Advances in Artificial Intelligence for Data Visualization: Automate Reading of Diagnostic Plots with Compute Vision Models

Weihao (Patrick) Li
PhD student

Report for
Progress review

11 April 2023

MONASH BUSINESS SCHOOL

Department of Econometrics & Business Statistics

(04) 0459 1219
weihao.li@monash.com

ABN: 12 377 614 012

keywords: AI, data visualization, residual plot, visual inference, hypothesis testing, computer vision

1 Overview of the thesis

1.1 Background and motivation

Model diagnostics play a critical role in evaluating the accuracy and validity of a statistical model. They enable the assessment of the model's assumptions, detection of outliers, evaluation of how well the model fits the data, and identification of possible approaches to improve the model's performance.

When conducting model diagnostics, despite the availability of numeric summaries endorsed by finite or asymptotic properties, graphical representations of data are often preferred or required by data analysts. The preference for visual diagnostics is attributed to its intuitive nature and the possibility of discovering unexpected abstract and unquantifiable insights. In the context of regression diagnostics, a common practice is to plot residuals against fitted values, which serves as a starting point for evaluating the adequacy of the fit and verifying the underlying assumptions. Other visualization techniques such as histograms, Q-Q plots and box plots can be used to identify potential issues with assumptions made about the data, such as linearity, normality, or homoscedasticity.

Recently, a novel statistical inferential framework known as visual inference (Buja et al. 2009) has been developed, which relies on the use of graphical representations of data. The visual inference approach makes use of the natural capability of the human visual system to identify patterns and deviations from expected patterns. It provides a more comprehensible way of interpreting data and conducting hypothesis testing compared to conventional statistical testing.

Practically, visual inference is conducted via the lineup protocol. The protocol is inspired by the police lineup technique employed in eyewitness identification of criminal suspects. It comprises m randomly positioned plots, where one of them represents the data plot, while the remaining $m-1$ plots represent the null plots with the same graphical structure, except that the data has been replaced with data consistent with the null hypothesis H_0. An example lineup is provided in Figure 1. To compute the p-value of the visual test, the lineup will be independently presented to a number of participants, asking them to pick the most different plot. Under H_0, the data plot is expected to be indistinguishable from the null plots, and the probability of correctly identifying the data plot by an observer is $1/m$. If a large number of participants correctly identify the data plot, the corresponding p-value will be small, indicating strong evidence against H_0.

This method has gained increasing traction in recent years and has already been integrated into data analysis of various topics, such as diagnostics of hierarchical linear models (Loy and Hofmann 2013), geographical research (Widen et al. 2016) and forensic examinations (Krishnan and Hofmann 2021).

