

- پرسش ۲. پیاده‌سازی یک سیستم طبقه‌بندی خودرو ۱
- ۱-۲- پیش پردازش داده ۱
- ۱-۱-۲- نمودار فراوانی داده‌ها و برجسپ عددی هر کلاس ۱
- ۲-۱-۲- بالانس کردن داده ۲
- ۳-۱-۲- گزارش ابعاد داده آموزش و تست ۲
- ۲-۲- آموزش و ارزیابی مدل ۲
- ۱-۲-۲- نحوه ی عملکرد معیارهای Recall, Accuracy, F1Score, Precision ۲
- ۲-۲-۲- گزارش معیارهای ارزیابی و نتایج VGG, AlexNet, SVM+VGG, CNN ۳
- ۳-۲-۲- ماتریس درهم‌ریختگی VGG, AlexNet, SVM+VGG, CNN ۴
- ۴-۲-۲- اصلی‌ترین تفاوت بین AlexNet و VGGNet ۶
- ۳-۲- تحلیل نتایج ۷
- ۴-۲- امتیازی ۱۰
- ۱-۴-۲- مدل SVM+VGG با کرنل RBF ۱۰
- ۲-۴-۲- تاثیر استخراج ویژگی بر نتیجه نهایی در طبقه بندی ۱۱

شکل‌ها

- شکل ۱۴. نمودار فراوانی دیتاست toyota ۱
- شکل ۱۵. ماتریس آشفتگی ۴ مدل مختلف برای دیتاست toyota ۵
- شکل ۱۶. نمودار ارزیابی عملکرد مدل‌های مختلف روی دیتاست toyota ۷
- شکل ۱۷. ماتریس آشفتگی مدل SVM با کرنل غیرخطی RBF ۱۱

جدول‌ها

جدول ۳. تقسیم‌بندی داده‌ها به دو مجموعه آموزشی و تستی به نسبت ۸۰:۲۰، شامل اندازه تصاویر و

برچسب‌ها..... ۲

جدول ۴. ارزیابی عملکرد مدل‌ها روی دیتاست toyota..... ۳

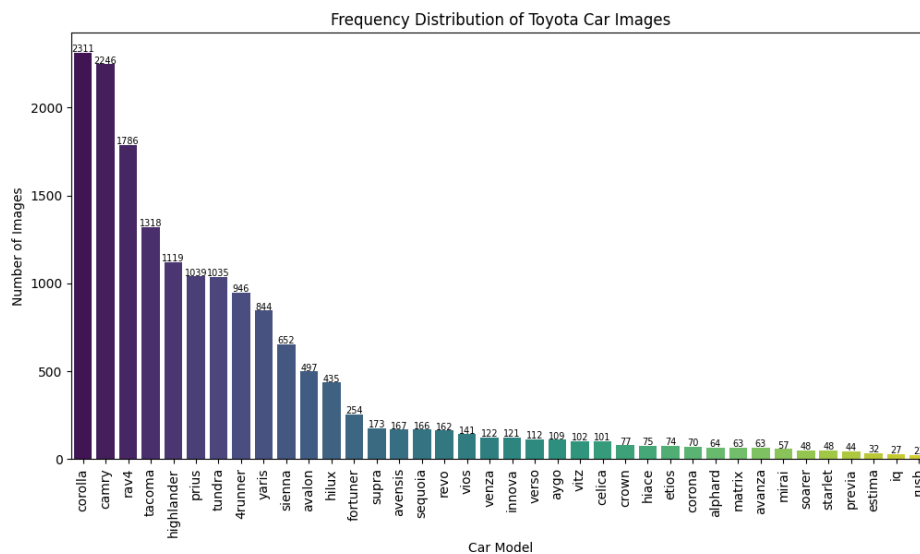
جدول ۵. ارزیابی عملکرد SVM با کرنل غیرخطی RBF روی دیتاست toyota..... ۱۰

پرسش ۲. پیاده‌سازی یک سیستم طبقه‌بندی خودرو

تمام کدها و نمودارها و نتایج حاصل در فایل VGG_SVM.ipynb واقع در پوشه appendix در مسیر اصلی فولدر موجود می‌باشد.

۱-۲- پیش پردازش داده

۱-۱-۲- نمودار فراوانی داده‌ها و برچسپ عددی هر کلاس



شکل ۱. نمودار فراوانی دیتاست toyota

۱۰ کلاس انتخاب شده توسط ما و لیبل و ایندکس هر کلاس. توجه شود که در ماتریس آشفتگی از

لیبل‌ها یا ایندکس استفاده شده است.

