Enhance the vehicle image segmentation performance of deep learning models in complex backgrounds

Technical University of Denmark

DTU Compute
Department of Applied Mathematics and Computer Science

Xuluo Cai s233485,Yuhang Pan s232282,Jiahe Liu s232276

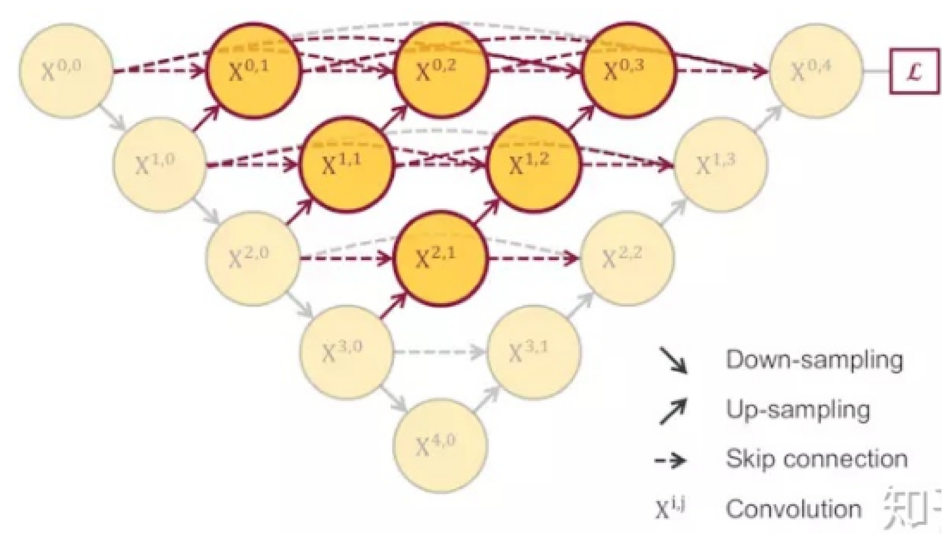
DTU Compute · Technical University of Denmark, Kgs. Lyngby, Denmark
DTU Compute · Technical University of Denmark, Kgs. Lyngby, Denmark
DTU Compute · Technical University of Denmark, Kgs. Lyngby, Denmark

1 Introduction

- Nowadays, cars have become an indispensable part of our daily lives, so rapid vehicle damage assessment after a car accident is increasingly important. Our project is to quickly cut vehicle pictures under the premise of complex background, so as to quickly identify different parts of different vehicles.
- UNet++ extends U-Net by adding densely connected decoders. The densely connected decoder consists of a series of U-Nets of different depths. Each U-Net in a densely connected decoder is connected to all other U-Nets in the decoder.

  Densely connected decoders allow UNet++ to learn between encoders to extract more complex features. This can improve the segmentation accuracy relationship This time we will also use UNet++ to cut and recognize vehicle photos..

