بسم الله الرحمن الرحيم

دانشگاه پیام نور

دانشکده ریاضی

مرکز تهران

پایان نامه

برای دریافت مدرک کارشناسی ارشد

رشته ریاضی کاربردی گرایش تحقیق در عملیات

گروه ریاضی

تشخیص انسان در تصاویر کامپیوتری

با استفاده از دسته بندی روی خمینه های ریمانی

نرگس یوسفی

استاد راهنما: دکتر علاءالدین ملک

استاد مشاور: دکتر بیژن زاده

اردیبهشت ۹۲

تاریخ: 92/4/1

شماره: .....

بسمه تعالی

صورتجلسه دفاع از پایان نامه دوره کارشناسی ارشد

جلسه دفاع از پایان نامه دوره کارشناسی ارشد

خانم **نرگس یوسفی**

دانشجوی رشته **ریاضی کاربردی** به شماره دانشجویی ۸۷۰۰۰۲۶۳٤

تحت عنوان «تشخیص انسان در تصاویر کامپیوتری با استفاده از دسته بندی روی خمینه های ریمانی» با حضور هیات داوران در روز شنبه مورخ 1/ 92./4 ساعت ۱۱-۱۲ در محل ساختمان دانشگاه پیام نور واحد تهران شرق برگزار شد و هیات داوران پس از بررسی، پایان نامه مذکور را شایسته نمره به عدد ۱۸ به حروف هجده با درجه بسیار خوب تشخیص داد.

ردیف	نام ونام خانوادگی	هیات داوران	مرتبه دانشگاهی	دانشگاه/ موسسه
۱	دکتر علاالدین ملک	استاد راهنما	دانشیار	دانشگاه تربیت مدرس
۲	دکتر محمد حسن بیژن زاده	استاد مشاور	استاد	پیام نور
۳	دکتر محمد رضا پیغامی	استاد داور	دانشیار	دانشگاه صنعتی خواجه نصیرالدین طوسی
٤	دکتر خدیجه احمدی آملی	نماینده علمی گروه	استاد یار	پیام نور

گواهی اصالت، نشر و حقوق مادی و معنوی اثر

اینجانب نرگس یوسفی دانشجوی ورودی سال ۱۳۸۷ مقطع کارشناسی ارشد رشته ریاضی کاربردی گرایش تحقیق در عملیات گواهی می نمایم چنانچه در پایان نامه خود از فکر، ایده و نوشته دیگری بهره گرفته ام با نقل قول مستقیم یا غیر مستقیم منبع و ماخذ آن را نیز در جای مناسب ذکر کرده ام. بدیهی است مسئولیت تمامی مطالبی که نقل قول دیگران نباشد بر عهده خویش می دانم و جوابگوی آن خواهم بود.

نام و نام خانوادگی دانشجو

تاریخ و امضاء

اینجانب نرگس یوسفی دانشجوی ورودی سال ۱۳۸۷ مقطع کارشناسی ارشد رشته ریاضی کاربردی گرایش تحقیق در عملیات گواهی می نمایم چنانچه براساس مطالب پایان نامه خود اقدام به انتشار مقاله، کتاب، و ... نمایم ضمن مطلع نمودن استاد راهنما، با نظر ایشان نسبت به نشر مقاله، کتاب، و ... و به صورت مشترک و با ذکر نام استاد راهنما مبادرت نمایم.

نام و نام خانوادگی دانشجو

تاریخ و امضاء

کلیه حقوق مادی مترتب از نتایج مطالعات ، آزمایشات و نو آوری ناشی از تحقیق موضوع این پایان نامه متعلق به دانشگاه پیام نور می باشد.

تیر ۱۳۹۲

تقدیم به

پدرم

و مادر عزیزم

که همواره دعایش گشاینده راهم بود

تقدیم به

همسر مهربانم

که بی یاریش این نامه به پایان نمی رسید

از استاد عزیز جناب دکتر ملک و نیز دوستان گرامی آقای دکتر انصاری و خانم سمانه خدایاری که در اجرای برنامه ها مرا یاری رساندند و خانم خاتمیان فر که در یافتن منابع به من کمک کردند، بسیار سپاسگزارم.

