
بسم الله الرحمن الرحيم

دانشکده مهندسی کامپیوتر و فناوری اطلاعات

گروه مهندسی کامپیوتر

رساله دکتری

شناسایی نوع و مدل وسیله نقلیه با استفاده از مدل‌های مبتنی بر بخش

نگارنده: محسن بیگلری

استاد راهنما

دکتر علی سلیمانی

استاد مشاور

دکتر حمید حسن‌پور

اردیبهشت ۱۳۹۶

شماره: ۷۴۸/ف/ک
تاریخ: ۹۹/۵/۲۱
ویرایش:

باسمه تعالی

مدیریت تحصیلات تکمیلی

فرم شماره ۱۲: صورت جلسه نهایی دفاع از رساله دکتری (Ph.D)

(ویژه دانشجویان ورودی های ۹۴ و ما قبل)

بدینوسیله گواهی می شود آقای **محسن بیگلری** دانشجوی دکتری رشته **مهندسی کامپیوتر / هوش مصنوعی** به شماره دانشجویی **۹۱۳۵۱۷۵** ورودی **مهر** ماه سال **۱۳۹۱** در تاریخ ۱۳۹۶/۰۴/۱۲ از رساله **نظری**☑ / **عملی**☐ خود با عنوان : شناسایی نوع و مدل وسیله نقلیه با استفاده از مدل‌های مبتنی بر بخش
دفاع و با اخذ نمره ۱۹ به درجه : عالی نائل گردید.

الف) درجه عالی: نمره ۲۰-۱۹ ☑	ب) درجه بسیار خوب: نمره ۱۸/۹۹ – ۱۷☐
ج) درجه خوب: نمره۱۶/۹۹– ۱۵☐	د) غیر قابل قبول و نیاز به دفاع مجدد دارد☐
ه) رساله نیاز به اصلاحات دارد☐	

ردیف	هیئت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱	دکتر سید علی سلیمانی	استاد/ اساتید راهنما	دانشیار	
۲	دکتر حمید حسن‌پور	مشاور/ مشاورین	استاد	
۳	دکتر اسدا... شاه‌بهرامی	استاد مدعو خارجی	دانشیار	
۴	دکتر علیرضا احمدی‌فرد	استاد مدعو داخلی	دانشیار	
۵	دکتر مرتضی زاهدی	استاد مدعو داخلی	دانشیار	
۶	دکتر وحید ابوالقاسمی	سرپرست (نماینده) تحصیلات تکمیلی دانشکده	استادیار	

مدیر محترم تحصیلات تکمیلی دانشگاه:

ضمن تأیید مراتب فوق مقرر فرمائید اقدامات لازم در خصوص انجام مراحل دانش آموختگی آقای محسن بیگلری بعمل آید.

نام و نام خانوادگی رئیس دانشکده :

تاریخ و امضاء و مهر دانشکده:

تقدیم به

همسر عزیزم،

که بدون فداکاری‌های او، انجام این رساله غیرممکن بود.

دخترم،

که با لبخندهایش، همه خستگی‌هایم را به فراموشی می‌سپرد.

پدر و مادر و پدر و مادر همسرم،

که زحمت‌ها و پشتیبانی‌هایشان مایه دلگرمی بود.

تشکر و قدردانی

سپاس خدایی را سزاست که به او تکیه می‌کنم و او گرامی‌ام می‌دارد و دست نوازش بر سرم می‌کشد و به مردم تکیه نمی‌کنم که اگر کنم خوارم می‌کنند و تنهایم می‌گذارند.

سپاس خدایی را سزاست که هر وقت بخواهم می‌توانم صدایش کنم و هرگاه در پی خلوتی با او باشم بی‌هیچ واسطه‌ای می‌توانم داشته باشم و او همیشه برآورنده خواسته‌های من است.

«فرازی از دعای امام زین العابدین (ع) در شب‌های ماه مبارک رمضان»

از زحمات بی‌دریغ استاد راهنمای ارجمندم، جناب آقای دکتر سلیمانی که در طول دوره دکتری، هیچ کمکی را از بنده دریغ نکردند، نهایت قدردانی را دارم. بی‌شک اگر حمایت‌های دلسوزانه ایشان نبود، به پایان رساندن این مهم به سادگی امکان‌پذیر نبود.

از جناب آقای دکتر حمید حسن‌پور که مشاوره این پایان‌نامه را بر عهده داشتند، بسیار سپاسگزارم. از تمامی دفعاتی که با دقت تمام برای بهبود کیفیت کار اینجانب، وسواس به خرج دادند، قدردانی می‌کنم.

از مرکز کنترل هوشمند ترافیک استان قم که در تهیه تصویر از دوربین‌های مداربسته با بنده همکاری کردند نیز کمال تشکر را دارم.

از همسر و فرزند عزیزم که در این مدت، سختی‌های زیادی را تحمل کرده‌اند، تشکر می‌کنم. انشاالله بتوانم بخشی از این فداکاری‌ها را جبران کنم.

همچنین از تمامی اعضای خانواده‌ام که در تمامی مراحل زندگی مرا حمایت کردند، سپاسگزارم. امیداورم بتوانم همیشه قدردان زحمت‌هایشان باشم.

تعهد نامه

اینجانب **محسن بیگلری** دانشجوی دوره دکتری رشته **مهندسی کامپیوتر، گرایش هوش مصنوعی** دانشکده **مهندسی کامپیوتر و فناوری اطلاعات** دانشگاه صنعتی شاهرود نویسنده پایان نامه **شناسایی نوع و مدل وسیله نقلیه با استفاده از مدل‌های مبتنی بر بخش** تحت راهنمائی **آقای دکتر علی سلیمانی** متعهد می شوم .

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است.

تاریخ ۱۳۹۶/۰۴/۲۵

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه ای ، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

چکیده

در شناسایی دانه‌ریز اشیاء (Fine-grained Recognition)، دسته کلی شیء مشخص بوده و هدف، شناسایی زیردسته‌ی دقیق یا دانه‌ریز شیء می‌باشد. شناسایی نوع و مدل وسیله نقلیه به عنوان یک مسئله شناسایی دانه‌ریز دشوار شناخته می‌شود. این دشواری به دلیل تعداد کلاس‌های زیاد، شباهت برون‌کلاسی بسیار کم و تفاوت درون‌کلاسی بسیار زیاد در این مسئله است. این موضوع در مواردی چون سیستم‌های نظارتی و امنیتی، پارکینگ هوشمند، دریافت عوارض هوشمند و تشخیص تقلب کاربرد دارد.

در این رساله، رویکردی مبتنی بر بخش برای شناسایی نوع و مدل وسیله نقلیه ارائه گشته که سعی بر غلبه بر چالش‌های جدی موجود دارد. این رویکرد برای هر دسته از وسایل نقلیه یک مجموعه بخش متمایزکننده یاد می‌گیرد. رویکرد پیشنهادی برای رسیدن به یک طبقه‌بند کارا از یک روال آموزش چند کلاسه بهره می‌برد که مراحل موردنیاز برای یادگیری مجموعه بخش‌های متمایزکننده، رابطه هندسی بین آن‌ها و داده‌کاوی روی نمونه‌های منفی را در خود دارد. برای آموزش سیستم و ارزیابی کارایی آن، دو مجموعه داده با ویژگی‌های متفاوت تهیه شده‌اند. مجموعه داده اول (BVMMR v1) دارای ۷۲۰ تصویر و مجموعه داده دوم (BVMMR v2) دارای ۵۰۰۰ تصویر می‌باشند. اخیراً نیز یک مجموعه داده بسیار بزرگ به نام CompCars با ۵۰۰۰۰ تصویر از بیش از ۲۰۰ مدل متفاوت از وسایل نقلیه ارائه شده است. دقت کسب شده توسط سیستم بر روی هر دو مجموعه داده BVMMR v2 و CompCars بیش از ٪۹۶ بوده است که نشان از کارایی مناسب روش پیشنهادی دارد.

سرعت سیستم پیشنهادی با افزایش تعداد طبقه‌بندها کاهش می‌یابد. علاوه‌براین، در مواقعی که از یک مجموعه داده نامتوازن برای آموزش سیستم استفاده می‌شود، احتمال رخداد بیش‌برازش و کاهش قدرت عمومیت طبقه‌بندها بیشتر می‌شود. برای حل این مشکل و بهبود سرعت سیستم، یک الگوریتم آبشاری پیشنهاد شده است. الگوریتم پیشنهادی با درنظر گرفتن دو معیار «اطمینان» و «فراوانی» طبقه‌بند، یک آبشار از طبقه‌بندها تشکیل می‌دهد. ارزیابی این الگوریتم بر روی دو مجموعه داده BVMMR v2 و CompCars نشان از تأثیرگذاری مناسب این روش داشته است. به صورتی که بدون کاهش دقت، سرعت سیستم تا ٪۳۰ روی این دو مجموعه داده افزایش پیدا کرده است. با چشم‌پوشی از کاهش دقت یک درصدی، سرعت سیستم تا ٪۸۰ نیز قابل افزایش است.

کلمات کلیدی: شناسایی دانه‌ریز اشیاء، شناسایی نوع و مدل وسیله نقلیه، VMMR، شناسایی مبتنی بر بخش، ماشین بردار پشتیبان مخفی

مقالات مستخرج از رساله

مقالات داخلی

ردیف	مقاله	وضعیت	تاریخ
۱	محسن بیگلری، علی سلیمانی، "مقایسه روش‌های شناسایی نوع و مدل وسیله نقلیه با رویکرد کلی‌نگر و جزئی‌نگر و ارائه یک رویکرد جدید"، مجله علمی-ترویجی محاسبات نرم	پذیرفته شده	فروردین ۹۶
۲	محسن بیگلری، علی سلیمانی، حمید حسن‌پور، "شناسایی نوع و مدل وسیله نقلیه با استخراج خودکار بخش‌های مشترک"، مجله علمی-پژوهشی ماشین بینایی و پردازش تصویر	چاپ شده	تیر ۹۶
۳	محسن بیگلری، علی سلیمانی، حمید حسن‌پور، "شناسایی نوع و مدل وسیله نقلیه با استفاده از مجموعه بخش‌های متمایزکننده"، مجله علمی-پژوهشی پردازش علائم و داده‌ها	پذیرفته شده	خرداد ۹۶
۴	محسن بیگلری، علی سلیمانی، حمید حسن‌پور، "ارائه یک سیستم آبشاری مبتنی بر بخش برای شناسایی نوع و مدل وسیله نقلیه"، مجله علمی-پژوهشی رایانش نرم و فناوری اطلاعات	در حال داوری	شهریور ۹۵
۵	محسن بیگلری، علی سلیمانی، حمید حسن‌پور، "بهبود سرعت و دقت یک سیستم شناسایی مدل خودرو با ارائه یک الگوریتم آبشاری"، مجله علمی-پژوهشی جهاد دانشگاهی	در حال داوری	مهر ۹۵

مقالات خارجی

Row	Paper	Status	Date
1	Mohsen Biglari, Ali Soleimani, Hamid Hassanpour, “Part-Based Recognition of Vehicle Make and Model”, IET Image Processing	Published	March 2017
2	Mohsen Biglari, Ali Soleimani, Hamid Hassanpour, “A Cascaded Part-based System for Fine-grained Vehicle Classification”, IEEE Transactions on Intelligent Transportation Systems	Accepted	July 2017
3	Mohsen Biglari, Hamid Hassanpour, Ali Soleimani, “A Cascading Scheme for Speeding up Multiple Classifier Systems”, Pattern Analysis and Applications	Under Review	December 2016

مقالات کنفرانسی

ردیف	مقاله	وضعیت	تاریخ
۱	محسن بیگلری، علی سلیمانی، حمید حسن‌پور، " شناسایی نوع و مدل وسیله نقلیه با ترکیبی از مدل‌های مبتنی بر بخش"، کنفرانس پردازش سیگنال و سیستم‌های هوشمند	پذیرفته و ارائه شده	آذر ۹۵
۲	محسن بیگلری، علی سلیمانی، حمید حسن‌پور، " Vehicle Make and Model Recognition using Auto Extracted Parts"، دومین همایش بین‌المللی حمل و نقل هوشمند	پذیرفته و ارائه شده	بهمن ۹۵

فهرست مطالب

فهرست شکل‌ها

فهرست جدول‌ها

فهرست الگوریتم‌ها

علائم اختصاری

BOW	Bag Of Words
CCTV	Closed-Circuit Television Camera
DoG	Difference of Gaussian
DCT	Discrete Cosine Transform
HOG	Histogram of Oriented Gradients
k-NN	k-Nearest Neighbors
LSVM	Latent SVM
LDA	Linear Discriminant Analysis
LESH	Local Energy Shape Histogram
QR Code	Quick Response Code
SPC	Number of Samples per Class
PCFM	Phase Congruency Feature Map
ROI	Region of Interest
SIFT	Scale-Invariant Feature Transform
SURF	Speeded Up Robust Features
SVM	Support Vector Machine
VMMR	Vehicle Make and Model Recognition

واژه‌نامه

0-9	
2D-LDA	تحلیل افتراقی خطی دوبعدی
3-Fold Cross Validation	ارزیابی ضربدری به پیمانه سه
A	
Accuracy	دقت
Anchor	لنگر
Annotation	حاشیه‌نویسی
B	
Bag of Words	کیسه‌ای از کلمات
Bag of Visual Words	کیسه‌ای از کلمات تصویری
Bias	بایاس
Blur	تار
C	
Calibration	کالیبره کردن
Classification	طبقه‌بندی
Closed-Circuit Television Camera	دوربین‌های مداربسته
Confusion Matrix	ماتریس درهم‌ریختگی
Contourlet Transform	تبدیل کانتورلت
Convex	محدب
Cosine	کسینوسی
Curvelet Transform	تبدیل پیچک
D	
Decision Tree	درخت تصمیم
Developmental Kit	بسته توسعه نرم‌افزاری
Discrete Cosine Transform	تبدیل کسینوسی گسسته
E	
Edge Histogram	هیستوگرام لبه
Edge Orientation	جهت لبه‌ها
Emblems	علائم (نوشته‌ها)
Entropy	بی‌نظمی
F	
Feature Map	نقشه ویژگی
Fit	برازش
G	
Gaussian	گوسی
Generalization	تعمیم
Grayscale	خاکستری
Grilles	جلوپنجره
H	
Harris Corner Detection	تشخیص گوشه‌ی هریس
Hinge Loss	خطای لولایی
Histogram of Oriented Gradients	هیستوگرام گرادیان‌های جهت دار
Homography Matrix	ماتریس هوموگرافی
I	
Image Pyramids	هرم تصاویر
Infrared	مادون قرمز
K	
K-L Divergence	واگرایی K-L
k-Nearest Neighbors	k-نزدیک‌ترین همسایگی‌ها

L

Latent SVM	ماشین بردار پشتیبان مخفی
L^2 Norm	نرم۲
Leave-one-out Cross Validation	ارزیابی ضربدری تک حذفی
Linear Discriminant Analysis	تحلیل افتراقی خطی
Local Energy Shape Histogram	هیستوگرام شکل انرژی محلی
Logo	نشان‌واره

M

Max-Voting	رای‌گیری به روش بزرگترین امتیاز
Morphological Operators	عملگرهای مورفولوژیکی
Multi-Layer Perceptron	شبکه عصبی پرسپترون چندلایه

N

Naïve Bayes	بیز ساده

O

Object Class Recognition	شناسایی کلاسِ شیء
Offline	خارجِ خط
One Against All	یک در برابر همه
Online	برخط
Overfitting	بیش‌برازش

P

Phase Congruency Feature Map	نگاشت ویژگی تناسب فاز
Pictorial Structure	ساختار تصویری
Platt Scaling	مقیاس‌گذاری پلت
Posterior Probability	احتمال پسین
Precision	صحت

Q

Quantize	تقطیع

R

Rank-3, Top 3	سه رتبه اول
Recall	بازیابی
Region of Interest	ناحیه مطلوب
Resolution	وضوح
Rigid	سخت

S

Scale-Invariant Feature Transform	تبدیل ویژگی مقاوم به اندازه
Segmentation	قطعه‌بندی
Sliding Windows	پنجره لغزان
Sobel Edge Response	پاسخ لبه سوبل
Soft-Margin	حاشیه نرم
Sparse Feature Coding	کدگذاری ویژگی پراکنده
Speeded Up Robust Features	ویژگی‌های مقاوم تسریع شده
Square Mapped Gradients	گرادیان‌های نگاشت شده‌ی مربعی
Star Model	مدل ستاره‌ای
State of the Art	مبنای امروزی
Support Vector Machine	ماشین بردار پشتیبان

T

Topological Sort	مرتب‌سازی توپولوژیکی
Tracking	ردگیری
Transformation	دگرسازی

U

Upsampling	نمونه‌برداری افزایشی

V

Validation Set	مجموعه داده اعتبارسنجی
Vehicle Detection	تشخیص وسیله نقلیه
Vehicle Make and Model Recognition	شناسایی نوع و مدل وسیله نقلیه
Vehicle Type Recognition	شناسایی دسته کلی وسلیه نقلیه

W

Wavelet Transform	تبدیل موجک

فصل

۱- مقدمه

در دهه جاری، استفاده از پردازش تصویر و بینایی ماشین در کاربردهای نظارتی و امنیتی به بلوغ مناسبی رسیده است؛ چنان‌که، سیستم‌های تجاری بسیاری با بکارگیری از الگوریتم‌های پردازش تصویر در حال استفاده در دنیای واقعی هستند. تشخیص اثر انگشت، شناسایی چهره و تشخیص پلاک تنها چند نمونه از این موارد می‌باشند.

در حال حاضر، اغلب سیستم‌های نظارتی و امنیتی (در حوزه حمل و نقل هوشمند) با استفاده از پلاک وسایل نقلیه عمل می‌کنند. تشخیص پلاک در سال‌های اخیر پیشرفت‌های بسیاری داشته و به دقت قابل قبولی رسیده است [۱–۴]. با این حال، جعل پلاک و مسدود کردن پلاک به روش‌های مختلف، از تکنیک‌های بسیار معمول برای منحرف کردن سیستم‌های مبتنی بر پلاک است؛ آن‌چنان‌که می‌بینیم، خودرویی بدون دلیل در یک مکان متفاوت از محل صدور جریمه و با مشخصات متفاوت، به دلیل تخلف رانندگی جریمه می‌شود. درصورتی‌که بتوان سیستمی طراحی کرد که قادر به شناسایی نوع و مدل دقیق وسیله نقلیه از روی تصویر آن باشد؛ می‌توان با تلفیق این دو سیستم، گستره‌ی جدیدی از قابلیت‌های کاربردی را تعریف کرد.

۱–۱– تعریف مسئله و اهداف

سیستم‌های تشخیص و شناسایی وسایل نقلیه را می‌توان به سه دستی کلی زیر تقسیم‌بندی کرد:

۱. **تشخیص وسیله نقلیه**[۱]: هدف در این سیستم‌ها، تشخیص مکان و محدوده‌ی وسایل نقلیه در تصویر است. روش‌های زیادی در این رابطه ارائه گشته و دقت‌های قابل قبولی نیز به دست آورده‌اند [۵–۹].
۲. **شناسایی دسته‌ی کلی وسیله نقلیه**[۲]: هدف در این سیستم‌ها، یافتن دسته‌ی کلی وسایل نقلیه است. یک نمونه دسته‌بندی کلی از وسایل نقلیه عبارتست از: وسایل نقلیه سنگین (کامیون، اتوبوس و مینی‌بوس)، سواری‌ها (تاکسی و وسیله نقلیه شخصی) و وسایل نقلیه سبک (موتور و دوچرخه). روش‌های متعددی برای این منظور نیز ارائه شده است [۱۰–۱۵].
۳. **شناسایی نوع و مدل وسیله نقلیه**[۳]: تلاش این سیستم‌ها، یافتن نوع و مدل دقیق وسیله نقلیه با استفاده از نشانه‌های بصری است. برای مثال خودروی پراید ۱۳۱ و یا پژو ۲۰۶.

---

[۱] Vehicle Detection
[۲] Vehicle Type Recognition
[۳] Vehicle Make and Model Recognition (VMMR)

مسئله‌ی هدف در این پایان‌نامه، دسته‌ی سوم یعنی شناسایی نوع و مدل وسیله نقلیه می‌باشد. این مسئله، یک مسئله‌ی شناسایی کلاسِ شیء[1] درون‌گروهی است؛ به شکلی که فرض می‌شود گروه کلی شیء مشخص است (وسیله‌ی نقلیه) و هدف یافتن کلاس دقیق شیء در گروه مربوطه می‌باشد. این مسئله نسبت به شناسایی شیء عمومی که در آن تنها گروه کلی اشیاء تعیین می‌گردد، چالش‌برانگیزتر است؛ زیرا شباهت کلاس‌ها در درون گروه بسیار اندک بوده و دارای ویژگی‌هایی هستند که بین همه‌ی کلاس‌ها مشترک است. از طرفی کلاس‌های مشابه به سادگی به دلیل تغییرات روشنایی، رنگ و زاویه اشتباه گرفته می‌شوند. درنتیجه، روش‌های معمول برای شناسایی اشیاء[2] ممکن است جوابگوی این مسائل نباشد.

در یک سیستم شناسایی نوع و مدل خودرو، فرض می‌شود که مکان و محدوده وسایل نقلیه موجود در تصویر تشخیص داده شده است. شناسایی نوع و مدل وسیله نقلیه موضوعی نوپا است که در دهه‌ی اخیر مورد توجه قرار گرفته است؛ ازاین‌رو، کارهای انجام شده بر روی آن محدود بوده و اغلب در محیط‌های کنترل شده آزمایش شده‌اند [۱۶–۱۹]. هدف این پژوهش، ارائه روشی برای شناسایی نوع و مدل خودروها با دقتی بالا بر روی وسایل نقلیه‌ی داخلی است؛ هرچند که تغییر مجموعه‌ی وسایل نقلیه تأثیری بر ساختار و عملکرد سیستم نخواهد داشت.

۱-۲- کاربردها

شناسایی نوع و مدل وسایل نقلیه کاربردهای متنوعی دارد. در ادامه به چند مورد از این کاربردها اشاره شده است.

- **سیستم‌های نظارتی**: شناسایی نوع و مدل وسیله نقلیه و تطبیق آن با شناسنامه‌ی به‌دست‌آمده از پلاک وسیله نقلیه؛ این کاربرد برای راستی آزمایی خروجی سیستم تشخیص پلاک و جلوگیری از بروز انواع تقلب‌ها (جعل پلاک، انسداد بخشی از پلاک و ...) مناسب است.
- **سیستم‌های امنیتی**: جستجو و یافتن نوع و مدلی خاص از وسایل نقلیه در محدوده‌ی مشخصی از شهر. برای مثال خودرویی متخلف در ناحیه‌ای مشخص از شهر گزارش شده است؛ می‌توان با فیلتر کردن وسایل بر اساس نوع و مدل گزارش شده، جستجوی پلاک‌ها را سرعت بخشید.

[1] Object Class Recognition
[2] Object Detection

- **تبلیغات هدفمند**: در دو دهه اخیر، مطالعاتی در رابطه با مشخصه‌های رانندگان وسایل نقلیه‌ی مختلف صورت گرفته است؛ از قبیل ارتباط جنسیت، درآمد و علایق شخص و خودرویی که سوار می‌شود [۲۰]. به این شکل می‌توان تبلیغات ارائه شده برای وسایل نقلیه را هوشمند سازی کرد، تا تبلیغات متناسب با مشخصه‌های راننده برای او ارائه گردد [۲۱].
- **آمارگیری**: با تلفیق یک سیستم شناسایی نوع و مدل و دوربین‌های موجود در چهارراه‌ها، خیابان‌ها، عوارضی‌ها و بزرگراه‌ها، می‌توان آمار ارزشمندی از فراوانی انواع و مدل‌های مختلف وسایل نقلیه در مناطق مختلف به دست آورد. این آمارها می‌تواند برای راهنمایی و رانندگی، شرکت‌های ساخت خودرو، فروشندگان لوازم جانبی خودرو و سایر شرکت‌های مرتبط بسیار کاربردی باشد.
- **تشخیص تقلب و خودروهای سرقتی**: بسیاری از خودروهای سرقتی پس از تعویض پلاک مورد استفاده سارقین قرار می‌گیرند. می‌توان با تطبیق شناسنامه بدست آمده از پلاک خودرو و مشخصات استخراج شده توسط سیستم شناسایی مدل، برخی از موارد جعل پلاک را تشخیص داده و گزارش کرد.
- **هوشمندسازی سرویس‌ها**: برخی از سرویس‌ها از قبیل پرداخت عوارض یا کارواش‌ها می‌توانند به وسیله شناسایی نوع و مدل وسایل نقلیه به صورت کاملاً خودکار انجام شوند. در حال حاضر این سیستم‌های از تکنولوژی‌هایی همچون GPS، QR Code و حسگرهای مادون قرمز[۱] بهره می‌برند [۲۲–۲۴]. استفاده از سیستم شناسایی مدل هزینه بسیار کمتری به همراه خواهد داشت.
- **کنترل ترافیک**: با اضافه کردن قابلیت شناسایی نوع و مدل وسیله نقلیه می‌توان عملکرد سیستم‌های کنترل ترافیک را بهبود بخشید. برای مثال می‌توان از ورود برخی از انواع خودروها به مسیرهای مشخصی جلوگیری کرد.

۱-۳- چالش‌ها

چالش‌های بسیاری در مسیر تشخیص خودرو و شناسایی نوع و مدل آن وجود دارد. در ادامه تنها به تعدادی از چالش‌های مرتبط با مسئله‌ی هدف اشاره شده است.

- تنوع بسیار زیاد در رنگ، اندازه و شکل وسایل نقلیه.

[۱] Infrared

- تغییرات روشنایی زیاد در شرایط مختلف زمانی و آب و هوایی، مانند صبح، ظهر، عصر و شب، بارندگی، برف و مه.
- انسداد بخشی از وسیله نقلیه به دلایل مختلف از قبیل ترافیک، وجود اجسام اضافی و مشکل‌دار بودن بخشی از خودرو.
- زوایای متفاوت به شکلی که انواع چرخش‌های افقی و عمودی محتمل است. البته در روش پیشنهادی تنها تغییرات زاویه محدود قابل قبول است.
- تنوع بسیار زیاد مدل‌های خودرو و در نتیجه تعداد کلاس‌های زیاد.
- فاصله‌ی بین کلاسی بسیار کم: تنوع در طراحی خودروهای مختلف چندان زیاد نیست.
- فاصله‌ی درون کلاسی قابل توجه: تغییرات رنگ، روشنایی، انسداد بخشی از خودرو، زوایای متفاوت، تعمیر یا تغییر بخشی از خودرو، همگی موجب ایجاد فاصله‌ی درون کلاسی قابل توجه می‌شوند.

بعضی از چالش‌ها در خودروهای داخلی بیشتر دیده می‌شود. برای مثال تغییر رنگ خودرو غیرقانونی است ولی در موارد نه‌چندان کمی دیده می‌شود. تغییر بعضی از بخش‌های خودرو نیز به وفور دیده می‌شود؛ تغییر چراغ‌های جلو/عقب، حذف نشان‌واره[۱]ی شرکت سازنده و تغییرات ظاهری ناشی از حوادث و تصادفات از این نمونه‌ها است.

۱-۴- مفروضات

برای تمرکز بیشتر در پیاده‌سازی هر چه بهتر سیستم پیشنهادی، مفروضات زیر در نظر گرفته شده‌اند.

۱. فرض می‌شود که وسیله یا وسایل نقلیه‌ی موجود در تصویر تشخیص داده شده و مکان و محدوده‌ی آن‌ها مشخص است. در واقع تشخیص و قطعه‌بندی[۲] از وظایف سیستم پیشنهادی محسوب نمی‌گردد.

۲. سیستم‌های شناسایی نوع و مدل وسیله نقلیه اغلب بر روی یک دوربین مداربسته نصب می‌شوند. بنابراین فرض بر این است که زاویه دوربین ثابت بوده و تصاویر دارای تغییرات زاویه محدودی (کمتر از ۱۵ درجه) هستند.

---

[۱] Logo

[۲] Segmentation

۳. از آنجا که نماهای جلو و عقب خودرو دارای ویژگی‌های متمایزکننده بیشتری نسبت به سایر نماها هستند، روش پیشنهادی بر روی یکی از این دو نما پیاده‌سازی و ارزیابی می‌شود.
۴. روش پیشنهادی برای استخراج ویژگی‌های مناسب، نیاز به تصاویری با وضوح[۱] قابل قبول دارد. به همین دلیل، فاصله دوربین تا تصویر ۳ تا ۱۵ متر در نظر گرفته شده است. این فاصله توسط اغلب دوربین‌های مداربسته پشتیبانی می‌شود [۲۵].
۵. سیستم پیشنهادی بر روی تصاویر ثابت آزمایش می‌شود. بهبود تصاویر متحرک و یا ردگیری[۲] شیء موجود در تصاویر متحرک از وظایف سیستم پیشنهادی نمی‌باشند.

۱-۵- نوآوری‌های رساله

در این رساله یک سیستم شناسایی نوع و مدل وسیله نقلیه ارائه شده است. سیستم پیشنهادی دارای چند بخش جدید است که مسیر طی شده تا تکامل آن به همراه مقالات استخراج شده از هر بخش در ادامه مختصراً توضیح داده شده‌اند.

ایده‌ی اصلی این رساله، بهره‌گیری از بخش‌ها برای شناسایی نوع و مدل وسیله نقلیه می‌باشد. منظور از بخش یک پنجره با مکان و ابعاد مشخص از تصویر خودرو است. می‌توان به جای استخراج ویژگی از کل تصویر و یا یک ناحیه مطلوب نسبتا بزرگ که کاهش سرعت و احتمالا کاهش دقت را به همراه خواهد داشت، تنها از تعدادی بخش کوچک بهره برد. برای آزمایش اولیه این ایده، یک مجموعه داده از ۷۰۰ تصویر از ۲۱ مدل متفاوت از نمای جلو و پشت خودروها تهیه شد. سپس تعدادی از بخش‌های معنی‌دار مانند چراغ‌های جلو و عقب و جلوپنجره که در ظاهر قدرت تمایز بیشتری نسبت به سایر بخش‌ها داشتند، به صورت دستی علامت‌گذاری شدند. آنگاه مقایسه‌ای بین طبقه‌بندی با استفاده از این بخش‌ها و روش-های معمول صورت گرفت (بخش ۶-۳-۱-). نتیجه‌ی این مقایسه به همراه یک مرور مناسب بر روی کارهای گذشته در قالب یک مقاله گزارش شد:

۱. محسن بیگلری، علی سلیمانی، "مقایسه روش‌های شناسایی نوع و مدل وسیله نقلیه با رویکرد کلی-نگر و جزئی‌نگر و ارائه یک رویکرد جدید"، مجله علمی-ترویجی محاسبات نرم، پذیرفته شده (فروردین ۹۶)

---

[۱] Resolution
[۲] Tracking

منظور از بخش معنی‌دار، بخشی با نام یا برچسب مشخص مانند چراغ‌ها است. یک بخش متمایز کننده لزوما یک بخش معنی‌دار نیست. بنابراین روش ارائه شده در [۲۶] برای استخراج خودکار بخش‌های متمایزکننده آزمایش شد. در این آزمایش، تعدادی بخش ثابت و مشترک برای همه خودروها به صورت خودکار استخراج گشت. آنگاه کارایی طبقه‌بندی با استفاده از این بخش‌ها با روش قبلی مورد مقایسه قرار گرفت. نتیجه‌ی این آزمایش منجر به تهیه مقاله‌ای دیگر شد:

۲. **محسن بیگلری، علی سلیمانی، حمید حسن‌پور، "شناسایی نوع و مدل وسیله نقلیه با استخراج خودکار بخش‌های مشترک"، مجله علمی-پژوهشی ماشین بینایی و پردازش تصویر، دوره ۴، شماره ۱، صفحه ۳۸-۲۹ (بهار و تابستان ۹۶)**

از آنجا که بخش‌های استخراج شده در آزمایش قبل بین همه خودروها مشترک بودند، انتظار رسیدن به حداکثر کارایی وجود نداشت. بنابراین رویکرد اصلی پیشنهاد شده در رساله از این نقطه آغاز شد. اولین نوآوری این رساله، ارائه یک الگوریتم آموزشی چندکلاسه است که به صورت خودکار مجموعه بخش‌های متمایز کننده هر مدل از خودروها را استخراج می‌کند. این الگوریتم پنج مرحله‌ای به صورت همزمان به استخراج بخش‌های متمایزکننده و رابطه هندسی بین آن‌ها پرداخته و یک طبقه‌بند نیز آموزش می‌دهد. فرآیند آموزش با استفاده از ترکیبی از ماشین بردار پشتیبان[۱] و ماشین بردار پشتیبان مخفی[۲] صورت می-گیرد. آزمایش این الگوریتم بر روی یک مجموعه داده جدید (BVMMR) متشکل از ۶۰۰۰ خودرو از تقریبا ۳۰ مدل متفاوت از خودروهای داخلی منجر به تهیه مقاله سوم گشت:

۳. **محسن بیگلری، علی سلیمانی، حمید حسن‌پور، "شناسایی نوع و مدل وسیله نقلیه با استفاده از مجموعه بخش‌های متمایزکننده"، مجله علمی-پژوهشی پردازش علائم و داده‌ها، پذیرفته شده (خرداد ۹۶)**

دومین نوآوری، ارائه یک الگوریتم حریصانه برای یافتن بخش‌های متمایز کننده هر کلاس از خودروها است. این الگوریتم، تعادلی بین قدرت تمایز یک بخش و فاصله بخش‌های انتخاب شده از هم برقرار می-کند. سومین نوآوری، پیشنهاد یک الگوریتم داده‌کاوی چندکلاسه است که در الگوریتم آموزشی مورد استفاده قرار می‌گیرد. با بهره‌گیری از الگوریتم داده‌کاوی مورد اشاره، می‌توان استفاده بهتری از نمونه‌های منفی کرد و در حین حال به سرعت آموزش بیشتر و فضای محاسباتی کمتری دست پیدا کرد. پس از

[۱] Support Vector Machine (SVM)

[۲] Latent SVM (LSVM)

اضافه کردن دو قسمت مورد اشاره به سیستم پیشنهادی، نتایج بدست آمده توسط سیستم نهایی در قالب دو مقاله جدید گزارش گشت:

۴. **محسن بیگلری، علی سلیمانی، حمید حسن‌پور، " شناسایی نوع و مدل وسیله نقلیه با ترکیبی از مدل‌های مبتنی بر بخش"، کنفرانس پردازش سیگنال و سیستم‌های هوشمند، پذیرفته و ارائه شده (آذر ۹۵)**

5. **Mohsen Biglari, Ali Soleimani, Hamid Hassanpour, "Part-Based Recognition of Vehicle Make and Model", IET Image Processing, vol. 11, no. 7, pp. 483-491 (March 2017)**

تا این مرحله از توسعه سیستم، دو مجموعه داده از خودروهای داخلی جمع‌آوری و حاشیه‌نویسی شده است. برای تهیه این مجموعه داده، یک نرم‌افزار جامع پیاده‌سازی گشت و از آن برای علامت‌گذاری تصاویر بهره گرفته شد. با استفاده از این نرم‌افزار کاربر قادر به علامت‌گذاری کلی یا جزئی خودروها خواهد بود.

نوآوری بعدی، پیشنهاد یک الگوریتم آبشاری است که با استفاده از آن می‌توان تعادل مناسبی بین دقت و سرعت سیستم ایجاد کرد. یک نسخه از این الگوریتم می‌تواند سرعت را بدون افت دقت به مقدار قابل قبولی بهبود بخشد. در حالی که نسخه‌ای دیگر از الگوریتم قادر به افزایش چندین برابری سرعت در ازای افت جزئی دقت است. بررسی تاثیر این الگوریتم بر روی عملکرد سیستم پیشنهادی منجر به مقاله‌ای دیگر شد:

۶. **محسن بیگلری، علی سلیمانی، حمید حسن‌پور، "بهبود سرعت و دقت یک سیستم شناسایی مدل خودرو با ارائه یک الگوریتم آبشاری"، مجله علمی-پژوهشی جهاد دانشگاهی، در حال داوری (مهر ۹۵)**

الگوریتم آبشاری پیشنهادی قابلیت استفاده در کاربردهای مشابه دیگر را نیز دارد. برای نشان دادن این قضیه، مقاله‌ای دیگر نگارش و ارائه شد:

7. **Mohsen Biglari, Hamid Hassanpour, Ali Soleimani, "A Cascading Scheme for Speeding up Multiple Classifier Systems", Pattern Analysis and Applications, Under Review (December 2016)**

نوآوری آخر، خودکارسازی تمامی مراحل مورد اشاره است که منجر به ارائه یک سیستم تمام خودکار برای شناسایی نوع و مدل وسایل نقلیه می‌گردد. مراحل آموزش و آزمایش در این سیستم به صورت کاملاً

خودکار صورت گرفته و نیاز به هیچ‌گونه تنظیمات دستی برای هر مرتبه اجرا ندارد. با تغییر مجموعه داده مورد استفاده برای آموزش، پارامترهای سیستم نیاز به تغییر ندارند. در مرحله آزمایش نیز به هیچ پیش-پردازشی نیاز نیست و نرمال‌سازی تصویر ورودی، تشخیص خودروهای موجود در تصویر و شناسایی نوع و مدل خودروها به صورت کاملاً خودکار صورت می‌گیرد. حتی در رابطه با تعداد خودروهای موجود در تصویر نیز محدودیتی وجود ندارد. ساختار این سیستم جامع در قالب دو مقاله تشریح شده است که مقاله دوم توضیحات بیشتری را در رابطه با بخش‌های کلیدی ارائه کرده است:

۸. محسن بیگلری، علی سلیمانی، حمید حسن‌پور، "ارائه یک سیستم آبشاری مبتنی بر بخش برای شناسایی نوع و مدل وسیله نقلیه"، مجله علمی-پژوهشی رایانش نرم و فناوری اطلاعات، در حال داوری (شهریور ۹۵)

9. Mohsen Biglari, Ali Soleimani, Hamid Hassanpour, "A Cascaded Part-based System for Fine-grained Vehicle Classification", IEEE Transactions on Intelligent Transportation Systems, Accepted (July 2017)

برای خوانایی و سادگی بیشتر، موارد توضیح داده شده در ادامه به صورت موردی فهرست شده‌اند:

- ارائه یک الگوریتم آموزشی چندکلاسه که به صورت توأمان، مجموعه بخش‌های متمایز کننده هر مدل از خودروها را استخراج کرده و با استفاده از بخش‌های استخراج شده و رابطه هندسی بین آن‌ها، یک طبقه‌بند آموزش می‌دهد.
- ارائه یک الگوریتم حریصانه برای یافتن بخش‌های متمایز کننده هر کلاس.
- ارائه یک الگوریتم داده‌کاوی چندکلاسه.
- ارائه یک الگوریتم آبشاری برای بهبود سرعت سیستم.
- ارائه یک سیستم تمام خودکار برای شناسایی نوع و مدل وسایل نقلیه که نسبت به تغییرات روشنایی و تغییرات اندک زاویه مقاوم است.
- جمع‌آوری و حاشیه‌نویسی دو مجموعه تصویر از خودروهای داخلی متشکل از ۷۰۰ تصویر از ۲۰ مدل و ۵۰۰۰ تصویر از تقریباً ۳۰ مدل متفاوت اشاره کرد.
- طراحی و پیاده‌سازی یک نرم‌افزار جامع برای علامت‌گذاری، حاشیه‌نویسی و پردازش خودکار تصاویر علامت‌گذاری شده در دو مجموعه داده تهیه شده.

۱-۶- ساختار رساله

رساله جاری از هفت فصل و یک پیوست تشکیل شده است. مقدمه‌ای در رابطه با موضوع رساله، اهداف و چالش‌های آن در فصل اول ارائه گشته است. در فصل دوم، کارهای پیشین به طور مفصل مورد بررسی قرار گرفته‌اند. روش پیشنهادی در سه فصل سوم، چهارم و پنجم ارائه شده است. در فصل سوم، الگوریتم-های مورد استفاده در روش پیشنهادی توضیح داده شده است. فصل چهارم و پنجم به معرفی سیستم پیشنهادی و بهبودهای صورت گرفته بر روی کارایی آن می‌پردازد. فصل ششم شامل ارزیابی و نتایج آزمایش‌های انجام شده بر روی سیستم پیاده‌سازی شده در رساله می‌باشد. در پایان و در فصل هفتم به نتیجه‌گیری مطالب مطرح شده پرداخته شده است. همچنین، مدل‌های آموزش دیده شده توسط سیستم پیشنهادی و چند نمونه از خروجی‌های سیستم در قالب یک پیوست ارائه گشته است.

۱-۷- جمع‌بندی

در این فصل ابتدا به تعریف مسئله و سپس بیان اهداف پرداخته شد. علاوه‌براین، کاربردهای متنوع موضوع مورد بحث و چالش‌های جدی آن مورد اشاره قرار گرفت. تعداد چالش‌های موجود بیشتر از چالش‌های نامبرده هستند؛ هرچند تنها چالش‌های نامبرده در رساله جاری مورد بررسی قرار گرفته است. در ادامه مفروضات مسئله مطرح گردید و در پایان، خلاصه‌ای از نوآوری‌های صورت گرفته در این رساله بیان شد. به دلیل مفصل بودن توضیحات مربوط به مجموعه داده‌ها، این بخش در فصل ۴-۲- ارائه گشته است.

فصل

۲- مروری بر کارهای پیشین

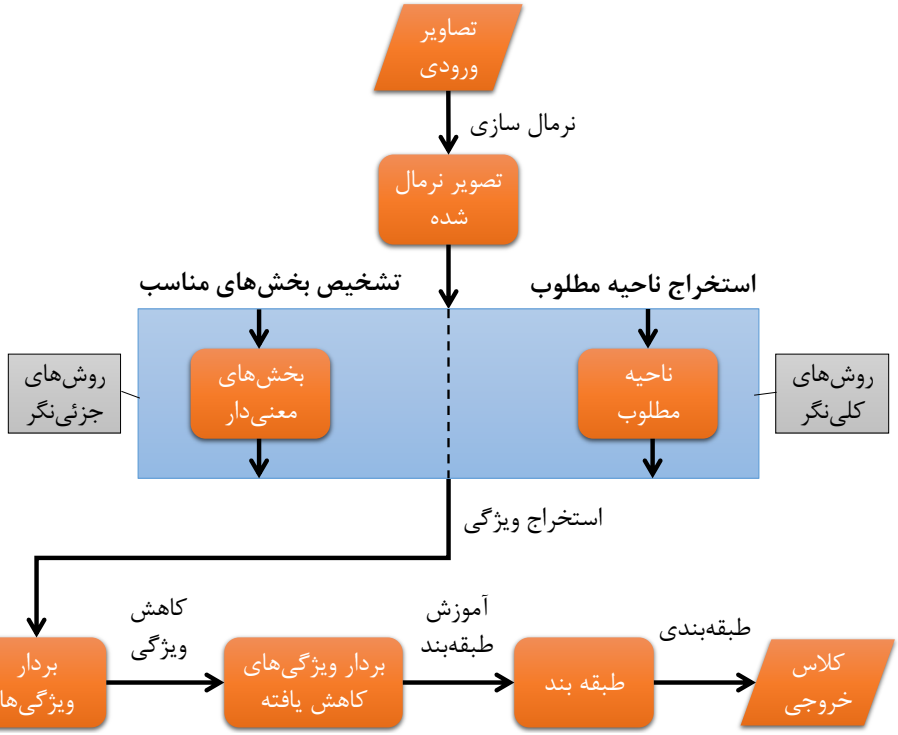
۲-۱- مقدمه

حوزه تشخیص و شناسایی وسایل نقلیه بسیار گسترده است. در زمینه‌ی تشخیص وسایل نقلیه، تحقیقات بسیار صورت گرفته و نتایج قابل قبولی نیز بدست آمده است [۵–۷, ۹, ۲۷]. در زمینه شناسایی دسته‌ی کلی وسایل نقلیه (سنگین، سواری، سبک و ...) نیز روش‌های متعددی ارائه شده است [۱۰, ۱۲, ۱۴, ۱۵, ۲۸–۳۰]. در این بخش تنها به مرور روش‌های مرتبط با موضوع مورد بحث رساله یعنی شناسایی نوع و مدل وسیله نقلیه پرداخته شده است.

روش‌های موجود برای شناسایی نوع و مدل خودرو را می‌توان به دو گروه کلی روش‌های «کلی‌نگر» و «جزئی‌نگر» تقسیم‌بندی کرد. شکل ۲-۱، فلوچارت مراحل کلی کار این روش‌ها را نمایش داده است. البته همه روش‌ها لزوماً همه‌ی بخش‌های نشان داده شده را شامل نمی‌شوند. برخی از روش‌های پیشین را می‌توان در هر دو گروه مورد اشاره قرار داد. در چنین مواردی، با توجه به رویکرد غالب روش تصمیم گرفته شده است.

روش‌های کلی‌نگر، کل تصویر ورودی و یا ناحیه مطلوبی[۱] از آن را مورد پردازش قرار داده و به دنبال یافتن توصیفی کلی از خودرو هستند. در این روش‌ها شناختی در مورد جزئیات ناحیه مورد پردازش وجود ندارد. در نتیجه به تغییرات اندک زاویه و اشتباه در تشخیص ناحیه مطلوب بسیار حساس هستند. از طرفی ناچار به استفاده از مفروضاتی ناقص برای استخراج ناحیه مطلوب نیز می‌باشند. هدف روش‌های جزئی‌نگر، یافتن بخش یا مجموعه بخش‌های متمایز کننده در تصویر است که به جداسازی کلاس‌ها کمک می‌کنند. برای مثال می‌دانیم که چراغ‌ها بخشی یکتا در بسیاری از مدل‌های خودرو هستند؛ درصورتی‌که بتوانیم چراغ‌های یک خودرو را مکان‌یابی کنیم، می‌توانیم تعداد زیادی از مدل‌ها را از یکدیگر تمیز دهیم. مهم‌ترین ضعف این روش‌ها در مکان‌یابی بخش‌ها است؛ درصورتی‌که بخش موردنظر به‌درستی تشخیص داده نشود، عملکرد کلیه‌ی مراحل بعدی زیر سؤال خواهد رفت. تغییر بخش «استخراج ناحیه مطلوب» به بخش «تشخیص بخش‌های معنی‌دار» در شکل ۲-۱ ممکن است موجب تغییر ساختار سایر بخش‌ها نیز بشود.

[۱] Region of Interest (ROI)



شکل ۲-۱: فلوچارت مراحل کار سیستم‌های کلی‌نگر و جزئی‌نگر در شناسایی نوع و مدل خودرو.

۲-۲- روش‌های کلی‌نگر

مرجع [۳۱] 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 نموده است. شکل ۲-۲ ضرایب استفاده شده را نشان می‌دهد. پس از تشخیص ناحیه مطلوب، از نسخه‌ی ساده‌شده‌ای از عملگر تبدیل ویژگی مقاوم به اندازه[۱] برای استخراج ویژگی بهره گرفته شده است [۳۲]. برای این منظور، ابتدا ناحیه مطلوب به تعدادی بخش هم‌پوشان تقسیم شده و از هر بخش، ویژگی استخراج گشته است. شکل ۲-۳ نمایی از شکل استخراج ویژگی در این سیستم را نشان می‌دهد. از فاصله اقلیدسی به عنوان معیار شباهت بردارهای ویژگی و از k-نزدیک‌ترین همسایگی‌ها[۲] به عنوان طبقه‌بند استفاده شده است. همچنین کارایی الگوریتم تحلیل افتراقی خطی[۳] برای کاهش ویژگی و سپس طبقه‌بندی[۴] نیز مورد

---

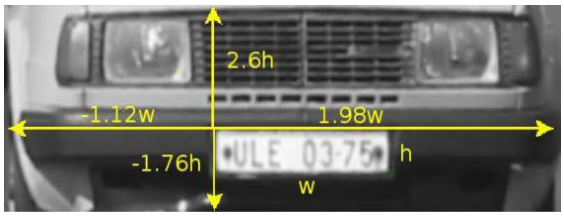
۱ Scale-Invariant Feature Transform (SIFT)

۲ k-Nearest Neighbors (k-NN)

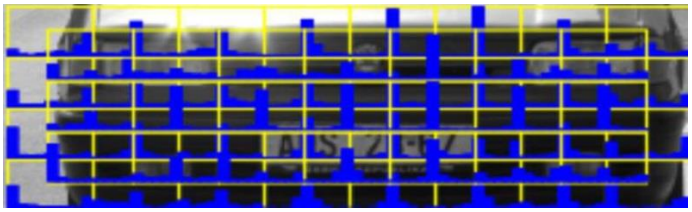
۳ Linear Discriminant Analysis (LDA)

۴ Classification

آزمایش قرار گرفته است [۳۳]. لاگنکوو [۳۴] و چند مرجع دیگر نیز روند مشابهی را برای استخراج ناحیه مطلوب به کار گرفته‌اند [۳۵–۳۹].



شکل ۲-۲: روش تشخیص ناحیه مطلوب در [۳۱].



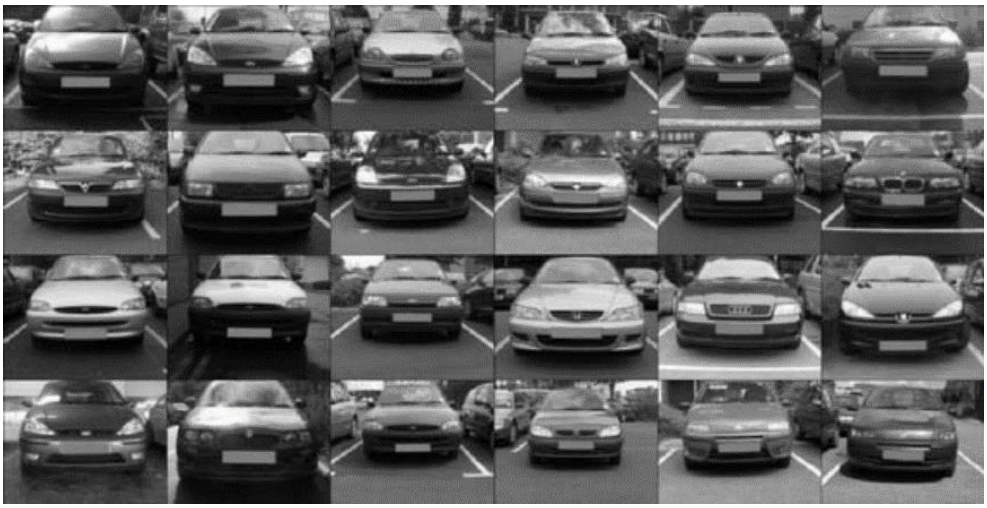
شکل ۲-۳: استخراج ویژگی با استفاده از عملگر تبدیل ویژگی مقاوم به اندازه به صورت چندبخشی در [۳۱].

در مقاله مورد اشاره از دو مجموعه داده استفاده شده است [۳۱]؛ یکی برای سواری‌ها که از ۱۰۰ تصویر در ده کلاس تشکیل شده و دیگری برای وسایل نقلیه سنگین که از ۵۰۰ تصویر در هفت کلاس تشکیل شده است. بهترین دقت گزارش شده برای این دو مجموعه داده، به ترتیب ٪۹۴ و ٪۹۲ می‌باشد. همچنین استفاده از تحلیل افتراقی خطی برای کاهش ویژگی باعث افزایش دقت سیستم نگردیده است. نمونه‌ای از مجموعه داده‌های تهیه شده در شکل ۲-۴ مشاهده می‌شود. وابستگی این سیستم به مکان پلاک، بزرگترین اشکال آن است که برخی از کاربردهای مهم تشخیص نوع و مدل خودرو را زیر سوال می‌برد. دقت کسب شده توسط این سیستم برای تعداد کلاس‌های محدودِ مورد استفاده، قابل قبول نمی‌باشد. علاوه بر این، روش مذکور حساسیت بالایی نسبت به زاویه و نویز تصویر دارد.

شکل ۲-۴: مجموعه داده‌ی تهیه شده توسط [۳۱].

پتروویچ و همکاران [۴۰] نیز پس از استخراج ناحیه مطلوب مشابه با روش قبل، روش‌های مختلف استخراج ویژگی از قبیل جهت لبه‌ها[۱]، روش تشخیص گوشه‌ی هریس[۲] [۴۱]، پاسخ لبه سوبل[۳] و گرادیان‌های نگاشت شده‌ی مربعی[۴] شکل (روش پیشنهادی مقاله) را برای شناسایی نوع و مدل خودرو مورد آزمایش قرار دادند. در نهایت، ویژگی گرادیان‌های نگاشت شده‌ی مربعی با استفاده از فاصله اقلیدسی برای سنجش فاصله‌ی بردارهای ویژگی و نزدیک‌ترین همسایگی برای طبقه‌بندی، بهترین نتیجه را که برابر با ٪۹۳ بوده است را بدست داد. این نتیجه بر روی مجموعه داده‌ای با ۱۱۳۲ تصویر از ۷۷ کلاس مختلف به دست آمده است. ۱۰۵ تصویر برای آموزش و ۱۰۲۷ تصویر برای آزمایش به کار گرفته شده‌اند. تصاویر همگی از نمای جلوی خودرو و با تغییرات زاویه بسیار کم تهیه شده‌اند. شکل ۲-۵ چند نمونه از این تصاویر را نشان می‌دهد. شکل ۲-۶ چند مورد از تصاویری که روش پیشنهادی موفق به طبقه‌بندی درست آن‌ها نشده را نمایش داده است که در این بین، نمونه‌های ساده‌ای هم دیده می‌شوند.

[۱] Edge Orientation
[۲] Harris Corner Detection
[۳] Sobel Edge Response
[۴] Square Mapped Gradients



شکل ۲-۵: مجموعه داده‌ی تهیه شده توسط [۴۰].

شکل ۲-۶: چند نمونه از تصاویری که روش پیشنهادی در [۴۰] موفق به طبقه‌بندی درست آن‌ها نشده است.

پیِرس و همکارش [۳۸] با پتروویچ هم‌عقیده نیستند؛ آن‌ها به این نتیجه رسیدند که روش تشخیص گوشه‌ی هریس اگر به‌صورت محلی و بخش‌بندی شده اعمال شود، عملکرد بهتری نسبت به گرادیان‌های نگاشت شده‌ی مربعی خواهد داشت. گرادیان‌های نگاشت شده‌ی مربعی با استفاده از پاسخ لبه‌های افقی و عمودی سوبل و به‌صورت نشان داده شده در فرمول (۲-۱) محاسبه می‌شوند. در طرف مقابل، روش تشخیص گوشه‌ی هریس با استفاده از مشتقات نسخه‌ی هموارشده‌ی تصویر محاسبه می‌شود (فرمول (۲-۲)).

$$g_y = \frac{2s_x s_y}{s_x^2 + s_y^2} \qquad gx = \frac{s_x^2 - s_y^2}{s_x^2 + s_y^2} \tag{۱-۲}$$

$$C_N = \frac{I_x^2 I_y^2 - \left(I_{xy}^2\right)}{I_x^2 + I_y^2} \tag{۲-۲}$$

در فرمول (۲-۱)، S_x و S_y پاسخ‌های لبه‌های افقی و عمودی سوبل و در فرمول (۲-۲)، I_x و I_y و I_{xy} مشتقات نسخه‌ی هموار شده‌ی تصویر می‌باشند. شکل ۲-۷، خروجی این دو روش را برای یک نمونه ناحیه مطلوب تصویر خودرو نشان می‌دهد [۳۸]. برای استخراج بردار ویژگی، پاسخ روش تشخیص گوشه‌ی هریس بر روی تصویر ورودی با ابعاد ۱۵۰ در ۱۵۰ محاسبه شده و خروجی حاصل در برداری به طول ۲۲۵۰۰ قرار داده می‌شود. برای افزایش کارایی، این فرآیند به صورت چندبخشی تکرار گشته و هر بار تصویر به چهار قسمت شکسته شده و بردار ویژگی از هر یک از آن‌ها استخراج می‌گردد؛ و این روال تا چند مرحله به صورت بازگشتی تکرار می‌شود. این مقاله عملکرد دو طبقه‌بند k-نزدیک‌ترین همسایگی‌ها و بیز ساده[۱] را مقایسه کرده و بیز ساده را ترجیح داده است؛ تئوری بیز در فرمول (۲-۳) آورده شده است. هدف یافتن احتمال پسین[۲] هر کلاس C با در دست داشتن بردار ویژگی X است. کلاسی که دارای بالاترین احتمال پسین باشد به عنوان کلاس خروجی انتخاب می‌گردد. $P(C)$ از فراوانی هر کلاس در مجموعه داده‌های آموزشی محاسبه می‌شود.

$$p(C|X) = \frac{P(X|C)P(C)}{P(X)} \qquad (۲-۳)$$

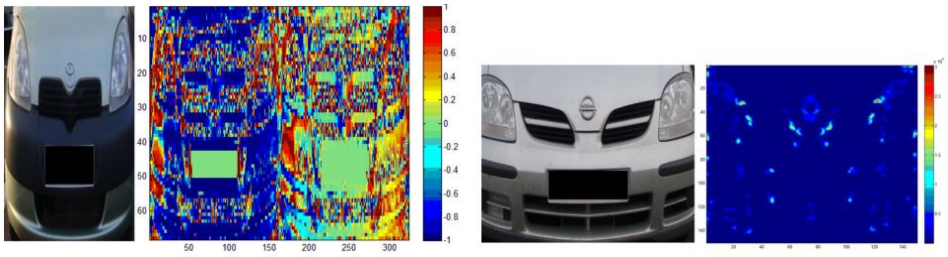
دقت ٪۹۶ برای ۲۶۲ تصویر از نمای جلوی ۷۴ کلاس مختلف و به روش ارزیابی ضربدری تک حذفی[۳]، گزارش شده است. ناحیه مطلوب با مکان‌یابی دستی پلاک‌ها استخراج گشته است. تعداد زیاد کلاس‌ها و تعداد تصاویر کم استفاده شده در این تحقیق با هم سازگاری ندارند. در این حالت، به اغلب کلاس‌ها تنها یک یا دو نمونه تعلق می‌گیرد که برای آزمایش عملکرد الگوریتم کافی نیست.

