|  |  |  |
| --- | --- | --- |
|  | **به نام خدا**  **دانشگاه تهران**  **دانشکده‌ مهندسی برق و کامپیوتر** |  |
| **درس شبکه‌های عصبی و یادگیری عمیق**  **تمرین دوم** | | |

فهرست

[پرسش2. پیاده‌سازی یک سیستم طبقه‌بندی خودرو 1](#_Toc203072139)

[1-2- پیش پردازش داده 1](#_Toc203072140)

[1-1-2- نمودار فراوانی داده‌ها و برچسپ عددی هر کلاس 1](#_Toc203072141)

[2-1-2- بالانس کردن داده 2](#_Toc203072142)

[3-1-2- گزارش ابعاد داده آموزش و تست 2](#_Toc203072143)

[2-2- آموزش و ارزیابی مدل 2](#_Toc203072144)

[1-2-2- نحوه ی عملکرد معیارهای Recall, Accuracy, F1Score, Precision 2](#_Toc203072145)

[2-2-2- گزارش معیارهای ارزیابی و نتایج VGG, AlexNet, SVM+VGG, CNN 3](#_Toc203072146)

[3-2-2- ماتریس درهم‌ریختگی VGG, AlexNet, SVM+VGG, CNN 4](#_Toc203072147)

[4-2-2- اصلی‌ترین تفاوت بین AlexNet و VGGNet 6](#_Toc203072148)

[3-2- تحلیل نتایج 7](#_Toc203072149)

[4-2- امتیازی 10](#_Toc203072150)

[1-4-2- مدل SVM+VGG با کرنل RBF 10](#_Toc203072151)

[2-4-2- تاثیر استخراج ویژگی بر نتیجه نهایی در طبقه بندی 11](#_Toc203072152)

**شکل‌ها**

[شکل 14. نمودار فراوانی دیتاست toyota 1](#_Toc203072128)

[شکل 15. ماتریس آشفتگی 4 مدل مختلف برای دیتاست toyota 5](#_Toc203072129)

[شکل 16. نمودار ارزیابی عملکرد مدل‌های مختلف روی دیتاست toyota 7](#_Toc203072130)

[شکل 17. ماتریس آشفتگی مدل SVM با کرنل غیرخطی RBF 11](#_Toc203072131)

**جدول‌ها**

[جدول 3. تقسیم‌بندی داده‌ها به دو مجموعه آموزشی و تستی به نسبت ۸۰:۲۰، شامل اندازه تصاویر و برچسب‌ها 2](#_Toc203072120)

[جدول 4. ارزیابی عملکرد مدل‌ها روی دیتاستtoyota 3](#_Toc203072121)

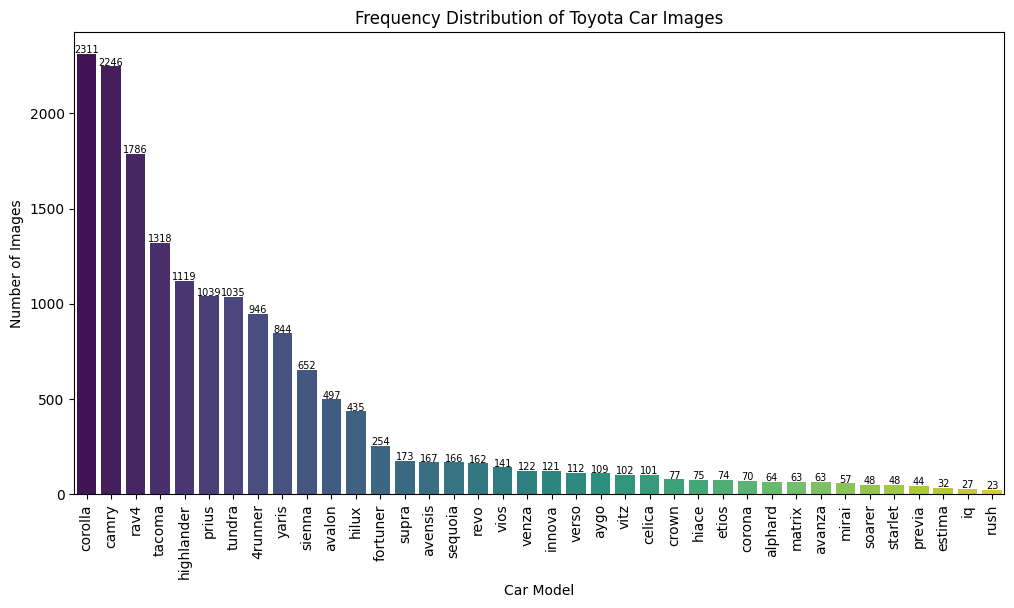
[جدول 5. ارزیابی عملکرد SVM با کرنل غیرخطی RBF روی دیتاستtoyota 10](#_Toc203072122)