نرگس یوسفی

خرداد۹۲

چکیده

مسئله‌ی تشخیص انسان[۱] که زیر شاخه ای از مسئله تشخیص شیء[۲] در تصاویر کامپیوتری است، به دلیل کاربرد آن در علم رباتیک، دستگاه های خودکار و خودگردان (مانند خودرو های اتوماتیک) و... در دو دهه اخیر مورد توجه مهندسین کامپیوتر و کارشناسان هوش مصنوعی و ریاضیدانان معاصر قرار گرفته است. در این پایان نامه الگوریتمی جدید برای تشخیص انسان ارائه می شود که از ماتریس های کوواریانس به عنوان توصیف گرهای شیء[۳] استفاده می کند. فضای این ماتریس‌های متقارن، معین مثبت و منفرد می‌تواند به عنوان یک خمینه ریمانی پیوسته بیان شود؛ در نتیجه عمده تلاش روش مورد استفاده در این پایان نامه دسته بندی نقاط (ماتریس‌های کوواریانس) روی این خمینه به دو دسته‌ی مثبت (انسان) و منفی (غیر انسان) به کمک هندسه ریمانی و یک روش قدرتمند دسته بندی به نام لجیت‌بوست می‌باشد؛ این روش در سال ۲۰۰۸ توسط پیتر میر[٤]، هانسل توزل[٥] و فاتح پوریکلی[٦]، ارائه شد. الگوریتم این روش، نه تنها برای انسان‌ها که برای تشخیص هر گروه دیگری از اشیاء می تواند به کار رود. در واقع روشی کلی برای تشخیص تمامی اشیاء در تصاویر دیجیتال از جمله ماشین‌ها، حیوانات و ... به حساب می‌آید. این الگوریتم در پایگاه داده انسانی اینریا[۷] آزمایش شده و برتری سرعت عملکرد آن نسبت به روش‌های پیشین مشاهده شده است [٥٦]. استخراج الگوریتم‌ها و اجرای برنامه تشخیص انسان برای نمونه ها و تصاویر مختص کشورمان ایران مطابق با پوشش‌های محلی و بومی آن در کنار نمونه‌های اروپایی، کاری است که در این پایان نامه انجام شده است.

واژگان کلیدی: تشخیص انسان، تشخیص شیء، ماشین بینایی، تشخیص الگو، هوش مصنوعی، یادگیری ماشین، پردازش تصویر، دسته بندی، لجیت بوست، آبشار پس زننده، توصیف گر کوواریانس، ماتریس معین مثبت، رگرسیون کمترین مربعات وزین، خمینه ریمانی.

---

[۱] Human detection
[۲] Object detection
[۳] Object descriptors
[٤] P. Meer
[٥] O. Tuzel
[٦] F. Porikli
[۷] INRIA

فهرست مطالب

مقدمه ۱

فصل اول: بررسی اجمالی پیشینه‌ی مطالعات و روش‌های به کار گرفته شده در مسئله‌ی تشخیص انسان ٤

