19-Oct-2025

RE: JCGS-25-535

Dear Dr Weihao Li:

Thank you for submitting the manuscript "Automated Assessment of Residual Plots with Computer Vision Models" for possible publication in the Journal of Computational and Graphical Statistics.

One referee, an Associate Editor (AE), and I have reviewed your paper. Although the topic is interesting, the referee raised several major concerns, which are detailed in their reports. The AE also indicated in the report that "the description of the work and the work itself require additional detail and rigor." As such, I cannot accept your paper in its current form. Nonetheless, given its potential, I would consider a revision that adequately addresses all the referees' comments.

Please also keep the following in mind for your revision:

1- Your revision should not increase the length of the manuscript.

2- I will download and run your code as part of the next round of reviews. I suggest, but don't require, the tidyverse style for your code and, in any case, ask that you submit code that is documented and as user-friendly as possible. Before resubmitting please run your code in a "clean" environment as issues such as required packages that aren't loaded or calls to files that don't exist are common and can delay publication (I'm not saying these exist in your case, just asking you to be aware!).

3- Your paper will be printed in black and white. The online version will be in color but to accommodate printing without color please be sure you don't refer to color in either the figure captions or text and that all figures are still interpretable in black and white.

4- Please be sure all figures have font size of labels and any text in the figure of at least 10pt.

Articles published in JCGS are ordinarily accompanied by online publication of the data and computer code used to illustrate the proposed methods. The goal is to enable readers to replicate much of the analysis and simulations described in a paper and to apply proposed methods in their own data without having to write code from scratch. When it is not practical to include a whole dataset in the manuscript, the manuscript should state how the complete dataset can be obtained. Please let me know if you would not be able to publish your data for reasons of security or confidentiality. Please provide the computer code and/or data that you hope to publish along with your revision. I have attached the instructions for including online materials here.

Any appendices should be removed and uploaded as a separate supplementary file.

Please bear in mind that your revision is due four months from today. A revision sent later than the deadline may be considered a new submission. When you provide your revision, please include a point-by-point response to each report in a single, blinded, attachment. When uploading your response file, please choose the "Response to reviewer comments" file designation from the drop-down menu. Please be sure to blind your response (if you chose double-blind review).

Please also have each author take the time to login to the system and use the Edit Account link (located in the upper right corner) to enter all of their detailed information (assuming they did not do so at some time previously).

Regarding color images: I encourage you to provide figures that can be printed in gray-scale but can be viewed electronically in color. If your paper is ultimately accepted and you wish your figures to be printed in color, we may need to ask your assistance to offset the additional cost of printing. Details and a payment form will be provided if needed. For an estimate please contact the production editor at Rebecca.Corpier@taylorandfrancis.com.

To submit a revision, go to https://aus01.safelinks.protection.outlook.com/?
url=https%3A%2F%2Frp.tandfonline.com%2Fsubmission%2Fflow%3FsubmissionId%3D257978497%26step%3D1&data=05%7C0
2%7Cpatrick.li%40anu.edu.au%7Cdab1bb3213dd48bcf73b08de0f7f8f58%7Ce37d725cab5c46249ae5f0533e486437%7C0%7C0
%7C638965237862292357%7CUnknown%7CTWFpbGZsb3d8eyJFbXB0eU1hcGkiOnRydWUsIlYiOiIwLjAuMDAwMCIsIlAiOiJXa
W4zMiIsIkFOIjoiTWFpbCIsIldUIjoyfQ%3D%3D%7C0%7C%7C%7C&sdata=jl3Jx5Klg%2Ba6fSEdHT%2BHVDVE7BXsIbfpqCfLuoh
0PS0%3D&reserved=0. If you decide to revise the work, please submit a list of changes or a rebuttal against each point which is being raised when you submit the revised manuscript.

If you have any questions or technical issues, please contact the journal's editorial office at jcgs.office@gmail.com.

This link will remain active until you have submitted your revised manuscript.

IMPORTANT: Your original files are available to you when you upload your revised manuscript. Please delete any redundant files before completing the submission.

Upon submitting your revision, you will be asked if the manuscript has been submitted to the journal previously. Click `Yes' and then cut-and-paste the manuscript number from the email you received.

Because we are trying to facilitate timely publication of manuscripts submitted to Journal of Computational and Graphical Statistics, your revised manuscript should be uploaded by 19-Apr-2026. If it is not possible for you to submit your revision by this date, we may have to consider your paper as a new submission.

Once again, thank you for submitting your manuscript to Journal of Computational and Graphical Statistics. I look forward to receiving your revision.

Sincerely,

Yuguo Chen

Editor, Journal of Computational and Graphical Statistics
yuguo@illinois.edu

Review Report for "Automated Assessment of Residual Plots with Computer Vision Models, JCGS-25-535"

October 2025

1 General Comments

The idea of the paper seems interesting: using computer vision to help identify residual plots which indicates model violation. For any computer vision task, the input is clearly the image. However, there should be a clear definition of the outcome variable that the algorithm wants to predict from images. From the current draft of this paper, it is unclear that what is the outcome variable that the computer vision is learning from the residual plots, how the training datasets are created for computer vision learning, and how the performance of the computer vision algorithm is assessed. Please see my detailed comments in items 10, 12,13. I suggest the authors make these basic setups explicitly defined from the very beginning of the article.

2 Specific Comments

1. p3, l9: What is "the lineup protocol"?

2. missing references: Loy and Hofmann (2013; 2014; 2015)

3. p8 l16: Why do we need to replace it by a full-rank covariance matrix?

4. eq (2), typically KL is specified by KL(p— q) due to asymmetry

5. p9 l 53: to solve eq 2: using "evaluate" for "solve" is better

6. p19 l30: What is "the data generating process"? We don't know the true distribution of y.

7. p12, sec 5.1: this sounds like a standard simulation of the sampling distribution of $\hat{D}$ in traditional statistics.

8. p13, Sec 5.2: How do you do bootstrapping? We need to know the distribution of $\hat{D}$ under the null. However, the given observed data y may not come from the null.

9. Tbl1: it is unclear what this table is measuring. What is the R^2 measuring? What's the response and what is the predictors?

10. Section 7 and 8: In computing $\hat{D}$, you need a P and Q for each targeted model violation. Is your computer vision learning algorithm targeted a particular model departure, eg, non-linearity or heteroskedasticity? What is your P and Q in generating the training data of $\hat{D}$ and residual plots for computer vision learning? However, your results also show your performance for different kinds of model violations. This is confusing. You need to specify clearly what is the "true" $\hat{D}$ and what the "predicted $\hat{D}$" using computer vision in generating training data of $\hat{D}$ and residual plots.

11. There is no numbering for your equations.