



# شناسایی نوع و مدل وسیله نقلیه با استفاده از

## مجموعه بخش‌های متمایزکننده

محسن بیگلری<sup>۱\*</sup>، علی سلیمانی<sup>۲</sup> و حمید حسن‌پور<sup>۳</sup>

<sup>۱</sup> دانشکده مهندسی کامپیوتر و آی‌تی، دانشگاه صنعتی شاهرود، شاهرود، ایران

<sup>۲</sup> دانشکده مهندسی برق و رباتیک، دانشگاه صنعتی شاهرود، شاهرود، ایران

### چکیده

طبقه‌بندی دقیق اشیاء (Fine-Grained Recognition) چالشی است که جامعهٔ بینایی ماشین در حال حاضر با آن روبه‌رو شده است. در این نوع طبقه‌بندی گروه کلی شیء مشخص بوده و هدف تعیین زیرگروه دقیق آن است؛ شناسایی نوع و مدل وسیله نقلیه (VMMR) نیز در این حوزه قرار می‌گیرد. این مسئله به دلیل وجود تعداد طبقه‌های زیاد، تفاوت درون طبقه‌ای بسیار و تفاوت بین طبقه‌ای کم از مسائل طبقه‌بندی دشوار به‌شمار می‌رود. در این مقاله روشی مبتنی بر بخش برای شناسایی نوع و مدل خودرو پیشنهاد شده است. این روش برای طبقه‌بندی طبقه‌های مختلف خودرو، ابتدا بخش‌های متمایزکنندهٔ هر یک را به صورت خودکار می‌یابد؛ سپس با استخراج ویژگی از این بخش‌ها و رابطه هندسی بین آن‌ها، یک مدل می‌آموزد. وزن بخش‌های مختلف هر مدل به صورت پویا و با استفاده از مجموعه داده‌های آموزشی یاد گرفته می‌شود. سامانهٔ پیشنهادی با ترکیب این مدل‌ها به شناسایی طبقهٔ خودرو می‌پردازد. برای آزمایش سامانهٔ پیشنهادی و به دلیل عدم وجود مجموعه داده به اشتراک گذاشته شده، یک مجموعه داده با بیش از ۵۰۰۰ خودرو از ۲۸ طبقهٔ مختلف تهیه و به صورت کامل علامت‌گذاری شده است. نتیجهٔ آزمایش‌های انجام شده بر روی این تصاویر که دارای تغییرات روشنایی زیاد و تغییرات زاویه اندک هستند، نشان از دقت بالای روش پیشنهادی دارد.

واژگان کلیدی: شناسایی دقیق اشیاء، شناسایی نوع و مدل وسیله نقلیه، VMMR، رویکرد مبتنی بر بخش

## Using Discriminative Parts for Vehicle Make and Model Recognition

Mohsen Biglari<sup>1\*</sup>, Ali Soleimani<sup>2</sup> & Hamid Hassanpour<sup>3</sup>

<sup>1,3</sup>Department of Computer Engineering and IT, Shahrood University of Technology,  
Shahrood, Iran

<sup>2</sup>Department of Electrical Engineering and Robotics, Shahrood University of Technology,  
Shahrood, Iran

### Abstract

In fine-grained recognition, the main category of object is well known and the goal is to determine the subcategory or fine-grained category. Vehicle make and model recognition (VMMR) is a fine-grained classification problem. It includes several challenges like the large number of classes, substantial inner-class and small inter-class distance. VMMR can be utilized when license plate numbers cannot be identified or fake number plates are used. VMMR can also be used when specific models of vehicles are required to be automatically identified by cameras. Few methods have been proposed to cope with limited lighting conditions. A number of recent studies have shown that latent SVM trained on a large-scale dataset using data mining can achieve impressive results on several object classification tasks. In this paper, a novel method has been proposed for VMMR using a modified version of latent SVM. This method finds discriminative parts of each class of vehicles automatically and then learns a model for each class using features extracted from these parts and spatial relationship between them. The parts weights of each model are tuned using training dataset. Putting this individual models together, our proposed system can classify vehicles make and model. All training and testing steps of the proposed system are done automatically. For training and testing the

\* Corresponding author

\* نویسندهٔ عهده‌دار مکاتبات

فصلنامه



performance of the system, a new dataset including more than 5000 vehicles of 28 different make and models has been collected. This dataset poses different kind of challenges, including variations in illumination and resolution. The experimental results performed on this dataset show the high accuracy of our system.

**Keywords:** Fine-grained recognition, vehicle make and model recognition, VMMR, part-based approach

## ۱- مقدمه

شناسایی نوع و مدل وسیله نقلیه<sup>۱</sup> (VMMR) از موضوعاتی است که به دلیل توجه کم پژوهشگران، مطالعات زیادی روی آن صورت نگرفته است؛ هم‌اکنون، اغلب سامانه‌های نظارتی و امنیتی از پلاک خودروها استفاده می‌کنند. با اینکه سامانه‌های مبتنی بر پلاک به دقت قابل قبولی رسیده‌اند، ولی پاسخ‌گوی تمامی نیازها نیستند. نیازهایی همچون تشخیص جعل پلاک و ردگیری یک خودرو با مدل مشخص به‌خصوص در حالتی که پلاک قابل تشخیص نیست. با اضافه کردن قابلیت شناسایی نوع و مدل وسیله نقلیه به سامانه‌های نظارتی و امنیتی، گستره جدیدی از قابلیت‌های کاربردی را می‌توان تعریف کرد.

سامانه‌های تشخیص و شناسایی وسیله نقلیه را به سه دسته کلی "تشخیص وسیله نقلیه"<sup>۲</sup>، "شناسایی دسته کلی وسیله نقلیه"<sup>۳</sup> و "شناسایی نوع و مدل وسیله نقلیه" می‌توان تقسیم‌بندی کرد. هدف در سامانه‌های تشخیص وسیله نقلیه، تشخیص مکان و محدوده خودرو در تصویر است؛ روش‌های بی‌شماری در این زمینه ارائه شده و به دقت‌های قابل قبولی دست یافته‌اند [1, 2]. یافتن دسته کلی وسیله نقلیه مانند وسایل نقلیه سنگین، سبک و سواری‌ها، هدف سامانه‌های دسته دوم است؛ روش‌های متعددی برای این منظور نیز ارائه شده است [3-6]. سامانه‌های دسته سوم با سخت‌ترین نوع مسئله در این سه دسته مواجهند؛ یعنی یافتن نوع و مدل دقیق وسیله نقلیه؛ برای مثال پراید ۱۳۱ و یا پژو ۲۰۶. وجود تعداد طبقه‌های زیاد، تفاوت درون طبقه‌ای بسیار و تفاوت بین طبقه‌ای کم این مسئله را نسبت به مسائل دسته دوم، دشوارتر می‌گرداند.

سامانه‌های VMMR کاربردهای متنوعی دارند. تشخیص تقلب در سامانه‌های نظارتی، جستجو برای خودرویی با نوع و مدل مشخص در سامانه‌های امنیتی دو نمونه از این کاربردها است. آمارگیری از فراوانی انواع و مدل‌های مختلف خودروها در محدوده‌های مختلف شهری و کشوری

کاربردهای مناسبی برای راهنمایی و رانندگی، فروشندگان خودرو و لوازم جانبی خودرو می‌تواند داشته باشد. در دهه اخیر مطالعاتی در رابطه با مشخصه‌های رانندگان وسایل نقلیه مختلف نیز صورت گرفته است، از قبیل رابطه جنسیت، درآمد و علائق شخص و خودرویی که سوار می‌شوند [7]. به این وسیله تبلیغات هدفمند برای وسایل نقلیه مختلف می‌توان ارائه کرد. در این مقاله رویکردی جدید برای شناسایی نوع و مدل خودرو پیشنهاد شده است. این رویکرد برای هر نوع و مدل از خودروها، به جستجوی مجموعه‌ای از بخش‌های متمایزکننده پرداخته و مدلی متشکل از این مجموعه بخش‌ها یاد می‌گیرد؛ بنابراین به‌ازای هر طبقه از وسایل نقلیه، یک مدل متناظر ایجاد می‌شود که شامل بخش‌های مختص به آن طبقه، وزن‌های مناسب برای فیلترهای مربوط به هر بخش و رابطه هندسی بخش‌ها نسبت به یکدیگر است؛ از این‌رو، هر مدل با استفاده از هر یک از این سه جنبه، قدرت تمایز خود را می‌تواند افزایش دهد. برای آزمایش رویکرد پیشنهادی، مجموعه داده‌ای بزرگ با بیش از ۵۰۰۰ خودرو از ۲۸ کلاس متفاوت تهیه و علامت‌گذاری شده است. آزمایش‌های انجام‌شده بر روی این مجموعه داده، نشان از کارایی بالای رویکرد پیشنهادی دارند.

در ادامه مقاله، در بخش دوم، مروری بر کارهای گذشته صورت گرفته است. رویکرد پیشنهادی در بخش سوم توضیح داده شده است. بخش چهارم شامل جزئیات مجموعه داده تهیه‌شده و نتیجه آزمایش‌های انجام‌شده بر روی آن است. نتیجه‌گیری و کارهای آینده نیز در بخش پنجم و آخر ارائه شده است.

