

Review Response for Paper ID 248

Dear Reviewers,

We would like to thank the reviewers for giving some insightful comments and suggestion which help us to improve our paper to be better. We have read the comments and suggestion carefully and we realized there are some points that we need to improve and revise. We have improved and revised the paper and address all the reviewer comments as following.

Comment Reviewer #1

Reviewer #1 has given some comments and suggestion about the part of the paper. We have addressed all the Reviewer #1 comments as following.

A. Abstract

Comment 1:

Introduction to the proposed field is not explained properly.

Response 1:

We would like to thank reviewer #1 for giving some comments about our introduction. We add the improvement of the introduction in the beginning sentence about the proposed field in the abstract.

Comment 2:

Author explicitly stated two limitations of predicting calories based on images. First, being the problem of occluded food and the lower R Squared value when predicting calories using linear regression model with only the area feature. What is the second limitation then? Whether that limitation is not considered in this work or out of scope for this work? Whether occluded food and lower R squared value are considered simultaneously or as two different problems. Please specify.

Response 2:

We would like to thank reviewer #1 for giving some question to explain more about the problem. We added the explanation about those two limitations. The second limitation is the low R Squared value in linear regression using area feature. So, the occluded food and low R Squared value is the different problem in our research. In this research, we do the research for the two limitations to improve the measurement.

Comment 3:

Reference in the abstract should be avoided.

Response 3:

We would like to thank reviewer #1 for the correction. We have removed the reference and the detail explanation about the method is in the paper.

Comment 4:

Data points for the linear regression model w.r.t, to food class is not mentioned.

Response 4:

We would like to thank reviewer #1 for the comments. We have added the number of data points in linear regression model in each class in the abstract.

Comment 5 :

How many total numbers of images for Indonesian street food dataset and images in individual classes are not mentioned.

Response 5 :

We would like to thank reviewer #1 for the comments. We have added the total number of images and total instance in individual classes.

Comment 6 :

Conclusion of the abstract is not clear.

Response 6 :

We would like to thank reviewer #1 for the comments. We have improved the wording of the conclusion in abstract.

B. Introduction**Comment 1 :**

Reference of WHO data is not mentioned.

Response 1 :

We would like to thank reviewer #1 for the comments. We have added reference from WHO official website.

Comment 2 :

Important reference is missing “He, K., Gkioxari, G., Dollár, P. and Girshick, R., 2017. Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).”

Response 2 :

We would like to thank reviewer #1 for the comments. We have added the sentence about the research that has been developed Mask R-CNN and give the reference about that in the second paragraph of introduction part.

C. Methods**Comment 1 :**

Figures are of very low quality.

Response 1 :

We would like to thank reviewer #1 for the corrections. We realized that the figure is not clear enough to understand so we have fixed the unclear figure in all parts.

Comment 2 :

No detail explanation given for hyperparameter tuning.

Response 2 :

We would like to thank reviewer #1 for the comments. We have improved our explanation in Mask R-CNN part in the point Methods. Also, we have detailed explanation about the hyperparameter model i.e. the backbone model in the experiment result.

D. Experimental Result

Comment 1 :

Details of the camera is not mentioned.

Response 1 :

We would like to thank reviewer #1 for the comments. We have included the detailed explanation about the camera device that we use to take the picture in Experimental Result point in sub point Datasets that has been mentioned in first paragraph.

Comment 2 :

Why the proposed model is not reducing the MAE score in some class? Please give reason

Response 2 :

We would like to thank reviewer #1 for giving us the question for our improvement. We have included the detailed reason about the cause of not reducing MAE in some class in Experimental Result point in sub point Final Model.

Comment Reviewer #2

Reviewer #2 has given some comments and suggestion about the paper. We have addressed all the Reviewer #2 comments as following.

Comment 1 :

The figs are not clear

Response 1 :

We would like to thank reviewer #1 for the corrections. We realized that the figure is not clear enough to understand so we have fixed the unclear figure in all parts.

Comment 2 :

The authors need to provide comparative analysis of other algorithms in order to show the effectiveness of the proposed one.

Response 2 :

We would like to thank Reviewer #2 for the suggestion. In this paper we have provided comparative analysis in linear regression with area as the independent variable that has been proposed by Chiang et al. and multiple linear regression with length and width as the independent variable that has been proposed by Abdelhady et al. We compare the linear regression and multiple linear regression model and assume if more features are used in the regression model, it will improve the performance. Also, we provided comparative analysis about the backbone model of Mask R-CNN using the ResNet and ResNeXt model, and we provided the Mask R-CNN model with FPN and without FPN (Backbone C4). Both of them have different architectures that have been explained in "He, K., Gkioxari, G., Dollár, P. and Girshick, R., 2017. Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).".