



پردیس علوم  
دانشکده ریاضی، آمار و علوم کامپیوتر

# بازشناسی اشیا به وسیله استخراج ویژگی‌های حاوی اطلاعات مفید

نگارنده

محمدامین مظفری

استاد راهنما

دکتر محمد گنج تابش

پایان‌نامه برای دریافت درجه کارشناسی  
در رشته علوم کامپیوتر

تیر ماه ۱۳۹۷

## چکیده

در این پروژه، راهکاری را بازگو می‌کنیم که برای بازشناسی اشیا بهینه می‌باشد که فقط از ویژگی‌هایی حاوی اطلاعات مفید در این راستا استفاده می‌شود. در قسمت طبقه‌بندی این مقاله، از طبقه‌بندهای خطی و بیز ساده استفاده شده است. نتایج به دست آمده حاکی از برتری این روش نسبت به روش استخراج ویژگی‌ها به صورت عمومی، به طور مثال موجک، دارد. در بخشی از مقاله روشی معرفی می‌شود که ویژگی‌های استخراج شده از آن علاوه بر دارا بودن اطلاعات مفید، باید نسبت به دیگر ویژگی‌های استخراج شده وابستگی کمتری داشته باشند که در پی آن کارایی روش مذکور بسیار بهبود پیدا خواهد کرد و سرعت دستیابی به نتیجه مورد قبول افزایش خواهد یافت. علاوه بر بهینه بودن زمانی و حافظه‌ای این روش نسبت به روش‌های عمومی، از طبقه‌بند های بسیار ساده‌تر و قابل فهم‌تر می‌توان برای حل این مسئله استفاده کرد که خود باعث کاهش پیچیدگی راه حل و افزایش سرعت آن خواهد شد. در قسمت پایانی این مقاله، دو روش معرفی شده در مورد استخراج ویژگی‌های حاوی اطلاعات مفید و به علاوه روش‌های عمومی آن به تفصیل مورد مقایسه قرار می‌گیرند تا کارایی این نظریه به خوبی نشان داده شود.

# فهرست مطالب

۱	مروری بر بازشناسی اشیاء	۱
۲	۱.۱ مراحل پایهای	۱.۱
۲	۲.۱ بازشناسی اشیاء به وسیله هوش ماشینی	۲.۱
۳	۱.۲.۱ تاریخچه هوش مصنوعی	۱.۲.۱
۵	۳.۱ کاربردها	۳.۱
۷	۲ شبکه‌های عصبی بازشناسی اشیاء	۲
۹	۱.۲ مقدمه	۱.۲
۹	۱.۱.۲ طبقه‌بند	۱.۱.۲
۱۳	۲.۱.۲ استخراج ویژگی	۲.۱.۲
۱۴	۲.۲ بازشناسی اشیاء و طبقه‌بند خطی	۲.۲
	۱.۲.۲ استخراج ویژگی‌های حاوی اطلاعات مفید	۱.۲.۲
۱۵		
۱۶	۲.۲.۲ شباهت‌سنجی و یافتن آستانه	۲.۲.۲
۱۷	۳.۲.۲ جسجتجوی حریصانه	۳.۲.۲
۱۷	۴.۲.۲ طبقه‌بندی به وسیله‌ی جداساز خطی	۴.۲.۲
۱۸	۵.۲.۲ آزمایش‌ها	۵.۲.۲
۲۱	آ A : آنتروپی شانون	آ

ب B : آنتروپی شرطی

۲۳

# فصل ۱

## مروری بر بازشناسی اشیاء

بازشناسی اشیاء عبارت است از توانایی درک خواص فیزیکی شیء (مانند شکل، رنگ و بافت) و سپس اعمال معانی به آن شیء (مانند شناسایی یک شیء به عنوان سیب). این فرآیند شامل درک استفاده از آن شیء، داشتن تجربه قبلی با آن شیء و چگونگی ارتباط آن با سایر اشیاء می‌باشد. در موجودات زنده پردازش شناخت بصری<sup>۱</sup> به طور معمول در یک سلسله مراتب پایین به بالا مورد بررسی قرار گرفته‌است که در آن اطلاعات به صورت پیوسته با پیچیدگی‌های در حال افزایش پردازش می‌شوند؛ در جایی که پردازنده‌های سطح پایین قشر بالایی مغز<sup>۲</sup> (مانند قشر بصری اولیه<sup>۳</sup>) در ابتدای این سلسله مراتب پردازش قرار می‌گیرند و پردازنده‌های سطح بالا قشر بالایی مغز (مانند قشر ناحیه ای<sup>۴</sup>) در بالای انتهای آن، جایی که کار تشخیص آسان می‌شود. شناخت‌شدترین نظریه سلسله مراتبی متعلق به دیوید مار<sup>۵</sup> می‌باشد. لازم به ذکر است که مدل‌هایی مانند

---

1- Visual Recognition

2- Cortex

3- Primary Visual Cortex

4- Inferotemporal Cortex (IT)

5-David Marr

مدل پیشنهادشده توسط مش بار<sup>۶</sup> ارائه شده است که به نظر قابل قبولتر در بین اندیشمندان قرار گرفته است.

## ۱.۱ مراحل پایهای

بازشناسی اشیاء بر پایه شواهد موجود در حوزه علوم اعصاب و روانشناختی<sup>۷</sup> به چهار مرحله ذیل تقسیم می شود:

۱. پردازش اجزای اساسی اولیه مانند رنگ، عمق و شکل
۲. این اجزای اساسی سپس به گروههای مجزا بر اساس شباهت تقسیم بندی می شوند و اطلاعات به طور بصری ارائه می شوند.
۳. نمایش بصری با توصیفهای موجود ساختاری در مغز تطبیق داده می شوند
۴. ویژگیهای معنایی به نمایش بصری ارتباط داده می شوند و در نتیجه آن شیء مورد شناسایی قرار می گیرد

## ۲.۱ بازشناسی اشیاء به وسیله هوش ماشینی

در طول سالهای اخیر به دلیل اهمیت پیشرفت در حوزه بازشناسی اشیاء دانشمندان و متخصصان زیادی شروع به کار کردند. تا به آن جا که به یک شاخه از هوش مصنوعی<sup>۸</sup> تبدیل گشته است. هوش مصنوعی یا هوش ماشینی به هوشی که یک ماشین در شرایط مختلف از خود نشان می دهد، گفته می شود. به عبارت دیگر هوش مصنوعی به سیستمهایی گفته می شود که می توانند واکنشهایی مشابه رفتارهای هوشمند انسانی

<sup>6</sup>- Moshe Bar

<sup>7</sup>- Neuropsychological

<sup>8</sup>- Artificial Intelligence (AI)

از جمله درک شرایط پیچیده، شبیه‌سازی فرایندهای تفکری و شیوه‌های استدلالی انسانی و پاسخ موفق به آنها، یادگیری و توانایی کسب دانش و استدلال برای حل مسایل را داشته‌باشند.

