The authors present a non-invasive method to extract iron contamination in silicon using the extracted ideality factor and neural networks.

There are existing experimental methods to extract iron contamination from wafers, however, the iron contamination can be different in solar cells due to the different fabrication processes. To fill this gap, the authors have used simulated data and deep neural network algorithms to develop their method.

Although the research question is clear, I do not think the paper meets the minimum requirements for publication. I will strongly encourage the authors to improve the study and resubmit the paper.

My main comments are as follows:

1. I think that a MUST requirement for publication is a demonstration of the method using experimental measurements. This can be easily done to confirm the method.
2. The method assumes only one recombination channel that is related to iron; however, this is often not the case. Can you discuss the more practical case of a few recombination channels? How this impacts your method?
3. The simulations do not consider edge recombination although it can have a strong impact on the ideality factor at low injection. Can you comment?
4. “The simulated IVCs were fitted by using double diode model [37] equation with neglecting of both series and shunt resistances”: Why neglecting? How this can be done in practice? IV measurements are often impacted by both.
5. Can you explain “Although the ideality factor n = 2 is often used to describe the trap related recombination”? n=2 is often used for edge recombination and recombination within the space charge region.
6. “A simple back surface field (BSF) n+–p–p+ structure is important from an applied point of view”: A BSF is not a relevant structure anymore. Can you use more relevant structures such as PERC or TOPCon?
7. Can you clarify regarding the temperature dependency of Auger recombination? I do not think Ref [30] provides such a dependency.
8. “We have assumed uniform iron atom distribution in both SC base (p–region) and BSF–layer (p+–region) with concentration NFe”: Is it a valid assumption?
9. It seems a constant n is used (see Fig 2), despite the injection-dependency of this parameter. Can you please comment?
10. Equation 3 (MSRE) seems to be wrong. Is it used to evaluate the model and present the results? Please note that Equation 3 is an error metric for model evaluation, not a loss function.
11. Figure 4:
    - Full of details, however, only a basic discussion. Please improve the figure and more importantly, the discussion.
    - What is the meaning of the blue line? The model should be evaluated on 100% of test datasets rather than evaluating its performance on parts of the data sets.
12. The test data is generated by either varying individual variables or by varying all the variables at the same time. Practically, any of these combinations are possible, therefore, the authors should combine all the cases to evaluate their model.
13. The details of the architecture are missing: loss function, activation function, optimizer, etc.
14. The authors should also show other error metrics: R-square, root mean square, etc. for model evaluation.

15. A supplementary material is mentioned (line 54 of page 2, second column), however, it has not been provided.

More comments:

1. The first paragraph (Introduction), please add references to support your claims.
2. IVC – is not commonly used. It is worth using common-used symbols as IV
3. Figure 2: nodes, marks, color for training, and test data are unclear and are hard to understand. Perhaps, the author should look for another way to show the variables.
4. The number of test samples is not shown for each case.
5. Add detailed information in Fig. 1 caption.
6. It is strongly recommended that the English will be reviewed by a professional editor before submitting it in the future.