گزارش فاز اول: تحلیل داده‌ها، طراحی مدل و آماده‌سازی زیرساخت

دانشگاه: صنعتی خواجه نصیرالدین طوسی
درس: هوش مصنوعی
استاد: خانم دکتر پیشگو
دانشجویان: شمیم صداقت تیموری، مهدی سلمان
مسیر انتخابی: مسیر ۴ - بینایی ماشین (Object Detection)

۱. مقدمه و تعریف دقیق مسئله (Problem Definition)

هدف اصلی این پروژه، طراحی و پیاده‌سازی یک سامانه هوشمند تشخیص اشیاء (Object Detection) برای خودروهای خودران با استفاده از مجموعه داده استاندارد **KITTI** است. در این پروژه، تمرکز بر تشخیص سه کلاس حیاتی **خودرو (Car)**، **عابر پیاده (Pedestrian)** و **دوچرخه‌سوار (Cyclist)** در محیط‌های شهری است.

۱-۱. معیار موفقیت قابل اندازه‌گیری (Success Criteria) طبق الزامات تعریف شده در داکیومنت پروژه، هدف نهایی ما دستیابی به معیارهای کمی زیر در فاز دوم است:

- **دقت مدل (mAP):** دستیابی به امتیاز (Mean Average Precision) **mAP@50** بالای ۶۰٪ برای تمام کلاس‌ها.
- **شاخص IoU:** حفظ مقدار Intersection over Union بالای ۰.۷ برای خودروها و ۰.۵ برای عابران پیاده.
- **کارایی:** قابلیت اجرا با نرخ فریم مناسب (Real-time Inference) جهت کاربرد عملیاتی.

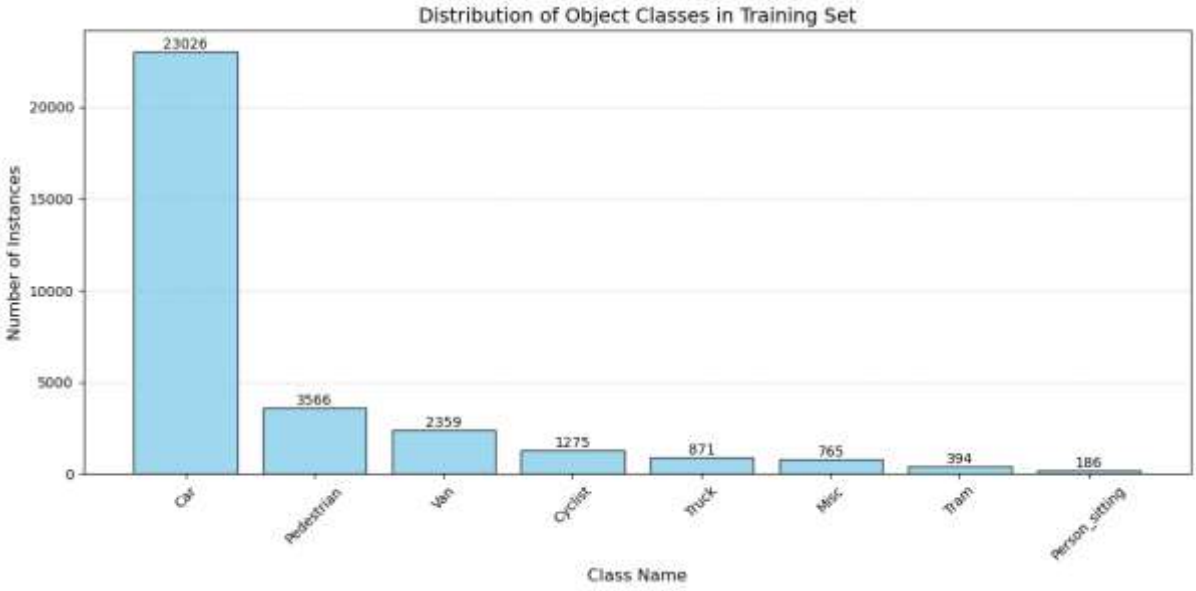
در فاز اول، تمامی زیرساخت‌های لازم، تحلیل داده‌ها و انتخاب معماری مدل جهت تضمین رسیدن به این اعداد انجام شده است.