## ۱.۲.۱ تاریخچه هوش مصنوعی

هوش مصنوعی توسط فلاسفه و ریاضی‌دانانی نظیر جرج بول که اقدام به ارائه قوانین و نظریه‌هایی در مورد منطق نمودند، مطرح شد. با اختراع رایانه‌های الکترونیکی در سال ۱۹۴۳، هوش مصنوعی، دانشمندان آن زمان را به چالشی بزرگ فراخواند. در این شرایط، چنین به نظر می‌رسید که این فناوری قادر به شبیه‌سازی رفتارهای هوشمندانه خواهد بود. با وجود مخالفت گروهی از متفکرین با هوش مصنوعی که با تردید به کارآمدی آن می‌نگریستند تنها پس از چهار دهه، شاهد تولد ماشین‌های شطرنج باز و دیگر سامانه‌های هوشمند در صنایع گوناگون شدیم. نام هوش مصنوعی در سال ۱۹۶۵ میلادی به عنوان یک دانش جدید ابداع گردید. البته فعالیت در این زمینه از سال ۱۹۶۰ میلادی شروع شد. بیشتر کارهای پژوهشی اولیه در هوش مصنوعی بر روی انجام ماشینی بازی‌ها و نیز اثبات قضیه‌های ریاضی با کمک رایانه‌ها بود. در آغاز چنین به نظر می‌آمد که رایانه‌ها قادر خواهند بود چنین فعالیت‌هایی را تنها با بهره گرفتن از تعداد بسیار زیادی کشف و جستجو برای مسیرهای حل مسئله و سپس انتخاب بهترین روش برای حل آنها به انجام رسانند. اصطلاح هوش مصنوعی برای اولین بار توسط جان مکارتی (که از آن به عنوان پدر علم و دانش تولید ماشین‌های هوشمند یاد می‌شود) استفاده شد. وی مخترع یکی از زبان‌های برنامه‌نویسی هوش مصنوعی به نام لیسپ<sup>۹</sup> است. با این عنوان می‌توان به هویت رفتارهای هوشمندانه یک ابزار مصنوعی پی برد. (ساخته

<sup>۹</sup>- lisp

دست بشر، غیرطبیعی، مصنوعی) حال آنکه هوش مصنوعی به عنوان یک اصطلاح عمومی پذیرفته شده که شامل محاسبات هوشمندانه و ترکیبی (مرکب از مواد مصنوعی) است. بازشناسی اشیاء امروزه بر عهده ماشین اعم از ربات، رایانه و وسایل مشابه انجام می پذیرد، چرا که انرژی و هزینه کمتری نسبت به انسانها صرف می کنند و امروزه حتی در بعضی موارد نتایج بهتری نیز نسبت به نمونه های انسانی به دست می دهند. از همین رو شرکتهای بزرگ مانند مایکروسافت، گوگل و آمازون به شدت در این حوزهها مشغول تلاش هستند. برای استفاده از بازشناسی اشیاء تا به الان تکنیکهای زیادی ارائه شده است که در فصل بعد به تفصیل در این مورد توضیح خواهیم داد.



## ۳.۱ کاربردها

بازشناسی اشیاء در طیف گسترده‌ای از حوزه‌ها کاربردهای زیادی دارد. از آنجایی که بازشناسی چهره یکی از شاخه‌های مهم در این حوزه است، مهمترین کاربرد آن را می‌توان به زمینه‌های قضایی و امنیتی کشورها اختصاص داد به این صورت که متهم‌های قضایی یا موردهای مشکوک امنیتی را به راحتی مورد بازشناسی قرار داد. یکی دیگر از کاربردهای مهم بازشناسی اشیاء می‌توان به امور حمل و نقل به ویژه کنترل ترافیک در خیابان‌ها و جاده‌ها اشاره کرد. که با شناسایی خودرو یا پلاک آن تصمیمات مورفنز را در جهت بهبود وضع خیابان‌ها اتخاذ نمود. یکی از مدرن‌ترین کاربردهای بازشناسی اشیاء که هنوز در کشور ما مهیا استفاده نشده‌است، استفاده از آن در فروشگاه‌های آنلاین است. به این صورت که خریدار تصویر جنس موردنظرش را بارگذاری می‌کند و فروشنده در صورت شناسایی و موجود بودن کالا آن را به خریدار عرضه می‌کند. همانطور که روشن است توسعه این حوزه می‌تواند بسیار پرکاربرد و مفید برای زندگی روزمره انسان واقع شود و طبق شواهد موجود، پیشرفت در این حوزه سرعت گرفته و حتما همه ما تاثیرش را بر زندگی‌های خود حس خواهیم نمود.

در ادامه گزارش به شرح شبکه های عصبی بازشناسی اشیاء و شرح دو روش بهینه برای انجام این کار و توضیح در مورد چگونگی این انجام این فرآیند در شبکه های عصبی عمیق<sup>۱۰</sup> می پردازیم و در قسمت پایانی این دو روش را مورد مقایسه قرار خواهیم داد.

