بینایی ماشینی که بتواند با سیستم بینایی انسان رقابت کند، توسعه پیدا نکرده‌است.

سیستم بینایی انسان توانایی فوق‌العاده‌ای در شناسایی سریع و مقاوم اشیای مختلف در دنیای واقعی دارد. حتّا حضور اشیا در پس‌زمینه‌های درهم‌ریخته‌ای که در آن‌ها تغییرات شدیدی در مقیاس، موقعیت، زاویه دید و شدت روشنایی اشیا وجود دارد، تأثیری بر کارایی این سیستم ندارد و این سیستم به سرعت قادر است خود را با شرایط جدید منطبق کند. به هرحال، عملکرد بهترین سیستم‌های بازشناسی اشیا قابل مقایسه با عملکرد سیستم بینایی انسان نیست. توانایی فوق‌العاده بینایی انسان، محققان حوزه علوم اعصاب و بینایی ماشین را بر آن داشته تا به دنبال کشف سازوکار حاکم بر سیستم بینایی انسان و طراحی مدلی محاسباتی برمبنای این سازوکار جهت استفاده در سیستم‌های بینایی ماشین باشند.

هابل[1] و ویزل[2] اوّلین محققانی بودند که موفق به کشف ساختار سلسله‌مراتبی قشر بینایی مقدماتی در پستانداران شدند و مدلی کیفی برای توصیف عملکرد آن ارائه دادند [۱، ۲]. در مدل آن‌ها چندین ردیف سلول قرار دارند که هر ردیف به محرک‌های خاصی در تصویر پاسخ می‌دهند. به طور خلاصه، مدل توصیفی آن‌ها با سلول‌های متقارن شعاعی شروع می‌شود. هریک از این سلول‌ها به یک نقطه نورانی در تصویر پاسخ می‌دهند (مانند سلول‌های هسته‌ی جانبی خمیده LGN[3]). در ادامه، سلول‌های ساده[4] قرار دارند که با ادغام ورودی‌های دریافتی از سلول‌های متقارن شعاعی به خطوط راست و لبه‌ها در تصویر پاسخ می‌دهند. در مرحله بعدی ساختار سلسله‌مراتبی، سلول‌های پیچیده[5] قرار دارند که به خطوط و لبه‌های جهت‌دار پاسخ می‌دهند و جهت خطوط و لبه‌ها را از تصویر استخراج می‌کنند. امّا این سلول‌ها به مکان حضور خطوط و لبه‌ها در تصویر و حالت آن‌ها (مستقیم، خمیده، منحنی و ...) هیچ حساسیتی ندارند. در مرحله نهایی ساختار سلسله‌مراتبی، سلول‌های فراپیچیده[6] قرار دارند که هم به وضعیت و حالت خطوط و لبه‌ها پاسخ می‌دهند و هم نسبت به طول آن‌ها حساس هستند [۳].

پس‌از کشف ساختار سلسله‌مراتبی سلول‌های قشر بینایی توسط هابل و ویزل، مدل‌های محاسباتی فراوانی با الگوبرداری از این ساختار سلسله‌مراتبی جهت استفاده در سیستم‌های بینایی ماشینی پیشنهاد شده‌اند. مدل‌های نئوکاگنیترون[7] [٤]، ویزنت[8] [٥] و لامینارت[9] [٦، ۷] نمونه‌های از این مدل‌ها هستند. امّا این مدل‌های محاسباتی برخلاف سیستم بینایی انسان که توانایی فوق‌العاده‌ای در شناسایی طیف متنوعی از اشیا در دنیای واقعی دارد، اغلب تنها جنبه‌ی محاسباتی داشته و تنها برای کاربردهای خاصی توسعه یافته‌اند و قابلِ اعمال به تصاویر دنیای واقعی نیستند. این در حالی است که مهم‌ترین هدف در مدل‌سازی عملکرد قشر بینایی انسان دست‌یابی به سیستمی چندمنظوره است که همانند سیستم بینایی انسان توانایی تشخیص آسان و سریع طیف گسترده‌ای از اشیا را در دنیای واقعی و تحت شرایط مختلفی مثل تغییر در اندازه، موقعیت، شدت روشنایی، زاویه دید و مانند آن را داشته باشد. این شکاف تا حدودی با ارائه مدل HMAX[10] مرتفع شده و این مدل تا حد زیادی توانسته است این مشخصات را برآورده کند. این مدل، ابتدا به وسیله‌ی رایزنهابر[11] و پوجیو[12] پیشنهاد شد [۸] سپس، توسعه پیدا کرد [۹]. عملکرد این مدل در طیف متنوعی از کاربردهای مبتنی بر بازشناسی اشیا و روی پایگاه داده‌های مختلفی مورد آزمایش قرار گرفته و نشان داده شده که این مدل در همه‌ی این

[1] Hubel
[2] Wiesel
[3] Lateral Geniculate Nucleus Cells
[4] Simple Cells

[5] Complex Cells
[6] Hypercomplex
[7] Neocognitron
[8] Visnet
[9] LAMIRANT
[10] Hierarchical Model and X
[11] Riesenhuber
[12] Poggio

© Copyright 2016 ISBME, http://www.ijbme.org

کاربردها از بازشناسی اشیای ثابت در محیط‌های شلوغ گرفته تا درک مناظر پیچیده و چالش برانگیزی مثل خیابان کارایی فوق‌العاده‌ای از خود نشان می‌دهد [۹].

معماری مدل HMAX از چهار لایه محاسباتی سلسله‌مراتبی (S1، C1، S2 و C2) تشکیل شده‌است. در لایه‌های S به تصویر ورودی یک تابع گاوسین شکل و در لایه‌های C عمل‌گر ماکزیمم اعمال می‌شود.

از نظر مفاهیم بازشناسی الگو، مدل HMAX یک مدل استخراج ویژگی از تصاویر است. ویژگی‌هایی که این مدل از تصاویر استخراج می‌کند، نسبت به تغییر موقعیّت و مقیاس اشیا در تصویر مقاوم و تغییرناپذیر هستند. در مرحله یادگیری مدل HMAX، تعداد زیادی از قسمت‌های تصاویر در اندازه‌های مختلف، که تکّه[1] نامیده می‌شوند، در موقعیت‌های کاملا تصادفی از مجموعه تصاویر آموزشی استخراج می‌شوند. این کار در لایه‌ی C1 و در جهت‌ها و اندازه‌های مختلفی صورت می‌گیرد. این تکّه‌ها در لایه‌ی S2 مدل مورد استفاده قرار می‌گیرند. هر ویژگی استخراج‌شده با مدل HMAX، متناظر با یک تکّه است. این مجموعه تکه‌ها نقش مهمی در عملکرد نهایی سیستم بازشناسی مبتنی بر مدل HMAX دارند.

استخراج تصادفی و بی‌هدف تکّه‌ها یکی‌از ضعف‌های اصلی مدل HMAX است که هم بر کارایی آن از نظر دقت رده‌بندی تأثیر منفی دارد و هم باعث افزایش بار محاسباتی مدل می‌شود؛ چراکه با استخراج تصادفی تکّه‌ها امکان استخراج تکه‌هایی با الگوهای تکراری یا مشابه، به شدت افزایش می‌یابد. ویژگی‌های استخراج شده با استفاده‌از این الگوهای تکراری یا مشابه، نه تنها هیچ اطلاعات جدیدی به سیستم اضافه نمی‌کنند، بلکه ممکن است به دلیل افزونگی اطلاعات باعث اخلال در عملکرد رده‌بند سیستم یادگیری هم بشوند. از طرفی، هرچه افزونگی اطلاعات ویژگی‌ها افزایش یابد، طراح سیستم مجبور است برای دست‌یابی به دقت رده‌بندی مطلوب خود، تعداد ویژگی‌های استخراج شده از نمونه‌های آموزشی و به تبع آن تعداد تکّه‌های استخراج‌شده از این نمونه‌ها را افزایش دهد. این افزایش باعث بالارفتن بار محاسباتی سیستم و افزایش زمان پردازشی آن خواهد شد. به ویژه، زمانی که تعداد نمونه‌های آموزشی کم باشد، این مشکلات تشدید می‌شود. به همین دلیل، توجه بخشی از تحقیقات توسعه‌ای روی مدل HMAX به برطرف کردن این ضعف اختصاص یافته‌است.

