

Data Appendix

Introduction

This project aims to utilize the methodology outlined in "Protest Activity Detection and Perceived Violence Estimation from Social Media Images" by Wondonghyeon et al. to analyze photos from the Hong Kong protests [1]. Through this analysis, we seek to gain insights into protest dynamics and the perceived levels of violence during these events and determine the validity of our hypothesis using a CNN model, images from the Hong Kong protests will depict *higher* levels of perceived violence compared to general social media posts

Original Datasets Overview

- Entries: The dataset consists of 60 photographs during the Hong Kong protests.
- Variables:
 - The dataset contains the ImageName, protest, violence, sign, photo, fire, police, children, group 20, group 100, flag, night, and shouting features.
- Time Frame: The dataset was collected during the Hong Kong protests in 2019.
- Dataset CSV: The dataset was taken from the internet archive and the images were downloaded as a jpeg file [2].

DF Variable Descriptions

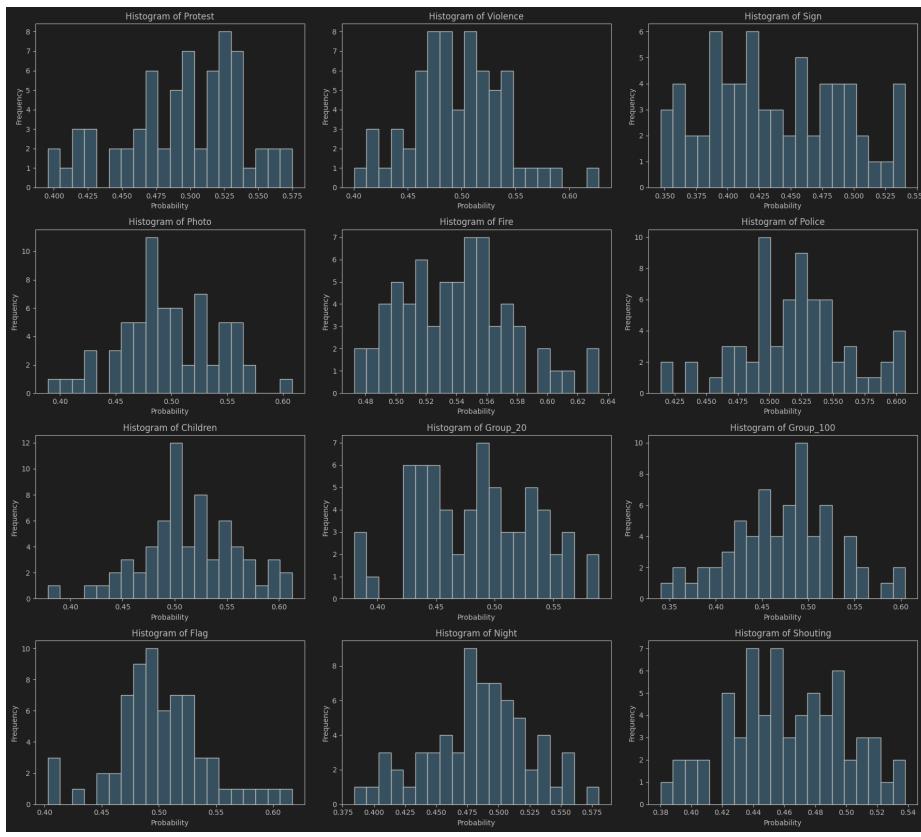
Data Dictionary:

Note: These are features so there was no variable type attributed to the data. The variable types below are in reference to the probability results.