Index	Car Model	Num of Imgs	Car Model Encoded
0	runner	946	0
1	alphard	64	1
2	camry	2246	6
3	hilux	435	16
4	iq	27	18
5	prius	1039	22
6	soarer	48	28
7	supra	173	30
8	tundra	1035	32
9	vitx	102	36

۲-۱-۲- بالانس کردن داده

در این مرحله، یکی از مهم‌ترین اقداماتی که می‌توان انجام داد، متعادل‌سازی توزیع داده‌ها میان کلاس‌های مختلف است. ما در این پیاده‌سازی، از روش Data Augmentation برای کلاس‌های دارای نمونه کمتر استفاده کرده‌ایم تا به سطح تعادل مطلوب‌تری (هر کلاس ۵۰۰ نمونه) در داده‌ها برسیم و کلاس‌هایی با داده‌های فراوان‌تر را نیز فقط به اندازه ۵۰۰ نمونه برمی‌داریم به این ترتیب کلاس‌های ما متعادل هستند.

به جای انتخاب ساده یا تکرار کورکورانه تصاویر، از تکنیک‌هایی مانند چرخش تصادفی، برش، وارونگی افقی، تغییر روشنایی و کنتراست بهره گرفته‌ایم. این روش باعث می‌شود نه تنها تعداد داده‌ها در کلاس‌های اقلیت افزایش پیدا کند، بلکه تنوع و غنای بصری نیز حفظ شود.

در مقایسه با روش‌هایی مانند Oversampling ساده (کپی مستقیم تصاویر)، استفاده از Data Augmentation منجر به تولید نمونه‌هایی با ویژگی‌های دیداری متفاوت و متنوع می‌شود که به مدل کمک می‌کند توانایی تعمیم‌دهی به داده‌های جدید را به شکل بهتری یاد بگیرد و توجه شود که هرچه داده‌ها متنوع‌تر باشد مدل مقاوم‌تر به شرایط واقعی خواهد بود.

۲-۱-۳- گزارش ابعاد داده آموزش و تست

جدول ۱. تقسیم‌بندی داده‌ها به دو مجموعه آموزشی و تستی به نسبت ۸۰:۲۰، شامل اندازه تصاویر و برچسب‌ها

مجموعه داده	تعداد نمونه	Image Shape	Label Shape
آموزشی (Train)	۴۰۰۰	(۳، ۲۲۴، ۲۲۴، ۴۰۰۰)	(۴۰۰۰،)
تستی (Test)	۱۰۰۰	(۳، ۲۲۴، ۲۲۴، ۱۰۰۰)	(۱۰۰۰،)

۲-۲- آموزش و ارزیابی مدل

۲-۲-۱- نحوه ی عملکرد معیارهای Recall, Accuracy, F1Score, Precision

- ❖ True Positive (TP): مدل درست تشخیص داده که ماشین corolla هست واقعاً هم corolla بوده.
- ❖ False Positive (FP): مدل گفته corolla هست، ولی اشتباه کرده و مثلاً ماشین supra بوده.
- ❖ False Negative (FN): ماشین واقعاً corolla بوده، ولی مدل اشتباه گفته نبوده.
- ❖ True Negative (TN): ماشین corolla نبوده و مدل هم درست تشخیص داده که نیست.

Accuracy: این معیار نسبت پیش‌بینی‌های درست از کل پیش‌بینی‌ها را بیان می‌کند، اما در داده‌های نامتوازن (مثلاً وقتی یک کلاس خیلی بیشتره) ممکنه گمراه‌کننده باشه.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision: این معیار بیان می‌کند که چه نسبتی از پیش‌بینی‌های مثبت تشخیص داده شده، واقعاً مثبت بوده‌اند. در صورتی که تعداد کلاس‌ها از دوتا بیشتر باشد، می‌توان میانگین precision ها را محاسبه کرد.

$$Precision = \frac{TP}{TP + FP}$$

Recall: این معیار نشان می‌دهد که چه نسبتی از نمونه‌های مثبت را به‌درستی شناسایی کرده‌ایم. مشابه شاخص قبل، در صورتی که تعداد کلاس‌ها از دوتا بیشتر باشد، می‌توان میانگین را گزارش نمود.

$$Recall = \frac{TP}{TP + FN}$$

F1-Score: این معیار تعادل بین Precision و Recall را بیان می‌کند و برای زمانی خوبه که هم Precision و هم Recall مهم باشند، مخصوصاً وقتی داده‌ها نامتوازن هستند و به تعادل بین Precision و Recall اهمیت می‌دهیم.