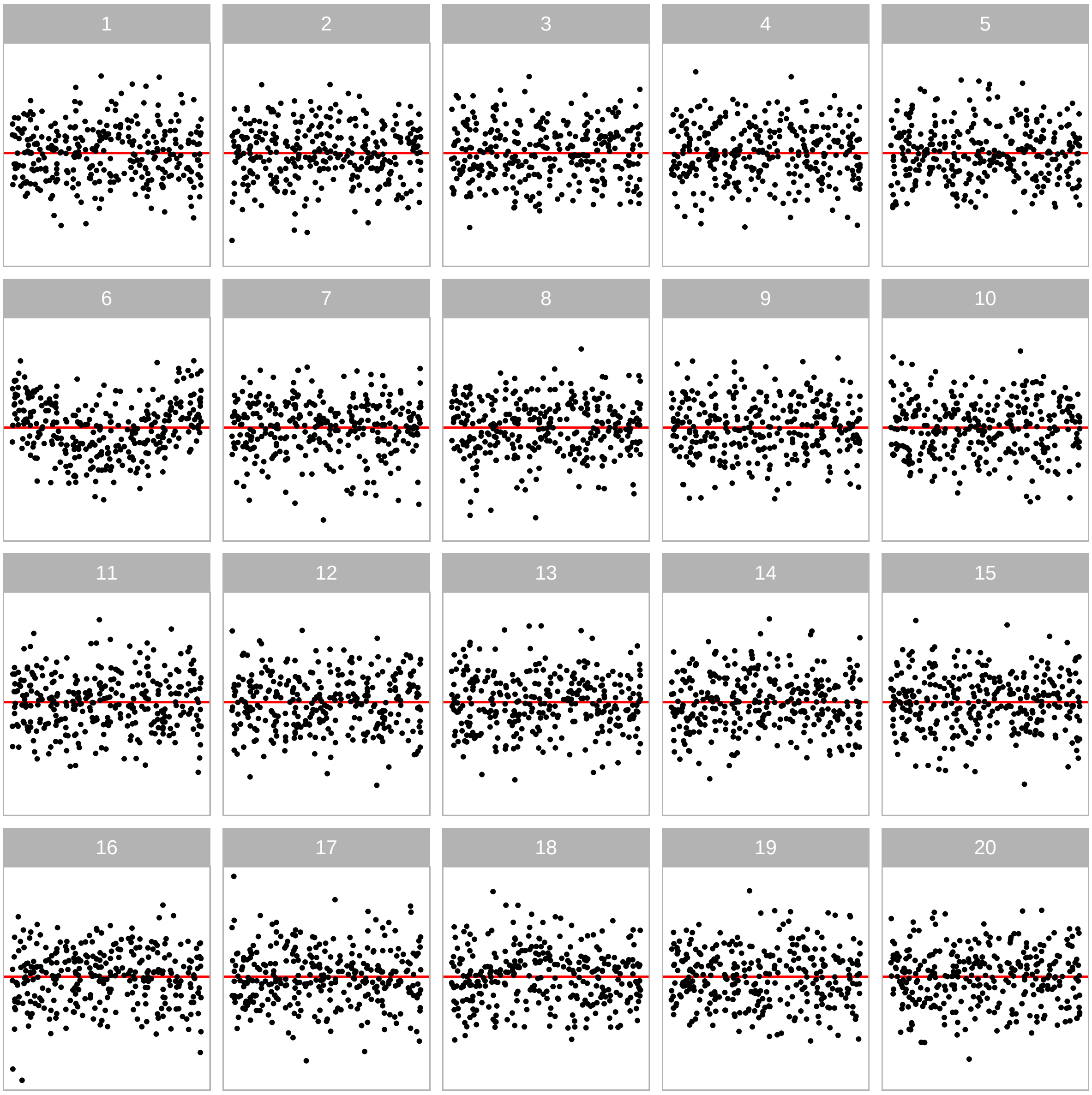

Figure 1: *Visual testing is conducted using a lineup, as in the example here. The residual plot computed from the observed data (plot $2^2 + 2$, exhibiting non-linearity) is embedded among 19 null plots, where the residuals are simulated from a standard error model. Computing the p-value requires that the lineup be examined by a number of human judges, each asked to select the most different plot. A small p-value would result from a substantial number selecting plot $2^2 + 2$.*

With the advent of sophisticated visualization techniques and tools, visual inference has the potential to provide an innovative alternative to traditional statistical approaches and enabling more effective communication of model findings.

The reliance of human assessment is a fundamental aspect of visual tests, but it may restrict its widespread usage. The lineup protocol is unsuitable for large-scale applications, due to its high labor costs and time requirements. Moreover, it presents significant usability issues for individuals with visual impairments, resulting in reduced accessibility.

Modern computer vision models offer a promising solution to this challenge. As a subfield of AI, computer vision with its modern deep learning architectures has successfully resolved numerous critical problems in automation. The development of the convolutional neural network (CNN) by Fukushima and Miyake (1982) was inspired by the vision processing in living organisms. The modern computer vision model typically utilizes deep neural networks with convolutional layers, which leverage the hierarchical pattern in data and provide regularized versions of fully-connected layers. This approach downscales and transforms images by summarizing information in a small space. Numerous studies have shown that it can effectively tackle vision tasks, such as image recognition (Rawat and Wang 2017), computer-aided diagnosis (Lee and Chen 2015), pedestrian detection (Brunetti et al. 2018), and facial recognition (Emami and Suciu 2012).

Utilizing computer vision models in reading data plot is not a common choice. Nevertheless, certain fields have adopted this idea by applying computer vision models in reading recurrence plots for time series regression (Ojeda, Solano, and Peramo 2020), time series classification (Chu et al. 2019; Hailesilassie 2019; Hatami, Gavet, and Debayle 2018; Zhang et al. 2020), anomaly detection (Chen, Su, and Yang 2020), and pairwise causality analysis (Singh et al. 2017). However, the assessment of lineups with computer vision models is a relatively novel area of study.

1.2 Research questions

The main objective of this research is to construct an automatic visual inference system capable of conducting visual tests on a large scale in the domain of model diagnostics, particularly regression diagnostics. The study will concentrate on three specific projects, namely

1. Exploring the application of visual inference in regression diagnostics and comparing it with conventional hypothesis tests.
2. Designing an automated visual inference system to assess lineups of residual plots of classical normal linear regression model.
3. Deploying the automatic visual inference system as an online application and publishing the relevant open-source software.

1.3 Current research outcomes

In order to examine the potential applicability of integrating visual inference techniques into regression diagnostics, a pilot study was conducted in the first year with the participation of 64 individuals, which was followed by a formal study involving 443 participants in the subsequent year. The participants were presented with lineups consisting of 20 plots, where a residual plot was embedded along with 19 null plots, drawn with data simulated using the residual rotation technique. The study considered two primary forms of residual departures in multiple linear regression, namely non-linearity and heteroskedasticity. To enrich the visual features, different fitted value distributions, including normal, uniform, lognormal, and discrete, were also incorporated.

