دانشگاه صنعتی خواجه نصیرالدین طوسی

به نام خدا

تشخیص و طبقه‌بندی علائم ترافیکی آلمان (GTSRB)

تحلیل عمیق و مقایسه‌ای شبکه های عصبی پرسپترون چندلایه (MLP)و CNN

دانشجویان: علی‌رضا دیوسالار(۴۰۲۱۷۴۹۳) و حامد بهروزی مقدم(۴۰۲۱۵۶۳۳)

استاد راهنما: آقای دکتر علیاری

پروژه نهایی درس هوش مصنوعی مقدماتی

دانشگاه صنعتی خواجه نصیرالدین طوسی

دانشگاه صنعتی خواجه نصیرالدین طوسی

پروژه تشخیص علائم ترافیکی آلمان (GTSRB)

خلاصه و ساختار کلی پروژه

۵برابر سبک‌تر

بهینه‌سازی پارامترها

CNNبا۳۵۶ هزار پارامتر در برابر MLP با ۱.۷۴میلیون پارامتر

٪۹۷.۲۶ دقت نهایی

عملکرد شبکه **CNN**

بر روی ۱۲,۶۳۰تصویر تست دست‌نخورده و دیده نشده

۴۳کلاس ترافیکی

مجموعه داده **GTSRB**

شامل بیش از ۵۰,۰۰۰تصویر رنگی با چالش‌های شدت نور و تاری

سرفصل‌های اصلی ارائه پروژه:

- فاز صفر : معرفی مجموعه داده **GTSRB** چالش‌های محیطی، عدم توازن کلاس‌ها و استراتژی تقسیم 80/20 Stratified
- فاز ۱ :پیاده‌سازی و ارزیابی شبکه **MLP** پیش‌پردازش (Flattening)، معماری ۳لایه، Callbacks، منحنی‌های یادگیری و ماتریس درهم‌ریختگی
- فاز ۲ : طراحی و پیاده‌سازی شبکه پیچشی **CNN** حفظ ساختار فضایی سه بعدی، معماری ۳بلوکی، کاهش شدید پارامترها و نتایج فوق‌العاده
- فاز ۳ : ارزیابی مستقل روی داده‌های **Unseen** و مقایسه نهایی تست مدل‌ها با Test.csv روی ۱۲,۶۳۰تصویر جدید و تحلیل مهندسی Inductive Bias

فاز صفر —معرفی مجموعه داده GTSRBو چالش‌های بینایی ماشین

تعداد نمونه‌ها	بخش داده‌ها
۳۹,۲۰۹ تصویر	مجموع کل داده آموزش
۳۱,۳۶۷ تصویر	مجموعه آموزش
۷,۸۴۲ تصویر	مجموعه اعتبارسنجی
۱۲,۶۳۰ تصویر	مجموعه تست دیده نشده (Test)

با استفاده از این استراتژی، نسبت تمام ۴۳کلاس ترافیکی در هر دو مجموعه Trainو Validationکاملاً یکسان حفظ گردید تا ارزیابی مدل کاملاً معتبر باشد.

ویژگی‌های اصلی مجموعه داده GTSRB:

✓**مشخصات داده‌ها:** شامل ۴۳کلاس متفاوت از علائم راهنمایی و رانندگی آلمان با بیش از ۵۰,۰۰۰تصویر رنگی RGB.

✓**تنوع شرایط محیطی:** تغییرات شدید نورپردازی (سایه‌های غلیظ، نور مستقیم آفتاب)، کنتراست پایین و تاریکی.

✓**تغییرات هندسی و فیزیکی:** تاری حرکتی (Motion Blur)، تغییر زاویه دید دوربین و آسیب‌دیدگی یا کثیفی تابلوها.

✓**عدم توازن کلاس‌ها (Class Imbalance):** تفاوت چشمگیر تعداد نمونه‌ها در کلاس‌های پرکاربرد (مانند محدودیت سرعت با ۲۰۰۰نمونه) در برابر کلاس‌های نادر (با ۲۰۰نمونه).

✓**تقسیم‌بندی Stratified (80/20):** تقسیم داده‌ها به ۳۱,۳۶۷نمونه آموزش (Train)و ۷,۸۴۲نمونه اعتبارسنجی (Validation)با حفظ دقیق تناسب کلاس‌ها.