---

<sup>10</sup>- Deep Neural Network

## فصل ۲

# شبکه‌های عصبی بازشناسی اشیاء

دهه‌ها است دانشمندان علوم اعصاب میکوشند شبکه‌های رایانه‌ای طراحی کنند که بتوانند مهارت‌های دیداری، همچون بازشناسی اشیاء را، که مغز انسان خیلی دقیق و سریع انجام می‌دهد، تقلید کنند. تاکنون هیچ نوع مدل رایانه‌ای نتوانسته‌است با مغز نخستیان<sup>۱</sup> در بازشناسی بی‌درنگ اشیاء برابری کند. اما دانشمندان علوم اعصاب دانشگاه ام‌آی‌تی (MIT) در تحقیقات خود دریافته‌اند که آخرین نسل به اصطلاح "شبکه‌های عصبی ژرف" با مغز نخستیان تطبیق می‌کنند. جیمز دیکارلو<sup>۲</sup>، استاد علوم اعصاب و نویسنده اصلی مقاله که در شماره ۱۸ مجله *Biology Computational PLoS* به چاپ رسیده‌است می‌گوید: «از آنجا که این شبکه‌های عصبی بر اساس یافته‌های اخیر مربوط به چگونگی عملکرد مغز در بازشناسی اشیاء طراحی شده‌است، موفقیت آنها بر این نکته دلالت می‌کند که دانشمندان فهم نسبتاً دقیقی از عملکرد بازشناسی مغز به دست

<sup>۱</sup>- primate

<sup>۲</sup>-James DiCarlo

آوردند. این واقعیت که مدل‌ها پاسخ‌های عصبی و فاصله اشیا را در فضای جمعیت نورونی پیش‌بینی می‌کنند نشانگر آن است که مدل‌های مزبور تبیین خوبی از آنچه در این بخش اسرارآمیز مغز اتفاق می‌افتد انجام می‌دهند». چنین درک اصلاح شده‌ی از کارکرد مغز نخستیان می‌تواند به هوش مصنوعی و شاید به راه‌های جدید بهبود نقص عملکرد بینایی کمک کند. در دهه هفتاد میلادی دانشمندان با الهام گرفتن از مغز به امید تقلید توانایی پردازش اطلاعات دیداری، بازشناسی کلامی و فهم زبان شروع به ساخت شبکه‌های عصبی نمودند. در پیشنهاد شبکه‌های عصبی مبتنی بر ادراک بینایی، دانشمندان از بازنمایی سلسله مراتبی اطلاعات دیداری در مغز الهام گرفتند. درونداهای دیداری از شبکه به قشر اولیه دیداری مخ و از آنجا به قشر اطلاعاتی مخ ۳ می‌رود و در هر لایه از پردازش، پله پله مشخص‌تر می‌شود که شیء قابل بازشناسی چیست. طراحان شبکه عصبی برای تقلید از عملکرد مغز، چندین لایه محاسباتی در مدل‌های جدید ایجاد کردند. در هر لایه بازنمایی‌های دیداری شیء پیچیده‌تر و پیچیده‌تر می‌شوند و اطلاعات غیرلازم مانند محل شیء یا حرکت شیء کنار گذاشته می‌شوند. در این پژوهش محققان ابتدا توانایی بازشناسایی شیء را اندازه گرفتند. آنها با قرار دادن یک دسته الکتروود در قشر IT و همچنین در منطقه  $V4$  ۴ توانستند بازنمایی عصبی هر شیء را که حیوان می‌بیند مشاهده کنند. سپس آنها را با بازنمایی‌های «شبکه‌های عصبی ژرف» که شامل ماتریس عددی حاصل از عناصر محاسباتی در سیستم بود، مقایسه کردند. هر تصویر مجموعه متفاوتی از اعداد تولید می‌کند. دقت مدل وابسته به این است که مدل بتواند در جریان بازنمایی، اشیا مشابه را در دسته‌های مشابه گروه‌بندی کند.

<sup>3</sup>- IT

<sup>۴</sup>- بخشی از سیستم بینایی است که به قشر IT مرتبط است

## ۱.۲ مقدمه

همانطور که گفته شد شبکه‌های عصبی باز شناسی اشیاء از ساختار سلسه مراتبی استفاده می‌کنند. از بارزترین این شبکه‌ها ، شبکه‌های عصبی پیچشی<sup>۵</sup> می‌باشد، که توضیح مفصل این نوع شبکه‌ها در آدرس وبسایت شماره ۵ بخش کتاب نامه قابل مطالعه می‌باشد. یکی از مهمترین لایه‌های شبکه‌های عمیق لایه استخراج ویژگی می‌باشد. بدین صورت که این لایه‌ها از ورودی شبکه عصبی شروع به استخراج ویژگی‌ها می‌کنند. می‌دانیم مهمترین بخش شبکه‌های عصبی بخش یادگیری آنهاست و این یادگیری با همین ویژگی‌های استخراج شده از لایه‌های شبکه صورت می‌گیرد. علاوه بر یادگیری شبکه ، کاربرد بسیار مهم دیگری که استخراج ویژگی‌ها دارند، در قسمت انتهایی شبکه‌ها و زمان دادن خروجی شبکه است. بدین صورت که خروجی شبکه با طبقه بندی‌های تعبیه شده در شبکه مورد تصمیم‌گیری واقع می‌شود.

### ۱.۱.۲ طبقه‌بند

در زمینه یادگیری ماشین ، هدف از طبقه‌بندی آماری مشخص کردن تعلقش به کلاس (یا گروه) موردنظر با توجه به ویژگی‌های شیء می‌باشد. تابع‌هایی که به منظور طبقه‌بندی بر روی بردارها اعمال می‌شوند را طبقه‌بند می‌نامیم که در ادامه با دو مورد از آن‌هایی که در این مقاله استفاده شده‌است آشنا می‌شویم.

<sup>5</sup>- Convolutional Neural Network

<sup>6</sup>- Classifier

## طبقه‌بند خطی

۷ یک طبقه‌بند خطی، یک طبقه‌بندی با استفاده از ترکیب خطی ویژگی‌های ورودی است که ورودی‌های مسئله را به دو گروه تقسیم می‌کند. یکی از مهمترین ابزار طبقه‌بندهای خطی ماشین بردار پشتیبان<sup>۸</sup> می‌باشد.