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

۲-۲-۲- گزارش معیارهای ارزیابی و نتایج VGG, AlexNet, SVM+VGG, CNN

جدول ۲. ارزیابی عملکرد مدل‌ها روی دیتاست toyota

مدل	Precision	Accuracy	F1 Score	Recall
SVM+VGG	۰.۸۴۰۶	۰.۸۴۲۰	۰.۸۴۰۷	۰.۸۴۲۰
VGG16	۰.۸۴۰۱	۰.۸۲۱۰	۰.۸۱۵۰	۰.۸۲۱۰
AlexNet	۰.۸۰۸۳	۰.۸۰۶۰	۰.۸۰۳۸	۰.۸۰۶۰
CNN	۰.۴۹۲۷	۰.۵۴۲۰	۰.۵۱۱۴	۰.۵۴۲۰

همانطور که مشهود است، مدل SVM+VGG با دستیابی به دقت ۰.۸۴۲۰، بالاترین عملکرد را در بین مدل‌های مورد بررسی داشته و در تمامی معیارها (Precision، Recall و F1) نیز بیشترین مقدار را کسب کرده است. این امر نشان‌دهنده تأثیر مثبت استخراج ویژگی‌های عمیق توسط VGG و سپس طبقه‌بندی با SVM است.

مدل VGG16 با دقت ۰.۸۲۱۰ در رتبه دوم قرار دارد و نشان می‌دهد استفاده مستقیم از شبکه‌های پیش‌آموزش‌دیده نیز می‌تواند عملکرد مناسبی ارائه دهد.

مدل AlexNet با دقت ۰.۸۰۶۰ عملکرد متوسطی داشته و F1 آن (۰.۸۰۳۸) نشان می‌دهد که در حفظ تعادل میان Precision و Recall عملکرد نسبتاً پایداری داشته است.

در مقابل، مدل CNN ساده با دقت ۰.۵۴۲۰ و مقدار پایین F1 (۰.۵۱۱۴) عملکرد ضعیفی از خود نشان داده و بیانگر این است که معماری‌های ساده‌تر در مواجهه با داده‌های پیچیده‌تر مانند Toyota، کارایی کافی ندارند.

۲-۳-۲- ماتریس درهم‌ریختگی VGG, AlexNet, SVM+VGG, CNN

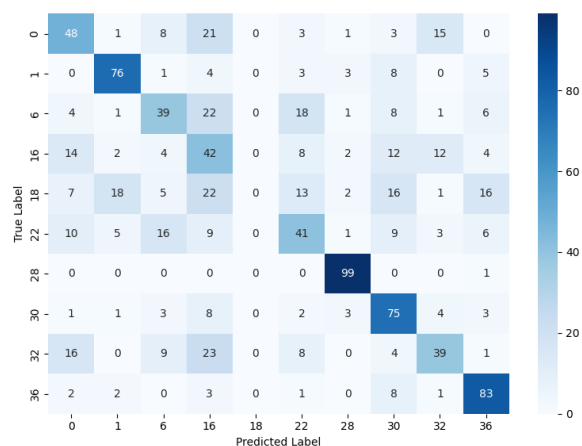
سوال: نحوه ی تحلیل مقادیر عددی ماتریس درهم‌ریختگی را توضیح دهید؟

در یک ماتریس $n \times n$ ، سطرها نمایانگر کلاس‌های واقعی (True Labels) و ستون‌ها نشان‌دهنده کلاس‌های پیش‌بینی‌شده (Predicted Labels) هستند.

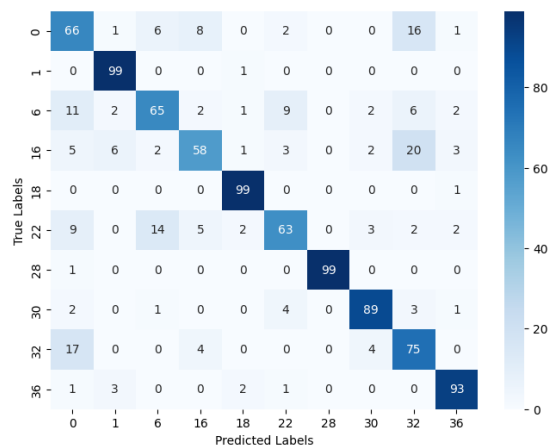
مقادیر قطر اصلی (Diagonal) بیانگر پیش‌بینی‌های درست (True Positives) برای هر کلاس‌اند، در حالی که مقادیر خارج از قطر به خطاهای مدل در طبقه‌بندی بین کلاس‌ها اشاره دارند.