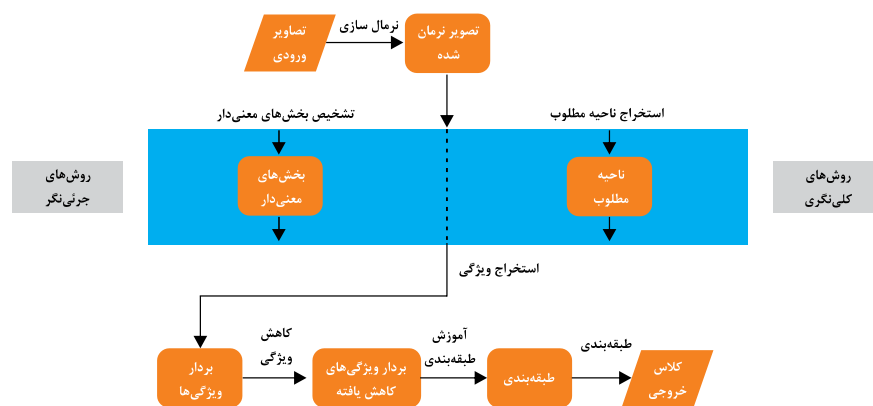
## ۲- مروری بر کارهای گذشته

از زاویه دید رویکرد پیشنهادی به مسئله، کارهای پیشین را به دو گروه کلی روش‌های کلی‌نگر و جزئی‌نگر می‌توان تقسیم‌بندی کرد. شکل (۱) روندنمای مراحل این روش‌ها را به نمایش گذاشته است. البته همه روش‌ها به‌الزام شامل همه بخش‌های نشان‌داده‌شده در شکل نیستند. در ادامه به توضیح روش‌های ارائه شده در هر یک از این دو گروه پرداخته شده است.

<sup>1</sup> Vehicle Make and Model Recognition

<sup>2</sup> Vehicle Detection

<sup>3</sup> Vehicle Type Recognition



(شکل-۱): روندنمای مراحل کار روش‌های کلی‌نگر و جزئی‌نگر

(Figure-1): Flowchart of Holistic and part-based approaches

## ۲-۱- روش‌های کلی‌نگر

اغلب روش‌های ارائه‌شده پیشین، در این دسته قرار می‌گیرند. این روش‌ها با پردازش کل تصویر و یا ناحیه مطلوبی از آن، به دنبال یافتن یک توصیف کلی از تصویر هستند؛ یکی از ضعف‌های عمده این روش‌ها، استفاده از فرضیات ناقص برای جستجو و استخراج ناحیه مطلوب است. از طرفی به دلیل استخراج ویژگی از یک ناحیه بزرگ، حساسیت به تغییرات اندک زاویه نیز از جمله ضعف‌های این روش‌ها به شمار می‌آید. کونوس [8] از مکان و اندازه پلاک برای تشخیص ناحیه مطلوب استفاده کرده است؛ به این صورت که ضرایب خاصی نسبت به ابعاد پلاک در نظر گرفته و محدوده اطراف آن را استخراج کرده است. این شیوه توسط افراد بسیاری به کار گرفته شده است [9-14]. کونوس سپس از نسخه ساده‌شده‌ای از عملگر [15] SIFT به صورت چندبخشی، برای استخراج ویژگی از ناحیه مطلوب بهره گرفته است. از فاصله اقلیدسی به عنوان معیار شباهت و از k-NN به عنوان طبقه‌بند استفاده شده است. در این مقاله، آزمایش‌ها بر روی دو مجموعه داده، یکی برای سواری‌ها که از یکصد تصویر در ده طبقه و دیگری برای وسایل نقلیه سنگین که از پانصد تصویر در هفت طبقه تشکیل شده، صورت گرفته است. بهترین دقت گزارش‌شده برای این دو مجموعه داده، به ترتیب ۹۴٪ و ۹۲٪ هستند. به دلیل شباهت ساختاری زیاد روش‌های ارائه‌شده و به منظور صرفه‌جویی در فضا، جدول (۱) عملکرد این روش‌ها را به همراه دقت گزارش‌شده برای هر یک، خلاصه کرده است.

## ۲-۲- روش‌های جزئی‌نگر

روش‌های اندکی در این دسته ارائه شده‌اند. این روش‌ها به دنبال یافتن بخش یا بخش‌های معنی‌دار یا متمایزکننده در تصویر هستند. برای مثال، یافتن مکان چراغ‌ها و استفاده از آن‌ها برای جداسازی طبقه‌ها. البته هر بخش معنی‌داری به‌الزام متمایزکننده نیست. ضعف عمده این روش‌ها در مکان‌یابی بخش‌های مناسب است؛ در صورتی که بخش موردنظر به درستی مکان‌یابی نشود و یا قدرت تمایز قابل قبولی نداشته باشد، عملکرد کلیه مراحل بعدی تحت تأثیر قرار خواهد گرفت. در صورتی که از الگوریتم‌های هوشمند برای مکان‌یابی بخش‌ها استفاده نشود، این روش‌ها نیز عملکردی مشابه با روش‌های کلی‌نگر خواهند داشت.

سایلس و همکاران [16] نشان‌واره<sup>۲</sup> را بخشی متمایزکننده برای شناسایی مدل خودروها در نظر گرفته‌اند. آن‌ها ابتدا با استفاده از مکان و اندازه پلاک، ناحیه مطلوب را استخراج و سپس به روش نگاشت ویژگی تناسب فاز<sup>۳</sup> نواحی برجسته مانند لبه‌ها و گوشه‌ها را استخراج کرده‌اند. آنگاه با نگاشت خروجی الگوریتم بر روی محور افقی و فیلتر کردن مقادیر کمتر از یک حد آستانه، محل تقریبی نشان‌واره را تشخیص داده‌اند. آن‌ها از روش SIFT برای استخراج ویژگی از نشان‌واره و از k-NN برای طبقه‌بندی تصاویر استفاده و دقت ۹۱٪ را برای مجموعه داده‌ای با ۱۲۰۰ نشان‌واره گزارش کرده‌اند.

جدول (۲) روش‌های ارائه‌شده در این گروه را دسته‌بندی و عملکرد هر یک را به صورت خلاصه ارائه کرده است.

<sup>۲</sup> Logo<sup>۳</sup> Phase Congruency Feature Map (PCFM)<sup>۱</sup> Scale-Invariant Feature Transform

(جدول-۱): خلاصه نتایج ارائه شده توسط روش های گذشته کلی نگر

(Table-1): Summarized results of previous holistic methods

مرجع	استخراج ویژگی از	روش استخراج ویژگی	روش طبقه بندی	مجموعه داده	نمای تصاویر	دقت گزارش شده
[8]	ناحیه مطلوب	SIFT چند بخشی	k-NN و فاصله اقلیدسی	۱۰۰ تصویر از ۱۰ کلاس سواری ۵۰۰ تصویر از ۷ کلاس سنگین	جلو	۹۴٪ ۹۲٪
[17]	ناحیه مطلوب	گرادیان های نگاشت شده ی مربعی <sup>۱</sup>	k-NN و فاصله اقلیدسی	۱۱۳۲ تصویر از ۷۲ کلاس	جلو	۹۳٪
[13]	ناحیه مطلوب	تشخیص گوشه ی هریس <sup>۲</sup> چند بخشی	بیز ساده	۲۶۲ تصویر از ۷۴ کلاس	جلو	۹۶٪
[18]	ناحیه مطلوب و چراغ ها	تشخیص لبه ی کنی <sup>۳</sup>	k-NN MLP	۱۸۰ تصویر از ۶ کلاس	جلو	۹۹٪/۹۹ ۹۹٪/۵۳
[19]	ناحیه مطلوب	تبدیل پیچک <sup>۴</sup>	SVM k-NN	۳۰۰ تصویر از ۵ کلاس	پشت	۹۹٪ ۹۳٪
[20]	ناحیه مطلوب	تبدیل کنتورلت <sup>۵</sup> و کاهش ویژگی با D-LDA <sup>۶</sup>	SVM	۳۰۰ تصویر از ۲۰ کلاس	جلو	۹۶٪
[21]	ناحیه مطلوب	SURF چند بخشی HOG <sup>۶</sup> چند بخشی	SVM چند گانه و رای گیری	۶۹۳۶ تصویر از ۲۹ کلاس	جلو	۹۹٪
[22]	ناحیه مطلوب	SIFT و کدگذاری ویژگی پراکنده <sup>۷</sup> برای کاهش ویژگی	SVM	۲۵۰۰ تصویر از ۱۰ کلاس	جلو	۹۳٪/۲۰
[14]	ناحیه مطلوب	هیستوگرام شکل انرژی محلی <sup>۸</sup>	SVM	۱۹۶ تصویر از ۲۲ کلاس (آموزش) ۳ ویدئو به طول ۱-۲ دقیقه (آزمایش)	جلو	۹۵٪/۸۳
[23]	ناحیه مطلوب	SIFT و SURF هیستوگرام لبه <sup>۹</sup>	k-NN	۳۸۵۹ تصویر از ۱۷ کلاس	جلو	۹۱٪/۷
[24]	ناحیه مطلوب	استفاده از پیکسل ها به عنوان ویژگی	شبکه عصبی عمیق <sup>۱۰</sup>	۳۲۱۰ تصویر از ۱۰۷ کلاس	جلو	۸۸٪/۲
[25]	ناحیه مطلوب	SURF چند بخشی و تشکیل کیسه ای از کلمات تصویری <sup>۱۱</sup>	SVM چند گانه و رای گیری	۱۶۶۰۲ از ۲۹ کلاس	جلو	۹۵٪/۷۷