فاز ۱ —پیش‌پردازش و معماری شبکه عصبی چندلایه (MLP)

لایه /نوع	شکل خروجی	تعداد پارامتر
Input Layer (Flatten)	(None, 3072)	۰
Dense 1 + Dropout(0.3)	(None, 512)	۱,۵۷۳,۳۷۶
Dense 2 + Dropout(0.3)	(None, 256)	۱۳۱,۳۲۸
Dense 3 (ReLU)	(None, 128)	۳۲,۸۹۶
Dense Output (Softmax)	(None, 43)	۵,۵۴۷

بیش از ۱.۵میلیون پارامتر فقط متعلق به اولین لایه متراکم است که نشان‌دهنده هزینه محاسباتی بالای شبکه MLP در پردازش تصاویر است.

مراحل پیش‌پردازش داده‌ها برای MLP:

تسطیح تصویر (Flattening)

تبدیل تصویر ۳۲×۳۲× ۳به یک بردار خطی یک‌بعدی با طول ۳۰۷۲پیکسل .(32*32*3 = 3072)

نرمالسازی (Normalization)

تقسیم مقادیر پیکسل‌ها [0, 255]بر ۲۵۵.۰جهت نگاشت به بازه [0, 1]برای سرعت در همگرایی گرادیان.

کدگذاری One-Hot Labels

تبدیل برچسب کدهای ۰تا ۴۲به بردارهای ۴۳تایی جهت سازگاری با تابع زیان Categorical Crossentropy.

فاز ۱ —استراتژی‌های کنترلی (Callbacks)و منحنی‌های آموزش MLP

استراتژی‌های کنترلی (Callbacks):

EarlyStopping

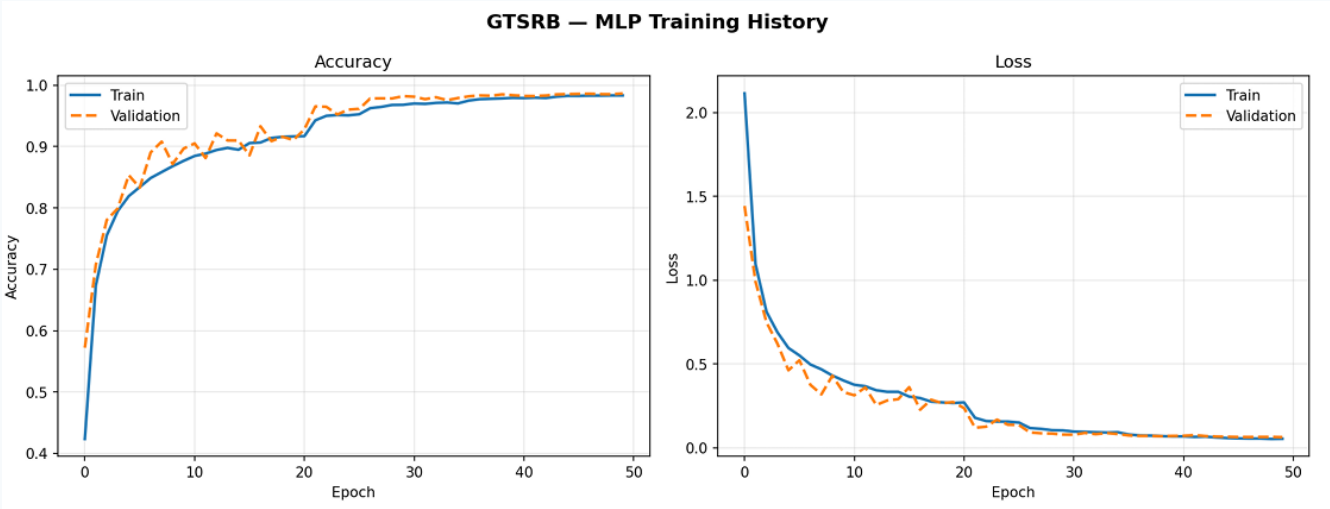
توقف آموزش در صورت عدم کاهش val_loss پس از ۱۰ اپوک متوالی (patience=10)برای جلوگیری از Overfitting.