2 Model architecture

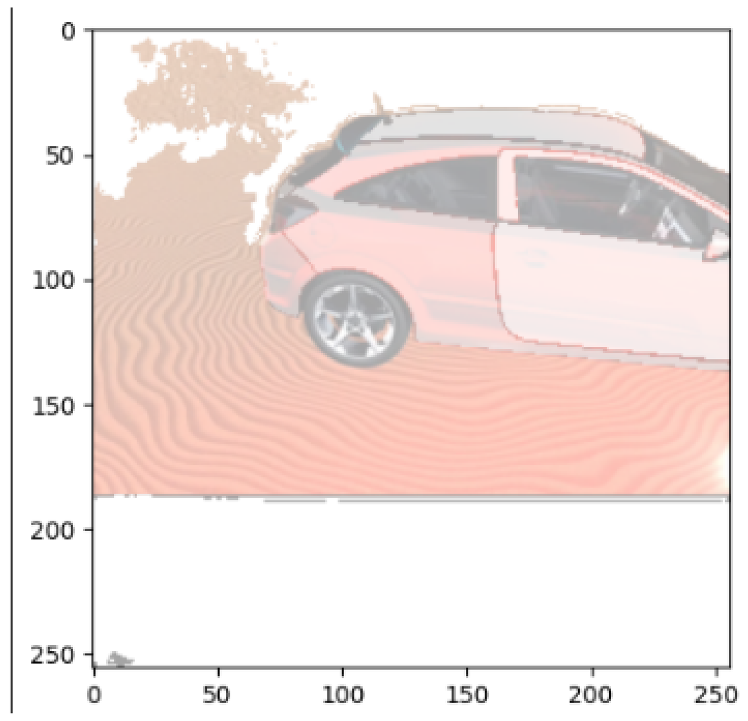


In this model, we chose the UNet++ architecture. The UNet++ model architecture consists of two main parts: encoder and decoder. The encoder is responsible for extracting features from the input image, and the decoder is responsible for reconstructing the output image from the features. The encoder consists of a series of convolutional layers, and the decoder consists of a series of upsampling layers.

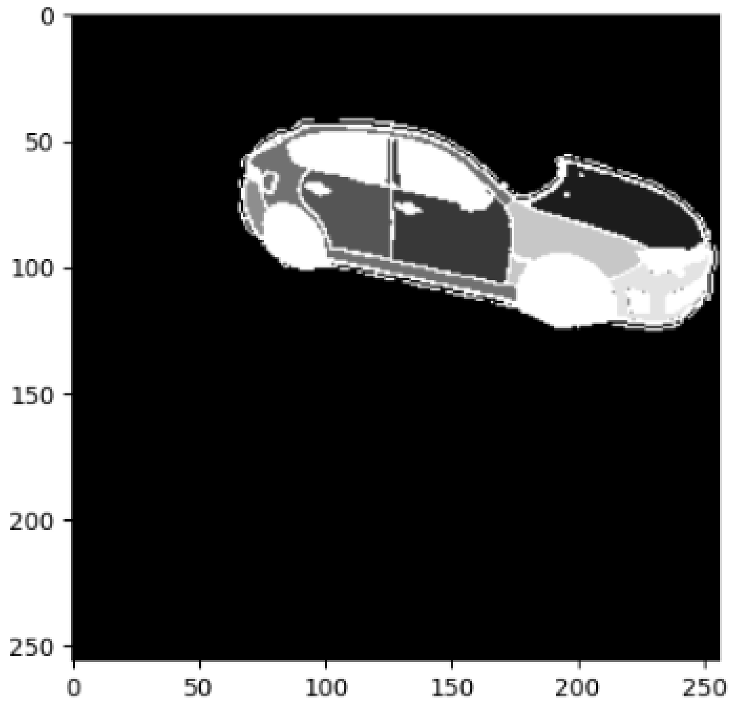
UNet++'s decoder is different from U-Net's decoder in that it uses a densely connected structure. A dense connection structure connects all U-Net modules in the decoder together. This allows UNet++ to learn relationships between features from different levels, thereby improving segmentation accuracy. Our project involves car image segmentation, so the advantages of UNet++ are a perfect fit for our project. UNet++ can provide higher segmentation accuracy, which is very important for our project. Furthermore, UNet++ can be adapted to our project by adapting the structure of the densely connected decoder. This means we can tune the model to our specific needs to get the best performance.

In our model, we set up four downsampling and upsampling layers separately, and at each step of the upsampling, the feature map is merged with the output of the corresponding downsampling layer using the concatenate function to improve the accuracy of the segmentation, and the softmax activation function is used so that the model generates a probability distribution for each pixel