اگر مقدار زیادی از نمونه‌ها در قطر اصلی قرار داشته باشند، نشان‌دهنده عملکرد مناسب مدل است. و اگر مقادیر زیادی در خارج از قطر مشاهده شود، باید بررسی شود که آیا مدل در تمایز بین کلاس‌های خاصی دچار سردرگمی است. تحلیل ستون‌ها و سطرها ی ماتریس می‌تواند به ما نشان دهد که کدام کلاس‌ها بیشتر با کدام کلاس‌ها اشتباه گرفته شده‌اند و این اطلاعات برای بهبود مدل بسیار مفید است.

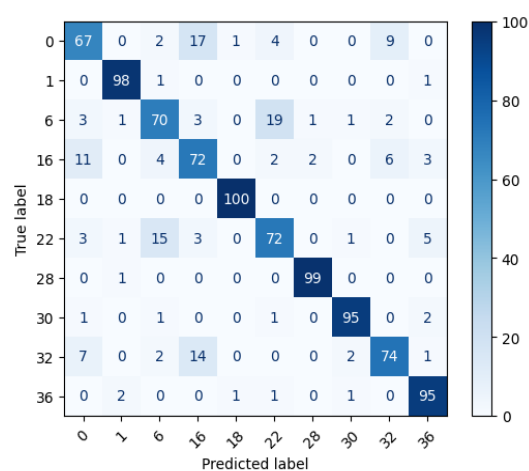
از این ماتریس، معیارهای مهمی چون Accuracy، Precision، Recall و F1-score برای ارزیابی عملکرد مدل به دست می‌آیند.



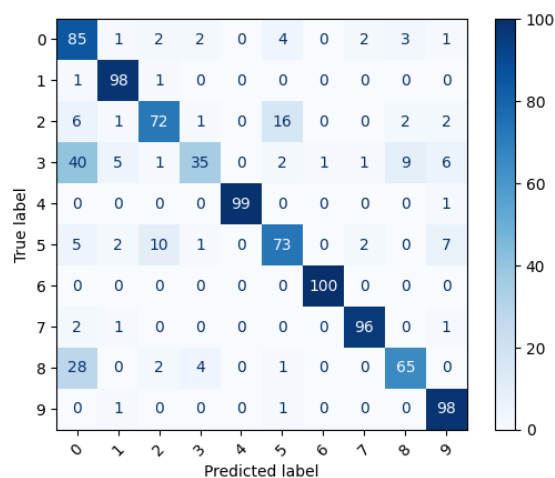
CNN Mode



AlexNet Model



Proposed (VGG + SVM) Model



VGG Model

شکل ۲. ماتریس آشفتگی ۴ مدل مختلف برای دیتاست toyota

مدل ترکیبی VGG+SVM :

این مدل با دقت بالا و کمترین نرخ اشتباه در طبقه‌بندی، عملکرد بسیار قابل قبولی از خود نشان داده است. کلاس‌هایی با فراوانی بالا نظیر Camry، Tundra و Runner با نرخ بالای True Positive شناسایی شده‌اند. علاوه بر این، سردرگمی بین کلاس‌هایی با شباهت ظاهری مانند Camry و Prius نیز به حداقل رسیده است. این موضوع حاکی از قدرت بالای مدل در تعیین مرزهای تصمیم‌گیری دقیق و مؤثر است.

مدل VGG16 :

مدل VGG16 نیز عملکرد مطلوبی در تفکیک کلاس‌های مختلف داشته و اختلاف آن با مدل ترکیبی در اکثر موارد جزئی بوده است. با این حال، در کلاس‌هایی چون Supra و Soarer که شباهت‌های بصری

دارند، اشتباهاتی مشاهده شده است. همچنین، در کلاس‌های کم‌نمونه قبل از اعمال تکنیک داده‌افزایی، نرخ خطای بیشتری دیده می‌شود که با به‌کارگیری داده‌افزایی بهبود یافته است.

مدل AlexNet :

در این مدل، میزان اشتباهات به‌ویژه در کلاس‌هایی مانند IQ، Hilux و Alphard افزایش یافته است. بخش قابل توجهی از این اشتباهات ناشی از شباهت بصری بین کلاس‌ها و همچنین تعداد کم نمونه‌ها در برخی کلاس‌هاست. معماری ساده‌تر این مدل نسبت به VGG باعث شده تا در استخراج ویژگی‌های ظریف عملکرد ضعیف‌تری داشته باشد.