ModelCheckpoint

ذخیره پیوسته بهترین وزن‌های مدل بر اساس کمترین val_loss در فایل best_mlp.keras.

ReduceLROnPlateau

کاهش نرخ یادگیری با فاکتور 0.5 در صورت عدم بهبود خطا طی ۳ اپوک برای خروج از مینیمم‌های محلی.



پروژه بازشناسی علائم ترافیکی آلمان (GTSRB)

فاز ۱ —ارزیابی عملکرد و تحلیل علل خطای بنیادی MLP

علت بنیادی خطاهای مدل MLP

نابودی ساختار فضایی (Spatial Structure)

عملیات Flattening و تبدیل تصویر ۲بعدی به بردار ۱بعدی موجب نابودی کامل همبستگی پیکسل‌های مجاور می‌شود.

اشتباه در کلاس‌های مشابه

شبکه MLP در تفکیک علائم دارای شباهت ظاهری (مانند تابلوهای محدودیتهای مختلف سرعت ۵۰ و ۸۰ km/h) دچار سردرگمی می‌شود.

واکنش به شدت نور منفرد

شبکه MLP اشکال هندسی (دایره، مثلث) یا لبه‌ها را درک نمی‌کند و صرفاً به شدت نور پیکسل‌های منفرد واکنش نشان می‌دهد.

نتایج ارزیابی MLP روی داده‌های Validation:

دقت اعتبارسنجی (Accuracy)

میزان دقت روی مجموعه اعتبارسنجی به ۹۸.۶۶٪ رسید.

شاخص‌های Classification Report

معیارهای Precision، Recall و F1-Score برای اکثر ۴۳ کلاس محاسبه گردید.

تحلیل ماتریس درهم‌ریختگی

تمرکز اصلی مقادیر روی قطر اصلی ماتریس است اما پراکندگی در خارج قطر وجود دارد.

فاز ۲ طراحی و پیاده‌سازی شبکه عصبی پیچشی (CNN)

بلوک /لایه	شکل خروجی	تعداد پارامتر
Block 1: Conv2D(32, 3x3) + Pool	(None, 14, 14, 32)	۸۹۶
Block 2: Conv2D(64, 3x3) + Pool	(None, 5, 5, 64)	۱۸,۴۹۶
Block 3: Conv2D(128, 3x3) + Pool	(None, 1, 1, 128)	۷۳,۸۵۶
Flatten + Dense(256) + Drop(0.5)	(None, 256)	۳۳,۰۲۴
Output Dense (43, Softmax)	(None, 43)	۱۱,۰۵۱

مراحل و ویژگی‌های معماری CNN

حفظ ساختار ۳بعدی ورودی:

پذیرش مستقیم ماتریس تصویر 32x32x3بدون تسطیح اولیه جهت استخراج ویژگی‌های فضایی.

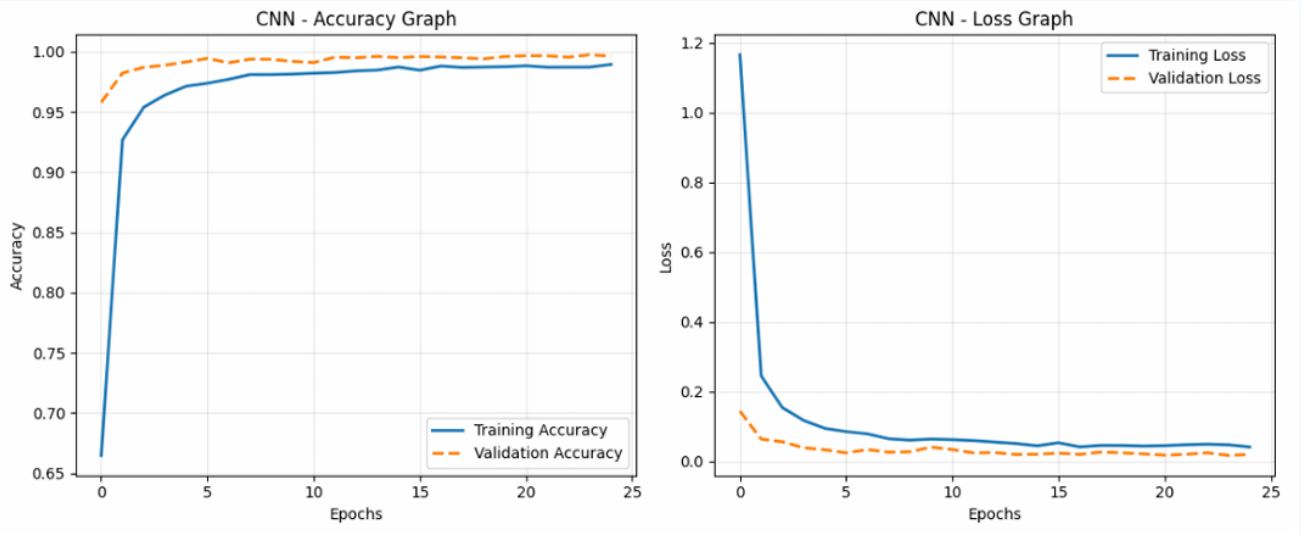
استخراج سلسله‌مراتبی ویژگی‌ها:

فیلترهای محلی ۳x ۳در لایه‌های ابتدایی لبه‌ها و رنگ‌ها را استخراج کرده و در لایه‌های عمیق‌تر اشکال و اعداد را بازشناسی می‌کنند.

ادغام با MaxPooling2D:

لایه‌های pooling 2x2 علاوه بر کاهش حجم محاسبات، مدل را نسبت به جابه‌جایی‌های کوچک تصویر مقاوم (Invariant) می‌کنند.

فاز ۲ —منحنی‌های یادگیری و همگرایی شبکه CNN



تحلیل رفتار آموزش شبکه CNN

همگرایی سریع و عالی:

منحنی‌های آموزش و اعتبارسنجی از اپوک‌های ابتدایی بسیار سریع صعود کرده و به دقت بالایی می‌رسند.

انطباق کامل منحنی‌ها:

چسبندگی کامل منحنی Validation به Train نشان‌دهنده عدم وجود بیش‌برازش (Overfitting) یا کم‌برازش (Underfitting) است.

پایداری نرخ خطا (Loss):

افت یکنواخت تابع زیان تا مقادیر بسیار نزدیک به صفر بدون نوسانات شدید.

پروژه بازشناسی علائم ترافیکی آلمان (GTSRB)

فاز ۲ —تحلیل عملکرد و ماتریس درهم‌ریختگی CNN

معیار ارزیابی	مدل MLP	مدل CNN
Accuracy	٪۹۸.۶۶	٪۹۹.۶۸
Loss	۰.۰۶۳۷	۰.۰۱۸۲
تعداد پارامترها	۱,۷۴۶,۷۳۱	۳۵۶,۹۳۹

نتایج Classification Report مدل CNN:

دقت نزدیک به ۱۰۰٪ (1.00)

شاخص‌های Precision، Recall و F1-Score برای قریب به اتفاق ۴۳ کلاس ترافیکی به عدد ۱.۰۰ رسیده‌اند.