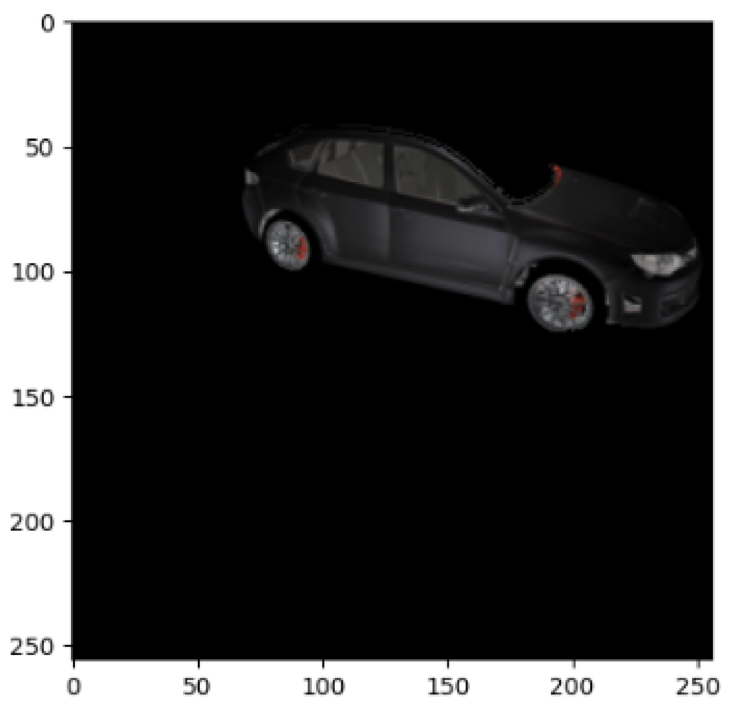
3 Data preprocessing



We first check and visualize image data stored in NumPy array format.



For each file, consider the first three channels as image data and the fourth channel as mask data. Process the mask data into different categories. Each grey level of the mask data (10, 20, ... , 90) is mapped to a category label (1, 2, ... , 9), transforming the image segmentation problem into a classification problem, where each pixel point needs to be assigned to a specific category to identify the location of each car part in the image.