(جدول-۲): خلاصه نتایج ارائه شده توسط روش های گذشته جزئی نگر

(Table-2): Summarized results of previous part-based methods

مرجع	استخراج ویژگی از	روش استخراج ویژگی	روش طبقه بندی	مجموعه داده	نمای تصاویر	دقت گزارش شده
[16]	نشان واره	SIFT	k-NN	۱۲۰۰ تصویر از ۱۰ کلاس	جلو	۹۱٪
[26]	نشان واره و پنجره ای در سمت چپ آن	DCT <sup>۱۲</sup>	SVM	۱۰۹۶ تصویر از ۱۲ کلاس	جلو	۹۷٪
[27]	چراغ های عقب و ناحیه مطلوب	لبه های ناحیه مطلوب، چراغ ها و زاویه چراغ ها با پلاک	k-NN	۴۰ تصویر از ۱۸ کلاس (آموزش) ۱۸ ویدئو از ورود ۱۸ خودرو به پارکینگ (آزمایش)	پشت	۸۹٪
[28]	بخش های پویا	LESH	k-NN و معیار شباهت K-L و آگرایی	۳۸ کلاس با حداقل ۱۵ نمونه برای هر یک	پشت جلو	۶۲٪ ۴۸٪
[29]	نشان واره	SIFT	شبکه عصبی احتمالی	۱۱۰ تصویر از ۱۱ کلاس	جلو	۸۵٪ (نوع) ۵۴٪ (مدل)
[30]	نوشته های <sup>۱۳</sup> پشت خودرو	اندازه و مکان نوشته ها و HOG	SVM	۱۳۴۲ تصویر از ۵۲ کلاس	پشت	۹۳٪/۷۵

<sup>7</sup> Sparse Feature Coding

<sup>8</sup> Local Energy Shape Histogram (LESH)

<sup>9</sup> Edge Histogram

<sup>10</sup> Deep Neural Network

<sup>11</sup> Bag of Visual Words

<sup>12</sup> Discrete Cosine Transform

<sup>13</sup> Emblems

<sup>1</sup> Squared Mapped Gradients

<sup>2</sup> Harris Corner Strength

<sup>3</sup> Canny Edge Detector

<sup>4</sup> Curvelet Transform

<sup>5</sup> Contourlet Transform

<sup>6</sup> Histogram of Oriented Gradients

## ۳- روش پیشنهادی

روش‌های تشخیص و شناسایی اشیا به‌مرور زمان و به‌دلیل مواجه شدن با چالش‌های جدید، به روش‌های جزئی‌نگر گرایش پیدا کرده‌اند [31-33]. برای مثال جهت تشخیص موفق انسان در تصویر باید مشکلات متعددی مانند تغییرات روشنایی، زوایای متفاوت و انسداد بخشی از بدن را پشت سر گذاشت. روش‌های مبتنی بر بخش<sup>۱</sup> نیز که در دسته روش‌های جزئی‌نگر قرار می‌گیرند، به همین منظور ارائه شده‌اند. این روش‌ها که اغلب در تشخیص اشیا به کار گرفته شده‌اند، با استخراج بخش‌های متمایزکننده از تصویر به طبقه‌بندی تصاویر می‌پردازند [32, 33].

روش پیشنهادی یک روش مبتنی بر بخش است که بر مبنای ایده مطرح شده در [32] شکل گرفته است. ایده مطرح‌شده در این مقاله، پیش از این در موضوع شناسایی نوع و مدل خودرو ارائه نشده است. از آنجا که همه خودروها دارای بخش‌های مشابه و مشخصی هستند، یک رویکرد مناسب برای تمایز آن‌ها از یکدیگر، تمرکز بر روی بخش‌های آن‌ها است. این رویکرد منطقی است، زیرا تفاوت ظاهری مدل‌های مختلف، در طراحی و شکل بخش‌ها و ارتباط بین آن‌ها نهفته است. برای مثال، با در نظر گرفتن تنها چراغ‌های جلوی خودروها تعداد مدل‌های بسیاری را از هم می‌توان تفکیک کرد. در اینجا یک سؤال مهم مطرح می‌شود. کدام بخش‌ها برای تفکیک مدل‌های مختلف خودرو مناسب‌تر هستند؟ ذهن انسان برای هر مدلی از خودرو، یک یا چند مشخصه یاد می‌گیرد. این مشخصه‌ها به‌الزام یک بخش معنی‌دار مثل چراغ نیستند؛ ممکن است یک خم یا پیچ در بخش مشخصی از بدنه خودرو باشند. روش پیشنهادی با توجه به طرز عملکرد ذهن انسان، برای هر مدل از خودرو، مشخصه‌های منحصربه‌فرد یاد می‌گیرد. این بخش‌ها یک بخش واقعی از خودرو (مثل جلوپنجره یا چراغ‌ها) و یا تنها بخش کوچکی از آن می‌توانند باشند.

## ۱-۳- مدل

روش ارائه شده در این مقاله، برای هر طبقه از خودروها یک مدل (طبقه‌بند) یاد می‌گیرد. هر مدل دارای چند بخش است: فیلتر ریشه، فیلتر بخش‌ها و رابطه بخش‌ها نسبت به یکدیگر. فیلترها همگی خطی هستند. یک فیلتر درواقع یک ماتریس سه‌بعدی است که بعد نخست و دوم آن، موقعیت و بعد سوم آن ضرایب را تعیین می‌کند. ضرایب مربوطه در موقعیت مشخص‌شده

<sup>۱</sup> Part-based Methods

اعمال می‌شوند. در مرحله آزمایش، ابتدا هرم اندازه‌های متفاوت تصویر تشکیل شده و از همه سطح‌های هرم ویژگی استخراج می‌شود (با هر روش دلخواهی)؛ سپس فیلتر ریشه و بخش‌ها به مکان‌های مختلف در ماتریس ویژگی‌ها در هر یک از سطح‌های هرم اعمال می‌شود. مکان‌هایی که مجموع امتیاز آن‌ها از آستانه مشخصی بیشتر است، به‌عنوان خروجی‌های مثبت علامت‌گذاری خواهند شد. برای اعمال کردن فیلتر بخش‌ها باید رابطه بخش‌ها با یکدیگر را نیز لحاظ کرد. یافتن چنین مدلی برای هر طبقه از خودروها، یک مسئله بهینه‌سازی دشوار به حساب می‌آید [32].

پارامترهای مدل  $M$  را به دو بخش تقسیم می‌کنیم. فیلترها ( $\beta$ ) و رابطه بین بخش‌ها ( $Z$ ). هدف، یافتن هر دوی این پارامترها با استفاده از مجموعه داده آموزشی برچسب‌دار  $D$  است به شکلی که  $D = \{ \langle x_1, y_1 \rangle, \dots, \langle x_n, y_n \rangle \}$  و  $y_i \in \{1, -1\}$  یک طبقه‌بند را در نظر بگیرید که نمونه  $x$  را با تابعی به فرم زیر امتیازدهی می‌کند:

$$f_{\beta}(x) = \max_{z \in Z(x)} \beta \cdot \phi(x, z) \quad (1)$$

تابع  $\phi(x, z)$  بردار ویژگی بخشی از نمونه  $x$  که متناسب با رابطه  $Z$  است را برمی‌گرداند.  $Z(x)$  مجموعه مقادیر ممکن  $Z$  برای نمونه  $x$  است. با اعمال کردن یک حد آستانه به خروجی تابع  $f_{\beta}$ ، عضویت نمونه به مدل  $M$  را می‌توان تعیین کرد. مشابه با ماشین بردار پشتیبان<sup>۲</sup> (SVM) استاندارد، برای به‌دست آوردن  $\beta$  با استفاده از مجموعه داده  $D$ ، تابع هدف زیر را کمینه می‌کنیم:

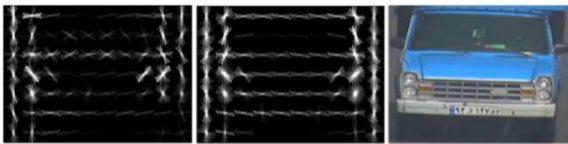
$$L_D(\beta) = \frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^n \max(0, 1 - y_i f_{\beta}(x_i)) \quad (2)$$

