

BT10 - Chuẩn hóa dữ liệu

Hoàng Tiến Dũng - MSSV: 19521388

1 Tại sao đối với giá trị của ảnh đầu vào, ta thường chuẩn hóa về 0-1 hoặc chuẩn hóa dạng zero-mean thay vì 0-255?

Trong thực tế đôi khi các điểm ảnh chênh lệch với nhau lớn, ví dụ có những điểm ảnh có giá trị 255 và có những điểm ảnh có giá trị 0 hoặc 1 hoặc 2 hoặc 3,...Như ta đã biết, các trọng số của mô hình được khởi tạo từ các giá trị ngẫu nhiên nhỏ và được cập nhật dựa trên lỗi dự đoán(loss) trong quá trình huấn luyện do đó nếu đầu vào không được chuẩn hóa có thể dẫn tới quá trình huấn luyện không ổn định. Với các giá trị điểm ảnh nằm được chuẩn hóa thì khi cho ảnh qua các mạng Deep Learning thì có thể tránh được trường hợp exploding gradient và giúp việc tính toán trở nên dễ dàng và nhanh hơn so với khoảng 0-255. Ngoài ra việc chuẩn hóa dữ liệu sẽ làm giảm một phần nhiễu không mong muốn nên mô hình sau huấn luyện sẽ có chất lượng tốt hơn.

2 Tại sao đối với giá trị bounding box đầu ra, ta thường chuẩn hóa theo 2 trục không gian về đoạn 0-1?

Với mỗi ảnh đầu vào ta có thể nhận được các ảnh với kích thước khác nhau. Việc chuẩn hóa dữ liệu giúp các bounding box tương đồng nhau dù cho độ phân giải của mỗi ảnh là khác nhau. Nếu không thực hiện chuẩn hóa thì khi thực hiện dự đoán thì ta sẽ nhận được các kết quả không mong muốn do độ phân giải của ảnh khác với ảnh khi huấn luyện. Lí do chuẩn hóa về đoạn 0-1 mà không phải đoạn khác là vì giúp tránh xảy ra exploding gradient.