# Response to Reviewer 2 Comments

**Point 1:** The sentence "In this paper, a combination of SOTA model such as DenseNet, 8 InceptionNet, ResNet, NasNet, and MobileNet and Soft-Attention is proposed" in the abstract is not correct. You didn't use combination of all them in a specific unique structure. Different combinations of these networks are used in your proposed approach. Discuss about it or correct this sentence.

**Response 1: We agree with your comment about the structure of the above sentence and fix it.**

**Point 2:** All of the equations should be numbered.

**Response 2: We have already labeled and numbered all of the equations.**

**Point 3:** How do you propose the weights formula? (Section 2.2.5). Is there any related reference? How much is the size of output W in this equation?

**Response 3: We have already added the citation to the paper that inspire the W formula.**

**Point 4:** As I know, the input size of the used CNNs such as mobileNetV2, mobileNetv3, ResNet, etc are not same. Do you resize all of the images to the same size to start process? Or do you run each CNN with different input size?

**Response 4: We have already indicated how image is preprocessed in the Input Schema subsection. It is at the 2. Material and Methods → 2.2 Methodology → 2.2.2 Input Schema**

**Point 5:** I think your proposed approach can be used widely in medical applications. For example, it can be used in DNA classification, etc. For example, I find a paper titled "DNA Repair Genes (APE1 and XRCC1) Polymorphisms–Cadmium interaction in Fuel Station Workers", which has enough relation. Cite this paper and discuss about it as one the advantages of your proposed approach.

**Response 5: We agree that our proposed method can be used in many different approaches. However the paper you suggested is about DNA Repair Genes so that their approach is to analyze the effect of Cadmium on Fuel Station Workers, therefore visulizing the genes is needed. DNA classification, on the other hand use another approach. Since those above image is not the real image of genes (hard to extract main pattern), DNA sequence ("AGXTTTATTX" for example) is applied.**

**Point 6:** It is suggested to discuss about the runtime of your proposed method briefly ( Compare performance with other methods is not needed)

**Response 6: We agree with your comment, in the paper we just briefly show an interesting point about the performance of mobile-based models which have pretty low number of parameters.**

**Point 7:** In scientific papers, usually, the title of the tables is written above them.

**Response 7:**

**Point 8:** The pre-process of skin lesion recognition is skin detection process. For example, I find a paper titled "An innovative skin detection approach using color based image retrieval technique", which has relation. Cit this paper and discuss about the necessary pre-process in this scope briefly.

**Response 8: We agree with your comment. The paper you suggest is really fantastic that we have not consider the way of approaching. We have already cite it and discuss a little bit about it and the need of image preprocessing before feeding into the model. It is added in the section 2. Materials and Method → 2.2 Methodology → 2.2.2 Input Schema.**