بخش نخست از فرمول (۲) برای بیشینه‌کردن فاصله بین بردارهای پشتیبان و بخش دوم تابع خطای لولایی<sup>۳</sup> استاندارد است. ثابت  $C$  نیز تاثیر خطای لولایی بر هزینه کل را تعیین می‌کند. چیزی که فرمول (۲) را از رابطه استاندارد SVM متفاوت می‌سازد، وجود مجموعه  $Z(x)$  است. در صورتی که طول این مجموعه برابر با یک باشد،  $f_{\beta}$  نسبت به  $\beta$  خطی بوده و فرمول (۲) تبدیل به رابطه SVM خطی خواهد شد. در غیر این صورت، با یک رابطه غیر محدب روبه‌رو هستیم. برای حل این رابطه از ماشین بردار پشتیبان با مقادیر مجهول<sup>۴</sup> (LSVM) استفاده می‌کنیم [32]. این الگوریتم از یک رابطه

<sup>۲</sup> Support Vector Machine<sup>۳</sup> Hinge Loss<sup>۴</sup> Latent SVM

یک نکته مهم در الگوریتم ارائه شده، شکل به کارگیری تصاویر آموزشی است. ترکیب پیشنهادی، تعادل مناسبی بین دقت بالا و سرعت قابل قبول در مرحله آموزش برقرار می کند. در مرحله نخست، فیلتر ریشه با استفاده از کلیه نمونه های مثبت مربوط به طبقه جاری و نمونه های منفی که شامل هیچ خودرویی نیستند، آموزش داده می شود. در پایان این مرحله، فیلتر ریشه، تفاوت بین خودرو و غیرخودرو را یاد می گیرد. در مرحله دوم، از تصاویر خودروهای متعلق به سایر طبقه ها به عنوان نمونه های منفی استفاده می شود. در نتیجه، ضرایب فیلتر ریشه برای تمایز دادن بین طبقه جاری و سایر طبقه ها، به سمت ساختار اختصاصی طبقه جاری متمایل می شود. شکل (۲)، فیلتر ریشه را در پایان مرحله نخست و دوم برای طبقه "وانت زامیاد" نمایش داده است. با مقایسه ضرایب فیلتر ریشه در پایان دو مرحله، بخش های منحصر به فرد خودروی وانت زامیاد مانند چراغ ها و جلوپنجره به وضوح قابل مشاهده هستند.

محل قرارگیری الگوریتم تعیین مکان بخش ها بسیار مهم است؛ پس از پایان مرحله دوم، فیلتر ریشه خود را با ساختار انحصاری طبقه مورد نظر تطبیق داده است؛ در نتیجه بخش هایی با قدرت تمایز بالاتر را در فیلتر ریشه می توان جستجو کرد. در صورتی که تعیین مکان بخش ها در پایان مرحله نخست اتفاق می افتاد، چنین امری میسر نمی شد.



(شکل-۲): مقایسه فیلتر ریشه آموزش دیده شده در پایان مرحله نخست (وسط) و دوم (چپ) برای طبقه "وانت زامیاد" (راست). از توصیف گر هیستوگرام گرادیان های جهت دار<sup>۱</sup> [31] برای استخراج ویژگی استفاده شده است.

(Figure-2): The comparison of a trained root filter at the end of stage 1 (middle), and stage 2 (left) for "Zamyad Truck" category (right). HOG descriptor is used for feature extraction [31].

در مسائلی که تعداد نمونه های منفی بسیار بیشتر از نمونه های مثبت است، استفاده از همه نمونه های منفی در فرآیند آموزش امکان پذیر نبوده و در عمل آموزش را غیرممکن می گرداند. با توجه به اثبات ارائه شده در [32]، با یک روال تکرار شونده و با تنها زیرمجموعه ای از نمونه های منفی، مدلی با قدرت یکسان می توان به دست آورد. در این روال تکرار شونده، در هر مرحله از آموزش، تعدادی از نمونه های

دومرحله ای برای بهینه سازی فرمول (۲) بهره می برد. در مرحله نخست، بهترین مقدار  $Z$  برای هر یک از نمونه های مثبت محاسبه می شود؛ سپس در مرحله دوم، مجموعه  $Z(x)$  را ثابت در نظر گرفته تا فرمول (۲) به یک مسئله محدب تبدیل شود؛ آنگاه  $L_D(\beta)$  را بهینه کرده و مقدار  $\beta$  را به دست می آوریم. این دو مرحله آنقدر تکرار می شوند تا تغییرات پاسخ ها از حد معینی کوچک تر شوند.

در مرجع [32] روشی مشابه با این روش برای تشخیص اشیاء به کار گرفته شده است. در مسئله تشخیص شیء، با دو طبقه مواجه هستیم که مسئله ساده تری است. در شناسایی نوع و مدل خودرو، علاوه بر افزایش تعداد طبقه ها، رویکرد نیز تغییر می کند؛ زیرا شباهت ساختاری بسیار زیاد مدل های مختلف خودروها نیاز به قدرت بسیار بالای طبقه بند در تفکیک صحیح آن ها دارد. در ادامه به تغییرات مورد نیاز منحصر به مسئله جاری پرداخته شده است.

در رویکرد پیشنهادی، برای هر طبقه  $c$  از خودروها یک مدل  $M_c$  یاد گرفته می شود که از سه بخش فیلتر ریشه، فیلتر بخش ها و رابطه هندسی بین بخش ها ( $S$ ) تشکیل شده است:  $M_c = (F_r, F_{parts}, S_{parts})$  الگوریتم (۱)، مراحل آموزش یک مدل را ارائه کرده است. الگوریتم پیشنهادی از پنج مرحله اصلی تشکیل شده است. در مرحله نخست و دوم تنها فیلتر ریشه از مدل مورد نظر آموزش داده می شود. در مرحله سوم، مکان بخش ها تعیین شده و فیلتر بخش ها مقداردهی اولیه می شوند. مرحله چهارم و پنجم به آموزش کلیه بخش های مدل می پردازد. مرحله سوم از الگوریتم در بخش ۲-۳- باز شده است.

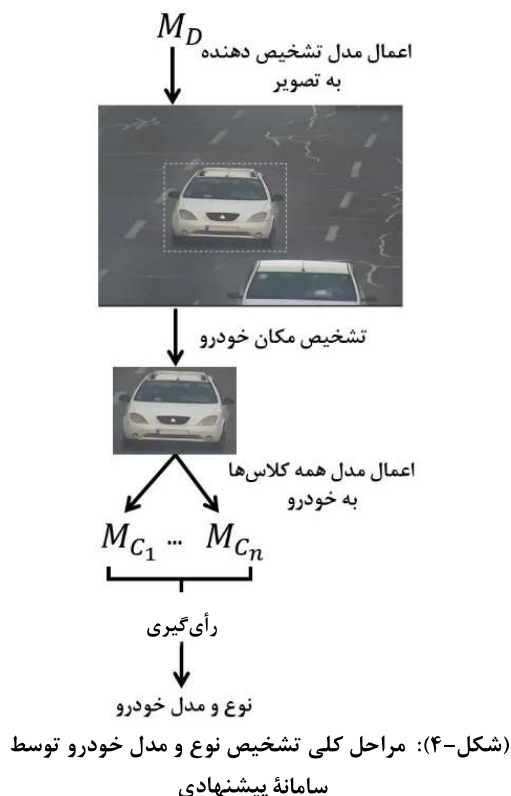
### Function $M_c = \text{Train}(\text{Class } c)$

$F_r$ :  $[0]_{d \times d}$   
 pos: positive samples (vehicle class  $c$ )  
 neg: negative samples (non-vehicle)  
 vehNeg: negative samples (other vehicle classes)  
 1. Train  $F_r$  with {pos,neg} samples using SVM  
 2. Retrain  $F_r$  with {pos,vehNeg} samples using SVM  
 3. Initialize parts filters ( $F_{parts}$ ) and positions ( $S_{parts}$ ) using  $F_r$   
 $M_c = \{F_r, F_{parts}, S_{parts}\}$   
 hardNeg: {x random negative examples}  
 4. Train  $M_c$  with {pos,hardNeg} samples using LSVM  
 5. Retrain  $M_c$  with {pos,vehNeg} samples using LSVM  
 Return  $M_c$   
 end

(الگوریتم-۱): مراحل کلی آموزش یک مدل برای یک طبقه از خودرو (Algorithm-1): The training algorithm of a model for a specific category of vehicles

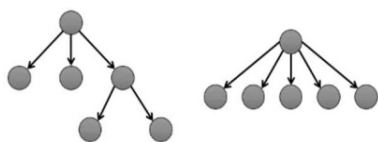
<sup>۱</sup>HOG





(Figure-4): The flowchart of proposed system for vehicle make and model recognition

درختی و ستاره‌ای شکل (۵) یک الگوریتم برنامه‌نویسی پویای بهینه می‌توان ارائه کرد که همه حالت‌های مختلف انطباق را در زمان چندجمله‌ای بررسی کند [32]. می‌دانیم که خودرو دارای ساختار سخت‌است؛ یعنی مکان بخش‌های آن قابل جابه‌جایی نیستند. در نتیجه برای مدل کردن رابطه هندسی بین بخش‌ها از ساختار تصویری ستاره‌ای استفاده کرده‌ایم. استفاده از این ساختار در مقایسه با ساختار درختی یک مزیت مهم دیگر نیز به همراه دارد. الگوی موجود در ساختار درختی، مدل را مجبور به یافتن بخش‌هایی می‌کند که شاید بهترین بخش‌ها برای تفکیک یک طبقه از سایر طبقه‌ها نباشند؛ ولی در ساختار ستاره‌ای هیچ رابطه‌ای بین خود بخش‌ها وجود ندارد و مکان همه بخش‌ها تنها نسبت به ریشه تعیین می‌شود. درضمن هر مدل تعداد بخش‌های متغیری می‌تواند داشته باشد.