---

۱ Naïve Bayes

۲ Posterior Probability

۳ Leave-one-out Cross Validation



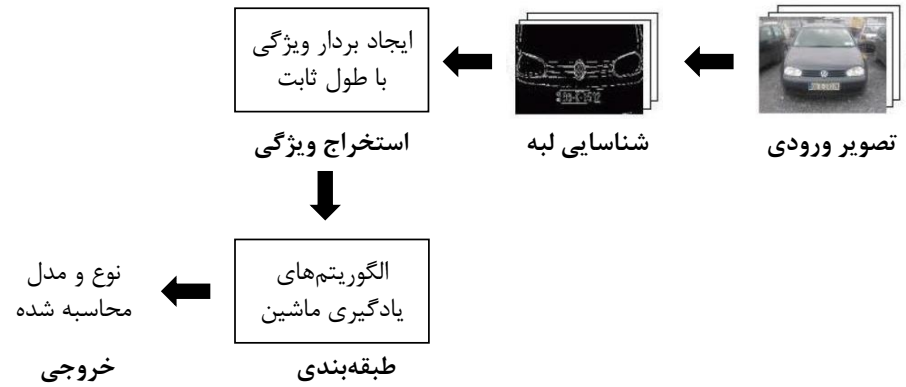
شکل ۲-۷: خروجی الگوریتم‌های تشخیص گوشه‌ی هریس (سمت راست) و گرادیان‌های نگاشت شده‌ی مربعی (سمت چپ) برای یک نمونه ناحیه مطلوب خودرو [۳۸].

مونرو و همکاران چند روش طبقه‌بندی تک کلاسه و چند کلاسه را مورد آزمایش قرار داده‌اند [۴۲]. آن‌ها با استفاده از الگوریتم تشخیص لبه‌ی کَنی[۱]، ناحیه مطلوب را تشخیص داده و از آن ویژگی استخراج کرده‌اند. برای تشخیص مکان چراغ‌های جلو، با استفاده از عملگرهای مورفولوژیکی[۲] حلقه‌های موجود در تصویر لبه‌ها پر گشته و ناحیه‌های پیوسته جستجو شده‌اند. برای استخراج ویژگی از ناحیه مطلوب و چراغ‌ها، فاصله‌ی لبه‌ها تا نقطه‌ی مرکزی ناحیه مربوطه در یک بردار یک‌بعدی نگاشت می‌شود. مجموعه داده‌ی مورد استفاده از ۱۵۰ تصویر از پنج کلاس مختلف، به اضافه ۳۰ تصویر از کلاس متفرقه تشکیل شده است. در حالت طبقه‌بندی چند کلاسه، عملکرد k-نزدیک‌ترین همسایگی‌ها و شبکه عصبی پرسپترون چندلایه[۳] برابر با ٪۹۹/۹۹ و ٪۹۵/۵۳ گزارش شده است؛ برای محاسبه دقت سیستم از روش ارزیابی ضربدری به پیمانه ده بهره گرفته شده است. در طبقه‌بندی تک کلاسه‌ی تعریف شده توسط این مقاله، طبقه‌بند، عضویت یا عدم عضویت ورودی را به تنها کلاس موجود اعلام می‌کند. برای این منظور از نزدیک‌ترین همسایگی استفاده شده و دقت ٪۹۷/۷۰ گزارش گشته است. فلوچارت مراحل سیستم در شکل ۲-۸ به تصویر کشیده شده است.

[۱] Canny Edge Detection

[۲] Morphological Operators

[۳] Multi-Layer Perceptron



شکل ۲-۸: فلوچارت مراحل سیستم پیشنهادی در [۴۲].

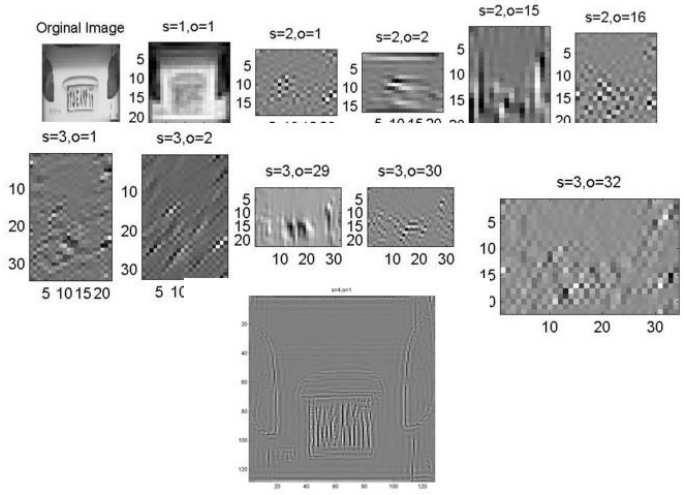
کاظمی و همکاران [۴۳] کارایی ویژگی جدیدی به نام تبدیل پیچک[۱] [۴۴] که نسخه‌ای از تبدیل موجک[۲] است را بررسی کرده‌اند. آن‌ها همچنین دو طبقه‌بند ماشین بردار پشتیبان و k-نزدیک‌ترین همسایگی‌ها را نیز مورد مقایسه قرار داده‌اند. شکل ۲-۹، عملکرد تبدیل پیچک را با اندازه و جهت‌های مختلف نشان داده است. مجموعه داده‌ی تهیه شده توسط آن‌ها تنها دارای ۳۰۰ تصویر از پنج کلاس متفاوت از نمای پشت وسایل نقلیه است. ۲۳۰ تصویر برای آموزش و ۷۰ تصویر برای آزمایش سیستم به کار گرفته شده است. دقت گزارش شده برای ماشین بردار پشتیبان برابر با ٪۹۹ و برای k-نزدیک‌ترین همسایگی‌ها برابر با ٪۹۲ بوده است.

ظفر و همکاران در [۴۵] از روش مشابهی به نام تبدیل کانتورلت[۳] [۴۶] برای استخراج ویژگی استفاده کرده‌اند. ادعا می‌شود که این تبدیل بهتر از دو روش قبل، خم‌های هموار را استخراج می‌کند. مقایسه‌ای از عملکرد این تبدیل و تبدیل موجک در شکل ۲-۱۰ به نمایش گذاشته شده است. ظفر و همکاران پس از استخراج ویژگی توسط تبدیل کانتورلت، به کمک الگوریتم تحلیل افتراقی خطی دوبعدی[۴] طول بردار ویژگی‌ها را کاهش داده‌اند؛ سپس طبقه‌بند ماشین بردار پشتیبان را آموزش داده و دقت ٪۹۶ را بر روی مجموعه داده‌ای با ۳۰۰ تصویر از نمای جلوی ۲۵ کلاس مختلف از خودروها گزارش کرده‌اند. از هر کلاس حداقل ۱۱ تصویر موجود بوده و تأثیر استفاده از تعداد تصاویر مختلف برای آموزش بررسی گردیده است. کاظمی و همکاران در مقاله‌ای دیگر و به شکلی جامع‌تر، این بار تبدیل‌های فوریه سریع، موجک و پیچک را برای استخراج ویژگی با یکدیگر مورد مقایسه قرار داده‌اند [۴۷]. تأکید بر روی عملکرد بهتر تبدیل

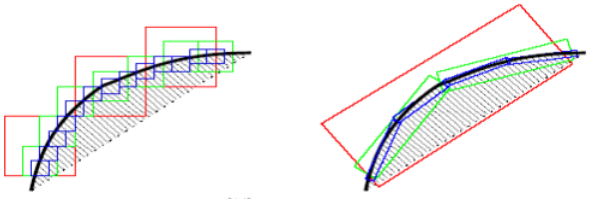
---

۱ Curvelet Transform
۲ Wavelet Transform
۳ Contourlet Transform
۴ 2D-LDA

پیچک نسبت به دو تبدیل دیگر، نتیجه‌ی آزمایشات صورت گرفته توسط آن‌ها بوده است. دقت گزارش شده بر روی مجموعه داده‌ی قبلی برای سه تبدیل نامبرده به ترتیب برابر با ٪۹۷، ٪۹۲ و ٪۱۰۰ می‌باشد.



شکل ۲-۹: عملکرد تبدیل پیچک با اندازه و جهت‌های مختلف [۴۳]. S اندازه و O جهت را نشان می‌دهد.

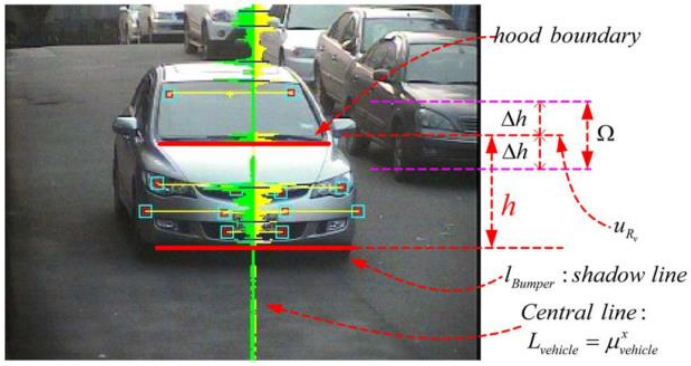


شکل ۲-۱۰: تشخیص کانتورها توسط تبدیل موجک و تبدیل کانتورلت [۴۴].

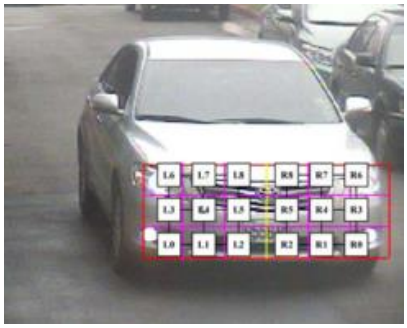
شیپِه و همکاران عملگر ویژگی‌های مقاوم تسریع شده[۱] را به شکلی تغییر داده‌اند که تنها نقاطی که دارای معادل متقارن هستند را نگه دارد [۴۸, ۴۹]. از این ویژگی می‌توان برای تشخیص اشیاء متقارن بهره برد. آن‌ها پس از تشخیص نقاط متقارن در تصویر، خط عمود بر مرکز خطوط متصل کننده نقاط متقارن را رسم کرده و با استفاده از این خط، ناحیه مطلوب را استخراج می‌کنند. محدوده‌ی پایینی ناحیه مطلوب توسط سایه‌ی موجود در زیر ماشین تشخیص داده می‌شود. طول سایه درواقع طول ناحیه مطلوب را مشخص می‌کند. محدوده‌ی بالایی ناحیه مطلوب توسط خط بین کاپوت و شیشه‌ی جلو تعیین می‌گردد. تشخیص خطوط توسط الگوریتم تشخیص لبه سوبل انجام می‌شود. شکل ۲-۱۱ تمامی توضیحات داده‌شده را به تصویر کشیده است.

[۱] Speeded Up Robust Features (SURF)

پس از تشخیص ناحیه مطلوب، محدوده‌ی مربوطه به 6 × 4 بخش تقسیم می‌شود (شکل ۲-۱۲). سپس از هر بخش، توسط الگوریتم‌های هیستوگرام گرادیان‌های جهت‌دار[۱] و ویژگی‌های مقاوم تسریع شده ویژگی استخراج شده و هر بخش به‌صورت جداگانه به یک طبقه‌بند ماشین بردار پشتیبان آموزش داده می‌شود. در مرحله‌ی آزمون، بین همه‌ی طبقه‌بندها رأی‌گیری شده و کلاس خروجی مشخص می‌گردد. دقت گزارش شده برای مجموعه داده‌ای با ۲۸۴۶ تصویر از ۲۹ کلاس مختلف برای آموزش و ۴۰۹۰ تصویر برای آزمایش، برابر با ٪۹۹ بوده است.



شکل ۲-۱۱: روش ارائه شده برای تشخیص خودرو و ناحیه مطلوب به صورت همزمان توسط [۴۸].



شکل ۲-۱۲: استخراج ویژگی ارائه شده به‌صورت چندبخشی از ناحیه مطلوب توسط [۴۸].

ناظمی و همکاران از تبدیل ویژگی مقاوم به اندازه برای استخراج ویژگی و از ماشین بردار پشتیبان برای طبقه‌بندی تصاویر خودروها در جاده‌ها بهره برده‌اند [۵۰]. نوآوری آن‌ها در استفاده از تکنیک‌های کدگذاری ویژگی پراکنده[۲] مبتنی بر روش کیسه‌ای از کلمات[۳] بر روی بردار ویژگی‌ها بوده است. پس از کدگذاری پراکنده‌ی بردار ویژگی‌ها، بردار ویژگی جدید به طبقه‌بند ماشین بردار پشتیبان آموزش داده

[۱] Histogram of Oriented Gradients (HOG)
[۲] Sparse Feature Coding
[۳] Bag of Words (BOW)

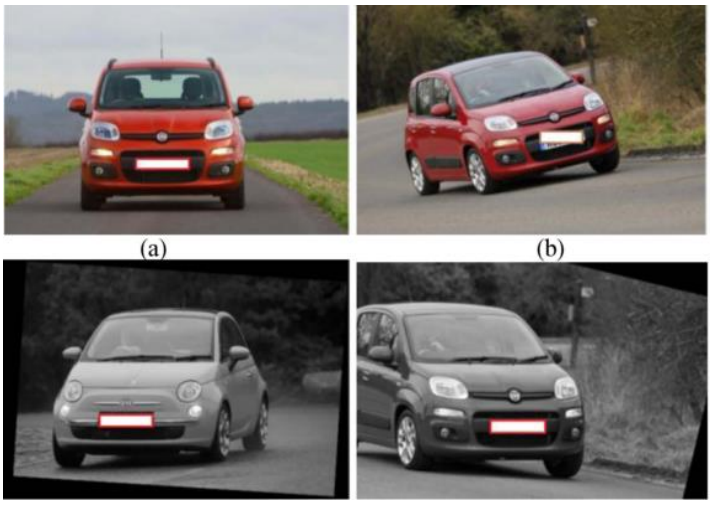
می‌شود؛ برای آموزش از ۲۰۰ تصویر به ازای هر یک از ده کلاس موجود و برای آزمایش، ۵۰ تصویر از هر کلاس به کار گرفته می‌شود. نمونه‌ای از تصاویر مورد استفاده در شکل ۲-۱۳ نمایش داده شده‌اند. دقت گزارش شده برابر با ٪۹۳/۲۰ بوده است.

شکل ۲-۱۳: نمونه تصاویری از مجموعه داده مورد استفاده در [۵۰].

روش کیسه‌ای از کلمات توسط [۵۱] نیز به کار گرفته شده است. در این روش ابتدا توسط عملگر ویژگی‌های مقاوم تسریع شده از همه تصاویر ویژگی استخراج می‌شود. سپس با استفاده از این مجموعه ویژگی‌ها، کیسه‌ای از کلمات بصری[۱] بدست آورده می‌شود. آنگاه ویژگی‌های استخراج شده از هر تصویر با استفاده از این کیسه به یک هیستوگرام تبدیل می‌شود. هر ستون از این هیستوگرام، فراوانی یک کلمه در بردار ویژگی‌های استخراج شده از تصویر را نشان می‌دهد. دقت بدست آمده از طبقه‌بندی این هیستوگرام‌ها توسط ماشین بردار پشتیبان برابر با ٪۹۵/۷۷ گزارش شده است. در این مقاله از مجموعه داده به تازگی منتشر شده‌ی NTOU-MMR، استفاده شده است [۵۲]. ۶۵۰۰ تصویر روبه‌جلو از ۲۹ نوع و مدل خودروی موجود در این مجموعه داده به صورت تصادفی (٪۸۰ به ٪۲۰) به تصاویر آموزشی و آزمایشی تقسیم گشته‌اند.

[۱] Bag of Visual Words

ساروی و همکاران از هیستوگرام شکل انرژی محلی[1] برای استخراج ویژگی از تصاویر به دست آمده از دوربین‌های مداربسته[2] استفاده کرده‌اند [۳۹]. این روش لبه‌ها و خم‌ها را به خوبی استخراج می‌کند. برای استحکام بیشتر ویژگی‌های استخراج شده، تصاویر پس از تشخیص پلاک، نسبت به پلاک، تراز می‌شوند. این امر باعث می‌شود، تغییرات زاویه در سر پیچ‌ها تأثیر کمتری بر روی ویژگی‌های استخراجی بگذارند. شکل ۲-۱۴ یک نمونه تصویر تراز نشده و تراز شده را نمایش می‌دهد. پس از تراز کردن، ناحیه مطلوب با استفاده از اندازه‌های ثابت نشان داده شده در شکل ۲-۱۵ استخراج گشته و سپس بردار ویژگی‌ها محاسبه می‌گردد. مجموعه داده‌های آموزشی از ۱۹۶ تصویر روبه‌جلو از ۲۲ کلاس مختلف تشکیل شده‌اند. برای آزمایش سیستم از سه ویدئو به طول یک الی دو دقیقه بهره برده شده که در آن‌ها خودروهایی از هر ۲۲ کلاس وجود دارند. طبقه‌بند ماشین بردار پشتیبان به پنج فریم متوالی اعمال شده و بین نتایج به دست آمده رأی‌گیری می‌شود. دقت ٪۹۵/۸۳ گزارش شده است.



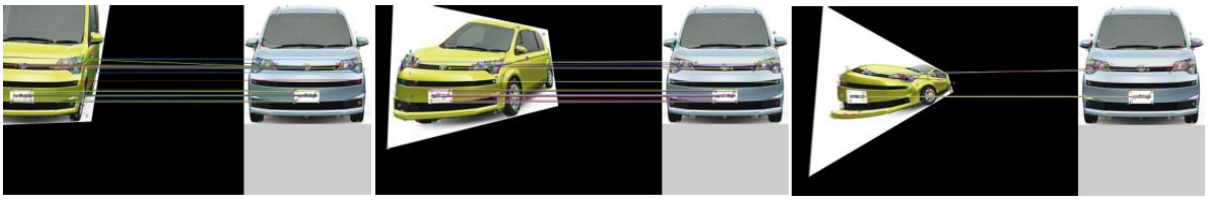
شکل ۲-۱۴: تراز کردن افقی تصویر نسبت به پلاک [۳۹]. ردیف بالا تصویر تراز نشده و ردیف پایین تصویر تراز شده را نمایش داده است.

[1] Local Energy Shape Histogram (LESH)

[2] Closed-Circuit Television Camera (CCTV)

شکل ۲-۱۵: روش استخراج ناحیه مطلوب در [۳۹].

شینوزوکا و همکاران سیستمی بر روی تلفن همراه پیاده‌سازی کرده‌اند که نسبت به چرخش تا ۶۰ درجه مقاوم است [۵۳]. تصاویر آموزشی در این سیستم همگی روبه‌جلو هستند. پس از دریافت تصویر ورودی، نقاط کلیدی توسط الگوریتم تبدیل ویژگی مقاوم به اندازه استخراج می‌شوند. با معیار قرار دادن چهارگوشه پلاک دو تصویر، تصویر ورودی توسط ماتریس هوموگرافی[۱] چرخانده شده تا به نمای جلو منتقل شود. سپس مقایسه ویژگی‌ها توسط معیار کسینوسی[۲] انجام شده و تصاویر ثبت شده در سیستم به ترتیب امتیازی که کسب کرده‌اند، مرتب شده و نمایش داده می‌شوند. ۶۰ تصویر آموزشی از نه کلاس که از هر یک دو نمونه با رنگ‌های متفاوت وجود دارد در سیستم ثبت شده‌اند. از سه تا ده تصویر از نه کلاس مختلف برای آزمایش سیستم بهره گرفته شده است. تصاویر همگی به‌صورت دستی بریده شده و کاملاً ایده‌آل هستند. نتیجه‌ی گزارش شده از آزمایش‌ها نشان می‌دهد که سیستم در پنج تصویر پیشنهادی اول، کلاس درست تصویر را تشخیص می‌دهد. شکل ۲-۱۶ نمونه‌ای از روش چرخش تصویر پیشنهاد شده را نمایش می‌دهد.



شکل ۲-۱۶: روش چرخش پیشنهادی توسط [۵۳] برای از بین بردن تغییرات زاویه در هنگام مقایسه دو تصویر.

باران و همکاران دو روش برای شناسایی نوع و مدل خودرو ارائه کرده‌اند که یکی برخط[۳] با تمرکز بر سرعت و دیگری خارجِ خط[۱] و با تمرکز بر دقت است [۵۴]. سیستم برخط از ویژگی‌های مقاوم تسریع

---

[۱] Homography Matrix

[۲] Cosine

[۳] Online

شده برای استخراج ویژگی و از ماشین بردار پشتیبان برای طبقه‌بندی بهره می‌برد. سیستم خارجِ خط از تبدیل ویژگی مقاوم به اندازه، ویژگی‌های مقاوم تسریع شده و هیستوگرام لبه[۲] به‌صورت ترکیبی برای استخراج ویژگی‌ها و از k-نزدیک‌ترین همسایگی‌ها برای طبقه‌بندی استفاده می‌کند. برای تشخیص ناحیه مطلوب از الگوریتم تشخیص شیء ارائه شده توسط ویولا و جونز استفاده شده است [۵۵]؛ ناحیه مطلوب شامل چراغ‌های جلو، نشان‌واره و جلوپنجره[۳] به این تشخیص دهنده آموزش داده می‌شود (شکل ۲-۱۷).

مجموعه داده تهیه شده توسط [۵۴] از ۱۳۶۰ تصویر آموزشی و ۲۴۹۹ تصویر آزمایشی که نیمی از آن‌ها از اینترنت و نیم دیگر به‌صورت دستی جمع‌آوری شده‌اند، تشکیل شده است. تعداد نمونه‌های مربوط به هر یک از ۱۷ کلاس در تصاویر آموزشی ۸۰ عدد بوده و در تصاویر آزمایشی از ۶۰ تا ۲۳۰ نمونه متفاوت است. دقت گزارش شده برای سیستم برخط برابر با ٪۹۱/۷ و برای سیستم خارجِ خط ٪۹۷/۲ می‌باشد. میانگین زمان پردازش یک تصویر با ابعاد ۷۰۴ در ۵۷۶ توسط سیستم برخط برابر با ۴۰ میلی‌ثانیه گزارش گردیده است. نوآوری این مقاله بیشتر بر روی پیاده‌سازی یک سیستم جامع بوده و از هیچ رویکرد جدیدی استفاده نشده است؛ چند نمونه‌ی دیگر ازاین‌دست نیز وجود دارند [۳۹, ۵۶].

شکل ۲-۱۷: ناحیه مطلوب آموزش داده شده به الگوریتم تشخیص شیء ویولا و جونز در [۵۴].

۲-۳- روش‌های جزئی‌نگر

سایلس و همکاران نشان‌واره را بخشی متمایزکننده برای شناسایی مدل خودروها در نظر گرفته‌اند [۵۷]. برای تشخیص نشان‌واره، ابتدا با استفاده از مکان و اندازه پلاک، ناحیه مطلوب خودرو استخراج می‌گردد. سپس روش نگاشت ویژگی تناسب فاز[۴] [۵۸] برای تشخیص نشان‌واره اعمال می‌شود. این الگوریتم، روشی مناسب برای استخراج ویژگی‌های برجسته مانند لبه‌ها و گوشه از ناحیه مطلوب می‌باشد. با نگاشت خروجی الگوریتم بر روی محور افقی و فیلتر کردن مقادیر کمتر از یک حد آستانه‌ی مشخص، می‌توان محل تقریبی نشان‌واره را تشخیص داد؛ البته برای این منظور، در نظر داشتن محل تقریبی نشان‌واره با

---

[۱] Offline

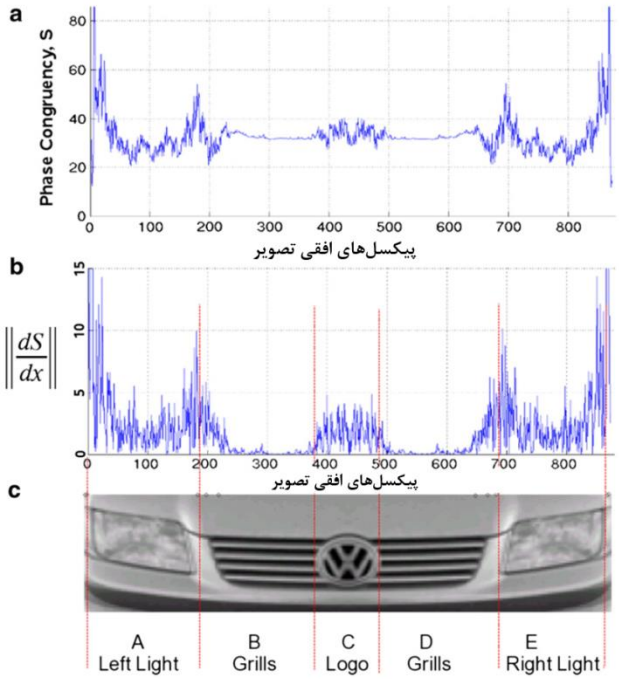
[۲] Edge Histogram

[۳] Grilles

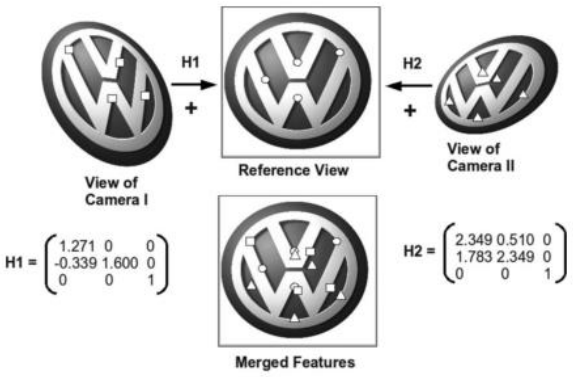
[۴] Phase Congruency Feature Map (PCFM)

توجه به مکان پلاک ضروری به نظر می‌رسد. شکل ۲-۱۸ توضیحات داده شده را به تصویر کشیده است. پس از تشخیص نشان‌واره، نشان‌واره‌های استخراجی از تعدادی تصویر هم‌کلاس جستجو شده و بهترین نشان‌واره به‌صورت دستی برگزیده می‌شود. سپس با مقایسه‌ی نشان‌واره‌ی سایر تصاویر نسبت به تصویر مبنا، ماتریس هوموگرافی همه‌ی آن‌ها محاسبه شده و ویژگی‌های استخراجی از همه نشان‌واره‌ها به عنوان بردار ویژگی ذخیره می‌گردد. به کمک ماتریس هوموگرافی، ویژگی سایر نشان‌واره‌ها به فضای ویژگی‌های نشان‌واره‌ی مبنا نگاشت می‌شود (شکل ۲-۱۹).

در مقاله مورد بحث [۵۷]، از تبدیل ویژگی مقاوم به اندازه برای استخراج ویژگی و از k-نزدیک‌ترین همسایگی‌ها برای طبقه‌بندی تصاویر بهره گرفته شده است. مجموعه داده‌ی مورد استفاده از ۱۲۰۰ نشان‌واره از ده کلاس تشکیل شده که ۴۰۰ نمونه از آن‌ها برای آموزش و باقی تصاویر برای آزمایش به کار گرفته شده‌اند. دقت به دست آمده برای سیستم مورد اشاره برابر با ٪۹۱ بوده است. همان‌طور که پیش‌تر در مورد کل روش‌های جزئی‌نگر اشاره شد، ضعف این روش نیز در یافتن مکان نشان‌واره است؛ برای یافتن نشان‌واره، فرض شده که جلوپنجره‌ها دارای الگوی افقی هستند، درصورتی‌که این فرض در بسیاری از موارد برقرار نیست. همچنین نشان‌واره به تنهایی قادر به تعیین مدل دقیق خودرو نمی‌باشد.



شکل ۲-۱۸: تشخیص محل نشان‌واره با استفاده از روش نگاشت ویژگی تناسب فاز در [۵۷].

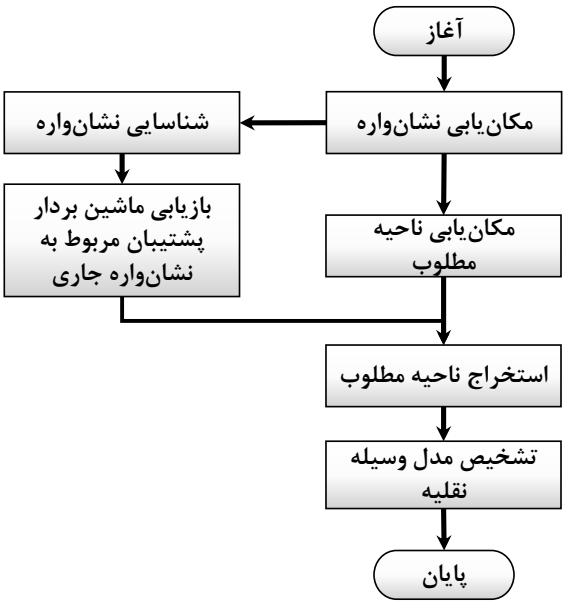


شکل ۲-۱۹: نگاشت ویژگی‌های استخراج شده از سایر نشان‌واره‌ها به فضای ویژگی‌های نشان‌واره‌ی مبنا [۵۷].

یانگ و همکاران نیز در رویکردی مشابه، سعی در شناسایی نوع خودرو با استفاده از نشان‌واره کرده‌اند (شکل ۲-۲۰) [۵۹]. در این مقاله، ابتدا بسته به مکان پلاک، ناحیه مطلوبِ مربوط به نشان‌واره استخراج می‌شود؛ برای این منظور، ناحیه‌ی مشخصی از بالای پلاک خودرو در نظر گرفته شده و لبه‌های عمودی آن استخراج می‌گردند. سپس با استفاده از فیلتر گوسی[۱]، تصویر هموار شده و به شکل دودویی درآورده می‌شود. ناحیه‌های پیوسته کوچک حذف شده و ناحیه‌های متصل ادغام می‌شوند. در نهایت تنها ناحیه‌ی باقی مانده توسط عملگرهای مورفولوژیکی پر شده و به عنوان ناحیه دربرگیرنده نشان‌واره در نظر گرفته می‌شود. در مرحله‌ی بعد، با استفاده از تبدیل کسینوسی گسسته[۲]، از نشان‌واره و یک پنجره با اندازه‌ی ثابت از سمت چپ آن، بردار ویژگی‌ها استخراج می‌گردد (شکل ۲-۲۱). نویسندگان این مقاله معتقد بوده‌اند که سمت راست پنجره ممکن است گاهی شامل متن‌های دیگری باشد و قدرت تمایز بالایی ندارد. ویژگی‌های استخراجی به یک طبقه‌بند ماشین بردار پشتیبان آموزش داده شده و بر روی مجموعه داده‌ای با ۱۰۹۶ تصویر از ۱۲ نوع مختلف خودرو آزمایش گردیده‌اند. ۶۲۵ تصویر برای آموزش و باقی تصاویر برای آزمایش به کار گرفته شده و دقت ٪۹۷ گزارش شده است.

[۱] Gaussian

[۲] Discrete Cosine Transform (DCT)



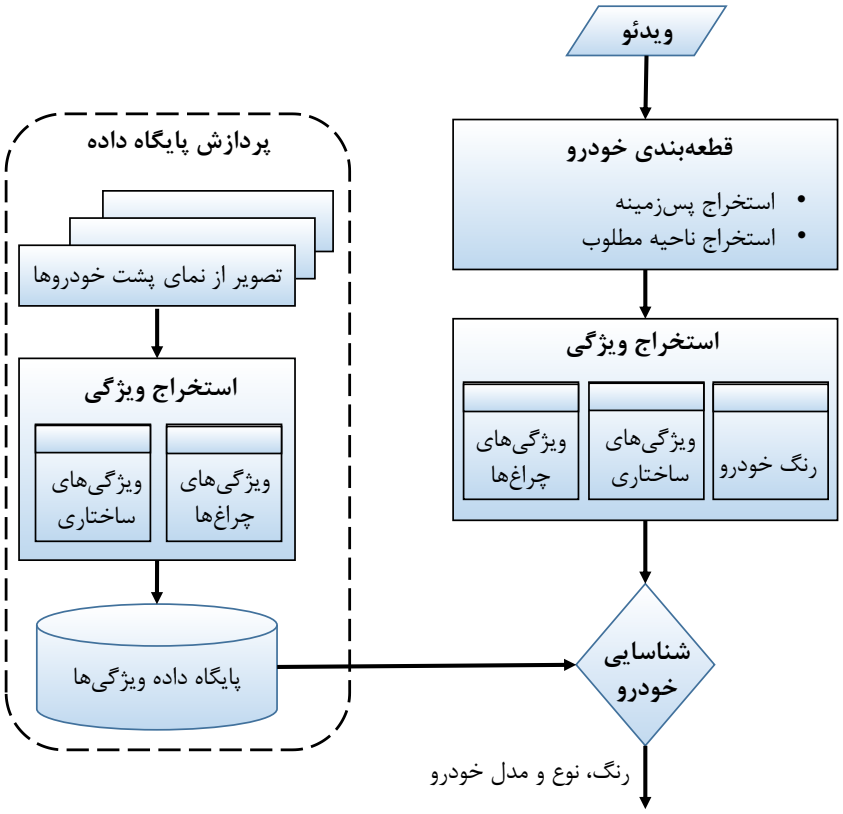
شکل ۲-۲۰: فلوچارت مراحل سیستم پیشنهادی در [۵۹].

۱. نشان‌واره **۲. ناحیه مطلوب مبتنی بر نشان‌واره** **۳. ناحیه مطلوب استخراج شده**

شکل ۲-۲۱: استخراج مکان نشان‌واره و یک ناحیه از سمت چپ آن در [۵۹].

سانتوس و همکاران روشی برای شناسایی نوع، مدل و رنگ خودرو پیشنهاد کرده‌اند که تا حدودی، جزئی‌نگر عمل می‌کند [۶۰]. این روش علاوه بر تشخیص ناحیه مطلوب (مانند روش‌های کلی‌نگر)، مکان چراغ‌های عقب را نیز استخراج می‌کند. برای این منظور، فرض می‌کند که رنگ چراغ‌های عقب در همه‌ی خودروها قرمز است. برای تشخیص رنگ خودرو، از رنگ‌های موجود در ناحیه مطلوب و در کانال رنگی HSV رأی‌گیری صورت می‌گیرد. پس از استخراج پلاک، ناحیه مطلوب و چراغ‌های عقب، با استفاده از لبه‌های ناحیه مطلوب و چراغ‌ها، بردار ویژگی‌ها ایجاد می‌شود. سپس شکل و زاویه‌ی چراغ‌ها نسبت به پلاک استخراج می‌گردد. از هیچ روش استخراج ویژگی خاصی استفاده نشده است (از لبه‌ها به‌صورت مستقیم استفاده می‌شود). شکل ۲-۲۲ فلوچارت مراحل سیستم را نمایش می‌دهد. برای آموزش سیستم، ۴۰ تصویر از نمای پشت ۱۸ مدل متفاوت از خودروها به کار گرفته شده است. برای آزمایش سیستم، ۱۸

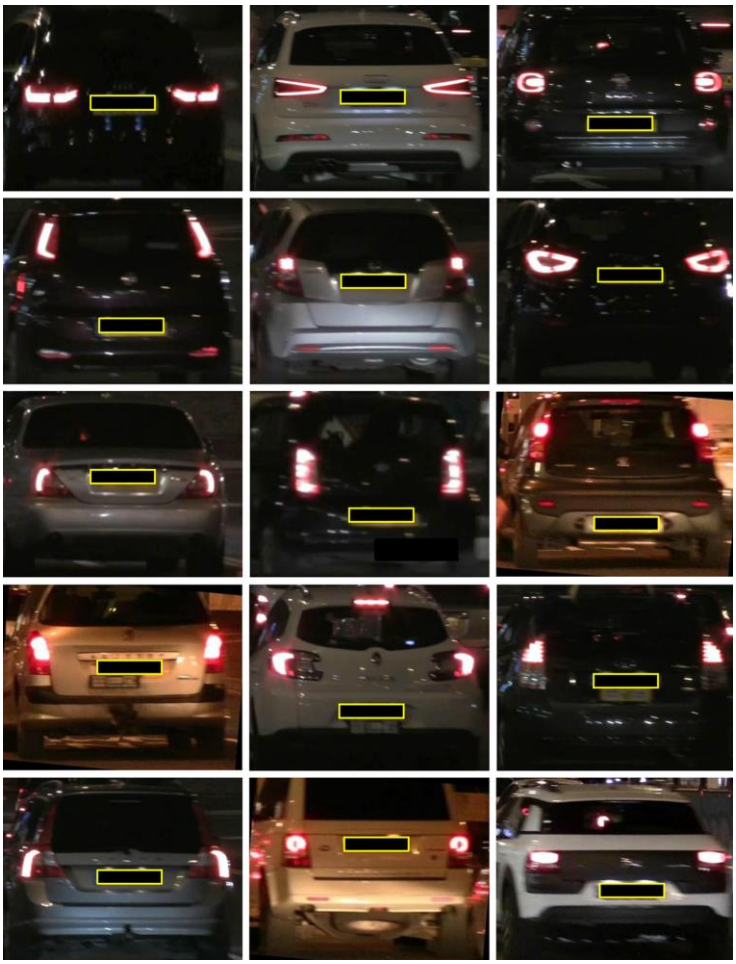
ویدئو از ورود ۱۸ خودرو به یک پارکینگ تهیه شده است (هر ویدئو یک خودرو). دقت گزارش شده با استفاده از طبقه‌بند k-نزدیک‌ترین همسایگی‌ها برابر با ٪۸۹ بوده است. دقت بخش تشخیص رنگ نیز برابر با ٪۱۰۰ گزارش شده است.



شکل ۲-۲۲: فلوچارت مراحل سیستم پیشنهادی در [۶۰].

بوسیم و همکارش نیز روشی مبتنی بر استفاده از چراغ‌های عقب ارائه کردند [۱۸]. آن‌ها روشی برای شناسایی نوع و مدل خودرو در شب پیشنهاد کرده‌اند و چون شب‌ها، چراغ‌های جلوی خودرو مانع استخراج ویژگی مناسب می‌گردد، از نمای عقب خودرو بهره برده‌اند. با این فرض که شکل چراغ‌های عقب، مکان پلاک و زاویه چراغ‌ها نسبت به پلاک برای هر مدل از خودروها یکتا است. آن‌ها همچنین از الگوریتم ژنتیک برای انتخاب زیرمجموعه‌ای از ویژگی‌ها استفاده کرده‌اند. از سه طبقه‌بند ماشین بردار پشتیبان، درخت تصمیم[۱] و k-نزدیکترین همسایگی‌ها به روش رأی اکثریت برای تعیین کلاس خروجی استفاده شده است. دقت گزارش شده روی ۷۶۶ تصویر از ۴۲۱ نوع و مدل متفاوت که نمونه‌های آن در شکل ۲-۲۳ نمایش داده شده، برابر با ٪۹۳/۸ بوده است.

[۱] Decision Tree



شکل ۲-۲۳: نمونه تصاویری از مجموعه داده مورد استفاده در [۱۸].

استفاده از چراغ‌ها در امر شناسایی مدل خودرو بسیار معمول است. در همین راستا، هی و همکارانش نیز از چراغ‌ها و پلاک بهره برده‌اند ولی نه برای استخراج ویژگی، بلکه برای تنظیم راستای درست خودرو در تصویر [۱۹]. آن‌ها پس از پیش‌پردازش تصویر به روش مورد اشاره، با استفاده از الگوریتم DoG[۱] به استخراج ویژگی پرداخته‌اند. بردارهای ویژگی استخراج شده به یک شبکه عصبی آموزش داده شده و دقت ٪۹۲/۴۷ برای ۱۱۹۶ خودرو از ۳۰ نوع و مدل متفاوت گزارش شده است. تصاویر همگی روبه‌جلو، در شرایط نوری متفاوت (روز و شب) و با تغییرات زاویه اندک بوده‌اند.

سرفراز و خان روشی مبتنی بر یافتن بخش‌های متمایزکننده ارائه کرده‌اند [۶۱]. این روش در مرحله‌ی آموزش، تعدادی زیرپنجره را جستجو کرده و از هر یک، ویژگی هیستوگرام شکل انرژی محلی را استخراج

[۱] Difference of Gaussian

می‌کند. سپس زیرپنجره‌هایی که بیشترین شباهت را در کلاس‌های مشابه و بیشترین تفاوت را در کلاس-های متفاوت به دست می‌دهند، نگه داشته می‌شوند (شکل ۲-۲۴). شباهت دو پنجره با استفاده از معیار شباهت واگرایی K-L[۱] محاسبه می‌گردد (فرمول (۲-۴)).

$$D_{KL}(x_1, x_2) = \sum_i x_1 log \frac{x_1}{x_2} \quad \textbf{(۴-۲)}$$

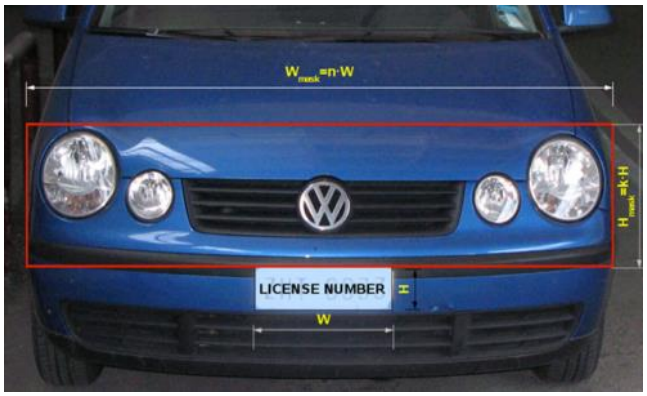
x_1 و x_2 بردارهای ویژگی استخراج شده توسط هیستوگرام شکل انرژی محلی از دو بخش مختلفِ تصویر می‌باشند. در مرحله‌ی آزمایش، مجموعه پنجره‌های منتخب از تصویر ورودی و تصاویر ثبت شده در سیستم با یکدیگر مقایسه می‌شوند؛ تصویری که مجموعه پنجره‌های آن بیشترین شباهت را به مجموعه پنجره‌های تصاویر ورودی داشته باشد، به عنوان خروجی انتخاب می‌گردد. آزمایشات بر روی مجموعه تصاویری متشکل از ۳۸ کلاس مختلف که از هر یک حداقل ۱۵ نمونه تصویر از نمای جلو و پشت وسیله نقلیه وجود دارد، صورت گرفته است. دقت گزارش شده برای نمای پشت، ٪۶۲ و برای نمای جلو، ٪۴۸ بوده است.

شکل ۲-۲۴: یک نمونه پنجره تأیید شده (سمت چپ) و یک نمونه پنجره‌ی رد شده (سمت راست) توسط روش پیشنهادی در [۶۱].

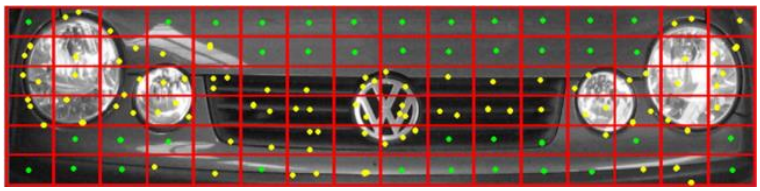
سایلس و همکاران، کار قبلی خود [۵۷] را بهبود بخشیده و یک بخش جدید به آن اضافه کردند [۶۲]. در این مقاله، علاوه بر تشخیص نشان‌واره برای تعیین نوع خودرو، از ناحیه مطلوب (شکل ۲-۲۵) برای شناسایی مدل خودرو استفاده شده است. برای استخراج ویژگی از تبدیل ویژگی مقاوم به اندازه به‌صورت چندبخشی بهره گرفته شده است (شکل ۲-۲۶). سپس نقاط کلیدی به شکلی فیلتر شده‌اند که در هر بخش تنها یک نقطه‌ی کلیدی باقی بماند. به مرکز بخش‌هایی که دارای هیچ نقطه‌ی کلیدی نیستند، یک نقطه کلیدی اضافه می‌شود (نقاط سبز رنگ در شکل ۲-۲۶). در نهایت، این نقاط کلیدی در کنار هم قرار

[۱] K-L Divergence

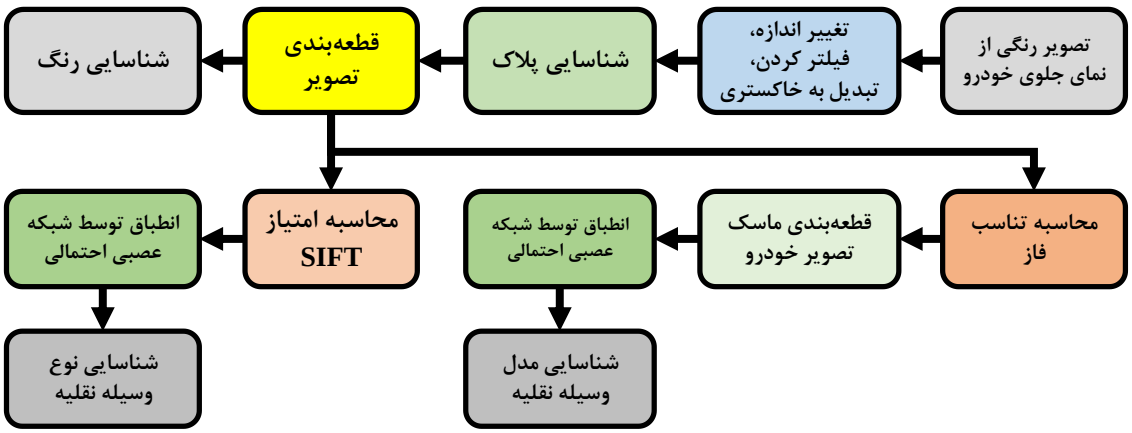
داده شده و بردار ویژگی‌های نهایی را تشکیل می‌دهند. از یک شبکه‌ی عصبی احتمالی [۶۳] برای طبقه-بندی استفاده شده است. آزمایشات بر روی مجموعه داده‌ای از ۱۱۰ تصویر از ده کلاس مختلف و یک کلاس ناشناخته صورت گرفته‌اند. دقت گزارش شده برای شناسایی نوع ٪۸۵ و برای شناسایی مدل ٪۵۴ بوده است. شکل ۲-۲۷ فلوچارت مراحل سیستم را نشان می‌دهد.



شکل ۲-۲۵: روش استخراج ناحیه مطلوب در [۶۲].

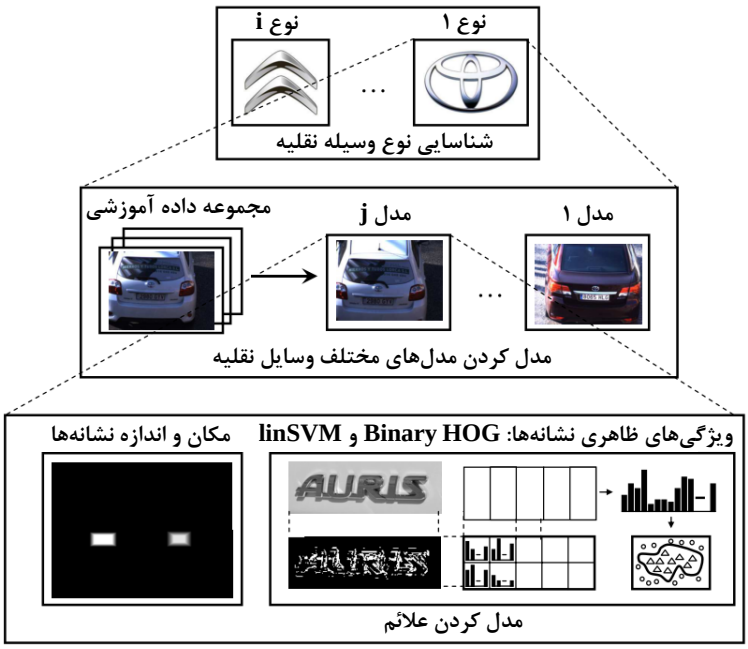


شکل ۲-۲۶: استخراج نقاط کلیدی توسط تبدیل ویژگی مقاوم به اندازه از ناحیه مطلوب [۶۲]. نقاط سبز رنگ به مرکز بخش‌هایی که دارای هیچ نقطه‌ی کلیدی نبوده‌اند، اضافه شده‌اند.



شکل ۲-۲۷: فلوچارت مراحل سیستم ارائه شده در [۶۲].

ایده‌ی یورکا و همکاران، استفاده از مکان، اندازه و مشخصه‌ی ظاهری علائم[1] (نوشته‌ها) پشت خودروها برای شناسایی نوع و مدل آن‌ها است [۶۴]. در این روش، برای علائم مربوط به هر کلاس از خودروها، یک مدل یاد گرفته می‌شود؛ هر مدل متشکل از اندازه و مکان میانگین علائم و بردار ویژگی‌های استخراج شده توسط روش هیستوگرام گرادیان‌های جهت‌دار است که به طبقه‌بند ماشین بردار پشتیبان آموزش داده می‌شود (شکل ۲-۲۸). در مرحله آزمایش، خروجی طبقه‌بندهای مختلف، محاسبه شده و بالاترین احتمال انتخاب می‌شود. برای تشخیص نوع خودرو از نشان‌واره و برای شناسایی مدل خودرو از علائم پشت خودروها استفاده می‌شود. مجموعه داده‌ی تهیه شده از ۱۳۴۲ تصویر از ۵۲ کلاس مختلف تشکیل می‌شود که ۹۱۰ نمونه از آن برای آموزش و ۴۳۲ نمونه از آن برای آزمایش به کار گرفته شده است. دقت به دست آمده برای این مجموعه داده برابر با ٪۹۳/۷۵ گزارش شده است. برای از بین بردن تأثیر مکان-یابی پلاک و علائم بر روی دقت سیستم نهایی، این دو بخش به‌صورت دستی انجام گرفته‌اند.



شکل ۲-۲۸: مراحل عملکرد سیستم پیشنهادی در [۶۴].

استفاده از شبکه‌های عصبی عمیق به دلیل قدرت بالای این ابزار، در سال‌های اخیر افزایش پیدا کرده است. به ویژه پس از ارائه چندین کتابخانه‌ی قدرتمند برای آموزش شبکه‌های عصبی عمیق مانند OverFeat [۶۵]، Caffe [۶۶] و Matconvnet [۶۷] و وجود شبکه‌های از پیش آموزش دیده شده.

[1] Emblems

رویکرد مبتنی بر بخش استفاده شده در این رساله از نظر ساختاری مانند یک شبکه عصبی عمیق عمل می‌کند [۶۸]. بنابراین می‌توان شبکه عصبی عمیق را نیز یک روش مبتنی بر بخش و در نتیجه یک روش جزئی‌نگر در نظر گرفت. چند روش مبتنی بر این ابزار در حوزه شناسایی مدل وسیله نقلیه ارائه شده‌اند. به عنوان نمونه، هوتونن و همکاران عملکرد شبکه عصبی عمیق را با ماشین بردار پشتیبان برای شناسایی دسته‌ی کلی خودرو مقایسه کرده‌اند [۱۰]. برای آموزش ماشین بردار پشتیبان از هیستوگرام حاصل از کیسه‌ای از کلمات بصری بهره برده شده است. کیسه‌ای از کلمات با استفاده از بردار ویژگی‌های بدست آمده از عملگر تبدیل ویژگی مقاوم به اندازه ایجاد گشته است. برای آموزش شبکه عصبی عمیق نیز از کل تصویر ورودی بدون استخراج ویژگی استفاده شده است. مجموعه داده مورد استفاده در این مقاله از ۶۵۵۵ تصویر روبه‌جلو از دو دوربین مدار بسته متفاوت تهیه شده که چهار دسته‌ی کلی اتوبوس، کامیون، سواری و وَن را شامل می‌شود (شکل ۲-۲۹). ۹۰ درصد تصاویر برای آموزش و مابقی برای آزمایش به کار گرفته شده‌اند. دقت گزارش شده برای شبکه عصبی عمیق برابر با ٪۹۸/۰۶ و برای ماشین بردار پشتیبان برابر با ٪۹۷/۷۵ بوده است که نشان از برتری شبکه عصبی دارد.

شکل ۲-۲۹: نمونه تصاویری از مجموعه داده مورد استفاده در [۱۰].

مرجع [۶۹] ابتدا با استفاده از تفاضل فریم‌ها، محدوده‌ی خودرو را در تصویر تشخیص می‌دهد. سپس با بهره‌گیری از یک فیلتر متقارن، ناحیه مطلوب استخراج می‌گردد. تصاویر دودویی حاصل از ناحیه مطلوب به شبکه عصبی عمیق آموزش داده می‌شوند. مجموعه داده استفاده شده در این مقاله از ۳۲۱۰ تصویر تشکیل شده که ۱۰۷ مدل مختلف از خودروها را شامل می‌شوند. از هر مدل، ۳۰ تصویر موجود است که ۲۹ تصویر برای آموزش و یک تصویر باقیمانده برای آزمایش به کار رفته است. دقت گزارش شده برابر با

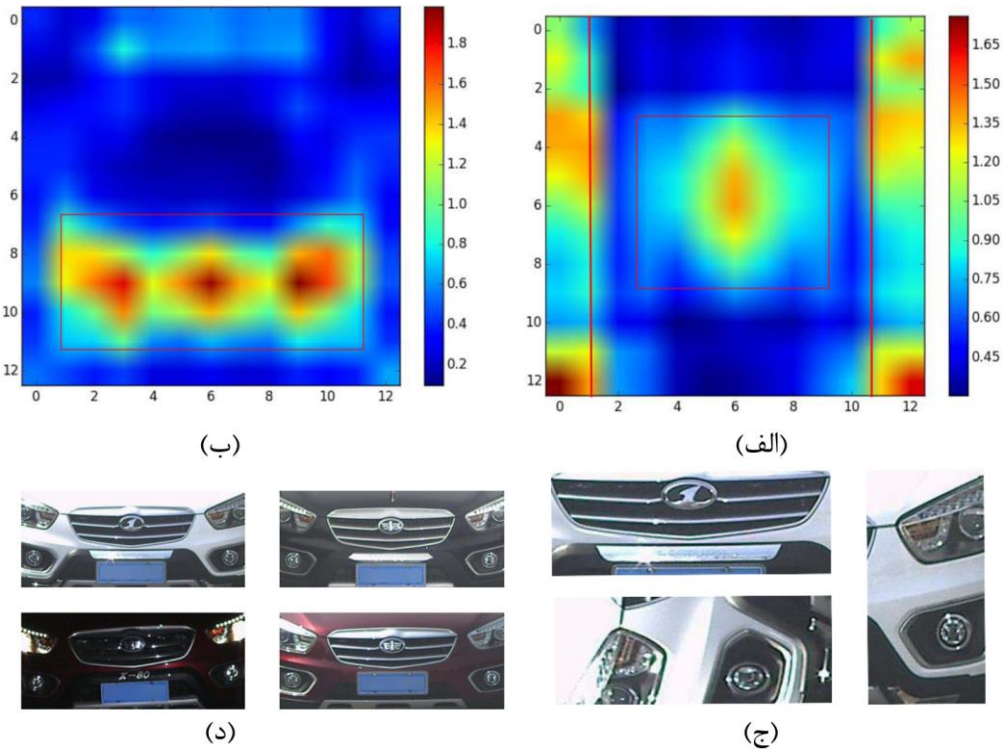
٪۸۸/۲ می‌باشد. نویسندگان این مقاله، کار قبلی خود را بهبود بخشیده و به جای آموزش مستقیم تصاویر دودویی ناحیه مطلوب به شبکه عصبی، ویژگی‌های استخراج شده با استفاده از عملگر هیستوگرام گرادیان‌های جهت‌دار را به شبکه آموزش داده‌اند [۷۰]. این امر موجب بهبود دقت به ٪۹۸/۵ به بهای کاهش دو برابری سرعت گشته است.

لی و همکاران نیز در رویکردی مشابه از شبکه عصبی عمیق بهره برده‌اند [۱۷]. آن‌ها پس از استخراج ناحیه مطلوب به روشی ساده، الگوریتم PCA را به تصویر اعمال کرده و یک ماتریس جدید بدست آورده‌اند. آنگاه از هر بلوک، یک هیستوگرام استخراج شده و مقادیر دهدهی ماتریس ورودی در قالب تعدادی ستون مشخص، دسته‌بندی می‌شوند. بردار ویژگی نهایی از اتصال این هیستوگرام‌ها حاصل می‌گردد. با ارسال بردار ویژگی برای شبکه عصبی عمیق، ضرایب یک لایه مشخص از شبکه استخراج شده و به ماشین بردار پشتیبان آموزش داده می‌شوند. این تحقیق از دو مجموعه داده خصوصی با ۴۰۰۰ تصویر از ۲۲ مدل و ۲۰۰۰ تصویر از ۲۱ مدل خودرو بهره برده و دقت‌های ٪۹۵/۴۸ و ٪۹۵/۸۴ را گزارش کرده است.