۱-۱- استخراج خصیصه ٥

۱-۱-۱- نمایش های پراکنده ٥

۱-۱-۲- نمایش های چگال ٦

۱-۱-۲-۱- آنالیز مؤلفه اصلی ۷

۱-۱-۲-۲- موجک های هار ۷

۱-۱-۲-۳- میدان های پذیرای موضعی (LRFs) ۸

۱-۲- تشخیص شئ ۹

۱-۲-۱-شبکه های عصبی ۱۰

۱-۲-۲-ماشین‌های بردار پشتیبان ۱۱

۱-۲-۳-بوستینگ ۱۲

فصل دوم:تعاریف و تئوری‌های ریاضی مرتبط با مسئله تشخیص انسان ۱٦

۲-۱آشنایی با مباحث و تعاریف مرتبط با مسئله تشخیص انسان ۱٦

۲-۲-یادگیری ماشین و ارتباط آن با تشخیص انسان ۲۰

۲-۳-رگرسیون موزون ۲٦

۲-٤-توصیف‌گر های کوواریانس ۳۱

۲-٥-هندسه‌ی ریمانی ۳۷

۲-٥-۱- ژئودزی و نمودار نمایی ۳۸

۲-۵-۲- فضای ماتریس‌های معین مثبت متقارن و تعریف نما، لگاریتم و جذر در این فضا ٤٠
۲-۵-۳- یک فاصله ناوردای آفین ٤١
۲-۵-٤- نگاشت های نمایی و لگاریتمی ٤٦
۲-۵-۵- نمایش مینیمال روی فضای مماسی ٤٨
۲-٦- میانگین نقاط روی خمینه‌های ریمانی ٤٨
۲-۷- دسته بندی روی خمینه های ریمانی ۵۲
۲-۷-۱- نگاشت های موضعی و بوستینگ ۵۲
۲-۷-۲- لجیت بوست روی خمینه های ریمانی ۵٤
۲-۷-۳- الگوریتم لجیت بوست ۵۷
۲-۸- تشخیص انسان ٦٠

فصل سوم: الگوریتم های ریاضی مرتبط با مسئله‌ی تشخیص انسان ٦٨
۳-۱- الگوریتم آبشار تعلیمی ٦٩
پایان الگوریتم آبشار تعلیمی ۷۲
۳-۱- الف: الگوریتم لجیت بوست تعلیمی ۷۲
پایان الگوریتم لُجیت بوستِ تعلیمی ۷۸
۳-۱- ب: الگوریتم آزمون ۷۸
پایان الگوریتم آزمون ۸۰
۳-۲- الگوریتم تشخیص انسان‌های درون یک تصویر ۸۰
پایان الگوریتم تشخیص ۸۳
۳-۳-الگوریتم محاسبه ماتریس کوواریانس نرمال هر زیرپنجره ۸۳
۳-۳-۱- محاسبه ماتریس کوواریانس هر پنجره : ۸۳

۳ـ۳ـ۱ـ۱ـ الگوریتم ساختن تصویر ویژگی ۸ بعدی ۸۳
۳ـ۳ـ۱ـ۲ـ الگوریتم محاسبه ماتریس‌های تصویرانتگرالی $integral\ F_k$ و $integral\ F_rF_s$ ۸۵
۳ـ۳ـ۱ـ۳ـ الگوریتم محاسبه ماتریس کوواریانس پنجره مشخص با استفاده از تصویر ویژگی F کل تصویر ۸۶
۳ـ۳ـ۲ـ نرمال سازی ماتریس کوواریانس زیرپنجرهٔ C_r با استفاده از ماتریس کوواریانس پنجره‌ی اصلی C_R ۸۷
۳-٤- الگوریتم نگاشتن ماتریس کوواریانس x در فضای مماسی در نقطه m و تبدیل آن به یک بردار ۳۶ بعدی x ۸۸
۳ـ٤ـ۱ـ تبدیل UDU^T ۸۸
۳ـ٤ـ۲ـ الگوریتم محاسبه لگاریتم یک ماتریس ۸۹
پایان الگوریتم ۳ـ٤ ..... ۹۰
۳-۵- الگوریتم محاسبه میانگین وزین ۹۰
۳-٦- الگوریتم برازش تابع خطی f با بردار ضرایب B با رگرسیون کمترین مربعات وزین از z_i ها به xp_i و xn_i ها ۹۱

فصل چهارم: آزمون و مقایسه ۹۳
٤-۱- آزمایشات روی مجموعه داده‌ی اینریا ۹۳
٤-۲-تشخیص نمونه ها ۹۹
٤-۳-پیچیدگی محاسبات ۱۰۰
٤-٤- مزایای روش ۱۰۰
فهرست منابع ۱۰۳

مقدمه

مسأله‌ی تشخیص انسان زیر شاخه‌ای از مسأله‌ی تشخیص شیء در تصاویر کامپیوتری است که در دو دهه‌ی اخیر مورد توجه مهندسین کامپیوتر، کارشناسان هوش مصنوعی و ریاضیدانان معاصر قرار گرفته است.