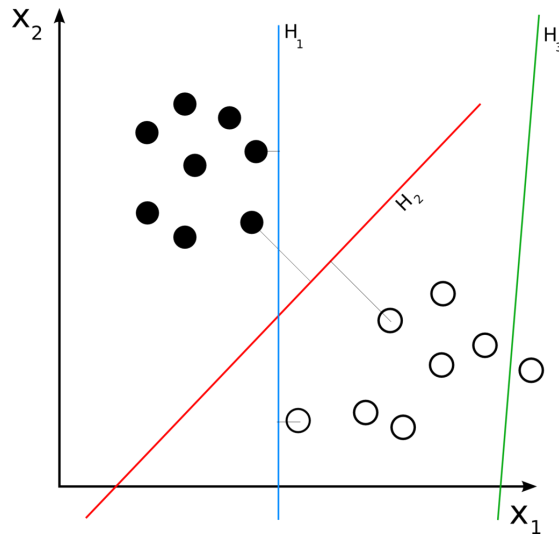
● SVM: این روش از جمله روش‌های نسبتاً جدیدی است که در سال‌های اخیر کارایی خوبی نسبت به روش‌های قدیمی‌تر برای طبقه‌بندی از جمله شبکه‌های عصبی پرسپترون نشان داده‌است. مبنای کاری دسته‌بندی کننده SVM دسته‌بندی خطی داده‌ها است و در تقسیم خطی داده‌ها سعی می‌کنیم خطی را انتخاب کنیم که حاشیه اطمینان بیشتری داشته‌باشد. حل معادله پیدا کردن خط بهینه برای داده‌ها به وسیله روش‌های QP که روش‌های شناخته‌شده‌ای در حل مسائل محدودیت‌دار هستند صورت می‌گیرد. قبل از تقسیم خطی برای اینکه ماشین بتواند داده‌های با پیچیدگی بالا را دسته‌بندی کند داده‌ها را به وسیله تابع phi به فضای با ابعاد خیلی بالاتر می‌بریم. اگر بردار ویژگی‌های ورودی که قرار است طبقه‌بندی شوند را  $\vec{x}$  در نظر بگیریم، خروجی طبقه‌بند خطی ما نیز به فرم زیر خواهد بود:

$$y = f(\vec{w} \cdot \vec{x}) = f\left(\sum_j w_j x_j\right)$$

<sup>7</sup>-Linear Classification

<sup>8</sup>- Support Vector Machine

که بردار وزنهای شبکه عصبی و بردار جداساز به منظور طبقه‌بندی ویژگی‌ها می‌باشد.



شکل ۱.۲: در این نمونه، نقطه‌های توخالی و توپر به راحتی می‌توانند توسط هر تعداد طبقه‌بند خطی به درستی طبقه‌بندی شوند. همانطور که می‌بینیم  $H_1$  و  $H_2$  به درستی این دو دسته را طبقه‌بندی کرده‌اند.  $H_2$  به علت انعطاف پذیری بیشتر در طبقه‌بندی مطلوب‌تر است. اما  $H_3$  نتوانسته این دو کلاس را از هم تفکیک کند پس مورد قبول نخواهد بود.

● بیز ساده<sup>۹</sup> : دستبندی‌کننده بیز ساده در یادگیری ماشین به گروهی از دستبندی‌کننده‌های ساده بر پایه احتمالات گفته می‌شود که با متغیرهای تصادفی مستقل ساده مفروض میان حالت‌های مختلف و براساس قضیه بیز کاربردی است. به‌طور ساده روش بیز روشی برای دستبندی پدیده‌ها، بر پایه احتمال وقوع یا عدم وقوع یک پدیده است. برای نمونه یک میوه ممکن است پرتقال باشد. اگر نارنجی و کروی با شعاع حدود ده سانتی‌متر باشد. اگر این احتمالات به درستی به همدیگر وابسته باشند نایو بیز در تشخیص اینکه این میوه پرتقال است یا نه بدرستی عمل خواهد کرد. برنامه‌های کاربردی بسیاری هستند که پارامترهای نایو بیز را تخمین می‌زنند، بنابراین افراد بدون سروکار داشتن با تئوری بیز می‌توانند از این امکان به منظور حل مسایل موردنظر بهره ببرند. با وجود مسایل طراحی و پیش فرض‌هایی که در خصوص روش بیز وجود دارد، این روش برای طبقه‌بندی کردن بیشتر مسایل در جهان واقعی، مناسب است. تابع این طبقه‌بند به شرح زیر می‌باشد:

$$\hat{y} = \operatorname{argmax}_{k \in \{1, \dots, K\}} p(C_k) \prod_{i=1}^n p(x_i | C_k)$$

که  $C_k$  کلاس‌های موجود در مسئله و  $x$  نیز بردارهای ورودی شبکه می‌باشند.

همانطور که گفتیم، یکی از مفهومی‌های مهم در بازشناسی اشیاء استخراج ویژگی می‌باشد. در ادامه توضیح مختصری برای آشنایی با این روش در مسائل هوش ماشینی می‌دهیم.

<sup>۹</sup>- Naive Bayes



## ۲.۱.۲ استخراج ویژگی

استخراج ویژگی<sup>۱۰</sup> فرایندی است که در آن با انجام عملیاتی بر روی داده‌ها، ویژگی‌های بارز و تعیین‌کننده آن مشخص می‌شود. هدف استخراج ویژگی این است که داده‌های خام به شکل قابل استفاده‌تری برای پردازش‌های آماری بعدی درآیند. بسیاری از اندیشمندان این حوزه معتقدند، استخراج ویژگی بهینه‌شده کلیدی برای ساخت مدلی باارزش است. روش‌های مختلف استخراج ویژگی بنا به فلسفه پشت سرشان ممکن است یک یا چند کار زیر را انجام دهند:

۱. حذف نویز داده‌ها

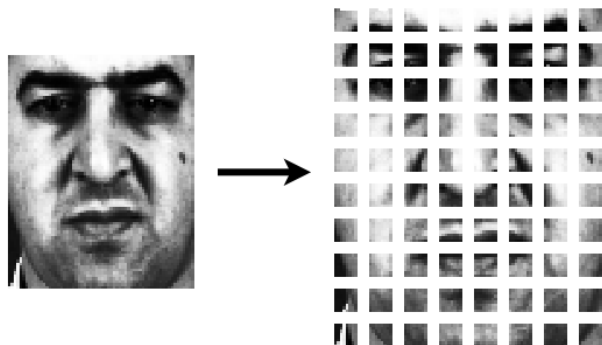
۲. جداسازی اجزای مستقل داده‌ها

۳. کاهش ابعاد داده برای تولید بازنمایی مختصرتر

۴. افزایش بعد برای تولید بازنمایی جدایی‌پذیرتر

انجام استخراج ویژگی فرایند بسیار متداولی در انواع مختلف پردازش داده‌ها چون پردازش تصویر، پردازش صوت و غیره است. در مقوله‌ی پردازش تصویر برای اینکه از روی الگوهای یک تصویر هویت یا خالق آن تصویر مشخص شود باید یک سری مشخصات عام یا خاص از دل تصویر بیرون کشیده شود که به این کار همان استخراج ویژگی گفته می‌شود. به عنوان مثال در تشخیص امضاء به وسیله پردازش تصویر یک سری ویژگی‌ها (مانند شیب خط‌ها) از تصویر اسکن شده امضاء بیرون کشیده می‌شود، که به وسیله آن می‌توان صاحب امضاء را تشخیص داد. در شکل زیر نمونه‌ای از استخراج ویژگی در پردازش تصویر نشان داده شده‌است.

<sup>10</sup>- Feature extraction



شکل ۲.۲: استخراج ویژگی از تصویر چهره [۷]

## ۲.۲ بازشناسی اشیاء و طبقه‌بند خطی

طرح‌های طبقه‌بندی بصری معمولاً در دو مرحله پیادسازی می‌شوند. اول ویژگی‌ها از عکس استخراج می‌شوند که شیء مورد طبقه‌بندی با استفاده از این ویژگی‌ها نمایش داده می‌شود. دوم، یک طبقه‌بند بر روی ویژگی‌های بدست‌آمده اعمال شده و تصمیم نهایی را لحاظ می‌کند. که ما در این بخش از طبقه‌بند خطی که بسیار کار را ساده می‌کند استفاده می‌کنیم. در بعضی موارد ویژگی‌های استخراج شده دارای ابعاد زیادی می‌باشند و کار با آن‌ها دشوار خواهد بود به عنوان مثال: زمانی که شدت<sup>۱۱</sup> تصویر به عنوان ویژگی‌های اساسی استخراج می‌شود؛ که در این شرایط با روش‌هایی مثلاً ماشین بردار پشتیبان ابعادها را کاهش داده و ویژگی‌ها قابل استفاده خواهند شد. نکته مهم قابل ذکر در این مبحث این است که اگر ویژگی‌های استخراج شده بسیار ساده و حاوی اطلاعات مفیدی نباشند، باعث پیچیدگی مسئله و در نتیجه ناکارآمدی در استفاده از طبقه‌بندهای خطی خواهند شد.

<sup>۱۱</sup>- intensity

## ۱.۲.۲ استخراج ویژگی‌های حاوی اطلاعات مفید

در این بخش فرآیند انتخاب ویژگی‌های حاوی اطلاعات مفید را شرح می‌دهیم. هدف انتخاب ویژگی‌های مختص یک کلاس است که بیشترین اطلاعات در مورد آن کلاس را انتقال می‌دهند. این نوع ویژگی‌ها از بین تعداد بسیار زیادی ویژگی‌های استخراج شده، برگزیده می‌شوند. معمولاً چند ده‌هزار ویژگی به صورت مستطیل‌هایی با ابعاد و مکان‌های مختلف از تصویر کلاس مورد نظر جده می‌شوند و بعد از آن توسط یک الگوریتم جستجوی حریصانه<sup>۱۲</sup> ویژگی‌های مفید و دارای اطلاعات استخراج می‌شوند. در ادامه خلاصه‌ای از مراحل اصلی برای پیدا کردن این ویژگی‌ها آورده شده است.

- تولید یک مجموعه بزرگ از ویژگی‌ها  $[F_i]$
- محاسبه آستانه بهینگی برای هر ویژگی که حداقل شباهت بصری را برای آن تعیین می‌کند که در یک تصویر شناسایی شود.
- انتخاب مجموعه شامل بهترین ویژگی‌های شیء

---

<sup>12</sup>- Greedy-search algorithm



شکل ۳.۲: مثال هایی از عکس های ماشین و غیر ماشین که برای آموزش شبکه به کار رفته است [۴]



شکل ۴.۲: مثال هایی از ویژگی های حاوی اطلاعات مفید که توسط ماشین استخراج شده اند [۴]

## ۲.۲.۲ شباهت سنجی و یافتن آستانه

حضور یک قطعه در تصویر توسط ترکیبی از اندازه گیری تشابه و تشخیص آستانه، تعیین می شود. اندازه گیری تشابه به حرکت دادن پنجره بر روی تصویر به راحتی قابل بدست آوردن است. و اما تعیین آستانه به شرح زیر می باشد. یک میجموعه ای از قطعات به نام  $X_i$  شامل متغیر های دودویی که نشان دهنده حضور یا عدم حضور قطعه در عکس می باشند، در نظر می گیریم. ایجاد این مجموعه نیاز به تشخیص آستانه یعنی مجموعه  $\theta_i$  که، نشان دهنده حداقل میزان شباهت برای هر قطعه است، دارد. حال مقدار  $X_i$  بسته به اینکه حداکثر میزان شباهت پیدا شده در عکس بزرگتر از  $\theta_i$  است یا خیر، مشخص می شود. مقدار آستانه  $\theta_i$  به طور خودکار با بیشینه کردن اطلاعات متقابل،  $I(X_i; C)$ ، میان قطعه  $X_i$  و کلاس متغیر دودویی  $C$ ، تنظیم می شود. احتمال های شرطی  $P(X_i(\theta_i) = 0|C)$  و  $P(X_i(\theta_i) = 1|C)$

در برآورد اطلاعاتی که از داده‌های آموزشی محاسبه می‌شوند، مورد نیاز هستند. آستانه تشخیص برای هر قطعه به عبارت زیر تعریف می‌شود:

$$\theta_i = \operatorname{argmax}_{\theta} I(X_i(\theta); C) = \operatorname{argmax}_{\theta} (H(C) - H(C|X_i(\theta))).$$

که  $H(x)$  و  $H(x|y)$  آنروپی شانون و آنروپی شرطی شانون هستند که  $x$  و  $y$  در اینجا عضو مجموعه  $0, 1$  می‌باشند. آنروپی شانون در بخش پیوست این پایان‌نامه آورده شده است.