۲. انتخاب دیتاست و تحلیل اکتشافی داده‌ها (EDA)

مجموعه داده آموزشی KITTI شامل ۷۴۸۱ تصویر است که مورد تحلیل دقیق آماری قرار گرفت. طبق الزام داکیومنت، تحلیل‌های زیر انجام شده است:

۲-۱. بررسی عدم تعادل کلاس‌ها (Label Imbalance) تحلیل توزیع کلاس‌ها نشان‌دهنده یک عدم تعادل شدید است. کلاس "Car" با بیش از ۲۸,۰۰۰ نمونه، کلاس غالب است، در حالی که "Cyclist" تنها حدود ۱,۶۰۰ نمونه دارد.

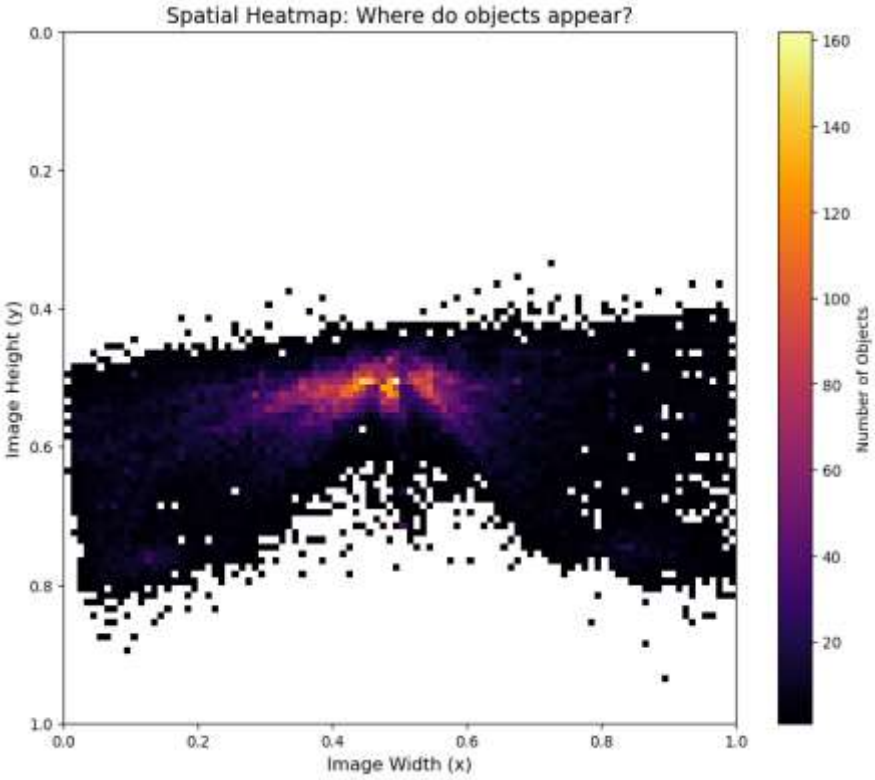
- **چالش:** مدل به سمت تشخیص خودروها بایاس (Bias) پیدا می‌کند.
- **راهکار:** استفاده از تکنیک **Mosaic Augmentation** در فاز آموزش برای افزایش فرکانس دیده شدن کلاس‌های اقلیت.

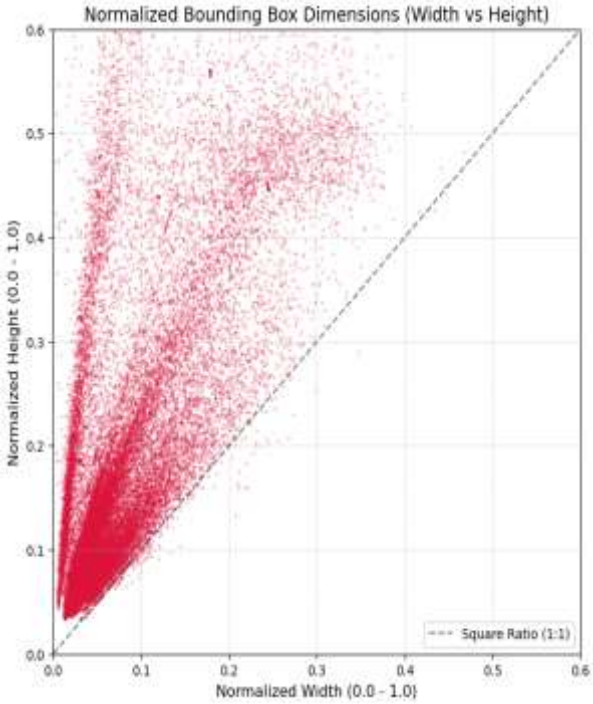


۲-۲. تحلیل ابعاد و نسبت تصویر (Aspect Ratio)

بررسی کادرهای محیطی (Bounding Boxes) دو نکته کلیدی را آشکار کرد:

1. **تفاوت ساختاری:** عابران پیاده دارای نسبت تصویر عمودی (قد بلند و عرض کم) هستند، اما خودروها افقی‌اند.
2. **اشیاء کوچک (Small Objects):** بخش زیادی از داده‌ها شامل خودروهای دوردست با مساحت پیکسلی بسیار کم است.