# پرسش2. پیاده‌سازی یک سیستم طبقه‌بندی خودرو

**تمام کدهای این سوال و نمودارها و نتایج حاصل در فایل** VGG\_SVM.ipynb **واقع در پوشه** appendix **در مسیر اصلی فولدر موجود می‌باشد.**

## 1-2- پیش پردازش داده

### 1-1-2- نمودار فراوانی داده‌ها و برچسپ عددی هر کلاس



شکل 14. نمودار فراوانی دیتاست toyota

10 کلاس انتخاب شده توسط ما و لیبل و ایندکس هر کلاس. توجه شود که در ماتریس آشفتگی از لیبل ها 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 vitz 102 36

### 2-1-2- 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 500 نمونه) در داده‌ها برسیم و کلاس‌هایی با داده‌های فراوان‌تر را نیز فقط به اندازه 500 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 نه‌تنها 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.

### 3-1-2- گزارش ابعاد داده آموزش و تست

جدول 3. 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) | 4000 | (3، 224، 224، 4000) | (،4000) |
| تستی (Test) | 1000 | (1000،224،224،3) | (،1000) |

## 2-2- آموزش و ارزیابی مدل

### 1-2-2- 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.

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

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.

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.

### 2-2-2- گزارش معیارهای ارزیابی و نتایج VGG, AlexNet, SVM+VGG, CNN

جدول 4. ارزیابی عملکرد مدل‌ها روی دیتاستtoyota

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| مدل | Precision | Accuracy | F1 Score | Recall |
| SVM+VGG | 0.8406 | 0.8420 | 0.8407 | 0.8420 |
| VGG16 | 0.8401 | 0.8210 | 0.8150 | 0.8210 |
| AlexNet | 0.8083 | 0.8060 | 0.8038 | 0.8060 |
| CNN | 0.4927 | 0.5420 | 0.5114 | 0.5420 |

همانطور که مشهود است، مدل SVM+VGG با دستیابی به دقت 0.8420، 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 نشان‌دهنده‌ی تأثیر مثبت استخراج ویژگی‌های عمیق توسط VGG و سپس طبقه‌بندی با SVM است.

مدل VGG16 با دقت 0.8210 در رتبه دوم قرار دارد و نشان می‌دهد استفاده مستقیم از شبکه‌های پیش‌آموزش‌دیده نیز می‌تواند عملکرد مناسبی ارائه دهد.

مدل AlexNet با دقت 0.8060 عملکرد متوسطی داشته و F1 آن (0.8038) نشان می‌دهد که در حفظ تعادل میان Precision و Recall عملکرد نسبتاً پایداری داشته است.

در مقابل، مدل CNN ساده با دقت 0.5420 و مقدار پایین F1 (0.5114) عملکرد ضعیفی از خود نشان داده و بیانگر این است که معماری‌های ساده‌تر در مواجهه با داده‌های پیچیده‌تر مانند Toyota، کارایی کافی ندارند.