در مدل پیشنهادی هوانگ[2] و هم‌کارانش [۱۰] که برای رده‌بندی صحنه در محیط‌های ویدیویی زنده با استفاده‌از مدل HMAX ارائه شده‌است، تکه‌ها به جای موقعیت‌های تصادفی از نواحی برجسته صحنه استخراج می‌شوند. میزان برجسته بودن یک ناحیه با اندازه‌گیری چگالی انرژی آن ناحیه به دست می‌آید. چگالی انرژی یک ناحیه با مجموع مقادیر پیکسل‌های آن ناحیه تعیین می‌شود.

حمیدی[3] و برجی[4] [۱۱] که مدل توسعه یافته‌ای از HMAX را برای بازشناسی ارقام دست‌نوشته به کار گرفتند، اگرچه تغییری در فرآیند استخراج تصادفی تکّه‌ها ایجاد نکردند، امّا با محدود کردن پاسخ لایه‌ی S2 به درصد مشخصی از اندازه‌ی تصویر ورودی که در محدوده‌ی اطراف مختصات ذخیره شده محل استخراج آن تکّه در مرحله یادگیری قرار دارد، موفق شدند اثرات نامطلوب استخراج تصادفی تکّه‌ها را تا حدودی خنثا کنند.

تریالت[5] و هم‌کارانش [۱۲] هم از شیوه‌ی مشابهی استفاده کردند. در روش پیشنهادی آن‌ها در لایه‌ی C2 برخلاف مدل اصلی HMAX که در آن برای هر تکّه، ماکزیمم سراسری طرح‌های خروجی لایه‌ی S2 محاسبه شده و به عنوان یک ویژگی در نظر گرفته می‌شود، برای هر تکه، ماکزیمم سراسری در ٦ ناحیه‌ی مختلف به مرکز مختصات مکانی آن تکّه در تصویر (که در مرحله‌ی یادگیری ذخیره شده) محاسبه می‌شود. شعاع این نواحی به ترتیب ۵، ۱۰، ۳۰، ۵۰ ، ۷۰ و ۱۰۰ درصد تصویر ورودی را پوشش می‌دهند. بنابراین، به ازای هر تکّه، ٦ ویژگی استخراج می‌شود.

کروپکا[6] و هم‌کارانش [۱۳] برای انتخاب تکّه‌های

[1] Patch

[2] Huang
[3] Hamidi
[4] Borji
[5] Theriault
[6] Krupka

© Copyright 2016 ISBME, http://www.ijbme.org

باکیفیت‌تر از میان تکّه‌های تصادفی، یک فرآیند یادگیری طراحی کردند. آن‌ها هر تکّه را با مجموعه‌ای از ویژگی‌ها توصیف کردند و با استفاده‌از یک الگوریتم یادگیری، تکّه‌ها را به دو دسته‌ی «خوب» و« بد» تقسیم کردند و نشان دادند که ویژگی‌های استخراجی با استفاده‌از تکّه‌های خوب، کارایی سیستم را افزایش می‌دهند.

قدرتی و همکارانش [۳] با استفاده‌از الگوریتم تکاملی ژنتیک (GA)[1]، سازوکاری برای انتخاب تکّه‌های بهینه از بین تکّه‌های تصادفی در مدل HMAX طراحی کردند. آن‌ها تصاویر موجود در پایگاه داده را به سه بخش مجموعه‌ی آموزشی، ارزیابی و آزمون تقسیم کردند و دقت رده‌بندی بر روی مجموعه‌ی ارزیابی را به عنوان تابع شایستگی الگوریتم ژنتیک در نظر گرفتند. در واقع در مدل پیشنهادی آن‌ها GA مجموعه‌ای از تکّه‌ها را انتخاب می‌کند که تابع شایستگی را حداکثر نماید. نتایج شبیه‌سازی نشان داد که مدل پیشنهادی آن‌ها، ضمن افزایش کارایی، به دلیل کاهش تعداد تکّه‌ها، باعث کاهش زمان پردازشی سیستم بازشناسی اشیای مبتنی بر مدل HMAX می‌شود.

تمرکز این مقاله نیز روی اصلاح فرآیند تصادفی استخراج تکّه‌ها در مدل HAMX قرار گرفته و هدف آن طراحی سازوکاری برای استخراج تکّه‌های بهینه و به تبع آن افزایش دقت رده‌بندی و کاهش بار محاسباتی در سیستم‌های بازشناسی اشیای مبتنی بر مدل HMAX است. برای دست‌یابی به این هدف، یک مدل جدید برمبنای روش‌های گزینش ویژگی برپایه‌ی نظریه‌ی اطلاعات پیشنهاد شد. در این مدل، با استفاده‌از یک روند بازگشتی، تکّه‌های بهینه از روی ویژگی‌های بهینه‌ای انتخاب شدند که به وسیله‌ی یک الگوریتم گزینش ویژگی، ازبین مجموعه ویژگی‌های تصاویر آموزشی گزینش شدند. برای گزینش ویژگی از الگوریتم حداکثرکردن اطلاعات متقابل (MIM)[2] استفاده شد.

در ادامه، ابتدا معماری مدل اصلی HMAX و سازوکار حاکم بر هریک از لایه‌های آن مورد بررسی قرار گرفته است. سپس، مدل پیشنهادی برای گزینش تکّه‌های بهینه معرفی شده است. در بخش آخر، عملکرد روش پیشنهادی روی چند پایگاه داده تصاویر مورد ارزیابی قرار گرفته و با عملکرد مدل اصلی HMAX مقایسه شده و کارایی روش پیشنهادی اثبات گردیده است.

۲- مواد و روش‌ها

۲-۱- ساختار مدل اصلی HMAX

در این بخش سازوکار حاکم بر واحدهای محاسباتی مدل اصلی HMAX معرفی می‌شود. مدل HMAX دارای چهار واحد محاسباتی S1، C1، S2 و C2 است که با الگوبرداری از عملکرد سلول‌های ساده و پیچیده در گذرگاه بطنی[3] قشر بینایی طراحی شده‌اند. ورودی مدل، یک تصویر خاکستری و خروجی آن یک بردار ویژگی C2 است که مشخصات تصویر ورودی را به صورت کمّی توصیف می‌کند و از آن برای آموزش سیستم یادگیری استفاده می‌شود.