### ۳.۲.۲ جستجوی حریصانه

در این مرحله قطعاتی که در قسمت قبل آستانه‌های آنها را مشخص کردیم به ترتیب نزولی مقادیر آستانه‌هایشان مرتب می‌کنیم. با الگوریتم حریصانه شروع به انتخاب قطعات مرتب شده می‌نماییم. تا زمانی این قطعات را برمی‌داریم که بودن قطعه‌ای دیگر به اطلاعات مجموعه ما اضافه نکند. که شرح معادله‌ی این کار در زیر آمده است:

$$X_k = \operatorname{arg max}_{X_i \in P_n} \min_{X_j \in S_n} (I(X_i, X_j, C) - I(X_j; C)).$$

و بعد از به روزرسانی مجموعه ویژگی‌های استخراج شده به شرح زیر خواهد بود:

$$S_{n+1} = S_n \cup \{X_k\}, \quad P_{n+1} = P_n \setminus X$$

که در اینجا مجموعه  $P$  تمام قطعات استخراج شده و مجموعه  $S$  ویژگی‌های حاوی اطلاعات مفید می‌باشند.

### ۴.۲.۲ طبقه‌بندی به وسیله‌ی جداساز خطی

در زمان طبقه‌بندی، ماشین یک بردار  $X = [X_1, X_2, \dots, X_N]$  تولید می‌کند که نشان‌دهنده رمزگذاری تصویر در فضای ویژگی‌ها می‌باشد. تصمیم

نهایی در مورد کلاس با جایگذاری بردار ویژگی در تابع طبقه‌بندی  $f(X)$  که ۰ یا ۱ برمی‌گرداند (بسته به تعیین وجود یا عدم وجود شیء)، گرفته می‌شود. یک قاعده ساده ترکیب ویژگی که در برنامه پیاده‌سازی شده، یک جداکننده خطی است، که توسط ماشین بردار پشتیبان خطی آموزش داده شده است. که این تابع جدا ساز خطی به فرم زیر تعریف می‌شود: که در اینجا  $x_i$  نشان‌دهنده مقادیر به دست‌آمده برای هر ویژگی به

$$f(X) = \begin{cases} 1 & \text{if } \sum_i \alpha_i X_i \geq \theta, \\ 0 & \text{otherwise.} \end{cases}$$

تنهایی است و  $\alpha_i$  وزن های ویژگی‌هاست که در طول آموزش به دست‌آمده‌اند و  $\theta$  بایاس<sup>۱۳</sup> می‌باشد.

## ۵.۲.۲ آزمایش‌ها

حال در این قسمت آزمایش های مختلف که با طبقه‌بندیهای مختلف و روش‌های متفاوت در استخراج ویژگی انجام شده‌اند را بیان می‌کنیم تا درک برتری این روش در نزد خواننده آسان‌تر شود. لازم به ذکر است از سه طبقه‌بند ماشین بردار پشتیبان خطی، ماشین بردار پشتیبان غیرخطی و درخت افزون‌شده نیو- ساده<sup>۱۴</sup> و استخراج ویژگی توسط موجک و قطعات حاوی اطلاعات مقایسه شده‌اند.

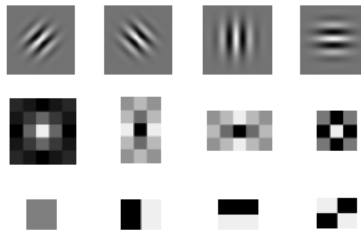
استخراج ویژگی با استفاده از موجک

اگ بخواهیم یک اشاره کوچک به این روش کنیم؛ این روش از فیلتر هایی استفاده می‌کند که معمولاً به شکل خط‌هایی در چهار جهت که در شکل زیر به خوبی نشان داده شده‌اند، می‌باشد. و با حرکت

<sup>13</sup>- bias

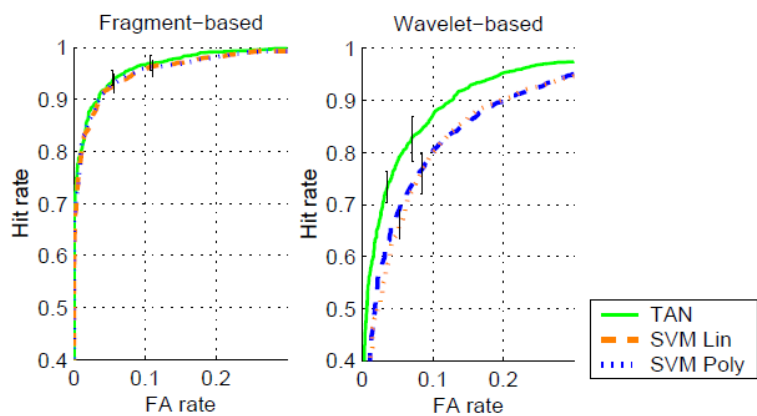
<sup>14</sup>- TAN

این فیلترها بر روی عکس لبه‌ها شناسایی شده و با استفاده از آن کلاس مربوطه را تشخیص می‌دهد.