مدل CNN ساده:

ضعیف‌ترین عملکرد در میان مدل‌های بررسی‌شده متعلق به این مدل است. مقادیر قابل توجهی از نمونه‌های کلاس‌های با داده‌ی کم مانند IQ، Vitz و Soarer به‌اشتباه به کلاس‌های پرتکرار نظیر Camry اختصاص یافته‌اند. این موضوع بیانگر تمایل مدل به کلاس‌های غالب (class imbalance bias) و ضعف در یادگیری ویژگی‌های متمایز است. این مدل برای کاربردهای واقعی نیازمند اصلاحات بنیادی در معماری و افزایش ظرفیت یادگیری می‌باشد.

۲-۲-۴- اصلی‌ترین تفاوت بین AlexNet و VGGNet

تفاوت اصلی میان دو مدل AlexNet و VGGNet در ساختار و طراحی معماری آن‌هاست؛ به‌ویژه در عمق شبکه، اندازه فیلترها، و یکنواختی لایه‌ها.

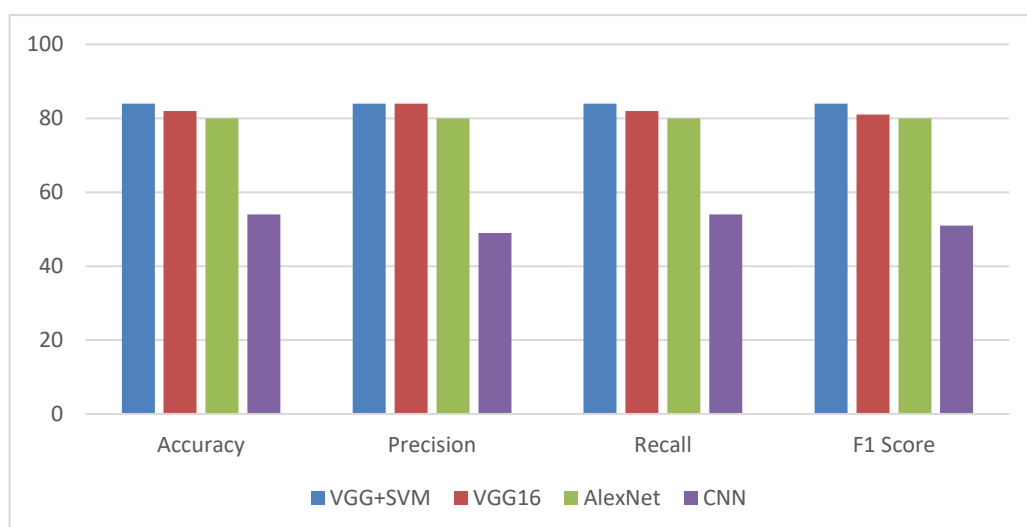
AlexNet از فیلترهای نسبتاً بزرگ (مانند 11×11 و 5×5) در لایه‌های ابتدایی استفاده می‌کند تا با تعداد نسبتاً کمی از لایه‌ها، ویژگی‌های سطح پایین تصویر را استخراج کند. همچنین این مدل برای نخستین بار از تابع فعال‌سازی ReLU به‌طور گسترده بهره‌برد، که نقش مهمی در کاهش مشکل *vanishing gradient* در شبکه‌های عمیق داشت. به‌طور کلی، معماری AlexNet بیشتر شبیه به یک طراحی کلاسیک است: لایه‌ها متنوع، درشت و کم‌عمق‌اند و ساختار کلی ساده‌تر و سریع‌تر پیاده‌سازی می‌شود.

در مقابل، VGGNet با رویکردی متفاوت به طراحی شبکه نگاه می‌کند. این مدل به‌جای استفاده از فیلترهای بزرگ، با بهره‌گیری پیاپی از فیلترهای کوچک 3×3 ، همان receptive field را پوشش می‌دهد.

این کار نه تنها موجب افزایش غیرخطی بودن مدل و قدرت تفکیک بهتر ویژگی‌ها می‌شود، بلکه تعداد پارامترها را هم در هر لایه کنترل می‌کند. ساختار VGGNet منظم‌تر، عمیق‌تر و یکنواخت‌تر است و معمولاً از چندین بلوک تکرارشونده شامل Convolution، ReLU و Pooling تشکیل می‌شود.