در شکل (۱)، معماری مدل HMAX به همراه تغییراتی که هریک از واحدها روی یک تصویر نمونه اعمال می‌کنند، نشان داده شده‌است. جزییات هریک از واحدها به طور خلاصه به شرح ذیل است:

واحدهای S1: در این واحدها به تصویر خاکستری ورودی مجموعه‌ای از فیلترهای آشکارساز لبه اعمال می‌شود. این فیلترها برمبنای توابع گابور دوبعدی ساخته شده‌اند که مدل مناسبی برای میدان پذیرش[4] سلول‌های ساده در قشر بینایی مغز انسان هستند. توابع گابور با رابطه‌های (۱)، (۲) و (۳) توصیف می‌شوند [۹]:

$$g(x,y) = \exp\left(-\frac{x_0^2+\gamma^2 y_0^2}{2\sigma^2}\right) \times \cos\left(\frac{2\pi}{\lambda}x_0\right) \qquad (۱)$$

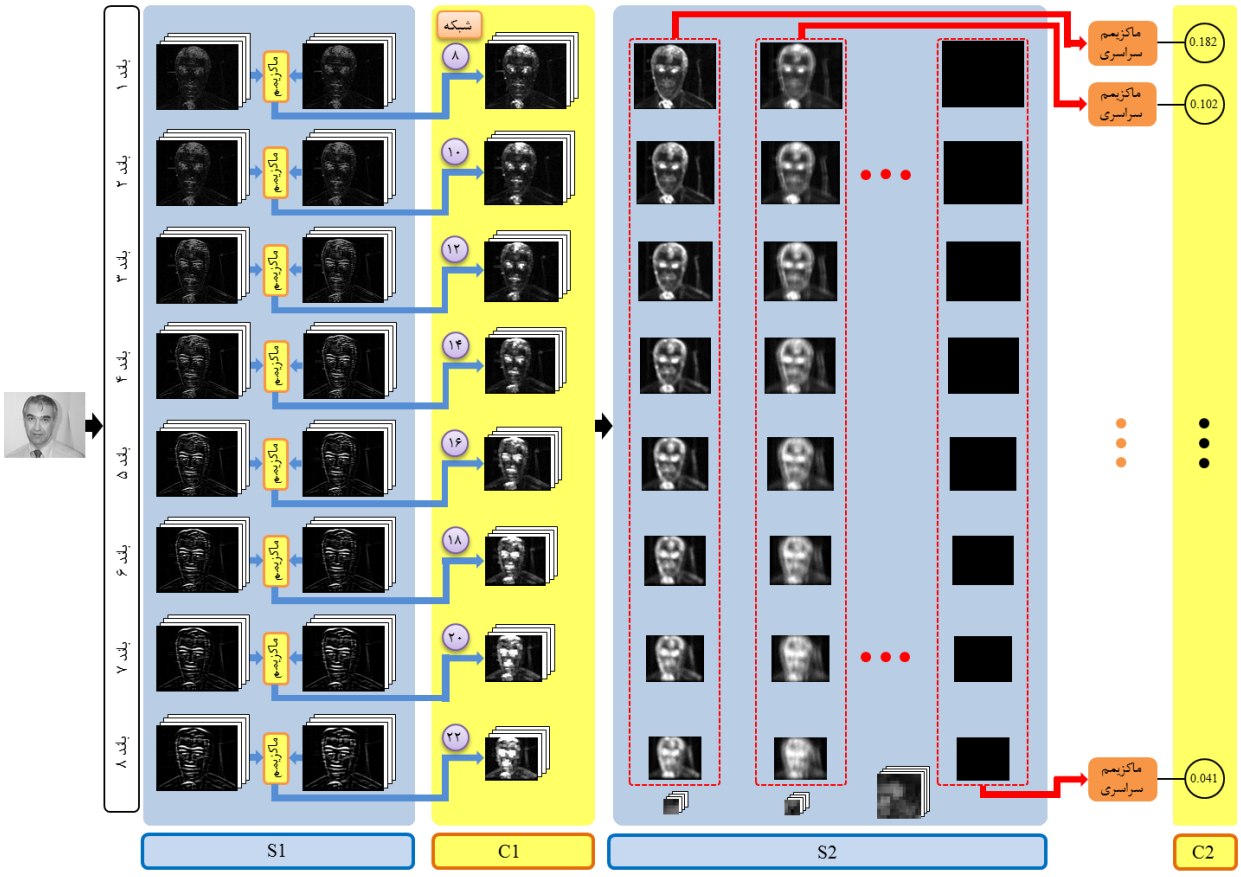
$$x_0 = x\cos\theta + y\sin\theta \qquad (۲)$$

$$y_0 = -x\sin\theta + y\cos\theta \qquad (۳)$$

که در آن γ ضریب جهت شیب، θ جهت (زاویه چرخش)، σ پهنای باند مؤثّر و λ طول موج فیلتر و x و y مختصات مکانی هستند. این پارامترها مشخصات مکانی میدان دریافت

[1] Genetic Algorithm
[2] Mutual Information Maximisation
[3] Ventral Stream
[4] Receptive Field

© Copyright 2016 ISBME, http://www.ijbme.org



شکل (۱)- معماری مدل HMAX

واحدها را تعیین می‌کنند. همه پارامترها براساس آزمایش‌های تجربی، به گونه‌ای تنظیم می‌شوند که مدل مناسبی برای سلول‌های ساده باشند.

پاسخ یک واحد S1 هنگامی بیشینه است که جهت محرّک ورودی (مثلا یک خط جهت‌دار یا یک لبه) با جهت فیلتر مطابقت داشته باشد. هنگامی که جهت محرک و فیلتر، تفاوت زیادی داشته باشند پاسخ صفر می‌شود. اندازه‌ی فیلترهای S1 به گونه‌ای انتخاب شده است که هرمی از مقیاس‌ها را ایجاد کند، این فیلترها از اندازه‌ی ۷×۷ شروع شده و با گام‌هایی دوتایی تا اندازه‌ی ۳۷×۳۷ ادامه می‌یابند؛ بنابراین، فیلترها دارای ۱۶ اندازه‌ی مختلف هستند. برای آنکه تعداد واحدها از کنترل خارج نشود، تنها چهار جهت ۰، ۴۵، ۹۰ و ۱۳۵ درجه برای فیلترها در نظر گرفته شده است. بدین ترتیب، در واحدهای S1 در مجموع ۶۴ فیلتر مکانی مختلف به تصویر ورودی اعمال می‌شود (۱۶ مقیاس × ۴ جهت) و ۶۴ تصویر جدید از تصویر ورودی ساخته می‌شود.

واحدهای C1: عملکرد واحدهای C1، با عملکرد سلول‌های پیچیده قشر بینایی مطابقت دارد. این واحدها نسبت به واحدهای S1 در مقابل تغییر اندازه و تغییر مکان شئ در تصویر مقاوم‌تر هستند. این مشخصه با استفاده‌از یک عمل‌گر ماکزیمم در لایه‌ی C1 برآورده می‌شود.

خروجی‌های لایه‌ی C2 با اعمال عمل‌گر ماکزیمم به خروجی‌های لایه‌ی S1 به دست می‌آیند. در هر واحد C1، دو طرح S1 با جهت یکسان و باند مقیاس مشابه، از طریق عمل‌گر ماکزیمم درهم ادغام شده و تبدیل به یک طرح C1 می‌شوند. هر باند مقیاس، شامل دو اندازه فیلتر مجاور است؛ بنابراین، به علّت وجود ۱۶ اندازه‌ی فیلتر مختلف، هشت باند مختلف به وجود می‌آیند. برای نمونه، باند مقیاس شماره‌ی ۱ شامل فیلترهای S1 با اندازه‌های ۷×۷ و ۹×۹ است. عمل ادغام برای هر باند مقیاس در پنجره‌ای با ابعاد متفاوت صورت می-

© Copyright 2016 ISBME, http://www.ijbme.org

گیرد. ابعاد این پنجره $N_s \times N_s$ است و براساس انـدیس بانـد مقیاس و با توجّه به جدول (۱) تعیین می‌شود [۹]. این فرآیند برای همه‌ی باند مقیـاس‌هـا و در هـر چهـار جهـت بـه طـور مستقل اجرا می‌شـود. ادغـام دو مقیـاس مجـاور، مقاومـت در مقابل تبدیلات دوبعدی را از لایه‌ی S1 تا C1 افزایش می‌دهد. بنابراین، ٦٤ تصویر ورودی از لایه‌ی S1 در ایـن لایـه بـه ۳۲ تصویر تبدیل می‌شود.