استفاده از شکل سه‌بعدی خودرو و همچنین جهت خودرو موجب بهبود دقت شبکه عصبی عمیق می‌گردد. این ادعایی است که هروت و همکاران در [۷۱] مطرح کرده‌اند. مجموعه داده مورد استفاده توسط این تحقیق متشکل از ۲۱۲۵۰ خودرو از ۱۴۸ نوع و مدل متفاوت بوده است (شکل ۲-۳۰). دقت گزارش شده برای این مجموعه داده برابر با ٪۷۳/۱ بوده است که بهبودی در حدود ٪۶ نسبت به حالتی که از شکل سه‌بعدی و جهت خودرو استفاده نمی‌شود، به همراه داشته است.

شکل ۲-۳۰: نمونه تصاویری از مجموعه داده مورد استفاده در [۷۱].

یک ایده‌ی مبتنی بر بخش با استفاده از شبکه عصبی عمیق در [۷۲] مطرح شده است. در این مقاله ابتدا به آموزش یک شبکه عصبی عمیق پرداخته می‌شود. سپس با استخراج نقشه ویژگی[۱] از آخرین لایه شبکه، بخش‌هایی که بیشترین تأثیر در طبقه‌بندی را داشته‌اند، شناسایی می‌شوند. آن بخش‌ها از تصاویر استخراج شده و مجددا به یک شبکه‌ی عمیق دیگر آموزش داده می‌شوند. این روال آنقدر تکرار می‌شود تا بخش مناسب دیگری یافت نشود. نقشه ویژگی به خروجی یک لایه از شبکه، اطلاق می‌شود. شکل ۲-۳۱ دو نمونه نقشه ویژگی و بخش‌های متناظر با آن‌ها را نمایش داده است.



شکل ۲-۳۱: روش تشخیص بخش‌های ارائه شده در [۷۲]. تصویر (الف) نقشه ویژگی استخراج شده از لایه آخر شبکه عصبی عمیق را نشان می‌دهد. تصویر (ب) نقشه ویژگی مربوط به شبکه‌ی آموزش دیده شده روی بخش‌های استخراج شده از تصویر (الف) را نمایش می‌دهد؛ تصاویر (ج) و (د)، بخش‌های متناظر با نقشه‌های ویژگی بخش‌های (الف) و (ب) را ارائه کرده‌اند.

مقاله مورد بحث پس از تشخیص بخش‌ها و آموزش یک شبکه عصبی عمیق به ازای هریک از آن‌ها، از ماشین بردار پشتیبان به شکل یک در برابر همه برای طبقه‌بندی بهره می‌برد. وزن‌های استخراج شده از لایه ماقبل آخر شبکه عصبی به عنوان بردارهای ویژگی در نظر گرفته شده و به ماشین بردار پشتیبان

[۱] Feature Map

آموزش داده می‌شوند. با استخراج ویژگی از تمامی شبکه‌های عصبی و کنار هم قرار دادن آن‌ها، بردار ویژگی نهایی تشکیل شده و به ۲۸۰ ماشین بردار پشتیبان آموزش داده می‌شود. علت استفاده از ۲۸۰ طبقه‌بند، وجود ۲۸۰ کلاس در مجموعه داده استفاده شده (CompCars) توسط مقاله است. دقت میانگین و مجموع گزارش شده روی این مجموعه داده برابر با ٪۹۸/۲۹ و ٪۹۸/۶۳ می‌باشد. منطق استفاده شده در این مقاله شباهت‌هایی به روش پیشنهادی در رساله‌ی جاری دارد ولی تفاوت‌های بارزی نیز دیده می‌شود. هر دو رویکرد در جستجوی بخش‌های متمایز کننده هستند؛ هر چند روش ارائه شده در [۷۲] تعدادی بخش مشترک بین همه کلاس‌ها می‌یابد؛ و برای یافتن بخش‌های متمایز کننده از شبکه عصبی عمیق بهره می‌برد.

۲-۴- جمع‌بندی

به منظور مقایسه‌ی بهتر روش‌های ارائه شده در این حوزه، خلاصه‌ی روش‌های معرفی شده در دو بخش قبل در قالب یک جدول ارائه شده‌اند. در جدول ۲-۱، بخش‌های اصلی روش‌های مختلف مانند روش استخراج ناحیه مطلوب، استخراج ویژگی، طبقه‌بندی، دقت کسب شده و مجموعه داده‌ی مورد استفاده، جمع‌آوری گشته و ارائه شده‌اند.

جدول ۲-۱: جمع‌بندی و مقایسه‌ی روش‌های پیشین.

مرجع	استخراج ناحیه مطلوب	استخراج ویژگی	طبقه‌بندی	مجموعه داده	دقت گزارش شده
کونوس [۳۱]	مکان و اندازه پلاک	تبدیل ویژگی مقاوم به اندازه به صورت چندبخشی و همپوشان	k-NN و فاصله اقلیدسی	نمای جلو سواری: ۱۰۰ تصویر از ده کلاس وسیله نقلیه سنگین: ۵۰۰ تصویر از هفت کلاس	سواری: ٪۹۴ سنگین: ٪۹۲
پتروویچ [۴۰]	مکان و اندازه پلاک	گرادیان‌های نگاشت شده‌ی مربعی	k-NN و فاصله اقلیدسی	نمای جلو ۱۱۳۲ تصویر از ۷۷ کلاس	٪۹۳
پیرس [۳۸]	مکان و اندازه پلاک	تشخیص گوشه‌ی هریس به صورت محلی و بخش‌بندی شده	بیز ساده و ارزیابی ضربدری Leave one out	نمای جلو ۲۶۲ تصویر از ۷۲ کلاس	٪۹۶
مونرو [۴۲]	تشخیص لبه‌ی کنی و چراغ‌های جلو	لبه‌ها و فاصله لبه‌ها تا مرکز ناحیه مطلوب	k-NN و شبکه عصبی پرسپترون چندلایه و ارزیابی ضربدری به پیمانه ده	نمای جلو ۱۵۰ تصویر از پنج کلاس ۳۰ تصویر از کلاس متفرقه	k-NN: ٪۹۹/۹۹ شبکه عصبی: ٪۹۹/۵۳

کاظمی [۴۳]	کل تصویر از نمای پشت	تبدیل پیچک	k-NN و ماشین بردار پشتیبان	نمای پشت ۳۰۰ تصویر از پنج کلاس	k-NN: ٪۹۲ SVM: ٪۹۹
ظفر [۴۵]	برش دستی	تبدیل کانتورلت	ماشین بردار پشتیبان	نمای جلو ۳۰۰ تصویر از ۲۵ کلاس	٪۹۶
شیبِه [۴۸, ۴۹]	ویژگی‌های مقاوم تسریع شده متقارن	هیستوگرام گرادیان‌های جهت‌دار و ویژگی‌های مقاوم تسریع شده به صورت چندبخشی و مستقل و سپس رأی‌گیری	ماشین بردار پشتیبان	نمای جلو ۶۹۳۶ تصویر از ۲۹ کلاس	٪۹۹
ناظمی [۵۰]	کل تصویر از نمای جلو	تبدیل ویژگی مقاوم به اندازه	ماشین بردار پشتیبان	نمای جلو ۲۵۰ تصویر از ده کلاس	٪۹۳/۲۰
صدیقی [۵۱]	برش دستی	کیسه‌ای از کلمات و ویژگی‌های مقاوم تسریع شده	ماشین بردار پشتیبان	نمای جلو ۶۵۰۰ تصویر از ۲۹ کلاس	٪۹۵/۷۷
ساروی [۳۹]	مکان و اندازه و جهت پلاک	هیستوگرام شکل انرژی محلی	ماشین بردار پشتیبان	نمای جلو ۱۹۶ تصویر از ۲۲ کلاس و سه ویدئو به طول ۱-۲ دقیقه	٪۹۵/۸۳
شینوزوکا [۵۳]	برش دستی	تبدیل ویژگی مقاوم به اندازه	k-NN و معیار کسینوسی	نمای جلو ۶۰ تصویر از نه کلاس	٪۱۰۰ با k=5
باران [۵۴]	الگوریتم تشخیص شیء ویولا و جونز	تبدیل ویژگی مقاوم به اندازه و ویژگی‌های مقاوم تسریع شده و هیستوگرام لبه	k-NN	نمای جلو ۳۵۸۹ تصویر از ۱۷ کلاس	٪۹۷/۲
سایلس [۵۷]	نشان‌واره	تبدیل ویژگی مقاوم به اندازه	ماشین بردار پشتیبان به صورت چندبخشی و همپوشان	۱۲۰۰ نشان‌واره از ده کلاس	٪۹۱
یانگ [۵۹]	نشان‌واره و ناحیه‌ای از سمت چپ آن	تبدیل کسینوسی گسسته	ماشین بردار پشتیبان	نمای جلو ۱۰۹۶ تصویر از ۱۲ کلاس	٪۹۷
سانتوس [۶۰]	ناحیه مطلوب و چراغ‌های عقب	لبه‌های ناحیه مطلوب، چراغ‌ها و زاویه چراغ‌ها با پلاک	k-NN و معیار شباهت سفارشی	نمای پشت ۴۰ تصویر از ۱۸ کلاس و ۱۸ ویدئو از ورود ۱۸ خودرو به پارکینگ	٪۸۹
بوسیم [۱۸]	مکان و اندازه پلاک و چراغ‌های عقب	شکل چراغ‌های عقب، مکان پلاک و زاویه چراغ‌ها نسبت به پلاک	رای‌گیری بین ماشین بردار پشتیبان، درخت تصمیم و k-نزدیکترین	نمای پشت ۷۶۶ تصویر از ۴۲۱ کلاس	٪۹۳/۸

		همسایگی‌ها			
هی [۱۹]	مکان پلاک و چراغ‌های جلو	روش DoG	شبکه عصبی	نمای جلو ۱۱۹۶ تصویر از ۳۰ کلاس	۹۲/۴۷٪
سرفراز [۶۱]	زیرپنجره‌های متمایز کننده	هیستوگرام شکل انرژی محلی	k-NN و معیار شباهت واگرایی K-L	نمای جلو و پشت حداقل ۵۷۰ تصویر از ۳۸ کلاس	جلو: ۴۸٪ پشت: ۶۲٪
سایلس [۶۲]	نشان‌واره و مکان و اندازه پلاک	تبدیل ویژگی مقاوم به اندازه چندبخشی	شبکه عصبی احتمالی	نمای جلو ۱۱۰ تصویر از ده کلاس	۵۴٪
یورکا [۶۴]	علائم پشت خودروها	هیستوگرام گرادیان‌های جهت‌دار	ماشین بردار پشتیبان	نمای پشت ۱۳۴۲ تصویر از ۵۲ کلاس	۹۳/۷۵٪
هوتونن [۱۰]	کل تصویر	کیسه‌ای از کلمات و تبدیل ویژگی مقاوم به اندازه	شبکه عصبی عمیق	نمای جلو ۶۵۵ تصویر از ۴ کلاس	۹۸/۰۶٪
جاو [۶۹]	تفاضل فریم‌ها	تصویر دودویی	شبکه عصبی عمیق	نمای جلو ۳۲۱۰ تصویر از ۱۰۷ کلاس	۸۸/۲٪
جاو [۷۰]	تفاضل فریم‌ها	هیستوگرام گرادیان‌های جهت‌دار	شبکه عصبی عمیق	نمای جلو ۳۲۱۰ تصویر از ۱۰۷ کلاس	۹۸/۵٪
لی [۱۷]	ناحیه مطلوب	ضرایب یک لایه مشخص از شبکه عصبی	شبکه عصبی عمیق	نمای جلو ۴۰۰۰ تصویر از ۲۲ کلاس ۲۰۰۰ تصویر از ۲۱ کلاس	۹۵/۴۸٪ ۹۵/۸۴٪
هروت [۷۱]	کل تصویر	پیکسل‌های خام و شکل سه‌بعدی و جهت خودرو	شبکه عصبی عمیق	همه نماها ۲۱۲۵۰ تصویر از ۱۴۸ کلاس	۷۳/۱٪
فنگ [۷۲]	کل تصویر و بخش‌های تشخیص داده شده	ضرایب لایه ماقبل آخر شبکه عصبی	شبکه عصبی عمیق	نمای جلو ۵۰۰۰۰ تصویر از ۲۸۰ کلاس	۹۸/۹۶٪

همان‌طور که در جدول ۲-۱ قابل مشاهده است، اغلب روش‌های موجود برای شناسایی نوع و مدل وسیله نقلیه، یک ناحیه از تصویر را به عنوان ناحیه مطلوب استخراج کرده و از آن به استخراج ویژگی می‌پردازند. ممکن است ناحیه مطلوب ابتدا به زیربخش‌هایی تقسیم شده و از هر یک مستقلاً ویژگی‌هایی استخراج گردد. سپس با استفاده از یک یا چند طبقه‌بند، طبقه‌بندی صورت می‌گیرد. ویژگی‌های متفاوت و طبقه-بندهای متفاوتی برای این منظور مورد آزمون قرار گرفته‌اند. مشکل عمده‌ی اغلب این روش‌ها، رویکرد مورد استفاده توسط آن‌ها است. برای پیشرفت قابل ملاحظه در این مسئله، نیاز به رویکردی با انطباق بیشتر بر ویژگی‌های خاص مسئله است. برخی از مشکلات ناشی از رویکرد روش‌های موجود در ادامه تشریح شده‌اند.

مشکل اول، استفاده از دانش قبلی انسان در تشخیص ناحیه مطلوب است. معمولاً ناحیه‌ی از پیش تعیین‌شده‌ای از اطراف پلاک به عنوان ناحیه مطلوب استخراج می‌گردد. این فرض در شرایطی منجر به نتایج اشتباه می‌گردد. برای مثال، پلاک در همه خودروها در وسط قرار ندارد. از طرفی، کلیت موضوع، یعنی کارایی استفاده از ناحیه مطلوب برای طبقه‌بندی وسایل نقلیه، دارای اشکال است. ناحیه مطلوب می‌تواند دارای بخش‌های غیرکاربردی باشد که منجر به کاهش دقت طبقه‌بندی و افزایش طول بردار ویژگی‌ها می‌گردد. در صورتی که بتوان بخش‌های مؤثر و کاربردی را بدون اعمال دانش قبلی انسان و به صورت پویا استخراج کرد، می‌توان از تأثیرات منفی این رویکرد جلوگیری کرد.

مشکل بعدی، حساسیت بالا نسبت به بخش‌های تغییریافته در مدل‌های یکسان وسایله نقلیه است. برای مثال، مالکین بعضی خودروها(ی داخلی)، چراغ‌های عقب را با مدل‌های غیرشرکتی تعویض می‌کنند و یا برخی بخش‌ها استعداد بیشتری برای ضربه خوردن در تصادفات و برخوردهای جزئی خودرو دارند؛ چنان‌که می‌بینیم وجود شکستگی در چراغ‌های جلو کم نیست. در نتیجه، تنها با مقایسه‌ی ویژگی‌های استخراج شده از ناحیه مطلوب تصویر ورودی و تصویر ثبت شده در سیستم، نمی‌توان تطابقی با دقت بالا بدست آورد. درصورتی‌که سیستم بتواند اهمیت چنین بخش‌هایی را کاهش داده و آن را در مرحله استخراج و انطباق بخش تأثیر دهد، می‌توان حساسیت نسبت به بخش‌های غیرهم‌شکل را کمتر کرد. چراغ‌های جلو در خودروهای «پراید ۱۳۱» را بخشی متمایز کننده فرض کنید. سیستم می‌تواند در مرحله پردازش تصاویر آموزشی این خودروها تشخیص دهد که تغییر شکل در چراغ‌های جلوی این مدل به قدری زیاد است که این بخش برای استفاده در طبقه‌بندی مناسب نیست؛ بنابراین آن را از مجموعه بخش‌های متمایز کننده حذف و یا اهمیت آن را کاهش می‌دهد.

تراز نبودن بخش‌های متناظر در دو تصویر، مشکل بعدی است. برای مثال روش‌هایی که از چراغ‌ها برای جداسازی کلاس‌های متفاوت بهره می‌برند، به تراز بودن ناحیه مستطیلی یافته شده برای هر چراغ دقت نمی‌کنند. انطباق و تراز کردن بخش‌ها قبل از مقایسه از اهمیت بالایی برخوردار است. در صورتی که دو بردار استخراج شده، برهم منطبق نباشند، مقایسه‌ی بردارهای ویژگی با استفاده از روش‌هایی که تشابه را به صورت یک‌به‌یک می‌سنجند، نتیجه‌ی مناسبی به دست نخواهد داد. این مشکل می‌تواند به حساسیت نسبت به تغییرات زوایه اندک تصویر نیز تعمیم داده شود؛ زیرا عدم مقابله با تغییرات زاویه اندک تصویر می‌تواند منجر به تراز نبودن بخش‌های متناظر شود.

حساسیت نسبت به تغییرات روشنایی نیز از مشکلات دیگر موجود در روش‌های ارائه شده است؛ که البته این مشکل ارتباط چندانی به رویکرد مورد استفاده توسط روش‌ها ندارد. تغییرات روشنایی از چالش‌های

بسیار معمول و درعین‌حال دشوار در این حوزه است. با استفاده از یک روش پیش‌پردازش مناسب و یا یک الگوریتم استخراج ویژگی مقاوم به تغییرات روشنایی، می‌توان تا حد زیادی، حساسیت نسبت به تغییرات روشنایی را کاهش داد.

فصل

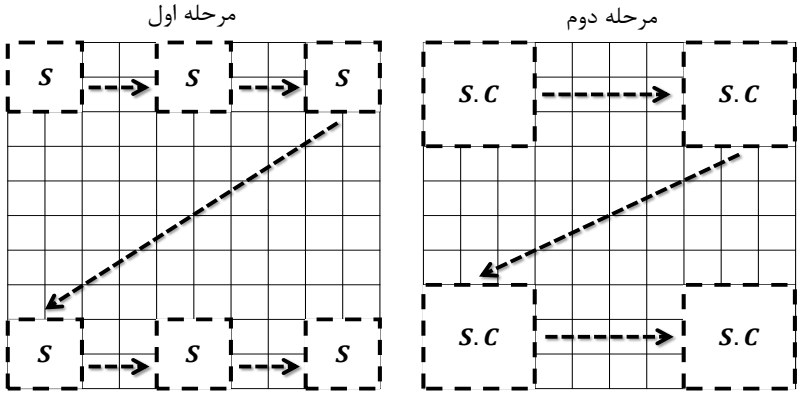
۳- تشریح الگوریتم‌های استفاده شده در روش پیشنهادی

۳-۱- مقدمه

در پیاده‌سازی سیستم مورد بحث در این رساله از تعدادی از الگوریتم‌های شناخته شده‌ی پردازش تصویر و بینایی ماشین بهره گرفته شده است. در این فصل به تشریح این الگوریتم‌ها پرداخته شده تا نیازی به توضیح عملکرد آن‌ها در فصل‌های پیش‌رو نباشد.

۳-۲- هرم تصاویر

برای تشخیص اشیاء در تصاویر، باید اندازه‌های متفاوت اشیاء را در نظر گرفت. برای تشخیص همه اشیاء موجود در تصویر با اندازه‌های متفاوت معمولاً از یکی از دو تکنیک پنجره‌های لغزان[۱] و هرم تصاویر[۲] بهره برده می‌شود. این دو تکنیک کاربردهای فراوانی داشته و در الگوریتم‌های متعددی به کار گرفته شده‌اند [۲۶, ۵۵, ۷۳, ۷۴]. در تکنیک پنجره‌های لغزان، یک پنجره با اندازه اولیه S بر روی تصویر لغزانده می‌شود. سپس این پنجره در مقدار ثابت و مثبت C ضرب شده و بزرگتر می‌گردد (شکل ۳-۱). این روال آنقدر تکرار خواهد شد تا اندازه پنجره از مقدار بیشینه M بزرگتر شود. در طرف مقابل، در تکنیک هرم تصاویر، این تصویر است که تغییر اندازه داده می‌شود و اندازه پنجره ثابت در نظر گرفته می‌شود. تصویر اصلی در پایین هرم که سطح صفر نامیده می‌شود، قرار می‌گیرد. با حرکت کردن از پایین هرم به سمت بالای آن، اندازه تصاویر نیز در مقدار مثبت و کوچکتر از یکِ $\frac{1}{C}$ ضرب می‌شود. معمولاً قبل از هر مرتبه کوچک کردن تصویر، ابتدا آن را تار[۳] می‌کنیم.



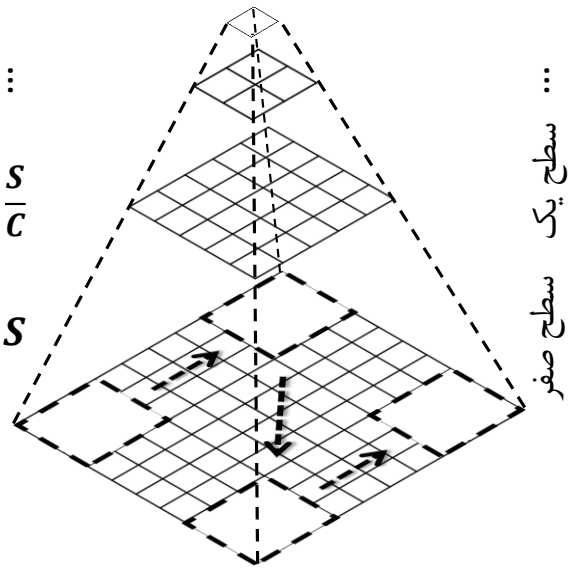
شکل ۳-۱: الگوریتم پنجره‌های لغزان برای پردازش تصویر ورودی در همه اندازه‌ها و همه مکان‌ها.

---

۱ Sliding Windows

۲ Image Pyramids

۳ Blur



شکل ۳-۲: الگوریتم هرم تصاویر برای پردازش تصویر ورودی در همه اندازه‌ها و همه مکان‌ها.

در سیستم پیاده‌سازی شده از الگوریتم هرم تصاویر استفاده شده است. در روش پیشنهادی، ثابت C در مرحله آموزش برابر با $2^{0.2}$ و در مرحله آزمایش برابر با $2^{0.1}$ در نظر گرفته شده است. به این ترتیب، در مرحله آموزش اندازه‌ی تصویر موجود در سطح پنجم هرم برابر است با نصف اندازه تصویر سطح صفرم. این اندازه‌ها به صورت تجربی و بر اساس نتایج گزارش شده در [۲۶, ۷۵] انتخاب شده است.

۳-۳- ماشین بردار پشتیبان مخفی

ماشین بردار پشتیبان استاندارد یک حالت خاص از ماشین بردار پشتیبان مخفی است. بنابراین ابتدا از فرمول ماشین بردار پشتیبان استاندارد آغاز کرده و سپس آن را به فرمول ماشین بردار پشتیبان مخفی بسط می‌دهیم.

مجموعه داده برچسب‌دار D را به صورت زیر در نظر می‌گیریم، به صورتی که x_i نمونه i ام و $y_i \in \{-1,1\}$ برچسب آن است:

$$D = \{< x_1, y_1 >, < x_2, y_2 >, \dots, < x_n, y_n >\}$$

ماشین بردار پشتیبان استاندارد خطی، فرمول (۳-۱) را کمینه می‌کند (در نسخه‌ی حاشیه نرم[۱]). با یافتن $\overrightarrow{w}$ و b مناسب، می‌توان نمونه‌های جدید را طبقه‌بندی کرد. بخش اول از این فرمول، اندازه حاشیه بین

[۱] Soft-Margin

بردارهای پشتیبان را بیشینه کرده و بخش دوم، تابع خطای لولایی[1] استاندارد است. تابع خطای لولایی به نمونه‌هایی که در سمت صحیح صفحه جداکننده قرار نگیرند، خطایی در بازه [۱-۰] نسبت می‌دهد. ثابت C نیز تأثیر خطای لولایی بر هزینه کل را تعیین می‌کند.

$$L = \frac{1}{2}\|\vec{w}\|^2 + C\sum_{i=1}^{n}\max(0,1 - y_i(\vec{w}.\vec{x}_i - b)) \quad (۱-۳)$$

تابع L نسبت به w و b محدب[2] است، بنابراین توسط انواع روش‌های بهینه‌سازی توابع محدب قابل حل می‌باشد. حال طبقه‌بند $f(\vec{x})$ را در نظر بگیرید که نمونه x را به شکل نشان داده شده در فرمول (۳-۲) امتیازدهی می‌کند. با آستانه‌گذاری روی امتیاز بدست آمده توسط این تابع، می‌توان برچسب نمونه x را تعیین کرد.

$$f(\vec{x}) = \max_{z\in Z(x)} \vec{w}.\phi(\vec{x},\vec{z}) \quad (۲-۳)$$

در تابع $f(\vec{x})$ دو مجهول وجود دارد. بردار $\vec{w}$ که صفحه جداکننده را تعیین کرده و بردار $\vec{z}$ که بردار مقادیر مخفی مربوط به نمونه x را تعیین می‌کند. در مسئله‌ی مورد بحث در این رساله که طبقه‌بندی مبتنی بر بخش است، مقادیر مخفی می‌توانند مکان یا تعداد بخش‌ها در نظر گرفته شوند. بنابراین $Z(x)$، مجموعه مقادیر ممکنی است که مکان/تعداد بخش‌ها می‌توانند داشته باشند. تابع $\phi(\vec{x},\vec{z})$ بردار ویژگی نمونه x را متناسب با مکان/تعداد بخش‌های تعیین شده توسط x برمی‌گرداند. برای بدست آوردن بردار $\vec{w}$، مشابه با ماشین بردار پشتیبان استاندارد، تابع زیر را کمینه می‌کنیم:

$$L = \frac{1}{2}\|\vec{w}\|^2 + C\sum_{i=1}^{n}\max(0,1 - y_i f(\vec{x}_i)) \quad (۳-۳)$$

وجود مجموعه $Z(x)$ چیزی است که فرمول (۳-۳) را از رابطه استاندارد ماشین بردار پشتیبان در فرمول (۳-۱) متفاوت می‌سازد. در صورتی که طول این مجموعه برابر با یک باشد، تابع f نسبت به w خطی بوده و فرمول (۳-۳) تبدیل به رابطه ماشین بردار پشتیبان استاندارد خواهد شد. در غیراینصورت، تابع خطای لولایی خطی نخواهد بود. برای حل این مسئله از روش ارائه شده در زمان۲۶] بهره برده و از یک الگوریتم دو مرحله‌ای استفاده می‌کنیم. تابع خطای لولایی برای نمونه‌های منفی، خطی است. به منظور خطی شدن این تابع برای نمونه‌های مثبت، باید مقدار z را برای تمامی نمونه‌های مثبت، ثابت در نظر

[1] Hinge Loss
[2] Convex

گرفت. در چنین حالتی، تابع خطای لولایی محدب بوده و فرمول (۳-۳) توسط انواع روش‌های بهینه‌سازی توابع محدب قابل حل خواهد بود. الگوریتم دومرحله‌ای مورد اشاره به صورت زیر عمل می‌کند:

۱. با طبقه‌بند فعلی (f)، بهترین مقدار z_i از مجموعه $Z(x_i)$ که منجر به بالاترین امتیاز برای نمونه مثبت x_i می‌شود را انتخاب کن.
۲. پس از جایگذاری مقدار z_i برای هر نمونه‌ی مثبت x_i، تابع محدب L را کمینه کن.

هر دو مرحله‌ی بالا موجب بهتر شدن (و یا عدم تغییر) مقدار تابع L خواهند شد [۲۶]. شرط پایان این الگوریتم می‌تواند تعداد تکرار مشخص و یا همگرایی تابع L در نظر گرفته شود. نکته حائز اهمیت در این الگوریتم، مقداردهی اولیه و با دقت به بردار $\vec{w}$ است؛ زیرا در غیراینصورت، مقادیر z_i انتخاب شده برای نمونه‌های مثبت نامناسب بوده و در نتیجه، طبقه‌بند نهایی بدست آمده نیز مناسب نخواهد بود.

۳-۴- هیستوگرام گرادیان‌های جهت‌دار

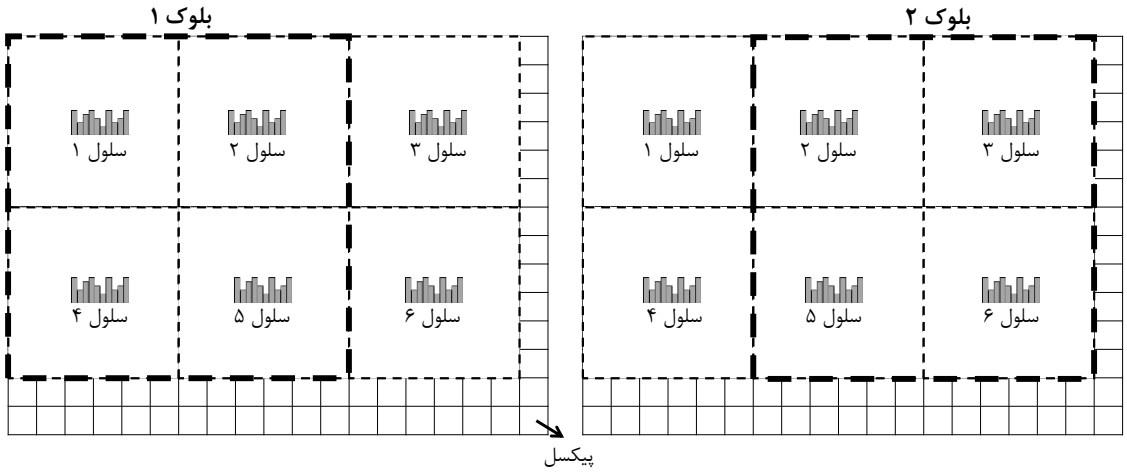
هیستوگرام گرادیان‌های جهت‌دار برای اولین بار توسط دلال و همکاران [۷۶] برای تشخیص افراد در تصاویر ثابت به کار گرفته شد. به دلیل نتایج بسیار مناسب این روش و همچنین سرعت بالای آن، به سرعت در زمینه‌های مختلف تشخیص و شناسایی اشیاء مورد استفاده قرار گرفت [۷۷–۸۰]. این الگوریتم، گرادیان‌های تصویر را به صورت محلی محاسبه کرده و یک هیستوگرام تشکیل می‌دهد. طول هیستوگرام برابر با تعداد جهت‌های لبه‌های موجود در تصویر است. در حالت بدون علامت معمولاً نه جهت در بازه [۱۸۰-۰] و در حالت علامت‌دار ۱۸ جهت در بازه [۳۶۰-۰] در نظر گرفته می‌شود. زوایای بدست آمده از گرادیان‌ها در هر پیکسل به یکی از این جهات تقطیع[۱] شده و بسته به اندازه‌ی (مقدار) گرادیان به یک یا چند ستون از هیستوگرام رأی می‌دهد. بنابراین نقاطی که دارای اندازه گرادیان بزرگتری هستند، تعداد رأی بیشتری خواهند داشت. ماسک یک‌بعدی $[-1,0,1]$ برای استخراج گرادیان-ها توصیه شده است.

هدف این الگوریتم در استخراج گرادیان‌ها و یا همان جهت لبه‌ها رسیدن به یک بردار ویژگی مقاوم به تغییرات روشنایی و دگرسازی‌های[۲] هندسی و فوتومتریک است. استفاده از نسخه علامت‌دار برای برخی از اشیاء از جمله خودروها و دوچرخه‌ها می‌تواند کارا باشد ولی برای تشخیص انسان به دلیل وجود انواع پوشش‌ها و لباس‌ها، تأثیر چندانی ندارد.

---

[۱] Quantize
[۲] Transformation

در این الگوریتم از دو مفهوم سلول و بلوک استفاده می‌شود. تعدادی از پیکسل‌های کنار هم، در قالب یک سلول در نظر گرفته شده و از آن‌ها هیستوگرام گرادیان‌ها استخراج می‌گردد. یک بلوک از چند سلول کنار هم تشکیل می‌شود. هیستوگرام سلول‌ها نسبت به بلوکی که در آن قرار دارند نرمال‌سازی می‌شوند. این عمل موجب مقاومت نسبت به انواع تغییرات روشنایی می‌گردد. بلوک‌ها معمولاً همپوشانی دارند تا سازگاری پس از نرمال‌سازی در کل تصویر حفظ شود. با همپوشانی بلوک‌ها، هر سلول میانی در چهار بلوک مشارکت خواهد کرد و در نتیجه به چهار شکل متفاوت، نرمال‌سازی خواهد شد. عملکرد این الگوریتم در شکل ۳-۳ نمایش داده شده است. یکی از روش‌های مناسب برای نرمال‌سازی بلوک‌ها، استفاده از نرم۲[۱] است که در فرمول (۳-۴) ارائه گشته است.



شکل ۳-۳: نمایش نحوه سلول‌بندی و بلوک‌بندی در الگوریتم هیستوگرام گرادیان‌های جهت‌دار.

$$L2-norm = \frac{v}{\sqrt{\|v\|_2^2 + \varepsilon^2}} \qquad (۳-۴)$$

بردار ویژگی v از اتصال هیستوگرام‌های بدست آمده از سلول‌های داخل بلوک، حاصل می‌شود و ε یک عدد مثبت و کوچک است. دلال و همکاران اندازه‌های مختلف سلول و بلوک را مورد آزمایش قرار داده‌اند و اندازه سلول 8×8 و اندازه بلوک 2×2 را مقادیر مناسبی گزارش کرده‌اند [۷۶]. علاوه‌براین، آن‌ها به این نتیجه رسیدند که به دلیل نرمال‌سازی انجام شده در بلوک‌ها، نیازی به نرمال‌سازی در مرحله پیش-پردازش وجود ندارد. این الگوریتم نسبت به جابجایی‌های محلی در تصویر که از حد سلول‌ها تجاوز نکنند، مقاوم است.

[۱] L^2 Norm

۳-۵- جمع‌بندی

در این فصل به تشریح الگوریتم‌های هرم تصاویر، ماشین بردار پشتیبان، ماشین بردار پشتیبان مخفی و هیستوگرام گرادیان‌های جهت‌دار پرداخته شد. این الگوریتم‌ها در سیستم پیشنهادی مورد استفاده قرار گرفته‌اند. مقادیر پارامترها و شکل استفاده از این روش‌ها در هنگام ارائه روش پیشنهادی توضیح داده شده‌اند.

در ابتدا الگوریتم هرم تصاویر ارائه شد که از الگوریتم‌های شناخته شده در تشخیص اشیاء است. این الگوریتم برای اعمال یک طبقه‌بند به تصویر در همه مکان‌ها و با اندازه‌های مختلف به کار گرفته می‌شود؛ آن هم در حالتی که طبقه‌بند تنها با یک اندازه‌ی ثابت آموزش دیده است. سپس ماشین بردار پشتیبان و ماشین بردار پشتیبان مخفی توضیح داده شدند. این الگوریتم‌ها به ترتیب برای آموزش یک طبقه‌بند بدون پارامترهای مخفی و با پارامترهای مخفی به کار می‌روند. در روش پیشنهادی نیز پارامترهای مخفی از قبیل تعداد یا مکان بخش‌ها وجود دارد. در پایان فصل نیز به تشریح الگوریتم هیستوگرام گرادیان‌های جهت‌دار پرداخته شد. این الگوریتم استخراج ویژگی نسبت به تغییرات روشنایی و دگرسازی‌های هندسی و فوتومتریک مقاومت خوبی دارد و از جمله روش‌های موفق در تشخیص و شناسایی اشیاء شناخته می‌شود. در روش پیشنهادی، از این الگوریتم برای استخراج ویژگی استفاده شده است؛ هر چند، استفاده از هر روش دیگری نیز تغییری در نحوه‌ی عملکرد سیستم ایجاد نخواهد کرد.

فصل

۴- شناسایی نوع و مدل وسیله نقلیه با استفاده از مدل‌های مبتنی بر بخش

۴-۱- مقدمه

در این فصل به توضیح بخش‌های اصلی از سیستم پیاده‌سازی شده در این رساله و نوآوری‌های صورت گرفته پرداخته شده است. تمرکز در این فصل بر روی سیستم پایه است. در فصل بعد، روشی برای بهبود توأمان سرعت و دقت سیستم ارائه شده است. در سراسر این رساله، واژگان «کلاس»، «دسته» و «نوع و مدل» به یک معنی مورد استفاده قرار گرفته‌اند. همچنین دو واژه «طبقه‌بند» و «مدل» نیز برای خوانایی بیشتر جملات، به منظور یکسان به کار گرفته شده‌اند؛ در سیستم پیشنهادی برای هر نوع و مدل از وسایل نقلیه، یک مدل (طبقه‌بند) یاد گرفته می‌شود؛ به منظور خوانایی و زیبایی بیشتر جملات، از آوردن دوباره‌ی واژه «مدل» در برخی مکان‌ها پرهیز شده است.

روش‌های تشخیص و شناسایی اشیاءِ مبنای امروزی[۱]، گرایشی نسبی به سمت روش‌های جزئی‌نگر پیدا کرده‌اند [۲۶, ۸۱, ۸۲]. حتی مشاهده می‌شود که شبکه‌های عصبی عمیق که در حال حاضر، یکی از قدرتمندترین ابزارهای یادگیری ماشین هستند، با شکستن اشیاء پیچیده به بخش‌های کوچک‌تر و ساده‌تر به یادگیری آن‌ها می‌پردازند [۶۸]. روش‌های جزئی‌نگر با استخراج بخش‌هایی (متمایزکننده) از تصویر به تشخیص و شناسایی آن اقدام می‌کنند. علت این گرایش، پیچیدگی مسائل شناسایی اشیاء است؛ برای مثال در تشخیص موفق انسان باید حالت‌های مختلف و چالش‌برانگیزی را پوشش داد؛ تغییرات نور، زاویه-های متفاوت و انسداد بخشی از بدن توسط اشیاء گوناگون تنها چند نمونه از این حالت‌ها هستند. روش-های پیشین که عموماً کلی‌نگر بوده‌اند در تصاویر پیچیده دچار مشکل می‌شوند. برای نمونه، تا چند سال قبل، روش‌های تشخیص شیء ارائه شده در مسابقه شناخته شده‌ی پاسکال، قادر به گذر از دقت مشخصی نبودند [۸۳]؛ تا اینکه در سال ۲۰۰۸ و با ارائه‌ی یک روش مبتنی بر بخش، دقت‌ها بهبود پیدا کرد و این روش پایه‌گذار روش‌های موفق دیگر شد [۷۵].

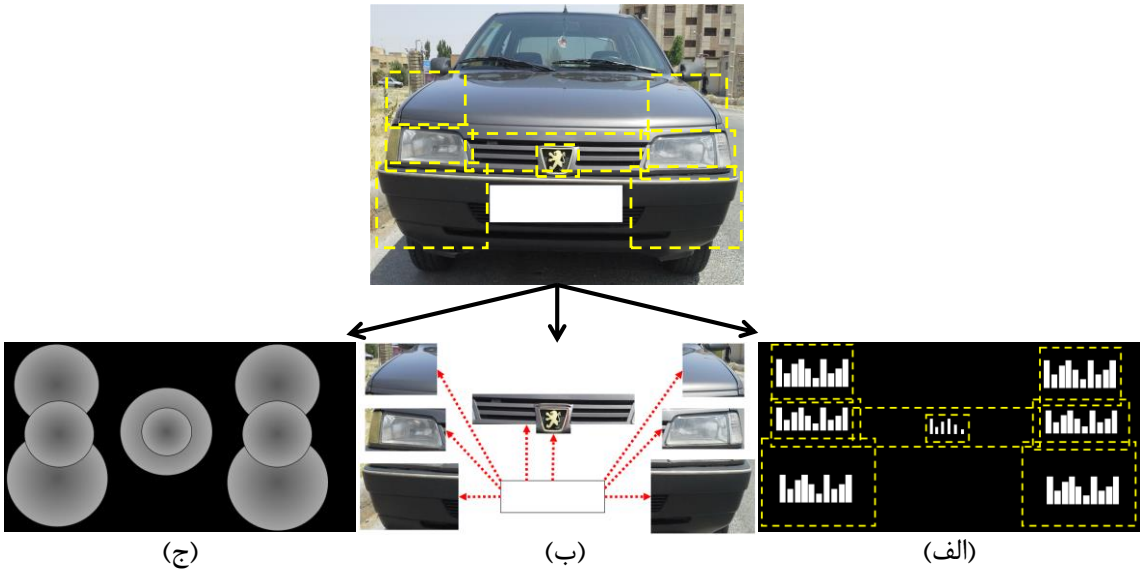
روش پیشنهادی که در دسته‌ی روش‌های جزئی‌نگر قرار می‌گیرد، رویکردی جدید برای شناسایی نوع و مدل وسیله نقلیه ارائه می‌دهد که سعی بر برطرف کردن مشکلات مورد اشاره در بخش ۱-۳- را دارد. ایده‌ی اصلی این رساله از [۷۵] گرفته شده که روشی مبتنی بر بخش برای تشخیص اشیاء ارائه کرده است. از آنجا که همه خودروها دارای بخش‌های مشابه و مشخصی هستند، یک رویکرد مناسب برای تمایز آن‌ها، تمرکز بر روی بخش‌های متمایز کننده آن‌ها و رابطه هندسی بین آن‌ها است. برای مثال با در نظر گرفتن تنها چراغ‌های جلوی خودروها می‌توان تعداد مدل‌های بسیاری را از هم تفکیک کرد؛ هر چند

[۱] State of the Art

چراغ‌ها به تنهایی کافی نیستند. با اضافه کردن رابطه هندسی بین بخش‌ها، می‌توان خودروهایی که دارای بخش‌های مشابه ولی چیدمان بخش‌های متفاوت هستند را نیز تمیز داد. علاوه‌براین، در صورتی که در مرحله‌ی انطباق بخش‌ها، امکان جابجایی بخش‌ها نیز فراهم شود، می‌توان مقاومت به تغییرات زاویه را افزایش داد؛ برای مثال برای جابجا شدن هر بخش از مکان ایده‌آل، خطای مشخصی درنظر گرفت.

نکته حائز اهمیت دیگر، مربوط به بخش‌های انتخاب شده برای هر نوع و مدل از خودروها است. کدام بخش‌ها دارای قدرت تمایز بیشتری هستند؟ چشم انسان برای هر مدل از خودرو یک یا چند مشخصه یاد می‌گیرد که این مشخصه‌ها لزوماً بخش‌های معنی‌دار (برچسب‌دار) نیستند؛ برای مثال وجود یک یا چند شیار بر روی کاپوت. بنابراین برای یافتن مجموعه بخش‌های متمایز کننده‌ی هر مدل از خودروها، نباید جستجو را محدود به بخش‌های برچسب‌دار کرد.

سیستم پیشنهادی همه موارد اشاره شده در بالا را مدنظر قرار داده است. این سیستم تلاش می‌کند تا برای هر مدل از خودروها یک طبقه‌بند آموزش دهد؛ این طبقه‌بند دارای مجموعه‌ای از بخش‌های متمایز کننده‌ی مختص به آن مدل از خودرو است. همچنین، رابطه هندسی بخش‌ها و خطای جابجایی بخش‌ها را نیز در بر دارد. یک نمونه از ساختار چنین طبقه‌بندی در شکل ۴-۱ نمایش داده شده است. بخش (ج) از این شکل، نقشه جابجایی هر بخش را نشان می‌دهد؛ رنگ روشن‌تر در این نقشه به معنی خطای بیشتر است. در این نقشه‌ی جابجایی، بازه‌ی جابجایی هر بخش محدود شده است. در مرحله‌ی آزمایش، هر طبقه‌بند ابتدا بهترین انطباق بخش‌ها (با در نظر گرفتن خطای جابجایی بخش‌ها) را می‌یابد و سپس امتیاز بهترین تشخیص را بازمی‌گرداند. سیستم پیشنهادی از مجموعه‌ای از طبقه‌بندها تشکیل می‌شود که هر یک قادر به شناسایی یک نوع و مدل خاص از خودروها هستند. هر تصویر ورودی برای همه طبقه-بندها ارسال شده و طبقه‌بندی که بیشترین امتیاز را به تصویر داده، نوع و مدل خودرو را تعیین می‌کند.



شکل ۴-۱: هر طبقه‌بند از ویژگی‌های ساختاری (لبه‌ها، خم‌ها و ...) بخش‌های متمایزکننده خودرو (الف)، رابطه ی هندسی بین بخش‌ها (ب) و خطای جابجایی هر بخش (ج) بهره می‌برد. رابطه نشان داده شده تنها یک مثال است.

در ادامه، ابتدا به توضیح مجموعه داده‌های تهیه شده در این رساله پرداخته شده است. سپس، زیربخش تشخیص وسیله نقلیه تشریح گشته است. آنگاه مراحل طی شده برای ایجاد سیستم شناسایی نوع و مدل وسیله نقلیه پیشنهادی توضیح داده شده‌اند. این مراحل عبارتند از ارائه یک الگوریتم کارا برای آموزش یک طبقه‌بند (مدل) که بتواند یک مدل از خودرو را با دقت بالا شناسایی کند؛ ارائه یک الگوریتم داده-کاوی چندکلاسه برای استخراج نمونه‌های منفی دشوار در فرآیند آموزش و ارائه یک الگوریتم برای یافتن بخش‌های متمایزکننده‌ی یک مدل در مقایسه با سایر مدل‌ها؛ و در نهایت، توضیح روش استفاده شده برای کالیبره کردن[۱] طبقه‌بندها به شکلی که بتوان امتیاز بدست آمده از آن‌ها را با هم مقایسه کرد. علاوه-علاوه‌براین، تغییرات کوچک ولی مهم صورت گرفته در طول این مسیر که منجر به بهبود دقت یا سرعت سیستم شده‌اند نیز مورد اشاره قرار گرفته‌اند.

۴-۲- مجموعه داده‌ها

در حین انجام این رساله، به دلیل عدم وجود مجموعه داده به اشتراک گذاشته شده، دو مجموعه داده با دو هدف متفاوت تهیه شده است. مجموعه داده اول به منظور آزمایش ایده‌ی اولیه رساله جمع‌آوری گشت؛ یعنی مقایسه کارایی رویکرد مبتنی بر بخش نسبت به رویکرد کلی‌نگر. مجموعه داده دوم که هم از

[۱] Calibration

نظر تعداد تصاویر و هم تعداد مدل‌ها بزرگ‌تر از مجموعه داده اول است، برای ارزیابی سیستم نهایی تهیه شد. این دو مجموعه داده به دو شکل متفاوت علامت‌گذاری و حاشیه‌نویسی شدند تا پاسخگوی نیازهای مسئله باشند. برای این منظور، نرم‌افزار جامعی طراحی شد که در پایان این بخش به صورت اجمالی معرفی شده است. دو مجموعه داده تهیه شده از طریق آدرس اینترنتی http://mbt925.ir/publications.htm به همراه کدهای موردنیاز برای کار با آن‌ها و فقط برای کاربردهای علمی و تحقیقاتی به صورت رایگان در دسترس هستند.

۴-۲-۱- مجموعه داده اول (نوآوری)

برای تهیه این مجموعه داده که آن را BVMMR v1 می‌نامیم، ۷۲۰ تصویر از ۲۱ مدل مختلف از وسایل نقلیه در مدت شش ماه جمع‌آوری شد. این تصاویر در سه بازه‌ی زمانی صبح، ظهر و عصر گرفته شدند تا انواع تغییرات روشنایی را شامل شوند. فاصله دوربین تا خودروها، سه تا پنج متر و کیفیت دوربین‌های مورد استفاده از پنج تا ده مگاپیکسل متغیر بوده است. از هر خودرو، دو تصویر از نمای جلو و پشت در زمان یکسان تهیه شده است. شکل ۴-۲، نمونه‌ای از تصاویر تهیه شده را به نمایش گذاشته است.

شکل ۴-۲: چند نمونه از تصاویر تهیه شده در مجموعه داده اول.

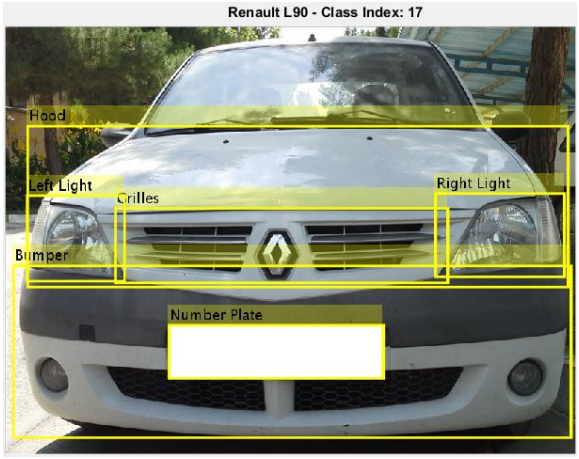
برای مقایسه رویکرد مبتنی بر بخش و رویکرد کلی‌نگر، نیاز به مجموعه داده‌ای بود که بخش‌های وسایل نقلیه در آن به صورت دقیق علامت‌گذاری شده باشد. به این ترتیب، امکان استخراج ویژگی هر یک از بخش‌ها و طبقه‌بندی تصاویر تنها با یک یا چند بخش فراهم می‌شد. به همین دلیل، بخش‌های زیر از هر نما به صورت دستی علامت‌گذاری شدند. علاوه‌براین، مدل و نمای خودروها نیز مشخص شدند.

- نمای جلو: پلاک، چراغ‌های چپ و راست، جلوپنجره، کاپوت و سپر

- نمای پشت: پلاک، چراغ‌های چپ و راست، صندوق و سپر

تصاویر پس از علامت‌گذاری، توسط نرم‌افزار طراحی شده، مورد پردازش قرار گرفته و محدوده‌ی پلاک آن‌ها برای حفظ حریم شخصی پاک‌سازی می‌گردد؛ همچنین، محدوده خودروها از تصاویر بریده شده و طول آن‌ها به ۵۰۰ پیکسل تغییر اندازه داده می‌شود. عرض تصاویر نیز به شکلی که نسبت ابعاد تصویر حفظ گردد، تغییر اندازه داده می‌شوند.

در نسخه منتشر شده از این مجموعه داده، کدهای موردنیاز (در نرم‌افزار متلب) برای خواندن فایل‌های حاشیه‌نویسی مربوط به هر تصویر نیز ارائه شده است. شکل ۴-۳ یک تصویر نمایش داده شده توسط همین کدها را در نرم‌افزار متلب نشان می‌دهد.



شکل ۴-۳: تصویر نمایش داده شده توسط کدهای مربوط به خواندن تصویر در نرم‌افزار متلب.

۴-۲-۲- مجموعه داده دوم (نوآوری)

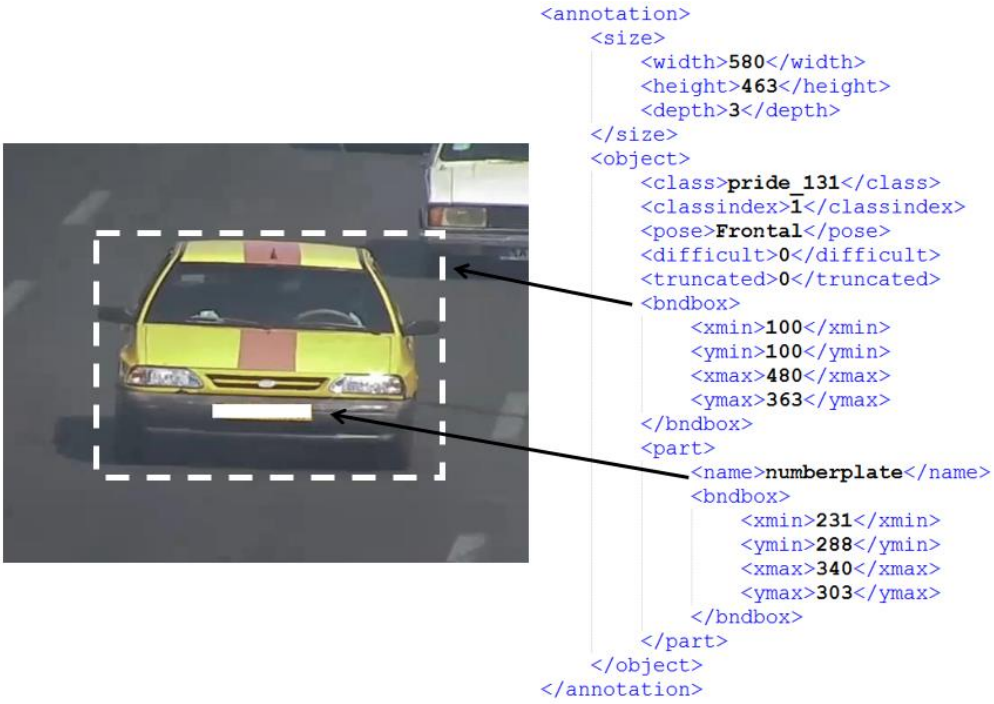
تنها چالش جدی موجود در مجموعه داده اول، تغییرات روشنایی بود؛ تغییرات زاویه موجود در آن تصاویر، چندان زیاد نبود. از این رو و به دلیل کم بودن تعداد تصاویر، ناچار به تهیه مجموعه داده‌ای بزرگ‌تر و با چالش‌های بیشتر به نام BVMMR v2 بودیم. برای این منظور، در بازه‌ی زمانی یک سال و با همکاری مرکز کنترل هوشمند ترافیک استان قم، فیلم‌هایی از چند دوربین مداربسته در چند جاده‌ی درون‌شهری، تهیه شد. نمای همه خودرها روبه‌جلو و فاصله دوربین‌ها تا خودروها از ۵ تا ۲۰ متر متغیر بوده است. در مجموع، ۴۸۵۸ تصویر از ۲۸ نوع و مدل مختلف از ۵۹۹۱ خودرو تهیه شد. تصاویر دارای تغییرات روشنایی زیاد و وضوح متفاوت هستند. علاوه‌براین، همان‌طور که در شکل ۴-۴ نیز قابل مشاهده

است، انتخاب خودروها با تنوع ظاهری بالا صورت گرفته است. در این رساله، هر جا سخن از مجموعه داده BVMMR بدون ذکر مجموعه داده اول یا دوم به میان آورده شود، منظور این مجموعه داده است.

شکل ۴-۴: یک نمونه از تنوع ظاهری یک مدل یکسان در مجموعه داده دوم.

در سیستم پیشنهادی، بخش‌های متمایزکننده هر خودرو به صورت خودکار و تنها با استفاده از تصاویر خودرو یاد گرفته می‌شود. در نتیجه، نیازی به علامت‌گذاری زمان‌بر بخش‌ها وجود ندارد. به همین دلیل، در این مجموعه داده، تنها پلاک و محدوده هر خودرو در تصویر علامت‌گذاری شده است.

به منظور سهولت بیشتر در استفاده از این مجموعه داده، علامت‌گذاری تصاویر منطبق با استاندارد مسابقه شناخته شده پاسکال صورت گرفته است [۸۳]. در این استاندارد، برای هر تصویر یک فایل XML تولید می‌شود که شامل اطلاعاتی از قبیل کلاس خودرو، نمای خودرو، محدوده اطراف خودرو و پلاک آن است (شکل ۴-۵).



شکل ۴-۵: یک نمونه تصویر حاشیه‌نویسی شده منطبق بر استاندارد مسابقه پاسکال از مجموعه داده دوم.

برای این مجموعه، یک بسته کدِ توسعه[۱]ی منطبق با استاندارد پاسکال و در نرم‌افزار متلب ارائه شده است. با استفاده از این بسته، می‌توان تصاویر آموزشی و آزمایشی تعیین شده را پردازش کرد؛ فایل-های حاشیه‌نویسی را خواند و دقت بدست آمده توسط سیستم را محاسبه کرد. تعداد نمونه‌های هر کلاس در کیفیت طبقه‌بندهای آموزش دیده شده مؤثر است. از این رو، نوع و مدل‌های موجود در این مجموعه داده و تعداد تصاویر آموزشی و آزمایشی هر یک از آن‌ها در جدول ۴-۱ ارائه شده است.

جدول ۴-۱: تعداد تصاویر آموزشی، آزمایشی و کل موجود در مجموعه داده دوم از هر نوع و مدل.