### 3-2-2- ماتریس درهم‌ریختگی VGG, AlexNet, SVM+VGG, CNN

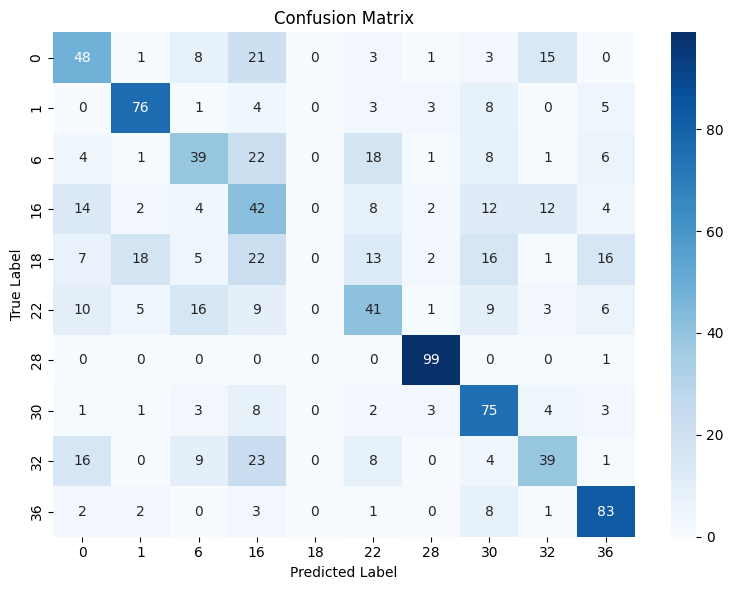
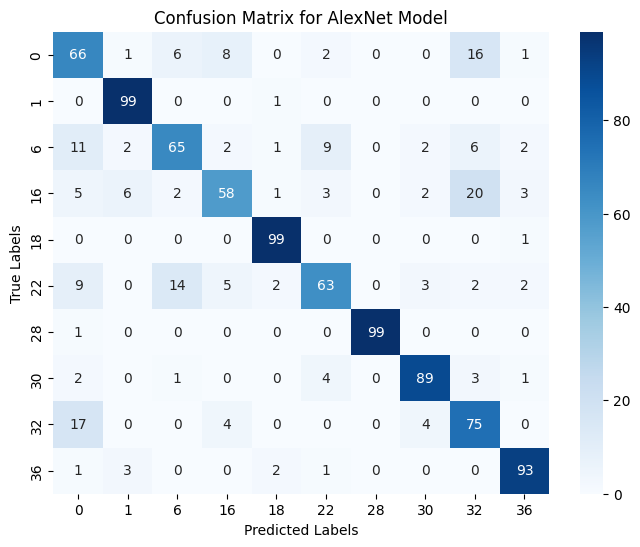
**سوال:** نحوه ی تحلیل مقادیر عددی ماتریس درهم‌ریختگی را توضیح دهید؟

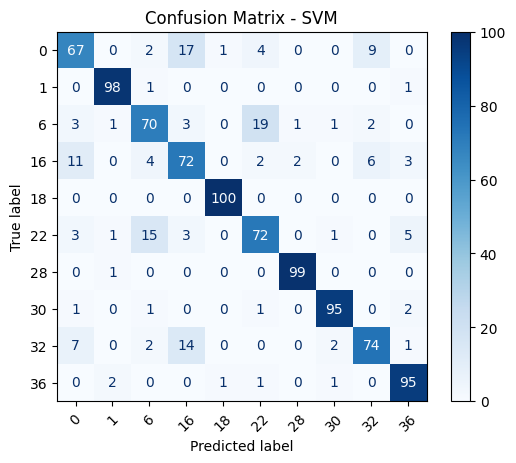
در یک ماتریس n×n، سطرها نمایانگر کلاس‌های واقعی (True Labels) و ستون‌ها نشان‌دهنده کلاس‌های پیش‌بینی‌شده (Predicted Labels) هستند.