در یک نگاه مفهومی، VGGNet را می‌توان مانند یک معمار مدرن دانست: طراحی ساده، یکنواخت و در عین حال کارآمد، با لایه‌های نازک ولی زیاد، که به دقت و قدرت تعمیم بالایی در یادگیری ساختارهای پیچیده تصویر منتهی می‌شود.

۳-۲- تحلیل نتایج



شکل ۳. نمودار ارزیابی عملکرد مدل‌های مختلف روی دیتاست toyota

تحلیل عملکرد مدل‌ها و کلاس‌های شناسایی شد با توجه به جدول ۲ و ماتریس‌های آشفتگی شکل ۲ و شکل ۳:

مدل ترکیبی VGG+SVM بهترین عملکرد را در میان تمامی مدل‌های مورد بررسی داشته و در تمامی معیارهای ارزیابی نظیر Accuracy، Precision، Recall و F1-Score بالاترین نتایج را کسب کرده است. این مدل با بهره‌گیری از قابلیت‌های استخراج ویژگی دقیق VGG16 و قدرت طبقه‌بندی بالای SVM، مرزهای تصمیم‌گیری شفافی ایجاد کرده و توانسته است در کلاس‌هایی با داده‌های پرتعداد مانند Camry (۲۲۴۶ تصویر)، Tundra (۱۰۳۵ تصویر) و Runner (۹۴۶ تصویر) نرخ تشخیص بسیار بالایی را ثبت کند.

در مقابل، در کلاس‌هایی با داده‌های اندک نظیر IQ (۲۷ تصویر)، Soarer (۴۸ تصویر) و تا حدودی Supra (۱۷۳ تصویر) به دلیل محدودیت یادگیری ویژگی‌های تمایزبخش، عملکرد ضعیف‌تری داشته است.

مدل VGG16 نیز به دلیل معماری عمیق خود، در استخراج ویژگی‌های تصویری موفق عمل کرده و در کلاس‌های پرتعداد مانند Camry و Tundra نتایج رضایت‌بخشی ارائه داده است. با این حال، در کلاس‌های کم‌نمونه مانند Vitz، IQ و Soarer و همچنین کلاس‌هایی با شباهت ظاهری بالا نظیر Hilux، با نرخ خطای بیشتری مواجه بوده که عمدتاً ناشی از افزایش FP در مرزبندی میان کلاس‌هاست.

مدل AlexNet با وجود ساختار ساده‌تر و سرعت بالاتر، در شناسایی کلاس‌هایی با نمونه‌های کافی مانند Camry عملکرد متوسطی داشته، اما در تشخیص کلاس‌های پیچیده‌تر یا کم‌نمونه مانند Hilux، Supra، Soarer، IQ و Alphard ضعف قابل توجهی نشان داده است. این افت عملکرد را می‌توان به عمق کم شبکه و قدرت محدود آن در استخراج ویژگی‌های سطح بالا نسبت داد.

در نهایت، مدل CNN ساده ضعیف‌ترین نتایج را در بین تمامی مدل‌ها ارائه کرده است. گرچه در کلاس Camry تا حدودی موفق بوده، اما در اکثر کلاس‌های کم‌نمونه مانند IQ، Soarer، Alphard و Vitz و حتی کلاس‌های با داده متوسط مانند Hilux و Supra دچار نرخ بالای FN و FP شده که نشان‌دهنده محدودیت جدی در یادگیری ویژگی‌های پیچیده است.

نتیجه‌گیری:

نتایج به وضوح نشان می‌دهند که استفاده از معماری‌های عمیق و طبقه‌بندهای پیشرفته نقشی کلیدی در بهبود عملکرد مدل‌های بینایی ماشین دارد. مدل ترکیبی VGG+SVM با بهره‌گیری از استخراج ویژگی‌های غنی و مرزبندی دقیق، در شناسایی کلاس‌های مختلف به‌ویژه در دیتاست‌های با کلاس‌های پرتعداد و ظرافت‌های تصویری بالا، بهترین عملکرد را ارائه داده است. همچنین، به‌کارگیری داده‌افزایی (Data Augmentation) در این پروژه نقشی مؤثر در متعادل‌سازی کلاس‌های دارای داده‌ی کم ایفا کرده و موجب بهبود نسبی عملکرد در شناسایی این دسته از کلاس‌ها شده است.

در مسائل پیچیده‌ای مانند شناسایی خودرو با کلاس‌های متعدد و مشابه، توصیه می‌شود از مدل‌های پیش‌آموزش‌دیده به همراه طبقه‌بندهای قدرتمند بهره گرفته شود تا هم ویژگی‌های سطح بالای تصویر به‌درستی استخراج شود و هم تصمیم‌گیری نهایی با دقت و تفکیک‌پذیری بیشتری انجام گیرد.