شماره کلاس	نوع و مدل	تعداد تصاویر	تعداد تصاویر آزمایشی	تعداد تصاویر آموزشی
۱	لیفان ۶۲۰	۲۰	۹	۱۱
۲	وانت مزدا	۴۲	۱۹	۲۳
۳	مگان	۲۲	۱۰	۱۲
۴	ام وی ام ۳۱۵	۲۶	۱۴	۱۲
۵	ام وی ام ۵۳۰	۳۶	۱۹	۱۷
۶	سایر	۱۹۸	۱۰۴	۹۴

[۱] Developmental Kit

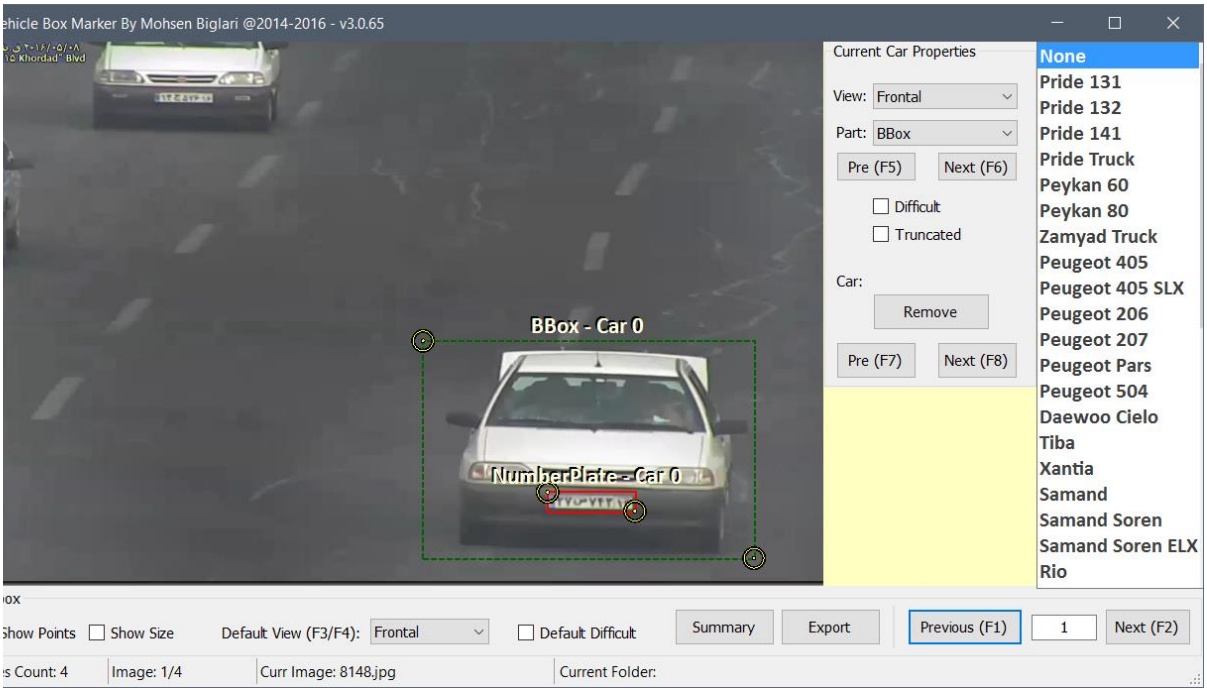
۷	پژو ۲۰۶	۲۴۷	۱۲۵	۱۲۲
۸	پژو ۴۰۵	۷۸۵	۳۸۳	۴۰۲
۹	پژو ۴۰۵ اس ال ایکس	۱۲۰	۶۳	۵۷
۱۰	پژو پارس	۳۲۷	۱۵۹	۱۶۸
۱۱	پیکان ۸۰	۴۸۶	۲۳۷	۲۴۹
۱۲	پراید ۱۳۱	۲۰۰۱	۱۰۰۴	۹۹۷
۱۳	پراید ۱۳۲	۲۴۰	۱۲۵	۱۱۵
۱۴	پراید ۱۴۱	۸۰	۸۰	۱۶۰
۱۵	ال ۹۰	۱۸۹	۹۸	۹۱
۱۶	ریو	۷۲	۳۸	۳۴
۱۷	رانا	۸۱	۳۸	۴۳
۱۸	سمند	۴۸۱	۲۴۲	۲۳۹
۱۹	سمند سورن	۵۰	۲۴	۲۶
۲۰	تیبا	۱۷۲	۹۰	۸۲
۲۱	زانتیا	۴۱	۱۸	۲۳
۲۲	وانت زامیاد	۱۲۵	۶۳	۶۲
۲۳	سمند سورن ای ال ایکس	۱۲	۷	۵
۲۴	ام وی ام ۱۱۰	۱۸	۹	۹
۲۵	پژو ۵۰۴	۸	۶	۲
۲۶	سیلو	۱۵	۸	۷
۲۷	پژو ۲۰۷	۱۰	۵	۵
۲۸	کیا اپتیما ۲۰۱۰	۷	۳	۴
مجموع	-	۵۹۹۱	۳۰۰۰	۲۹۹۱

۴-۲-۳- نرم‌افزار علامت‌گذاری تصاویر (نوآوری)

برای علامت‌گذاری تصاویر دو مجموعه داده مورد اشاره، نرم‌افزاری به زبان برنامه‌نویسی C# طراحی گشت. قابلیت‌های این نرم‌افزار به طور خلاصه در ادامه فهرست شده‌اند. شکل ۴-۶ نمایی از صفحه‌ی اصلی نرم‌افزار را به نمایش گذاشته است.

- تعیین کلاس مربوط به هر خودرو
- تعیین نمای مربوط به هر خودرو (جلو یا پشت)
- امکان تعریف و تعیین محدوده‌های مستطیلی برای هر تعداد بخش دلخواه
- امکان تغییر و حذف محدوده‌های تعیین شده در آینده

- بریدن تصویر بر اساس محدوده خودرو و حذف محدوده پلاک از تصاویر (برای حفظ حریم شخصی)
- قابلیت نمایش و ویرایش محدوده‌های تعریف شده در هر سیستمی با هر وضوحی
- ساخت فایل‌های متنی (XML) منطبق با استاندارد پاسکال شامل جزئیات هر خودرو (حاشیه نویسی[۱])



شکل ۴-۶: نمایی از صفحه اصلی نرم‌افزار طراحی شده برای علامت‌گذاری تصاویر مجموعه داده.

۴-۲-۴- مجموعه داده CompCars

اخیراً یک مجموعه داده بزرگ مخصوص آموزش شبکه‌های عصبی عمیق به نام CompCars ارائه شده است [۸۴]. این مجموعه دارای دو بخش است؛ بخش اول شامل تصاویری است با نما و زاویه‌های متفاوت که از اینترنت جمع‌آوری گشته‌اند و بخش دوم شامل تصاویری است که از دوربین‌های مداربسته جمع-آوری شده‌اند. بخش دوم از این مجموعه در این رساله مورد استفاده قرار گرفته است.

این مجموعه داده دارای ۵۰۰۰۰ تصویر از ۲۸۰ نوع و مدل متفاوت خودرو است. نمای همه تصاویر روبه-جلو است. تصاویر این مجموعه هم در شب و هم در روز جمع‌آوری شده‌اند، بنابراین دارای تغییرات

[۱] Annotation

روشنایی بسیار زیاد می‌باشند.

شکل ۴-۷: چند نمونه تصویر از مجموعه داده بزرگ CompCars [۸۴].

۴-۲-۵- مجموعه داده NTOU-MMR

این مجموعه داده نیز به تازگی معرفی شده و شامل ۶۵۰۰ تصویر روبه‌جلو از ۲۹ نوع و مدل خودروی متفاوت است [۵۲]. تصاویر این مجموعه داده دارای حاشیه‌نویسی نیستند، بنابراین آزمایش سیستم روی این مجموعه به سادگی قابل انجام نیست و بسیار زمان‌بر خواهد بود. چند نمونه از تصاویر این مجموعه داده در شکل ۴-۸ نمایش داده شده‌اند.

شکل ۴-۸: چند نمونه تصویر از مجموعه داده NTOU-MMR [۵۲].

۴-۳- تشخیص وسیله نقلیه

اعمال طبقه‌بندها به کل تصویر از چند جهت منطقی نیست؛ پردازش‌های پیچیده طبقه‌بندها در بخش-هایی از تصویر که خودرو وجود ندارد، موجب افزایش زمان اجرای سیستم می‌گردد. از طرفی این امر منجر به افزایش تعداد تشخیص‌های اشتباه نیز خواهد شد. به این دلایل، بخش تشخیص وسیله نقلیه در مرحله پیش‌پردازش اعمال می‌گردد. این بخش که دارای سرعت بسیار بالایی است، ناحیه‌هایی از تصویر که در آن‌ها خودرو وجود ندارد را فیلتر می‌کند. در نتیجه تنها بخش‌های کوچکی از تصویر برای طبقه-بندهای نهایی ارسال می‌شوند.

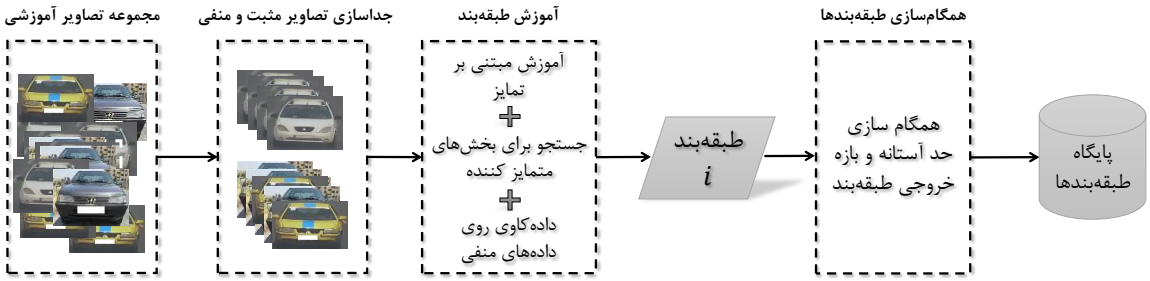
برای آموزش طبقه‌بندِ تشخیص دهنده خودرو از مجموعه داده‌های آموزشی تهیه شده به عنوان نمونه‌های مثبت و از تصاویر جمع‌آوری شده از اینترنت به عنوان نمونه‌های منفی استفاده شده است. تعداد نمونه-های مثبت شامل تقریباً ۳۰۰۰ خودرو و تعداد نمونه‌های منفی در حدود ۲۰۰۰۰ تصویر بوده‌اند. از الگوریتم موفق ارائه شده در [۷۶] برای آموزش این طبقه‌بند بهره گرفته شده است. در این الگوریتم از ماشین بردار پشتیبان به عنوان طبقه‌بند و از هیستوگرام گرادیان‌های جهت‌دار برای استخراج ویژگی استفاده می‌شود.

برای پیاده‌سازی ماشین بردار پشتیبان از کتابخانه LIBSVM [۸۵] استفاده شده است. تمامی پارامترهای مورد استفاده، منطبق با مرجع مربوطه تنظیم شده‌اند؛ با این تفاوت که از نسخه‌ی علامت‌دار هیستوگرام گرادیان‌های جهت‌دار در بازه‌ی [۳۶۰-۰] و با ۱۸ جهت به دلیل عملکرد بهتر آن برای خودروها بهره برده شده است (نسبت به نسخه بدون علامت).

روش‌های متعددی برای تشخیص خوردو و به صورت کلی تشخیص شیء ارائه شده است [۵, ۷, ۸۶–۸۸]؛ که اغلب آن‌ها برای رسیدن به دقت بالا نیاز به پیچیدگی محاسباتی یا الگوریتمی بالا دارند. به همین دلیل، روش ساده و در عین حال سریع ارائه شده در [۷۶] برگزیده شده است. از طرفی، نتایج آزمایشات انجام شده در این رساله نشان از دقت بالای این روش نیز دارند.

۴-۴- شناسایی نوع و مدل وسیله نقلیه (نوآوری)

همان‌طور که پیش‌تر اشاره شد، سیستم پیشنهادی برای هر نوع و مدل از خودروها یک طبقه‌بند (مدل) آموزش می‌دهد. مراحل آموزش یک طبقه‌بند در شکل ۴-۹ نمایش داده است. تمامی بخش‌های نشان داده شده در این شکل، در بخش جاری توضیح داده خواهند شد.



شکل ۴-۹: مراحل آموزش یک طبقه‌بند در سیستم پیشنهادی.

۴-۴-۱- مدل

یک مدل دارای سه بخش اصلی است: ریشه، بخش‌ها و رابطه بین بخش‌ها. ریشه شکل کلی شیء را می‌آموزد. بخش‌ها سعی می‌کنند به صورت دقیق‌تر، اجزای تشکیل دهنده ریشه را یاد بگیرند. رابطه بین بخش‌ها در سیستم پیشنهادی توسط یک مدل ستاره‌ای[۱] پیاده‌سازی می‌شود. در مدل ستاره‌ای، مکان بخش‌ها تنها نسبت به ریشه اهمیت دارد و رابطه‌ای بین مکان بخش‌ها نسبت به یکدیگر وجود ندارد. از آنجا که خودرو دارای یک ساختار سخت[۲] است و مکان بخش‌ها نسبت بهم غیرقابل تغییرند، استفاده از مدل ستاره‌ای کاملاً منطقی است؛ چنان‌که در بخش‌های آتی نیز راجع به این انتخاب بحث شده است.

ریشه و بخش‌ها در واقع از یک ماسک تشکیل شده‌اند که برای جمله‌بندی بهتر آن‌ها را در ادامه رساله، فیلتر خطاب می‌کنیم. این فیلترها در یک ماتریس ویژگی‌ها ضرب خواهند شد. ماتریس ویژگی‌ها، یک آرایه سه‌بعدی است که بعد اول و دوم آن موقعیت در تصویر و بعد سوم آن، بردار ویژگی‌های استخراج شده از آن موقعیت در تصویر است. ابعاد ماتریس ویژگی‌ها به تعداد سلول‌ها (در الگوریتم هیستوگرام گرادیان‌های جهت‌دار) در راستای افقی و عمودی اشاره دارد. یک فیلتر نیز دقیقاً یک آرایه سه‌بعدی با ساختار مشابه ولی با بعد اول و دوم متفاوت است. امتیاز یا پاسخ فیلتر F با ابعاد اول و دوم $w \times h$ در موقعیت (x, y) از ماتریس ویژگی‌های G به صورت زیر محاسبه می‌شود:

$$Score_F(x,y) = \sum_{x'=0}^{w} \sum_{y'=0}^{h} F(x',y').G(x+x',y+y') \qquad (۱-۴)$$

در فرمول (۴-۱)، ابتدا گوشه بالا و سمت چپ از فیلتر F در مکان (x, y) از ماتریس G قرار داده شده و سپس ضرب نقطه‌ای بین عناصر بعد سوم دو ماتریس انجام می‌گیرد. برای محاسبه امتیاز یک فیلتر در همه‌ی مکان‌ها و اندازه‌ها، از یک هرم ویژگی‌ها استفاده می‌شود. هرم ویژگی‌ها، در واقع همان هرم تصاویر

---

[۱] Star Model

[۲] Rigid

است که از همه‌ی سطوح آن ویژگی استخراج شده است؛ بنابراین در هر سطح از هرم ویژگی‌ها، بجای یک تصویر، یک ماتریس ویژگی‌ها قرار دارد. برای سادگی بیشتر، فرمول (۴-۱) را به صورت زیر می‌نویسیم، زیرا زیرپنجره‌ای از G که F در آن ضرب می‌شود به صورت ضمنی از اندازه فیلتر F قابل محاسبه است.

$$Score_F(x,y) = F.G(x,y) \quad \textbf{(۴-۲)}$$

از آنجا که ریشه نیاز به جزئیات تصویر ندارد و بخش‌ها روی جزئیات تصویر عمل می‌کنند، اگر فیلتر ریشه در سطح l از هرم ویژگی‌ها اعمال شود، فیلتر بخش‌ها در سطح $l-\theta$ از هرم اعمال خواهند شد. ابعاد ماتریس (وضوح تصویر) در سطح $l-\theta$ دو برابر سطح l است. با استفاده از این رویکرد، ریشه روی ویژگی‌های مربوط به کلیات شیء و بخش‌ها روی ویژگی‌های مربوط به جزئیات شیء تمرکز می‌کنند. برای مثال، ریشه، شکل و ساختار کلی خودرو را یاد گرفته و بخش‌ها، جزئیات ظاهری هر بخش از ریشه را یاد می‌گیرند.

هر بخش، علاوه‌بر یک فیلتر، دارای یک مکان نسبی به ریشه (مکان لنگر[۱]) و ضرایب یک تابع درجه دو است. مکان لنگر از سه‌تایی (x,y,θ) تشکیل می‌شود که (x,y) مکان بخش نسبت به ریشه را تعیین می‌کند و θ تعداد سطوحی که باید در هرم به پایین حرکت کرد تا به سطح مناسب برای این بخش رسید. خطای جابجایی نسبت به مکان لنگر توسط یک تابع درجه دو محاسبه می‌گردد. امتیاز بخش p برابر است با امتیاز فیلتر مربوط به آن (F_p) منهای خطای جابجای آن نسبت به مکان لنگر:

$$Score_p(x,y) = F_p.G(x,y) - d_p.\phi(dx_p, dy_p) \quad \textbf{(۴-۳)}$$

dx_p و dy_p که مقدار جابجای بخش p نسبت به ریشه هستند، به صورت نشان داده شده در فرمول (۴-۴) محاسبه می‌شوند. در این فرمول، x_r و y_r مکان پنجره‌ی ریشه می‌باشند. فاصله یک بخش نسبت به ریشه از مرکز آن محاسبه می‌شود؛ دو ثابت β_{xp} و β_{yp} به همین دلیل اضافه شده‌اند. وجود عدد دو به عنوان ضریب نیز به دلیل اختلاف دو برابری ابعاد فیلتر ریشه و بخش‌ها است.

$$dx_p = x_p - 2x_r + \beta_{xp} \qquad dy_p = y_p - 2y_r + \beta_{yp} \quad \textbf{(۴-۴)}$$

تابع ϕ یک تابع درجه دوم با ساختار زیر است که بردار d_p به طول چهار در آن ضرب می‌شود. در واقع d_p ضرایب این چند جمله‌ای را مشخص می‌کند. برای نمونه، در صورتی که $d_p = (1,0,1,0)$ باشد، خطای جابجایی بخش p برابر خواهد بود با مربع فاصله اقلیدسی مکان فعلی و مکان لنگر آن.

[۱] Anchor

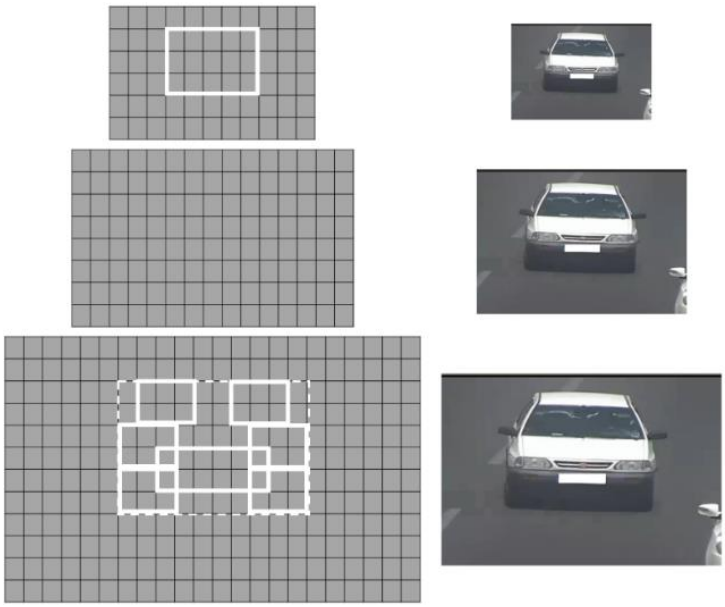
$$\phi(dx_p, dy_p) = (dx_p^2, dx_p, dy_p^2, dy_p) \quad (۵-۴)$$

حال می‌توانیم امتیاز یک مدل را محاسبه کنیم؛ به صورتی که ریشه در مکان (x, y) قرار داده شده و سایر بخش‌ها هریک در بهترین مکان ممکن در محدوده پنجره‌ی ریشه قرار گرفته‌اند. ثابت M تعداد بخش‌های مدل را تعیین می‌کند و r بیانگیر ریشه است.

$$Score = F_r . G(x, y) + \sum_{i=1}^{M} Score_{p_i} \quad (۶-۴)$$

$$Score_{p_i} = \max_{x_i, y_i} \; Score_p(x_i, y_i) \quad (۷-۴)$$

نیاز به یادآوری است که بخش دوم از فرمول (۴-۶) در سطحی از هرم ویژگی‌ها محاسبه می‌شود که وضوحی دو برابری نسبت به بخش اول از فرمول دارد. شکل ۴-۱۰ یک مدل نمونه با هفت بخش را در هرم ویژگی‌ها نمایش می‌دهد که در آن، پنجره ریشه و بخش‌ها در مکان‌های اولیه خود قرار گرفته‌اند.



شکل ۴-۱۰: نمایش نحوه قرارگیری یک مدل نمونه با هفت بخش در سطوح مناسب از هرم ویژگی‌ها.

۴-۴-۲- استخراج ویژگی

آموزش مدل‌ها با استفاده از هرم ویژگی‌ها انجام می‌گیرد. برای استخراج ویژگی از الگوریتم هیستوگرام گرادیان‌های جهت‌دار بهره برده‌ایم که در این بخش به جزئیات آن پرداخته شده است. برای استخراج

ویژگی از اندازه سلول 8×8 و بلوک 2×2 استفاده شده است. این اندازه‌ها بر اساس نتایج گزارش شده توسط دلال و همکاران و چند مرجع دیگر و به صورت تجربی انتخاب گشته‌اند [۷۵, ۷۶].

دلال و همکاران برای اشیائی مانند خودرو و دوچرخه، استفاده از نسخه علامت‌دار عملگر هیستوگرام گردیان‌های جهت‌دار را توصیه کرده‌اند. پس از انجام برخی آزمایش‌ها به این نتیجه رسیدیم که استفاده از هر دو نسخه علامت‌دار و بدون علامت از این عملگر در کنار یکدیگر، بهترین نتیجه را بدست می‌دهد. از این رو، بردار ویژگی نهایی بدست آمده از هر سلول از دو بردار ویژگی علامت‌دار به طول ۱۸ (در بازه [۳۶۰-۰]) و بدون علامت به طول نه (در بازه [۱۸۰-۰]) تشکیل می‌شود. پس از استخراج ویژگی‌ها از هر سلول، بلوک‌ها به روش نرم۲، نرمال‌سازی می‌شوند. با چنین روشی، برای یک تصویر نمونه به ابعاد 800×600 بردار ویژگی به طول ۴۷۲۵ خواهیم داشت:

$$\left(\frac{800}{8}+\frac{600}{8}\right) * (18+9) = 4725$$

۴-۴-۳- داده‌کاوی چند کلاسه در داده‌های منفی

در مسائل یادگیری ماشین که مبتنی بر داده‌های مثبت و منفی است، معمولاً تعداد داده‌های منفی بسیار بیشتر از داده‌های مثبت است. در یک مسئله چندکلاسه که هر طبقه‌بند به صورت یک در برابر همه، آموزش داده می‌شود، عدم توازن بین داده‌های منفی و مثبت، شدیدتر نیز خواهد شد. وارد کردن همه داده‌های منفی در محاسبات از نظر فضایی و زمانی ممکن نیست. از طرفی، استفاده از همه داده‌های منفی لزوماً منجر به بهبود دقت سیستم نهایی نخواهد شد؛ زیرا عدم توازن شدید داده‌های مثبت و منفی احتمال بروز بیش‌برازش[۱] و یا کاهش قدرت تعمیم[۲] طبقه‌بند را افزایش می‌دهد [۸۹]. در چنین مواقعی یک راه‌حل معمول، استفاده از زیرمجموعه‌ای از داده‌های منفی است. برای این منظور، داده‌های دشوار داده‌کاوی شده و آموزش تنها بر روی این داده‌ها صورت می‌گیرد. داده‌کاوی برای بهبود دقت انواع طبقه-بندها در مسائلی که با داده‌های نامتوازن و یا برچسب خورده‌ی معیوب طرف هستیم، نیز به کار گرفته شده است [۹۰–۹۲].

در حالتی که با مسئله‌ی چندکلاسه مواجه هستیم، شکل انتخاب این زیرمجموعه نیز اهمیت پیدا می‌کند. در [۲۶] روشی کارا برای داده‌کاوی دوکلاسه ارائه شده که دقتی برابر با دقت واقعی سیستم در حالتی که

[۱] Overfitting
[۲] Generalization

روی کل داده‌ها آموزش دیده باشد به دست می‌دهد. روش مذکور توسعه داده شده و در این بخش به ارائه‌ی نسخه‌ی چندکلاسه‌ی آن پرداخته شده است.

در یک مسئله یادگیری چند کلاسه، دو نوع داده منفی وجود دارد؛ داده‌های منفی که شامل اشیائی غیر از شیء هدف هستند و داده‌های منفی که شامل شیء هدف هستند ولی در دسته‌ی کلاس در حال آموزش قرار نمی‌گیرند. در مسئله شناسایی نوع و مدل وسیله نقلیه، این دو دسته به صورت زیر تعریف می‌شوند:

۱. نمونه‌های منفی تشکیل شده از همه اشیاء غیر از وسایل نقلیه.
۲. نمونه‌های منفی تشکیل شده از وسایل نقلیه که مدلی متفاوت نسبت به وسلیه نقلیه در حال آموزش دارند.

برای سادگی در نوشتار، نمونه‌های منفی دسته اول را دسته الف (NV) و نمونه‌های منفی دسته دوم را دسته ب (V) می‌نامیم. در الگوریتم پیشنهادی، از این دو مجموعه به دو شکل متفاوت برای آموزش طبقه‌بند بهره برده می‌شود. برای آموزش طبقه‌بند با داده‌های دسته الف، به صورت نشان داده شده در الگوریتم ۴-۱ عمل می‌کنیم:

الگوریتم ۴-۱: الگوریتم آموزش یک طبقه‌بند با داده‌کاوی از داده‌های دسته الف.

```
Function negSmall = Datamining1(M_C)
    pos: positive samples (vehicle class c)
    1. neg: a random subset of NV (neg ⊆ NV)
    2. negSmall: a random subset of neg (negSmall ⊆ neg)
    for dm = 1:dmIterations
        3. Train(c) with {pos, negSmall}
        4. Update negSmall:
          for ∀(easy sample X in neg)      negSmall = negSmall − X
          for ∀(hard sample X in neg)      negSmall = negSmall ∪ X
    end
end
```

M_C در الگوریتم Datamining1، طبقه‌بند مورد آموزش است. داده‌کاوی با استفاده از این طبقه‌بند انجام می‌گیرد. مجموعه pos شامل نمونه‌های مثبت و مجموعه NV شامل نمونه‌های منفی دسته الف می‌باشند.

خروجی یا حاصل این الگوریتم، مجموعه negSmall است که در یک روال تکرار شونده، کوچک‌تر یا بزرگ‌تر می‌شود. در هر تکرار از روال آموزش یک طبقه‌بند، تعدادی نمونه به این مجموعه اضافه و یا حذف می‌شود. مراحل الگوریتم ارائه شده در ادامه به تفکیک توضیح داده شده‌اند:

۱. زیرمجموعه‌ای تصادفی (neg) از کل نمونه‌های منفی دسته الف (NV) در نظر گرفته می‌شود.
۲. تعدادی از تصاویر این زیرمجموعه به صورت تصادفی انتخاب و در مجموعه negSmall قرار داده می‌شوند.
۳. آموزش مدل با استفاده از کل نمونه‌های مثبت (pos) و مجموعه negSmall انجام می‌گیرد.
۴. پس از پایان آموزش، مجموعه negSmall بروز می‌شود. به این شکل که، نمونه‌های آسان از آن حذف و نمونه‌های دشوار به آن اضافه می‌شوند.
۵. مراحل سه و چهار به تعداد dmIterations مرتبه تکرار می‌شوند. در رابطه با تعداد تکرار الگوریتم داده‌کاوی، در بخش‌های بعد بحث خواهد شد.

منظور از نمونه آسان، نمونه‌ای است که امتیاز آن توسط طبقه‌بند فعلی، کوچکتر از ۱- محاسبه شود و نمونه دشوار، نمونه‌ای است که امتیاز آن بزرگتر از ۱- باشد؛ یعنی یا در درون حاشیه‌های طبقه‌بند قرار می‌گیرد و یا به اشتباه به عنوان نمونه مثبت در نظر گرفته می‌شود.

ذکر این نکته نیز حائز اهمیت است که الگوریتم داده‌کاوی و الگوریتم آموزش یک طبقه‌بند به صورت توأمان اجرا می‌شوند و جداسازی کامل این دو الگوریتم امکان‌پذیر نیست. چنان‌که در بخش‌های بعد، اجزای الگوریتم داده‌کاوی در الگوریتم آموزشی ادغام شده و به صورت توأمان ارائه شده است.

اندازه‌ی زیرمجموعه‌ی neg که در مرحله اول انتخاب می‌شود با توجه به حافظه و زمان در دسترس تعیین می‌گردد. در صورتی که اندازه این زیرمجموعه برابر با کل داده‌های منفی باشد، ادعای مرجع [۲۶] مبنی بر رسیدن به دقت ایده‌آل صحیح خواهد بود؛ در غیراینصورت، دقت طبقه‌بند بدست آمده با دقت ایده‌آل، اختلاف اندکی خواهد داشت. دقت ایده‌آل برابر است با دقت سیستمی که با استفاده از کل داده‌های منفی آموزش دیده است. البته باید این نکته را نیز مدنظر قرار داد که افزایش دائمی تعداد نمونه‌های منفی لزوماً منجر به افزایش دقت طبقه‌بند نخواهد شد.

در ادامه به توضیح داده‌کاوی روی داده‌های منفی دسته ب پرداخته می‌شود. یک مجموعه داده‌ی فرضی با M نوع و مدل متفاوت از خودروها در نظر بگیرید. وقتی در حال آموزش یک طبقه‌بند برای نوع و مدل

m_i هستیم، نمونه تصاویر مربوط به $M-1$ نوع و مدل دیگر را داده‌های منفی دسته ب می‌نامیم. برای آموزش مدل با دسته ب، به صورت نشان داده شده در الگوریتم ۴-۲ عمل می‌کنیم:

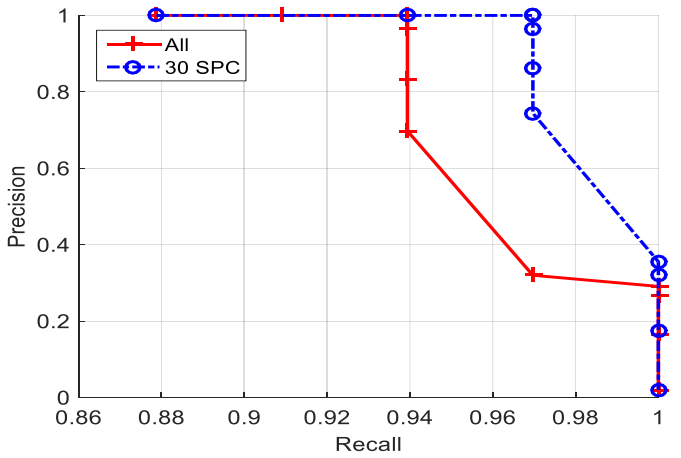
الگوریتم ۴-۲: الگوریتم آموزش یک طبقه‌بند با داده‌کاوی از داده‌های دسته ب.

```
Function negSmall = Datamining2(M_C)
    pos: positive samples (vehicle class c)
    1. neg: a K-distributed subset of V (neg ⊆ V)
       for ∀(negative class nc)      pick K samples {X_1, X_2, ..., X_K} from V (X_i ∈ nc)
    2. negSmall: a L-distributed subset of neg (negSmall ⊆ neg)
    for dm = 1:dmIterations
        3. Train(c) with {pos, negSmall}
        4. Update negSmall:
          for ∀(easy sample X in neg)      negSmall = negSmall − X
          for ∀(hard sample X in neg)      negSmall = negSmall ∪ X
    end
end
```

مراحل الگوریتم ارائه شده به تفکیک در ادامه توضیح داده شده‌اند. مقدار ثابت K و L با توجه به حافظه و زمان در دسترس تعیین می‌گردند. انتخاب این دو ثابت باید به گونه‌ای صورت گیرد که بارگذاری کل تصاویر موردنظر در حافظه‌ی RAM امکان‌پذیر باشد. نمونه‌های منتخب در مراحل یک و دو به صورت تصادفی برگزیده می‌شوند.

۱. زیرمجموعه‌ای تصادفی از کل داده‌های منفی به شکلی انتخاب می‌شود که از هر نوع و مدل حداقل K نمونه وجود داشته باشد. این شکل انتخاب زیرمجموعه را «انتخاب توزیع شده به طول K» می‌نامیم.
۲. تعدادی از تصاویر این زیرمجموعه به صورت توزیع شده به طول L انتخاب و در مجموعه negSmall قرار داده می‌شود.
۳. آموزش مدل با استفاده از کل نمونه‌های مثبت (pos) و مجموعه negSmall انجام می‌گیرد.
۴. پس از پایان آموزش، مجموعه negSmall بروز می‌شود. به این شکل که، نمونه‌های آسان از آن حذف و نمونه‌های دشوار به آن اضافه می‌شوند.
۵. مراحل سه و چهار به تعداد dmIterations مرتبه تکرار می‌شوند. در رابطه با تعداد تکرار الگوریتم داده‌کاوی، در بخش‌های بعد بحث خواهد شد.

برای آموزش سیستم بر روی مجموعه داده BVMMR، از کل داده‌های منفی در مرحله اول استفاده شده و L برابر با ۱۰ در نظر گرفته شده است؛ در حالی که برای آموزش سیستم بر روی مجموعه داده CompCars، K برابر با ۳۰ و L برابر با ۱۰ قرار داده شده است. به دلیل تعداد زیاد نمونه‌های مجموعه داده CompCars، استفاده از کل نمونه‌های منفی در مرحله اول، علاوه‌بر اینکه موجب کاهش بسیار زیاد سرعت آموزش می‌گردد، منجر به کاهش دقت سیستم حاصل نیز خواهد شد. هر کلاس از این مجموعه داده به صورت میانگین دارای ۱۵۰ نمونه تصویر است. فرض کنید قصد داریم یک سیستم با ۱۸۰ طبقه-بند را برای ۱۸۰ کلاس از این مجموعه داده آموزش دهیم. در این حالت، برای آموزش هر کلاس، ۱۵۰ نمونه مثبت و ۲۶۸۵۰ (150×179) نمونه منفی از دسته ب خواهیم داشت. عدم توازن داده‌ها در این حالت بسیار شدید است و آموزش یک طبقه‌بند مانند ماشین بردار پشتیبان با استفاده از این داده‌ها، به احتمال زیاد منجر به بیش‌برازش خواهد شد و یا طبقه‌بندی با قدرت تعمیم بالا حاصل نخواهد کرد. برای آزمایش این فرضیه، یک طبقه‌بند را برای کلاس «Besturn x80 366» از مجموعه داده CompCars یک مرتبه با کل نمونه‌های منفی دسته ب و یک مرتبه با انتخاب توزیع شده به طول ۳۰ (SPC[۱]) آموزش دادیم. البته در سیستم پیشنهادی که فرآیند آموزش با استفاده از داده‌کاوی روی داده‌های منفی صورت می‌گیرد، تعداد داده‌ها کاهش یافته و در نتیجه عدم توازن نیز کمتر می‌شود. شکل ۴-۱۱ نمودار صحت و بازیابی مربوط به این دو طبقه‌بند را با هم مقایسه کرده است. مشاهده می‌شود که عملکرد طبقه‌بند آموزش دیده با داده‌کاوی روی تعداد نمونه‌های منفی کمتر، بهتر از طبقه‌بند دیگر است.



شکل ۴-۱۱: نمودار صحت و بازیابی یک طبقه‌بند که یک مرتبه با استفاده از کل نمونه‌های منفی و یک مرتبه با انتخاب توزیع شده به طول ۳۰ آموزش دیده است.

[۱] Number of Samples per Class

انتخاب توزیع شده تصاویر در بین نوع و مدل‌های متفاوت در مرحله نخست، برای رسیدن به طبقه‌بندی قدرتمند، ضروری است. الگوریتم پیشنهادی برای داده‌های دسته ب موجب بهبود دقت طبقه‌بند بدست آمده نسبت به حالتی که از الگوریتم ۴-۱ برای این داده‌ها استفاده می‌شد، می‌گردد (شکل ۴-۱۱)؛ زیرا عدم وجود نمونه از همه‌ی نوع و مدل‌ها، ممکن است منجر به حذف مجموعه‌ای از نمونه‌های منفی دشوار شود. الگوریتم داده‌کاوی پیشنهادی از ترکیبی از دو الگوریتم ۴-۱ و الگوریتم ۴-۲ بهره می‌برد که شکل استفاده از آن در بخش بعد تشریح شده است.

۴-۴-۴- الگوریتم آموزش یک مدل

در رویکرد پیشنهادی، برای هر کلاس c از خودروها یک مدل M_c یاد گرفته می‌شود که از سه بخش فیلتر ریشه، فیلتر بخش‌ها و رابطه‌ی هندسی (S) بین بخش‌ها تشکیل شده است: $M_c = \mathcal{F}(F_r, F_{parts}, S_{parts})$. رابطه‌ی هندسی بین بخش‌ها همان‌طور که در بخش‌های قبل مورد اشاره قرار گرفت، از ساختار ستاره‌ای بهره می‌برد و دارای دو بخش است: مکان لنگر هر بخش نسبت به مکان فیلتر ریشه (x, y, θ) به طول سه و خطای جابجایی نسبت به مکان لنگر (d_p) که برداری به طول چهار می‌باشد.

برای آموزش یک مدل با استفاده از ماشین بردار پشتیبان مخفی، بخش رابطه‌ی هندسی بین بخش‌ها (S_{parts}) به عنوان متغیر مخفی و دو بخش دیگر، پارامترهای بهینه‌سازی در نظر گرفته می‌شوند. بنابراین، هدف یافتن مکان و خطای جابجایی بخش‌ها در حین فرآیند آموزش طبقه‌بند است. در صورت متغیر در نظر گرفتن تعداد و ابعاد بخش‌ها، دشواری مسئله افزایش زیادی پیدا خواهد کرد. به همین دلیل، تعداد و ابعاد بخش‌ها در فرآیند آموزش، ثابت در نظر گرفته شده و تنها خطای جابجایی بخش‌ها آموزش داده می‌شود. مراحل کلی آموزش یک مدل (M_c) به روش پیشنهادی در ادامه ارائه شده است.

الگوریتم ۴-۳: الگوریتم مراحل کلی آموزش یک مدل به روش پیشنهادی برای کلاس c از خودروها.

Function $M_c = \textbf{Train}(\text{Class c})$
F_r: $[0]_{R_1 \times R_2}$ $F_{parts} = \{[0]_{P_1 \times P_2}\}_{1 \times M}$ $S_{parts} = \{[0]_{1\times 3}, [0]_{1\times 4}\}_{1\times M}$
pos: positive samples (vehicle class c)
neg: a random subset of negative non-vehicle samples ($neg \subseteq NV$)
vehNeg: a K-distributed subset of negative vehicle samples ($vehNeg \subseteq V$)
1. Training(F_r, pos, neg, 1, SVM)
2. Training(F_r, pos, a L-distributed subset of vehNeg, 15, SVM)
3. Initialize parts filters (F_{parts}) and positions (S_{parts}) using F_r
$M_c = \{F_r, F_{parts}, S_{parts}\}$
4. Training(M_c, pos, a L-distributed subset of vehNeg, 80, LSVM)
5. Training(M_c, pos, vehNeg, 5, LSVM)
end

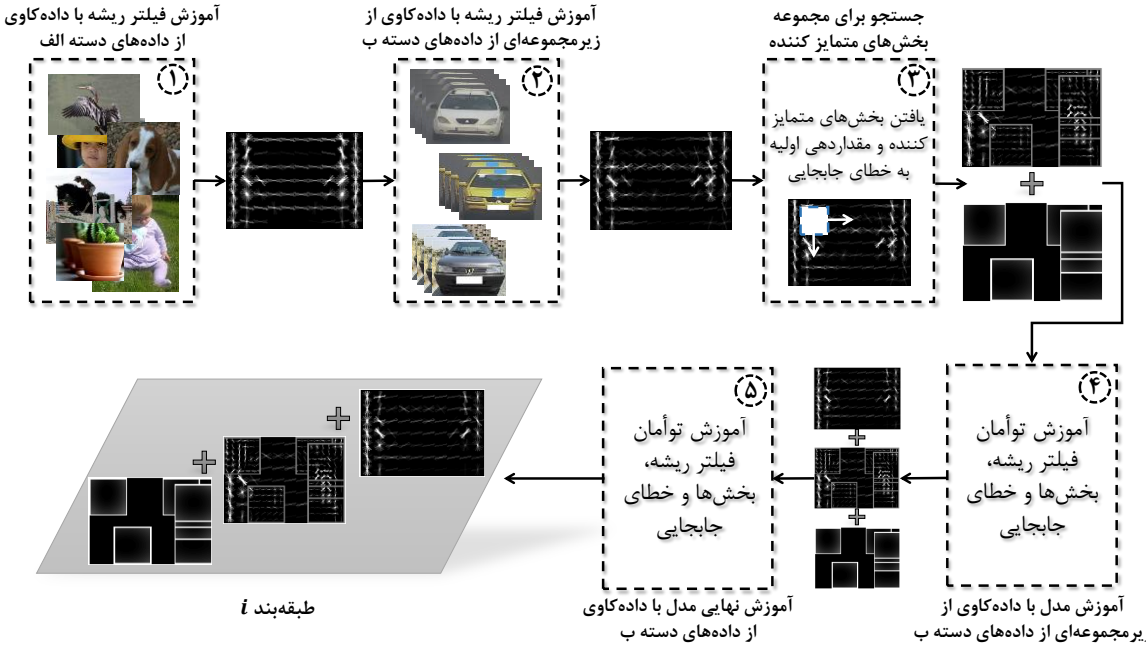
ثابت‌های $(R_1 \times R_2)$ و $(P_1 \times P_2)$ ابعاد فیلترهای ریشه و بخش‌ها هستند که بر اساس ابعاد داده‌های آموزشی تعیین شده و در طول فرآیند آموزش ثابت در نظر گرفته می‌شوند. ثابت M بیانگر تعداد بخش‌ها است. رابطه‌ی هندسی بین بخش‌ها (S_{parts}) از دو زیربخش مکان لنگر و خطای جابجایی تشکیل شده که در بخش‌های قبل تشریح شده است.

مجموعه داده pos، نمونه‌های مثبت و دو مجموعه داده‌ی neg و $vehNeg$، نمونه‌های منفی را تشکیل می‌دهند. مجموعه داده‌ی neg، نمونه‌های منفی دسته الف و مجموعه داده‌ی $vehNeg$، نمونه‌های منفی دسته ب می‌باشند. این دو مجموعه بسته به محدودیت حافظه و زمان برای مرحله آموزش، زیرمجموعه‌ای از کل نمونه‌های منفی خواهند بود. انتخاب مقادیر K و L نیز بر همین اساس تعیین می‌شود. روش داده کاوی روی این دو مجموعه داده در بخش ۴-۴-۳- تشریح شده است.

الگوریتم ۴-۳ برای آموزش یک مدل از پنج مرحله تشکیل می‌شود که در شکل ۴-۹ نمایش داده شده و در ادامه نیز توضیح داده شده‌اند:

۱. آموزش فیلتر ریشه با استفاده از ماشین بردار پشتیبان استاندارد بر روی نمونه داده‌های آموزشی دسته الف.
۲. آموزش فیلتر ریشه با استفاده از ماشین بردار پشتیبان استاندارد بر روی زیرمجموعه‌ای تصادفی و توزیع شده از نمونه‌های داده‌های آموزشی دسته ب.

۳. یافتن مجموعه بخش‌های متمایزکننده برای مدل جاری، تعیین مکان لنگر برای آن‌ها و مقداردهی اولیه به فیلتر بخش‌ها.
۴. آموزش کل اجزای مدل با استفاده از ماشین بردار پشتیبان مخفی بر روی زیرمجموعه‌ای تصادفی و توزیع شده از نمونه داده‌های دسته ب.
۵. آموزش کل مدل با استفاده از ماشین بردار پشتیبان مخفی بر روی نمونه داده‌های دسته ب.



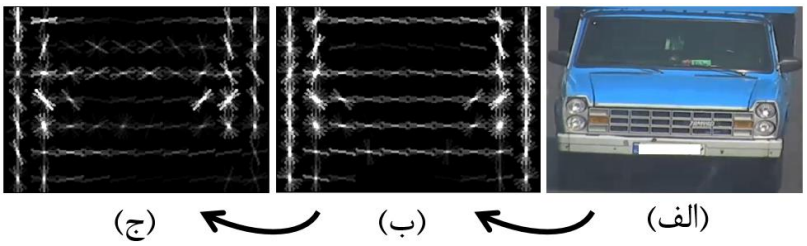
شکل ۴-۱۲: مراحل پنج‌گانه الگوریتم آموزش یک مدل.

آموزش فیلترها و مدل‌ها با استفاده از انواع ماشین‌های بردار پشتیبان، به صورت توأمان همراه با داده-کاوی صورت می‌گیرد. فرآیند آموزش که توسط تابع Training انجام می‌شود به تعداد مشخصی تکرار می‌شود (nIterations). در حالتی که آموزش روی زیرمجموعه‌ای از کل نمونه‌های منفی صورت می-گیرد، تعداد تکرارهای حلقه‌ی آموزشی بزرگتر تعیین می‌شود. در مرحله چهارم که یکی از مهترین مراحل آموزش طبقه‌بند است، تعداد تکرارهای حلقه فرآیند آموزش، بیشتر از سایر مراحل تعیین شده است. بین تعداد دور آموزشی و اندازه مجموعه داده‌های منفی یک رابطه معکوس وجود دارد. هرچقدر که اندازه مجموعه داده‌های منفی کوچک‌تر باشد، تعداد دور آموزشی بیشتر تعیین می‌شود. تعداد تکرارهای مراحل دوم، چهارم و پنجم به ترتیب برابر با ۱۵، ۸۰ و ۵ قرار داده شده است. این مقادیر با آزمون و خطا بدست آمده‌اند.

تابع Training که چند مرتبه در الگوریتم ۴-۳ فراخوانی شده، توسط الگوریتم ۴-۴ ارائه شده است. این تابع $nIterations$ مرتبه تکرار شده و در هر تکرار یک مرحله داده‌کاوی و یک مرحله آموزش مدل را انجام می‌دهد. در صورتی که خطای مدل بدست آمده در تکرار جاری نسبت به مدل قبلی، تغییر چندانی نکرده باشد، الگوریتم متوقف می‌گردد. پارامتر ورودی آخر تابع Training، نوع الگوریتم آموزشی است که یکی از دو مقدار SVM و LSVM را می‌پذیرد.

در مرحله اول از الگوریتم ۴-۳، فیلتر ریشه با استفاده از کلیه نمونه‌های منفی دسته الف که شامل هیچ خودرویی نیستند، آموزش داده می‌شود. در پایان این مرحله، فیلتر ریشه تفاوت بین خودرو و غیرخودرو را یاد می‌گیرد. این مرحله در واقع یک مقداردهی اولیه مناسب برای فیلتر ریشه است. در مرحله دوم و با استفاده از تصاویر آموزشی دسته ب، فیلتر ریشه به سمت ساختار اختصاصی کلاس جاری (c) متمایل شده و قدرت تمایز بین کلاس c و سایر کلاس‌ها را می‌آموزد. شکل ۴-۱۳، فیلتر ریشه را در پایان مرحله اول و دوم برای کلاس «وانت زامیاد» نمایش داده است. با مقایسه دو فیلتر ریشه در این شکل، مشاهده می‌کنیم که تأثیر چرخ‌ها در فیلتر دوم کمتر شده و تأثیر چراغ‌ها و شکل خاص کاپوت این خودرو افزایش یافته است. یک نکته مهم در الگوریتم ۴-۳، شکل بکارگیری تصاویر آموزشی است. ترکیب پیشنهادی، تعادل مناسبی بین دقت بالا و سرعت قابل قبول در مرحله آموزش برقرار می‌کند. به شکلی که مدت زمان آموزش یک مدل با ۲۰۰۰ نمونه مثبت، ۱۰۰۰۰ نمونه منفی دسته الف و ۵۰۰۰ نمونه منفی دسته ب، در حدود یک ساعت و نیم است.

محل قرارگیری مرحله تعیین مکان بخش‌ها بسیار مهم است. پس از پایان مرحله دوم از الگوریتم ۴-۳، فیلتر ریشه خود را با ساختار انحصاری کلاس موردنظر، تطبیق داده است؛ در نتیجه می‌توان بخش‌هایی با قدرت تمایز بالاتر را در فیلتر ریشه جستجو کرد. در صورتی که تعیین مکان بخش‌ها در پایان مرحله اول اتفاق می‌افتاد، چنین امری میسر نمی‌شد. مرحله تعیین مکان بخش‌ها در بخش بعد تشریح شده است.



شکل ۴-۱۳: نمایش فیلتر ریشه پس از آموزش توسط الگوریتم ۴-۳ در پایان مرحله اول (ب) و دوم (ج) برای کلاس «وانت زامیاد» (الف). برای سادگی بیشتر، تنها بخش بدون علامت از عملگر هیستوگرام گرادیان‌های جهت‌دار نمایش داده شده است.

تفاوت مراحل چهارم و پنجم در الگوریتم ۴-۳، اندازه‌ی مجموعه داده‌های منفی و تعداد تکرارهای تابع Training است. مرحله چهارم، داده‌کاوی را روی زیرمجموعه‌ای توزیع شده از کل داده‌های منفی دسته ب انجام می‌دهد؛ در صورتی که مرحله پنجم، داده‌کاوی را روی کل داده‌های منفی انجام می‌دهد. به دلیل کوچک‌تر بودن اندازه مجموعه در مرحله چهارم، تعداد تکرارهای این مرحله بیشتر است. به همین دلیل است که در مرحله پنجم، داده‌کاوی به تعداد بسیار کمتری نسبت به مرحله قبل انجام می‌شود.

الگوریتم ۴-۴: الگوریتم مراحل جزئی آموزش یک مدل به همراه داده‌کاوی از نمونه‌های منفی.

Function M_c = **Training**(Model M_c, posSet, negSet, nIterations, algorithm)
  $M_{0..nIterations} = \{M_c \dots M_c\}$
  $loss_0 = \infty$
  for i = 1: nIterations
    negs: data mining hard negatives by M_c
    M_i: Train M_c with {posSet, negs} using algorithm procedure
    $loss_{dm}$ = compute hinge loss on negs
    if $(loss_{i-1} - loss_i) < \varepsilon$
      $M_{nIterations} = M_i$
      break
    end
    $loss_{i-1} = loss_i$
  end
  $M_c = M_{nIterations}$
end

تابع Training از یک حلقه اصلی تشکیل شده است. در هر تکرار از حلقه ، یک مرتبه آموزش طبقه‌بند و یک مرتبه داده‌کاوی انجام می‌گیرد. پس از هربار اجرای الگوریتم داده‌کاوی و آموزش، بهبود خطای طبقه‌بند جاری نسبت به طبقه‌بند دور قبل بررسی شده و در صورت ناچیز بودن بهبود (ε)، حلقه داخلی متوقف می‌شود. در نهایت، طبقه‌بند بدست آمده در دور پایانی به عنوان طبقه‌بند مطلوب بازگردانده می-شود.

۴-۴-۵- مجموعه بخش‌های متمایزکننده برای هر مدل

برای مدل کردن رابطه هندسی بین بخش‌ها، روش‌های متعددی ارائه شده است [۹۳]. برای تطابق یک الگو با تصویر نیز روش‌هایی وجود دارد که به دلیل پیچیدگی زیاد، کارایی یا سرعت بالایی ندارند. برای مثال روش‌های [۹۴, ۹۵] برای یافتن یک الگوی منطبق بهینه، مکان بخش‌ها را به مجموعه کوچکی از مکان‌ها محدود می‌کنند. در حالی که، با استفاده از ساختارهای تصویری[۱] درختی و ستاره‌ای (شکل ۴-۱۴) می‌توان یک الگوریتم برنامه‌نویسی پویای بهینه ارائه کرد که همه حالت‌های مختلف انطباق را در زمان چندجمله‌ای بررسی کند [۲۶, ۹۳]. به همین دلیل، در سیستم پیشنهادی از مدل ستاره‌ای برای پیاده‌سازی رابطه بین بخش‌ها استفاده کرده‌ایم. این انتخاب چند مزیت به همراه دارد:

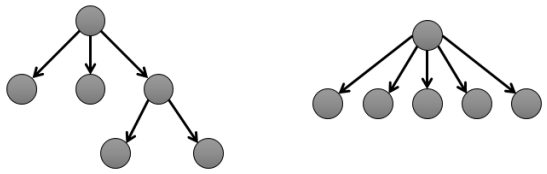
- امکان جستجوی همه حالات انطباق در زمان معقول
- عدم وابستگی بخش‌ها نسبت به یکدیگر
- سادگی در پیاده‌سازی

مزیت اول در رابطه با مدل‌های درختی نیز صادق است ولی دو مزیت دیگر مختص مدل‌های ستاره‌ای هستند. عدم وابستگی بخش‌ها نسبت به یکدیگر به سیستم کمک می‌کند تا انطباق هر بخش را مستقل از سایر بخش‌ها انجام دهد؛ در این حالت می‌توان از موازی‌سازی برای افزایش سرعت پردازش نیز بهره برد. مرجع [۹۶] از مدل درختی برای تشخیص اشیاء استفاده کرده است. در مدل درختی، مکان هر بخش ممکن است وابسته به بخش دیگری باشد. بنابراین نمی‌توان بخش‌ها را مستقل از هم پردازش کرد. مرجع مربوطه نیز ناچاراً در مرحله انطباق بخش‌ها، ابتدا با استفاده از مرتب‌سازی توپولوژیکی[۲] ترتیبی از بخش‌ها را به شکلی بدست می‌آورد که ابتدا بخش‌های مستقل پردازش شوند. چنین سرباری به دلیل استفاده از مدل درختی به سیستم تحمیل شده است. سادگی در پیاده‌سازی مدل ستاره‌ای، در تبدیل هر طبقه‌بند به معادل آبشاری آن که در فصل آینده توضیح داده شده نیز کمک شایانی خواهد کرد.

---

[۱] Pictorial Structure

[۲] Topological Sort



(ب) (الف)

شکل ۴-۱۴: دو ساختار تصویری ستاره‌ای (الف) و درختی (ب) برای توصیف رابطه هندسی بین بخش‌ها.

همان‌طور که در بخش‌های قبل توضیح داده شد، هر بخش از مدل پیشنهادی شامل سه قسمت است که عبارتند از فیلتر بخش، مکان لنگر و خطای جابجایی نسبت به مکان لنگر. در ادامه، روش پیشنهادی برای یافتن بخش‌های متمایزکننده‌ی یک مدل و مکان لنگر مناسب برای هر یک از بخش‌ها تشریح شده است.

در ابتدا معیار تمایز یک بخش را تعریف می‌کنیم. یک بخش از طبقه‌بند کلاس c را متمایزکننده می‌خوانیم اگر شباهت کمی نسبت به بخش‌های مشابه از طبقه‌بند سایر کلاس‌ها داشته باشد. منظور از شباهت کم، در مقایسه با سایر بخش‌های طبقه‌بند کلاس c است. طبقه‌بند کلاس c دارای یک مجموعه بخش‌های ممکن می‌باشد. هدف، یافتن n بخش از این مجموعه است که نسبت به سایر اعضای مجموعه، تفاوت بیشتری با بخش‌های متناظر از طبقه‌بند سایر کلاس‌ها داشته باشند. در فرآیند جستجو تنها نیاز به فیلتر ریشه‌ی همه M طبقه‌بند داریم که آن‌ها را $F_{rc_1}, F_{rc_2}, \dots, F_{rc_M}$ نامگذاری می‌کنیم. طبقه‌بند کلاس c را جدا کرده و در مجموعه F_{rc} قرار داده و سایر طبقه‌بندها را در مجموعه F_{ro} قرار می‌دهیم:

$$F_{rc} = \{F_{rc_i}\}$$

$$F_{ro} = \{F_{rc_1}, F_{rc_2}, \dots, F_{rc_{i-1}}, F_{rc_{i+1}}, \dots F_{rc_M}\}$$

وضوح فیلتر بخش‌ها دو برابر وضوح فیلتر ریشه است. بنابراین، اندازه فیلترهای ریشه ابتدا با استفاده از نمونه‌برداری افزایشی[۱] دو برابر شده و سپس مورد استفاده قرار می‌گیرند. الگوریتم ۴-۵ مجموعه بخش‌های متمایزکننده‌ی کلاس c را با داشتن دو مجموعه F_{rc} و F_{ro} به صورت حریصانه جستجو کرده و بازمی‌گرداند. تعداد بخش‌ها (n) را برای همه کلاس‌ها به صورت ثابت و برابر با هشت در نظر گرفته‌ایم. ابعاد بخش‌ها را نیز برابر با 6×6 قرار داده‌ایم. این مقادیر به صورت تجربی بهترین نتایج را بدست داده‌اند.

[۱] Upsampling

الگوریتم ۴-۵: الگوریتم حریصانه یافتن مجموعه بخش‌های متمایزکننده.

```
Function P = FindPartsAnchor(F_rc, F_ro)
   n : number of parts,   S: fixed size of parts
   P : ∅
   for i = 1:n
        sims :[∞]_{1×|F_rc|}
        for r_c of size S ∈ F_rc
             sims[r_c] = Similarity(r_c, F_ro)
        end
        r_c = find r so that  min_r sims[r]
        P = P ∪ r_C
   end
end
```

در حلقه‌ی داخلی، تمامی بخش‌های ممکن در فیلتر F_{rc} که ابعادی برابر با S دارند، پیمایش شده و شباهت آن‌ها با بخش‌های مشابه از فیلتر سایر کلاس‌ها محاسبه می‌شود. منظور از طول فیلتر F_{rc} ($|F_{rc}|$)، تعداد بخش‌هایی هستند که شرایط توضیح داده شده را دارند. نتایج محاسبه شده را در آرایه sims قرار داده و در پایان حلقه، بخشی که دارای کمترین شباهت بوده را انتخاب کرده و در مجموعه P قرار می‌دهیم. طرز عملکرد تابع Similarity در فرمول (۴-۸) ارائه شده است.

