Group 47 CW1 Feedback

1 Introduction & Dataset
- The introduction is clear about the task at hand.
- Data description and the cross-validation process is well defined.

Points for improvement:
- The introduction could be improved by setting some context of the two different image classification approaches, i.e., classic computer vision vs deep learning.
- The introduction can also include motivation for the robustness exploration and comparison between the two models.
- To be as least ambiguous as possible, be explicit about when and how the hyperparameter tuning fits in the data splitting process.

2 Methodology
- ResNet procedure is well explained with clear descriptions of the data transforms, optimisation, and regularisation.
- The implementation of the Bag of visual words procedure in the SIFT-BOW classifier is very well-defined.
- Method evaluation is well explained for both models — a good illustration of model-performance vs hyperparameter setting with the line chart.

Points for improvement
- ResNet model description could be improved by providing more details on the existing architecture and the modifications made, such as replacing the last fully connected layer. Also, it might be beneficial to include explanations on why the data transforms would lead to better network performance and whether such random transforms are applicable to the validation set.
- No hyperparameter tuning found for ResNet. Potentially could explore different learning rates, mini-batch size, number of layers as a hyperparameter.
- As an alternative, explore dense sift to boost the performance of SVM.
- The regularisation parameter in SVM, lambda, could be better explained. If this is referring to the misclassification penalty coefficient, then this should be explicitly addressed. It is possible to include this as a hyperparameter for the model too, however, given that the kernel type is also selected to be a hyperparameter, it is fair to use a single value to avoid overcomplicating the fine-tuning.
- The impact (computational, overfitting etc.) of using a more extensive vocabulary size can also be discussed.
- For perturbation of HSV Hue Noise Increase, the value greater than one should be minus by one rather than set to 1.
- It is advisable to refer to code when discussing the implementation to make it easier for the reader to validate the method with code.

3 Results & Discussion
- The discussion addresses the trends observed in the results and provides support through referenced literature.

Points for improvement
- To extend the explanation of performance by providing theoretical support/reasoning for the observations, i.e., why one model exhibits stronger robustness to a particular perturbation than another. Possibly link this to the differences in characteristics between a deep model vs a classical computer vision model.
- Graphs can be better represented by standardising both the Y and X axis scales and label for all graphs, so they are comparable. Consider including one model's performance for all perturbations in one graph and arrange side by side with the other model for better comparison and visualisation.
- To include +/- one standard deviation in all result graphs.
- Consider placing all graphs of perturbation results in the discussion section rather than in the methodology section to make it easier for readers to refer to these graphs while reading the explanation.