



تمرین سری سیزدهم
درس تصویر پردازش رقمی

نام مدرس: دکتر محمدرضا محمدی
دستیار آموزشی مرتبط: محمد فرهمند
مهلت تحویل: ۱۴۰۱/۱۰/۱۷

۱- در جریان درس با مدل‌های زیر برای تشخیص اشیاء آشنا شده‌اید. این مدل‌ها را از منظر اجزای تشکیل‌دهنده، نحوه کارکرد، سرعت پردازش، دقت عملکرد و خاصیت چندمقیاسی (Multi-Scale) با یکدیگر مقایسه کنید.

- R-CNN
- Fast R-CNN
- Faster R-CNN
- RetinaNet
- YOLO
- SSD

۲- نوت‌بوک پیوست‌شده به این تمرین پیاده‌سازی مدل R-CNN است که با استفاده از TensorFlow 2 پیاده‌سازی شده است. قسمت‌های مشخص‌شده از نوت‌بوک مربوطه را مطابق توضیحات ارائه‌شده تکمیل کنید و مدل را آموزش دهید. برای راهنمایی شما در مورد کارکرد کد، خروجی بلوک‌ها پاک نشده است. دقت داشته باشید که عملکرد نهایی مدل شما نباید از مدل پایه ضعیف‌تر باشد اما بهبود عملکرد مدل با تنظیم آبرپارامترها یا تغییر اجزای مدل ۲۵٪ نمره امتیازی خواهد داشت.

۳- مقاله زیر مربوط به چهارمین نسخه از خانواده مدل‌های You Only Look Once یعنی مدل YOLOv4 است.

Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao.
"YOLOv4: Optimal speed and accuracy of object detection."
arXiv preprint arXiv:2004.10934 (2020).

الف) در این مقاله یک ساختار کلی برای شبکه‌های تشخیص اشیاء معرفی شده است که از سه بخش Neck, Backbone و Head تشکیل می‌شود. وظایف و وجه تمایز این سه بخش را کوتاه توضیح دهید و از معماری‌های معرفی‌شده برای هر بخش یک مثال بزنید.

ب) در این مقاله دو گروه از راهکارها برای بهبود عملکرد این خانواده از مدل‌ها معرفی شده‌اند که به Bag of Freebies و Bag of Specials شهرت دارند. وجه تمایز این دو گروه را شرح داده و راهکارهای ذکرشده را کوتاه و خلاصه توضیح دهید.

لطفاً سند قوانین انجام تمرین را ملاحظه و رعایت فرمایید.

موفق باشید.