شکل ۵.۲: فیلترهای موجک مورد استفاده معمول، در زمینه بازشناسی اشیاء [۴]

در این آزمایش در هر تکرار فرآیند اعتبارسنجی، از ۲۰۰ عکس ماشین و ۲۰۰ عکس غیر ماشین استفاده شده است و با این‌ها طبقه‌بندهای LSVM و TAN آموزش داده شدند. نمودارهای این آزمایش‌ها را به صورت زیر به دست آورده‌ایم: لازم به ذکر است که هر چه منحنی بالاتر رود یعنی درصد موفقیتش بیشتر شده و بهتر است. پس به وضوح استفاده از طبقه‌بند TAN موفقیت بیشتری نسبت به SVM ایجاد می‌کند، چرا که از پیچیدگی بیشتری نیز برخوردار است. به طور مثال در false-alarm rate، ۵٪ نرخ تشخیص در قسمت استخراج ویژگی قطعی‌تر از ۹۲٪ است، اما این نرخ در استخراج ویژگی به صورت موجک با طبقه‌بند LSVM حدود ۷۰٪ می‌باشد و با همان طبقه‌بند TAN، این نرخ به ۸۰٪ می‌رسد.



شکل ۶.۲: منحنی‌های مربوط به طبقه‌بندها و روش استخراج‌های مختلف [۴]



# پیوست آ

## A : آنتروپی شانون

آنتروپی متغیر تصادفی گسسته  $X$  با تابع جرم احتمال  $P(X)$  را با نماد  $H(X)$  نمایش می‌دهند و به صورت زیر تعریف می‌شود: در رابطه بالا

$$H(X) = E[I(X)] = E[-\log_b(P(X))].$$

$E$  امید ریاضی و  $I$  تابه میزان اطلاعات رویداد است.  $I(X)$  تابعی از یک متغیر تصادفی است.  $b$  پایه لگاریتم است و مقادیر مختلف آن آنتروپی را در واحدهای متفاوتی محاسبه می‌کند. متداولترین مقادیر برای  $b$ ، ۲ و  $e$  و ۱۰ هستند که به ترتیب، آنتروپی را در واحدهای بیت و nat و hartley محاسبه می‌کند.

می‌توان آن‌تروپی  $X$  را به صورت باز هم نوشت :

$$H(X) = \sum_{i=1}^n P(x_i) I(x_i) = - \sum_{i=1}^n P(x_i) \log_b P(x_i).$$

## پیوست ب

# B : آنتروپی شرطی

آنتروپی متغیر تصادفی  $X$  به شرط  $Y$  با توزیع احتمال مشترک  $P(X, Y)$  نیز به صورت زیر تعریف می‌شود:

$$H(X|Y) = - \sum_{i,j} P(x_i, y_j) \log \frac{P(x_i, y_j)}{P(y_j)}$$

$H(X|Y)$  میانگین اطلاعات حاصل از مشاهده  $X$  به شرط اطلاع از  $Y$  را نشان می‌دهد.

# Bibliography

- [1] Dahua Lin, Shuicheng Yan, Xiaoou Tang "Pursuing Informative Projection on Grassmann Manifold" Computer Vision and Pattern Recognition 2006 IEEE Computer Society Conference on, vol. 2 pp. 1734–1737 .2006
- [2] Grant Schindler, Matthew Brown, Richard Szeliski "City-Scale Location Recognition" Computer Vision and Pattern Recognition .2007 CVPR .07 IEEE Conference on, pp. 47–1 .2007
- [3] Sinisa Todorovic, Narendra Ahuja "Learning subcategory relevances for category recognition" Computer Vision and Pattern Recognition .2008 CVPR .2008 IEEE Conference on, pp. 48–1 .2008
- [4] Michel Vidal-Naquet Shimon Ullman "Object Recognition with Informative Features and Linear Classification" Faculty of Math-

ematics and Computer Science ,The Weizmann Institute of Science ,Rehovot ,76100 Israel .(2003)

- [5] Francois Fleuret “ Fast Binary Feature Selection with Conditional Mutual Information ” Journal of Machine Learning Research 5 (2004) 1555–1531
- [6] Ryan Gomes, Max Welling, Pietro Perona ”Incremental learning of nonparametric Bayesian mixture models” Computer Vision and Pattern Recognition .2008 CVPR .2008 IEEE Conference on. pp. 8–1 .2008
- [7] Miguel A. Vega-Rodriguez ”Feature Extraction and Image Processing” The Computer Journal, Volume: 47 Issue: 2 Jan. 2004
- [8] Bilgin Eşme, Bülent Sankur, Emin Anarvm ”Facial feature extraction using genetic algorithms” 1996 14th European Signal Processing Conference (EUSIPCO (1996 , Pages: 1 – 4 1996.
- [9] Yang Wu , Nanning Zheng , Qubo You , Shaoyi Du ”Object Recognition by Learning Informative, Biologically Inspired Visual Features” 2007 IEEE International Conference on Image Processing 16, Sept.–19 Oct. 2007
- [10] Johannes Schels, Joerg Liebelt, Rainer Lienhart ”Learning an object class representation

on a continuous viewsphere” Computer Vision and Pattern Recognition (CVPR) 2012 IEEE Conference on, pp. 3177–3179. 2012

- [11] Song Cao, Noah Snavely ”Minimal Scene Descriptions from Structure from Motion Models” Computer Vision and Pattern Recognition (CVPR) 2014 IEEE Conference on, pp. 468–469. 2014
- [12] Kevin Lin, Jiwen Lu, Chu-Song Chen, Jie Zhou ”Learning Compact Binary Descriptors with Unsupervised Deep Neural Networks” Computer Vision and Pattern Recognition (CVPR) 2016 IEEE Conference on, pp. 1192–1183. 2016
- [13] Marco Andreetto, Lih Zelnik-Manor, Pietro Perona ”Unsupervised learning of categorical segments in image collections” Computer Vision and Pattern Recognition Workshops . 2008 CVPRW . 08’ IEEE Computer Society Conference on, pp. 8–9. 2008
- [14] Nayyar A. Zaidi, David Suter ”Object Detection Using a Cascade of Classifiers” Digital Image Computing: Techniques and Applications (DICTA) 2008 pp. 605–609. 2008
- [15] S. Agarwal and D. Roth ”Learning a sparse representation for object detection” In Pro-