حل چالش‌های نور و زاویه دید

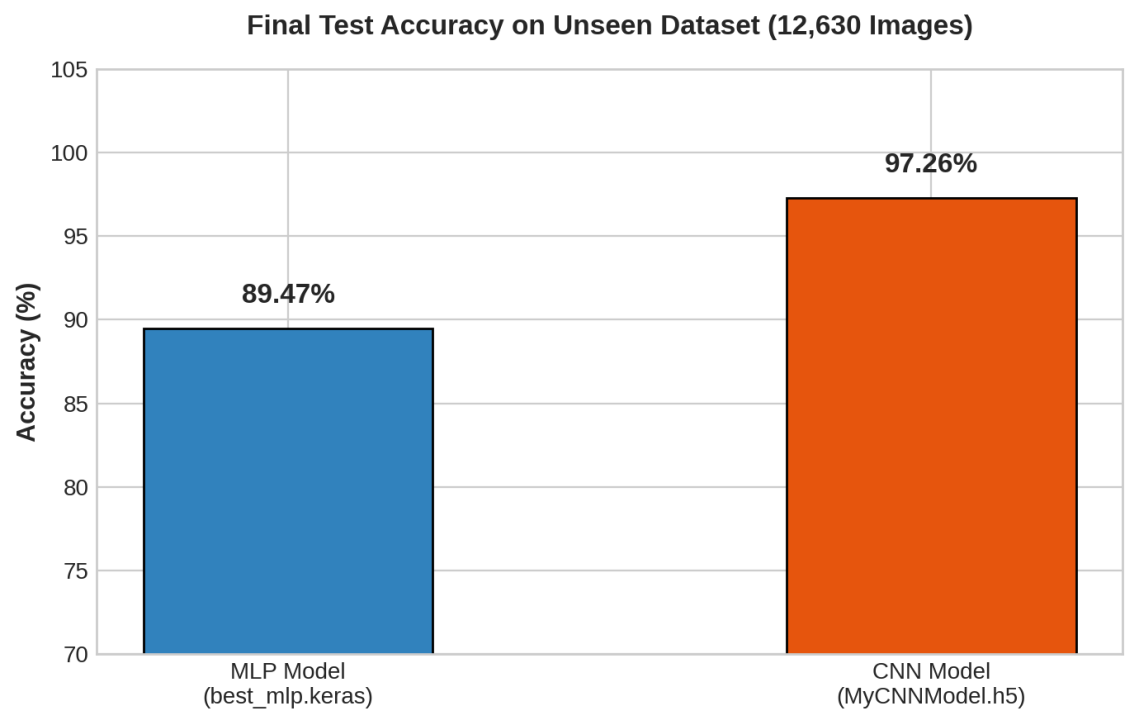
شبکه پیچشی به خوبی تفاوت بین تابلوهای مشابه با اعداد متفاوت را درک کرده است.

تمرکز بر قطر اصلی ماتریس

توزیع داده‌ها تماماً بر روی قطر اصلی ماتریس درهم‌ریختگی تمرکز دارد و خطای خارج از قطر به صفر نزدیک شده است.

فاز ۳ —ارزیابی مستقل روی داده‌های تست دیده نشده (Unseen Test)

```
==================================================
FINAL TEST ACCURACY ON UNSEEN DATA
==================================================
MLP Model ('best_mlp.keras'): 89.47%
CNN Model ('MyCNNModel.h5'):  97.26%
==================================================
```



اسکریپت ارزیابی نهایی (Test.csv)

ارزیابی واقعی تعمیم‌پذیری:

سنجش مدل‌ها بر روی ۱۲,۶۳۰ تصویر کاملاً جدید از فایل Test.csv که در هیچ یک از مراحل آموزش و تنظیم فراپارامترها نقشی نداشته‌اند.

فراخوانی مدل‌های ذخیره‌شده:

ارزیابی مستقیم فایل‌های best_mlp.keras و MyCNNModel.h5.

نتایج نهایی تست مستقل:

مدل MLP به دقت ۸۹.۴۷٪ رسید در حالی که مدل CNN دقت فوق‌العاده ۹۷.۲۶٪ را ثبت کرد.

پروژه بازشناسی علائم ترافیکی آلمان (GTSRB)

تحلیل مقایسه‌ای و دستاوردهای مهندسی پروژه

معیار مقایسه	شبکه عصبی MLP	شبکه عصبی پیچشی CNN	برتری مهندسی
تعداد کل پارامترها	(M۱.۷۴)۱,۷۴۶,۷۳۱	(K۳۵۶)۳۵۶,۹۳۹	CNNحدود ۵برابر سبک‌تر است
حفظ ساختار فضایی	خیر (D۱Flattening)	بله (ماتریس D RGB۳)	حفظ پیکسل‌های مجاور در CNN
دقت روی داده Unseen Test	٪۸۹.۴۷	٪۹۷.۲۶	افزایش ٪۷.۸ دقت نهایی
مقاومت در برابر جابه‌جایی	ضعیف	بسیار بالا (Pooling)	Invariantبه تغییر زاویه دید

نتایج این پروژه اثبات می‌کند که انتخاب معماری صحیح متناسب با ساختار داده (CNNبرای تصویر)، بسیار مهم‌تر از افزایش کورکورانه پارامترها یا عمق شبکه است.

تشکر از توجه شما