تشخیص گروه‌های مختلف اشیاء در تصویر کامپیوتری (دیجیتال) و یا در ویدئو و فیلم ها یکی از ابتدایی‌ترین و پایه‌ای‌ترین کارهایی است که در زمینه‌ی تحقیقات ماشین بینایی[۱]، تاکنون انجام شده است. موفقیت کارکرد بسیاری از کاربردهای آن مانند تجسس دیداری[۲]، بازیابی تصویر، علم رباتیک، دستگاه‌های خودکار و خودگردان (مانند خودرو های اتوماتیک) و حتی دوربین های کوچک، براساس دقت عملیات تشخیص بدست آمده است. نمونه‌ای از آن کاربرد تشخیص چهره در افزایش وضوح تصویر دوربین های دیجیتال در حالت خودکارِ این دوربین هاست که به صورت قرار دادن چهره‌ی تشخیص داده شده در داخل یک مربع هنگام گرفتن عکس، خود را نمایان می‌کند.

تشخیص انسان در تصاویر ثابت در میان نمونه مسائل تشخیص شیء[۳]، از مشکل ترین مسائل مورد توجه توجه تا به امروز بوده است. اسکلت مفصل دار و تصاویر متفاوت از حالت های مختلف بدن انسان و در آمیختن آن با سایه روشن ها و تغییرات حالت و ژست در پیچیدگی مسأله نقش دارند. در این پایان نامه در میان روش‌های مختلف تشخیص انسان در تصاویر ثابت کامپیوتری، الگوریتمی جدید مورد توجه قرار گرفته است که از ماتریس‌های کوواریانس به عنوان توصیف‌گرهای شیء[۴] استفاده می‌کند و فضای این ماتریس‌های متقارن و معین مثبت یک خمینه‌ی ریمانی خواهد بود؛ از آنجا که این توصیف‌گرها در یک فضای برداری قرار نمی‌گیرند، تکنیک‌های یادگیری ماشینِ[۵] شناخته‌شده‌ی معمول، برای آموزش دسته‌کننده‌ها[۶] مناسب نیستند. در نتیجه عمده تلاش روش مورد استفاده در این پایان نامه دسته‌بندی نقاط

---

[۱] Computer Vision
[۲] Visual Surveillance
[۳] Object Detection
[۴] Object descriptors
[۵] Machine Learning
[۶] Classifiers

(ماتریس‌های کوواریانس) روی این خمینه به دو دسته‌ی مثبت (انسان) و منفی (غیر انسان) با روش جدیدی به کمک هندسه‌ی ریمانی و یک روش قدرتمند دسته‌بندی به نام لجیت بوست می‌باشد[۵]. در الگوریتم این روش، ماشین، توصیف‌گرهای انسانی را خود می‌یابد و سپس با مقایسه‌ی آن ها با نمونه‌های منفی، به اصطلاح خود را تعلیم می‌دهد، اما مسأله‌ی مهم تر آن است که این عملیات را نه تنها برای انسان‌ها که برای هر گروه دیگری از اشیاء می‌تواند انجام دهد و در واقع این روش تنها مخصوص مسأله‌ی تشخیص انسان نیست و روشی کلی برای تشخیص تمامی اشیاء در تصاویر دیجیتال از جمله ماشین ها، حیوانات و ... به حساب می‌آید.

هدف ما در این پایان نامه ارائه‌ی تشخیص دهنده‌ای است که در داخل کشور نیز مورد استفاده قرار گیرد و در نتیجه حجم زیادی از داده‌های تعلیم خود را از نمونه‌های ملی و بومی انتخاب کردیم که این‌کار برای اولین بار در کشور ایران انجام می‌شود، امید است که این روند ادامه یافته و ارتقاء یابد تا در آینده‌ی نه چندان دور شاهد استفاده تشخیص دهنده‌های شیء مختلف، قدرتمند و پرسرعت در تصاویر دیجیتال در کشورمان باشیم و از کاربردهای آن بهره ببریم.

این پایان نامه به این صورت سازمان یافته است:

در فصل اول، معرفی مختصری از روش‌های مختلف تشخیص انسان در تصاویر دیجیتال، که تا کنون ارائه شده اند، داریم.