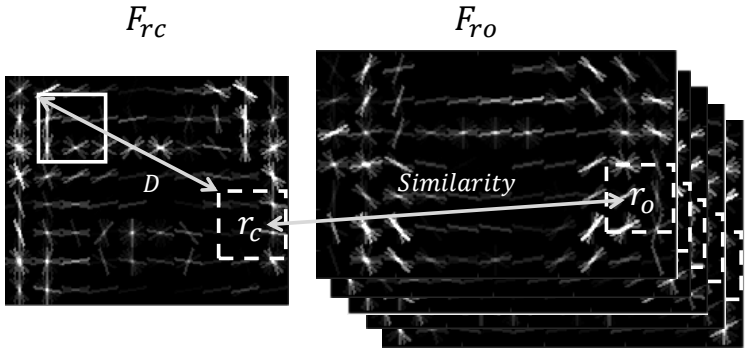
$$Similarity(r_c, \mathcal{F}_{ro}) = \frac{1}{|\mathcal{F}_{ro}|}\sum_{1}^{|\mathcal{F}_{ro}|} cosine(r_c, r_o) + \alpha \cdot (1 - \min_i D(l_r, l_i)) \qquad (۴-۸)$$

هر بخش (r_c) یک ناحیه 6×6 از فیلتر ریشه است. این تابع مجموع شباهت ناحیه r_c را با تمامی ناحیه-هایی از سایر فیلترها (F_{ro}) که در مکان مشابه قرار دارند و اندازه یکسان دارند، محاسبه می‌کند. مقداری که برای شباهت دو ناحیه r_c و r_o بدست می‌آید از دو بخش تشکیل شده است (شکل ۴-۱۵):

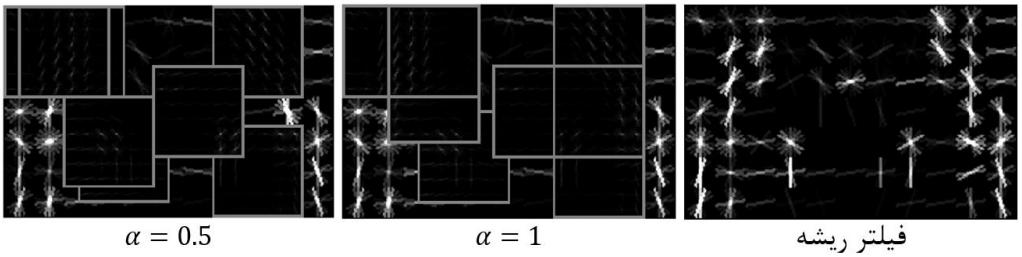
۱. فاصله کسینوسی دو ناحیه موردنظر (تابع $cosine$)
۲. فاصله اقلیدسی ناحیه r_c و نزدیک‌ترین ناحیه‌ی انتخاب شده به ناحیه r_c در مراحل قبل (تابع D)

ثابت α، میزان تأثیر دو بخش بالا در مقدار نهایی بدست آمده را تعیین می‌کند. بخش دوم در واقع یک پنالتی است تا بخش‌های نزدیک به هم انتخاب نشوند. علت وارد کردن فاصله اقلیدسی در معیار شباهت پیشنهادی، تمایل بخش‌ها به هم‌پوشانی بوده است. وقتی یک بخش دارای تمایز بالا بوده و به عنوان

بخش مناسب انتخاب گردد، احتمالا ناحیه‌های اطراف آن نیز شرایط مشابهی دارند. به این ترتیب، بدون وجود پنالتی، بخش‌های نزدیک بهم انتخاب خواهند شد. هرچقدر که مقدار ثابت α بزرگتر بوده و به یک نزدیک‌تر شود، فاصله بین بخش‌های انتخاب شده بیشتر خواهد بود (شکل ۴-۱۶). مقادیر $\alpha > 0.5$ تعادل قابل قبولی بین عدم شباهت بخش‌ها و فاصله اقلیدسی بین بخش‌های انتخاب شده برقرار می‌نمایند. تابع D فاصله اقلیدسی بدست آمده را توسط روش Min-Max به بازه [۱-۰] نگاشت کرده و برمی‌گرداند. l_r مختصات ناحیه r می‌باشد.



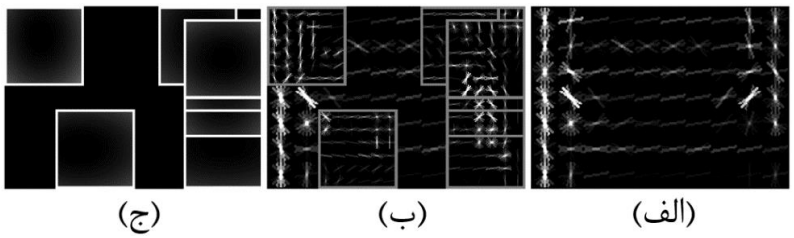
شکل ۴-۱۵: الگوریتم پیشنهادی سعی بر بیشینه کردن فاصله بین بخش‌های انتخاب شده (D) و کمینه کردن شباهت ($Similarity$) دارد.



شکل ۴-۱۶: مجموعه بخش‌های انتخاب شده برای کلاس «پراید ۱۳۱» به ازای دو مقدار متفاوت α.

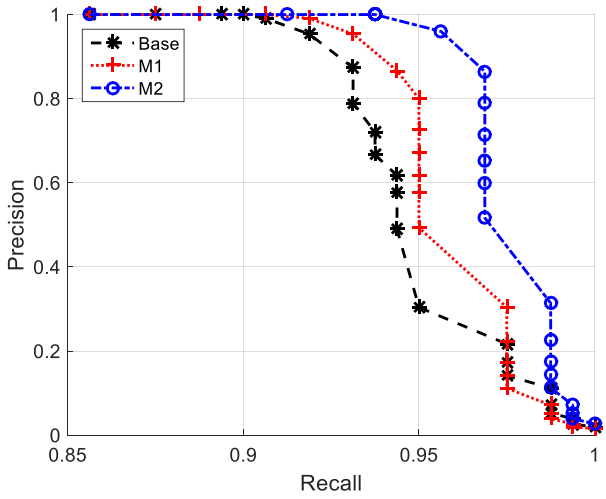
بردار ویژگی هر بخش در واقع یک هیستوگرام است که هر ستون آن یک زاویه است. هدف، یافتن هیستوگرامی است که کمترین شباهت را با هیستوگرام‌های متناظر از سایر کلاس‌ها دارد. با در نظر گرفتن زاویه بین دو بردار حاصل از دو هیستوگرام موردنظر، می‌توان تخمین مناسبی از شباهت بین آن‌ها ارائه کرد. برای مثال در صورتی که یک هیستوگرام، ترکیب خطی از هیستوگرام دیگر باشد، زاویه بین بردار حاصل از آن‌ها صفر خواهد بود؛ بنابراین شباهت بسیار زیادی خواهند داشت. علت استفاده از معیار کسینوسی نیز همین امر بوده است.

از مقادیر فیلتر ریشه به عنوان مقادیر اولیه فیلتر بخش‌ها استفاده می‌شود. این نوع مقداردهی اولیه بسیار مؤثرتر از یک مقداردهی اولیه تصادفی است. علاوه‌براین، مقدار اولیه ضرایب تابع خطای جابجایی برابر با [0.1,0,0.1,0] قرار داده می‌شود. شکل ۴-۱۷، مدل نهایی بدست آمده در پایان الگوریتم ۴-۳ را برای کلاس «وانت زامیاد» نمایش داده است.



شکل ۴-۱۷: مدل نهایی آموزش دیده شده برای کلاس «وانت زامیاد». فیلتر ریشه (الف). فیلتر بخش‌ها (ب). هزینه جابجایی بخش‌ها نسبت به مکان لنگر (ج): رنگ تیره‌تر به معنی خطای کمتر است.

برای نمایش بهبودهای حاصل شده توسط الگوریتم‌های پیشنهادی در این بخش و بخش قبل، یک مسئله دو کلاسه را در نظر گرفتیم. در این مسئله کلاس «وانت زامیاد» را به عنوان نمونه‌های مثبت و سایر کلاس‌ها را به عنوان نمونه‌های منفی تعیین کردیم. ابتدا یک مدل را توسط روش ارائه شده در [۲۶] که روش پایه این رساله است، آموزش دادیم؛ سپس الگوریتم ۴-۳ را اعمال کرده و آموزش را مجددا انجام دادیم (بدون الگوریتم یافتن مجموعه بخش‌های متمایزکننده). در پایان، الگوریتم ۴-۵ را اعمال کرده و آموزش را تکرار کردیم. شکل ۴-۱۸ تفاوت عملکرد این سه طبقه‌بند را نشان داده است.



شکل ۴-۱۸: نمودار صحت و بازیابی مدل بدست آمده توسط روش [۲۶] (Base)، پس از اعمال الگوریتم ۴-۳ (M1) و پس از اعمال الگوریتم ۴-۵ (M2).

۴-۴-۶- همگام‌سازی مدل‌ها

در سیستم پیشنهادی، مدل‌ها مستقل از هم آموزش می‌بینند و هر مدل یک حد آستانه مناسب برای خود انتخاب می‌کند. حد آستانه‌ای مناسب است که کمترین خطای ممکن را روی نمونه‌های آموزشی بدست بدهد. برای یافتن چنین حد آستانه‌ای می‌توان یک پیمایش خطی روی امتیاز بدست آمده برای همه نمونه‌های آموزشی انجام داد. به دلیل مستقل بودن فرآیند آموزش مدل‌ها، حد آستانه و محدوده‌ی خروجی (فاصله بین حاشیه‌ها در ماشین بردار پشتیبان) آن‌ها بر هم منطبق نیست.

برای یکسان‌سازی حد آستانه مدل‌ها می‌توان از یک ثابت بایاس[۱] استفاده کرد، ولی این روش تأثیری روی محدوده خروجی مدل‌ها ندارد؛ بنابراین نمی‌توان بین امتیاز بدست آمده از مدل‌ها، رأی‌گیری انجام داد. برای رفع این مشکل و اصطلاحاً کالیبره کردن مدل‌ها، از الگوریتم مقیاس‌گذاری پلت[۲] استفاده شده است [۹۷]. در این الگوریتم، یک تابع نمایی به شکل نشان داده شده در فرمول (۴-۹) به خروجی هر مدل $f(x)$ برازش[۳] داده می‌شود.

$$P(y=1|x)=\frac{1}{1+e^{Af(x)+B}} \qquad (۴-۹)$$

ضرایب A و B با استفاده از مجموعه داده‌های آموزشی تعیین می‌شوند. البته برای این منظور، پلت استفاده از مجموعه داده‌های اعتبارسنجی[۴] را توصیه کرده است؛ ولی به دلیل تعداد کم نمونه‌ها برای برخی کلاس‌ها، امکان تعریف چنین مجموعه‌ای میسر نبود. از طرفی، دقت حاصل از محاسبه ضرایب به کمک نمونه‌های آموزشی قابل قبول بوده است. $f(x)$ امتیاز خروجی قبلی مدل و $P(x)$ امتیاز کالیبره شده مدل برای نمونه ورودی x می‌باشد.

پلت همچنین پیشنهاد کرده است که به جای استفاده از مقادیر ۱ و ۱- به عنوان برچسب خروجی نمونه-های مثبت و منفی، از دو مقدار زیر استفاده شود. ثابت‌های N_+ و N_- بیانگر تعداد نمونه‌های مثبت و منفی هستند.

$$y_+=\frac{N_++1}{N_++2}$$

---

۱ Bias

۲ Platt Scaling

۳ Fit

۴ Validation Set

$$y_- = \frac{1}{N_- + 2}$$

در سیستم پیشنهادی، پس از پایان آموزش همه مدل‌ها، همگام‌سازی یا کالیبره کردن به روش توضیح داده شده، صورت می‌گیرد. تمامی نتایج ارائه شده در این رساله، توسط سیستم کالیبره شده به دست آمده‌اند.

۴-۵- جمع‌بندی

در این فصل، سیستم پیاده‌سازی شده برای شناسایی نوع و مدل وسیله نقلیه تشریح شد. اجزای مورد استفاده توسط سیستم از قبیل مجموعه داده‌ی تهیه شده، تشخیص وسیله نقلیه، الگوریتم آموزش مدل-ها، داده‌کاوی چند کلاسه بر روی نمونه‌های منفی، جستجو برای مجموعه بخش‌های متمایزکننده برای هر مدل و الگوریتم همگام‌سازی مدل‌ها مورد بحث قرار گرفتند. نشان داده شد که با اضافه شدن هر یک از این بخش‌ها به روش پیشنهادی، عملکرد آن بهبود پیدا کرد.

تا این مرحله، سیستم قادر است پس از دریافت تصویر ورودی، وسیله یا وسایل نقلیه موجود در تصویر را تشخیص داده و استخراج کند. سپس این تصاویر را به تک‌تک طبقه‌بندهای آموزش دیده شده ارسال می-نماید. آنگاه نتایج بدست آمده از هر یک از مدل‌ها را جمع‌آوری می‌کند. در نهایت، کلاس مربوط به مدلی که بالاترین امتیاز را گزارش کرده است، به عنوان خروجی صحیح در نظر گرفته می‌شود.

با توجه به نتایج آزمایشات انجام شده، دقت سیستم در حالت فعلی بسیار مناسب است ولی سرعت آن با وجود تعداد مدل‌های زیاد (۱۸۰ طبقه‌بند در مجموعه داده CompCars)، کند است. به منظور بهبود سرعت سیستم، یک الگوریتم آبشاری برای تغییر شکل اعمال طبقه‌بندها به تصویر ورودی ارائه کرده‌ایم که در فصل بعد توضیح داده شده است. این الگوریتم که تنها در مرحله آزمایش به کار گرفته می‌شود، تعادلی بین سرعت و دقت سیستم ایجاد می‌کند.

فصل

۵- آبشاری از مدل‌های آبشاری (نوآوری)

سیستم پیشنهادی برای شناسایی نوع و مدل تصویر ورودی، وسیله نقلیه موجود در تصویر را برای همه n طبقه‌بند $\{M_{C_1} \dots M_{C_n}\}$ ارسال کرده و از بین امتیازهای خروجی، طبقه‌بند دارای امتیاز بزرگتر را برمی-گزینند. اگر همه‌ی امتیازها از یک حد آستانه مشخص کمتر باشند، کلاس «سایر» به عنوان خروجی تعیین می‌گردد. شکل ۵-۱ مراحل کلی سیستم از تشخیص تا شناسایی وسیله نقلیه را نمایش داده است. اعمال همه طبقه‌بندها به تصویر یک فرآیند زمان‌بر است. با افزایش تعداد طبقه‌بندها، زمان پردازش نیز به صورت خطی افزایش خواهد یافت. در این فصل، الگوریتمی برای بهبود سرعت سیستم پیشنهاد شده است.

هرچند که اعمال مدل‌ها به تصویر به صورت موازی توسط CPU با هزینه‌ی کم قابل انجام است؛ با این حال، زمان پردازش سیستم هنوز جا برای بهبود دارد. اعمال موازی مدل‌ها توسط یک کامپیوتر چهار هسته‌ای در حالت ایده‌آل، زمان پردازش سیستم را به یک چهارم آن کاهش خواهد داد. از طرفی، در برخی کاربردها، این امکان وجود دارد که سرعت سیستم را افزایش داد و از کاهش جزئی دقت چشم-پوشی کرد؛ برای مثال در سیستمی که از خروجی چند الگوریتم طبقه‌بندی متفاوت با استفاده از رأی-گیری بین آن‌ها بهره می‌برد، دقت کل سیستم با کاهش جزئی در دقت یکی از الگوریتم‌ها، تأثیر نمی-پذیرد؛ زیرا رای اکثریت، اثر تصمیم اشتباه یک طبقه‌بند را خنثی خواهد کرد.

شکل ۵-۱: مراحل کلی شناسایی نوع و مدل خودرو توسط سیستم غیرآبشاری.

در مسئله تشخیص نوع و مدل وسیله نقلیه، وجود تعداد کلاس‌های زیاد (بیش از ۳۰) محتمل است. روش‌هایی برای ترکیب طبقه‌بندهای مستقل وجود دارند که اغلب آن‌ها مانند شکل ۵-۱، همه طبقه‌بندها را به تصویر اعمال می‌کنند [۳۰, ۹۸–۱۰۱]. علت این امر در هدف چنین الگوریتم‌هایی نهفته است که افزایش دقت طبقه‌بند نهایی است. به همین دلیل، زمان پردازش طبقه‌بندهای ترکیبی نه تنها کاهش نمی‌یابد، بلکه افزایش چندبرابری خواهد داشت.

در بخش جاری، تمرکز اصلی روی بهبود سرعت با حفظ دقت جاری سیستم و یا کاهش جزئی در آن است. برای این منظور، یک الگوریتم آبشاری ارائه شده است. این الگوریتم دارای دو بخش است که هر یک در زیربخش‌های جداگانه توضیح داده شده‌اند:

۱. **آبشاری از مدل‌ها**: اعمال ترتیبی مدل‌ها به تصویر با یک ترتیب از پیش تعیین شده.
۲. **مدل‌های آبشاری**: جایگزین کردن هر مدل با معادل آبشاری آن.

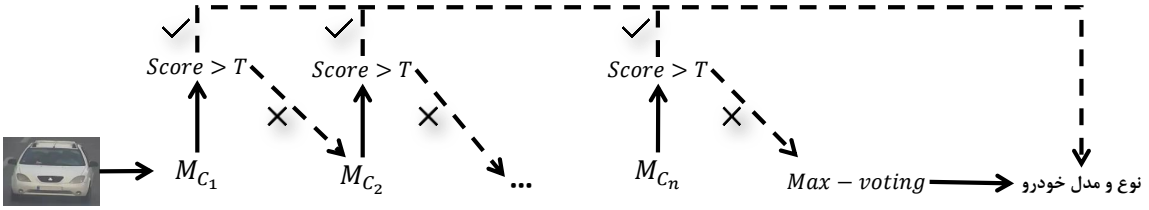
۵-۱- آبشاری از مدل‌ها

هر وسیله نقلیه دارای یک نوع و مدل مشخص است، بنابراین تنها یک طبقه‌بند از مجموعه طبقه‌بندهای سیستم، قادر به تشخیص نوع و مدل صحیح هر خودرو است. اگر اطلاعاتی در رابطه با نوع و مدل تصویر ورودی وجود داشت، می‌توانستیم طبقه‌بند صحیح را پیش‌بینی کرده و ابتدا آن را به تصویر اعمال کنیم؛ و یا حداقل آن را زودتر از قبل به تصویر اعمال کنیم. چنین اطلاعاتی را می‌توان به صورت نسبی و با بررسی نمونه‌های آموزشی، بدست آورد. علاوه‌بر این، با تخمین عملکرد هر طبقه‌بند، می‌توان بر رفتار در آن زمان اجرای سیستم نظارت کرد.

الگوریتم پیشنهادی به جای اعمال موازی طبقه‌بندها، یک ترتیب خطی از آن‌ها (آبشاری از مدل‌ها) و یک حد آستانه سراسری T در نظر می‌گیرد. طبقه‌بند اول در این صف، ابتدا به تصویر اعمال می‌شود؛ اگر امتیاز خروجی این طبقه‌بند بزرگتر از T باشد، نوع و مدل وسیله نقلیه، تعیین شده و فرآیند شناسایی متوقف می‌شود. در نتیجه تصویر برای سایر طبقه‌بندها ارسال نمی‌گردد. در صورتی که امتیاز خروجی طبقه‌بند اول کوچکتر از T باشد، تصویر برای طبقه‌بند دوم و به همین ترتیب تا آخرین طبقه‌بند در صف ارسال خواهد شد. در حالتی که امتیاز خروجی هیچ طبقه‌بندی بزرگتر از T نباشد، از الگوریتم رأی‌گیری بزرگترین امتیاز[۱] استفاده خواهد شد (شکل ۵-۲).

در تمامی آزمایش‌های این فصل از مجموعه داده دوم (BVMMR) بهره برده شده است. این انتخاب به دلیل توزیع غیریکنواخت تصاویر در کلاس‌های مختلف این مجموعه داده بوده که برای نمایش کارایی الگوریتم آبشاری پیشنهادی مناسب‌تر است. البته در بخش پایانی این فصل، نتایج حاصل از اعمال الگوریتم پیشنهادی به مجموعه داده CompCars نیز گزارش شده است.

[۱] Max-voting

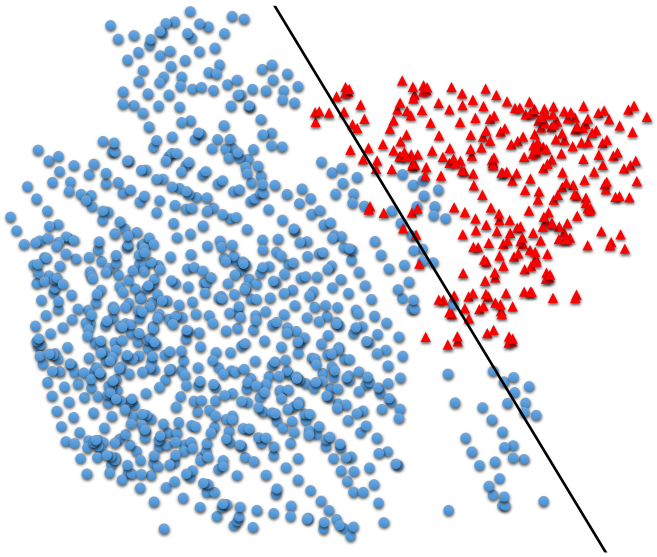


شکل ۵-۲: نحوه عملکرد الگوریتم آبشاری برای تعیین نوع و مدل خودرو. $\{M_{C_1}, M_{C_2}, \dots, M_{C_n}\}$ طبقه‌بندهای سیستم هستند. $Score$، امتیاز خروجی طبقه‌بند مربوطه و T حد آستانه سراسری است.

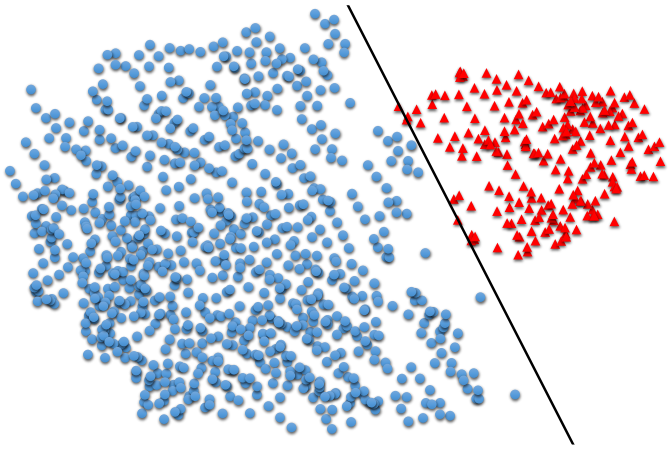
دو فاکتور سرعت و دقت در هر سیستم شناسایی شیء حائز اهمیت است. هدف اصلی در چنین سیستم-هایی، داشتن دقت و سرعت بالا در کنار هم می‌باشد. چطور می‌توان ترتیبی از اعمال طبقه‌بندها به تصویر ارائه کرد که سرعت را افزایش دهد و در عین حال کاهش زیادی به دقت اعمال نکند و یا حتی دقت را بهبود بخشد. برای ارائه چنین ترتیبی باید ابتدا به تحلیل عملکرد سیستم پرداخت.

۵-۱-۱- اطمینان

یکی از ضعف‌های ذاتی یک سیستم با چند طبقه‌بند، تفاوت اطمینان طبقه‌بندهای مختلف است که از تنوع در نمونه‌های آموزشی نشأت می‌گیرد. ممکن است یک کلاس دارای تعداد نمونه‌های بسیار بیشتری نسبت به سایر کلاس‌ها باشد؛ و یا ممکن است چالش‌های موجود در تصاویر یک کلاس بیشتر از سایر کلاس‌ها باشد. شکل ۵-۳ و شکل ۵-۴ به ترتیب دو طبقه‌بند ماشین بردار پشتیبان را که با تعداد نمونه-های چالش‌برانگیر بیشتر و کمتر آموزش دیده شده‌اند را نمایش داده است. برای مثال، طبقه‌بندی که برای مدل «پراید ۱۳۱» با بیشترین فراوانی در جاده‌های کشور آموزش داده می‌شود، یک مدل قدرتمند است. این مدل با تعداد تصاویر بسیار زیاد که دارای انواع رنگ‌ها و حالت‌ها بوده‌اند، آموزش دیده است. از طرفی، این امر موجب می‌شود درصد اشتباه این طبقه‌بند برای مدل‌های شبیه به «پراید ۱۳۱» افزایش یابد. در نقطه‌ی مقابل، طبقه‌بندی که برای مدل «رانا» با تعداد تصاویر کمتر و با چالش کمتر آموزش داده شده، به قدرتمندی مدل «پراید ۱۳۱» نیست ولی درصد اشتباه کمتری دارد زیرا نمونه‌های کمتری دیده و در نتیجه، سختگیری بیشتری دارد. وقتی امتیاز همه طبقه‌بندها برای یک تصویر ورودی درخواست می‌شود، احتمال اعلام اشتباه یک امتیاز بالا توسط مدل «پراید ۱۳۱» بیشتر از مدل «رانا» است. این مورد یک عامل منفی در جهت افزایش دقت سیستم جاری است.



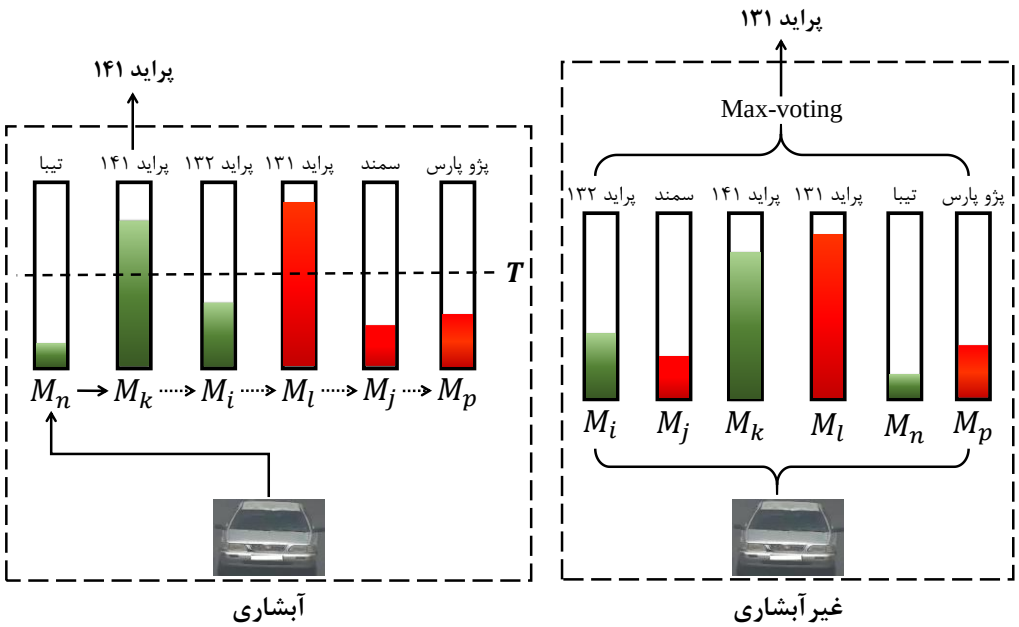
شکل ۵-۳: نمونه‌ای از یک طبقه‌بند که با نمونه‌های چالش‌برانگیز بیشتری آموزش دیده است.



شکل ۵-۴: نمونه‌ای از یک طبقه‌بند که با نمونه‌های چالش‌برانگیر کمتری آموزش دیده است.

برای کاهش تأثیر این عامل بر دقت سیستم می‌توان مدلی که دارای اطمینان بیشتری است را در ابتدای آبشار قرار داد. مدلی را مطمئن خطاب می‌کنیم که نمونه‌های مثبت را با امتیاز بالا پذیرفته و نمونه‌های منفی را با امتیاز پایین رد کند. با اعمال چنین ترتیبی، دقت سیستم بهبود پیدا خواهد کرد. علاوه‌براین، به دلیل اعمال خطی مدل‌ها، سرعت نسبت به حالت غیرآبشاری (و غیرموازی) افزایش پیدا خواهد کرد. در شکل ۵-۵ یک مثال از حالتی که عملکرد نسخه‌ی آبشاری منجر به بهبود نرخ شناسایی درست سیستم نسبت به نسخه‌ی غیرآبشاری می‌گردد، به تصویر کشیده شده است. در این شکل، تصویری از خودروی «پراید ۱۴۱» به عنوان ورودی ارائه شده است. در نسخه‌ی غیرآبشاری که همه طبقه‌بندها به تصویر ورودی اعمال می‌شوند، طبقه‌بند «پراید ۱۳۱» به دلیل شباهت بالایی که به مدل «پراید ۱۴۱»

دارد، امتیاز بالایی را گزارش کرده است. این امتیاز که بالاترین امتیاز گزارش شده در بین طبقه‌بندها می‌باشد، منجر به شناسایی اشتباه گشته است. در صورتی که در نسخه آبشاری، طبقه‌بندها بر اساس معیار اطمینان مرتب شده‌اند. بنابراین شانس بیشتری به طبقه‌بندهای مربوط به مدل‌های «تیبا»، «پراید ۱۴۱» و «پراید ۱۳۲» داده می‌شود. در نتیجه، طبقه‌بند مربوط به مدل «پراید ۱۴۱» با گزارش کردن امتیازی بزرگتر از حد آستانه‌ی سراسری (T)، از رقابت بین طبقه‌بندها پیروز خارج شده است. M_i در این تصویر، طبقه‌بند i ام است که خودروی متناظر با آن در بالای مستطیل نمایش داده شده است.



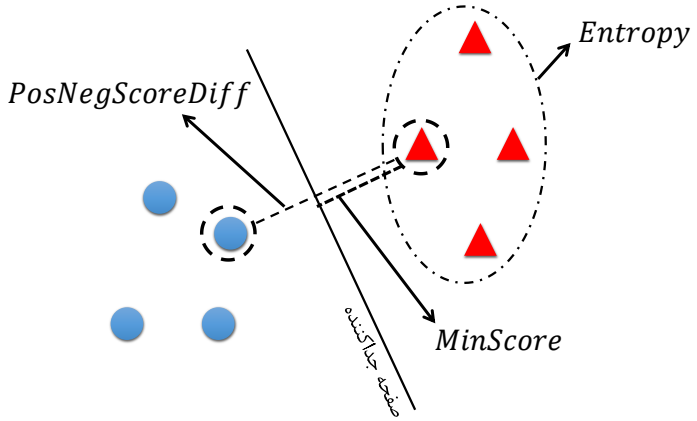
شکل ۵-۵: یک مثال از عملکرد بهتر نسخه‌ی آبشاری نسبت به نسخه‌ی غیرآبشاری. مدل تصویر ورودی «پراید ۱۴۱» است. رنگ قرمز و سبز به ترتیب بیانگیر اطمینان کمتر و اطمینان بیشتر هستند (در حالت خاکستری رنگ قرمز و سبز به خاکستری تیره و روشن تبدیل خواهند شد). مقدار پر شده از هر میله نشانگر امتیاز طبقه‌بند مربوطه است.

برای تخمین درصد اطمینان یک طبقه‌بند، سه معیار زیر را ارائه کرده و عملکرد آن‌ها مورد مقایسه قرار داده‌ایم. این سه معیار برای توصیف بهتر، در شکل ۵-۶ نیز به تصویر کشیده شده‌اند.

- بی‌نظمی[۱] ($Entropy$): طبقه‌بندی که دارای بی‌نظمی کمتری بین امتیازهای خروجی‌اش باشد، دارای اطمینان بیشتری است. برای محاسبه اطمینان مبتنی بر بی‌نظمی از فرمول (۵-۱) استفاده شده است.

[۱] Entropy

- کمترین امتیاز یک نمونه مثبت ($MinScore$): طبقه‌بندی که یک نمونه‌ی مثبت را با امتیاز پایین (نزدیک به صفر) می‌پذیرد، اطمینان کمی دارد. برای محاسبه اطمینان به این روش، فرمول (۵-۲) ارائه شده است.
- اختلاف کمترین امتیاز یک نمونه مثبت و بیشترین امتیاز یک نمونه منفی ($PosNegScoreDiff$): طبقه‌بندی مطمئن است که اختلاف بین کمترین امتیازی که برای یک نمونه‌ی مثبت بدست می‌دهد و بیشترین امتیازی که برای یک نمونه منفی بدست می‌دهد، زیاد باشد. فرمول (۵-۳) برای محاسبه اطمینان به این روش ارائه شده است.



شکل ۵-۶: سه معیار ارائه شده برای تخمین اطمینان یک طبقه‌بند.

$$E_i = -\frac{1}{\sum_{j=1}^{K} M_i(x_j)\log(M_i(x_j))} \quad (۵-۱)$$

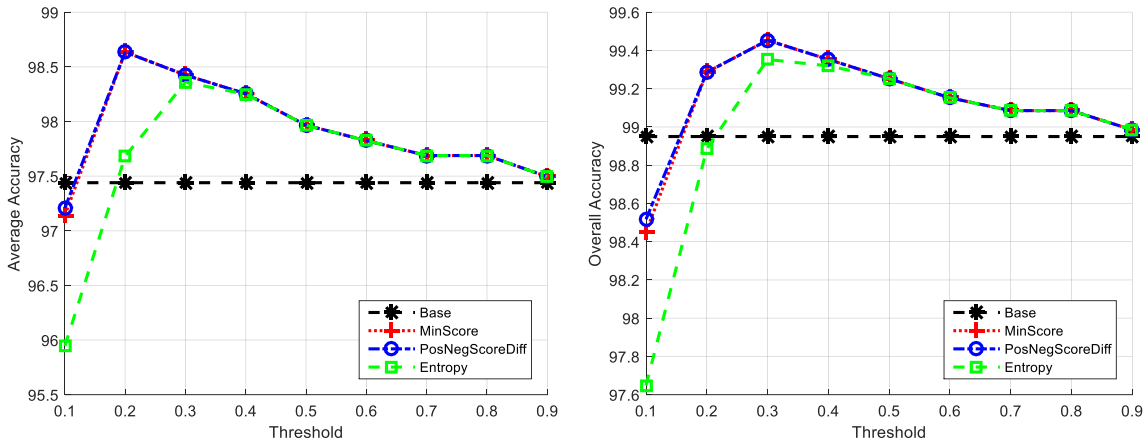
$$MS_i = \min_{x \in Pos(x)} M_i(x) \quad (۵-۲)$$

$$PNSD_i = \min_{x \in Pos(x)} M_i(x) - \max_{x \in Neg(x)} M_i(x) \quad (۵-۳)$$

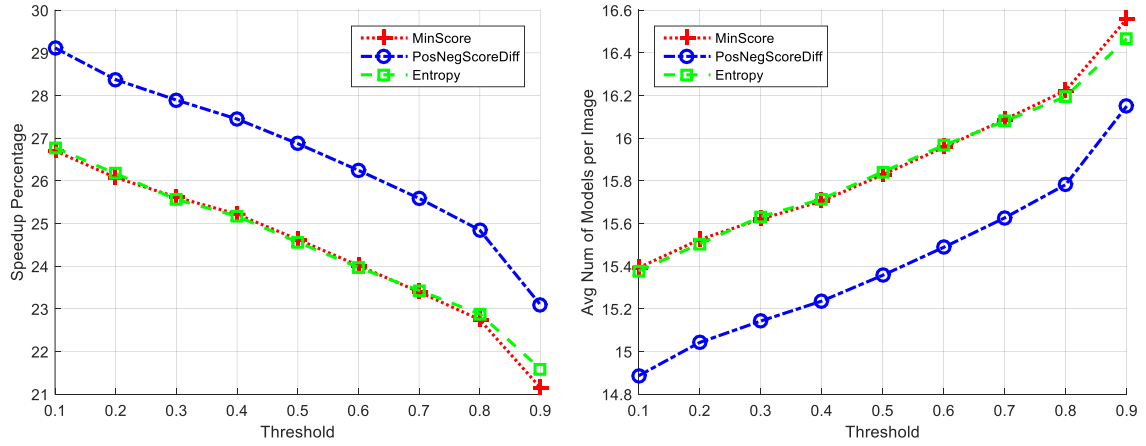
در فرمول (۵-۱)، M_i امتیاز خروجی طبقه‌بند i ام و E_i درصد اطمینان این طبقه‌بند است. برای محاسبه سه تعریف بالا از نمونه‌های آموزشی استفاده می‌شود. K تعداد نمونه‌های آموزشی و x_j یک نمونه‌ی آموزشی است. مجموعه $Pos(x)$ شامل نمونه‌های آموزشی مثبت و $Neg(x)$ شامل نمونه‌های آموزشی منفی است. با افزایش هر سه مقدار E_i، MS_i و $PNSD_i$، درصد اطمینان طبقه‌بند i ام افزایش می‌یابد.

برای آزمون عملکرد سه معیار ارائه شده، یک آزمایش ترتیب داده شده است. در این آزمایش، طبقه-بندهای از قبل آموزش داده شده بر روی ۲۱ مدل متفاوت از مجموعه داده BVMMR را در نظر گرفته-ایم. این آزمایش و سایر آزمایش‌های انجام شده در این فصل همگی بر روی نمونه‌های آموزشی صورت

گرفته‌اند. با استفاده از سه ترتیب موردنظر، آبشاری از مدل‌ها تشکیل داده و عملکرد سیستم را به ازای حدآستانه‌های سراسری متفاوت در بازه‌ی [۰/۹ – ۰/۱] مورد بررسی قرار داده‌ایم. در این بررسی چهار معیار دقت میانگین، دقت مجموع، افزایش سرعت (درصد) و میانگین تعداد طبقه‌بند اعمال شده به هر تصویر ورودی، اندازه گرفته شده است (شکل ۵-۷ و شکل ۵-۸). دقت میانگین با میانگین‌گیری از دقت همه کلاس‌ها بدست می‌آید. دقت مجموع نیز برابر است با نسبت تعداد نمونه‌های طبقه‌بندی شده درست به کل نمونه‌ها.



شکل ۵-۷: مقایسه دقت میانگین و دقت مجموع ترتیب‌های مختلف پیشنهادی به ازای حد آستانه‌های متفاوت.



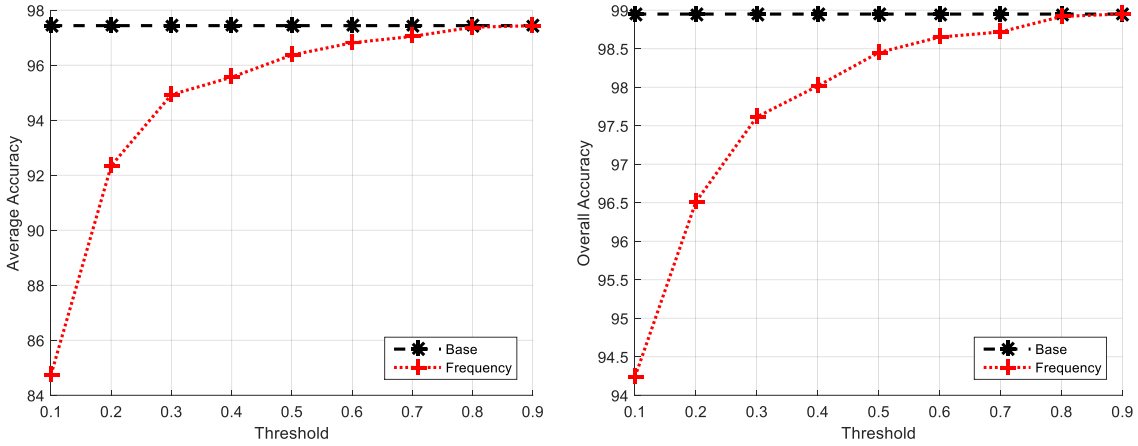
شکل ۵-۸: مقایسه درصد افزایش سرعت و میانگین تعداد طبقه‌بندهای اعمال شده به هر تصویر برای ترتیب‌های مختلف پیشنهادی به ازای حد آستانه‌های متفاوت.

در آزمایش بالا، ۲۱ طبقه‌بند وجود دارد؛ بنابراین نسخه غیرآبشاری از سیستم همه این ۲۱ طبقه‌بند را به هر تصویر ورودی اعمال خواهد کرد. نسخه غیرآبشاری در شکل ۵-۷ با عبارت Base مشخص شده است. از نتایج بدست آمده می‌توان نتیجه گرفت که هر سه روش، ترتیب مناسبی به ویژه برای حد آستانه‌های بزرگتر از 0.5 بدست می‌دهند. استفاده از حد آستانه‌ی کوچکتر از 0.5 در سیستم آبشاری پیشنهادی توصیه نمی‌شود، زیرا وابستگی سیستم به ترتیب طبقه‌بندها را به شکلی غیرمنطقی افزایش می‌دهد. با در نظرگرفتن توأمان سرعت و دقت، معیار $PosNegScoreDiff$ بهتر از دو معیار دیگر عمل کرده است. این معیار در حد آستانه 0.5 موجب افزایش چند دهم درصدی دقت میانگین و مجموع و افزایش تقریباً ۳۰ درصدی سرعت سیستم گشته است. به همین دلیل، این معیار در سیستم پیشنهادی مورد استفاده قرار گرفته است.

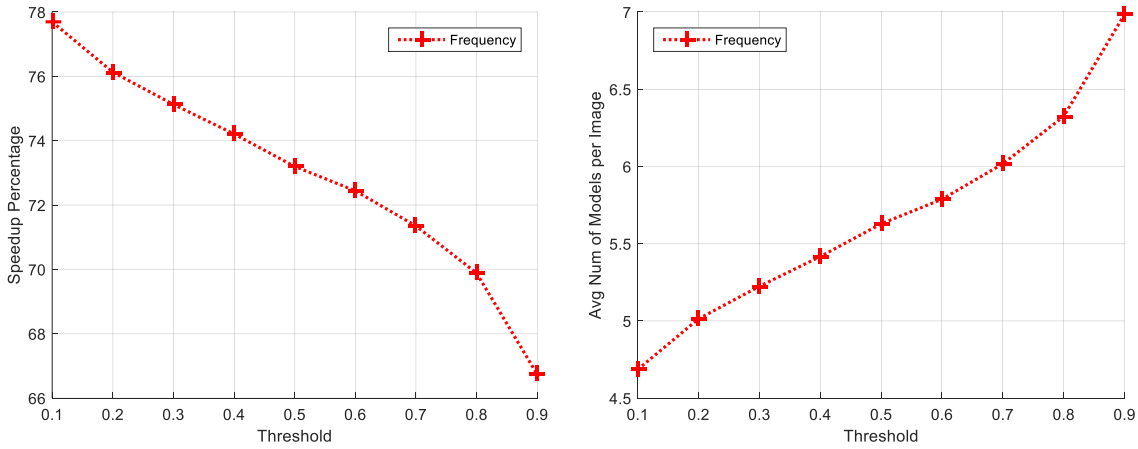
۵-۱-۲- فراوانی

توزیع مدل‌های مختلف خودرو در خیابان‌های هر کشور متفاوت است. معمولاً فراوانی مدل‌هایی که دارای بازه‌ی قیمتی پایین تا متوسط هستند بیشتر از سایر مدل‌ها است. با در نظر گرفتن توزیع مدل خودروها در مجموعه‌ی آموزشی، می‌توان یک تخمین مناسب از فراوانی انواع مدل‌ها بدست آورد. البته با این فرض که توزیع مدل‌ها در مجموعه آموزشی مشابه با توزیع مدل‌ها مجموعه آزمایشی و یا در دنیای واقعی است. با قرار دادن طبقه‌بندهای مربوط به مدل‌هایی که دارای فراوانی بیشتری هستند در ابتدای آبشار، می‌توان سرعت سیستم آبشاری را افزایش داد. برای مثال در صورتی که طبقه‌بند مربوط به کلاس «پراید ۱۳۱» را در ابتدای آبشار قرار دهیم، افزایش سرعت چندین برابری سیستم را تضمین خواهیم کرد؛ زیرا فراوانی این مدل در کشورمان همان‌طور که در مجموعه داده BVMMR (جدول ۴-۱) نیز مشاهده می‌شود، بسیار بیشتر از سایر مدل‌ها است. این معیار تنها در رابطه با مسئله شناسایی مدل خودرو صادق نیست؛ بلکه از آن می‌توان در سایر مسائل طبقه‌بندی چندکلاسه نیز بهره برد. برای مثال، در مسئله شناسایی حیوانات نیز واضح است که فراوانی برخی از انواع حیوانات از سایر آن‌ها بیشتر است.

غالباً طبقه‌بندهایی که تعداد تصاویر آموزشی چالش‌برانگیز کمتری داشته‌اند، اطمینان بیشتری دارند. در نتیجه ترتیب حاصل از اطمینان طبقه‌بندها معمولاً عکس ترتیب حاصل از فراوانی عمل می‌کند. در صورتی که طبقه‌بندها را به ترتیب فراوانی مدل‌هایشان در آبشار قرار دهیم، افزایش سرعت مناسبی بدست خواهیم آورد. شکل ۵-۹ و شکل ۵-۱۰، چهار معیار بررسی شده در آزمایش قبل را برای ترتیب حاصل از فراوانی نمایش داده‌اند.



شکل ۵-۹: مقایسه دقت میانگین و دقت مجموع ترتیب حاصل از فراوانی به ازای حد آستانه‌های متفاوت.



شکل ۵-۱۰: مقایسه درصد افزایش سرعت و میانگین تعداد طبقه‌بندهای اعمال شده به هر تصویر برای ترتیب حاصل از فراوانی به ازای حد آستانه‌های متفاوت.

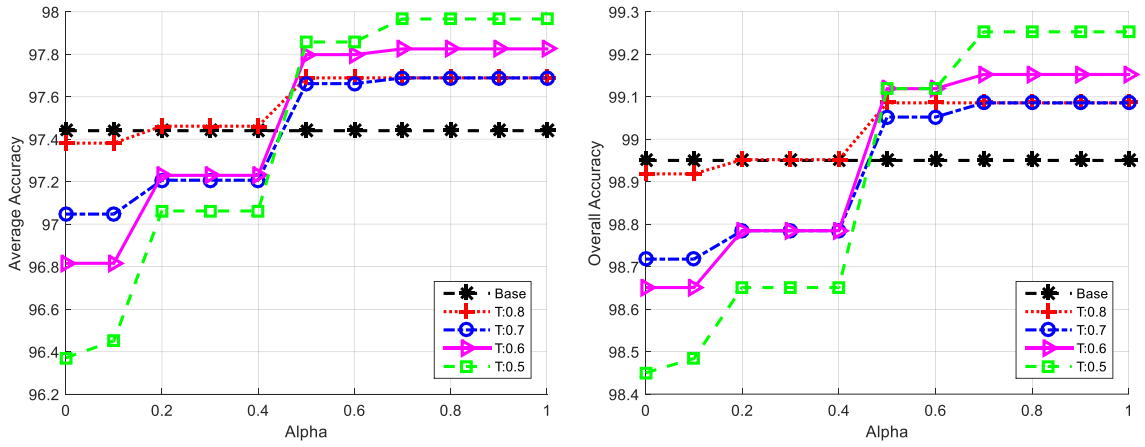
نتیجه این آزمایش همان‌طور که انتظار می‌رفت، افزایش سرعت ۶۵ الی ۸۰ درصدی سیستم را نشان می‌دهد. هر چند با افزایش هر چه بیشتر سرعت، دقت نیز کاهش پیدا کرده است. برای بدست آوردن ترکیبی از دقت و سرعت، می‌توان دو معیار معرفی شده را برای بدست آوردن یک ترتیب متعادل، توأمان به کار گرفت.

۵-۱-۳- ترکیب اطمینان و فراوانی

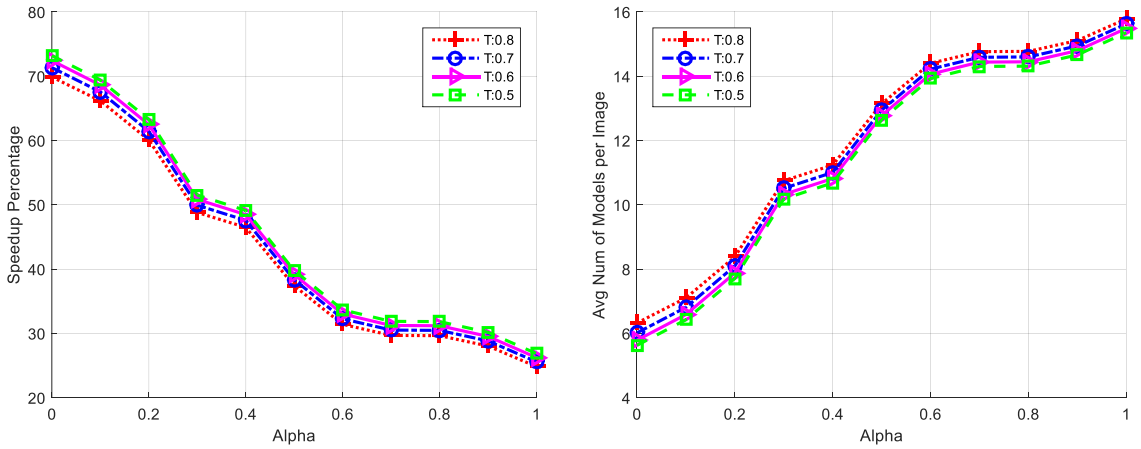
برای ترکیب دو معیار معرفی شده، دو تابع E_i و F_i را مقدار محاسبه شده برای اطمینان و فراوانی مدل i فرض می‌کنیم. مقدار خروجی این دو تابع در بازه [۰-۱] نرمال شده‌اند. فرمول (۵-۴) یک معیار ترکیبی برای ترتیب قرارگیری طبقه‌بندها در آبشار ارائه کرده است:

$$C_i = \alpha E_i + (1 - \alpha) F_i \qquad (۴-۵)$$

با تغییر ثابت α در بازه [۱-۰] می‌توان تأثیر هر یک از دو معیار در ترتیب خروجی را تعیین کرد. نتایج بدست آمده به ازای مقادیر مختلف α در شکل ۵-۱۱ و شکل ۵-۱۲ نمایش داده شده است. برای انتخاب حد آستانه سراسری باید کاملاً سختگیرانه عمل کرد؛ زیرا انتخاب یک حد آستانه کوچک به معنی احتمال پذیرفته شدن اشتباه تصویر ورودی در مراحل اولیه آبشار است که دقت سیستم را به شدت کاهش می-دهد. در طرف مقابل، مقادیر بزرگ حد آستانه منجر به کاهش کارایی آبشار شده و عملکرد سیستم آبشاری را بسیار نزدیک به سیستم غیرآبشاری می‌گرداند. به همین دلیل تنها مقادیر حد آستانه در بازه [۰/۸-۰/۵] مورد بررسی قرار گرفته‌اند.



شکل ۵-۱۱: مقایسه دقت میانگین و دقت مجموع ترتیب حاصل از معیار ترکیبی به ازای مقادیر مختلف α و حد آستانه‌ی سراسری.



شکل ۵-۱۲: مقایسه درصد افزایش سرعت و میانگین تعداد طبقه‌بندهای اعمال شده به هر تصویر برای ترتیب حاصل از معیار ترکیبی به ازای مقادیر مختلف α و حد آستانه‌ی سراسری.

در هر دو شکل ۵-۱۱ و شکل ۵-۱۲، با حرکت از سمت چپ به راست نمودارها، تأثیر فراوانی روی ترتیب نهایی کم شده و تأثیر اطمینان بیشتر می‌شود. نکته مهمی که در شکل ۵-۱۱ دیده می‌شود، تغییرات زیاد دقت سیستم در حد آستانه‌های کوچکتر از ۰٫۷ است. وقتی $\alpha < 0.5$ است، ترتیب بدست آمده وابستگی بیشتری به فراوانی دارد؛ در نتیجه، یک حد آستانه کوچک موجب افت دقت خواهد شد. دو مقدار مناسب برای α و حد آستانه عبارتند از $\alpha = 0.8$ و $T = 0.7$ که منجر به افزایش چند دهم درصدی دقت سیستم و افزایش ۳۰ درصدی سرعت سیستم خواهند شد.

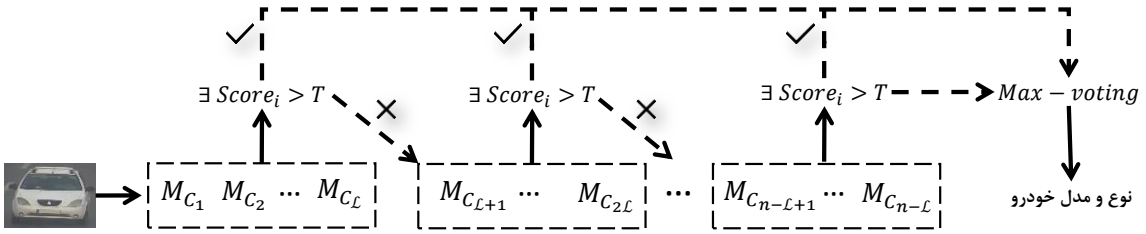
۵-۱-۴- پنجره لغزان

در الگوریتم ارائه شده در بخش قبل، وقتی حد آستانه‌ی کوچکی انتخاب می‌شود، دقت سیستم کاملاً وابسته به ترتیب طبقه‌بندها خواهد بود. برای کاهش این تأثیر ناخوشایند، نسخه عمومی‌تری از الگوریتم آبشاری ارائه شده است که نسخه غیرآبشاری و آبشاری ارائه شده در بخش‌های قبل یک حالت خاص از آن هستند. این نسخه را «آبشاری از مدل‌ها با پنجره لغزان» نامیده‌ایم.

الگوریتم آبشاری بخش قبل را در نظر بگیرید. در هر مرحله از آن الگوریتم، یک طبقه‌بند به تصویر اعمال شده و با توجه به امتیاز خروجی آن، در مورد ادامه یا توقف روند آبشار تصمیم‌گیری می‌شد. در نسخه‌ی پنجره لغزان، یک پنجره به طول $\mathscr{L}$ در ابتدای آبشار قرار داده می‌شود. در مرحله نخست آبشار، $\mathscr{L}$ طبقه-بندی که درون این پنجره قرار گرفته‌اند به تصویر ورودی اعمال می‌شوند. اگر امتیاز خروجی یک یا چند طبقه‌بند بیشتر از حد آستانه سراسری باشد، طبقه‌بندی که امتیاز بزرگتری دارد، برنده خواهد بود. در

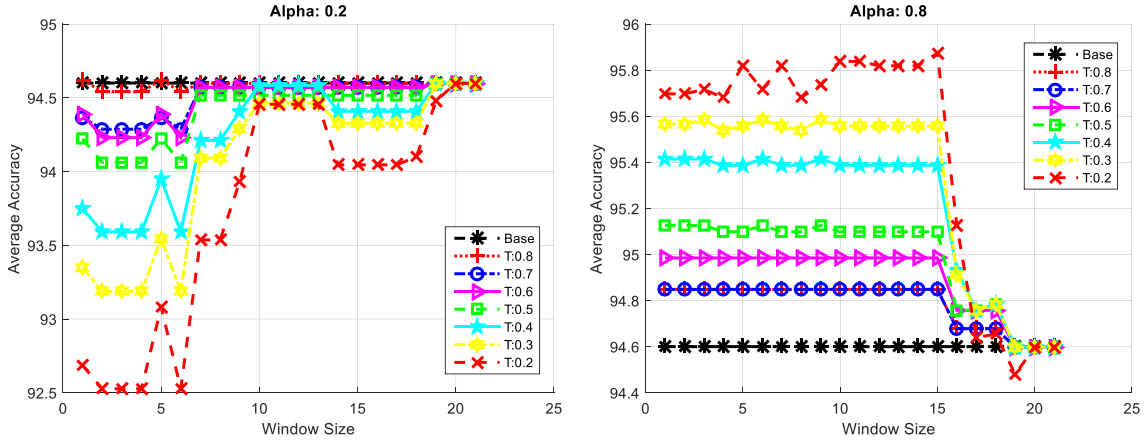
غیراینصورت، پنجره روی $\mathscr{L}$ طبقه‌بند بعدی قرار داده می‌شود (شکل ۵-۱۳). پنجره‌ها با هم همپوشانی ندارند.

در صورتی که $\mathscr{L}=1$ قرار داده شود، الگوریتم آبشاری بخش قبل را خواهیم داشت و در صورتی که $\mathscr{L}=n$ قرار داده شود، الگوریتم غیرآبشاری را بدست خواهیم آورد. اندازه پنجره، تعادلی بین سرعت و دقت برقرار خواهد کرد. هرچقدر که اندازه پنجره بزرگتر باشد، دقت بیشتر، سرعت کمتر و وابستگی به ترتیب در حد آستانه‌های کوچک، کمتر خواهد شد.

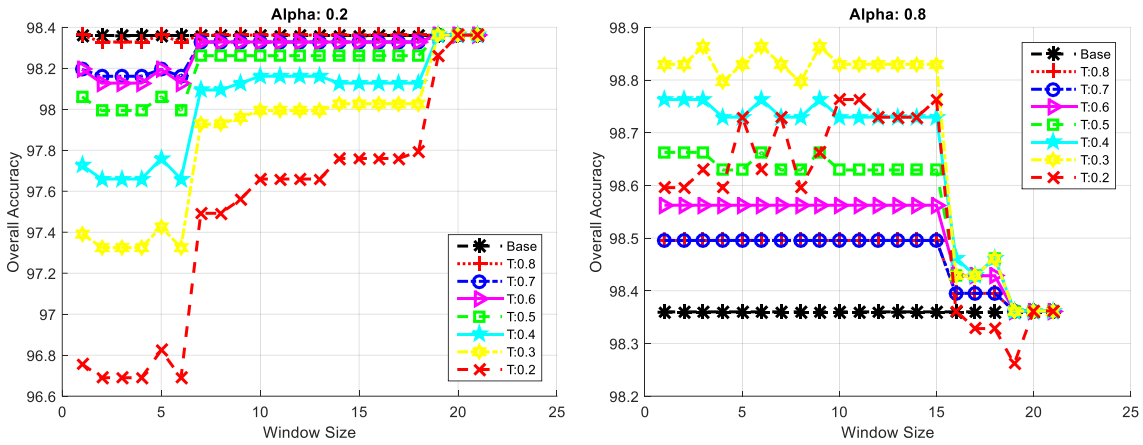


شکل ۵-۱۳: نحوه عملکرد الگوریتم آبشاری با پنجره لغزان برای تعیین نوع و مدل خودرو. پنجره‌ها با خط‌چین های مستطیلی نمایش داده شده‌اند. در هر پنجره، $\mathcal{L}$ طبقه‌بند قرار دارد.