واحدهای S2: واحدهای موجود در لایـه‌ی S2، ورودی‌هـای دریافتی از واحدهای C1 را در امتداد هر چهـار جهـت ادغـام می‌کنند. واحـدهای S2 رفتـاری مشـابه واحـدهای تـابع پایـه شعاعی (RBF)[1] دارند. پاسخ خروجی هر واحـد S2 نتیجـه‌ی اعمال یک روش گوسین مانند بر فاصله‌ی اقلیدسی بـین یـک ورودی از لایه‌ی قبل و یک تکّه تصویر از پیش ذخیـره شـده است. بدین معنا که برای وصله‌ی تصویر X از لایه‌ی قبلی C1 در مقیاس مشخص S (هر وصله شامل چهار طرح C1 در یک باند و چهار جهت مختلف است)، پاسخ واحد متنـاظر S2 کـه با r نمایش داده می‌شود، از رابطه‌ی (٤) به دست می‌آید:

$$r = \exp(-\beta \|X - P_i^2\|) \qquad (٤)$$

که در آن β میزان تیزی تابع گوسین را کنترل مـی‌کنـد و P_i یکی‌از N تکّه‌ای است که در طول مرحله یـادگیری اسـتخراج شده‌است. ابعاد هر طرح S2 برابر ابعاد طرح‌هـای C2 متنـاظر است. برای هر وصله تصویر ورودی X از لایه قبل، طرح‌هـای S2 برای همـه‌ی تکّـه‌هـای P_i و در هـر هشـت باند مقیـاس محاسبه می‌شود. بنـابراین، خروجـی واحـدهای S2 شـامل N دسته است که در هر دسـته، هشـت تصـویر وجـود دارد. هـر تصویر در هر دسته، در ارتباط با یک باند مقیاس اسـت و هـر دسته به یک تکّه P_i مربوط می‌شود.

واحدهای C2: در هریک از این واحدها کـه آخـرین مرحلـه سلسله‌مراتب هستند، ماکزیمم سراسری بین هر هشت تصـویر موجود در یک دسته محاسبه می‌شـود. بـدین صـورت کـه در لایه‌ی S2 میزان تطـابق هـر یـک از تکّـه‌هـای P_i بـا تصـویر ورودی در همه‌ی مقیاس‌ها محاسبه می‌شـود و در لایـه‌ی C2 بهترین تطابق نگه داشته شده و بقیه کنار گذاشته می‌شـوند تـا مقاومت مدل نسبت به تغییـر مکـان و تغییـر مقیـاس شـئ در تصویر تضمین شود. خروجی این لایـه، بـرداری بـا طـول N است که N تعداد تکّه‌هایی است که در طول مرحله یـادگیری استخراج شده‌اند. در واقع بردار خروجـی C2، بـردار ویژگـی استخراج شده به وسیله‌ی مدل HMAX برای تصـویر ورودی است.

مرحله یـادگیری: فرآینـد یـادگیری در مـدل HMAX شـامل انتخاب مجموعه‌ای از N تکّه P_i برای واحدهای S2 است. این کار با استفاده‌از یک فرآیند نمونه‌برداری ساده انجام مـی‌شـود، بدین صورت که تعداد زیادی از تکّه‌ها در اندازه‌های مختلـف و در موقعیـت‌هـای مکـانی تصـادفی از تصـاویر آموزشـی استخراج می‌شوند. استخراج این تکّه‌ها در سطح لایـه‌ی C1 و در همه‌ی جهت‌ها صورت مـی‌گیـرد، یعنـی یـک تکّـه P_0 بـا اندازه‌ی $n \times n$ دارای $n \times n \times 4$ عنصر است.

در مدل HMAX ارائه شده [۹] که مبنای کار ایـن مقالـه اسـت، تکّـه‌هـا در چهـار انـدازه‌ی مختلـف ٤، ۸، ۱۲ و ۱٦ استخراج شدند.

۲-۲- مدل پیشنهادی برای استخراج تکّه‌های بهینه

اسـتخراج تصـادفی تکّـه‌هـا در مرحلـه‌ی یـادگیری، یکـی از ضعف‌های اصلی مدل HMAX است کـه تـأثیر نـامطلوبی بـر کـارایی و بـار محاسـباتی آن مـی‌گـذارد. اسـتخراج تصـادفی احتمال حضور تکّه‌های زائد و نامرتبط را در مجموعه تکّه‌هـا افزایش می‌دهد. منظور از تکّه‌ی زائد، تکّه‌ای است که ویژگـی C2 حاصل از آن هیچ اطلاعات جدیدی بـه فرآینـد یـادگیری اضافه نمی‌کند. زائد بودن تکّه‌ها ناشی از استخراج تکّـه‌هـایی با الگوهای تکراری یا مشابه است که باعث می‌شود، ویژگـی-های حاصل از آن‌ها هم تکراری یا مشابه باشند. منظور از تکّه نامرتبط، تکّه‌ای است که ویژگی C2 حاصل از آن هیچ کمکی به سیسـتم یـادگیری در جداسـازی و تفکیـک بـین نمونـه‌هـا نمی‌کند. نامرتبط بودن تکّه‌ها ناشی از استخراج تکـه‌هـایی بـا الگوهای یکنواخت از نواحی پس‌زمینه است که ویژگـی‌هـای حاصل از آن‌ها برای همه‌ی نمونه‌ها تقریباً یکسان است و بین

[1] Radial Basis Function

© Copyright 2016 ISBME, http://www.ijbme.org

نمونه‌ها تفکیکی قائل نمی‌شوند.

افزایش افزونگی اطلاعات ویژگی‌ها هم می‌تواند باعث اخلال در عملکرد رده‌بند سیستم یادگیری شود و هم طراح سیستم مجبور است، برای دست‌یابی به دقت رده‌بندی مطلوب خود، تعداد ویژگی‌های استخراج شده از نمونه‌های آموزشی و به تبع آن تعداد تکّه‌های استخراج شده از این نمونه‌ها را افزایش دهد. این افزایش باعث بالارفتن بار محاسباتی سیستم و افزایش زمان پردازشی آن خواهد شد. به ویژه، زمانی که تعداد نمونه‌های آموزشی کم باشد، این مشکلات تشدید می‌شود.

در سیستم‌های بازشناسی اشیایی که از مدل HMAX برای استخراج ویژگی استفاده می‌کنند، هر ویژگی استخراجی C2 متناظر با تکّه‌ای ذخیره شده در لایه‌ی S2 است که قبلاً در مرحله‌ی یادگیری به طور تصادفی از بین همه‌ی نمونه‌های آموزشی استخراج شده‌است. به علّت وجود این تناظر یک به یک، به جای جست‌وجوی تکّه‌های بهینه، می‌توان ویژگی‌های بهینه را جست‌وجو کرد و از روی مجموعه‌ی ویژگی‌های بهینه به مجموعه‌ی تکّه‌های بهینه رسید. منظور از ویژگی‌های بهینه، آن دسته از ویژگی‌هایی هستند که بین الگوهای موجود در کلاس‌های متفاوت بیش‌ترین تمایز و بین الگوهای موجود در یک کلاس مشابه، بیش‌ترین تشابه را ایجاد کنند. با استفاده‌از ویژگی‌های بهینه، سیستم یادگیری قادر است مجموعه تصاویر آموزشی را به بهترین شکل یاد بگیرد. مدل پیشنهادی این مقاله برمبنای این ایده توسعه یافته است.