در فصل۲، ابتدا با روش ذخیره‌ی یک تصویر کامپیوتری (دیجیتال) آشنا می‌شویم و سپس به مبحث یادگیری ماشین و استفاده از اصول آن از جمله دسته‌بندی و تعلیم دسته‌کننده‌ها، در تشخیص انسان در تصویر دیجیتال می‌پردازیم. در ادامه تمامی مباحث، تعاریف و تئوری‌های ریاضی مرتبط با تشخیص انسان در روش خود از جمله رگرسیون لجستیک جمعی، توصیف‌گرهای کوواریانس، هندسه‌ی ریمانی و فضای ماتریس های معین مثبت متقارن، نگاشت های نمایی و لگاریتمی در این فضا، میانگین نقاط روی خمینه‌ی ریمانی و الگوریتم روش لجیت بوست را شرح داده و در انتها چگونگی عملکرد آبشار

تشخیص‌دهنده‌ی انسان را، با به‌کارگیری الگوریتم لجیت بوست برای تعلیم دسته‌کننده‌های لجیت بوست که نقاط روی خمینه‌ی ریمانی را دسته‌بندی می‌کنند، توضیح می‌دهیم.

در فصل۳ الگوریتم های تمامی مراحل روش تشخیص انسان خود را با جزئیات کامل از نمونه گیری نمونه‌های تعلیم تا تعلیم دسته‌کننده‌ها و سپس جستجو در تصویر ورودی برای یافتن انسان ها، آورده ایم.

در فصل ٤ نحوه‌ی عملکرد روش خود را در تشخیص انسان محک خواهیم زد.

خدا را شاکرم که مرا در فهم تئوری‌ها و استخراج الگوریتم‌های مشکل مرتبط با این پایان نامه و برنامه‌نویسی آن یاری داد. پاسخگوی سوالات احتمالی خوانندگان به آدرس اینترنتی yousefi.narges@yahoo.com هستم.

و من الله التوفیق

نرگس یوسفی

بهار ۹۲

فصل اول

بررسی اجمالی پیشینه‌ی مطالعات و روش‌های به کار گرفته شده در مسأله‌ی تشخیص انسان

همان طور که در مقدمه گفته شد، مسأله‌ی تشخیص انسان که در حوزه تشخیص شیء قرار دارد، در دو دهه‌ی اخیر مورد توجه مهندسین کامپیوتر، کارشناسان هوش مصنوعی و ریاضیدانان معاصر قرار گرفته است. ماشین یا ربات از طریق یک برنامه‌ی نرم افزاری باید بتواند انسانهای درون یک تصویر را تشخیص دهد و این یعنی هوشمند سازی ماشین حسابگر غول‌پیکری به نام کامپیوتر. پس وارد حوزه‌ی هوش مصنوعی و یادگیری ماشین[۱] می‌شویم. قاعدتاً باید ابتدا به ماشین، توصیفی از خصوصیات شیء مورد نظر تعلیم داده شود و سپس از او انتظار تشخیص داشت. به طور کلی دو مرحله‌ی پردازش مهم و اصلی در الگوریتم تشخیص شیء از هر نوعی، می‌توان شناخت: اول «استخراج خصیصه[۲]» که در آن بسیاری از توصیف‌گرهای شیء حاوی اطلاعات مفید در خصوص پروسه‌ی تشخیص، از متن تصویر بدست می‌آیند. دومین مرحله، مرحله‌ی «تشخیص» است که در آن توصیف‌گرهای شیء بدست آمده در

[۱] Machine Learning
[۲] Feature Extraction

یک چارچوب دسته‌بندی[1] برای تشخیص شیء مورد نظر، (به وسیله‌ی تعلیم دسته‌کننده‌ها) مورد استفاده قرار می‌گیرند.

۱-۲-استخراج خصیصه

روش‌های استخراج خصیصه بیشتر به دو دسته تقسیم می‌شوند:

۱- نمایش های پراکنده[2] ۲- نمایش های چگال[3]