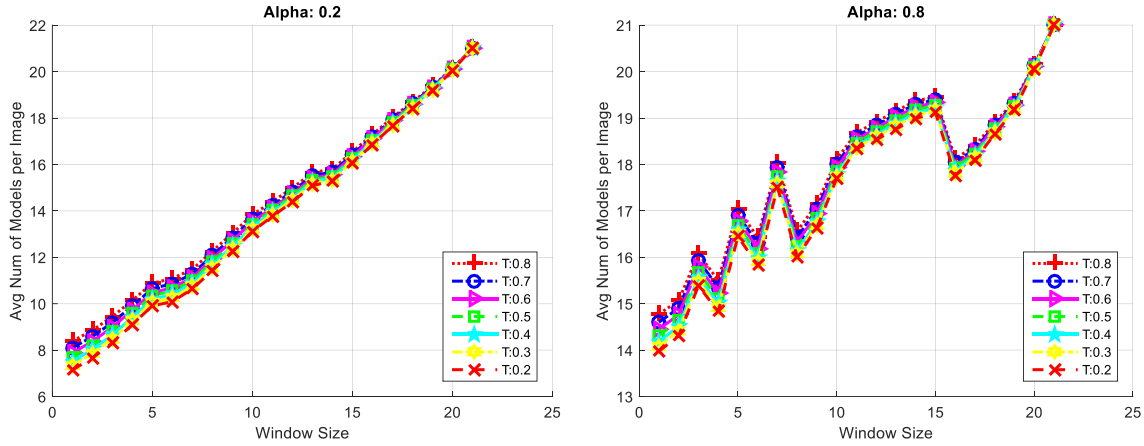
برای بررسی تأثیر اندازه پنجره بر عملکرد سیستم، آزمایش انجام شده در بخش قبل را به ازای اندازه‌های پنجره‌ی متفاوت تکرار کرده‌ایم. نتایج آزمایش برای دو مقدار $\alpha = 0.2$ و $\alpha = 0.8$ در شکل ۵-۱۴، شکل ۵-۱۵، شکل ۵-۱۶ و شکل ۵-۱۷ ارائه شده‌اند.



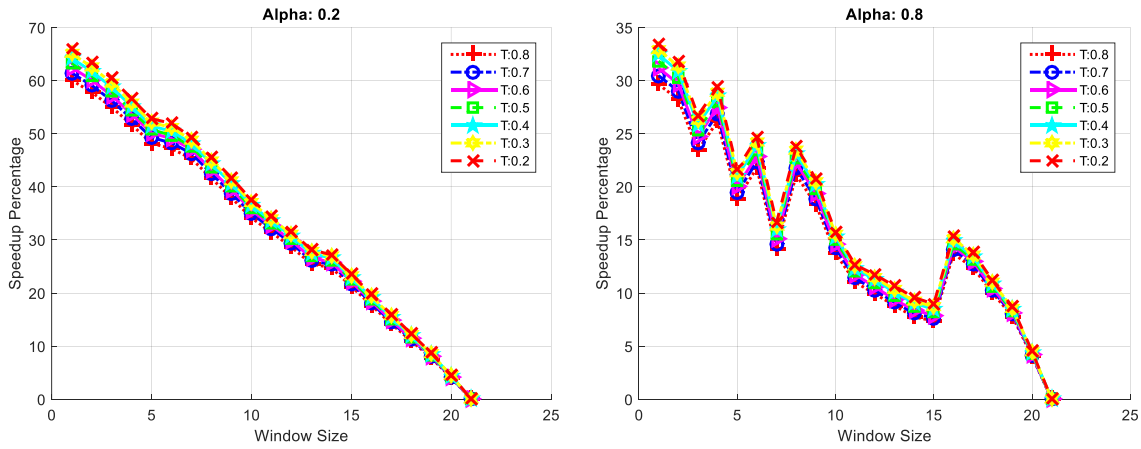
شکل ۵-۱۴: مقایسه دقت میانگین به ازای مقادیر مختلف اندازه پنجره و حد آستانه‌ی سراسری برای $\alpha = 0.2$ و $\alpha = 0.8$.



شکل ۵-۱۵: مقایسه دقت مجموع به ازای مقادیر مختلف اندازه پنجره و حد آستانه‌ی سراسری برای $\alpha = 0.2$ و $\alpha = 0.8$.



شکل ۵-۱۶: مقایسه میانگین تعداد طبقه‌بندهای اعمال شده به هر تصویر به ازای مقادیر مختلف اندازه پنجره و حد آستانه‌ی سراسری برای $\alpha = 0.2$ و $\alpha = 0.8$.



شکل ۵-۱۷: مقایسه درصد افزایش سرعت به ازای مقادیر مختلف اندازه پنجره و حد آستانه‌ی سراسری برای $\alpha = 0.2$ و $\alpha = 0.8$.

نتایج بدست آمده در شکل‌های بالا را برای دقت و سرعت به صورت مجزا تحلیل می‌کنیم. این نکته را نیز مدنظر قرار می‌دهیم که هدف از ارائه نسخه‌ی پنجره لغزان، کاهش حساسیت سیستم به ترتیب طبقه-بندها در حد آستانه‌های کوچک است. بنابراین انتظار داریم که عملکرد سیستم در آزمایش $\alpha = 0.2$ نسبت به $\alpha = 0.8$ بهبود بیشتری پیدا کند؛ زیرا ترتیب بدست آمده از $\alpha = 0.2$ تاکید کمتری روی اطمینان طبقه‌بندها دارد و در نتیجه پنجره‌گذاری تأثیر بیشتری روی آن خواهد داشت. چنان‌که در شکل ۵-۱۴ و شکل ۵-۱۵ نیز این بهبود را مشاهده می‌کنیم. در نقطه مقابل، عملکرد سیستم برای $\alpha = 0.8$ همانطور که در شکل ۵-۱۶ و شکل ۵-۱۷ نمایش داده شده است، تغییر چندانی نکرده است.

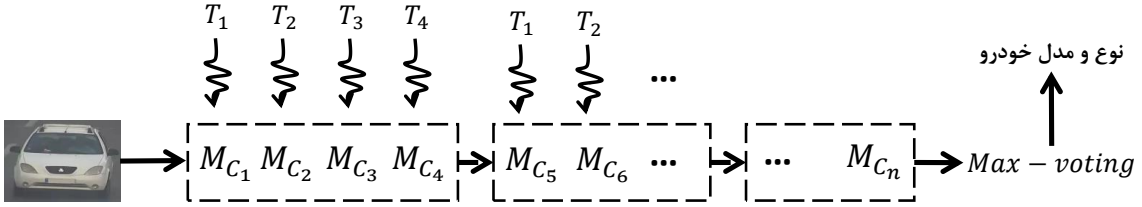
افزایش سرعت سیستم با پنجره‌گذاری کاهش می‌یابد؛ طبیعتاً هرچقدر که اندازه پنجره بزرگتر باشد، افزایش سرعت نیز کمتر می‌شود. مشاهده می‌شود که نرخ افزایش سرعت برای $\alpha = 0.8$ نوسانی و برای $\alpha = 0.2$ تقریباً ثابت است. علت نوسان نرخ افزایش سرعت برای $\alpha = 0.8$، وجود برخی طبقه‌بندها با اطمینان کمتر نسبت به سایر طبقه‌بندهای موجود در پنجره است. یک طبقه‌بند با اطمینان کمتر به ازای تصاویر ورودی بیشتر، امتیازی بزرگتر از حد آستانه سراسری اعلام کرده و منجر به افزایش سرعت بیشتری می‌گردد.

در سیستم پیشنهادی از حد آستانه‌ی سراسری بزرگتر از 0.5 استفاده می‌شود. بنابراین استفاده از اندازه پنجره‌ی بزرگتر از یک موجب افزایش دقت سیستم نخواهد شد. به همین دلیل، اندازه پنجره یک به

عنوان انتخاب نهایی در نظر گرفته شده است. محاسبه دقت سیستم برای سه رتبه اول[1] و یا پنج رتبه اول به ترتیب با استفاده از نسخه پنجره لغزان با طول پنجره سه و پنج صورت می‌گیرد.

یک کاربرد دیگر روش پنجره لغزان در موازی‌سازی الگوریتم آبشاری است. برای این منظور، طبقه‌بندهای موجود در پنجره‌ی جاری به صورت موازی به تصویر اعمال می‌شوند. پس از اجرای موازی این پنجره، در صورتی که تصویر به مراحل بعدی آبشار راه پیدا کرده باشد، پنجره بر روی طبقه‌بندهای بعدی قرار داده می‌شود. در نسخه‌ی موازی الگوریتم پنجره لغزان، یک انتخاب مناسب برای اندازه پنجره، تعداد هسته‌های پردازنده است. بالاترین کارایی در این حالت حاصل خواهد شد. اگر تعداد هسته‌های پردازنده را C عدد فرض کنیم، در اولین مرحله از الگوریتم، پنجره‌ای به طول C بر روی C طبقه‌بند ابتدای آبشار قرار داده می‌شود. سپس این پنجره حرکت کرده و بر روی C طبقه‌بند بعدی قرار می‌گیرد. این روال تا متوقف شدن آبشار ادامه پیدا می‌کند. الگوریتم به کار گرفته شده برای موازی‌سازی نسخه آبشاری در شکل ۵-۱۸ نمایش داده شده است. در این شکل، تعداد هسته‌های پردازنده برابر با چهار فرض شده است ($C = 4$).

به صورت کلی، نسخه موازی سیستم آبشاری کارایی برابری نسبت به نسخه موازی سیستم غیرآبشاری نخواهد داشت؛ به دلیل ماهیت ترتیبی آبشار، نمی‌توان حداکثر کارایی حاصل از موازی‌سازی را با چنین الگوریتم ساده‌ای و یا سایر الگوریتم‌های مشابه بدست آورد.



شکل ۵-۱۸: الگوریتم موازی‌سازی سیستم آبشاری. تعداد هسته‌های پردازنده، چهار عدد فرض شده است.

۵-۲- مدل‌های آبشاری

می‌دانیم که هر مدل از یک فیلتر ریشه و چند فیلتر بخش تشکیل شده است. برای محاسبه امتیاز یک پنجره، این فیلترها به صورت همزمان به تصویر اعمال شده و نتایج آن‌ها با هم ترکیب می‌شوند. این روال دارای پیچیدگی محاسباتی بالایی است و به زمان زیادی نیاز دارد. از طرفی می‌دانیم که در یک تصویر

[1] Rank-3, Top 3

ورودی که به روش پنجره لغزان پیمایش می‌گردد، تعداد پنجره‌هایی که منجر به یک شناسایی درست می‌شوند و یا حتی در آن‌ها یک خودرو وجود دارد بسیار کم است. کنار گذاشتن تعداد زیادی از این پنجره‌ها با پردازش کم، افزایش سرعت چندبرابری را در پی خواهد داشت.

به صورت کلی، اغلب طبقه‌بندها با استفاده از الگوریتم مناسب قابل تبدیل به نسخه‌ی آبشاری‌شان هستند. دو نمونه از این الگوریتم‌ها در [۵۵] و [۱۰۲] ارائه گشته است. همان‌طور که روشی مناسب برای تبدیل یک مدل مبتنی بر بخش به یک مدل آبشاری در [۱۰۳] ارائه شده است. الگوریتم پیشنهادی که در ادامه تشریح شده، نسخه‌ای مشابه با همین روش است. در الگوریتم مبتنی بر هرم تصاویر یا پنجره لغزان، یک پنجره در تمامی مکان‌ها و اندازه‌های تصویر قرار داده شده و وجود یا عدم وجود شیء در آن پنجره بررسی می‌گردد. یک مدل مبتنی بر بخش نیز شبیه به همین الگوریتم ولی پیچیده‌تر عمل می‌کند؛ در چنین مدلی، یک فیلتر ریشه و چند فیلتر بخش وجود دارد. پنجره‌هایی توسط این مدل پذیرفته می‌شوند که اعمال توأمان فیلتر ریشه و بخش‌ها به آن‌ها، امتیازی بالاتر از یک حد آستانه مشخص تولید نماید. در واقع پنجره‌های ریشه‌ای، شامل خودرو در نظر گرفته می‌شوند که مجموعه‌ی بخش‌ها با امتیاز بالا در آن پنجره تطبیق پیدا کرده باشند. نیازی به انجام این پردازش سنگین در همه پنجره‌ها نیست، زیرا درصد بسیار کمی از کل پنجره‌ها شامل خودرو هستند.

یک مدل که از یک فیلتر ریشه r و N بخش $\{p_1, \dots, p_N\}$ تشکیل شده را دارای $N+1$ بخش در نظر می‌گیریم. رابطه هندسی بخش‌های این مدل توسط ساختار ستاره‌ای توصیف شده‌اند. در چنین ساختاری، مکان هر بخش تنها به مکان ریشه وابسته است و بین خود بخش‌ها هیچ وابستگی وجود ندارد. بنابراین می‌توان بخش‌ها را به صورت مستقل به تصویر اعمال کرد. برای ارائه الگوریتم آبشاری، ابتدا نحوه محاسبه امتیاز یک پنجره توسط مدل غیرآبشاری M را مرور می‌کنیم (فرمول‌های (۵-۵) و (۵-۶)):

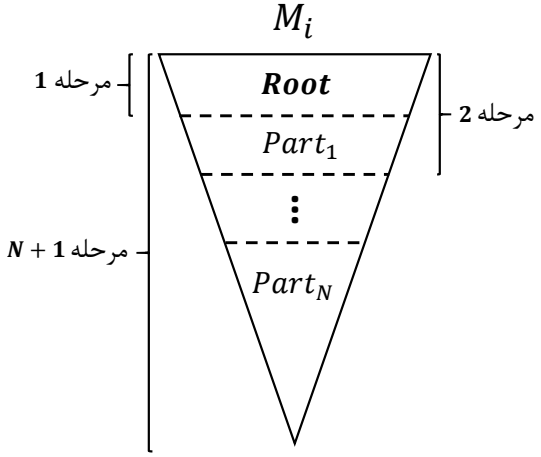
$$Score_M(x,y) = m_r(x,y) + \sum_{i=1}^{N} \max_{dx_{pi}, dy_{pi}} Score_{pi}(x,y,dx_{pi},dy_{pi}) \qquad (۵-۵)$$

$$Score_{pi}(x,y,dx,dy) = m_{pi}(x,y,dx,dy) - d_{pi}.\phi(dx,dy)) \qquad (۵-۶)$$

x و y مکان قرار گیری پنجره ریشه هستند. $m_r(x,y)$ امتیاز فیلتر ریشه در مکان (x,y) از تصویر را برمی‌گرداند. $m_{pi}(x,y,dx_{pi},dy_{pi})$ امتیاز بهترین قرارگیری بخش pi در پنجره ریشه‌ای که در مکان (x,y) از تصویر قرار دارد را بدست می‌دهد. بهترین قرارگیری بخش pi در مکانی با (dx_{pi},dy_{pi})

جابجایی نسبت به مکان ریشه حاصل می‌شود. تابع درجه دوم ϕ خطای جابجایی بخش ϕ و بردار d_{pi} ضرایب این تابع هستند.

برای بدست آوردن یک مدل آبشاری، این $N+1$ بخش را با یک ترتیب مشخص در کنار هم قرار داده و آن‌ها را به ترتیب به تصویر اعمال می‌کنیم. فیلتر ریشه را در مرحله‌ی اول از آبشار و سایر بخش‌ها را در مراحل بعدی قرار می‌دهیم. در این حالت، در واقع بجای یک مدل، $N+1$ مدل داریم. مدل اول، بدون بخش. مدل دوم با یک بخش. مدل سوم با دو بخش و مدل آخر با $N+1$ بخش (شکل ۵-۱۹). تصاویری که توسط مدل‌های ضعیف‌تر آغازین امتیاز قبولی را کسب می‌کنند برای مدل‌های قوی‌تر ارسال می‌شوند. به دلیل استفاده از ساختار ستاره‌ای، نیازی به محاسبه مجدد امتیازات محاسبه شده در مراحل قبلی آبشار نیست. برای این $N+1$ مرحله از آبشار به $2N$ حد آستانه نیاز داریم. در هر مرحله از آبشار با استفاده از دو حد آستانه به فیلتر کردن پنجره‌های منفی می‌پردازیم که به صورت زیر محاسبه می‌شوند:



شکل ۵-۱۹: ساختار کلی یک مدل آبشاری.

$$x_i = m_r(x,y) + \sum_{j=1}^{i-1} Score_{pj}(x,y) \qquad \textbf{(۷-۵)}$$

$$x_i' = t_i - d_{pi}.\phi(dx_{pi}, dy_{pi}) \qquad \textbf{(۸-۵)}$$

x_i امتیاز مدل i ام با یک ریشه و $i-1$ بخش است و x_i' همان x_i بدون احتساب خطای جابجایی بخش i ام می‌باشد. در صورتی که امتیاز یک پنجره توسط مدل i ام کوچکتر از x_i باشد، آن پنجره کنار گذاشته شده و به مراحل بعدی آبشار راه نمی‌یابد. علاوه‌براین، در مرحله جستجو برای مکان بهینه بخش i ام از مدل مدل $i+1$ ام، در صورتی که خطای جابجای این بخش موجب کوچک‌تر شدن امتیاز فعلی مدل از حد آستانه x_i' شود، مکان موردنظر نادیده گرفته شده و جستجو برای سایر مکان‌ها ادامه می‌یابد. وقتی

یک پنجره به مدل موجود در مرحله آخر آبشار می‌رسد، امتیاز کسب شده توسط آن با حد آستانه اصلی مدل یعنی T مقایسه خواهد شد (حد آستانه مدل غیرآبشاری).

برای بدست آوردن مجموعه حد آستانه‌های میانی $\{t_1, t_1', t_2, t_2', \dots, t_N, t_N'\}$ از مجموعه داده‌های آموزشی استفاده می‌کنیم. یک زوج حد آستانه میانی t_i و t_i' با خطای کم، حد آستانه‌هایی است که هیچ نمونه‌ی مثبتی را در مراحل میانی آبشار، فیلتر نکند. به بیانی دیگر، برای هیچ نمونه‌ی مثبتی، $x_i < t_i$ و $x_i' < t_i'$ نباشد. یک زوج حد آستانه میانی مناسب که دو شرط بالا را برای همه نمونه‌های آموزشی رعایت کنند، به صورت زیر محاسبه می‌شوند:

$$t_i = \min_{x \in X} x_i \qquad t_i' = \min_{x \in X} x_i' \tag{۹-۵}$$

X مجموعه داده‌ی آموزشی است. در صورتی که مجموعه X به اندازه کافی بزرگ باشد، حد آستانه‌های مناسبی بدست خواهند آمد. الگوریتم ۵-۱ طرز عملکرد یک مدل آبشاری با $2N$ حد آستانه میانی و یک حد آستانه سراسری را ارائه کرده است.

الگوریتم ۵-۱: الگوریتم تشخیص توسط یک مدل آبشاری که $2N$ حد آستانه میانی و یک حد آستانه سراسری را دریافت کرده و مجموعه‌ای از مکان‌های تشخیص داده شده را بازمی‌گرداند.

```
Function D = cascadeDetection(t_1, t_1', t_2, t_2', ..., t_N, t_N', T)
   D: ∅
   for ∀ (x, y) in all locations and scales of the image
      score = m_r(x, y)
      for i = 1:N
         if score < t_i      break
         p = −∞
         for ∀ (dx_pi, dy_pi) in root window
            if (score − d_pi. ϕ(dx_pi, dy_pi)) < t_i'      continue
            p = max(p, Score_pi(x, y, dx_pi, dy_pi))
         end
         score = score + p
      end
      if score ≥ T        D = D ∪ (x, y)
   end
end
```

الگوریتم ۵-۱ مجموعه مکان‌هایی از تصویر که توسط مدل آبشاری به عنوان نمونه مثبت تشخیص داده شده‌اند را در مجموعه D قرار داده و بازمی‌گرداند. از آنجا که اندازه فیلترها ثابت است، اندازه پنجره مشخص بوده و قید نشده است. در دو شرط اول از این الگوریتم، فیلتر کردن بر اساس حد آستانه‌های میانی t_i و t_i' صورت می‌گیرد.

برای مقایسه افزایش سرعت حاصل شده، نسخه پایه سیستم (شکل ۵-۱) را در نظر گرفته و هر طبقه‌بند را با نسخه آبشاری آن جایگزین کردیم. در آزمایش انجام شده بر روی نسخه پایه، ابتدا ۲۱ طبقه‌بند آموزش دیده شده با استفاده از مجموعه داده BVMMR که هر یک دارای نه بخش (یک ریشه و هشت بخش) هستند، را به تصاویر آزمایشی اعمال کردیم. بنابراین $21 \times 9 = 189$ فیلتر به هر تصویر ورودی اعمال می‌شود. نسخه آبشاری به صورت میانگین، ۳۹ فیلتر را به تصویر اعمال می‌کند؛ یعنی تقریباً برابر با اعمال تنها چهار طبقه‌بند به تصویر (به جای ۲۱ طبقه‌بند). در واقع، نسخه آبشاری تقریباً ٪۸۰ فیلترها را به تصویر اعمال نمی‌کند. سپس ۱۸۰ طبقه‌بند آموزش دیده شده بر روی مجموعه داده CompCars را به تصاویر آزمایشی اعمال کردیم. در این حالت، $180 \times 9 = 1620$ فیلتر به هر تصویر ورودی اعمال می‌شود. نسخه آبشاری از طبقه‌بندها به صورت میانگین، ۴۶۵ فیلتر را به هر تصویر اعمال می‌کند؛ یعنی ٪۷۱ صرفه‌جویی در اعمال فیلترها صورت می‌گیرد. این مقدار صرفه‌جویی در اعمال فیلترها در هر دو آزمایش، افزایش سرعت بالای ٪۵۰ را به همراه خواهد داشت. کمتر بودن درصد افزایش سرعت نسبت به درصد صرفه‌جویی گزارش شده، به دلیل زمان‌برتر بودن اعمال فیلتر ریشه نسبت به فیلتر بخش‌ها است. فیلتر ریشه در تمامی اندازه‌ها به تمامی مکان‌های تصویر اعمال می‌شود.

موردی که در [۱۰۳] به آن پرداخته نشده است، ترتیب قرارگیری بخش‌ها در آبشار است. البته با توجه به الگوریتم آبشاری ارائه شده در این بخش، انتظار نداریم که ترتیب‌های متفاوت منجر به تفاوت بارز در دقت بدست آمده شوند. برای ارائه ترتیب بخش‌ها، فرض می‌کنیم که هر بخش به تمامی تصاویر آموزشی اعمال و امتیازش محاسبه شده است. سه ترتیب زیر را برای بخش‌ها ارائه کردیم:

۱. بخشی را در ابتدای آبشار قرار بده که واریانس امتیازش از سایر بخش‌ها کمتر باشد.
۲. بخشی را در ابتدای آبشار قرار بده که میانگین امتیازش از سایر بخش‌ها بیشتر باشد.
۳. بخشی را در ابتدای آبشار قرار بده که میانگین امتیازش از سایر بخش‌ها کمتر باشد.

ترتیب اول، بخشی را در ابتدای آبشار قرار می‌دهد که اختلاف امتیازش با میانگین امتیاز بخش‌های آن مدل بیشینه باشد. هدف از ترتیب دوم، فیلتر کردن پنجره‌های بیشتر در مراحل اولیه آبشار است، زیرا

بخشی برای مراحل اولیه انتخاب می‌شود که سخت‌گیرتر است. ترتیب سوم معکوس ترتیب دوم عمل کرده و بخش ساده‌گیرتر را در ابتدا قرار می‌دهد. انتظار داریم که دقت بدست آمده از ترتیب سوم کمی بهتر از ترتیب دوم و سرعت آن کمی کمتر باشد. برای مقایسه این سه ترتیب، طبقه‌بندهای آبشاری بدست آمده از این سه ترتیب را در سیستم پایه جایگذاری کرده و دقت و سرعت را برای داده‌های آموزشی مجموعه داده BVMMR بدست آوردیم. نتایج این آزمایش در جدول ۵-۱ ارائه شده‌اند.

جدول ۵-۱: مقایسه تأثیر ترتیب‌های مختلف قرارگیری بخش‌ها در مدل آبشاری بر روی دقت و سرعت سیستم.

شماره	ترتیب	دقت میانگین	دقت مجموع	میانگین تعداد فیلترهای اعمال شده به هر تصویر	درصد تعداد فیلترهای اعمال نشده به کل فیلترها
۱	واریانس	۹۱/۷۷٪	۹۷/۸۰٪	۳۹/۲۴	۷۹/۲۴٪
۲	بیشینه	۹۲/۱۰٪	۹۷/۴۶٪	۳۶/۲۰	۸۰/۸۴٪
۳	کمینه	۹۱/۴۴٪	۹۷/۷۳٪	۵۰/۰۲	۷۳/۵۳٪

همان‌طور که پیش‌بینی شده بود، ترتیب کمینه، بالاترین دقت و ترتیب بیشینه، بالاترین سرعت را بدست دادند. هر چند تفاوت دقت و سرعت ترتیب‌های مختلف بسیار ناچیز است. در سیستم پیشنهادی از ترتیب بیشینه به دلیل سرعت بیشتر و سادگی نسبت به ترتیب واریانس بهره برده شده است.

دقت‌های گزارش شده کمتر از دقت‌های مورد انتظار هستند. می‌دانیم که یک مدل آبشاری با $N+1$ بخش به $2N$ حد آستانه میانی نیاز دارد. به دلیل کم بودن تعداد نمونه‌های برخی کلاس‌ها، محاسبه حد آستانه‌های میانی با دقت بالا قابل حصول نیست. از این رو، دقت مدل آبشاری روی این کلاس‌ها کاهش می‌یابد که موجب افت دقت میانگین سیستم می‌گردد.

۵-۳- ایجاد آبشاری از مدل‌های آبشاری

در بخش‌های قبل، الگوریتم‌های «آبشاری از مدل‌ها» و «مدل‌های آبشاری» تشریح شدند. برای افزایش هر چه بیشتر سرعت سیستم، این دو الگوریتم را در هم ادغام می‌کنیم. با تشکیل آبشاری از مدل‌های آبشاری می‌توانیم در دو سطح به صرفه‌جویی در حجم محاسبات انجام شده بپردازیم؛ در سطح اول، آبشار تشکیل شده از این مدل‌ها تنها کسری از کل مدل‌ها را به تصویر ورودی اعمال می‌کند. در سطح دوم، هر مدل آبشاری مقدار زیادی از محاسباتش را در اغلب مکان‌های تصویر ورودی اعمال نمی‌نماید. انتظار نداریم که پس از ترکیب این دو الگوریتم، افزایش سرعتی به اندازه مجموع افزایش سرعت‌های منفرد این

دو بدست آوریم. زیرا در آبشاری از مدل‌ها، تنها زیرمجموعه‌ای از کل مدل‌ها به تصویر اعمال شده و در نتیجه افزایش سرعت حاصل از مدل‌های آبشاری تنها در این زیرمجموعه اثرگذار خواهد بود.

سیستم پایه به صورت موازی عمل می‌کند؛ یعنی طبقه‌بندها به صورت موازی به تصویر ورودی اعمال می‌شوند. آزمایش‌ها بر روی کامپیوتری با پردازنده Corei7-4710HQ 2.5 GHz انجام شده است. بنابراین حداکثر افزایش سرعت حاصل از موازی‌سازی در حالت ایده‌آل، چهار برابر خواهد بود. به منظور آزمایش عادلانه سرعت نسخه آبشاری و غیرآبشاری سیستم، زمان اجرای نسخه موازیِ الگوریتم آبشاری نیز ارائه شده است.

جدول ۵-۲ و جدول ۵-۳ نتیجه آزمایش‌های انجام شده برای سنجش دقت و سرعت نسخه‌های موازی و غیرموازی سیستم‌های آبشاری و غیرآبشاری را بر روی دو مجموعه داده BVMMR و CompCars ارائه کرده است. این نتایج با استفاده از نمونه‌های آزمایشی اندازه‌گیری شده‌اند. زمان‌های گزارش شده، میانگین زمان پردازش هر تصویر توسط سیستم است. از آنجا که برخی تصاویر شامل بیش از یک خودرو بوده‌اند، برای قابل قیاس بودن زمان‌ها، زمان کل بدست آمده نسبت به تعداد کل خودروها، میانگین‌گیری شده است. علاوه‌براین، برای نسخه‌ی آبشاری سیستم، دو مقدار α آزمایش شده است تا تأثیر ترتیب طبقه‌بندها روی سرعت و دقت سیستم بهتر دیده شود. حد آستانه‌ی سراسری مورد استفاده برای نسخه‌های آبشاری برابر با $T = 0.7$ بوده است. سیستم عامل مورد استفاده Windows 10 بوده و هیچ تنظیمات خاصی به سیستم پیشنهادی اعمال نشده است.

جدول ۵-۲: مقایسه دقت و سرعت نسخه‌های مختلف سیستم پیشنهادی بر روی مجموعه داده BVMMR.

شماره	سیستم	α	زمان (ثانیه)	دقت میانگین	دقت مجموع
۱	غیرآبشاری – غیرموازی	-	۱۱	۹۶/۸۲٪	۹۸/۶۶٪
۲	غیرآبشاری – موازی	-	۳,۱	۹۶/۸۲٪	۹۸/۶۶٪
۳	آبشاری با مدل‌های غیرآبشاری – غیرموازی	۰,۸	۵,۸	۹۷/۰۱٪	۹۸/۸۰٪
۴	آبشاری با مدل‌های غیرآبشاری - غیرموازی	۰	۱,۷	۹۶/۴۲٪	۹۸/۶۶٪
۵	**آبشاری با مدل‌های غیرآبشاری – موازی**	**۰,۸**	**۲,۶**	**۹۷/۰۱٪**	**۹۸/۸۰٪**
۶	آبشاری با مدل‌های غیرآبشاری – موازی	۰	۰,۸	۹۶/۴۲٪	۹۸/۶۶٪
۷	آبشاری با مدل‌های آبشاری – غیرموازی	۰,۸	۳,۵	۹۲/۱۵٪	۹۷/۸۳٪
۸	آبشاری با مدل‌های آبشاری – غیرموازی	۰	۰,۸	۹۱/۴۶٪	۹۷/۷۰٪
۹	آبشاری با مدل‌های آبشاری - موازی	۰,۸	۱,۲	۹۲/۱۵٪	۹۷/۸۳٪
۱۰	**آبشاری با مدل‌های آبشاری - موازی**	**۰**	**۰,۲**	**۹۱/۴۶٪**	**۹۷/۷۰٪**

زمان تشخیص خودرو در زمان‌های گزارش شده در جدول ۵-۲ لحاظ نگردیده است. بهترین دقت توسط نسخه شماره پنج از سیستم حاصل شده است. این نسخه با استفاده از الگوریتم آبشاری، دقت میانگین را در حدود ٪۰/۲ بهبود داده و سرعتی بیشتر از حتی نسخه موازی سیستم غیرآبشاری دارد. در صورتی که سرعت اولویت بالاتری نسبت به دقت داشته باشد، می‌توان با افت جزئی دقت، از نسخه شماره ده استفاده کرد که سرعتی ۲۰۰ میلی‌ثانیه‌ای برای پردازش هر تصویر بدست می‌دهد. این سرعت با توجه به پیچیدگی طبقه‌بندها، بسیار مناسب است. علت اختلاف زیاد بین دقت میانگین نسخه شماره پنج و ده، محاسبه ضعیف حد آستانه‌های میانی (برای مدل‌های آبشاری) برای کلاس‌هایی با تعداد نمونه‌های کم آموزشی در مجموعه داده BVMMR بوده است. این مشکل در مجموعه داده CompCars وجود ندارد. درضمن، به دلیل کم بودن تعداد تصاویر این کلاس‌ها، با طبقه‌بندی اشتباه یک تصویر، دقت میانگین کاهش زیادی خواهد داشت؛ در حالی که دقت مجموع تغییر چندانی نخواهد کرد. چنان‌که می‌بینیم دقت مجموع نسخه شماره پنج و ده اختلاف اندکی با یکدیگر دارند.

جدول ۵-۳: مقایسه دقت و سرعت نسخه‌های مختلف سیستم پیشنهادی بر روی مجموعه داده CompCars.

شماره	سیستم	α	زمان (ثانیه)	دقت میانگین	دقت مجموع
۱	غیرآبشاری – غیرموازی	-	۳۰۰	٪۹۶/۷۲	٪۹۶/۷۲
۲	غیرآبشاری - موازی	-	۹۰	٪۹۶/۷۲	٪۹۶/۷۲
۳	آبشاری با مدل‌های غیرآبشاری – غیرموازی	۰٫۸	۱۴۵	٪۹۶/۷۴	٪۹۶/۷۱
۴	آبشاری با مدل‌های غیرآبشاری - غیرموازی	۰	۱۳۵	٪۹۶/۵۴	٪۹۶/۵۶
۵	**آبشاری با مدل‌های غیرآبشاری – موازی**	**۰٫۸**	۶۰	**٪۹۶/۷۴**	**٪۹۶/۷۱**
۶	آبشاری با مدل‌های غیرآبشاری – موازی	۰	۵۰	٪۹۶/۵۴	٪۹۶/۵۶
۷	آبشاری با مدل‌های آبشاری – غیرموازی	۰٫۸	۹۵	٪۹۵/۶۳	٪۹۵/۹۳
۸	آبشاری با مدل‌های آبشاری – غیرموازی	۰	۸۵	٪۹۵/۵۲	٪۹۵/۷۷
۹	آبشاری با مدل‌های آبشاری - موازی	۰٫۸	۲۵	٪۹۵/۶۳	٪۹۵/۹۳
۱۰	**آبشاری با مدل‌های آبشاری - موازی**	**۰**	۱۵	**٪۹۵/۵۲**	**٪۹۵/۷۷**

با افزایش تعداد طبقه‌بندها، زمان پردازش سیستم نیز افزایش پیدا خواهد کرد. سیستم آموزش داده شده بر روی مجموعه داده BVMMR، دارای ۲۱ طبقه‌بند است؛ در حالی که سیستم آموزش داده شده بر روی مجموعه داده CompCars دارای ۱۸۰ طبقه است. از طرفی ابعاد تصاویر مجموعه داده CompCars تقریباً ۱ و نیم برابر تصاویر مجموعه داده BVMMR می‌باشند. نسخه پنجم از الگوریتم آبشاری پیشنهادی موجب بهبود جزئی دقت میانگین و افزایش تقریباً ۳۰ درصدی سرعت سیستم گشته است. مشابه با جدول قبل، نسخه دهم از الگوریتم آبشاری بهترین سرعت را حاصل کرده که تقریباً ۸۰ درصد

نسبت به سیستم پایه سریع‌تر و تنها یک درصد دقت کمتری نسبت به آن دارد. زمان ۱۵ ثانیه برای پردازش یک تصویر زیاد بنظر می‌رسد؛ هرچند یک ماشین بردار پشتیبان استاندارد که به صورت چندکلاسه (یک در برابر همه) بر روی این مجموعه داده آموزش دیده شده باشد، زمانی در حدود ۳۰ ثانیه برای پردازش هر تصویر نیاز دارد.

۵-۴- جمع‌بندی

در این فصل به ارائه یک الگوریتم آبشاری برای بهبود سرعت سیستم پرداخته شد. الگوریتم آبشاری پیشنهادی به جای اعمال همزمان همه طبقه‌بندها به تصویر، آن‌ها را به ترتیب اعمال می‌کند. در این حالت، ترتیب اعمال طبقه‌بندها اهمیت پیدا خواهد کرد. بنابراین چند معیار برای محاسبه ترتیب‌های کارا برای بهبود سرعت و دقت پیشنهاد گشت. الگوریتم آبشاری با استفاده از ترتیب بدست آمده از ترکیبی از این معیارها و یک حد آستانه سراسری، سرعت پردازش تصویر را بدون افت دقت بهبود می‌بخشد. برای افزایش هرچه بیشتر سرعت، هر طبقه‌بند با معادل آبشاری خود جایگزین می‌شود. به این ترتیب، سرعت الگوریتم آبشاری چندین برابر افزایش پیدا می‌کند؛ در صورتی که تنها یک افت جزئی در دقت سیستم اتفاق می‌افتد.

پس از معرفی نسخه اولیه الگوریتم آبشاری، یک نسخه عمومی‌تر به نام «آبشاری از مدل‌ها با پنجره لغزان» معرفی شد. الگوریتم غیرآبشاری و نسخه اولیه الگوریتم آبشاری همگی یک حالت خاص از الگوریتم پنجره لغزان هستند. این الگوریتم امکان موازی‌سازی روش را نیز فراهم می‌آورد. برای این منظور می‌توان طبقه‌بندهای موجود در هر پنجره را به صورت موازی به تصویر اعمال کرد.

همه نسخه‌های معرفی شده با استفاده از مجموعه داده BVMMR مورد ارزیابی قرار گرفتند. در پایان، زمان پردازش و دقت شناسایی سیستم با استفاده و بدون استفاده از الگوریتم آبشاری برای دو مجموعه داده BVMMR و CompCars محاسبه و گزارش شده است. نتایج مناسب الگوریتم آبشاری بر روی این دو مجموعه داده‌ی کاملاً متفاوت، نشان از کارایی بالای آن دارد.

فصل

۶- ارزیابی و نتایج آزمایش‌ها

۶-۱- مقدمه

در این فصل به ارزیابی جامع بخش‌های مختلف سیستم پیاده‌سازی شده در رساله پرداخته شده است. آزمایش‌ها از بخش پیش‌پردازش یعنی تشخیص وسیله نقلیه آغاز می‌شوند. سپس تفاوت عملکرد رویکرد مبتنی بر بخش و رویکرد کلی‌نگر با یک آزمایش نشان داده می‌شود. در مسیر حرکت به سمت استخراج خودکار بخش‌ها، ابتدا استخراج دستی بخش‌ها با یک روش ساده برای استخراج خودکار مورد مقایسه قرار گرفته است. آنگاه سیستم نهایی که شامل استخراج مجموعه بخش‌های متمایزکننده‌ای برای هر مدل از خودروها است، بر روی دو پایگاه داده BVMMR و CompCars مورد ارزیابی قرار می‌گیرد. در پایان این فصل، مقایسه‌ای بین روش پیشنهادی و چند رویکرد پراستفاده در روش‌های پیشین انجام می‌شود. انتظار داریم که در پایان این فصل بتوان شناخت مناسبی از عملکرد همه بخش‌های سیستم پیشنهادی پیدا کرد.

۶-۲- ارزیابی سیستم تشخیص وسیله نقلیه

بخش تشخیص وسیله نقلیه در مرحله پیش‌پردازش به تصویر اعمال شده و خودروهای موجود در تصویر را استخراج می‌نماید. هدف از اضافه کردن این بخش به سیستم، کاهش پردازش‌های انجام شده توسط بخش‌های بعدی است؛ به این ترتیب، طبقه‌بندهای شناسایی کننده نوع و مدل خودرو تنها به بخش‌هایی از تصویر که شامل خودرو هستند اعمال خواهند شد. شکل ۶-۱ چند نمونه از نتایج این سیستم را به نمایش گذاشته است.

شکل ۶-۱: چند نمونه از نتایج سیستم تشخیص خودرو بر روی دو تصویر از دو دوربین متفاوت از مجموعه داده BVMMR.

برای اندازه‌گیری دقت بخش تشخیص خودرو، سیستم آموزش دیده شده بر روی ۲۹۹۱ تصاویر آموزشی از مجموعه داده BVMMR را در نظر گرفته و آن را به ۳۰۰۰ تصاویر آزمایشی همین مجموعه داده

اعمال کردیم. تعداد ۵۰۰۰ نمونه منفی که توسط سیستم در مرحله آموزش دیده نشده‌اند نیز در این آزمایش مورد استفاده قرار گرفته است. حد آستانه‌ی استفاده شده در این آزمایش برابر با $0.2-$ قرار داده شده است. بنابراین پنجره‌هایی که امتیاز آن‌ها از این مقدار بیشتر باشد به عنوان خودرو شناخته می‌شوند. نتیجه این آزمایش در جدول ۶-۱ ارائه گشته است.

جدول ۶-۱: ارزیابی سیستم تشخیص خودرو بر روی مجموعه داده BVMMR.

	کل	طبقه‌بندی به عنوان مثبت	طبقه‌بندی به عنوان منفی
تعداد نمونه‌های مثبت	۳۰۰۰	(TP) ۳۰۰۰	(FN) ۰
تعداد نمونه‌های منفی	۵۰۰۰	(FP) ۳	(TN) ۴۹۹۷

جدول ۶-۱ عملکرد بسیار عالی سیستم تشخیص خودرو را نشان می‌دهد. در ادامه دو معیار صحت[۱] و بازیابی[۲] با توجه به فرمول‌های (۶-۱) و (۶-۲) برای این سیستم محاسبه شده است. صحت سیستم برابر با ٪۹۹/۹۰ و بازیابی برابر با ٪۱۰۰ می‌باشد.

$$Precision = \frac{TP}{TP + FP} \quad (۶-۱)$$

$$Recall = \frac{TP}{TP + FN} \quad (۶-۲)$$

۶-۳- مقایسه رویکرد مبتنی بر بخش و رویکرد کلی‌نگر

۶-۳-۱- علامت‌گذاری دستی بخش‌ها

این آزمایش در مراحل اولیه توسعه سیستم انجام گرفت و هدف از آن، نشان دادن کارایی یک سیستم مبتنی بر بخش به ساده‌ترین روش ممکن بود. در این آزمایش، برخی از بخش‌های تشکیل دهنده خودرو در نمای جلو و پشت به صورت دستی علامت‌گذاری شدند. این بخش‌ها در ادامه فهرست شده‌اند که البته پلاک در آزمایش‌ها وارد نشده و تنها برای حذف محدوده پلاک از خودروها مورد استفاده قرار گرفته است.

- نمای جلو: پلاک، چراغ‌های چپ و راست، جلوپنجره و سپر
- نمای پشت: پلاک، چراغ‌های چپ و راست، صندوق و سپر

[۱] Precision
[۲] Recall

برای مقایسه کارایی رویکرد مبتنی بر بخش، از تک‌تک بخش‌های علامت‌گذاری شده و یا ترکیبی از آن‌ها برای طبقه‌بندی نوع و مدل وسایل نقلیه استفاده می‌شود. در طرف مقابل، از ویژگی‌های استخراج شده از ناحیه مطلوب و کل تصویر نیز به عنوان دو راه‌حل پرکاربرد توسط روش‌های پیشین، بهره خواهیم برد. این دو راه‌حل در واقع نمایندگان رویکرد کلی‌نگر در نظر گرفته شده‌اند. برای نمایش بهتر تفاوت دو رویکرد، یک بار، کل ناحیه مطلوب برای طبقه‌بندی و یک بار، بخش‌های تشکیل دهنده ناحیه مطلوب به صورت مجزا برای طبقه‌بندی استفاده گشته است. شکل ۶-۲ تفاوت این دو حالت را به تصویر کشیده است. معقول بنظر می‌رسد که ردیف پایینی از شکل ۶-۲ منجر به طبقه‌بندی با دقت بالاتری گردد، زیرا کاپوت دارای ویژگی یا ویژگی‌های متمایزکننده‌ی مناسبی نیست.

شکل ۶-۲: تفاوت رویکرد کلی‌نگر که از کل ناحیه مطلوب (ردیف بالا) برای استخراج ویژگی بهره می‌برد و رویکرد مبتنی بر بخش که از زیرمجموعه‌ای از اجزای تشکیل دهنده این ناحیه (ردیف پایین) استفاده می‌کند.

در این آزمایش، از مجموعه داده اول (BVMMR v1) به دلیل وجود بخش‌های علامت‌گذاری شده استفاده کرده‌ایم. تعداد نه مدل از خودروهایی که دارای بیش از ۲۴ نمونه (۱۲ نمونه از نمای جلو و ۱۲ نمونه از نمای پشت) بوده‌اند، برای انجام آزمایش انتخاب شده‌اند. برای محاسبه دقت[۱] شناسایی از روش ارزیابی ضربدری به پیمانه سه[۲] و به صورت نشان داده شده در فرمول (۶-۳) بهره گرفته شده است. p تعداد نمونه‌های درست طبقه‌بندی شده و N تعداد کل نمونه‌ها است.

$$Accuracy = \frac{p}{N} \qquad (۶-۳)$$

برای نمای جلو و پشت، دو آزمایش مجزا ترتیب داده شده است. در هر دو آزمایش، از توصیفگر هیستوگرام گرادیان‌های جهت‌دار با اندازه سلول 12×12 و اندازه بلوک 2×2 بهره برده شده است. این مقادیر به صورت تجربی و با توجه به نتایج ارائه شده در [۷۶] انتخاب گشته‌اند. با توجه به ابعاد بزرگتر تصاویر مجموعه داده اول نسبت به آنچه در [۷۶] مورد بررسی قرار گرفته است، اندازه سلول 12×12 عملکرد بهتری در مقایسه با اندازه سلول 8×8 دارد. برای طبقه‌بندی از ماشین بردار پشتیبان به صورت

[۱] Accuracy

[۲] 3-Fold Cross Validation

یک در برابر همه[1] و کتابخانه LIBSVM [۸۵] استفاده شده است. جدول ۶-۲، نتیجه آزمایش انجام شده بر روی نمای جلو و جدول ۶-۳، نتیجه آزمایش انجام شده بر روی نمای پشت را ارائه می‌دهند. دقت‌های گزارش شده، میانگین دقت حاصل از ارزیابی ضربدری هستند.

جدول ۶-۲: ارزیابی رویکرد مبتنی بر بخش در مقایسه با رویکرد کلی‌نگر بر روی تصاویر مجموعه داده اول از نمای جلو.

بخش‌های مورد استفاده	دقت
سپر	۸۷/۳٪
جلوپنجره	۹۳/۵٪
چراغ چپ	۹۶/۲٪
چراغ راست	۹۶/۲٪
چراغ چپ و راست	۹۷/۲٪
چراغ چپ و راست و جلوپنجره	۹۸/۸٪
چراغ چپ و راست، جلوپنجره و سپر	۹۷/۲٪
ناحیه مطلوب (کاپوت، چراغ‌ها، جلوپنجره و نشان‌واره)	۹۴/۴٪
کل تصویر	۷۸/۷٪

جدول ۶-۳: ارزیابی رویکرد مبتنی بر بخش در مقایسه با رویکرد کلی‌نگر بر روی تصاویر مجموعه داده اول از نمای پشت.

بخش‌های مورد استفاده	دقت
سپر	۸۸/۸٪
چراغ چپ	۹۶/۲٪
چراغ راست	۹۶/۸٪
چراغ چپ و راست	۹۸/۱٪
چراغ چپ و راست و سپر	۹۷/۲٪
ناحیه مطلوب (صندوق، چراغ‌ها، فضای بین دو چراغ و سپر)	۹۷/۲٪
کل تصویر	۸۴/۲٪

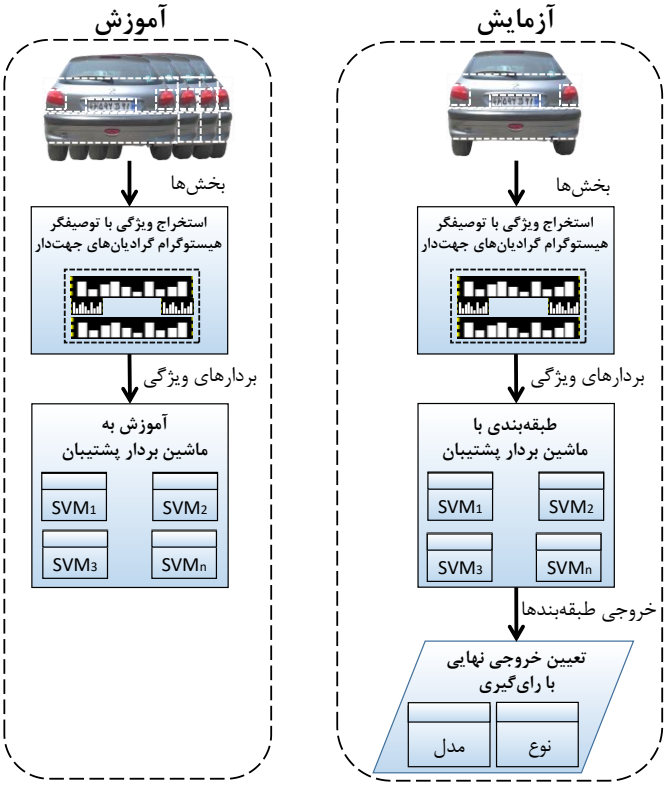
در سطرهایی از جدول ۶-۲ و جدول ۶-۳ که دو یا چند بخش در کنار هم قرار گرفته‌اند، بردارهای ویژگی این بخش‌ها به هم چسبانده شده و به صورت یکجا به ماشین بردار پشتیبان آموزش داده شده‌اند. منظور از ناحیه مطلوب یک ناحیه‌ای مستطیلی شکل است که بخش‌های قرار گرفته در آن در داخل پرانتز قید شده‌اند.

[1] One Against All

اولین نکته‌ای که از نتیجه آزمایش‌ها می‌توان استخراج کرد، عملکرد مناسب بخش‌های منفرد است. چراغ‌ها به تنهایی توانسته‌اند نزدیک به ٪۹۷ مدل‌ها را به درستی از یکدگیر تفکیک نمایند. در نمای جلو، به دلیل وجود کاپوت و سپر در ناحیه مطلوب، دقت بدست آمده از چراغ‌ها به تنهایی بهتر از دقت حاصل از ناحیه مطلوب است. سپر در هر دو آزمایش کمترین دقت را بدست آورده که کاملاً قابل پیش‌بینی بود.

نکته‌ای دوم، عملکرد بهتر اجزای تشکیل دهنده ناحیه مطلوب به صورت مجزا نسبت به کل ناحیه مطلوب است. در نمای جلو، دقت حاصل از اجزای ناحیه مطلوب برابر با ٪۹۸/۸ و دقت حاصل از خود ناحیه مطلوب برابر با ٪۹۴/۴ گزارش شده که نشان از برتری رویکرد مبتنی بر بخش دارد. وجود کاپوت در ناحیه مطلوب موجب ایجاد چنین اختلافی گشته است. چنان‌که می‌بینیم در نمای پشت، دقت ناحیه مطلوب فاصله کمتری نسبت به دقت بخش‌های تشکیل دهنده آن دارد.

برای استفاده بهتر از قدرت تمایز بخش‌های مختلف، آزمایش دیگری ترتیب داده شد. در این آزمایش، سه ماشین بردار پشتیبان برای سه بخش بهتر از نمای جلو و پشت آموزش داده شدند. در مرحله‌ی آزمایش، بین خروجی سه طبقه‌بند، رأی‌گیری انجام شده و خروجی غالب انتخاب می‌گردد. شکل ۶-۳ مراحل الگوریتم مورد استفاده در این آزمایش را در مرحله آموزش و آزمایش نمایش داده است. در این شکل، SVM_i با استفاده از ویژگی‌های استخراج شده از بخش i آموزش داده می‌شود. پارامترهای آزمایش قبلی در این آزمایش نیز بدون هیچ تغییری به کار گرفته شده‌اند. جدول ۶-۴ نتیجه این آزمایش را ارائه کرده است.



شکل ۶-۳: مراحل الگوریتم مورد استفاده برای ترکیب قدرت تمایز بخش‌های مختلف.

جدول ۶-۴: استفاده از چند طبقه‌بند برای چند بخش متفاوت و رأی‌گیری بین آن‌ها برای استفاده بهتر از قدرت تمایز هر بخش.

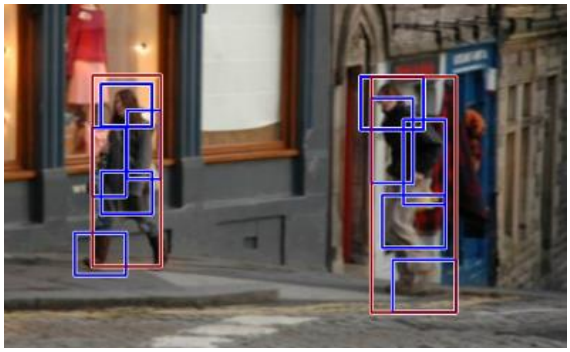
نما	بخش‌های مورد استفاده	دقت
جلو	چراغ‌های چپ و راست و جلوپنجره	۱۰۰
پشت	چراغ‌های چپ و راست و سپر	۱۰۰

همان‌طور که در جدول ۶-۴ مشاهده می‌شود، دقت در هر دو نما بهبود پیدا کرده است. می‌توان نتیجه گرفت که شکل استفاده از بخش‌ها نیز در بدست آوردن دقت مناسب، اهمیت ویژه‌ای دارد. البته رسیدن به دقت ٪۱۰۰ بیشتر به دلیل تعداد کم کلاس‌ها در مجموعه داده اول بوده است تا بدون نقص بودن روش پیشنهادی. هدف از این آزمایش تنها نشان دادن بهبود دقت سیستم نسبت به آزمایش قبل می‌باشد.

۶-۳-۲- علامت‌گذاری خودکار بخش‌های مشابه

در بخش قبل، برخی از بخش‌هایی که بنظر قدرت تمایز مناسبی داشتند به صورت دستی علامت‌گذاری شده و برای طبقه‌بندی مورد استفاده قرار گرفتند. در این بخش، تعدادی از بخش‌های مشابه بین همه وسایل نقلیه به صورت خودکار استخراج گشته و آزمایش قبل برای آن‌ها تکرار شده است. هدف از این آزمایش، مقایسه دقت طبقه‌بندی حاصل از استخراج دستی و خودکار بخش‌ها است.

برای استخراج خودکار بخش‌های مشابه، یک طبقه‌بند مبتنی بر بخش را به روش ارائه شده در [۲۶] و با استفاده از مجموعه داده BVMMR آموزش می‌دهیم. این طبقه‌بند در واقع تنها یک تشخیص دهنده خودرو است و بخش‌های مناسب برای تشخیص یک خودرو را نیز آموزش می‌بیند (شکل ۶-۴). این بخش‌ها در واقع، تأثیرگذارترین بخش‌ها برای تفکیک یک خودرو و هر شیء غیرخودرو است (از دیدگاه طبقه‌بند آموزش دیده شده). بنابراین کاملاً واضح است که این مجموعه از بخش‌ها برای جداسازی مدل-های مختلف خودرو ایده‌آل نیستند، زیرا شباهت زیادی بین آن‌ها وجود دارد. این قضیه در رابطه با آزمایش‌های انجام شده در بخش قبل هم صادق است. به همین دلیل، مقایسه این روش و روش آزمایش شده در بخش قبل، مقایسه‌ی عادلانه‌ای است.

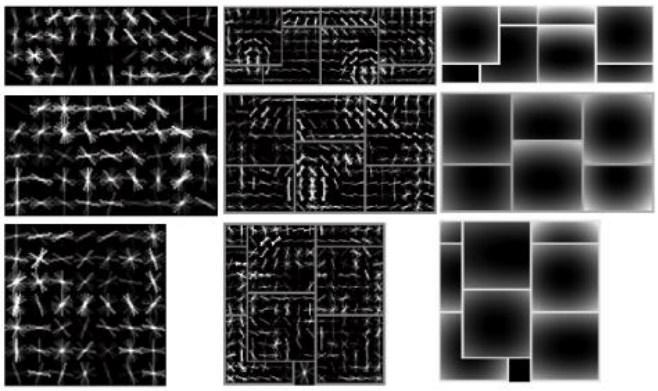


شکل ۶-۴: روش تشخیص شیء ارائه شده در [۲۶]، علاوه‌بر تشخیص شیء، بخش‌های متمایزکننده آن را نیز استخراج می‌نماید.

روش ارائه شده در [۲۶] برای اینکه بتواند اشیائی با اندازه و زوایای مختلف در تصویر را بیابد، برای هر کلاس از اشیاء، سه مدل (برای سه زاویه‌ی متفاوت) با هشت بخش آموزش می‌بیند (شکل ۶-۵). برای تشخیص زاویه خودروها، آن‌ها را بر اساس نسبت ابعادشان به سه دسته تقسیم می‌کند. روش مذکور با تغییرات زیر در این آزمایش به کار گرفته شده است.

استفاده از سه مدل برای استخراج بخش‌ها ممکن نیست. زیرا مجموعه بخش‌های استخراج شده از خودروها باید قابل انطباق با یکدیگر باشند. در صورت استفاده از سه مدل، ممکن است در یک خودرو، مدل اول بر تصویر منطبق شده و در خودرویی دیگر، مدل دوم یا سوم. از آنجا که زاویه وسایل نقلیه موجود در مجموعه داده دوم دارای تغییرات اندکی است، استفاده از یک مدل، اختلالی در دقت سیستم ایجاد نخواهد کرد. بنابراین برای نمای جلو، یک مدل با پنج بخش و برای نمای پشت یک مدل با چهار بخش آموزش داده شد. این بخش‌ها در ادامه فهرست شده‌اند. مکان بخش‌ها به صورت خودکار تعیین شده‌اند.