(شکل-۵): دو ساختار تصویری ستاره‌ای (راست) و درختی (چپ)

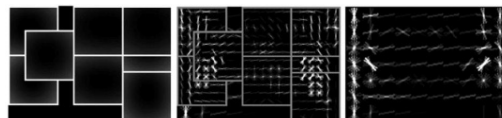
برای توصیف رابطه هندسی بین بخش‌ها

(Figure-5): Star (right), and tree (left) pictorial structures for describing the spatial relationship between parts.

<sup>2</sup>Rigid

منفی دشوار (که توسط مدل آموزش دیده شده تا این مرحله، قابل طبقه‌بندی نیستند) انتخاب شده و به نمونه‌های دشوار قبلی اضافه می‌شوند؛ مرحله بعدی آموزش به جای کل نمونه‌های منفی، تنها با همین نمونه‌ها انجام می‌شود. به همین دلیل، در مرحله چهارم تنها از نمونه‌های دشوار استفاده شده است. در مرحله پنجم که مرحله پایانی است، از همه نمونه‌های منفی با تعداد تکرار بسیار کمتر (نسبت به مرحله چهارم) بهره گرفته می‌شود تا مدل نهایی به دست آید. شکل (۳) مدل نهایی به دست آمده در پایان الگوریتم را برای طبقه "وانت زامیاد" نمایش داده است. در این شکل، تعداد بخش‌ها برابر با هشت قرار داده شده است. تمرکز بخش‌ها در اطراف چراغ‌ها و ساختار بخش‌ها از موارد قابل توجه در مدل به دست آمده است.

سامانه پیشنهادی ابتدا به صورت خودکار مکان خودرو در تصویر را تشخیص داده و سپس نوع و مدل آن را تعیین می‌کند. یک مدل که تنها از فیلتر ریشه تشکیل شده ( $M_D$ )، برای تشخیص خودرو آموزش داده شده است. پس از تشخیص مکان خودرو، محدوده موردنظر به صورت موازی به مدل‌های هریک از طبقه‌ها ( $\{M_{C_1} ... M_{C_n}\}$ ) ارسال می‌شود. طبقه‌ای که دارای بیشترین امتیاز است با رأی‌گیری به عنوان طبقه خروجی انتخاب می‌شود. اگر هیچ‌یک از مدل‌ها قادر به تشخیص مدل خودرو نباشد، طبقه ناشناخته به عنوان خروجی تعیین می‌شود. شکل (۴) مراحل توضیح داده شده را به نمایش می‌گذارد.



(شکل-۳): مدل نهایی آموزش دیده شده برای طبقه "وانت زامیاد". فیلتر ریشه (راست). فیلتر بخش‌ها (وسط) و هزینه جابه‌جایی بخش‌ها نسبت به ریشه (چپ): رنگ تیره‌تر به معنی خطای کمتر است.

(Figure-3): The final trained model for "Zamyad Truck" category. Right: root filter, Middle: part filters, Left: a spatial model for location of each part relative to the root. Black color means zero error.

## ۳-۲ - مجموعه بخش‌های متمایز کننده

برای مدل کردن رابطه هندسی بین بخش‌ها، روش‌های متعددی ارائه شده است [34]. برای تطابق یک الگو با تصویر نیز روش‌هایی وجود دارد که به دلیل پیچیدگی، کارایی یا سرعت بالایی ندارند. برای مثال روش‌های [35,36] برای یافتن یک الگوی منطبق بهینه، مکان بخش‌ها را به مجموعه کوچکی از مکان‌ها محدود می‌کنند. با استفاده از ساختارهای تصویری<sup>۱</sup>

<sup>1</sup>Pictorial Structure

## ۴- نتیجه آزمایش‌ها

### ۴-۱- مجموعه داده

پیش از این، مجموعه داده‌ای در این حوزه به اشتراک گذاشته نشده است. تنها در سال جاری دو مجموعه داده ارائه شده‌اند [37, 38]. مجموعه داده نخست دارای حاشیه‌نویسی نبوده و در نتیجه قابلیت استفاده به صورت خودکار را ندارد [38]. مجموعه داده دوم علاوه بر دشواری‌هایی که برای به دست آوردنش وجود دارد، به تازگی ارائه شده و در محدوده اطلاعات ما، هیچ آزمایشی بر روی آن گزارش نشده است [37]؛ در نتیجه امکان مقایسه روش پیشنهادی با سایر روش‌ها وجود ندارد.

برای آزمایش عملکرد روش پیشنهادی بر روی داده‌های واقعی، در بازه زمانی یک سال، فیلم‌هایی از چند دوربین مدار بسته که بر روی بزرگراه‌ها نصب شده بودند، تهیه شد. نمای خودروها رو به جلو است و دوربین‌ها به صورت میانگین در فاصله ده متری قرار گرفته‌اند. در مجموع، ۴۸۵۸ تصویر از ۲۸ نوع و مدل مختلف از ۵۹۹۱ خودرو تهیه شد. تصاویر دارای تغییرات روشنائی بسیار زیاد و وضوح متفاوت هستند که در شکل (۶) قابل مشاهده است. محدوده، نما، نوع و مدل تمامی خودروهای موجود در تصاویر علامت‌گذاری و منطبق با استاندارد مسابقه پاسکال [39] حاشیه‌نویسی شده است. دشواری تشخیص نوع و مدل هر خودرو نیز تعیین شده است. شکل (۷) اطلاعات حاشیه‌نویسی شده یک نمونه خودرو را نمایش داده است. این مجموعه داده با نام BVMMR Database v2 به همراه مجموعه کدهای مورد نیاز برای پردازش آن، از نشانی <http://mbt925.ir/publications.html> به صورت رایگان قابل دریافت است.



(شکل-۶): چند نمونه از شکل‌های متفاوت یک مدل یکسان در

مجموعه داده تهیه شده

(Figure-6): Different samples of a single vehicle model in our dataset

در مرحله آموزش مدل‌ها برای هر طبقه از خودروها، مکان بخش‌ها تغییر نمی‌کند، بلکه ضرایب یک تابع درجه دوم یاد گرفته می‌شوند که به عنوان یک تابع خطا عمل می‌کند. با این ترفند، به هر بخش اجازه می‌دهیم تا از مکان بهینه خود دور شود، ولی برای آن یک پناستی در نظر می‌گیریم. برای تعیین مکان بخش‌ها از الگوریتم (۲) بهره‌برده‌ایم. برای سادگی بیشتر در روش پیشنهادی، تعداد ( $n$ ) و اندازه بخش‌ها ( $S$ ) در همه مدل‌ها، ثابت و یکسان در نظر گرفته شده است.

```

Fr: root filter
n: number of parts
S: fixed size of parts
P: {}
for i=1:n
    r: find a region of size S in Fr with
        maximum energy (norm of weights)
    P: P ∪ r
    Fr (r)=0
end
bestLocations: [false]n×1
for a fixed number of iterations
    if bestLocations(1:n)==false then
        break
    end
    p: a random part in P
    P: P - p
    Fr (p)=initial value
    r: find a region of size S in Fr with
        maximum energy
    if r==p then
        bestLocations(index of p)=true
        continue
    end
    P: P ∪ r
    Fr (r)=0
end

```

(الگوریتم-۲): الگوریتم حریصانه تعیین مکان بخش‌ها

(Algorithm-2): The greedy algorithm used for initializing part location

الگوریتم (۲) از دو حلقه تشکیل شده است. در حلقه نخست،  $n$  بخش از فیلتر ریشه را که دارای بیشترین انرژی هستند، می‌یابیم. در حلقه دوم، سعی بر بهتر کردن نتیجه حاصل شده داریم. به این صورت که یکی از بخش‌ها را به صورت تصادفی انتخاب کرده و به دنبال مکان بهتری برای آن هستیم. تا زمانی که قادر به بهتر کردن مجموع انرژی به دست آمده از بخش‌ها باشیم (و به تعداد تکرار مشخص نرسیده باشیم)، این روال را تکرار می‌کنیم.



برای انجام آزمایش‌ها، به‌طور تقریبی نیمی از تصاویر برای آموزش و نیمی دیگر برای آزمایش در نظر گرفته شده‌اند. برای قابل مقایسه بودن کارهای آینده با نتایج ارائه شده در این مقاله، مجموعه آموزشی و آزمایشی استفاده شده نیز به همراه مجموعه داده، منتشر شده است. در مرحله آزمایش از حاشیه‌نویسی تصاویر هیچ استفاده‌ای نشده و مکان و محدوده خودروها همان‌طور که در شکل (۴) نشان داده شده است، به‌صورت خودکار (با دقت ۱۰۰٪) توسط سامانه پیشنهادی تشخیص داده می‌شود.