در مدل پیشنهادی، فرآیند یادگیری سیستم یک فرآیند بازگشتی است. بدین ترتیب که ابتدا با استفاده‌از مدل HMAX تعداد زیادی ویژگی از مجموعه تصاویر آموزشی استخراج می‌شود، سپس با استفاده‌از یک الگوریتم گزینش ویژگی، مجموعه ویژگی‌های بهینه از بین این ویژگی‌ها گزینش می‌شوند. درنهایت، سیستم با استفاده‌از این ویژگی‌ها آموزش می‌بیند. مجموعه تکّه‌های متناظر با این ویژگی‌های بهینه هم در لایه‌ی S2 مدل HMAX ذخیره می‌شوند تا برای استخراج ویژگی از نمونه‌های ورودی جدید مورد استفاده قرار گیرند. در واقع با استفاده‌از گزینش ویژگی، مجموعه‌ی ویژگی‌های زائد و نامرتبطی که در مجموعه‌ی ویژگی‌ها وجود دارند و از تکّه‌های زائد و نامرتبط به وجود آمده‌اند، از سیستم یادگیری حذف می‌شوند و تنها ویژگی‌های بهینه و به تبع آن، تکّه‌های بهینه در سیستم باقی می‌مانند.

معماری مدل پیشنهادی مقاله در شکل (۲) نشان داده شده است. مراحل آموزش سیستم بازشناسی اشیایی با K ویژگی به صورت ذیل است:

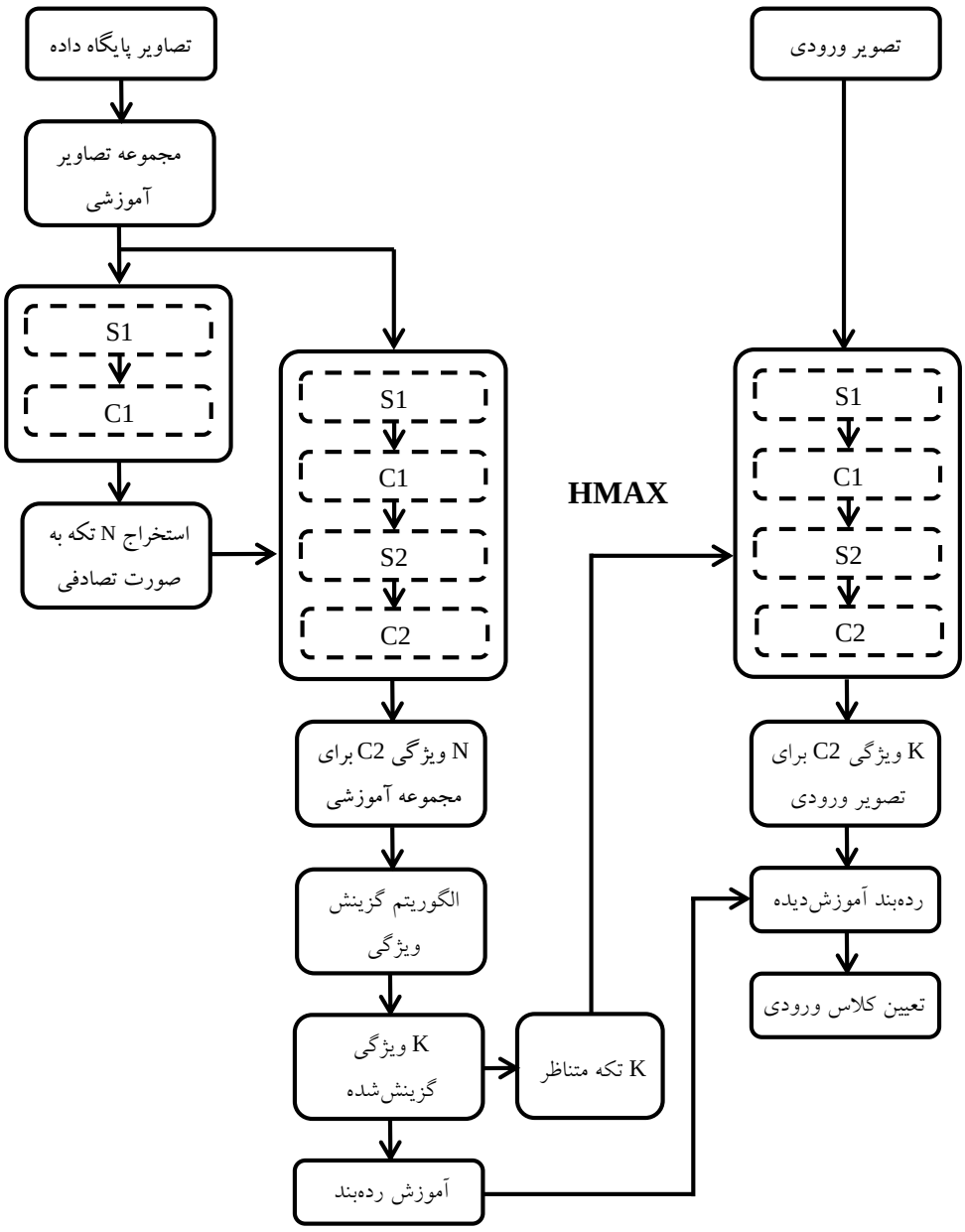
۱- *استخراج مخزنی از تکه‌ها از لایه‌ی C1*: در ابتدا، همه‌ی تصاویر مجموعه‌ی آموزشی با مجموعه‌ای از فیلترهای گابور تجزیه و تحلیل شده و خروجی‌های S1 ایجاد می‌شوند. خروجی‌های S1 به لایه‌ی C1 اعمال شده و طرح‌های C1 ساخته می‌شوند. مرحله بعدی استخراج تکّه است، در این مرحله، تعداد N تکّه که $N \gg K$ در چهار اندازه (٤، ٨، ١٢، ١٦) و در چهار جهت (°۰، °٤٥، °۹۰، °۱۳۵) از موقعیت‌های تصادفی در طرح‌های C1 استخراج می‌شوند.

۲- *محاسبه‌ی ویژگی‌های C2*: با استفاده‌از تکّه‌های انتخاب شده در مرحله یک، پاسخ‌های S2 و سپس ویژگی‌های C2 برای همه تصاویر موجود در مجموعه آموزشی محاسبه می‌شوند. چون تعداد N تکه در مرحله یک انتخاب شده، برای هر تصویر ورودی تعداد ویژگی‌های C2 برداری با طول $1 \times N$ خواهد بود. اگر M تصویر در مجموعه‌ی آموزشی وجود داشته باشد، ابعاد ماتریس C2 برابر $M \times N$ خواهد شد. هر سطر از این ماتریس متناظر با پاسخ یک تکّه خاص است.

۳- *اعمال الگوریتم گزینش ویژگی*: با استفاده‌از الگوریتم گزینش ویژگی، تعداد K ویژگی از میان N ویژگی استخراجی در مرحله دو گزینش می‌شود. با این کار ابعاد ماتریس C2 از $M \times N$ به $M \times K$ تقلیل می‌یابد.

٤- *ذخیره تکّه‌های متناظر*: K تکّه‌ی متناظر با K ویژگی گزینش شده، برای استخراج ویژگی‌های C2 از نمونه‌های ورودی جدید در لایه‌ی S2 مدل ذخیره می‌شود.