- نمای جلو: بخش پایین-راست (شامل چراغ راست)، بخش پایین-چپ (شامل چراغ چپ)، بخش بالا-راست، بخش بالا-چپ و بخش میانی (شامل جلوپنجره)
- نمای پشت: بخش پایین-راست (شامل چراغ راست)، بخش پایین-چپ (شامل چراغ چپ)، بخش بالا-راست، بخش بالا-چپ

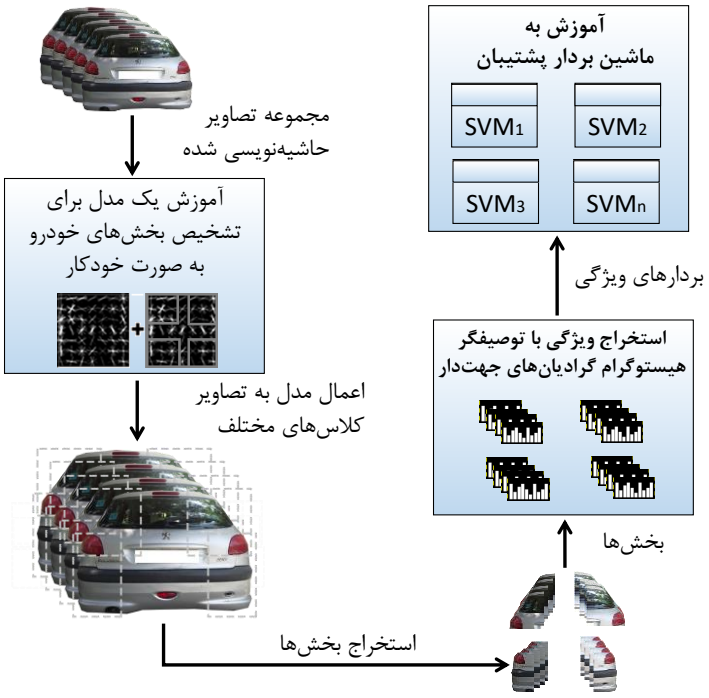


شکل ۶-۵: سه مدل آموزش دیده شده برای تشخیص خودرو از زاویه‌های متفاوت (هر سطر یک مدل) در [۲۶]. ستون اول: فیلتر ریشه، ستون دوم: فیلتر بخش‌ها، ستون سوم: خطای جابجایی مکان نسبی بخش نسبت به ریشه.

تغییر دیگری که در روش موردنظر اعمال شده است، اضافه کردن شماره هر بخش به خروجی الگوریتم است. هر بخش یک شماره یکتا خواهد داشت تا بخش‌های استخراج شده از خودروهای مختلف، قابل انطباق برهم باشند. شکل ۶-۶ چند نمونه از بخش‌های تشخیص داده شده توسط دو مدل آموزش دیده شده را به نمایش گذاشته است.

شکل ۶-۶: بخش‌های استخراج شده توسط دو مدل آموزش دیده شده از چند نمونه تصویر از نمای جلو و پشت.

برای استفاده از بخش‌های مورد اشاره در طبقه‌بندی نوع و مدل خودروها، ابتدا دو مدل آموزش دیده شده به همه تصاویر آموزشی اعمال می‌شوند. همه بخش‌های مشابه در یک مجموعه مشترک قرار گرفته و از همه آن‌ها توسط توصیفگر هیستوگرام گرادیان‌های جهت‌دار، ویژگی استخراج می‌گردد. آنگاه، هر مجموعه از ویژگی‌ها به یک ماشین بردار پشتیبان مجزا آموزش داده می‌شوند. شکل ۶-۷ مراحل الگوریتم آموزش هر طبقه‌بند را از ابتدا تا انتها به تصویر کشیده است. در این شکل، مجموعه ویژگی‌های استخراج شده از بخش i به طبقه‌بند SVM_i آموزش داده می‌شود.



شکل ۶-۷: مراحل الگوریتم آموزش یک طبقه‌بند به ازای هر یک از مجموعه بخش‌های مشترک استخراج شده به صورت خودکار.

مانند آزمایش‌های انجام شده در بخش قبل، دقت حاصل از هر بخش به صورت منفرد و ترکیبی از بخش-ها نیز گزارش شده است. برای استخراج ویژگی و طبقه‌بندی نیز دقیقاً از پارامترهای بخش قبل بهره گرفته شده است. جدول ۶-۵ و جدول ۶-۶، نتیجه آزمایش‌ها را در مقایسه با استخراج دستی بخش‌ها ارائه کرده است.

جدول ۶-۵: مقایسه دقت حاصل از استخراج دستی و خودکار بخش‌ها بر روی تصاویر مجموعه داده اول از نمای جلو.

بخش‌های مورد استفاده (دستی)	دقت	بخش‌های مورد استفاده (خودکار)	دقت
سپر	۸۷/۳٪	بالا-چپ	۷۸/۵٪
-	-	بالا-راست	۷۸/۷٪
جلوپنجره	۹۳/۵٪	میانه (جلوپنجره)	۹۵/۱٪
چراغ چپ	۹۶/۲٪	پایین-چپ (چراغ چپ)	۹۵/۴٪
چراغ راست	۹۶/۲٪	پایین-راست (چراغ راست)	۹۵/۴٪
چراغ چپ و راست	۹۷/۲٪	پایین-چپ، پایین-راست	۹۵/۷٪
چراغ چپ و راست و جلوپنجره	۹۸/۸٪	پایین-چپ، پایین-راست و میانه	۹۵/۹٪
چراغ چپ و راست، جلوپنجره و سپر	۹۷/۲٪	پایین-چپ، پایین-راست، بالا-چپ، بالا-راست و میانه	۹۶/۲٪

جدول ۶-۶: مقایسه دقت حاصل از استخراج دستی و خودکار بخش‌ها بر روی تصاویر مجموعه داده اول از نمای پشت.

بخش‌های مورد استفاده (دستی)	دقت	بخش‌های مورد استفاده (خودکار)	دقت
سپر	۸۸/۸٪	بالا-چپ	۸۹/۳٪
-	-	بالا-راست	۸۹/۴٪
چراغ چپ	۹۶/۲٪	پایین-چپ (چراغ چپ)	۹۳/۵٪
چراغ راست	۹۶/۸٪	پایین-راست (چراغ راست)	۹۳/۸٪
چراغ چپ و راست	۹۸/۱٪	پایین-چپ، پایین-راست	۹۵/۷٪
چراغ چپ و راست و سپر	۹۷/۲٪	پایین-چپ، پایین-راست، بالا-چپ، بالا-راست	۹۷/۵٪

مقایسه نتایج بدست آمده از دو روش استخراج بخش‌ها، نشان از کارایی مناسب روش استخراج خودکار دارد. با وجود ایده‌آل نبودن بخش‌های خودکار استخراج شده، دقت بدست آمده از این روش، فاصله چندانی نسبت به حالت دستی ندارد. با بهبود روش استخراج بخش‌ها به شکلی که برای هر مدل از خودروها، یک مجموعه بخش مجزا آموزش دیده شود، می‌توان به دقت بسیار بهتری دست یافت.

۶-۴- ارزیابی سیستم پیشنهادی

در این بخش به ارزیابی جامع سیستم پیاده سازی شده در این رساله پرداخته شده است. آزمایش بر روی دو مجموعه داده متفاوت انجام شده تا عملکرد سیستم در دو شرایط متفاوت سنجیده شود. مجموعه داده نخست، نسخه دوم از مجموعه داده تهیه شده توسط ما است (BVMMR). مجموعه داده دوم، مجموعه داده بزرگ CompCars است که در ابتدای فصل ۴ معرفی گشت.

در هر آزمایش، سه فاکتور گزارش می‌شود. دقت میانگین، دقت مجموع و ماتریس درهم‌ریختگی[۱]. همانطور که در فصل‌های پیشین هم اشاره شد، دقت میانگین با میانگین‌گیری از دقت همه کلاس‌ها بدست می‌آید. دقت مجموع نیز برابر است با نسبت تعداد نمونه‌های طبقه‌بندی شده درست به کل نمونه-ها. در صورتی که تعداد نمونه‌های همه‌ی کلاس‌ها نزدیک به هم باشند، دقت میانگین و مجموع نیز مشابه خواهند بود. دقت میانگین و مجموع برای رتبه اول، سه رتبه اول و پنج رتبه اول گزارش می‌شوند. این شکل از گزارش دقت در حوزه طبقه‌بندی اشیاء معمول است. سه رتبه اول و پنج رتبه اول عملکرد سیستم در پیشنهادهای بعدی را نیز مورد سنجش قرار می‌دهند.

در هر دو مجموعه داده، نمونه‌های آموزشی و آزمایشی مشخص شده‌اند؛ بنابراین برای قابل مقایسه بودن نتایج ارائه شده در این بخش و کارهای آینده‌ای که بر روی همین مجموعه داده‌ها انجام خواهد گرفت، ما

[۱] Confusion Matrix

نیز از همین بخش‌بندی بهره برده‌ایم. در مجموعه داده BVMMR، تقریباً نیمی از داده‌ها برای آموزش و نیمه دیگر برای آزمایش تعیین شده‌اند. در مجموعه داده CompCars نسبت داده‌های آموزشی به آزمایشی تقریباً برابر با ۷۰ به ۳۰ درصد است.

تنظیمات سیستم برای هر دو مجموعه داده یکسان است. برای استخراج ویژگی از توصیفگر هیستوگرام گرادیان‌های جهت‌دار با اندازه سلول 8×8 و اندازه بلوک 2×2 استفاده شده است. این مقادیر به دلیل نتایج مناسب ارائه شده در [۷۶] و به صورت تجربی انتخاب گشته‌اند. خروجی همه‌ی طبقه‌بندهای آموزش دیده شده، با استفاده از الگوریتم مقیاس‌گذاری پلت [۹۷] و بر روی نمونه‌های آموزشی همان مجموعه داده، کالیبره شده‌اند. بنابراین، خروجی همه طبقه‌بندها در بازه [۱-۰] قرار دارد. از حد آستانه ۰,۱ نیز به عنوان حد آستانه داخلی هر طبقه‌بند استفاده شده است؛ در صورتی که امتیاز یک طبقه‌بند از این مقدار کمتر باشد، فرض بر عدم شناسایی تصویر توسط طبقه‌بند مورد اشاره گذاشته می‌شود. این فرض تنها در مجموعه داده BVMMR که دارای کلاس «سایر» می‌باشد، مورد استفاده قرار گرفته است.

۶-۴-۱- ارزیابی روی مجموعه داده BVMMR

ارزیابی بر روی مجموعه داده BVMMR، بر روی ۲۱ کلاس که دارای بیش از ۲۰ نمونه تصویر بوده‌اند، صورت گرفته است. آموزش طبقه‌بند مناسب برای کلاس‌هایی با کمتر از این تعداد نمونه امکان‌پذیر نیست. سایر تصاویر در کلاس «سایر» قرار داده شده‌اند تا دشواری مسئله افزایش یابد. فراوانی نمونه‌های هر کلاس در جدول ۴-۱ ارائه شده است. با این تفاوت که تعداد تصاویر آزمایشی کلاس «سایر» به ۱۴۲ عدد افزایش یافته است.

دقت شناسایی برای سه نسخه متفاوت از سیستم گزارش شده است:

۱. نسخه پایه (غیرآبشاری) سیستم که از هیچ‌گونه آبشاری استفاده نمی‌کند. این نسخه دارای دقت بالا و کمترین سرعت در بین سه نسخه است.
۲. نسخه «آبشاری از مدل‌های غیرآبشاری» که طبقه‌بندهای غیرآبشاری را به صورت آبشاری به تصویر اعمال می‌کند. این نسخه دارای بالاترین دقت و سرعت مناسب است.
۳. نسخه «آبشاری از مدل‌های آبشاری» که طبقه‌بندهای آبشاری را به صورت آبشاری به تصویر اعمال می‌کند. این نسخه دارای کمترین دقت و بالاترین سرعت است.

جدول ۶-۷ دقت بدست آمده برای هر یک از کلاس‌ها توسط سیستم غیرآبشاری را نمایش داده است. جدول ۶-۸ دقت میانگین و مجموع نسخه غیرآبشاری را برای رتبه اول، سه رتبه اول و پنج رتبه اول گزارش کرده است.

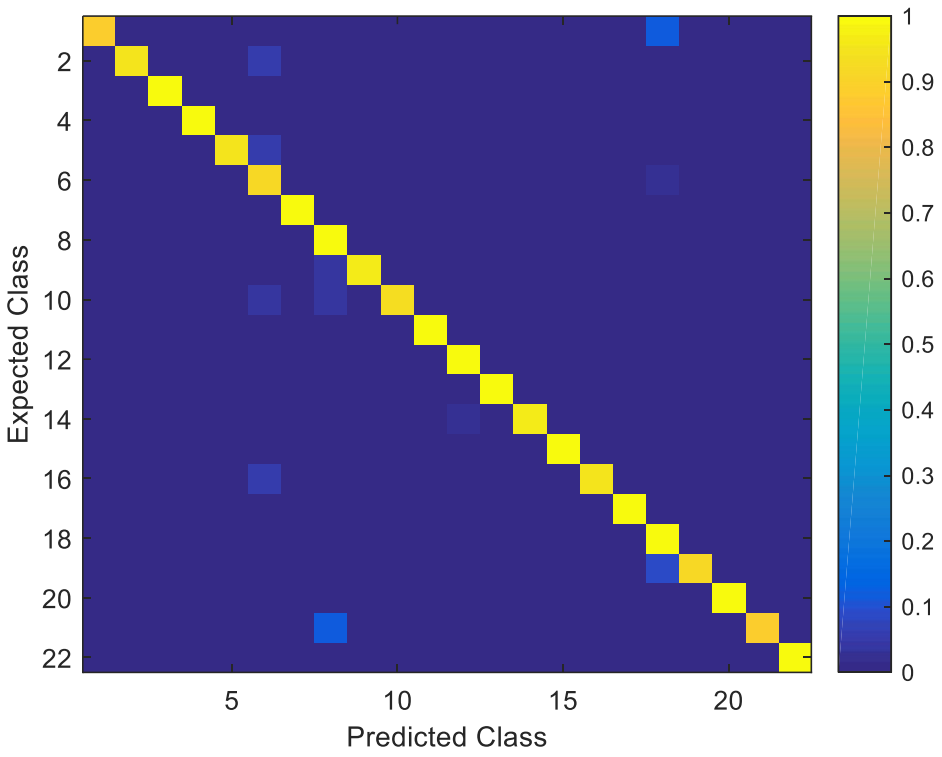
جدول ۶-۷: دقت بدست آمده (٪) برای هر یک از کلاس‌های مجموعه داده BVMMR توسط سیستم غیرآبشاری.

شماره کلاس	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
شماره کلاس	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	۲۰	۲۱	۲۲
تعداد نمونه‌ها	۱۰۰۴	۱۲۵	۸۰	۹۸	۳۸	۲۴۲	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

جدول ۶-۸: دقت مجموع و میانگین بدست آمده توسط سیستم غیرآبشاری بر روی مجموعه داده BVMMR.

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

نتایج گزارش شده برای نسخه غیرآبشاری نشان از عملکرد بسیار مناسب این سیستم دارند. علت اصلی اختلاف زیاد بین دقت مجموع و میانگین، کم بودن تعداد نمونه‌های برخی از کلاس‌ها (از جمله: ۱، ۲، ۵، ۱۹ و ۲۱) می‌باشد. دقت مجموع برای پنج رتبه اول تقریباً به ٪۱۰۰ رسیده است. نکته حائز اهمیت در جدول ۶-۷، طبقه‌بندی درست تعداد زیادی از نمونه‌های متعلق به کلاس «سایر» برای رتبه اول می‌باشد که این درصد شناسایی برای سه رتبه اول به ٪۱۰۰ رسیده است. شکل ۶-۸ ماتریس درهم‌ریختگی را تنها برای رتبه اول نمایش داده است. در این شکل، ثبات عملکرد سیستم در شناسایی نمونه‌های مربوط به همه کلاس‌ها را با توجه به قطر اصلی ماتریس مشاهده می‌کنیم.



شکل ۶-۸: ماتریس درهم‌ریختگی سیستم غیرآبشاری بر روی مجموعه داده BVMMR. رنگ روشن‌تر به معنی درصد بالاتر است.

دقت بدست آمده توسط نسخه «آبشاری از مدل‌های غیرآبشاری» برای دو مقدار متفاوت از $\boldsymbol{\alpha}$ ارائه شده است. این مقادیر همان دو مقداری هستند که در فصل ۵، سرعت سیستم به ازای آن‌ها گزارش شده است، یعنی $\boldsymbol{\alpha = 0}$ و $\boldsymbol{\alpha = 0.8}$. جدول ۶-۹ و جدول ۶-۱۰ نتایج بدست آمده برای $\boldsymbol{\alpha = 0.8}$ و جدول ۶-۱۱ و جدول ۶-۱۲ نتایج حاصل از $\alpha = 0$ را نمایش داده‌اند. برای محاسبه سه رتبه اول و پنج رتبه اول برای تمامی نسخه‌های آبشاری از الگوریتم پنجره لغزان به طول سه و پنج استفاده می‌شود.

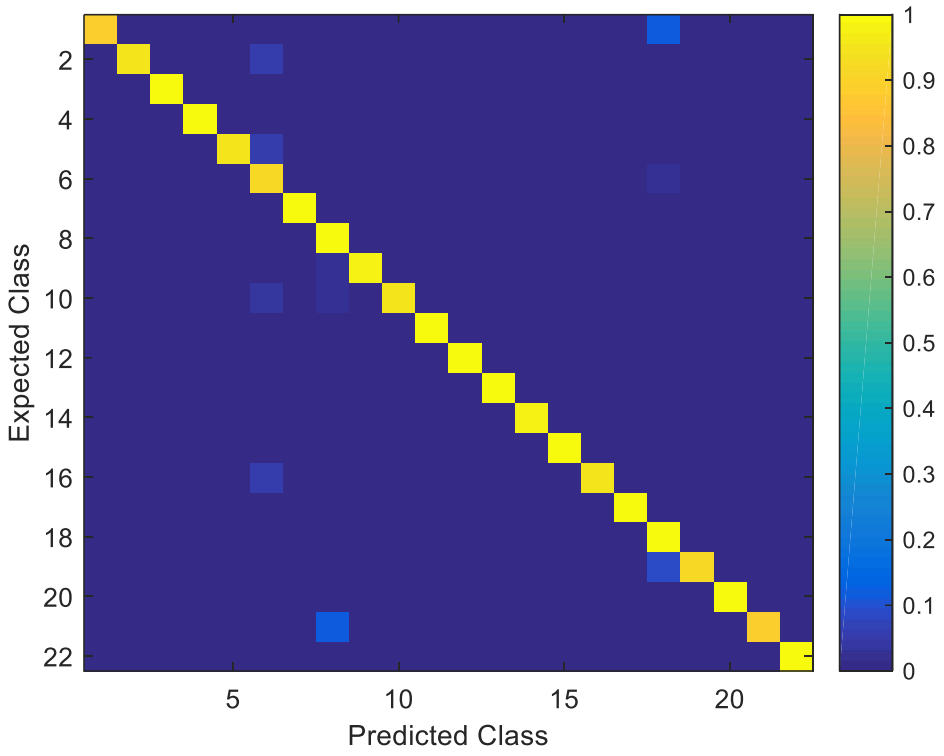
جدول ۶-۹: دقت بدست آمده (٪) برای هر یک از کلاس‌های مجموعه داده BVMMR توسط آبشاری از مدل های غیرآبشاری با $\alpha = 0.8$.

شماره کلاس	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	۹۸/۴۱	۹۴/۳۳	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
شماره کلاس	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	۹۷/۵۰	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

جدول ۶-۱۰: دقت مجموع و میانگین بدست آمده توسط آبشاری از مدل‌های غیرآبشاری بر روی مجموعه داده BVMMR با $\alpha = 0.8$.

	رتبه اول	سه رتبه اول	پنج رتبه اول
دقت مجموع	۹۸/۸۰٪	۹۹/۷۳٪	۹۹/۸۶٪
دقت میانگین	۹۷/۰۱٪	۹۸/۴۱٪	۹۹/۳۴٪

نسخه آبشاری بدون استفاده از مدل‌های آبشاری منجر به بهبود جزئی دقت نسبت به نسخه غیرآبشاری می‌شود. این نسخه همان نسخه شماره ۵ در جدول ۵-۲ است که بهترین دقت را با افزایش قابل توجه سرعت بدست می‌دهد. ماتریس درهم‌ریختگی بدست آمده توسط این نسخه در شکل ۶-۹ ارائه شده است.



شکل ۶-۹: ماتریس درهم‌ریختگی بدست آمده توسط آبشاری از مدل‌های غیرآبشاری بر روی مجموعه داده BVMMR با $\alpha = 0.8$.

جدول ۶-۱۱: دقت بدست آمده (٪) برای هر یک از کلاس‌های مجموعه داده BVMMR توسط آبشاری از مدل‌های غیرآبشاری با $\alpha = 0$.

شماره کلاس	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	۸۹/۴۷	۱۰۰	۱۰۰	۹۴/۷۴	۹۱/۵۴	۱۰۰	۱۰۰	۹۶/۸۳	۹۴/۹۶	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
شماره کلاس	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	۹۲/۱۰	۹۷/۳۶	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

جدول ۶-۱۲: دقت مجموع و میانگین بدست آمده توسط آبشاری از مدل‌های غیرآبشاری بر روی مجموعه داده BVMMR با $\alpha = 0$.

	رتبه اول	سه رتبه اول	پنج رتبه اول
دقت مجموع	۹۸/۶۶٪	۹۹/۷۳٪	۹۹/۸۶٪
دقت میانگین	۹۶/۴۲٪	۹۸/۴۱٪	۹۹/۳۴٪

نتایج ارائه شده در جدول ۶-۱۱ و جدول ۶-۱۲ متعلق به نسخه شماره ۶ در جدول ۵-۲ می‌باشند. این نسخه که سرعتی تقریباً ۴ برابری نسبت به نسخه غیرآبشاری دارد، دقتی مشابه با سیستم غیرآبشاری بدست داده است.

در پایان این بخش، دقت‌های کسب شده توسط نسخه «آبشاری از مدل‌های آبشاری» ارائه شده است. دو جدول ۶-۱۳ و جدول ۶-۱۴ عملکرد این نسخه را برای ترتیب حاصل از $\alpha = 0.8$ و جدول ۶-۱۵ و جدول ۶-۱۶ عملکرد این نسخه را برای ترتیب حاصل از $\alpha = 0$ گزارش کرده‌اند.

جدول ۶-۱۳: دقت بدست آمده (٪) برای هر یک از کلاس‌های مجموعه داده BVMMR توسط آبشاری از مدل های آبشاری با $\alpha = 0.8$.

شماره کلاس	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	۶۶/۶۷	۷۳/۶۸	۸۰	۸۵/۷۱	۸۴/۲۱	۹۰/۸۵	۱۰۰	۹۹/۴۸	۹۲/۰۶	۹۶/۸۶	۱۰۰
سه رتبه اول	۷۷/۷۸	۹۷/۵۰	۹۰	۹۲/۸۶	۸۴/۲۱	۹۹/۳۰	۱۰۰	۹۹/۴۸	۹۸/۴۱	۹۹/۳۷	۱۰۰
پنج رتبه اول	۸۸/۸۹	۸۹/۴۷	۹۰	۹۲/۸۶	۸۹/۴۷	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰
شماره کلاس	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	۹۹/۸۰	۹۶	۹۷/۵۰	۹۶/۹۳	۸۹/۴۷	۹۴/۷۴	۹۹/۵۹	۸۷/۵۰	۹۷/۷۸	۱۰۰	۹۸/۴۱
سه رتبه اول	۹۹/۹۰	۹۸/۴۰	۹۷/۵۰	۹۷/۹۶	۹۴/۷۴	۱۰۰	۱۰۰	۹۵/۸۳	۱۰۰	۱۰۰	۱۰۰
پنج رتبه اول	۱۰۰	۹۹/۲۰	۹۷/۵۰	۱۰۰	۹۷/۳۶	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰

جدول ۶-۱۴: دقت مجموع و میانگین بدست آمده توسط آبشاری از مدل‌های آبشاری بر روی مجموعه داده BVMMR با $\alpha = 0.8$.

	رتبه اول	سه رتبه اول	پنج رتبه اول
دقت مجموع	۹۷/۸۳٪	۹۹/۱۰٪	۹۹/۶۳٪
دقت میانگین	۹۲/۱۵٪	۹۵/۴۱٪	۹۷/۳۰٪

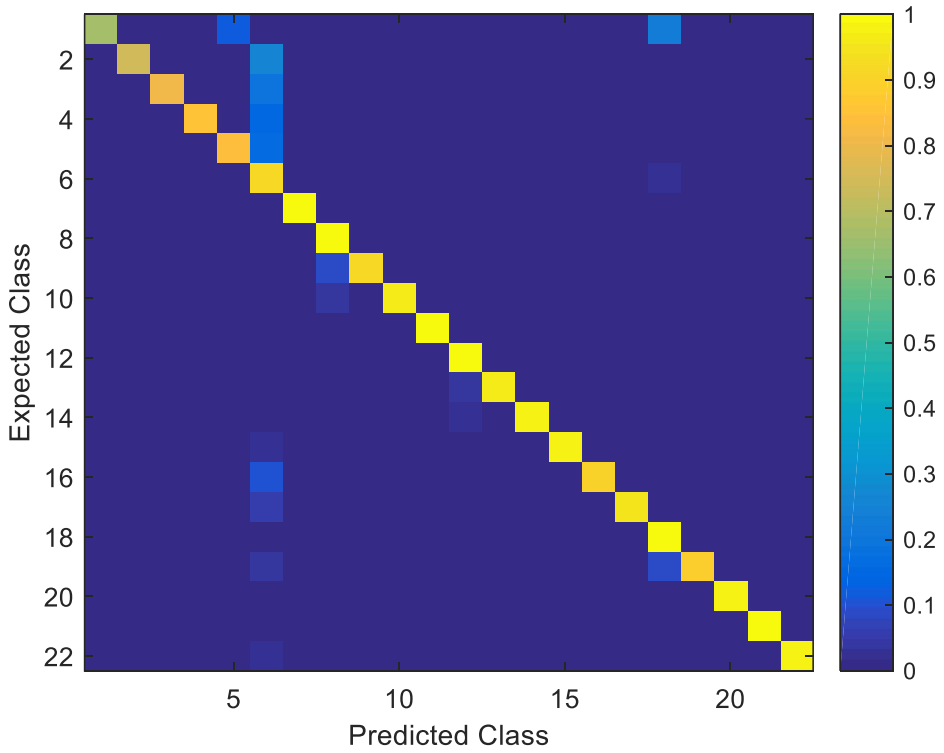
جدول ۶-۱۵: دقت بدست آمده (٪) برای هر یک از کلاس‌های مجموعه داده BVMMR توسط آبشاری از مدل های آبشاری با $\alpha = 0$.

شماره کلاس	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	۶۶/۶۷	۷۳/۶۸	۸۰	۸۵/۷۱	۸۴/۲۱	۹۰/۸۵	۱۰۰	۹۹/۴۸	۹۲/۰۶	۹۶/۸۶	۱۰۰
سه رتبه اول	۷۷/۷۸	۷۳/۶۸	۹۰	۹۲/۸۶	۸۴/۲۱	۹۹/۳۰	۱۰۰	۹۹/۴۸	۹۸/۴۱	۹۹/۳۷	۱۰۰
پنج رتبه اول	۸۸/۸۹	۸۹/۴۷	۹۰	۹۲/۸۶	۸۹/۴۷	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰
شماره کلاس	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	۹۹/۸۰	۹۶	۹۶/۲۵	۹۶/۹۳	۸۹/۴۷	۹۲/۱۰	۹۹/۵۹	۸۷/۵۰	۹۷/۷۸	۸۸/۸۸	۹۸/۴۱
سه رتبه اول	۹۹/۹۰	۹۸/۴۰	۹۷/۵۰	۹۷/۹۶	۹۴/۷۴	۱۰۰	۱۰۰	۹۵/۸۳	۱۰۰	۱۰۰	۱۰۰
پنج رتبه اول	۱۰۰	۹۹/۲۰	۹۷/۵۰	۱۰۰	۹۷/۳۶	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰

جدول ۶-۱۶: دقت مجموع و میانگین بدست آمده توسط آبشاری از مدل‌های آبشاری بر روی مجموعه داده BVMMR با $\alpha = 0$.

	رتبه اول	سه رتبه اول	پنج رتبه اول
دقت مجموع	۹۷/۷۰٪	۹۹/۱۰٪	۹۹/۶۳٪
دقت میانگین	۹۱/۴۶٪	۹۵/۴۱٪	۹۷/۳۰٪

دو نسخه‌ای که عملکرد آن‌ها در جدول ۶-۱۴ و جدول ۶-۱۶ گزارش شده در واقع نسخه‌های شماره نه و ده در جدول ۵-۲ هستند. این دو نسخه از سریعترین نسخه‌های سیستم می‌باشند که افزایش سرعت تقریباً ۱۵ برابری نسبت به سیستم غیرآبشاری به همراه دارند. علت اصلی کاهش دقت جزئی در این سیستم‌ها نسبت به سیستم پایه، محاسبه ضعیف حدآستانه‌های میانی در برخی مدل‌های آبشاری بوده است؛ این امر به دلیل وجود بعضی کلاس‌ها با تعداد نمونه‌های کم در مجموعه داده BVMMR رخ داده است. شکل ۶-۱۰ که ماتریس درهم‌ریختگی بدست آمده توسط نسخه شماره نه از سیستم را نمایش می-دهد نیز گویای این مطلب است. ستون ششم در این ماتریس که متعلق به کلاس «سایر» است، بیشترین طبقه‌بندی اشتباه را نشان می‌دهد. مدل‌های آبشاری ضعیف که دارای حد آستانه‌های میانی نامناسب هستند، تعدادی از نمونه‌های مثبت را در مراحل اولیه آبشار به اشتباه نپذیرفته و در کلاس «سایر» جای می‌دهند. به همین دلیل است که ستون ششم برای کلاس‌هایی که نمونه‌های کمتری دارند، پررنگ‌تر است.



شکل ۶-۱۰: ماتریس درهم‌ریختگی بدست آمده توسط آبشاری از مدل‌های آبشاری بر روی مجموعه داده BVMMR با $\alpha = 0.8$.

۶-۴-۲- ارزیابی روی مجموعه داده CompCars

ارزیابی بر روی مجموعه داده CompCars، تنها بر روی ۱۸۰ کلاس که دارای بیش از ۱۰۰ نمونه تصویر بوده‌اند، صورت گرفته است. آموزش طبقه‌بند مناسب برای کلاس‌هایی با کمتر از این تعداد نمونه امکان-پذیر نیست. فراوانی نمونه‌های هر کلاس به دلیل تعداد زیاد کلاس‌ها قابل ارائه نیست. هر کلاس از این مجموعه داده به صورت میانگین دارای ۱۵۰ نمونه است.

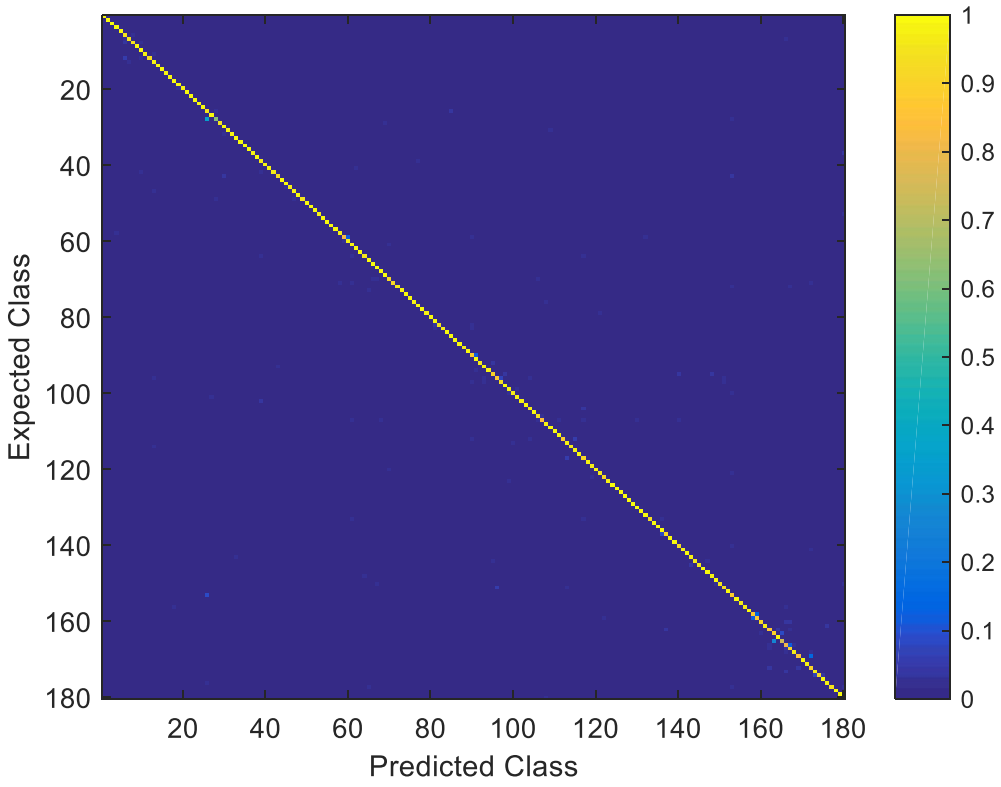
دقت شناسایی سیستم آموزش داده شده بر روی این مجموعه داده نیز مشابه با بخش قبل، برای سه نسخه متفاوت از سیستم گزارش شده است: نسخه غیرآبشاری، نسخه «آبشاری از مدل‌های غیرآبشاری»، نسخه «آبشاری از مدل‌های آبشاری». با این تفاوت که به دلیل تعداد زیاد کلاس‌ها و شباهت بالای ماتریس‌های درهم‌ریختگیِ این سه نسخه از سیستم، ماتریس درهم‌ریختگی تنها برای نسخه غیرآبشاری ارائه شده است.

جدول ۶-۱۷ دقت میانگین و مجموع مربوط به نسخه غیرآبشاری را برای رتبه اول، سه رتبه اول و پنج رتبه اول گزارش کرده است. شکل ۶-۱۱ نیز ماتریس درهم‌ریختگی این نسخه از سیستم را نمایش داده است.

جدول ۶-۱۷: دقت مجموع و میانگین بدست آمده توسط سیستم غیرآبشاری بر روی مجموعه داده CompCars.

	رتبه اول	سه رتبه اول	پنج رتبه اول
دقت مجموع	۹۶/۷۲٪	۹۸/۷۳٪	۹۹/۰۸٪
دقت میانگین	۹۶/۷۲٪	۹۸/۸۱٪	۹۹/۱۰٪

نتایج گزارش شده برای نسخه غیرآبشاری با توجه به دشواری مجموعه داده CompCars، نشان از عملکرد بسیار مناسب روش پیشنهادی دارند. به دلیل تعداد نمونه‌های مناسب هر کلاس در مجموعه داده CompCars، دقت مجموع و میانگین نزدیک به هم قرار دارند. با در نظر گرفتن تعداد کلاس‌های این مجموعه داده و ماتریس درهم‌ریختگی شکل ۶-۱۱، دقت سیستم کاملاً قابل قبول بنظر می‌رسد. در این مجموعه داده، تعدادی از مدل‌های خودرو کاملاً مشابه یکدیگر هستند و تفکیک آن‌ها برای ناظر انسانی نیز ممکن نیست. یک نمونه از این موارد در شکل ۶-۱۲ نمایش داده شده است. سایر موارد مشابه در پیوست اول قرار گرفته‌اند.



شکل ۶-۱۱: ماتریس درهم‌ریختگی سیستم غیرآبشاری بر روی مجموعه داده CompCars.

volvo_s60_1702

volvo_v60_1705

شکل ۶-۱۲: دو نمونه از کلاس‌های بسیار شبیه به هم در مجموعه داده CompCars که سیستم پیشنهادی در تفکیک آن‌ها خطای بیشتری داشته است.

دقت بدست آمده توسط نسخه «آبشاری از مدل‌های غیرآبشاری» برای دو مقدار متفاوت از α ارائه شده است. این مقادیر همان دو مقداری هستند که در فصل ۵، سرعت سیستم به ازای آن‌ها گزارش شده است، یعنی $\alpha = 0$ و $\alpha = 0.8$. جدول ۶-۱۸ نتایج بدست آمده برای $\alpha = 0.8$ و جدول ۶-۱۹ نتایج حاصل از $\alpha = 0$ را نمایش داده‌اند.

جدول ۶-۱۸: دقت مجموع و میانگین بدست آمده توسط آبشاری از مدل‌های غیرآبشاری بر روی مجموعه داده CompCars با $\alpha = 0.8$.

	رتبه اول	سه رتبه اول	پنج رتبه اول
دقت مجموع	۹۶/۷۱٪	۹۸/۴۰٪	۹۸/۷۰٪
دقت میانگین	۹۶/۷۴٪	۹۸/۴۵٪	۹۸/۷۲٪

جدول ۶-۱۹: دقت مجموع و میانگین بدست آمده توسط آبشاری از مدل‌های غیرآبشاری بر روی مجموعه داده CompCars با $\alpha = 0$.

	رتبه اول	سه رتبه اول	پنج رتبه اول
دقت مجموع	۹۶/۵۶٪	۹۸/۱۷٪	۹۸/۴۷٪
دقت میانگین	۹۶/۵۴٪	۹۸/۲۲٪	۹۸/۴۹٪

نسخه آبشاری بدون استفاده از مدل‌های آبشاری با مقدار $\alpha = 0.8$ منجر به بهبود جزئی دقت میانگین برای رتبه اول شده است. این نسخه همان نسخه شماره ۵ در جدول ۵-۳ است که بهترین دقت را با افزایش قابل توجه سرعت بدست می‌دهد. نتایج مربوط به سه و پنج رتبه اول نسبت به نسخه غیرآبشاری با افت جزئی همراه بوده است. نتایج نسخه شماره ۶ از جدول ۵-۳ در جدول ۶-۱۹ ارائه شده که دقت کمتر و سرعت بیشتری نسبت به نسخه قبل دارد.

در پایان این بخش، دقت‌های کسب شده توسط سریعترین نسخه‌های سیستم ارائه شده است. این دو نسخه همان نسخه‌های شماره نه و ده از جدول ۵-۳ می‌باشند. نتایج آزمایش‌های مربوط به این دو نسخه در دو جدول ۶-۲۰ و جدول ۶-۲۱ ارائه شده‌اند. نتایج نشان می‌دهند که به ازای کاهش یک درصدی دقت سیستم، سرعتی ۳ تا ۴ برابری نسبت به سیستم غیرآبشاری حاصل شده است.

جدول ۶-۲۰: دقت مجموع و میانگین بدست آمده توسط آبشاری از مدل‌های آبشاری بر روی مجموعه داده CompCars با $\alpha = 0.8$.

	رتبه اول	سه رتبه اول	پنج رتبه اول
دقت مجموع	۹۵/۹۳٪	۹۷/۶۱٪	۹۷/۸۸٪
دقت میانگین	۹۵/۶۳٪	۹۷/۲۸٪	۹۷/۵۰٪

جدول ۶-۲۱: دقت مجموع و میانگین بدست آمده توسط آبشاری از مدل‌های آبشاری بر روی مجموعه داده CompCars با $\alpha = 0$.

	رتبه اول	سه رتبه اول	پنج رتبه اول
دقت مجموع	٪۹۵/۷۷	٪۹۷/۳۶	٪۹۷/۶۲
دقت میانگین	٪۹۵/۵۲	٪۹۷/۰۸	٪۹۷/۲۸

۶-۵- مقایسه سیستم پیشنهادی و چند روش پرکاربرد

برای بررسی عملکرد سیستم پیشنهادی، آن را با چند روش دیگر مورد مقایسه قرار داده‌ایم. مقایسه مستقیم دقت‌های گزارش شده توسط کارهای گذشته به دلیل عدم استفاده از یک مجموعه داده مشترک امکان‌پذیر نمی‌باشد. از طرفی پیاده‌سازی این روش‌ها با توجه به پیچیدگی آن‌ها و توضیحات ناقص ارائه شده در مقالات مربوطه، به سادگی ممکن نیست. بنابراین یک راه‌حل منطقی، پیاده‌سازی رویکرد مشترک به کار رفته در اغلب کارهای گذشته است. در بیشتر کارهای گذشته از ماشین بردار پشتیبان به صورت یک‌بخشی و یا چندبخشی استفاده شده است. در حالت چندبخشی، تصویر به تعدادی بخش تقسیم شده و از هر بخش به صورت مجزا ویژگی استخراج می‌گردد. سپس برای هر یک از بخش‌ها، یک طبقه‌بند مستقل آموزش داده می‌شود. به همین دلیل، ما نیز از طبقه‌بند ماشین بردار پشتیبان با استفاده از کتابخانه LIBSVM بهره برده‌ایم [۸۵]. برای قابل مقایسه بودن روش‌های پیاده‌سازی شده و سیستم پیشنهادی، از یک روش استخراج ویژگی مشترک یعنی هیستوگرام گرادیان‌های جهت‌دار استفاده شده است. اندازه سلول و بلوک نیز به صورت یکسان تنظیم شده‌اند (اندازه سلول 8×8 و اندازه بلوک 2×2). در ادامه این بخش، روش مورد اشاره را «رویکرد مشترک» می‌نامیم.

برای استخراج ناحیه مطلوب، تصویر خودرو را در جهت عمودی به دو نیم تقسیم کرده و نیمه‌ی پایینی را به عنوان ناحیه مطلوب انتخاب کرده‌ایم. کلیه پردازش‌های آتی بر روی این ناحیه انجام می‌گیرد. دو نسخه از این سیستم به صورت تک‌بخشی و چندبخشی پیاده‌سازی شده است. در حالت چندبخشی، ناحیه مطلوب ابتدا به تعدادی سطر و ستون تقسیم شده و از هر سلول، ویژگی استخراج می‌گردد. ویژگی‌های استخراج شده از هر بخش به یک ماشین بردار پشتیبان مستقل آموزش داده می‌شوند. برای تعیین کلاس خروجی نیز از رأی‌گیری اکثریت استفاده می‌شود. شکل ۶-۱۳ مراحل طبقه‌بندی در حالت چندبخشی با دو سطر و دو ستون رو نمایش داده است.

شکل ۶-۱۳: مراحل طبقه‌بندی با استفاده از ماشین بردار پشتیبان به صورت چندبخشی در رویکرد مشترک.

رویکرد مشترک با استفاده از مجموعه داده‌های BVMMR و CompCars به صورت مجزا آموزش داده شده و با سیستم پیشنهادی مقایسه شده است. شکل آموزش سیستم برای هر مجموعه داده و نتایج بدست آمده از آن‌ها در ادامه تشریح گشته است.

۶-۵-۱- ارزیابی روی مجموعه داده BVMMR

تعداد نمونه‌های کلاس‌های مختلف در مجموعه داده BVMMR متوازن نیستند. برای مثال کلاس "پراید ۱۳۱" دارای ۲۰۰۱ نمونه و کلاس "لیفان ۶۲۰" دارای تنها ۲۰ نمونه است. اگر در چنین حالتی از یک ماشین بردار پشتیبان استاندارد استفاده کنیم، نتیجه مطلوبی بدست نخواهد آمد؛ زیرا احتمالا طبقه‌بند همه نمونه‌ها را تحت عنوان کلاس "پراید ۱۳۱" طبقه‌بندی و دقت مناسبی نیز گزارش خواهد کرد. برای رفع این مشکل از ماشین بردار پشتیبان وزن‌دار استفاده کرده‌ایم؛ با دادن وزن بیشتر به کلاس‌های دارای نمونه‌های کمتر، ارزش طبقه‌بندی درست آن‌ها را بالاتر خواهیم برد. اگر تعداد کل نمونه‌های آموزشی را برابر با N و تعداد نمونه‌های آموزشی کلاس i را برابر با N_i فرض کنیم؛ برای برابر کردن ارزش همه کلاس‌ها، از وزن $\frac{N}{N_i}$ برای کلاس i استفاده می‌کنیم. برای طبقه‌بندی وزن‌دار از الگوریتم C_SVM در کتابخانه LIBSVM استفاده شده است [۸۵].

در این آزمایش، کلاس «سایر» کنار گذاشته شده و دقت‌ها برای ۲۱ کلاس که دارای بیش از ۲۰ نمونه بوده‌اند، محاسبه شده است. جدول ۶-۲۲ نتایج این آزمایش را نشان می‌دهد. منظور از SINGLE_SVM_1_1، طبقه‌بندی تک‌بخشی و منظور از MULTI_SVM_2_3، طبقه‌بندی چندبخشی با دو سطر و سه ستون می‌باشد. شکل ۶-۱۳عملکرد MULTI_SVM_2_2 را به عنوان نمونه نمایش داده است.

جدول ۶-۲۲: مقایسه سیستم پیشنهادی با چند نسخه از رویکرد مشترک بر روی مجموعه داده BVMMR.

روش	دقت میانگین	دقت مجموع
SINGLE_SVM_1_1	۹۱/۲۸٪	۹۳/۷۷٪
MULTI_SVM_2_2	۸۹٪	۹۲/۴۸٪
MULTI_SVM_2_3	۸۸/۰۸٪	۸۹/۲۶٪
MULTI_SVM_2_4	۸۸/۲۰٪	۸۹/۷۱٪
آبشاری از مدل‌های آبشاری	۹۲/۲۱٪	۹۸/۱۸٪
سیستم غیرآبشاری	۹۷/۰۷٪	۹۹/۰۵٪
آبشاری از مدل‌های غیرآبشاری	۹۷/۲۷٪	۹۹/۱۹٪

نتایج جدول ۶-۲۲ دربرگیرنده دو نکته است. نکته اول، برتری روش‌های پیشنهادی با اختلاف زیاد نسبت به سایر روش‌های ارائه شده در جدول است. دو سیستم «آبشاری از مدل‌های آبشاری» و «آبشاری از مدل‌های غیرآبشاری» در این جدول همان نسخه‌های پنج و ده در جدول ۵-۲ می‌باشند. نکته دوم، عملکرد بهتر ماشین بردار پشتیبان در حالت تک‌بخشی نسبت به حالت 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.

۶-۵-۲- ارزیابی روی مجموعه داده CompCars

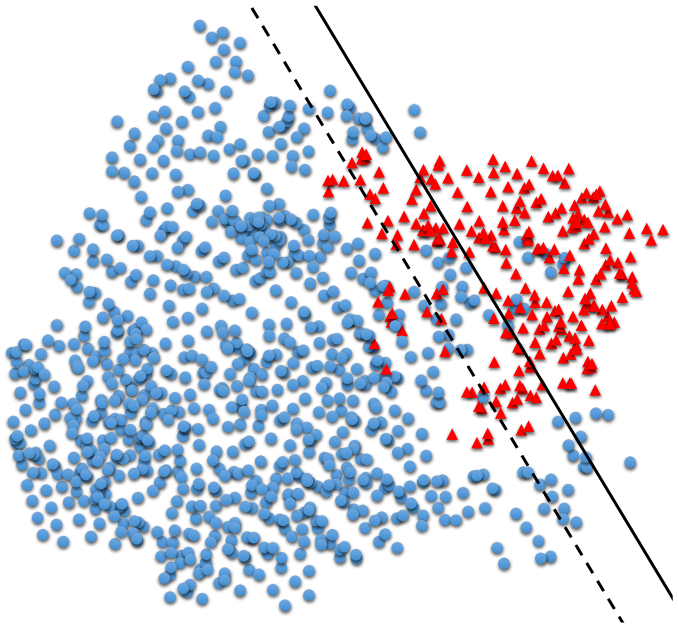
همانطور که در بخش ۴-۴-۳- اشاره شد، برای آموزش سیستم پیشنهادی بر روی مجموعه داده CompCars از داده‌کاوی چندکلاسه بر روی نمونه‌های منفی با انتخاب توزیع شده بهره گرفته شده است. در این حالت، برای آموزش طبقه‌بند i، نمونه‌های مربوط به کلاس i به عنوان نمونه‌های مثبت و تعداد مشخصی از نمونه‌های مربوط به سایر کلاس‌ها به عنوان نمونه‌های منفی درنظر گرفته می‌شوند. برای آموزش سیستم پیشنهادی از ۳۰ نمونه از هر یک از ۱۷۹ کلاس منفی به عنوان مجموعه نمونه‌های منفی استفاده شده است ($179 \times 30 = 5370$). این تعداد، تعادل مناسبی بین سرعت آموزش، پیچیدگی فضایی فرآیند آموزش و دقت سیستمِ حاصل ایجاد می‌کند. برای انجام یک مقایسه عادلانه، از همین روش داده‌کاوی برای آموزش الگوریتم مبتنی بر رویکرد مشترک بهره برده شده است. در ضمن، علاوه‌بر استفاده از انتخاب توزیع شده به طول ۳۰، طول‌های ۲۰ و ۲۵ نیز مورد آزمایش قرار گرفته‌اند.

به تازگی نتایجی بر روی مجموعه داده CompCars منتشر شده است [۷۲]. مقاله مورد اشاره از شبکه عصبی عمیق بهره برده است. علاوه‌براین روش‌های ارائه شده در [۳۰] و [۴۸] را نیز پیاده‌سازی کرده است. هرچند در رابطه با نحوه و صحت پیاده‌سازی نظری نمی‌توان داد. نتایج گزارش شده توسط این مقاله، روش پیشنهادی و رویکرد مشترک در جدول ۶-۲۳ مورد مقایسه قرار گرفته‌اند.

جدول ۶-۲۳: مقایسه سیستم پیشنهادی و چند روش دیگر بر روی مجموعه داده CompCars. موارد ستاره‌دار توسط [۷۲] پیاده‌سازی شده‌اند.

شماره	روش	دقت میانگین	دقت مجموع
۱	SINGLE_SVM_1_1 (20 SPC)	۷۹/۵۴٪	۸۰/۸۴٪
۲	SINGLE_SVM_1_1 (25 SPC)	۸۲/۳۹٪	۸۳/۳۴٪
۳	SINGLE_SVM_1_1 (30 SPC)	۸۴/۲۴٪	۸۵/۷۲٪
۴	MULTI_SVM_2_2 (30 SPC)	۸۳/۸۵٪	۸۱/۹۵٪
۵	MULTI_SVM_2_3 (30 SPC)	۸۳/۸۱٪	۸۲/۰۵٪
۶	MULTI_SVM_2_4 (30 SPC)	۸۳/۸۹٪	۸۲/۲۲٪
۷	* شیبِه [۴۸]	۵۱/۷۰٪	-
۸	* ژنگ [۳۰]	۸۳/۷۸٪	-
۹	شبکه عصبی عمیق [۷۲]	۹۰/۷۸٪	-
۱۰	شبکه عصبی عمیق + بخش‌ها ز	۹۸/۲۹٪	۹۸/۶۳٪
۱۱	آبشاری از مدل‌های آبشاری	۹۵/۶۳٪	۹۵/۹۳٪
۱۲	سیستم غیرآبشاری	۹۶/۷۲٪	۹۶/۷۲٪
۱۳	آبشاری از مدل‌های غیرآبشاری	۹۶/۷۴٪	۹۶/۷۱٪

منظور از SPC در جدول ۶-۲۳، طول مورد استفاده در انتخاب توزیع شده می‌باشد. دو سیستم «آبشاری از مدل‌های آبشاری» و «آبشاری از مدل‌های غیرآبشاری» در این جدول همان نسخه‌های پنج و ده در جدول ۵-۳ می‌باشند. مشاهده می‌شود که عملکرد تمامی نسخه‌های سیستم پیشنهادی نسبت به رویکرد مشترک، بسیار بهتر بوده است. نسخه چندبخشی از رویکرد مشترک در این آزمایش نیز نتوانسته دقت بهتری نسبت به نسخه تک‌بخشی کسب کند. علاوه‌براین، با افزایش طول انتخاب توزیع شده، دقت نیز بهبود یافته است. هر چند با افزایش بیشتر این طول به سه مشکل احتمالی برخورد می‌کنیم. مشکل اول، نیاز به حافظه RAM بیش از ۱۶ گیگا بایت برای آموزش یک طبقه‌بند است. مشکل دوم، احتمال بروز بیش‌برازش و مشکل احتمالی سوم، کاهش قدرت تعمیم طبقه‌بند حاصل است؛ زیرا تعداد نمونه‌های منفی بسیار بیشتر نسبت به تعداد نمونه‌های مثبت، طبقه‌بند را وادار به رسم صفحه جداکننده در مکانی مناسب می‌نماید. یک نمونه از چنین حالتی در شکل ۶-۱۴ به تصویر کشیده شده است. در این شکل، خط ممتد، یک صفحه (خط) جداکننده نامناسب و خط‌چین، یک صفحه جداکننده مناسب را نمایش داده است.



شکل ۶-۱۴: نمایش یک نمونه خط جداکننده مناسب (خط‌چین) و نامناسب (خط ممتد) در حالتی که تعداد نمونه‌های مثبت و منفی شدیدا نامتوازن هستند.

بر اساس نتایج ارائه شده در [۷۲]، روش‌های [۳۰] و [۴۸] که هر یک در مقالات خود دقتی در حدود ٪۹۹ را گزارش کرده‌اند، عملکرد مناسبی بر روی مجموعه داده چالش‌برانگیز CompCars نداشته‌اند. شبکه عصبی عمیق به دلیل اندازه بسیار بزرگ و حجم محاسباتی بسیار بالا توانسته در اغلب مسائل بینایی ماشین و پردازش تصویر عملکرد خیره‌کننده‌ای از خود به نمایش بگذارد. نیاز به تعداد تصاویر بسیار زیاد، زمان آموزش طولانی و سرعت کم در زمان اجرا از ضعف‌های این ابزار قدرتمند است. علاوه‌براین، آزمایش جزئی عملکرد سیستم و اشکال‌یابی آن به راحتی امکان‌پذیر نیست. با همه این تفاسیر، مشاهده می‌شود که روش پیشنهادی دقت بالاتری نسبت به شبکه عصبی عمیق بدست آورده است (روش شماره نه در جدول ۶-۲۳). تنها روشی که موفق به کسب دقت بالاتری نسبت به روش پیشنهادی شده است، از چهار شبکه عصبی عمیق و ۲۸۰ ماشین بردار پشتیبان بهره می‌برد. مدیریت چنین سیستمی با این حجم پردازشی و حافظه موردنیاز، نیاز به کامپیوتری با حافظه RAM بالا، پردازنده و پردازنده گرافیکی قدرتمند دارد.

۶-۶- جمع‌بندی

در این فصل به ارزیابی کامل سیستم پیشنهادی پرداخته شد. برای این منظور بیش از ۲۰ آزمایش بر روی سیستم پیشنهادی صورت گرفت. نتایج بدست آمده، گزارش شده و مورد تحلیل قرار گرفتند.

آزمایش‌ها بر روی دو مجموعه داده BVMMR و CompCars انجام گرفتند. عملکرد سیستم روی مجموعه داده BVMMR که متشکل از خودروهای داخلی بود، نشان از عملکرد بسیار مناسب سیستم پیشنهادی داشت؛ به صورتی که دقت مجموعِ نزدیک به ٪۹۹ برای این مجموعه داده حاصل شد. این مجموعه داده دارای چالش‌هایی از قبیل تغییرات روشنایی، وضوح تصویر متفاوت و تغییرات ظاهری (تغییر چراغ‌ها، آسیب‌دیدگی) بوده است. تعداد کلاس‌های مجموعه داده CompCars نزدیک به نه برابر مجموعه داده BVMMR می‌باشد. تعداد بسیار زیاد کلاس‌ها و شباهت برون‌کلاسی بسیار کم بین آن‌ها مهمترین چالش‌هایی بودند که سیستم پیشنهادی در این مجموعه داده با آن‌ها مواجه بود. با این وجود، دقت کسب شده توسط سیستم یعنی تقریباً ٪۹۷ نشان از کارایی بالای رویکرد پیشنهادی دارد.

در شکل ۶-۱۵ و شکل ۶-۱۶ چند نمونه از تشخیص‌های درست و اشتباه انجام شده توسط سیستم پیشنهادی بر روی این دو مجموعه داده به تصویر کشیده شده است. در این شکل‌ها، پنج رتبه اول پیشنهاد شده برای هر تصویر و امتیاز آن نمایش داده شده است. البته برای برخی از تصاویر کمتر از پنج تشخیص، پیشنهاد شده است. کلاس درست با رنگ سبز و کلاس‌های اشتباه با رنگ قرمز نشان داده شده‌اند (رنگ سبز و قرمز در حالت خاکستری[۱] به ترتیب به رنگ‌های خاکستری روشن و تیره تبدیل می‌شوند). در مجموعه داده BVMMR ابتدا مکان و محدوده خودرو تشخیص داده شده و سپس نوع و مدل، شناسایی می‌گردد. برای تعداد خودروهای موجود در تصویر، محدودیتی وجود ندارد.

[۱] Grayscale

شکل ۶-۱۵: چند نمونه از خروجی‌های سیستم بر روی مجموعه داده BVMMR. برای هر تصویر، پنج تشخیص اول سیستم و امتیاز هر یک نمایش داده شده است.

شکل ۶-۱۶: چند نمونه از خروجی‌های سیستم بر روی مجموعه داده CompCars. برای هر تصویر، پنج تشخیص اول سیستم و امتیاز هر یک نمایش داده شده است.