از ۲۸ طبقه موجود در مجموعه داده، ۲۱ طبقه را که دارای بیش از ۲۰ نمونه تصویر بوده‌اند را جدا کردیم. تعداد نمونه‌های هر طبقه در آورده شده است. برای دشوارتر کردن آزمایش، طبقه "سایر" را نیز اضافه کرده‌ایم. خودروهایی که شامل این ۲۱ طبقه نیستند، در طبقه "سایر" قرار گرفته‌اند. سامانه پیشنهادی در صورت عدم شناسایی یک نمونه ورودی، آن را در طبقه "سایر" طبقه‌بندی خواهد کرد.

برای آموزش مدل‌ها، از توصیف‌گر هیستوگرام گرادینان‌های جهت‌دار به‌دلیل قدرت بالا در توصیف و مقاومت در برابر تغییرات روشنایی بهره برده‌ایم [31]. برای استخراج ویژگی با استفاده از این توصیف‌گر، اندازه سلول  $8 \times 8$  و اندازه بلوک  $2 \times 2$  به کار گرفته شده است؛ این مقادیر به‌دلیل نتایج مناسب ارائه‌شده در [31] و به‌صورت تجربی انتخاب شده‌اند. جدول (۴)، دقت به‌دست آمده برای هر طبقه و جدول (۵)، دقت مجموع و میانگین را ارائه کرده است. نتایج با در نظر گرفتن رتبه نخست، سه رتبه نخست<sup>۱</sup> و پنج رتبه نخست محاسبه شده‌اند. دقت میانگین با میانگین‌گیری از دقت همه طبقه‌ها به‌دست می‌آید. دقت مجموع نیز برابر است با نسبت تعداد نمونه‌های درست طبقه‌بندی شده به کل نمونه‌ها. شکل (۸) ماتریس درهم‌ریختگی<sup>۲</sup> را تنها برای سه رتبه نخست نمایش داده است.



(شکل-۷): یک نمونه تصویر حاشیه نویسی شده از مجموعه داده

BVMMR 2016

(Figure-7): An example of annotated image in BVMMR 2016 dataset

(جدول-۳): تعداد نمونه‌های طبقه‌های انتخاب شده برای انجام

آزمایش‌ها

(Table-3): Samples count per classes in experiments

شماره کلاس	نوع و مدل	تعداد نمونه‌ها	شماره کلاس	نوع و مدل	تعداد نمونه‌ها
۱	لیفان ۶۲۰	۲۰	۱۲	پراید ۱۳۱	۲۰۰۱
۲	وانت مزدا	۴۲	۱۳	پراید ۱۳۲	۲۴۰
۳	مگان	۲۲	۱۴	پراید ۱۴۱	۱۶۰
۴	ام وی ام ۳۱۵	۲۶	۱۵	ال ۹۰	۱۸۹
۵	ام وی ام ۵۳۰	۳۶	۱۶	ریو	۷۲
۶	سایر	۲۶۸	۱۷	رانا	۸۱
۷	پژو ۲۰۶	۲۴۷	۱۸	سمند	۴۸۱
۸	پژو ۴۰۵	۷۸۵	۱۹	سمند سورن	۵۰
۹	پژو ۴۰۵ اس ال ایکس	۱۲۰	۲۰	تیبا	۱۷۲
۱۰	پژو پارس	۳۲۷	۲۱	زانتیا	۴۱
۱۱	پیکان ۸۰	۴۸۶	۲۲	وانت زامیاد	۱۲۵
تعداد کل نمونه‌ها		۵۹۹۱			

(جدول-۴): دقت به‌دست آمده (%) برای هر یک از طبقه‌ها توسط سامانه پیشنهادی

(Table-4): The proposed system accuracy (%) for each class

تعداد نمونه‌ها	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱
تعداد نمونه‌ها	۹	۱۹	۱۰	۱۴	۱۹	۱۴۲	۱۲۵	۳۸۳	۶۳	۱۵۹	۲۳۷
رتبه اول	۸۸/۸۹	۹۴/۷۴	۱۰۰	۱۰۰	۹۴/۷۴	۹۱/۵۴	۱۰۰	۱۰۰	۹۶/۸۳	۹۳/۰۸	۱۰۰
سه رتبه اول	۸۸/۸۹	۹۴/۷۴	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰
پنج رتبه اول	۸۸/۸۹	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰
تعداد نمونه‌ها	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	۲۰	۲۱	۲۲
تعداد نمونه‌ها	۱۰۰۴	۱۲۵	۸۰	۹۸	۳۸	۳۸	۲۴۲	۲۴	۹۰	۱۸	۶۳
رتبه اول	۹۹/۹۰	۹۹/۲۰	۹۶/۲۵	۱۰۰	۹۴/۷۴	۱۰۰	۹۹/۵۹	۹۱/۶۶	۱۰۰	۸۸/۸۹	۱۰۰
سه رتبه اول	۱۰۰	۹۹/۲۰	۹۷/۵۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۹۵/۸۳	۱۰۰	۸۸/۸۹	۱۰۰
پنج رتبه اول	۱۰۰	۹۹/۲۰	۹۷/۵۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰

<sup>۱</sup> Rank-3

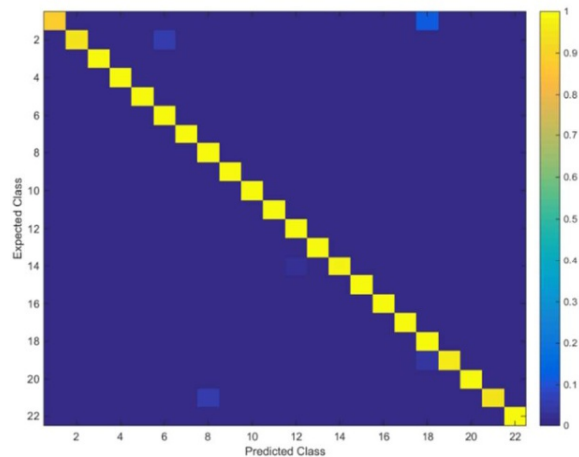
<sup>۲</sup> Confusion Matrix

(جدول-۵): دقت (٪) مجموع و دقت میانگین به دست آمده توسط

#### سامانه پیشنهادی

(Table-5): The proposed system overall and average accuracy (%)

رتبه نخست	سه رتبه نخست	پنج رتبه نخست	
۹۸/۶۶	۹۹/۷۳	۹۹/۸۶	دقت مجموع
۹۶/۸۲	۹۸/۴۱	۹۹/۳۴	دقت میانگین



(شکل-۸): ماتریس درهم ریختگی به دست آمده توسط سامانه پیشنهادی برای سه رتبه نخست

(Figure-8): The confusion matrix of the proposed system at rank 3

با توجه به دشواری تصاویر مجموعه داده، دقت کسب شده بسیار مناسب است. یکی از جنبه های مهم روال پیشنهاد شده برای آموزش یک مدل، عدم نیاز به تعداد تصاویر زیاد است؛ چنان که می بینیم طبقه های ۱ تا ۵ با تعداد تصاویر کم، با دقت بالایی طبقه بندی شده اند. نکته قابل توجه دیگر در نتایج ارائه شده، طبقه بندی درست اغلب تصاویر طبقه ۶ ("سایر") است؛ این مهم به دلیل حد آستانه مناسبی است که هر مدل آموزش دیده است. طبقه هایی که دارای تعداد نمونه های مناسبی هستند، منجر به یادگیری مدل قدرتمندتری شده اند. برای مثال دقت طبقه های ۸، ۱۲ و ۱۸ در حالت رتبه نخست و سه رتبه نخست تفاوت چندانی نکرده اند.

علت نتایج مناسب کسب شده توسط سامانه پیشنهادی در شکل آموزش مبتنی بر بخش آن نهفته است. سامانه برای یادگیری بخش های هر طبقه از خودروها بر روی دو جنبه هر بخش تکیه می کند. جنبه نخست، مکان هر بخش است؛ دو مدل متفاوت از خودروها که دارای بخش های به ظاهر مشابه هستند، ممکن است از نظر مکان قرارگیری بخش ها تفاوت داشته باشند. جنبه دوم، ضرایب فیلتر هر بخش است؛ دو طبقه

متفاوت که دارای مکان بخش های متمایز کننده مشابه هستند، ممکن است دارای ضرایب فیلترهای متفاوت باشند. این دو جنبه قدرت سامانه در تمایز مدل های شبیه به هم را افزایش می دهد.

مدت زمان تقریبی مورد نیاز برای پردازش یک تصویر با ابعاد  $800 \times 600$  توسط سامانه پیشنهادی بر روی رایانه ای با پردازنده Corei7-4710HQ 2.50 GHz برابر با  $\frac{2}{5}$  ثانیه است. این زمان برای کاربردهای برخط ایده آل نیست؛ هر چند با توجه به دقت به دست آمده و پیچیدگی الگوریتم در تشخیص بخش ها و استفاده از آن ها برای طبقه بندی، زمان بسیار مناسبی است. این زمان با موازی سازی بخش های مختلفی از سامانه حاصل شده است.