© Copyright 2016 ISBME, http://www.ijbme.org



شکل (۲)- معماری مدل پیشنهادی

۵-*آموزش رده‌بند*: با استفاده‌از K ویژگی C2 استخراج شده از تصاویر آموزشی، رده‌بند مورد استفاده در سیستم آموزش داده می‌شود.

در این تحقیق، برای گزینش ویژگی از الگوریتم بیشینه‌کردن اطلاعات متقابل (MIM) استفاده شد. این الگوریتم مبتنی بر نظریه‌ی اطلاعات است. به همین دلیل، برای توضیح ساختار الگوریتم ابتدا برخی مفاهیم نظریه‌ی اطلاعات معرفی می‌شود.

جزء اساسی نظریه‌ی اطلاعات، آنتروپی یک متغیر تصادفی است. آنتروپی متغیر تصادفی X که با H(X) نمایش داده می‌شود، در واقع میزان عدم قطعیت موجود در توزیع X را کمّی‌سازی می‌کند و با رابطه‌ی (۵) تعریف می‌شود [۱٤]:

$$H(X) = \sum_{x\epsilon\mathcal{X}} p(x) \log p(x) \qquad (۵)$$

که در آن x نشان دهنده‌ی یک مقدار ممکن برای متغیر X است که می‌تواند از مجموعه‌ی الفبای $\mathcal{X}$ انتخاب گردد. با مبانی نظریه‌ی احتمال، آنتروپی را می‌توان به پیشامدهای دیگر هم مشروط کرد. آنتروپی X به شرط Y با رابطه‌ی (٦) تعریف

© Copyright 2016 ISBME, http://www.ijbme.org

می‌شود:

$$H(X|Y) = \sum_{y\in Y} p(y) \sum_{x\in \mathcal{X}} p(x|y) \log p(x|y) \quad (٦)$$

که در آن X و Y دو متغیر تصادفی روی مجموعه‌ی الفباهای $\mathcal{X}$ و $\mathcal{Y}$ هستند. آنتروپی شرطی را می‌توان به عنوان مقدار عدم قطعیت باقی‌مانده در X بعد از مشخص شدن نتیجه آزمایش Y در نظر گرفت.

با استفاده‌از دو تعریف فوق می‌توان اطلاعات متقابل بین X و Y را که در واقع همان اطلاعات به اشتراک گذاشته با X و Y است، به صورت رابطه‌ی (۷) تعریف کرد [۱٤]:

$$I(X;Y) = H(X) - H(X|Y)$$

$$= \sum_{x\in \mathcal{X}} \sum_{y\in Y} p(xy) \log \frac{p(xy)}{p(x)\, p(y)} \quad (٧)$$

روند الگوریتم MIM در شکل (۳) نشان داده شده‌است. با فرض اینکه تعداد ویژگی‌های مورد نظر برای آموزش سیستم K ویژگی باشد، ابتدا N تکه که $N \gg K$ به طور تصادفی از تصاویر آموزشی استخراج می‌شود. سپس، اطلاعات متقابل بین هر ویژگی و مجموعه برچسب کلاس‌ها C محاسبه می‌شود. درنهایت، K ویژگی با بالاترین اطلاعات متقابل با مجموعه‌ی برچسب کلاس‌ها انتخاب می‌شود و K تکّه‌ی متناظر با این ویژگی‌ها در لایه‌ی S2 مدل HMAX ذخیره می‌شود [۱٤].

۳- یافته‌ها و بحث

روش پیشنهادی همانند اغلب تحقیقات صورت گرفته در این زمینه، روی مسایل دوکلاسی تشخیص حضور یا عدم حضور یک شئ در تصویر ورودی ارزیابی شده است. در این نوع مسایل، دو دسته تصویر در پایگاه داده وجود دارد: تصاویر مثبت و تصاویر منفی. تصاویر مثبت، تصاویری هستند که شئ مورد نظر به فرم‌های گوناگون در آن‌ها حضور دارد و تصاویر منفی، تصاویری هستند که شئ مورد نظر در آن‌ها حضور ندارد. معمولاً تصاویر منفی شامل تصاویر گوناگونی از مناظر و پس زمینه‌ها هستند که به صورت تصادفی جمع‌آوری می‌شوند و از الگوی خاصی پیروی نمی‌کنند.

ورودی: N ویژگی تصادفی
خروجی: K تکه انتخاب‌شده

۱) مقداردهی اولیه:
a) قرار بده: $F = \{f_i \mid i = 1, \dots, N\}$، مجموعه‌ی اولیه با N ویژگی؛
b) قرار بده: $S = \{\emptyset\}$، مجموعه‌ی خالی.
۲) محاسبه‌ی اطلاعات متقابل:
a) برای هر $f_i \in F$، مقدار $I(f_i; C)$ را محاسبه کن؛
۳) N ویژگی موجود در مجموعه ویژگی‌ها را بر حسب مقدار $I(f_i; C)$ آن‌ها به ترتیب نزولی مرتب کن؛
۴) K ویژگی با بالاترین مقدار $I(f_i; C)$ را انتخاب کن و در S قرار بده.

شکل (۳)- الگوریتم حداکثرکردن اطلاعات متقابل

برای ارزیابی و مقایسه‌ی عملکرد مدل پیشنهادی با مدل اصلی HMAX در مسایل دوکلاسی از مجموعه‌ی داده‌های CalTech5 و Caltech101 استفاده شد. از مجموعه‌ی داده CalTech5، دو پایگاه داده هواپیما و برگ به عنوان تصاویر مثبت و پایگاه داده پس‌زمینه به عنوان تصاویر منفی انتخاب شدند. از مجموعه‌ی داده CalTech101 دو پایگاه داده پیانو و دارسان[1] به عنوان تصاویر مثبت و پایگاه داده پس‌زمینه به عنوان تصاویر منفی به کار رفتند.

قبل‌از انجام آزمایش‌ها، همه تصاویر به تصاویر خاکستری تبدیل شدند و ارتفاع همه‌ی آن‌ها به ۱٤۰ پیکسل تغییر اندازه داده شد. عرض تصاویر هم به نسبت ارتفاع تغییر یافت، به گونه‌ای که نسبت ابعاد تصویر ثابت باقی بماند.

برای همه‌ی آزمایش‌ها تعداد تصاویر آموزشی در کلاس منفی ۵۰ تصویر، تعداد تصاویر آزمون در کلاس منفی ۵۰ تصویر و تعداد تصاویر آزمون در کلاس مثبت هم ۵۰ تصویر بود که در هر آزمایش به طور تصادفی از پایگاه داده انتخاب شدند. تعداد تصاویر آموزشی در کلاس مثبت در هر آزمایش بین ۳، ٦، ۱۵ و ۳۰ تصویر تغییر کرد. تکّه‌ها مشابه مدل اصلی HMAX [۹] تنها از تصاویر آموزشی مثبت استخراج شدند.

برای مقایسه‌ی عملکرد روش پیشنهادی و مدل اصلی HMAX از معیار نرخ بازشناسی روی تصاویر آزمون، یعنی

[1] Bonsai

© Copyright 2016 ISBME, http://www.ijbme.org

نسبت تعداد تشخیص‌های درست نسبت به تعداد کل تشخیص‌ها استفاده شد. برای تأیید صحت نتایج، هر آزمایش به طور تصادفی ۱۰ بار تکرار شده و میانگین نرخ بازشناسی در این ۱۰ تکرار گزارش شده است. برای رده‌بندی نمونه‌ها از رده‌بند SVM دوکلاسی استفاده و شبیه‌سازی‌ها با نرم‌افزار متلب انجام شد.

