Response to Comments

Dear Editor/Reviewer,

Thank you very much for your letter and the comments from the reviewers about our paper submitted to MDPI-electronics(Manuscript ID:electronics-2209383; Title: A Deep Ensemble Learning Based approach for Skin Cancer Diagnosis).

We gratefully appreciate the constructive and thorough comments from the reviewers that helped us in clarifying and improving the quality of our research work. After carefully going through reviewers' comments, we have revised and improved the paper according to reviews as shown below (the **BLUE** comments are from the reviewers and each comment is followed by a response from the authors). We submit hereby the revised manuscript.

If you have any question about this paper, please don't hesitate to let me know.

Sincerely Yours,

Engr. Ijaz Ahmad

Institute of Computer Science and Information Technology,
Agriculture University, Peshawar, Pakistan
Email: engr.ijazahmad@aup.edu.pk

Reviewer 1:

Question1. A language check is required.

RESPONSE:
We thank the reviewer for pointing out this problem. To make sure that we correct all language mistakes and grammar glitches in the current revision, we have carefully proofread the manuscript several times and significantly improve the language check.

Question 2. There are no references to some figures in the text (some are incorrectly numbered).

RESPONSE:
Thanks for pointing this issue. Thus, in our current revision, we have added the missing references of figures and also corrected the wrongly labeled figures.

Question 3. In Fig. 11, however, it looks as if the base-line model is better focused on the target part (column B and E).

RESPONSE:
We thank the reviewer for the comment. In most of the images in Figure 11, the base-line model focuses only on a portion of the disease images, whereas the proposed ensemble model considers the complete disease portion of the entire image. Thus, in the revised version, we have provided the details about the focus of base-line model and the proposed model. To address this comment, we rephrased and rewrote the description as follows.

After the results, it is clearly shown in Fig 10 that the proposed hybrid model more actively focuses on the target part of the images rather than base-line EfficientNetV2S and proved that transformer blocks have a greater capacity to see its decision-making process.

Note: Numbering of Fig. 11 has been changed as Fig. 10.

Question 4. There's no need to present the idea of the confusion matrix (Fig. 7).

RESPONSE:
Thanks for the comment. We agree with the reviewer that the idea of confusion matrix is not required. Thus, in the revised version, we have removed the figure of confusion matrix (i.e., Fig. 7).