سوال: برای بهتر عمل کردن مدل ها چه راهکارهایی پیشنهاد می کنید تا قدرت تشخیص مدل ها افزایش پیدا کند و به طور کلی مدل ها ارتقا پیدا کنند؟

افزایش تعداد نمونه ها به حداقل ۱۰۰۰ نمونه ممکن است تاثیر قابل توجهی بر افزایش عملکرد مدل ها داشته باشد که در این تمرین به دلیل محدودیت سخت افزاری از تعداد فقط ۵۰۰ نمونه استفاده شده است.

۱. بهبود کیفیت و تنوع داده ها

- داده افزایی هدفمند: اعمال تکنیک هایی مانند چرخش، بزرگ نمایی، تغییر روشنایی و نویز، به ویژه برای کلاس های کم نمونه، سبب افزایش تنوع داده ها و تقویت قابلیت تعمیم مدل می شود.
- متعادل سازی کلاس ها (Class Balancing): استفاده از روش هایی مانند Oversampling یا تولید داده ی مصنوعی برای کلاس های دارای نمونه ی کمتر، باعث کاهش bias مدل به سمت کلاس های پرتکرار می شود.
- پاک سازی داده ها: حذف تصاویر بی کیفیت، تار یا دارای برچسب اشتباه به بهبود آموزش مدل کمک می کند.

۲. ارتقای معماری مدل ها و تنظیمات یادگیری

- استفاده از معماری های پیشرفته تر: به کارگیری مدل هایی مانند ResNet ، DenseNet ، EfficientNet یا Vision Transformer می تواند دقت مدل را در تشخیص ویژگی های پیچیده افزایش دهد.
- تنظیم مجدد مدل های پیش آموزش دیده: (Fine-Tuning) تنظیم مجدد لایه های انتهایی مدل های مانند VGG16 یا ResNet می تواند تطبیق پذیری آن ها را با داده های خاص مجموعه افزایش دهد.
- تنظیم های پیرامونها: بهره گیری از روش هایی مانند Grid Search یا Random Search برای انتخاب بهینه ی پارامترهایی چون نرخ یادگیری، اندازه ی Batch و تعداد لایه ها، تاثیر چشم گیری در بهبود عملکرد دارد.
- استفاده از مدل های ترکیبی (Ensemble Learning): ترکیب خروجی چند مدل از طریق Voting یا Stacking می تواند دقت نهایی را افزایش داده و خطاها را کاهش دهد.

۳. پیشگیری از بیش‌برازش و بهینه‌سازی عملکرد

- به‌کارگیری Regularization : استفاده از تکنیک‌هایی مانند Dropout ، Batch Normalization و L2 Regularization می‌تواند از بیش‌برازش جلوگیری کرده و عملکرد مدل را در داده‌های دیده‌نشده ارتقا دهد.
- استفاده از توابع هزینه‌ی مناسب : انتخاب توابعی مانند Focal Loss یا Weighted Cross Entropy در شرایط عدم توازن داده‌ها، تمرکز مدل را بر روی نمونه‌های دشوار افزایش می‌دهد.

۴. تحلیل دقیق‌تر خطاها و یادگیری از اشتباهات

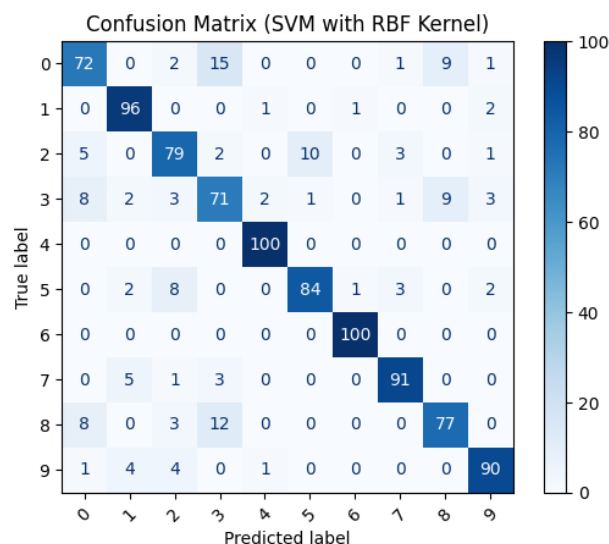
- بررسی دقیق ماتریس درهم‌ریختگی : شناسایی کلاس‌هایی با بیشترین خطای مدل و تحلیل ویژگی‌های آن‌ها می‌تواند به بازطراحی معماری یا تنظیمات کمک کند.
- آموزش بر اساس نمونه‌های دشوار (Hard Example Mining) : تمرکز بر نمونه‌هایی که مدل بیشترین نرخ خطا را روی آن‌ها دارد، به افزایش قدرت تمایز مدل کمک می‌کند.