در جدول (۱) نرخ بازشناسی در مدل پیشنهادی و مدل اصلی HMAX برای پایگاه داده‌های برگ، هواپیما، پیانو و دارسان مقایسه شده‌است.

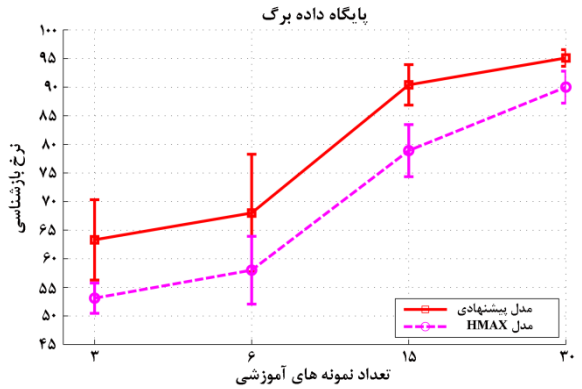
جدول (۱)- مقایسه نرخ بازشناسی در مدل پیشنهادی و مدل اصلی HMAX

پایگاه داده	نرخ بازشناسی در مدل اصلی HMAX	نرخ بازشناسی در مدل پیشنهادی
برگ	۹۵٫٤	۹۱٫۵
هواپیما	۹۲٫۸	۸۹٫۰
پیانو	۸٦٫۷	۸۲٫۱
دارسان	۷٤٫۱	۷۱٫۳

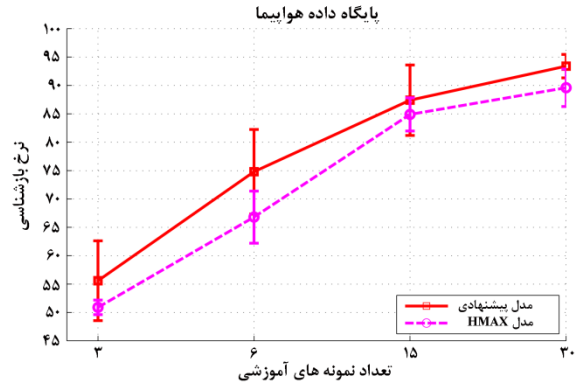
تعداد تصاویر آموزشی مثبت ۳۰ تصویر بود. ابتدا ۱۰۰۰۰ تکّه به طور تصادفی از تصاویر آموزشی مثبت استخراج شد (N=10000). سپس، تعداد ٦۰ ویژگی (K=60) در مدل پیشنهادی با الگوریتم MIM و در مدل اصلی HMAX به طور تصادفی انتخاب شد. نتایج نشان داد که عملکرد مدل پیشنهادی در هر چهار پایگاه داده بر عملکرد مدل اصلی HMAX برتری دارد.

برای ارزیابی دقیق‌تر کارایی مدل پیشنهادی، تأثیر تغییر تعداد نمونه‌های آموزشی و نیز تأثیر تغییر تعداد ویژگی‌ها بر عملکرد مدل پیشنهادی بررسی شد و با شرایط مشابه در مدل اصلی HMAX مقایسه گردید.

در ابتدا تأثیر تغییر تعداد نمونه‌های آموزشی روی عملکرد مدل پیشنهادی مورد بررسی قرار گرفت و با مدل اصلی HMAX مقایسه شد. نتایج شبیه‌سازی در تعداد نمونه‌های آموزشی مثبت مختلف ۳، ٦، ۱۵، ۳۰ و با انتخاب $N = 1000$ و $K = 50$ برای پایگاه داده برگ در شکل (٤) و برای پایگاه داده هواپیما در شکل (۵) نشان داده شده‌است. محور افقی تعداد نمونه‌ها و محور عمودی درصد نرخ بازشناسی را نشان می‌دهد. برای هر تعداد نمونه آموزشی، میانگین و انحراف معیار درصد نرخ بازشناسی برای ۱۰ اجرای مختلف گزارش شده است. همان طور که مشاهده می‌شود، نرخ بازشناسی در مدل پیشنهادی در همه موارد نسبت به مدل اصلی HMAX برتری دارد.



شکل (٤)- مقایسه عملکرد روش پیشنهادی و مدل اصلی HMAX در پایگاه داده برگ با تغییر تعداد نمونه‌های آموزشی



شکل (۵)- مقایسه عملکرد روش پیشنهادی و مدل اصلی HMAX در پایگاه داده هواپیما با تغییر تعداد نمونه‌های آموزشی

در پایگاه داده برگ بیش‌ترین بهبود برای ۱۵ نمونه آموزشی به دست آمد و مقدار آن ۱۲ درصد است. در پایگاه داده هواپیما بیش‌ترین بهبود برای ٦ نمونه آموزشی به دست آمد و مقدار آن ۹ درصد است.

همان طور که در نمودارها مشاهده می‌شود، با افزایش

© Copyright 2016 ISBME, http://www.ijbme.org

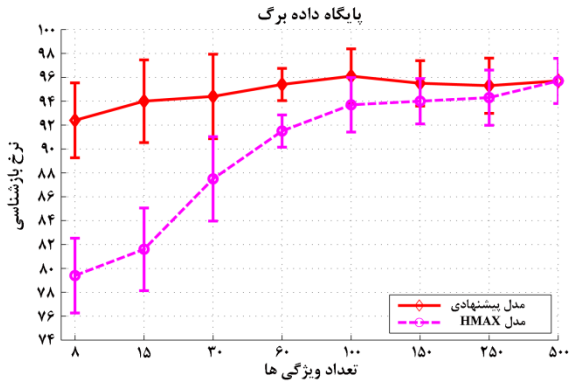
تعداد نمونه‌های آموزشی، میانگین نرخ بازشناسی هم در مدل پیشنهادی و هم در مدل HMAX افزایش یافته‌است؛ چراکه با افزایش تعداد نمونه‌های آموزشی رده‌بند امکان بیش‌تری برای یادگیری و جداسازی نمونه‌ها دارد. هم‌چنین با افزایش تعداد نمونه‌های آموزشی اختلاف نرخ بازشناسی در مدل پیشنهادی و مدل اصلی HMAX کم‌تر می‌شود؛ چراکه با افزایش تعداد نمونه‌ها احتمال استخراج تکّه‌های تکراری و به تبع آن تعداد ویژگی‌های زائد کم‌تر شده و گزینش ویژگی‌های بهینه تأثیر کم‌تری بر کارایی سیستم بازشناسی خواهد داشت.

در ادامه تأثیر تغییر تعداد ویژگی‌ها بر عملکرد مدل پیشنهادی بررسی شده و با شرایط مشابه در مدل HMAX مقایسه شده است.

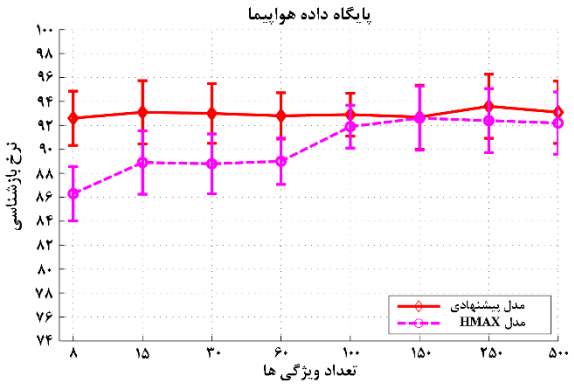
شکل‌های (٦) و (٧) به ترتیب نتایج شبیه‌سازی با N = 1000 روی پایگاه داده برگ و هواپیما را نشان می‌دهد که در آن K بین ٨، ١٥، ٣٠، ٦٠، ١٠٠، ١٥٠، ٢٥٠ و ٥٠٠ تغییر کرده‌است. محور افقی تعداد ویژگی‌ها و محور عمودی درصد نرخ بازشناسی است.