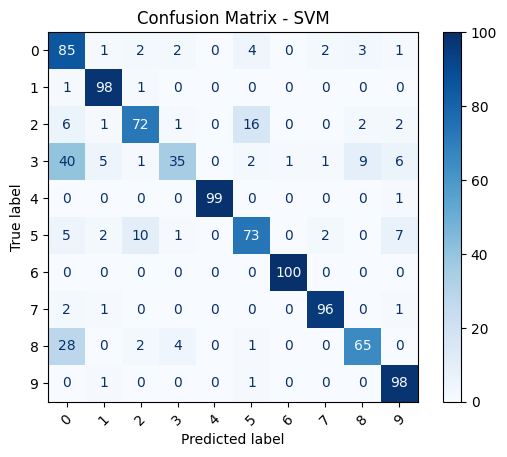
مقادیر قطر اصلی (Diagonal) بیانگر پیش‌بینی‌های درست (True Positives) برای هر کلاس‌اند، در حالی‌که مقادیر خارج از قطر به خطاهای مدل در طبقه‌بندی بین کلاس‌ها اشاره دارند.

اگر مقدار زیادی از نمونه‌ها در قطر اصلی قرار داشته باشند، نشان‌دهنده‌ی 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 سطرهای 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.



AlexNet Model CNN Mode



VGG Model Proposed (VGG + SVM) Model

شکل 15. ماتریس آشفتگی 4 مدل مختلف برای دیتاست toyota

**مدل ترکیبی VGG+SVM:**

این مدل با دقت بالا و کمترین نرخ اشتباه در طبقه‌بندی، عملکرد بسیار قابل‌قبولی 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.

### 4-2-2- 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 بهره برد، 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 غیرخطی‌بودن 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 درعین‌حال کارآمد، با لایه‌های نازک ولی زیاد، که به دقت و قدرت تعمیم بالایی در یادگیری ساختارهای پیچیده تصویر منتهی می‌شود.

## 3-2- تحلیل نتایج

شکل 16. نمودار ارزیابی عملکرد مدل‌های مختلف روی دیتاست toyota

تحلیل عملکرد مدل‌ها و کلاس‌های شناسایی‌شد با توجه به جدول 4 و ماتریس‌های آشفتگی شکل 15 و شکل 16:

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 (2246تصویر)، Tundra (1035 تصویر) و Runner (946 تصویر) نرخ تشخیص بسیار بالایی را ثبت کند. در مقابل، در کلاس‌هایی با داده‌های اندک نظیر IQ (27 تصویر)، Soarer (48 تصویر) و تا حدودی Supra (173 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 قابل‌توجهی 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 طورکلی مدل ها ارتقا پیدا کنند؟

افزایش تعداد نمونه‌ها به حداقل 1000 نمونه ممکن است تاثیر قابل توجهی بر افزایش عملکرد مدل‌ها داشته باشد که در این تمرین به دلیل محدودیت سخت‌افزاری از تعداد فقط 500 نمونه استفاده شده است.

1. بهبود کیفیت و تنوع داده‌ها

* داده‌افزایی هدفمند: اعمال تکنیک‌هایی مانند چرخش، بزرگ‌نمایی، تغییر روشنایی و نویز، به‌ویژه برای کلاس‌های کم‌نمونه، سبب افزایش تنوع داده‌ها و تقویت قابلیت تعمیم مدل می‌شود.
* متعادل‌سازی کلاس‌ها (Class Balancing): استفاده از روش‌هایی مانند Oversampling یا تولید داده‌ی مصنوعی برای کلاس‌های دارای نمونه‌ی کمتر، باعث کاهش bias مدل به سمت کلاس‌های پرتکرار می‌شود.
* پاک‌سازی داده‌ها: حذف تصاویر بی‌کیفیت، تار یا دارای برچسب اشتباه به بهبود آموزش مدل کمک می‌کند.