برای سنجش عملکرد روش پیشنهادی، چند پیاده سازی از روش های پرکاربردی که توسط کارهای پیشین استفاده شده اند، ارائه کرده ایم. این روش ها از ناحیه مطلوب استفاده می کنند. از این رو، محدوده خودرو را به دو نیم تقسیم کرده و نیمی پایینی را به عنوان ناحیه مطلوب در نظر گرفته ایم. برای عادلانه بودن مقایسه ها، از یک روش استخراج ویژگی مشترک با پارامترهای یکسان بهره برده ایم؛ یعنی توصیف گر هیستوگرام گرادیان های جهت دار با اندازه سلول  $8 \times 8$  و اندازه بلوک  $2 \times 2$ . از ماشین بردار پشتیبان نیز برای طبقه بندی استفاده کرده ایم (الگوریتم C\_SVM در کتابخانه LIBSVM [40]).

تعداد نمونه های طبقه های مختلف در مجموعه داده تهیه شده متوازن نیستند. برای مثال طبقه "پراید ۱۳۱" دارای ۲۰۰۱ نمونه آزمایشی و طبقه "لیفان ۶۲۰" دارای تنها بیست نمونه آزمایشی است. اگر در چنین حالتی از یک ماشین بردار پشتیبان استاندارد استفاده کنیم، نتیجه مطلوبی به دست نخواهد آمد؛ زیرا به احتمال زیاد طبقه بند همه نمونه ها را تحت عنوان طبقه "پراید ۱۳۱" طبقه بندی و دقت مناسبی نیز گزارش خواهد کرد. برای رفع این مشکل از ماشین بردار پشتیبان وزن دار استفاده می کنیم؛ با دادن وزن بیشتر به طبقه های دارای نمونه های کمتر، ارزش طبقه بندی درست آن ها را بالاتر خواهیم برد. اگر تعداد کل نمونه های آموزشی را برابر با  $N$  و تعداد نمونه های آموزشی طبقه  $i$  را برابر با  $N_i$  فرض کنیم؛ برای برابر کردن ارزش همه طبقه ها، از وزن  $\frac{N}{N_i}$  برای طبقه  $i$  استفاده کرده ایم.

جدول (۶) نتایج به دست آمده را ارائه کرده است. برای ماشین های بردار پشتیبان، طبقه "سایر" را کنار گذاشته ایم. روش SINGLE\_SVM\_1\_1 تنها از یک ماشین بردار

بخش‌های متمایزکننده طبقه مربوطه را تشخیص داده و با استفاده از آن‌ها، نوع و مدل خودرو را تعیین می‌کند. برای بررسی دقت سامانه پیشنهادی و به دلیل عدم وجود مجموعه داده به اشتراک گذاشته شده، مجموعه داده‌ای با بیش از پنج هزار خودرو تهیه و به صورت کامل حاشیه‌نویسی شده است. عملکرد مناسب سامانه روی این مجموعه داده نشان از قدرت بالای آن و مقاومت نسبت به تغییرات روشنایی زیاد و تغییرات زاویه اندک تصاویر دارد.

یکی از ضعف‌های سامانه پیشنهادی، زمان اجرای آن است؛ با ترکیب مدل‌های هر طبقه در یک مدل واحد و عدم تکرار برخی از پردازش‌های مشترک بین مدل‌ها، کاهش چشم‌گیری در زمان کل سامانه می‌توان ایجاد کرد. علاوه بر این، استفاده از حد آستانه در هر سامانه‌ای، تقابلی بین دقت و یادآوری<sup>۲</sup> ایجاد می‌کند. علت اختلاف زیاد بین دقت گزارش شده برای رتبه نخست و سه رتبه نخست نیز همین موضوع است. با بهبود شکل استفاده از حد آستانه دقت سامانه را بهتر از پیش می‌توان کرد.

## 6-References

## ۶- مراجع

- [1] Sun, Z., George, B., Ronald, M. "On-Road Vehicle Detection: A Review.", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 694-711, 2006.
- [2] Li, X., Guo, X. "A HOG Feature and SVM Based Method for Forward Vehicle Detection with Single Cam-era.", *5th International Conference on Intelligent Human-Machine Systems and Cybernetics*, pp. 263-266, 2013.
- [3] Yousaf, K., Ifthikhar, A., Javed, A. "Comparative Analysis of Automatic Vehicle Classification Techniques: A Survey.", *International Journal of Image, Graphics and Signal Processing*, vol. 4, no. 9, pp. 52, 2012.
- [4] Ambardekar, A., Nicolescu, M., Bebis, G., Nicolescu, M. "Vehicle Classification Framework: A Comparative Study.", *EURASIP Journal on Image and Video Processing*, vol. 2014, no. 1, pp. 1-13, 2014.
- [5] Zhang, B. "Reliable Classification of Vehicle Types Based on Cascade Classifier Ensembles.", *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 1, pp. 322-332, 2013.
- [6] Dong, Z., Jia, Y. "Vehicle Type Classification Using Distributions of Structural and Appearance-Based Features.", *IEEE International Conference on Image Processing*, pp. 4321-4324, 2013.

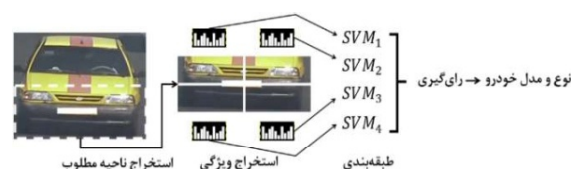
<sup>1</sup>Precision

<sup>2</sup>Recall

پشتیبان به صورت یک در برابر همه استفاده می‌کند. 1\_1 یعنی از کل ناحیه مطلوب به صورت یکجا استفاده شده است (تعداد سطر و ستون). روش MULTI\_SVM\_2\_3 ناحیه مطلوب را به دو سطر و سه ستون تقسیم کرده و از هر بخش، ویژگی استخراج می‌کند؛ آنگاه یک ماشین بردار پشتیبان برای هر بخش آموزش داده و برای تعیین خروجی، رأی‌گیری می‌کند. شکل (۹) مراحل عملکرد روش MULTI\_SVM\_2\_2 را به عنوان یک نمونه نمایش داده است.

(جدول-۶): مقایسه نتایج پیشنهادی با چند روش مبتنی بر بخش  
(Table-6): The comparison of the proposed system and some other part-based methods

روش	دقت میانگین (%)
SINGLE_SVM_1_1	۹۱/۲۸
MULTI_SVM_2_2	۸۸/۸۰
MULTI_SVM_2_3	۸۹/۱۵
MULTI_SVM_2_4	۹۰/۵۲
روش پیشنهادی	۹۷/۰۷



(شکل-۹): مراحل عملکرد روش MULTI\_SVM\_2\_2  
(Figure-9): The flowchart of MULTI\_SVM\_2\_2 method

نتایج جدول (۶) دربرگیرنده دو نکته است. نکته نخست، برتری روش پیشنهادی با اختلاف زیاد نسبت به سایر روش‌های ارائه شده در جدول است. این اختلاف، تأثیر استفاده از بخش‌های متفاوت برای هر دسته از خودروها را به روشنی نشان می‌دهد. استفاده از بخش‌های مشابه در طبقه‌های شبیه به هم به دقت بالایی نمی‌تواند دست یابد. نکته دوم، عملکرد بهتر ماشین بردار پشتیبان در حالت تک‌بخشی نسبت به چندبخشی است. استفاده ترکیبی از چند طبقه‌بند ضعیف‌تر که هر یک تنها روی یک بخش از ناحیه مطلوب آموزش دیده‌اند، منجر به دقت کمتری شده است. برای رسیدن به دقتی بالاتر در این حالت، انتخاب مناسب بخش‌ها بسیار مهم است.

## ۵- نتیجه‌گیری و کارهای آینده

در این مقاله روشی جدید برای تشخیص نوع و مدل خودرو ارائه شده شد. این روش برای هر طبقه از خودروها یک مدل منحصربه‌فرد یاد می‌گیرد. این مدل به صورت خودکار،