فصل

۷- نتیجه‌گیری و کارهای آینده

شناسایی نوع و مدل وسیله نقلیه یکی از چالش‌هایی است که در دهه اخیر توجه محققان را به خود جلب کرده است. این مسئله کاربردهای ضروری و متنوعی دارد که نیاز به پیاده‌سازی یک سیستم کاربردی را افزایش می‌دهند. در این رساله، روشی برای شناسایی نوع و مدل وسایل نقلیه ارائه شده است. سیستم پیاده‌سازی شده، همه بخش‌های موردنیاز برای استفاده در دنیای واقعی را شامل می‌شود؛ یعنی فرآیند آموزش و آزمایش کاملاً خودکار. تصاویر آموزشی مورد استفاده توسط سیستم پیشنهادی نیازی به علامت‌گذاری جزئی بخش‌ها نداشته و مشخص بودن محدوده خودرو در تصاویر کفایت می‌کند. در مرحله آزمایش نیز، سیستم به صورت خودکار مکان و محدوده خودروها در تصویر را تشخیص داده و نوع و مدل هر یک را گزارش می‌کند. حتی محدودیتی روی تعداد خودروهای موجود در تصویر نیز وجود ندارد. در دنیای واقعی، تعداد مدل‌های وسایل نقلیه دائما رو به افزایش است. سیستم پیشنهادی بعد از نصب و راه-اندازی می‌تواند به مرور خود را با تغییرات جدید وفق دهد. زیرا سیستم مورد بحث از تعدادی طبقه‌بند مستقل تشکیل شده است؛ بنابراین می‌توان در هر زمان دلخواه یک طبقه‌بند را از سیستم حذف و یا به آن اضافه کرد. اجزای سیستم پیشنهادی در بخش بعد به صورت مختصر مورد اشاره قرار گرفته‌اند.

۷-۱- جمع‌بندی روش‌های پیشنهادی در رساله

در این رساله یک سیستم جامع برای شناسایی نوع و مدل خودرو ارائه شد. سیستم پیاده‌سازی شده برای هر مدل از خودروها یک طبقه‌بند یاد می‌گیرد. هر طبقه‌بند شامل مجموعه‌ای از بخش‌های متمایزکننده مختص به آن مدل از خودروها است. یک طبقه‌بند با استفاده از ویژگی‌های استخراج شده از این بخش‌ها، رابطه هندسی بین آن‌ها و ضرایب یاد گرفته شده برای هر بخش می‌تواند یک مدل خودرو را از سایر مدل‌ها تمیز دهد. آموزش چنین طبقه‌بندی نیاز به یک فرآیند یادگیری با مقادیر مخفی دارد؛ زیرا طبقه-بند علاوه‌بر یادگیری شکل جدا کردن دو کلاس از یکدیگر، باید مکان مناسب برای بخش‌ها را نیز بیاموزد. این شکل از آموزش با استفاده از الگوریتم ماشین بردار پشتیبان مخفی پیاده‌سازی شده است. برای استخراج بخش‌های متمایزکننده‌ی مربوط به هر مدل از خودروها، از یک الگوریتم حریصانه بهره برده شده است. این الگوریتم با مقایسه کردن شباهت بین یک بخش از مدل جاری و بخش‌های مشابه از سایر مدل‌ها، کم‌شباهت‌ترین بخش‌ها را برمی‌گزیند. مجموعه بخش‌های یاد گرفته شده برای یک مدل از خودروها یا از نظر مکان بخش‌ها با سایر مدل‌ها تفاوت دارند و یا از نظر ضرایب فیلترهای مربوط به بخش‌ها؛ خطای جابجایی بخش‌ها نسبت به ریشه نیز می‌تواند برای هر مدل از خودروها یکتا باشد. این ترکیب سه‌تایی از ویژگی‌های مختص به یک مدل منجر به ایجاد یک طبقه‌بند قدرتمند می‌گردد.

عدم پیاده‌سازی بهینه فرآیند آموزش به سادگی منجر به افزایش نمایی مدت زمان آموزش و یا کارایی پایین طبقه‌بند حاصل می‌گردد. به همین دلیل از یک الگوریتم داده‌کاوی چندکلاسه برای استخراج نمونه‌های منفی دشوار استفاده شده است. پس از داده‌کاوی، تعداد کمی از نمونه‌های منفی در فرآیند آموزش شرکت خواهند کرد که این امر موجب تسریع روند آموزش، کاهش احتمال بیش‌برازش و یا کاهش قدرت تعمیم طبقه‌بند خواهد شد. تأثیر مثبت الگوریتم داده‌کاوی پیشنهادی و انواع متفاوت اعمال آن با چند آزمایش نشان داده شد.

برای بهبود سرعت سیستم در زمان اجرا، یک الگوریتم آبشاری پیشنهاد شده است. الگوریتم مذکور با اعمال ترتیبی طبقه‌بندها به تصویر، سرعت پردازش سیستم را بهبود می‌بخشد. برای دست یافتن به یک ترتیب کارا از طبقه‌بندها که منجر به افزایش سرعت و عدم کاهش دقت سیستم شود، یک معیار ترکیبی پیشنهاد شد. این معیار با توجه به اطمینان و فراوانی هر طبقه‌بند در مورد ترتیب قرارگیری آن‌ها تصمیم-گیری می‌کند. با قرار دادن طبقه‌بندهای مطمئن‌تر در ابتدای آبشار، می‌توان دقت را بهبود بخشید. در طرف مقابل، با قرار دادن طبقه‌بندی که دارای فراوانی بیشتر است، سرعت افزایش خواهد یافت. نسخه‌ی نهایی الگوریتم آبشاری که «آبشاری از مدل‌های آبشاری» نامیده می‌شود، می‌تواند سرعت سیستم را تا ٪۸۰ افزایش داده و تنها منجر به افت جزئی در دقت گردد. این نسخه از الگوریتم هر پنجره از طبقه‌بندها را به صورت موازی به تصویر اعمال می‌کند. برای جلوگیری از افت دقت، می‌توان بجای استفاده از مدل-های آبشاری از مدل‌های غیرآبشاری استفاده کرد. در این حالت، افزایش سرعت کمتری خواهیم داشت ولی دقت کاهش نیافته و حتی می‌تواند بهبود یابد.

مجموعه داده مورد استفاده برای آموزش سیستم نیاز به هیچ ویژگی خاصی ندارد و علامت‌گذاری خودروهای موجود در تصاویر تنها شرط لازم است. در همین راستا و به منظور توسعه سیستم پیشنهادی، دو مجموعه داده از خودروهای داخلی به نام BVMMR، متشکل از ۷۰۰ خودرو از تقریباً ۲۰ مدل و ۶۰۰۰ خودرو از تقریباً ۳۰ مدل متفاوت جمع‌آوری گشته است. علاوه‌بر حاشیه‌نویسی همه تصاویر این دو مجموعه داده، یک بسته توسعه نرم‌افزاری نیز برای استفاده از آن‌ها تهیه شده است. این دو مجموعه داده، به همراه بسته توسعه نرم‌افزاری برای کاربردهای علمی به اشتراک گذاشته و به صورت رایگان قابل دریافت هستند.

به دلیل آموزش مستقل هر طبقه، حد آستانه و محدوده خروجی آن‌ها بر هم منطبق نیستند. برای یکسان‌سازی حد آستانه و محدوده خروجی طبقه‌بندها، از الگوریتم مقیاس‌گذاری پلت استفاده می‌شود. این الگوریتم با استفاده از مجموعه داده آموزشی، یک تابع نمایی را به خروجی هر طبقه‌بند برازش می-

نماید. به این ترتیب، خروجی همه طبقه‌بندها در بازه [۱-۰] قرار می‌گیرد. بدون انجام این عملیات، دقت سیستم مورد بحث و سایر سیستم‌های مشابه، افت محسوسی خواهند داشت.

برای ارزیابی عملکرد سیستم، آزمایش‌های زیادی بر روی هر دو نسخه از مجموعه داده BVMMR و همچنین مجموعه داده بزرگ CompCars انجام گرفته است. نتایج آزمایش‌های مختلف صورت گرفته، نشان از کارایی سیستم پیشنهادی دارند. این آزمایش‌ها شامل سنجش دقت و سرعت سیستم می‌باشند. تأثیر الگوریتم آبشاری با تنظیمات مختلف نیز مورد بررسی قرار گرفته است. طبقه‌بندهای آموزش داده شده توسط سیستم و تعدادی از خروجی‌های سیستم بر روی مجموعه داده BVMMR در قالب یک پیوست و در فصل ۰ ارائه شده‌اند.

۷-۲- کارهای آینده

یکی از فرض‌های در نظر گرفته شده در پیاده‌سازی سیستم پیشنهادی، تغییرات اندک زاویه بوده است. هر دو مجموعه داده BVMMR v2 و CompCars که در آزمایش‌ها از آن‌ها استفاده شده است نیز این فرض را رعایت کرده‌اند. می‌توان سیستم حاضر را به شکلی تغییر داد که برای زوایای متفاوت عملکرد مناسبی داشته باشد. برای این منظور می‌توان برای هر طبقه‌بند (مدل) چند زیرمدل در نظر گرفت که هر یک برای یک محدوده زاویه‌ی مشخص آموزش دیده شده باشند. هر طبقه‌بند در زمان اجرا، همه زیرمدل‌هایش را به تصویر اعمال کرده و زیرمدلی که دارای امتیاز بیشتری است را به عنوان رأی نهایی خود برمی‌گزینند. یک رویکرد دیگر می‌تواند به این صورت باشد که ابتدا زاویه به هر روش دلخواه تخمین زده شود و سپس تصویر به زیرمدلی که روی آن زاویه آموزش دیده شده، ارسال گردد. این دو روش می‌توانند ایده‌های مناسبی برای کارهای آینده باشند.

در سیستم پیشنهادی از الگوریتم هیستوگرام گردیان‌های جهت‌دار برای استخراج ویژگی استفاده شده است. علت انتخاب این الگوریتم در بخش‌های مربوطه تشریح گشت. هرچند کارایی مناسب این روش نمی‌تواند دلیلی بر عدم وجود روش‌های بهتر باشد. برای مثال می‌توان از ویژگی‌های استخراج شده از یک شبکه عصبی عمیق بهره برد و آن را جایگزین هیستوگرام گرادیان‌های جهت‌دار کرد. در این رویکرد، ضرایب یک لایه مشخص از یک شبکه عصبی عمیق از پیش آموزش داده شده، استخراج شده و به عنوان بردار ویژگی مورد استفاده قرار می‌گیرند. مرجع [۶۸] یک رویکرد مشابه را پیشنهاد کرده و دقت سیستم را به مقدار قابل توجهی بهبود بخشیده است. در صورتی که شبکه عصبی با استفاده از مجموعه داده هدف آموزش داده شود، ویژگی‌های بسیار بهتری را بدست خواهد داد [۷۲]. بنابراین می‌توان انتظار داشت که

رویکرد مورد اشاره، کارایی سیستم جاری را نیز افزایش دهد. پیاده‌سازی این ایده نیاز به میزان قابل توجهی زمان دارد و می‌تواند شروع مناسبی برای کارهای آینده باشد.

ساختار مبتنی بر بخش مورد استفاده در سیستم پیشنهادی، پتانسیل مناسبی برای افزایش سرعت هر چه بیشتر دارد. اعمال موازی فیلتر ریشه و بخش‌ها و انجام موازی عملیات ضرب ماتریس‌ها را که در حال حاضر توسط CPU به صورت موازی صورت می‌گیرد، می‌توان توسط پردازنده‌های گرافیکی (GPU) انجام داد. به این ترتیب می‌توان افزایش سرعت چندین برابری نسبت به حالتی که از CPU برای موازی‌سازی استفاده می‌شود، بدست آورد. تبدیل کدهای جاری برای اجرای موازی روی GPU نیاز به مهارت و دقت بالایی دارد و یک ایده مناسب برای کارهای آینده است.

ایده‌ی اصلی استفاده شده در این رساله استفاده از رویکرد مبتنی بر بخش بوده است. این ایده می‌تواند به شکل‌های مختلفی پیاده‌سازی شود. در این رساله از مدل‌های مبتنی بر بخش برای این منظور بهره گرفته شد. می‌توان از شبکه‌های عصبی عمیق نیز در این جهت استفاده کرد. چنان‌که [۷۲] به شکلی ساده سعی بر پیاده‌سازی این ایده داشته است؛ و با وجود سادگی، توانسته بهبود مناسبی بدست آورد. این مرجع برای کلیه وسایل نقلیه، تعدادی بخش مشترک در نظر گرفته است. در صورتی که مشابه با رساله جاری برای هر مدل از وسایل نقلیه، یک مجموعه بخش متمایز کننده استخراج گردد، می‌توان بهبود بیشتری را انتظار داشت. این ایده که کاملا هم‌راستا با تحقیقات در حال انجام در این حوزه است، شروعی بسیار مناسب برای کارهای آینده می‌باشد.

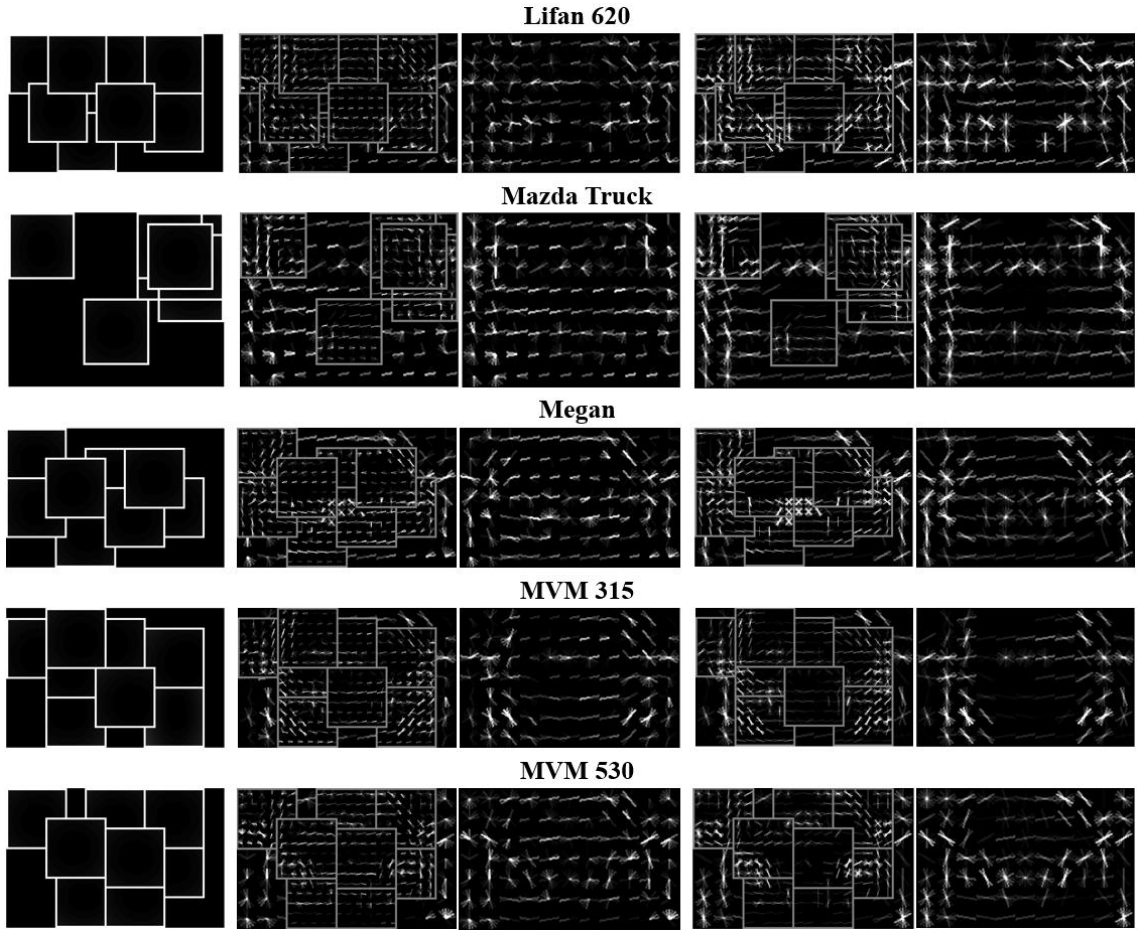
پیوست اول

خروجی‌های سیستم پیاده‌سازی شده

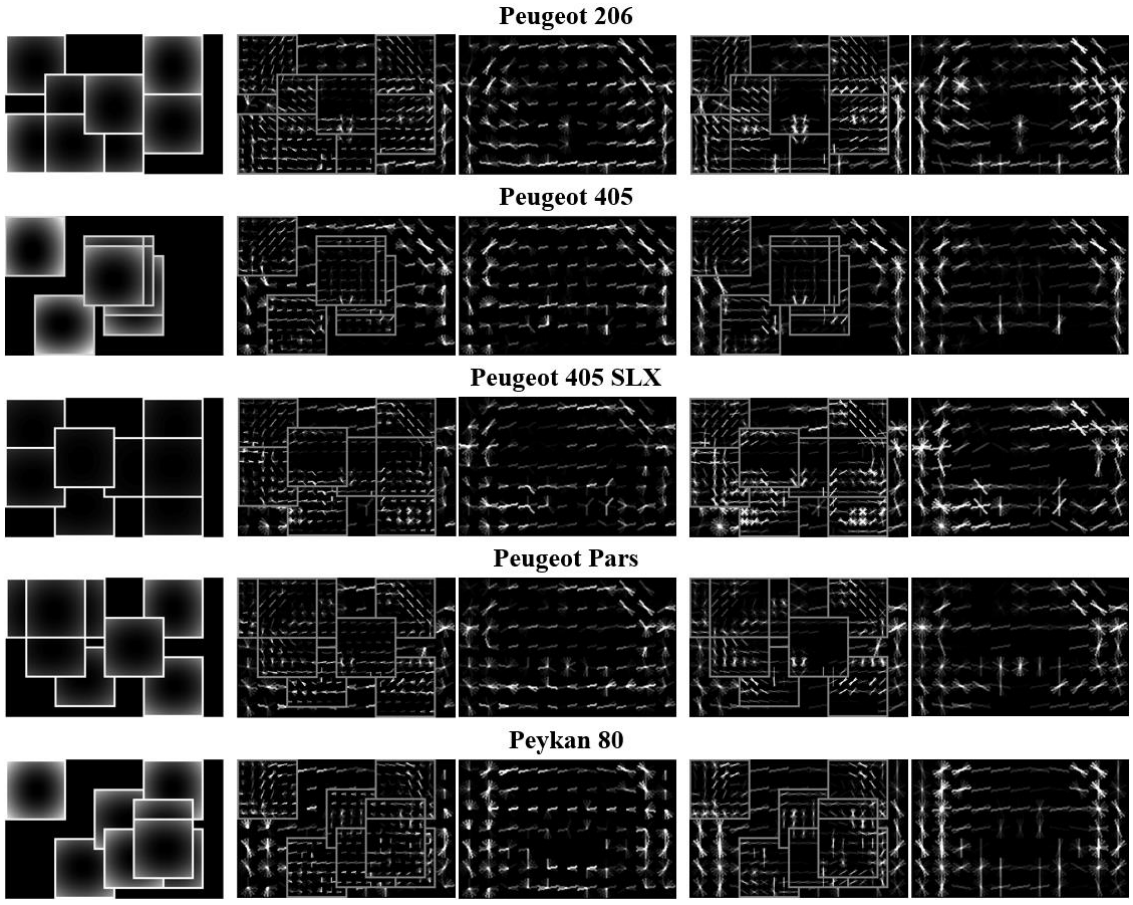
در این فصل به ارائه دو مجموعه تصاویر می‌پردازیم. مجموعه اول شامل طبقه‌بندهای آموزش داده شده برای هر یک از مدل‌های خودروی موجود در مجموعه داده BVMMR v2 است. مجموعه دوم نیز از شناسایی‌های انجام شده توسط سیستم پیشنهادی بر روی همین مجموعه داده تشکیل می‌شود.

مشاهده ساختار طبقه‌بندها به درک بهتر سیستم پیشنهادی کمک می‌کند. هر طبقه‌بند در سه بخش ارائه می‌شود: ضرایب فیلتر ریشه، ضرایب فیلتر بخش‌ها و رابطه هندسی بین بخش‌ها. این شکل ارائه در سایر بخش‌های رساله نیز مورد استفاده قرار گرفته است. قدرت تمایز هر طبقه‌بند در فیلتر ریشه، فیلتر بخش‌ها، مکان بخش‌ها و رابطه هندسی بین بخش‌ها نهفته است. از آنجا که بردارهای ویژگی با استفاده از الگوریتم هیستوگرام گرادیان‌های جهت‌دار استخراج می‌شوند، ضرایب فیلترها یک نمای کلی از خودروها را نمایش می‌دهند که در تصاویر مشخص است. علاوه‌براین، می‌دانیم که از هر دو نسخه‌ی علامت‌دار و بدون علامت الگوریتم هیستوگرام گردیان‌های جهت‌دار در بردار ویژگی‌های استفاده شده و نمایش هر دو نسخه به صورت توأمان ممکن نیست؛ به همین دلیل، هر طبقه‌بند یک مرتبه با ضرایب مربوط به نسخه علامت‌دار و یک مرتبه با ضرایب بدون علامت نمایش داده شده است. در سایر بخش‌های رساله، برای سادگی بیشتر و اشغال فضای کمتر، تنها نسخه‌ی بدون علامت نمایش داده می‌شد.

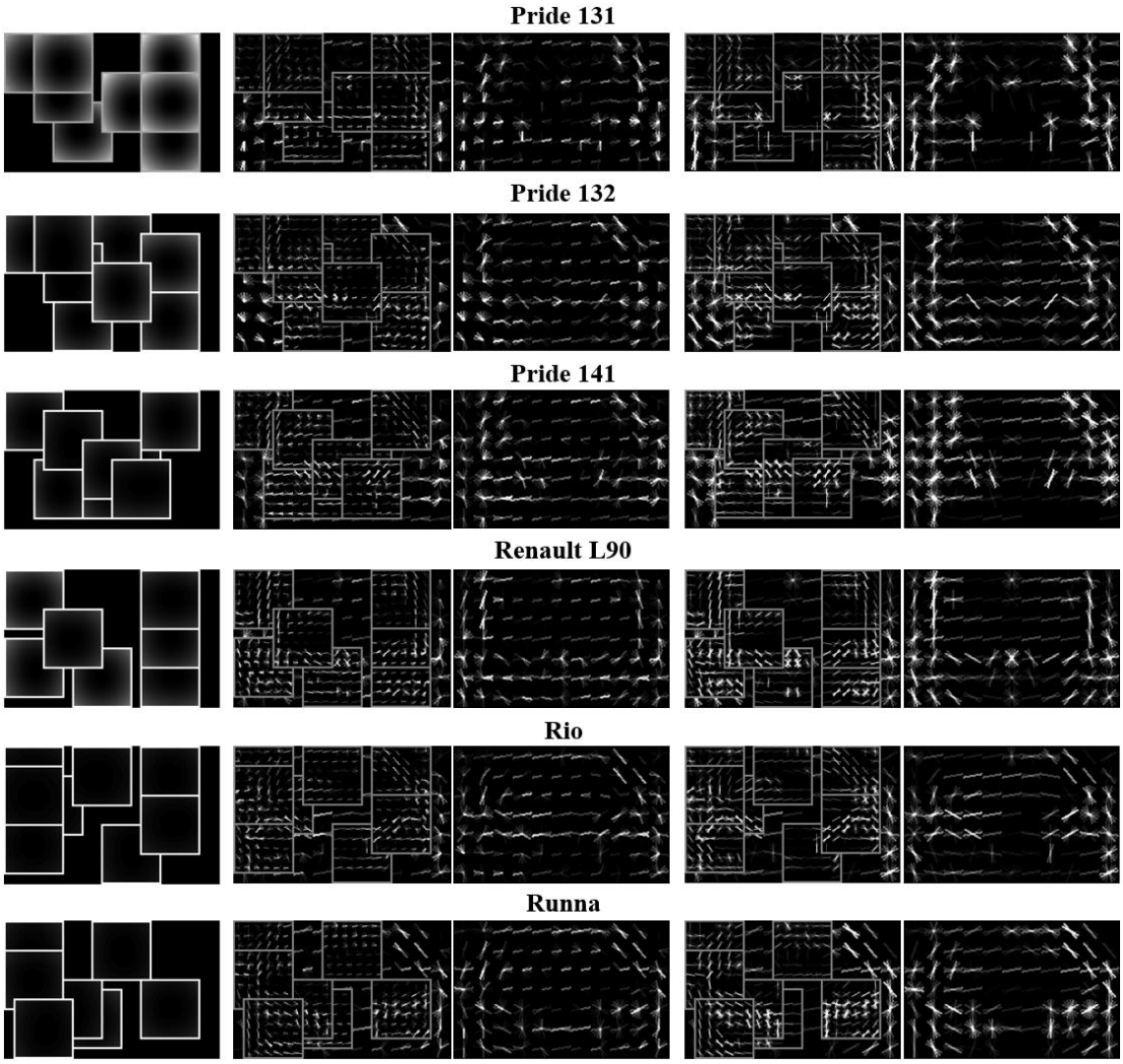
چهار شکل ۷-۵ و شکل ۷-۶ و شکل ۷-۷ و شکل ۷-۸، ۲۱ طبقه‌بند آموزش دیده شده را نمایش داده‌اند. مدل خودروی مربوط به هر طبقه‌بند در بالای آن نوشته شده است. هر طبقه‌بند در تصویر دارای ۵ بخش است که از سمت راست عبارتند از: فیلتر ریشه بدون علامت، فیلتر بخش‌های بدون علامت، فیلتر ریشه علامت‌دار، فیلتر بخش‌های علامت‌دار و خطای جابجایی بخش‌ها نسبت به ریشه. در بخش آخر، محدوده جابجایی هر بخش از طبقه‌بند توسط یک مربع مشخص شده است؛ رنگ‌های تیره‌تر در این مربع به معنی خطای جابجایی کمتر هستند.



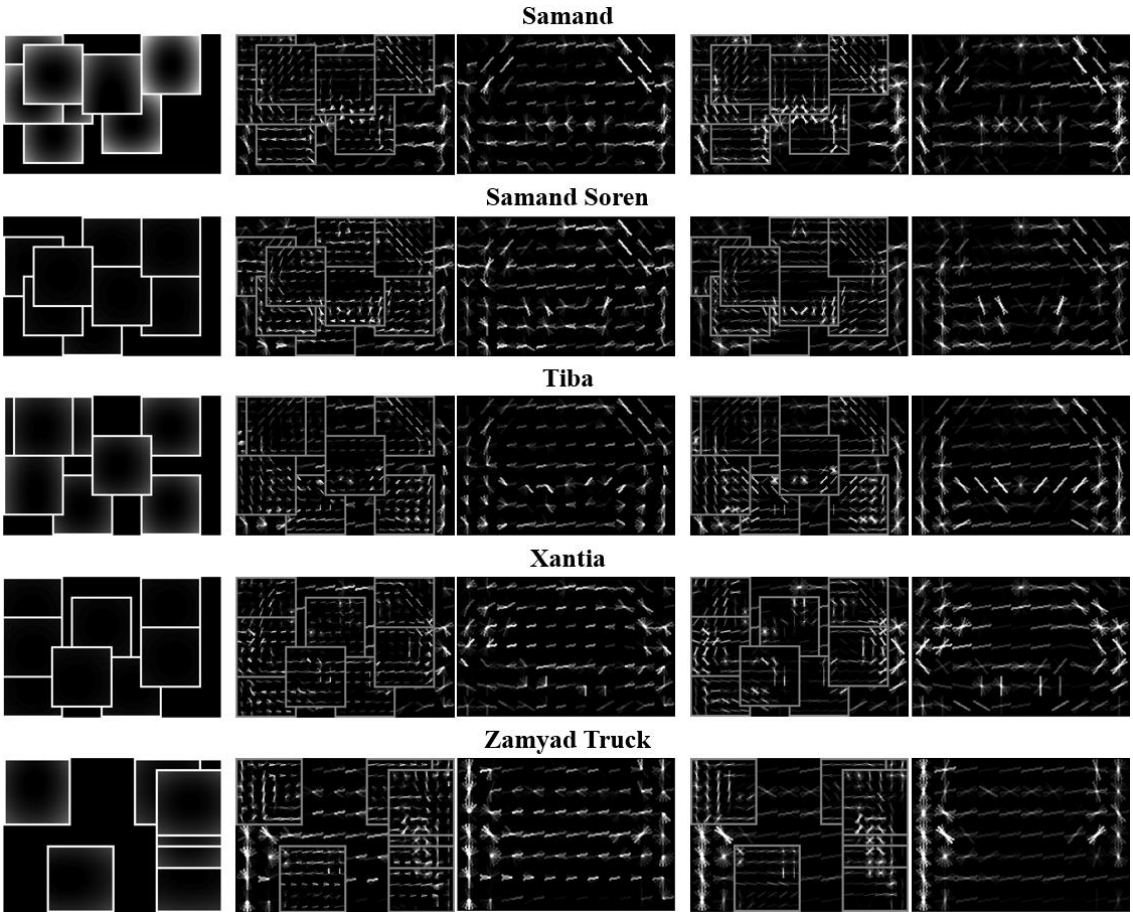
شکل ۷-۱: طبقه‌بندهای آموزش داده شده برای مدل‌های لیفان ۶۲۰، وانت مزدا، مگان، ام‌وی‌ام ۳۱۵ و ام‌وی‌ام ۵۳۰.



شکل ۷-۲: طبقه‌بندهای آموزش داده شده برای مدل‌های پژو ۲۰۶، پژو ۴۰۵، پژو ۴۰۵ SLX، پژو پارس و پیکان ۸۰.

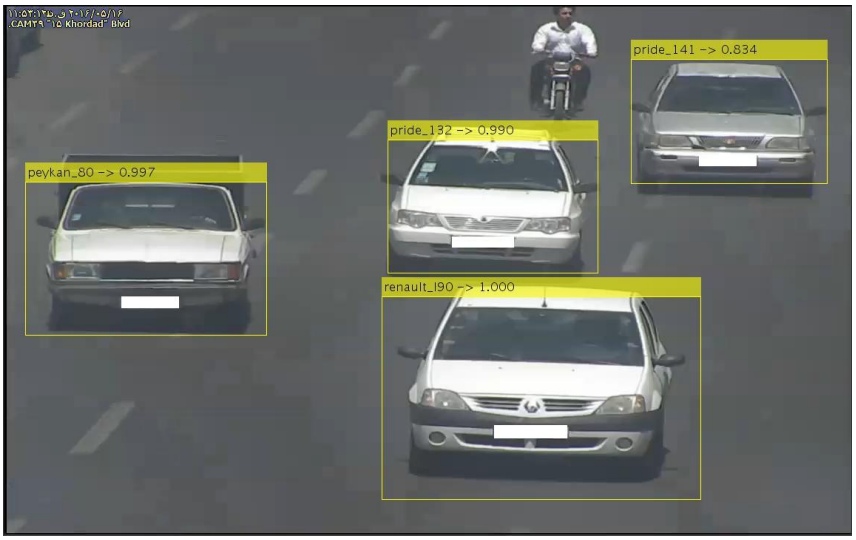


شکل ۷-۳: طبقه‌بندهای آموزش داده شده برای مدل‌های پراید ۱۳۱، پراید ۱۳۲، پراید ۱۴۱، ال ۹۰، ریو و رانا.

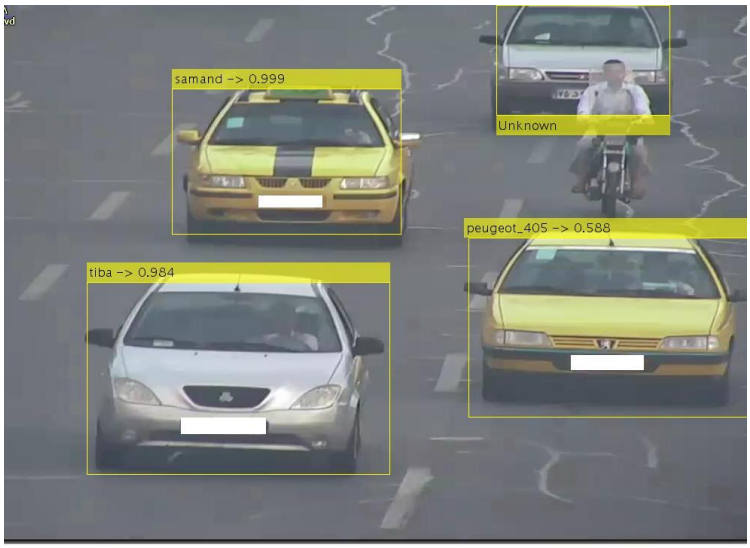


شکل ۷-۴: طبقه‌بندهای آموزش داده شده برای مدل‌های سمند، سمند سورن، تیبا، زانتیا، وانت زامیاد.

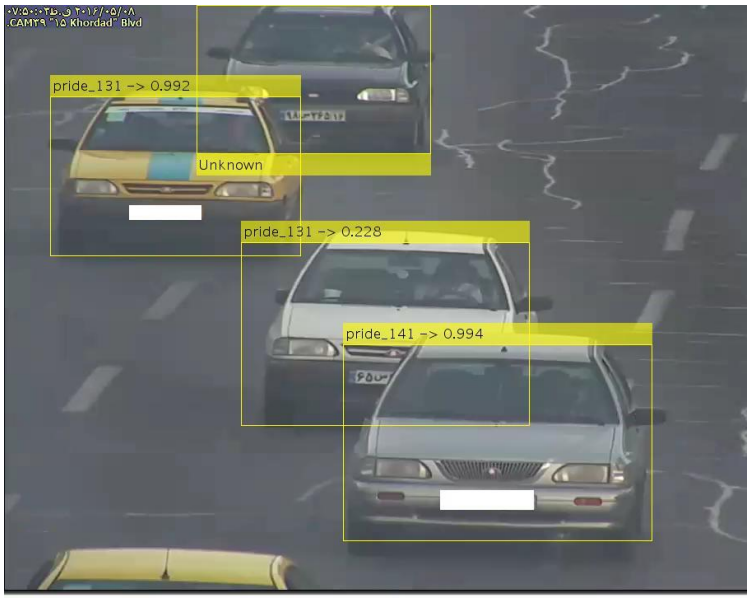
در ادامه چند نمونه از خروجی‌های سیستم بر روی مجموعه داده BVMMR v2 ارائه شده است (شکل ۷-۵ تا شکل ۷-۱۳). برای تولید این خروجی‌ها، سیستم نهایی به تصاویر اعمال شده و پس از تشخیص خودروهای موجود در تصاویر، هر یک از آن‌ها را با نوع و مدل‌شان برچسب‌گذاری می‌نماید. علت ترجیح مجموعه داده BVMMR به مجموعه داده CompCar برای این بخش، وجود چند خودرو در یک تصویر در مجموعه داده BVMMR بوده است. به این ترتیب، می‌توان در فضای کمتر، تعداد شناسایی‌های بیشتری را به نمایش گذاشت.



شکل ۷-۵: خروجی سیستم بر روی یک نمونه تصویر از مجموعه داده BVMMR. مدل همه خودروهای موجود در تصویر به درستی تشخیص داده شده است.



شکل ۷-۶: خروجی سیستم بر روی یک نمونه تصویر از مجموعه داده BVMMR. مدل همه خودروهای موجود در تصویر به درستی تشخیص داده شده است.



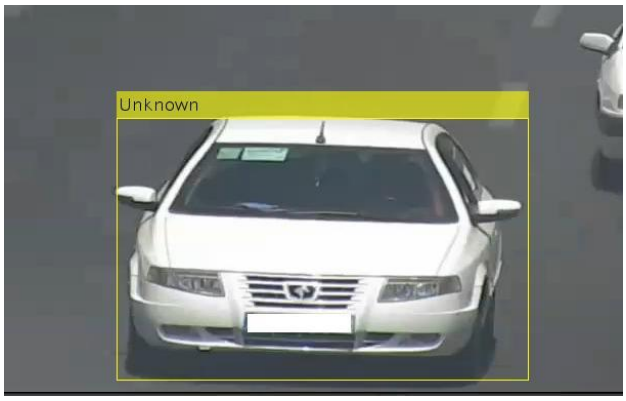
شکل ۷-۷: خروجی سیستم بر روی یک نمونه تصویر از مجموعه داده BVMMR. مدل همه خودروهای موجود در تصویر به درستی تشخیص داده شده است.

شکل ۷-۸: خروجی سیستم بر روی یک نمونه تصویر از مجموعه داده BVMMR. مدل همه خودروهای موجود در تصویر به درستی تشخیص داده شده است.

شکل ۷-۹: خروجی سیستم بر روی یک نمونه تصویر از مجموعه داده BVMMR. مدل همه خودروهای موجود در تصویر به درستی تشخیص داده شده است.

شکل ۷-۱۰: خروجی سیستم بر روی یک نمونه تصویر از مجموعه داده BVMMR. مدل همه خودروهای موجود در تصویر به درستی تشخیص داده شده است.

شکل ۷-۱۱: خروجی سیستم بر روی یک نمونه تصویر از مجموعه داده BVMMR. مدل خودروی پژو ۴۰۵ GLX به اشتباه پژو ۴۰۵ تشخیص داده شده است.



شکل ۷-۱۲: خروجی سیستم بر روی یک نمونه تصویر از مجموعه داده BVMMR. مدل خودروی سمند سورن توسط سیستم تشخیص داده نشده است. علت این امر، تعداد تصاویر کم آموزشی این مدل از خودرو بوده است.

شکل ۷-۱۳: خروجی سیستم بر روی یک نمونه تصویر از مجموعه داده BVMMR. مدل خودروی پژو پارس به اشتباه پژو ۴۰۵ تشخیص داده شده است.

در پایان، طبقه‌بندهای آموزش دیده شده بر روی مجموعه داده CompCars بررسی شده و کلاس‌هایی که به صورت دو به دو، بیشترین دشواری را در شناسایی ایجاد کرده‌اند، تشخیص داده شده‌اند. در ادامه چند نمونه تصویر از این کلاس‌ها ارائه گردیده است. شباهت بسیار زیاد کلاس‌ها در این تصاویر قابل مشاهده است؛ به شکلی که حتی در مواردی، یافتن یک تفاوت توسط انسان به سادگی قابل انجام نیست.

haima_3_1440

haima_haydo_1446

شکل ۷-۱۴: دو نمونه از کلاس‌های بسیار شبیه به هم در مجموعه داده CompCars که سیستم پیشنهادی در تفکیک آن‌ها خطای بیشتری داشته است.

byd_l3_302

byd_f3_314

شکل ۷-۱۵: دو نمونه از کلاس‌های بسیار شبیه به هم در مجموعه داده CompCars که سیستم پیشنهادی در تفکیک آن‌ها خطای بیشتری داشته است.

volkswagen_gran_lavida_439

volkswagen_lavide_440

شکل ۷-۱۶: دو نمونه از کلاس‌های بسیار شبیه به هم در مجموعه داده CompCars که سیستم پیشنهادی در تفکیک آن‌ها خطای بیشتری داشته است.

venucia_r50_1859

venucia_d50_1863

شکل ۷-۱۷: دو نمونه از کلاس‌های بسیار شبیه به هم در مجموعه داده CompCars که سیستم پیشنهادی در تفکیک آن‌ها خطای بیشتری داشته است.

مراجع

[1] Bhardwaj, D., Mahajan, S. “Review Paper on Automated Number Plate Recognition Techniques.”, *International Journal of Emerging Research in Management &Technology*., vol. 4, no. 5, pp. 319–324, 2015.

[2] Du, S., Ibrahim, M., Shehata, M., Badawy, W. “Automatic License Plate Recognition (ALPR): A State-of-the-Art Review.”, *IEEE Transactions on Circuits and Systems for Video Technology*., vol. 23, no. 2, pp. 311–325, 2013.

[3] Patel, C., Shah, D., Patel, A. “Automatic Number Plate Recognition System (ANPR): A Survey.”, *International Journal of Computer Applications*., vol. 69, no. 9, pp. 21–33, 2013.

[4] Sonavane, K., Soni, B., Majhi, U. “Survey on Automatic Number Plate Recognition (ANR).”, *International Journal of Computer Applications*., vol. 125, no. 6, pp. 6–9, 2015.

[5] California, S., Angeles, L. “Car Detection Using Deformable Part Models with Composite Features.”, *IEEE International Conference on Image Processing*, pp. 3812–3816, 2016.

[6] Al-Smadi, M., Abdulrahim, K., Salam, R.A. “Traffic Surveillance: A Review of Vision Based Vehicle Detection, Recognition and Tracking.”, *International Journal of Applied Engineering Research*., vol. 11, no. 1, pp. 713–726, 2016.

[7] Ren, S., He, K., Girshick, R., Sun, J. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.”, *Advances in neural information processing systems*, pp. 91–99, 2015.

[8] Sivaraman, S., Trivedi, M.M. “Looking at Vehicles on the Road: A Survey of Vision-Based Vehicle Detection, Tracking, and Behavior Analysis.”, *IEEE Transactions on Intelligent Transportation Systems*., vol. 14, no. 4, pp. 1773–1795, 2013.

[9] 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.

[10] Huttunen, H., Yancheshmeh, F.S., Chen, K. “Car Type Recognition with Deep Neural Networks.”, *arXiv preprint arXiv:1602.07125*., 2016.

[11] Wang, S., Liu, F., Gan, Z., Cui, Z. “Vehicle Type Classification via Adaptive Feature Clustering for Traffic Surveillance Video.”, *8th IEEE International Conference on Wireless Communications & Signal Processing*, pp. 1–5, 2016.

[12] Zhou, Y., Liu, L., Shao, L., Mellor, M. “DAVE: A Unified Framework for Fast Vehicle Detection and Annotation.”, *arXiv:1607.04564v2*., pp. 1–16, 2016.

[13] Dong, Z., Wu, Y., Pei, M., Jia, Y. “Vehicle Type Classification Using a Semisupervised Convolutional Neural Network.”, *IEEE Transactions on Intelligent Transportation Systems*., vol. 16, no. 4, pp. 2247–2256, 2015.

[14] 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.

[15] Yousaf, K., Iftikhar, 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.

[16] Yu, S., Wu, Y., Li, W., Song, Z., Zeng, W. “A Model for Fine-Grained Vehicle Classification Based on Deep Learning.”, *Neurocomputing*, 2017.

[17] Li, B., Dong, Y., Zhao, D., Wen, Z., Yang, L. “A PCANet Based Method for Vehicle Make Recognition.”, *19th IEEE International Conference on Intelligent Transportation Systems*, pp. 2404–2409, 2016.

[18] Boonsim, N., Prakoonwit, S. “Car Make and Model Recognition under Limited Lighting Conditions at Night.”, *Pattern Analysis and Applications*., no. DOI:10.1007/s10044-016-0559-6, pp. 1–13, 2016.

[19] He, H., Shao, Z., Tan, J. “Recognition of Car Makes and Models From a Single Traffic-Camera Image.”, *IEEE Transactions on Intelligent Transportation Systems*., vol. 16, no. 6, pp. 3182–3192, 2015.

[20] 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.

[21] Ziegler, C. “Targeted Ads Are Coming to Your Car’s Dashboard,” [Online] 2015, https://www.yahoo.com/tech/targeted-ads-are-coming-to-your-cars-dashboard-107164709024.html (Accessed: 25 September 2016).

[22] Anandaraj, S., Mohana Arasi, M. “Intelligent Toll Path System Using GPS and GSM.”, *International Journal of Advanced Research in Computer and Communication Engineering*., vol. 5, no. 5, pp. 44–47, 2016.

[23] Mahajan, S. “Microcontroller Based Automatic Toll Collection System.”, *International Journal of Information and Computation Technology*., vol. 3, no. 8, pp. 793–800, 2013.

[24] Shah, K., Joshi, P., Garg, D. “Automatic Toll Collection Using QR Code.”, *International Research Journal of Engineering and Technology*., vol. 3, no. 10, pp. 1321–1323, 2016.

[25] A. Kurdi, H. “Review of Closed Circuit Television (CCTV) Techniques for Vehicles Traffic Management.”, *International Journal of Computer Science and Information Technology*., vol. 6, no. 2, pp. 199–206, 2014.

[26] 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.

[27] Sun, Z., George, B., Ronald, M. “On-Road Vehicle Detection Using Optical Sensors: A Review.”, *7th IEEE International Conference on Intelligent Transportation Systems*, pp. 585–590, 2004.

[28] Dalka, P., Czyżewski, A. “Vehicle Classification Based on Soft Computing Algorithms.”, *International Conference on Rough Sets and Current Trends in Computing*, pp. 70–79, 2010.

[29] 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.

[30] 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.

[31] Conos, M. “Recognition of Vehicle Make from a Frontal View.”, Master Thesis, Czech Tech, 2007.

[32] Lowe, D.G. “Object Recognition from Local Scale-Invariant Features.”, *17th IEEE International Conference on Computer Vision*, pp. 1150–1157, 1999.

[33] Fisher, R.A. “The Use of Multiple Measurements in Taxonomic Problems.”, *Annals of eugenics*., vol. 7, no. 2, pp. 179–188, 1936.

[34] Dlagnekov, L. “Video-Based Car Surveillance: License Plate, Make, and Model Recognition.”, Master Thesis, University of California, San Diego, 2005.

[35] 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.

[36] 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, 2008.

[37] Huang, H., Zhao, Q., Jia, Y., Tang, S. “A 2DLDA Based Algorithm for Real Time Vehicle Type Recognition.”, *11th IEEE International Conference on Intelligent Transportation Systems*, pp. 298–303, 2008.

[38] 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.

[39] Saravi, S., Edirisinghe, E. a. “Vehicle Make and Model Recognition in CCTV Footage.”, *18th International Conference on Digital Signal Processing*, pp. 1–6, 2013.

[40] Petrovic, V., Cootes, T. “Analysis of Features for Rigid Structure Vehicle Type Recognition.”, *British Machine Vision Conference*, pp. 587–596, 2004.

[41] Harris, C., Stephens, M. “A Combined Corner and Edge Detector.”, *Alvey vision conference*, p. 50, 1988.

[42] Munroe, D.T., Madden, M.G. “Multi-Class and Single-Class Classification Approaches to Vehicle Model Recognition from Images.”, *16th Irish Conference on Artificial Intelligence and Cognitive Science*, pp. 93–102, 2005.

[43] 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.

[44] Candes, E.J., Donoho, D.L. “Curvelets: A Surprisingly Effective Nonadaptive Representation for Objects with Edges.”, Stanford Univ Ca Dept of Statistics, pp. 1-16, 2000.

[45] Zafar, I., Edirisinghe, E. a., Acar, B.S. “Localised Contourlet Features in Vehicle Make and Model Recognition.”, *SPIE-IS&T Electronic Imaging*, pp. 725105–725115, 2009.

[46] Do, M.N., Vetterli, M. “The Contourlet Transform: An Efficient Directional Multiresolution Image Representation.”, *IEEE Transactions on Image Processing.*, vol. 14, no. 12, pp. 2091–2106, 2005.

[47] Kazemi, F.M., Samadi, S., Poorreza, H.R., Akbarzadeh-T, M.-R. “Vehicle Recognition Based on Fourier, Wavelet and Curvelet Transforms - a Comparative Study.”, *4th International Conference on Information Technology*, pp. 939–940, 2007.

[48] 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.

[49] Hsieh, J., Chen, L. “Vehicle Make and Model Recognition Using Symmetrical SURF.”, *Advanced Video and Signal Based Surveillance*, pp. 472–477, 2013.

[50] 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.

[51] Siddiqui, A.J.A.M., Boukerche, A. “Towards Efficient Vehicle Classification in Intelligent Transportation Systems.”, *5th ACM Symposium on Development and Analysis of Intelligent Vehicular Networks and Applications*, pp. 19–25, 2015.

[52] “NTOU-MMR Dataset,” http://mmplab.cs.ntou.edu.tw/mmplab/MMR/MMR.html (Accessed: 22 October 2016).

[53] Shinozuka, Y., Miyano, R., Minagawa, T., Saito, H. “Vehicle Make and Model Recognition by Keypoint Matching of Pseudo Frontal View.”, *IEEE International Conference on Multimedia and Expo Workshops*, pp. 1–6, 2013.

[54] Baran, R., Glowacz, A., Matiolanski, A. “The Efficient Real- and Non-Real-Time Make and Model Recognition of Cars.”, *Multimedia Tools and Applications.*, vol. 74, no. 12, pp. 4269–4288, 2015.

[55] Viola, P., Jones, M. “Rapid Object Detection Using a Boosted Cascade of Simple Features.”, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 511–518, 2001.

[56] Jang, D., Turk, M. “Car-Rec: A Real Time Car Recognition System.”, *IEEE Workshop on Applications of Computer Vision*, pp. 599–605, 2011.

[57] Psyllos, A.P., Anagnostopoulos, C.-N.E., Kayafas, E. “Vehicle Logo Recognition Using a SIFT-Based Enhanced Matching Scheme.”, *IEEE Transactions on Intelligent Transportation Systems.*, vol. 11, no. 2, pp. 322–328, 2010.

[58] Kovesi, P. “Image Features from Phase Congruency.”, *Videre: Journal of computer vision research.*, vol. 1, no. 3, pp. 1–26, 1999.

[59] 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.

[60] 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.

[61] Sarfraz, M.S., Khan, M.H. “A Probabilistic Framework for Patch Based Vehicle Type Recognition.”, *VISAPP*, pp. 358–363, 2011.

[62] Psyllos, a., Anagnostopoulos, C.N., Kayafas, E. “Vehicle Model Recognition from Frontal View Image Measurements.”, *Computer Standards & Interfaces.*, vol. 33, no. 2, pp. 142–151, 2011.

[63] 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.

[64] 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.

[65] Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., LeCun, Y. “OverFeat: Integrated Recognition, Localization and Detection Using Convolutional Networks.”, *arXiv preprint arXiv:1312.6229.*, 2013.

[66] Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, D.T. “Caffe: Convolutional Architecture for Fast Feature Embedding.”, *22nd ACM International Conference on Multimedia*, pp. 675–678, 2014.

[67] Vedaldi A, L.K. “Matconvnet: Convolutional Neural Networks for Matlab.”, *23rd ACM International Conference on Multimedia*, pp. 689–692, 2015.

[68] Girshick, R., Iandola, F., Darrell, T., Malik, J. “Deformable Part Models Are Convolutional Neural

Networks.", *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 437–446, 2015.

[69] 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.

[70] Gao, Y., Lee, H. "Local Tiled Deep Networks for Recognition of Vehicle Make and Model.", *Sensors*., vol. 16, no. 2, pp. 226, 2016.

[71] Sochor, J., Herout, A., Havel, J. "BoxCars: 3D Boxes as CNN Input for Improved Fine-Grained Vehicle Recognition.", *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3006–3015, 2016.

[72] Fang, J., Zhou, Y., Yu, Y., Sidan, D. "Fine-Grained Vehicle Model Recognition Using A Coarse-to-Fine Convolutional Neural Network Architecture.", *IEEE Transactions on Intelligent Transportation Systems*., pp. 1–11, 2016.

[73] Grauman, K., Darrell, T. "The Pyramid Match Kernel: Discriminative Classification with Sets of Image Features.", *IEEE International Conference on Computer Vision*, pp. 1458–1465, 2005.

[74] Lazebnik, S., Schmid, C., Ponce, J. "Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories.", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2169–2178, 2006.

[75] Felzenszwalb, P., McAllester, D., Ramanan, D. "A Discriminatively Trained, Multiscale, Deformable Part Model.", *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2008.

[76] 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.

[77] Kobayashi, T. "BFO Meets HOG: Feature Extraction Based on Histograms of Oriented P.d.f. Gradients for Image Classification.", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 747–754, 2013.

[78] Gan, L., Liu, P., Wang, L. "Rotation Sliding Window of the Hog Feature in Remote Sensing Images for Ship Detection.", *8th International Symposium on Computational Intelligence and Design*, pp. 401–404, 2015.

[79] Chen, P.-Y., Huang, C.-C., Lien, C.-Y., Tsai, Y.-H. "An Efficient Hardware Implementation of HOG Feature Extraction for Human Detection.", *IEEE Transactions on Intelligent Transportation Systems*, pp. 656–662, 2014.

[80] Czajkowska, J., Pycinski, B., Pietka, E. "HoG Feature Based Detection of Tissue Deformations in Ultrasound Data.", *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6326–6329, 2015.

[81] Krause, J., Jin, H., Yang, J., Fei-Fei, L. "Fine-Grained Recognition without Part Annotations.", *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5546–5555, 2015.

[82] Savalle, P.-A., Tsogkas, S., Papandreou, G., Kokkinos, I. "Deformable Part Models with CNN Features.", *European Conference on Computer Vision, Parts and Attributes Workshop*, 2014.

[83] "The PASCAL Visual Object Classes," [Online] 2008, http://pascallin.ecs.soton.ac.uk/challenges/VOC/ (Accessed: 10 March 2015).

[84] Yang, L., Luo, P., Loy, C.C., Tang, X. "A Large-Scale Car Dataset for Fine-Grained Categorization and Verification.", *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3973–3981, 2015.

[85] 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.

[86] Wan, F., Wei, P., Han, Z., Fu, K., Ye, Q. "Weakly Supervised Object Detection with Correlation and Part Suppression.", *IEEE International Conference on Image Processing*, pp. 3638–3642, 2016.

[87] Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S. "Feature Pyramid Networks for Object Detection.", *arXiv preprint arXiv:1612.03144.*, 2016.

[88] Girshick, R., Donahue, J., Darrell, T., Malik, J. "Region-Based Convolutional Networks for Accurate Object Detection and Segmentation.", *IEEE transactions on pattern analysis and machine intelligence*., vol. 38, no. 1, pp. 142–158, 2016.

[89] Zughrat, A., Mahfouf, M., Yang, Y.Y., Thornton, S. "Support Vector Machines for Class Imbalance Rail Data Classification with Bootstrapping-Based Over-Sampling and Under-Sampling.", *IFAC Proceedings Volumes*., vol. 47, no. 3, pp. 8756–8761, 2014.

[90] Reed, S., Lee, H., Anguelov, D., Szegedy, C., Erhan, D., Rabinovich, A. "Training Deep Neural Networks on Noisy Labels with Bootstrapping.", *arXiv preprint arXiv:1412.6596.*, 2014.

[91] Kalal, Z., Matas, J., Mikolajczyk, K. "P-N Learning: Bootstrapping Binary Classifiers by Structural Constraints.", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 49–56, 2010.

[92] Wang, L.W.L., Chan, K.L.C.K.L., Zhang, Z.Z.Z. "Bootstrapping SVM Active Learning by Incorporating Unlabelled Images for Image Retrieval.", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 629–634, 2003.

[93] Felzenszwalb, P.F., Huttenlocher, D.P. "Pictorial Structures for Object Recognition.", *International Journal of Computer Vision*., vol. 61, no. 1, pp. 55–79, 2005.

[94] 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.

[95] Weber, M., Welling, M., Perona, P. "Towards Automatic Discovery of Object Categories.", *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 101–108, 2000.

[96] Girshick, R. "Object Detection with Grammar Models.", *Advances in Neural Information Processing Systems*, pp. 442–450, 2011.

[97] Platt, J. "Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods.", *Advances in large margin classifiers*., vol. 10, no. 3, pp. 61–74, 1999.

[98] Angelova, A., Krizhevsky, A., Vanhoucke, V., Ogale, A., Ferguson, D. "Real-Time Pedestrian Detection with Deep Network Cascades.", *26th British Machine Vision Conference*, pp. 1–12, 2015.

[99] Zhuang, X., Kang, W., Wu, Q. "Real-Time Vehicle Detection with Foreground-Based Cascade Classifier.", *IET Image Processing*., vol. 10, no. 4, pp. 289–296, 2016.

[100] Cai, Z., Vasconcelos, N. "A Real-Time Cascade Pedestrian Detection Based on Heterogeneous Features.", *International SoC Design Conference*, pp. 187–188, 2015.

[101] Tulyakov, S., Jaeger, S., Govindaraju, V., Doermann, D. "Review of Classifier Combination Methods.", *Studies in Computational Intelligence*., vol. 90, pp. 361–386, 2008.

[102] Bourdev, L., Brandt, J. "Robust Object Detection via Soft Cascade.", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 236–243, 2005.

[103] Felzenszwalb, P.F., Girshick, R.B., McAllester, D. "Cascade Object Detection with Deformable Part Models.", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2241–2248, 2010.

Abstract

In fine-grained recognition, the main category of the object is known and the goal is to determine the subcategory or fine-grained category. Vehicle Make and Model Recognition (VMMR) is a hard fine-grained classification problem, due to the large number of classes, substantial inner-class and small inter-class distance. Vehicle analysis is widely used in various applications such as intelligent traffic and transport systems, intelligent parking, automatic toll collection and number-plate forgery detection.

In this paper, a novel approach has been proposed for VMMR which tries to overcome the existing challenges. This approach automatically finds a set of discriminative parts for each class of vehicles. To obtain a powerful classifier, the proposed approach employs a multi-class training procedure which includes a novel greedy parts localization algorithm and a practical multi-class datamining method. For training and evaluation purposes, two new datasets with different properties are collected. These datasets include 720 and 5000 images respectively. Recently, comprehensive CompCars dataset has been released with 50000 images of more than 200 makes and models. The experimental results on BVMMR datasets and CompCars dataset show accuracies more than 96% which indicates the outstanding performance of our approach.

By increasing the number of classifiers in the proposed system, the processing speed will degrade. Moreover, an unbalanced training data may lead to overfitting problems or classifiers with low generality. A practical cascading scheme is proposed for reducing these effects and boosting the system speed. The proposed cascading scheme construct a cascade of classifiers by introducing two criteria, namely "Confidence" and "Frequency". The experiments done on BVMMR and CompCars datasets confirm the effectiveness of our scheme. One of the configurations of the proposed scheme results in up to 30% speedup in comparison to the baseline VMMR system with analogous recognition rate. Another configuration of the proposed cascading scheme achieves up to 80% speedup with just a minor drop in accuracy.

Keywords: Fine-grained Recognition, Make and Model Recognition, VMMR, Part-based Recognition, Latent SVM

Shahrood University of Technology

Faculty of Computer Engineering

PhD Dissertation in Artificial Intelligence

Vehicle Make and Model Recognition using Part-based Models

By: **Mohsen Biglari**

Supervisor:

Dr. Seyyed Ali Soleimani

Advisor:

Dr. Hamid Hassanpour

July 2017