1. ارتقای معماری مدل‌ها و تنظیمات یادگیری

* استفاده از معماری‌های پیشرفته‌تر: به‌کارگیری مدل‌هایی مانند ResNet، DenseNet، EfficientNet یا Vision Transformer می‌تواند دقت مدل را در تشخیص ویژگی‌های پیچیده افزایش دهد.
* تنظیم مجدد مدل‌های پیش‌آموزش‌دیده (Fine-Tuning): تنظیم مجدد لایه‌های انتهایی مدل‌های مانند VGG16 یا ResNet می‌تواند تطبیق‌پذیری آن‌ها را با داده‌های خاص مجموعه افزایش دهد.
* تنظیم هایپرپارامترها: بهره‌گیری از روش‌هایی مانند Grid Search یا Random Search برای انتخاب بهینه‌ی پارامترهایی چون نرخ یادگیری، اندازه‌ی Batch و تعداد لایه‌ها، تأثیر چشم‌گیری در بهبود عملکرد دارد.
* استفاده از مدل‌های ترکیبی (Ensemble Learning): ترکیب خروجی چند مدل از طریق Voting یا Stacking می‌تواند دقت نهایی را افزایش داده و خطاها را کاهش دهد.

1. 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.

1. تحلیل دقیق‌تر خطاها و یادگیری از اشتباهات

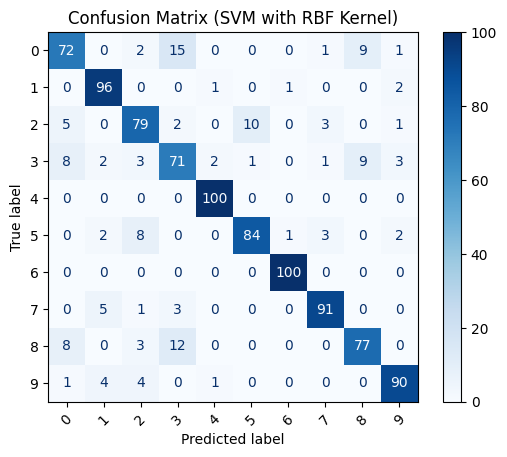
* بررسی دقیق ماتریس درهم‌ریختگی: شناسایی کلاس‌هایی با بیشترین خطای مدل و تحلیل ویژگی‌های آن‌ها می‌تواند به بازطراحی معماری یا تنظیمات کمک کند.
* آموزش بر اساس نمونه‌های دشوار (Hard Example Mining): تمرکز بر نمونه‌هایی که مدل بیشترین نرخ خطا را روی آن‌ها دارد، به افزایش قدرت تمایز مدل کمک می‌کند.

## 4-2- امتیازی

### 1-4-2- مدل SVM+VGG با کرنل RBF

جدول 5. ارزیابی عملکرد SVM با کرنل غیرخطی RBF روی دیتاستtoyota

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| مدل | Accuracy | Precision | Recall | F1 Score |
| SVM with RBF Kernel | 0.8600 | 0.8591 | 0.8600 | 0.8591 |



شکل 17. ماتریس آشفتگی مدل SVM با کرنل غیرخطی RBF

نتایج ارزیابی نشان می‌دهد که SVM با کرنل RBF عملکرد بهتری نسبت به نسخه خطی(جدول 4) از خود نشان داده و در ایجاد توازن میان Precision و Recall موفق‌تر عمل کرده است. کرنل RBF به‌دلیل ماهیت غیرخطی خود قادر است داده‌هایی با توزیع پیچیده و مرزهای نامنظم را به‌خوبی تفکیک کند؛ در حالی که SVM با کرنل خطی تنها در شرایطی که داده‌ها در فضای ویژگی به‌صورت 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 نگاشت داده‌ها به فضای ویژگی با ابعاد بالاتر، مرزهای تصمیم‌گیری دقیق‌تری ایجاد می‌کند. بنابراین، با توجه به نتایج حاصل‌شده، SVM با کرنل RBF می‌تواند انتخاب مناسب‌تری برای این مسئله محسوب می‌شود.

### 2-4-2- تاثیر استخراج ویژگی بر نتیجه نهایی در طبقه بندی

در مسائل طبقه‌بندی، به‌ویژه زمانی که داده‌های آموزشی محدود یا غیرموازنه (imbalanced) هستند، استخراج ویژگی به‌عنوان یک گام میانی مؤثر عمل می‌کند و می‌تواند به شکل قابل‌توجهی 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.