۱-۱-۱-نمایش های پراکنده

دراین روش‌ها مجموعه‌ای از ناحیه‌های موضعیِ نماینده، به عنوان برآیند یک الگوریتم تشخیص نقطه‌ی مورد نظر[4]، بدست می‌آیند. نقاط مورد نظرِ قابل اتکا، باید به صورت اطلاعات عددی فشرده و مفید در مورد محتوای آن تصویر موضعی، درآیند و نسبت به تغییراتی مانند نقطه‌ی دید، انتقال‌ها و یا روشنایی (سایه روشن) در خطوط کناره‌ی نمای شیء پایدار باشند. به این نقاط، اصطلاحاً نقاط کلیدی نیز می‌گویند. تشخیص‌دهنده‌های پایانی براساس بردارهای ویژگیِ محاسبه شده از توصیف‌گرهای این نقاط کلیدی تعلیم یافته‌اند. ازاین‌رو عملکرد تشخیص‌دهنده‌ی کل به این بستگی خواهد داشت که کدام یک از این نقاط کلیدیِ یافت شده می‌تواند از نظر قابل اطمینان بودن و تکرارپذیر بودن برای کلاس شیء داده شده قابل اتکا باشد. در زمینه‌ی تشخیص‌دهنده‌های نقطه‌ای نشریات فراوانی وجود دارد و مجموعه‌ی مقالات [٢٤]، [٢٨]، [٣١]، [٣٦] و[٣٨] تنها از تعداد کمی از کاربردهای بسیار معمول روش‌هایی که روی گستره‌ی وسیعی از شرایط عملیاتی پایدار می‌مانند، استفاده می‌کنند.

روش‌های اولیه‌ی توصیف‌گرهای پاره‌ای[5]، از خصیصه‌های (ویژگی‌های) شدت- محور[6] استفاده می‌کردند که در این روش‌ها ویژگی‌ها براساس نقاط یا ناحیه‌های موضعیِ برجسته بودند؛ مانند کار هریس و

---

[1] Classification
[2] Spare representations
[3] Dense representations
[4] Interest Point Detection
[5] Part descriptors
[6] Intensity-based features

استفانز[1] در سال ۱۹۸۸ ومیکلاوسکی و اشمیت[2] در سال ۲۰۰۲ و نیز نمایش های هیستوگرام- محور از گرادیان ها و لبه‌های تصویر در بافت سه بعدی (فضایی)، مانند تبدیل ویژگی ثابت- مقیاس[3] (SIFT) که توسط لویی در سال‌های ۱۹۹۹ و۲۰۰٤ [۳٦] ارائه شد.

در سال‌های ۲۰۰۰ تا ۲۰۰۵ چندین الگوریتم تشخیص شیء با روش های متمایز کننده [۱٤] و [٤۲] و یا با انطباق شکل ها و اجزاء در کنارهم [۳۵] و[۱۷]، ارائه شدند.

۱-۱-۲-نمایش های چگال

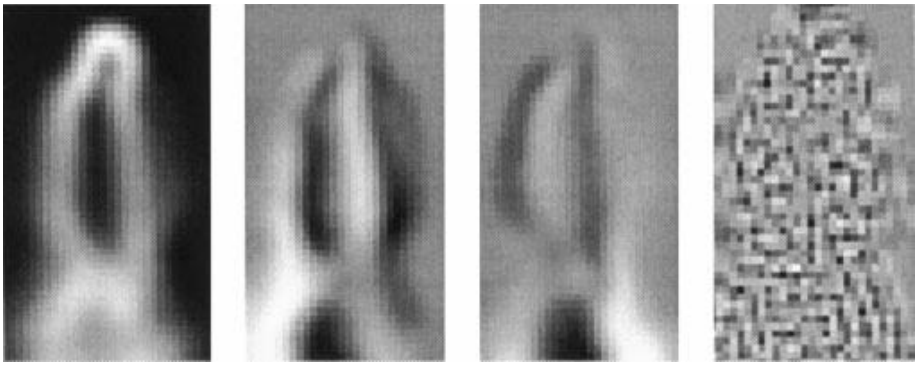
گروه دومِ روش های استخراج خصیصه، «نمایش های چگال» است، که توصیف‌گرهای شیء از داخل یک پنجره‌ی مشخص بدست می‌آیند. تصویر ورودی به صورت پرتراکم (اگر ممکن باشد پیکسل به پیکسل) به دنبال پنجره‌ها جستجو می‌شود و یک دسته‌کننده‌ی تعلیم یافته[4] از مدل شئ، مقدار دهی می‌شود. در مطالعات اولیه در زمینه‌ی تشخیص الگو[5] از شدت های تصویر مانند قالب[6] های تصویر [۵۲] [۵۲] و یا آنالیز مؤلفه‌ی اصلی (PCA)[7] [۵۵] و[۵۱]، برای نمایش مدل شئ استفاده می‌شد. اخیراً، توصیف‌گرهای موجک‌هار- محور[8] که مجموعه‌ای هستند از توابع پایه که تفاضل های شدت های بین دو ناحیه را کدگذاری می‌کنند، بسیار معروف شدند؛ به علت محاسبات مؤثر و برتری که در کدگذاری الگوهای بصری داشتند. اما روش استخراج خصیصه‌ی دیگری که ویژگی های انطباقی[9] را استخراج می‌کند، میدان های پذیرای موضعی[10] است. در اینجا توصیفی از سه روش مهم و کارآمد در میان دیگر روش‌ها یعنی آنالیز مؤلفه‌ی اصلی، موجک های هار و میدان های پذیرای موضعی را ارائه می‌دهیم.

