**Before answering your comments related to the children vs adults paper, I would like to clarify few things.**

* There were many ambiguities in the paper even with the contribution of the paper and in the central theme of the paper. For instance, your paper title is a comparison of facial expression recognition in children vs adults. The title reflects that you will compare the behavior of classifying the facial expression in adults vs children. But in the introduction section you say, “*We explore why children’s emotion detection is harder than that of adults and present a comparative analysis using the VGGFace2 model on two widely used datasets: CAFE and CK+”* and the reason why, supported by your results was completely missing.
* Moreover, the state-of-the-art studies showed a 99.50% accuracy score of the CAFÉ dataset in Table 1 but your accuracy score of the VGGFace2 model on the CAFÉ dataset was 94% in Table 2. How can you assume that the paper will publish with a 94% accuracy score while the 99.50% accuracy is already reported on the same dataset (CAFÉ in Table 1)?
* Lastly, the introduction section of your paper was based on just one paragraph. After this, you start the contribution, and the rest of the introduction section was discussing your contribution. The background and introduction of the paper were not satisfactory in this way. Secondly, our contribution is not too big or novel to take the complete space of the introduction section. Right?

To overcome all of these, firstly I remove the state-of-the-art table and add the ablation study to compare the proposed model with VGG, VGG16, VGG-19 and VGGFace2. Secondly, I write the related work in a broader aspect rather than writing it specifically for adults and children. If you are willing to still add the state-of-the-art table and related work as in the previous version, let me know I will add it.

Now I am answering your comments, read them carefully and write your final decision.

**Comment:** **1st point: the paper is about proving that children’s expressions are hard to recognize when compared to adults. The related section should be addressing the work done on this!**

**Ans:** No, the paper is not about proving that expression detection is harder in children compared to adults. The proposed work is simply about a comparison of detecting facial expressions in children and adults. However, experiments and results showed that the recognition of FE in children is harder than the recognition of FE in adults.

If we discuss this point in the related section, then we must need to discuss the papers published in 2023 with a 99.50% accuracy score (Table 1) and our paper must have a similar or greater accuracy score (the decision is yours)

**Comment: 2nd point: state of the art for ck+ is 99% and the result I got is the same! Without comparing to the state of the art, any model cant be judged**

**Ans:** You are right, the model is judged with the state-of-the-art model reported in the literature on a similar or another dataset. But for the CK+ dataset, your accuracy score is similar to the score reported in the literature which can be claimed as a contribution with robustness analysis. But the accuracy score of your model with CAFÉ is 97% while the reported is 99.50%. We cannot claim it as a contribution in any sense.

Comment: 3rd point the paper I sent before had the headings I was talking about. Please take a look again!

**Ans:** The heading in the previous section was completely irrelevant. I added the relevant heading now. When you take the decision about the accuracy score for the CAFÉ model, I will add the literature about the children’s FE in these headings.

**Comment: 4th point: comparison table is there in the paper that I gave. Table 1 is compared with state-of-the-art.**

**Ans:** see previous answers.

**Comment: Are we proposing an ensemble model?**

**Ans:** Yes, we proposed the deep ensemble model which is based on two variants of VGG (VGG-19 and VGGFace2)

**Comment: If we are, why are we putting the results of vggface2 instead of the ensemble model?**

**Ans:** We perform the ablation study to show the readers that the VGG-19 and VGGFace2 models did not perform well individually while our deep ensemble approach with customized dense layers achieved these results.