- [20] Hsieh, J.-W., Chen, L.-C., Chen, D.-Y. "Symmetrical SURF and Its Applications to Vehicle Detection and Vehicle Make and Model Recognition.", *IEEE Transactions on Intelligent Transportation Systems.*, vol. 15, no. 1, pp. 6–20, 2014.
- [21] Nazemi, A., Shafiee, M., Azimifar, Z. "On Road Vehicle Make and Model Recognition via Sparse Feature Coding.", *8th Iranian Conference on Machine Vision and Image Processing*, pp. 436–440, 2013.
- [22] Baran, R., Glowacz, A., Matiolanski, A. "The Efficient Real- and Non-Real-Time Make and Model Recognition of Cars.", *Multimedia Tools and Applications.*, no. June, pp. 1–20, 2013.
- [23] Gao, Y., Lee, H.J. "Moving Car Detection and Model Recognition Based on Deep Learning.", *Advanced Science and Technology Letters.*, vol. 90, no. Multimedia, pp. 57–61, 2015.
- [24] Siddiqui, A.J.A.M., Boukerche, A. "Towards Efficient Vehicle Classification in Intelligent Transportation Systems.", *Proceedings of the 5th ACM Symposium on Development and Analysis of Intelligent Vehicular Networks and Applications*, pp. 19–25, 2015.
- [25] Psyllos, A. "Vehicle Logo Recognition Using a SIFT-Based Enhanced Matching Scheme.", *Intelligent Transportation Systems*, *IEEE Transactions on.*, vol. 11, no. 2, pp. 322–328, 2010.
- [26] Yang, H., Zhai, L., Liu, Z., Li, L., Luo, Y., Wang, Y., Lai, H., Guan, M. "An Efficient Method for Vehicle Model Identification via Logo Recognition.", *International Conference on Computational and Information Sciences*, pp. 1080–1083, 2013.
- [27] Santos, D., Correia, P.L. "Car Recognition Based on Back Lights and Rear View Features.", *10th International Workshop on Image Analysis for Multimedia Interactive Services*, pp. 137–140, 2009.
- [28] Sarfraz, M.S., Saeed, A., Khan, M.H., Riaz, Z. "Bayesian Prior Models for Vehicle Make and Model Recognition.", *Proceedings of the 6th International Conference on Frontiers of Information Technology - FIT '09*, p. 6, ACM Press, New York, New York, USA, 2009.
- [29] Psyllos, A., Anagnostopoulos, C.N., Kayafas, E., Loumos, V. "Image Processing & Artificial Neural Networks for Vehicle Make and Model Recognition.", *10th international conference on applications of advanced technologies in transportation*, pp. 4229–4243, 2008.
- [7] Choo, S., Mokhtarian, P.L. "What Type of Vehicle Do People Drive? The Role of Attitude and Lifestyle in Influencing Vehicle Type Choice.", *Transportation Research Part A: Policy and Practice.*, vol. 38, no. 3, pp. 201–222, 2004.
- [8] Conos, M. "Recognition of Vehicle Make from a Frontal View.", Master Thesis, Czech Tech, 2007.
- [9] Dlagnekov, L. "Video-Based Car Surveillance: License Plate, Make, and Model Recognition.", Master Thesis, University of California, San Diego, 2005.
- [10] Negri, P., Clady, X., Milgram, M., Poulenard, R. "An Oriented-Contour Point Based Voting Algorithm for Vehicle Type Classification.", *18th International Conference on Pattern Recognition*, pp. 574–577, 2006.
- [11] Clady, X., Negri, P., Milgram, M., Poulenard, R. "Multi-Class Vehicle Type Recognition System.", *Artificial Neural Networks in Pattern Recognition, Lecture Notes in Computer Science*, pp. 228–239, Springer Berlin Heidelberg, 2008.
- [12] Huang, H., Zhao, Q., Jia, Y., Tang, S. "A 2DLDA Based Algorithm for Real Time Vehicle Type Recognition.", *11th International IEEE Conference on Intelligent Transportation Systems*, pp. 298–303, 2008.
- [13] Pearce, G., Pears, N. "Automatic Make and Model Recognition from Frontal Images of Cars.", *8th IEEE International Conference on Advanced Video and Signal Based Surveillance*, pp. 373–378, 2011.
- [14] Saravi, S., Edirisinghe, E. a. "Vehicle Make and Model Recognition in CCTV Footage.", *18th International Conference on Digital Signal Processing*, pp. 1–6, 2013.
- [15] Lowe, D.G. "Object Recognition from Local Scale-Invariant Features.", *17th IEEE International Conference on Computer Vision*, pp. 1150–1157, 1999.
- [16] Petrovic, V., Cootes, T. "Analysis of Features for Rigid Structure Vehicle Type Recognition.", *British Machine Vision Conference*, pp. 587–596, 2004.
- [17] Munroe, D.T., Madden, M.G. "Multi-Class and Single-Class Classification Approaches to Vehicle Model Recognition from Images.", *AICS '05*, pp. 93–102, 2005.
- [18] Kazemi, F.M., Samadi, S., Poorreza, H.R., Akbarzadeh-T, M.-R. "Vehicle Recognition Using Curvelet Transform and SVM.", *4th International Conference on Information Technology*, pp. 516–521, 2007.
- [19] Zafar, I., Edirisinghe, E. a., Acar, B.S. "Localised Contourlet Features in Vehicle Make and Model Recognition.", *Image Processing: Machine Vision*



در حال حاضر نیز، فارغ‌التحصیل مقطع دکترا از دانشگاه صنعتی شاهرود در گرایش هوش مصنوعی است. موضوع پایان‌نامه کارشناسی ارشد ایشان، "تشخیص و شناسایی چهره با استفاده از الگوریتم‌های موازی، توسط کودا" و موضوع پایان‌نامه دکترای او، "شناسایی نوع و مدل خودرو با استفاده از مدل‌های مبتنی بر بخش" است. علایق پژوهشی وی برنامه‌نویسی موازی و پردازش تصویر هستند.

mbt925@gmail.com



**علی سلیمانی** کارشناسی ارشد و دکترای خود را در رشته مهندسی برق- الکترونیک در سال‌های ۱۳۷۳ و ۱۳۷۹ از دانشگاه علم و صنعت ایران دریافت کرد. پروژه کارشناسی ارشد خود را در زمینه پردازش سیگنال با منطق فازی و پروژه دکترای خود را در زمینه کنترل سامانه‌های چندرودوی چندخروجی با استفاده از شبکه‌های عصبی فازی به انجام رسانید. از سال ۱۳۸۰ ایشان عضو هیئت علمی دانشگاه صنعتی شاهرود بوده و زمینه پژوهشی ایشان پردازش سیگنال‌های دیجیتال، شبکه‌های عصبی فازی و الگوریتم‌های تکاملی است.

solimani\_ali@shahroodut.ac.ir



**حمید حسن‌پور** استاد تمام دانشکده مهندسی کامپیوتر دانشگاه شاهرود می‌باشند. ایشان در سال ۱۳۷۲ مدرک کارشناسی مهندسی کامپیوتر خود را از دانشگاه علم و صنعت و در سال ۱۳۷۵

مدرک کارشناسی ارشد خود را در گرایش هوش ماشین از دانشگاه صنعتی امیرکبیر دریافت کرد. در سال ۱۳۸۳ موفق به اخذ مدرک دکترای خود از دانشگاه صنعتی کوئینزلند استرالیا در گرایش پردازش سیگنال شد. از سال ۱۳۸۴ الی ۱۳۸۶ نامبرده به‌عنوان عضو هیئت علمی در دانشکده مهندسی برق و کامپیوتر دانشگاه صنعتی بابل فعالیت داشت؛ سپس به دانشکده مهندسی کامپیوتر دانشگاه شاهرود انتقال یافت. زمینه‌های علمی مورد علاقه ایشان پردازش سیگنال، پردازش تصویر، داده‌کاوی، و پردازش متن است.

نشانی رایانامه ایشان عبارت است از:

h.hassanpour@shahroodut.ac.ir

- [30] Llorca, D., Colas, D., Daza, I. "Vehicle Model Recognition Using Geometry and Appearance of Car Emblems from Rear View Images.", *17th IEEE International Conference on Intelligent Transportation Systems*, pp. 3094–3099, 2014.
- [31] Dalal, N., Triggs, B. "Histograms of Oriented Gradients for Human Detection.", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 886–893, 2005.
- [32] Felzenszwalb, P.F., Girshick, R.B., McAllester, D., Ramanan, D. "Object Detection with Discriminatively Trained Part-Based Models.", *IEEE transactions on pattern analysis and machine intelligence.*, vol. 32, no. 9, pp. 1627–45, 2010.
- [33] Lampert, C., Nickisch, H., Harmeling, S. "Attribute-Based Classification for Zero-Shot Learning of Object Categories.", *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, vol. 36, no. 3, pp. 453 – 465, 2014.
- [34] Felzenszwalb, P.F., Huttenlocher, D.P. "Pictorial Structures for Object Recognition.", *International Journal of Computer Vision.*, vol. 61, no. 1, pp. 55–79, 2005.
- [35] Fergus, R., Perona, P., Zisserman, A. "Object Class Recognition by Unsupervised Scale-Invariant Learning.", *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 264–271, 2003.
- [36] Weber, M., Welling, M., Perona, P. "Towards Automatic Discovery of Object Categories.", *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 101–108, 2000.
- [37] Yang, L., Luo, P., Loy, C.C., Tang, X. "A Large-Scale Car Dataset for Fine-Grained Categorization and Verification.", *Proc. IEEE Conference on Computer Vision and Pattern Recognition.*, vol. 1, pp. 3973–3981, 2015.
- [38] "NTOU-MMR Dataset," <http://mmplab.cs.ntou.edu.tw/mmplab/MMR/MMR.html> (Accessed: 8 July 2016).
- [39] "The PASCAL Visual Object Classes," [Online] 2008, <http://pascallin.eecs.soton.ac.uk/challenges/VOC/> (Accessed: 10 March 2015).
- [40] Chang, C., Lin, C. "LIBSVM: A Library for Support Vector Machines.", *ACM Transactions on Intelligent Systems and Technology.*, vol. 2, no. 3, pp. 1–27, 2011.



**محسن بیگلری** مدرک کارشناسی خود را در گرایش نرم‌افزار از دانشگاه کاشان در سال ۸۹ دریافت کرد؛ سپس در سال ۹۱ موفق به اخذ مدرک کارشناسی ارشد در گرایش نرم‌افزار از همان دانشگاه شد.