برای هر تعداد ویژگی، میانگین و انحراف معیار درصد نرخ بازشناسی برای ١٠ اجرای مختلف گزارش شده است. همان طور که مشاهده می‌شود، عملکرد مدل پیشنهادی در همه‌ی موارد نسبت به عملکرد مدل اصلی HMAX برتری دارد. به طوری که برای پایگاه داده برگ، نرخ بازشناسی مدل پیشنهادی در ٦٠ ویژگی با نرخ بازشناسی نهایی در مدل HMAX با ٥٠٠ ویژگی برابر شده و این بار محاسباتی را یک دهم برابر کاهش داده‌است. برای پایگاه داده هواپیما نرخ بازشناسی در ١٥ ویژگی با نرخ بازشناسی نهایی در مدل HMAX با ١٥٠ ویژگی برابر شده‌است.

همان طور که در نمودارها مشاهده می‌شود، هرچه تعداد ویژگی‌ها افزایش می‌یابد اختلاف نرخ بازشناسی در مدل پیشنهادی و مدل اصلی HMAX کم‌تر می‌شود؛ چراکه با ثابت بودن مقدار N و افزایش تعداد ویژگی‌ها، فضای جست‌وجوی ویژگی‌های مؤثّرتر برای الگوریتم گزینش ویژگی کوچک‌تر می‌شود و عملکرد دو مدل به هم نزدیک‌تر است.



شکل (٦)- مقایسه عملکرد روش پیشنهادی و مدل اصلی HMAX در پایگاه داده برگ با تغییر تعداد ویژگی‌ها



شکل (٧)- مقایسه عملکرد روش پیشنهادی و مدل اصلی HMAX در پایگاه داده هواپیما با تغییر تعداد ویژگی‌ها

٤- نتیجه‌گیری

یک مدل جدید برای اصلاح فرآیند استخراج تصادفی تکّه‌ها در مدل HMAX و گزینش تکّه‌های بهینه از میان تکّه‌های تصادفی پیشنهاد شد. استخراج تصادفی تکّه‌ها در مدل HMAX باعث کاهش کارایی و افزایش بار محاسباتی این مدل می‌شود.

مدل پیشنهادی، یک روش بازگشتی است که در آن تکّه‌های بهینه از روی ویژگی‌های بهینه‌ای انتخاب می‌شوند که با الگوریتم گزینش ویژگی حداکثرکردن اطلاعات متقابل (MIM) ازبین مجموعه‌ی ویژگی‌های تصاویر آموزشی گزینش شده‌اند.

با فرض اینکه تعداد ویژگی‌های مورد نظر برای آموزش سیستم K ویژگی باشد، ابتدا N تکّه که $N \gg K$ است به طور تصادفی از تصاویر آموزشی استخراج می‌شود. سپس،

© Copyright 2016 ISBME, http://www.ijbme.org

اطلاعات متقابل بین هر ویژگی و مجموعه‌ی برچسب کلاس-های C محاسبه می‌شود. درنهایت، K ویژگی با بالاترین اطلاعات متقابل با مجموعه‌ی برچسب کلاس‌ها انتخاب می-شود و K تکّه متناظر با این ویژگی‌ها در لایه‌ی S2 مدل HMAX ذخیره می‌شود.

مدل پیشنهادی، روی پایگاه داده هواپیما و برگ از مجموعه‌ی داده CalTech5 و پایگاه داده پیانو و دارسان از مجموعه‌ی داده CalTech101 مورد ارزیابی قرار گرفت و عملکرد آن با مدل اصلی HMAX مقایسه شد. نتایج شبیه‌سازی‌ها نشان داد که عملکرد مدل پیشنهادی هم در تعداد ویژگی‌ها و هم در تعداد نمونه‌های آموزشی مختلف از عملکرد مدل اصلی HMAX بهتر است. هم‌چنین نتایج شبیه‌سازی‌ها نشان داد که هرچه تعداد نمونه‌های آموزشی یا تعداد ویژگی‌ها افزایش می‌یابد، عملکرد مدل پیشنهادی و مدل اصلی HMAX به هم نزدیک‌تر می‌شود؛ چراکه تعداد تکّه‌های زائد و نامرتبط در مجموعه‌ی تکّه‌ها کم‌تر می‌شود.

۵- مرجع‌ها

[1] D. H. Hubel T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex" *The Journal of physiology* Vol 160, pp 106–154, 1962.

[2] D. H. Hubel, T. N Wiesel, "Receptive fields and functional architecture of monkey striate cortex" *The Journal of physiology* Vol 195, pp 215-243, 1968.

[3] M. Ghodrati, S. M. Khaligh Razavi, R. Ebrahimpour, K. Rajaei, M. Pooyan, "How Can Selection of Biologically Inspired Features Improve the Performance of a Robust Object Recognition Model?" *PLOS ONE* Vol 7, No 2, pp 1-12, 2012.

[4] K. Fukushima, "Neocognitron: A Self-Organizing Neural Network Model for A Mechanism of Pattern Recognition Unaffected by Shift in Position" *Biological Cybernetics* Vol 36, No 4, pp 193–202, 1980.

[5] G. Wallis G, E. T. Rolls, "A Model of Invariant Object Recognition in the Visual System" *Progress in Neurobiology* Vol 51, pp 167-194, 1996.

[6] S. Grossberg, E. Mingolla, W. D. Ross, "Visual Brain and Visual Perception: How Does the Cortex Do Perceptual Grouping?" *Trends in Neurosciences* Vol 20, pp 106-111, 1997.

[7] S. Grossberg, "Towards A Unified Theory of Neocortex: Laminar Cortical Circuits for Vision and Cognition" *Progress in Brain Research* Vol 165, pp 79-104, 2007.

[8] M. Riesenhuber, T. Poggio, "Hierarchical Models of Object Recognition in Cortex" *Nature Neuroscience* Vol 2, No 11, pp 1019–1025, 1999.

[9] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, T. Poggio, "Robust Object Recognition with Cortex-Like Mechanisms" *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol 29, No 3, pp 411–426, 2007.

[10] K. Huang, T. Dacheng, Y. Yuan, L. Xuelong, T. Tieniu, "Biologically Inspired Features for Scene Classification in Video Surveillance" *IEEE Transactions on Systems Man and Cybernetics-Part B: Cybernetics* Vol 41, No 1, pp 307-313, 2011.

[11] M. Hamidi, A. Borji, "Invariance Analysis of Modified C2 Features: Case Study-Handwritten Digit Recognition," *Machine Vision and Applications* Vol 21, No 6, pp 969-979, 2010.

[12] C. Theriault, N. Thome, "Extended Coding and Pooling in the HMAX Model" *IEEE Transaction on Image Processing* Vol 22, No 2, 764-777, 2013.

[13] E. Krupka, A. Navot, N. Tishby, "Learning to Select Features using their Properties" *Journal of Machine Learning Research* Vol 9, No 10, pp 2349-2376, 2008.

[14] G. Brown, A. Pocock, M. J. Zhao, M. Lujan, "Conditional Likelihood Maximisation: A Unifying Framework for Information Theoretic Feature Selection" *Journal of Machine Learning Research* Vol 13, No 1, 2012.

© Copyright 2016 ISBME, http://www.ijbme.org