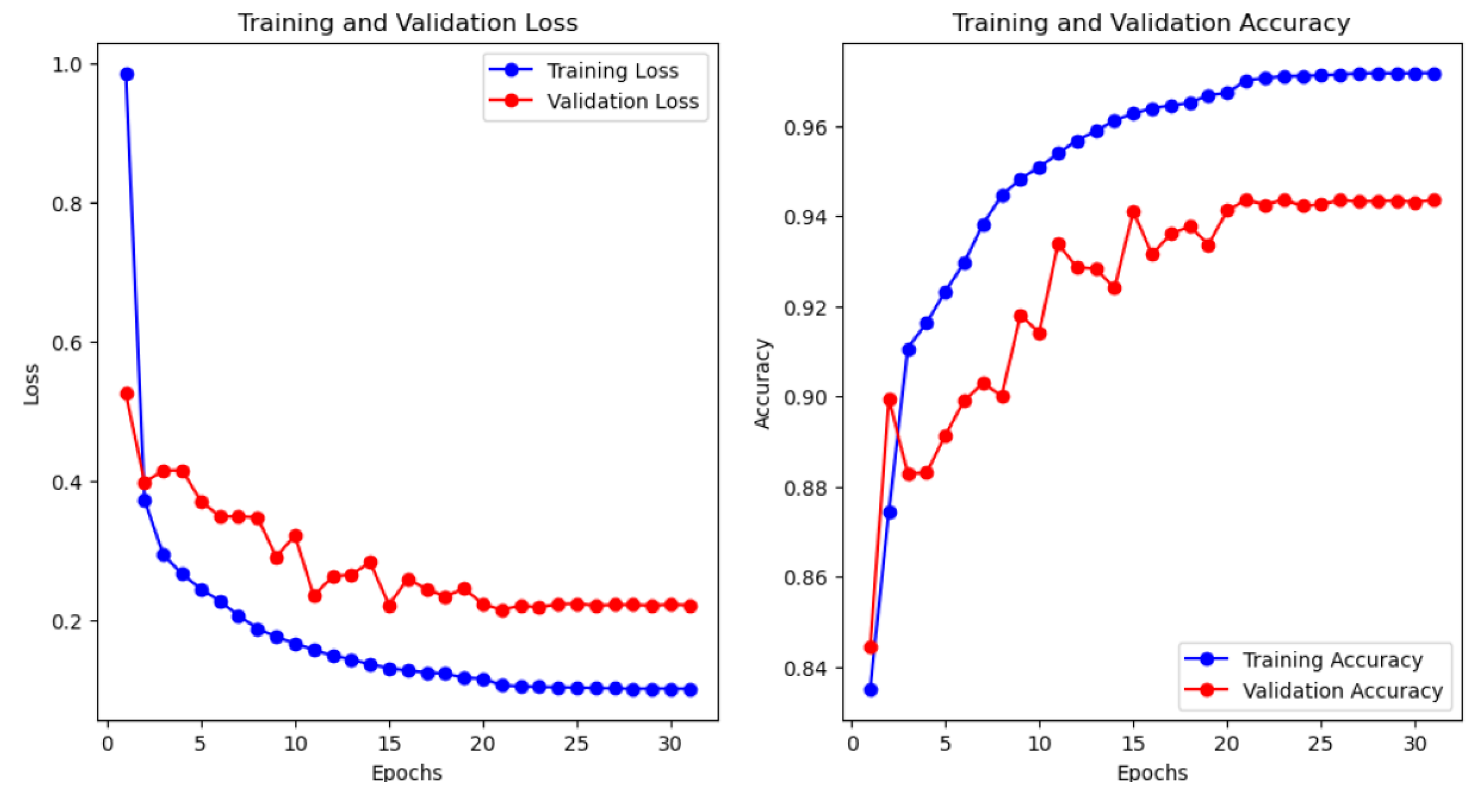
The image data is then normalised to between 0 and 1 to make it suitable for most deep learning models. And the masked data is converted to one-hot coding, (the mask for each pixel point is converted to a vector with a length of the number of categories (10))to calculate losses during training and interpret network predictions . The training and validation sets are then divided, with the validation set accounting for 20%.