ceedings of ECCV ,2002 volume ,4 pages  
 ,130-113.2002

- [16] Y. Amit and D. Geman "A computational model for visual selection" *Neural Computation*, 1715-1691:(7)11 .1999
- [17] M. S. Bartlett and T. J. Sejnowski "Viewpoint invariant face recognition using independent component analysis and attractor networks" In *Advances in Neural Information Processing Systems*, volume ,9 page .817 The MIT Press. .1997
- [18] C. J. C. Burgess "A tutorial on support vector machines for pattern recognition" *Data Mining and Knowledge Discovery*, -121:(2)2 ,167 .1998
- [19] C. K. Chow and C. N. Liu "Approximating discrete probability distributions with dependence trees" *IEEE Transactions on Information Theory*, IT14(3):462-467, May .1968
- [20] T. M. Cover and J. Thomas "Elements of Information Theory" *Wiley Series in Telecommunications*. John Wiley and Sons, New-York, NY, USA, .1991
- [21] R. Duda and P. Hart "attern Classification and Scene Analysis" P. John Wiley and Sons, Inc., .1973

- [22] N. Friedman, D. Geiger, and M. Goldszmidt "Bayesian network classifiers" *Machine Learning*, 41:163–29:131, 1997
- [23] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel "Backpropagation applied to handwritten zip code recognition" *Neural Computation*, 1(4):541–551, Winter, 1989
- [24] MN. Littlestone "Learning quickly when irrelevant attributes abound: a new linear-threshold algorithm" *Machine Learning*, 2, 1988
- [25] B. Mel. Seemore "Combining color, shape, and texture histogramming in a neurally inspired approach to visual object recognition" *Neural Computation*, 8:4–9:777, 1997
- [26] M. L. Minsky and S. Papert "Perceptrons: An Introduction to Computational Geometry" MIT Press, Cambridge, Massachusetts, 1988
- [27] A. Mohan, C. Papageorgiou, and T. Poggio "Example-based object detection in images by components" *IEEE Trans. PAMI*, 23, 4, 2001
- [28] C. Papageorgiou, M. Oren, and T. Poggio "A general framework for object detection" In *Proceedings of International Conference on Computer Vision*, 1998



- [29] J. Pearl "Probabilistic reasoning in intelligent systems: Networks of Plausible Inference" Morgan Kaufmann, California, 1988.
- [30] H. Rowley, S. Baluja, and T. Kanade "Neural network-based face detection" IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(1):20, January, 1998
- [31] S. Russel and P. Norvig "Artificial Intelligence: A Modern Approach" Prentice Hall Series in Artificial Intelligence, Upper Saddle River, New Jersey, 1995
- [32] H. Schneiderman and T. Kanade "A statistical approach to 3d object detection applied to faces and cars" In Proceedings of the Eighth IEEE International Conference on Computer Vision, (2000) June, 2000
- [33] M. Turk and A. Pentland "Eigenfaces for recognition" Journal of Cognitive Neuroscience, 3(1):71, 1991
- [34] S. Ullman, E. Sali, and M. Vidal-Naquet "A fragment-based approach to object representation and classification" In Proc. 4th IWVF, (2001) May, 2001
- [35] V. Vapnik "The Nature of Statistical Learning Theory" Springer-Verlag, New York, 1995

- [۳۶] P. Viola and M. Jones "Rapid object detection using a boosted cascade of simple features" In Proceedings IEEE Conf. on Computer Vision and Pattern Recognition. .۲۰۰۱
- [۳۷] M. Weber, M. Welling, and P. Perona "Un-supervised learning of models for recognition" In Proc. ۶ th Europ. Conf. Comput. Vision, June .۲۰۰۰
- [۳۸] L. Wiskott, J.-M. Fellous, N. Krüger, and C. von der Malsburg "Face recognition by elastic bunch graph matching" IEEE Trans. on Pattern Analysis and Machine Intelligence. ,۷۷۹–۷۷۵:(۷) ۱۹ . ۱۹۹۷

# واژه‌نامه

Feature extraction . . . . .	استخراج ویژگی
Naive Bayes . . . . .	بیز ساده
Complexity . . . . .	پیچیدگی
Greedy-Search . . . . .	جستجوی حریصانه
informative . . . . .	دارای اطلاعات بالا
Convolutional Neural Network . . . . .	شبکه‌های عصبی پیچشی
Deep Neural Network . . . . .	شبکه عصبی عمیق
intensity . . . . .	شدت
visual recognition . . . . .	شناخت بصری
Classifier . . . . .	طبقه‌بند

Linear Classification . . . . .	طبقه‌بندی خطی
Neuropsychological . . . . .	علوم اعصاب و روانشناختی
cortex . . . . .	قشر بالای مغز
cortex visual primary . . . . .	قشر بصری اولیه
Cortex Inferotemporal . . . . .	قشر ناحیه ای
edge . . . . .	لبه
Support VectorMachine . . . . .	ماشین بردار پشتیبان
IT . . . . .	مخ
wavelet . . . . .	موجک
primate . . . . .	نخستیان
feature . . . . .	ویژگی
Artificial Intelligence . . . . .	هوش مصنوعی

# Abstract

We show that efficient object recognition can be obtained by combining informative features with linear and Naive-Bayes classification. The results demonstrate the superiority of informative class-specific features, as compared with generic type features such as wavelets, for the task of object recognition. We show that information rich features can reach optimal performance with simple linear separation rules, while generic feature based classifiers require more complex classification schemes. This is significant because efficient and optimal methods have been developed for spaces that allow linear separation. To compare different strategies for feature extraction, we trained and compared classifiers working in feature spaces of the same low dimensionality, using two feature types (image fragments vs. wavelets) and two classification rules (linear hyperplane and a Bayesian Network). The results show that by maximizing the individual information of the features, it is possible to obtain efficient classification by a simple linear separating rule, as well as more efficient learning.



Faculty of Science  
School of mathematics, statistics and computer  
science

# Object Recognition With Informative Features

By

Mohammad Amin Mozaffari

Supervisor

Dr. Mohammad Ganjtabesh

Project for receiving bachelor degree  
Computer Science and its applications

July 2018