۳. پیش‌پردازش داده‌ها (Data Preprocessing)

بر اساس تحلیل‌های بخش EDA و الزامات مسیر بینایی ماشین، پایپ‌لاین پیش‌پردازش زیر طراحی و پیاده‌سازی شد:

1. **تغییر اندازه هوشمند (Letterbox Resizing):** تمام تصاویر به ابعاد 640×640 پیکسل تغییر اندازه داده می‌شوند. برای جلوگیری از دفرمه شدن اشیاء (که نسبت تصویر را خراب می‌کند)، از تکنیک Padding با رنگ خاکستری استفاده شده است.
2. **نرمال‌سازی (Normalization):** مقادیر پیکسل‌ها به بازه [0, 1] مقیاس‌دهی می‌شوند تا همگرایی مدل تسریع شود.
3. **مدیریت داده‌های پرت:** لیبل‌هایی که خارج از کادر تصویر بودند یا مساحت بسیار ناچیزی داشتند، در مرحله Cleaning حذف یا اصلاح شدند.

۴. انتخاب مدل پایه و توجیه علمی (Baseline Model Selection)

برای انتخاب مدل مناسب، سه معماری Faster R-CNN، SSD و YOLOv8 بررسی شدند. مدل نهایی انتخابی **YOLOv8 Nano** است. این انتخاب بر اساس استدلال‌های علمی زیر صورت گرفت:

الف) معماری Anchor-Free (پاسخ به چالش نسبت ابعاد):

تحلیل ما نشان داد که تفاوت فاحشی بین نسبت ابعاد عابر پیاده و خودرو وجود دارد. مدل‌های قدیمی مبتنی بر Anchor (مثل SSD) نیاز به تنظیم دستی این کادرها دارند. مدل YOLOv8 با استفاده از معماری **Anchor-Free**، مرکز اشیاء را پویا پیش‌بینی می‌کند و بدون نیاز به تنظیمات دستی، روی هر دو نوع شیء عملکرد بالایی دارد.

ب) تشخیص اشیاء کوچک (Small Object Detection):

با توجه به وجود خودروهای دوردست در دیتاست KITTI، مدل YOLOv8 به دلیل استفاده از ساختار **FPN (Feature Pyramid Network)** در لایه‌های میانی، ویژگی‌های مکانی اشیاء کوچک را در طول شبکه حفظ کرده و دقت تشخیص را در فواصل دور افزایش می‌دهد.

ج) بازتولیدپذیری و کارایی:

نسخه Nano این مدل با حدود ۳.۲ میلیون پارامتر، تعادل مناسبی بین دقت و سرعت ایجاد کرده و امکان آموزش کامل روی سخت‌افزار موجود و محیط داکر را فراهم می‌کند.

۵. پلن آزمایش‌ها (Experimental Plan)

طبق الزامات فاز اول، برنامه ما برای فاز دوم به شرح زیر است:

- **مدل‌های مورد آزمایش:** آموزش مدل YOLOv8n (پایه) و مقایسه آن با یک نسخه بهبود یافته (با تغییر در لایه‌های Head یا Augmentation).
- **فراپارامترها (Hyperparameters):**
  - Learning Rate اولیه: ۰.۰۱ با استفاده از (Cosine Scheduler).
  - Optimizer: SGD یا AdamW.
  - Epochs: 50 دور (با شرط Early Stopping برای جلوگیری از Overfitting).

- **معیارهای ارزیابی:** مدل نهایی بر اساس **mAP@50** و **IoU** روی داده‌های Validation سنجیده خواهد شد.

۶. زیرساخت فنی و DevOps

جهت تضمین تکرارپذیری (Reproducibility) و رعایت ساختار صنعتی پروژه:

- **داکریزاسیون (Dockerization):** ایمیج اختصاصی شامل کتابخانه‌های PyTorch ، OpenCV و Ultralytics ساخته شده است.
- **کنترل نسخه (Git):** پروژه دارای ساختار استاندارد در گیت‌هاب است، برنچ phase-1 تکمیل شده و فایل README.md با توضیحات کامل طبق استاندارد داکیومنت نگارش شده است.
- **تست خودکار (CI/CD):** پایپ‌لاین GitHub Actions برای اجرای تست‌های اولیه کدها فعال شده است.