---

[1] M. Stephens
[2] C. Schmid
[3] Scale-invariant feature transform
[4] Learned classifier
[5] Pattern Recognition
[6] Mask
[7] Principal Component Analysis
[8] Haar wavelete-based
[9] Adaptive features
[10] Local Receptive Fields (LRFs)

۱–۱–۲–۱– آنالیز مؤلفه‌ی اصلی

یکی از بهترین روش های استخراج خصیصه‌ی (خطی) شناخته شده، روش آنالیز مؤلفه‌ی اصلی (PCA) می‌باشد که توسط جین[1]، دووین[2] و مائو[3] در سال ۲۰۰۰ ارائه گردید. این روش آماری به طور مؤثری ابعاد را با استفاده از تشخیص ویژگی های بسیار گویا یعنی بردارهای ویژه با بزرگ ترین مقادیر ویژه، کاهش می‌دهد، درحالی‌که بردارهایی که مقادیر ویژه‌ی کوچکی دارند فرض می‌شود که شامل نویز (پارازیت) هستند و نتیجتاً حذف می‌شوند. در واقع این روش شباهت ها و تفاوت های داده‌ها را برجسته‌تر می‌کند. روش آنالیز مؤلفه‌ی اصلی می‌تواند از نظر ویژگی‌های کلی مورد توجه قرار گیرد؛ چرا که هر مؤلفه یک خصوصیت مهم و اصلی از کل الگوی ورودی را توصیف می‌کند؛ درحالی که جزئیات موضعی با کاهش بعد هموارتر می‌گردند (شکل۱–۱). تعداد مولفه‌های اصلی باقیمانده نوعاً توسط کاربر تعیین می‌شود.



شکل ۱–۱: مثالی از مولفه‌های اصلی بدست آمده روی مجموعه‌ی تعلیم، که به ترتیب نزولیِ مقادیر ویژه‌ی متناظر آن‌ها مرتب شده اند.

۱–۱–۲–۲- موجک های هار

[1] Jain
[2] R. Duin
[3] J. Mao

از معروف ترین ویژگی ها برای دسته‌بندی انسان که در مقالات یافت می‌شود، موجک های هار یا تعمیم‌های مربوط به آن [٤٤] و [٥٩] هستند. انگیزه‌ی استفاده از آن‌ها به این حقیقت برمی‌گردد که آن‌ها ویژگی های موضعی تصویر و به عبارت دیگر تفاضل‌های شدت را در چندین مقیاس مختلف، کدگذاری می‌کنند و در نتیجه موجب برقراری تعادل بین گویایی و تراکم می‌شوند. این روش یک بردار از داده‌ها مثل $\mathbf{x}$ را به یک بردار $\mathbf{x}'$ از ضرایب موجک تبدیل می‌کند. داده‌ی تبدیل شده می‌تواند «هرس» شود به این معنی که موجک های ظریف مقیاس فرض می‌شود که نویزها را نمایش می‌دهند و از این رو دور ریخته می‌شوند، به همین ترتیب بسیاری موجک های درشت مقیاس که به بزرگی شئ هستند از آنجا که اطلاعات مفیدی را کد نمی‌کنند، نیز دور ریخته می‌شوند (صفر در نظر گرفته می‌شود). سپس با مجموعه‌ی ضرایب باقیمانده می‌توان تخمینی از داده‌های اصلی را با استفاده از معکوس تبدیل به موجک‌های هار بدست آورد. در سال‌های ۱۹۹۸ تا ۲۰۰۱ *پاپاژورژیو و پاژیو و موهان* یک موجک هار مبتنی بر سیستم تشخیص انسان ارائه کردند.