The study revealed that conventional residual-based statistical tests are more sensitive to weak residual departures from model assumptions compared to visual tests evaluated by humans, leading to excessive rejections even when downstream analysis and outcomes would not be significantly impacted by the small departures from a good fit. Specifically, conventional tests conclude that the model fit has issues nearly twice as frequently as humans would, and they frequently reject departures in the form of non-linearity and heteroskedasticity that are not visible to humans. These findings emphasize the crucial role of graphical diagnostics and support the integration of visual inference into regression diagnostics. Moreover, they provide a compelling rationale for the development of an automated visual inference system to evaluate lineups of residual plots. The comprehensive details of the research methodology, along with other intriguing findings can be found in the attached paper.

2 Thesis structure

topic

chapter 1: Introduction

chapter 2 the first paper

chapter 3 build a computer vision system to evaluate residual plot

chapter 4? software publication

chapter 5 discussion and future work

3 Timetable

- April: submit abstract to ASC
- May: submit paper
- June: submit poster/short paper to IEEE vis conf, start working on computer vision model
- July: leave for few weeks
- Aug
- Sep: finalize the computer vision model
- Oct: IEEE vis conf
- Nov: web interface development

- Dec: ASC
- Jan, 2024:
- Feb, 2024:
- Mar, 2024: submit paper
- Aug, 2024: submit thesis

4 Difficulties

things that are actually quite serious

1. identified difficulties
2. suggestions for overcoming these difficulties

References

Brunetti, A, D Buongiorno, GF Trotta, and V Bevilacqua (2018). Computer vision and deep learning techniques for pedestrian detection and tracking: A survey. *Neurocomputing* **300**. Publisher: Elsevier, 17–33.

Buja, A, D Cook, H Hofmann, M Lawrence, EK Lee, DF Swayne, and H Wickham (2009). Statistical inference for exploratory data analysis and model diagnostics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **367**(1906), 4361–4383.

Chen, Y, S Su, and H Yang (2020). Convolutional neural network analysis of recurrence plots for anomaly detection. *International Journal of Bifurcation and Chaos* **30**(01). Publisher: World Scientific, 2050002.

Chu, H, X Liao, P Dong, Z Chen, X Zhao, and J Zou (2019). An automatic classification method of well testing plot based on convolutional neural network (CNN). *Energies* **12**(15). Publisher: Multidisciplinary Digital Publishing Institute, 2846.

Emami, S and VP Suciu (2012). Facial recognition using OpenCV. *Journal of Mobile, Embedded and Distributed Systems* **4**(1), 38–43.

Fukushima, K and S Miyake (1982). “Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition”. In: *Competition and cooperation in neural nets*. Springer, pp.267–285.

Hailesilassie, T (2019). Financial Market Prediction Using Recurrence Plot and Convolutional Neural Network.

Hatami, N, Y Gavet, and J Debayle (2018). Classification of time-series images using deep convolutional neural networks. In: *Tenth international conference on machine vision (ICMV 2017)*. Vol. 10696. International Society for Optics and Photonics, pp.106960Y.

Krishnan, G and H Hofmann (2021). Hierarchical Decision Ensembles- An inferential framework for uncertain Human-AI collaboration in forensic examinations. *arXiv:2111.01131 [cs, stat]*.

Lee, H and YPP Chen (2015). Image based computer aided diagnosis system for cancer detection. *Expert Systems with Applications* **42**(12), 5356–5365.

Loy, A and H Hofmann (2013). Diagnostic tools for hierarchical linear models. *Wiley Interdisciplinary Reviews: Computational Statistics* **5**(1), 48–61.

Ojeda, SAA, GA Solano, and EC Peramo (2020). Multivariate time series imaging for short-term precipitation forecasting using convolutional neural networks. In: *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*. IEEE, pp.33–38.

Rawat, W and Z Wang (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural computation* **29**(9), 2352–2449.

Singh, K, G Gupta, L Vig, G Shroff, and P Agarwal (2017). Deep convolutional neural networks for pairwise causality. *arXiv preprint arXiv:1701.00597*.

Widen, HM, JB Elsner, S Pau, and CK Uejio (2016). Graphical Inference in Geographical Research. en. *Geographical Analysis* **48**(2), 115–131.

Zhang, Y, Y Hou, S Zhou, and K Ouyang (2020). Encoding time series as multi-scale signed recurrence plots for classification using fully convolutional networks. *Sensors* **20**(14). Publisher: Multidisciplinary Digital Publishing Institute, 3818.