۲-۴- امتیازی

۲-۴-۱- مدل SVM+VGG با کرنل RBF

جدول ۳. ارزیابی عملکرد SVM با کرنل غیرخطی RBF روی دیتاست toyota

مدل	Accuracy	Precision	Recall	F1 Score
SVM with RBF Kernel	۰.۸۶۰۰	۰.۸۵۹۱	۰.۸۶۰۰	۰.۸۵۹۱



شکل ۴. ماتریس آشفته‌گی مدل SVM با کرنل غیرخطی RBF

نتایج ارزیابی نشان می‌دهد که SVM با کرنل RBF عملکرد بهتری نسبت به نسخه خطی (جدول ۲) از خود نشان داده و در ایجاد توازن میان Precision و Recall موفق‌تر عمل کرده است. کرنل RBF به دلیل ماهیت غیرخطی خود قادر است داده‌هایی با توزیع پیچیده و مرزهای نامنظم را به خوبی تفکیک کند؛ در حالی که SVM با کرنل خطی تنها در شرایطی که داده‌ها در فضای ویژگی به صورت خطی از یکدیگر جدا باشند، عملکرد مطلوبی دارد. در مسائلی که داده‌ها از ساختار پیچیده‌تری برخوردارند – مانند طبقه‌بندی تصاویر خودروها با شباهت‌های ظاهری بالا – مدل مبتنی بر RBF با نداشتن داده‌ها به فضای ویژگی با ابعاد بالاتر، مرزهای تصمیم‌گیری دقیق‌تری ایجاد می‌کند. بنابراین، با توجه به نتایج حاصل شده، SVM با کرنل RBF می‌تواند انتخاب مناسب‌تری برای این مسئله محسوب می‌شود.

۲-۴-۲- تأثیر استخراج ویژگی بر نتیجه نهایی در طبقه‌بندی

در مسائل طبقه‌بندی، به‌ویژه زمانی که داده‌های آموزشی محدود یا غیرموازنه (imbalanced) هستند، استخراج ویژگی به عنوان یک گام میانی مؤثر عمل می‌کند و می‌تواند به شکل قابل توجهی بر کیفیت نهایی مدل تأثیر بگذارد.

تحلیل تأثیر استخراج ویژگی بر نتیجه‌ی نهایی:

- کاهش بُعد و نویز داده‌ها: با استخراج ویژگی‌های غنی و فشرده از تصاویر، مدل نهایی با اطلاعات مهم‌تر و تمیزتری آموزش می‌بیند و از پیچیدگی محاسباتی کاسته می‌شود.

- افزایش تعمیم‌پذیری مدل: ویژگی‌های استخراج‌شده، نمایی انتزاعی‌تر و معنادارتر از تصاویر هستند و مدل نهایی را قادر می‌سازند تا بهتر به داده‌های دیده‌نشده پاسخ دهد.
 - تقویت طبقه‌بند ساده‌تر: در این مقاله، طبقه‌بندی با SVM انجام شده است؛ این طبقه‌بند زمانی که ورودی‌اش بردارهای ویژگی بهینه و استخراج‌شده باشد (مثلاً توسط VGG16)، عملکرد بسیار بهتری نسبت به استفاده مستقیم از تصاویر خام دارد.
 - انعطاف‌پذیری در ترکیب با مدل‌های کلاسیک: استخراج ویژگی امکان استفاده از قدرت مدل‌های عمیق را در کنار سادگی و سرعت طبقه‌بندهایی مثل SVM فراهم می‌کند؛ یک مزیت استراتژیک در شرایطی که منابع پردازشی محدودند.
- بنابراین، استخراج ویژگی به‌عنوان پلی مؤثر میان شبکه‌های عمیق و الگوریتم‌های سنتی طبقه‌بندی، نقش حیاتی در بهبود دقت و پایداری نتایج ایفا می‌کند و دقیقاً در پاسخ به این مسئله که "چگونه می‌توان با داده‌ی محدود به دقت بالا رسید؟" راه‌حل قدرتمندی ارائه می‌دهد.