آن‌ها جواب های تصحیح شده از سه موجک هار فوق کامل مرتبه‌ی اول یا همپوشانِ $\begin{bmatrix} ۱ & -۱ \\ ۱ & -۱ \end{bmatrix}$ و $\begin{bmatrix} ۱ & -۱ \\ -۱ & ۱ \end{bmatrix}$ و $\begin{bmatrix} ۱ & ۱ \\ -۱ & -۱ \end{bmatrix}$ را به عنوان یک مجموعه‌ی ویژگی استفاده کردند. بعدها مطالعات بیشتر نشان داد که ویژگی های کامل، جواب‌های تصحیح شده روی دو مقیاس ۹×۹ و ۱۲×۱۲ (ابعاد ماتریس موجک) هستند.

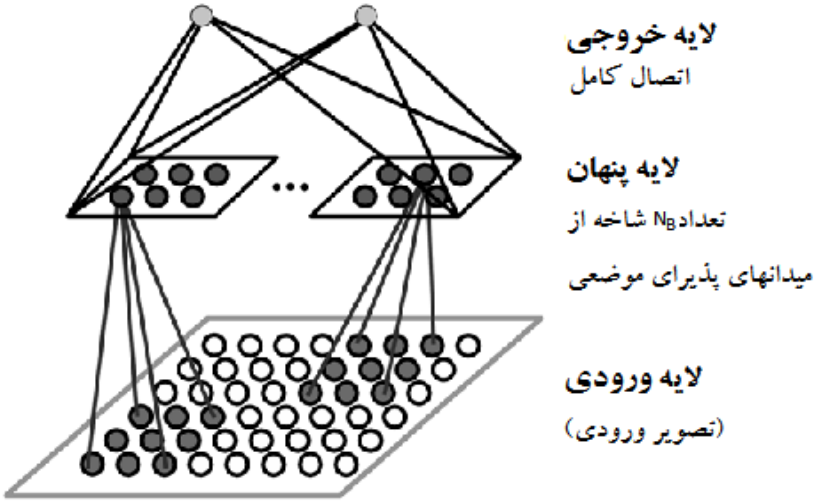
۱-۱-۲-۳- میدان های پذیرای موضعی (LRFs)

به جای انتقال شناور و دستی یک مجموعه از ویژگی ها، ادراک کننده‌های چند لایه[۱] یک روش انطباقی برای استخراج خصیصه به وسیله‌ی میانگین لایه‌های پنهان خود می‌سازند که در آن، ویژگی ها با داده در طول عملیات تعلیم، وفق داده می‌شوند. شبکه‌های عصبی پیش خور با میدان های پذیرای موضعی

[۱] Multilayer perceptrons

(NN/LRF) که توسط فوکوشیما[1] و دیگران [۲۶] ساخته شدند و سپس برای دسته‌بندی انسان بکار گرفته شدند، یک روش مخصوص جالب برای دسته‌بندی تصاویر دوبعدی هستند. در بخش ۱-۲-۱ توضیحی اجمالی از شبکه‌های عصبی ارائه شده است.

در مقایسه با ادراک کننده‌های چندلایه‌ی استاندارد، نورون ها در لایه‌ی پنهانی تنها به یک ناحیه‌ی موضعی محدود از تصویر ورودی معطوف به میدان های پذیرای موضعی شان (شکل۱-۲)، متصل می‌شوند. منشاء ویژگی های LRF، به طور ذاتی با تعلیم یک شبکه‌ی عصبی در آمیخته است. لایه‌ی پنهانی به تعدادی انشعاب تقسیم می‌شود که تمام نورون های داخل یک انشعاب، وزن یکسانی دارند. هر انشعابی تعدادی ویژگی موضعی تصویر را کدگذاری می‌کند. اتصال موضعی و هم وزنی انشعاب ها به طور مؤثری تعداد وزن هایی که در طول مرحله‌ی تعلیم تعیین می‌شوند، را کم می‌کند و در نتیجه موجب می‌شود که مجموعه‌های تعلیم کوچک متناظر برای موارد بالا، فراهم شود.



شکل ۱-۲: نمایش طرح کلی ساختار میدان های پذیرای موضعی

۱-۲- تشخیص شئ

[1] Fukushima