Finally, we use the ImageDataGenerator class for data augmentation, which helps the model learn how to process images under a variety of different conditions.

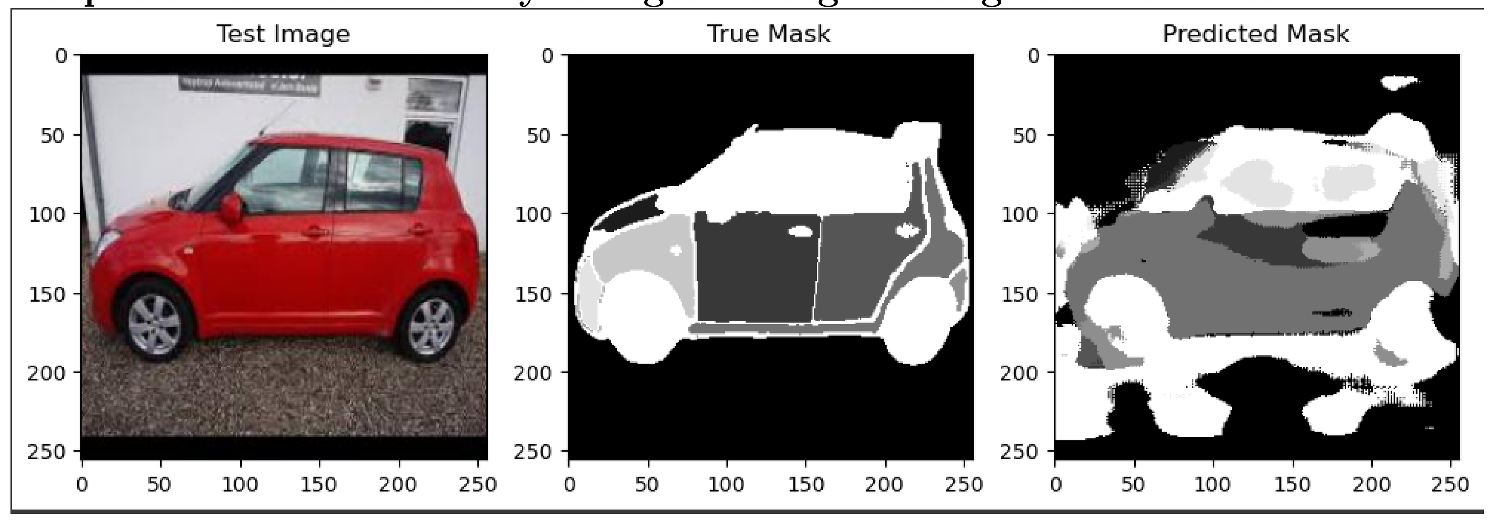
4 Performance and Result Discussion

4.1 Performance

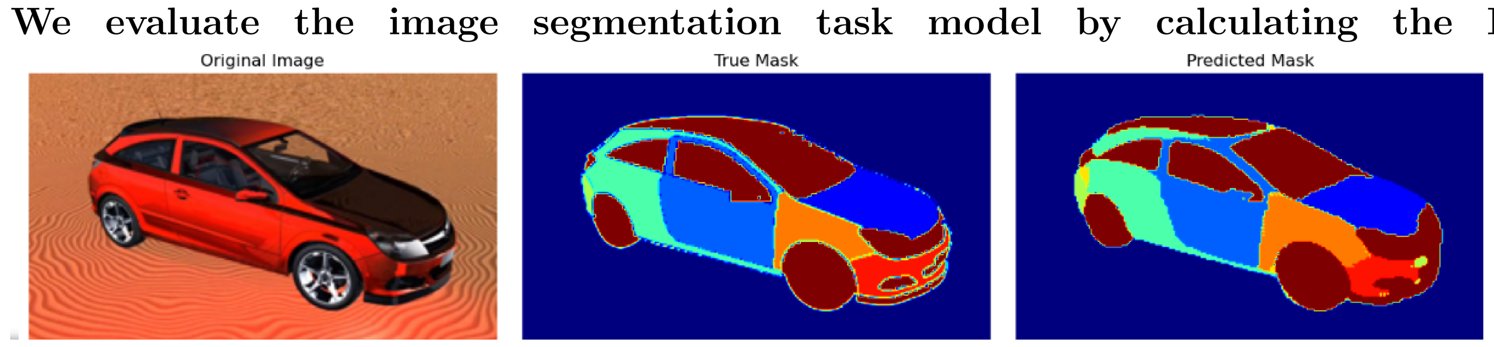
The model performs well in the training (0.9717) and validation set (0.9435), but performs poorly on the test set (IoU: 0.6678040397333116), and we consider that the model may be overfitting and that the model's generalisation is inadequate



Graphs of loss and accuracy changes during training and validation



We evaluate the image segmentation task model by calculating the IoU



Evaluation of the validation set

4.2 Result discussion

- IoU score (0.6678): This score indicates that the model has a moderate level of segmentation on this particular sample. While this is a reasonable score, there may still be room for improvement for applications that require high accuracy.
- We have made several hyperparameter adjustments and added drop layers to avoid overfitting, but the results are still unsatisfactory. In addition to overfitting, we are also considering the possibility that the model cannot be effectively generalised to the test data due to the uneven distribution of the dataset, as well as the lack of data augmentation. We are going to try the following methods to improve it: re-distribute the dataset, continue to adjust the hyperparameters, batch size, and change the optimiser to adjust the data enhancement.
- Direction of improvement: Based on the visualization results, the performance of the model on a specific area or a specific type of image can be further analyzed, and the model architecture or training data can be adjusted accordingly to improve the overall performance.

5 Conclusions & Outlook

During the development of the model, we found that the model performed 90on the training set during training, but it was lacking on the validation set. In the future we will further increase the number of samples in the training set. It may also be that the image background is too complex, and too much noise is still retained after preprocessing, affecting the accuracy of the model, so a more effective image data preprocessing will also be taken into consideration.

In future development directions, we hope to not only detect different parts of the car but also detect abnormal or damaged parts of the vehicle.