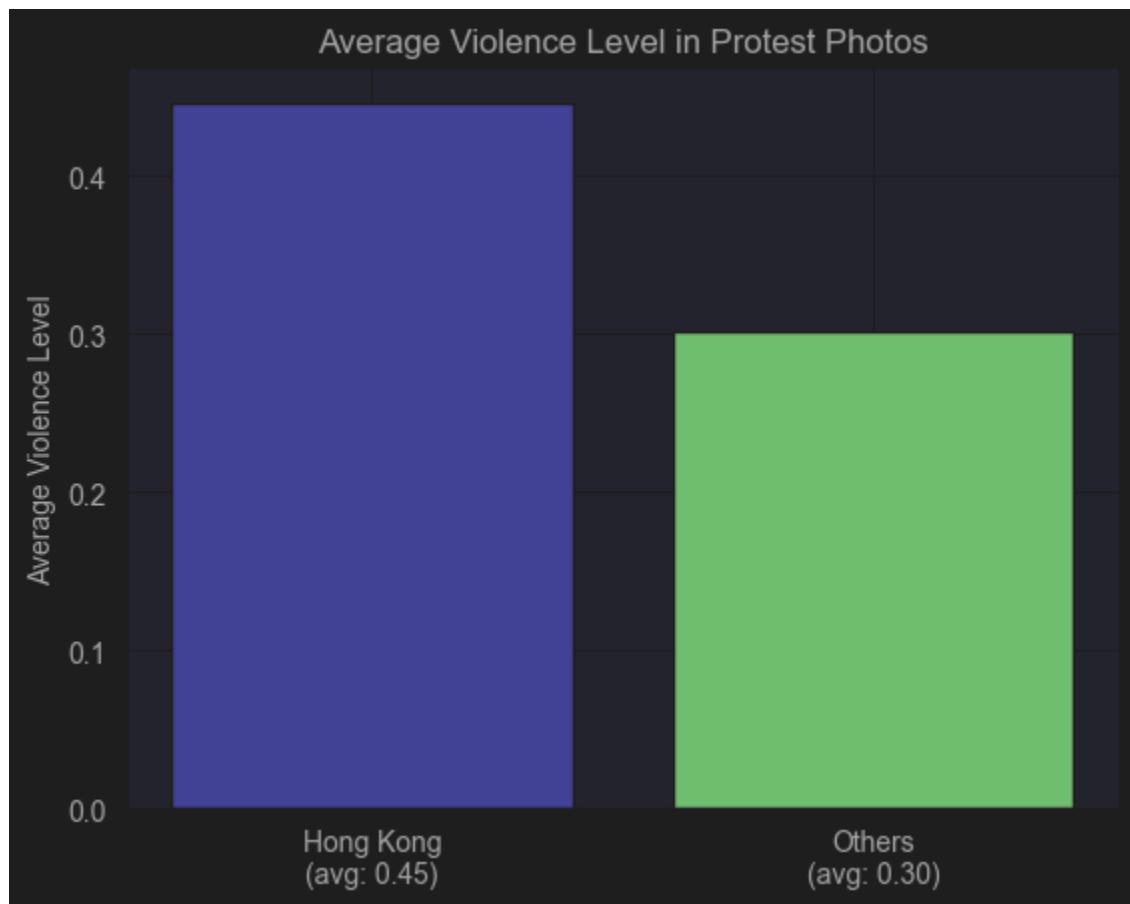
| Variable Name | Variable Type | Description |
|---------------|-------------------------|---|
| ImageName | pandas Series / String | The name of the image |
| Protest | pandas Series / Float64 | The probability of a protest in the image |
| Violence | pandas Series / Float64 | The probability of violence in the image |
| Sign | pandas Series / Float64 | The probability of signs in the image |
| Photo | pandas Series / Float64 | The probability of photos in the image |
| Fire | pandas Series / Float64 | The probability of fire in the image |

| | | |
|-----------|-------------------------|--|
| Police | pandas Series / Float64 | The probability of police in the image |
| Children | pandas Series / Float64 | The probability of children in the image |
| Group_20 | pandas Series / Float64 | The probability of there being more than 20 people in the image |
| Group_100 | pandas Series / Float64 | The probability of there being more than 100 people in the image |
| Flag | pandas Series / Float64 | The probability of flags in the image |
| Night | pandas Series / Float64 | The probability of it being night in the image |
| Shouting | pandas Series / Float64 | The probability of shouting in the image |

Detailed Statistical Analysis**Histogram representing the distribution of predicted probabilities**



The provided graph exhibits a series of histograms representing the distribution of predicted probabilities for a range of factors associated with Hong Kong protest images, as determined by an AI model. Each histogram corresponds to a specific attribute such as Protest, Violence, Sign, and so on, showcasing the frequency at which certain probability scores occur within the dataset. Peaks within a histogram indicate commonality in the model's certainty regarding the presence of a particular attribute in numerous images. For instance, a peak in the 'Violence' histogram suggests a significant number of images were confidently identified by the model as depicting violence. Conversely, the broader spread in the 'Sign' histogram suggests a varied interpretation of signs presence in the images, indicating either a range of different signs captured or varying levels of model certainty. This distribution pattern allows for an understanding of which attributes are most predominantly and clearly represented in the dataset, according to the model's assessment, providing insights into the nature of the protests and the model's interpretation capabilities.



The bar graph provides a clear visual representation of the differences in perceived violence levels between images from the Hong Kong protests and general social media protest posts. The data shows that the average perceived violence level in Hong Kong protest photos (0.446) is markedly higher than in the other category (0.302). This significant disparity supports the hypothesis that Hong Kong protest images are generally seen as more violent.

The t-test statistical analysis strengthens this observation:

- The average violence level for Hong Kong protest images stands at approximately 0.446, while for other social media protest images, it is around 0.302.
- The t-statistic is 8.974, which is considerably high, indicating a strong difference between the two groups.
- The p-value is extraordinarily low, at 3.01e-19, far surpassing the conventional threshold for statistical significance ($p < 0.05$). This implies that the probability of the results occurring by random chance is virtually nonexistent.

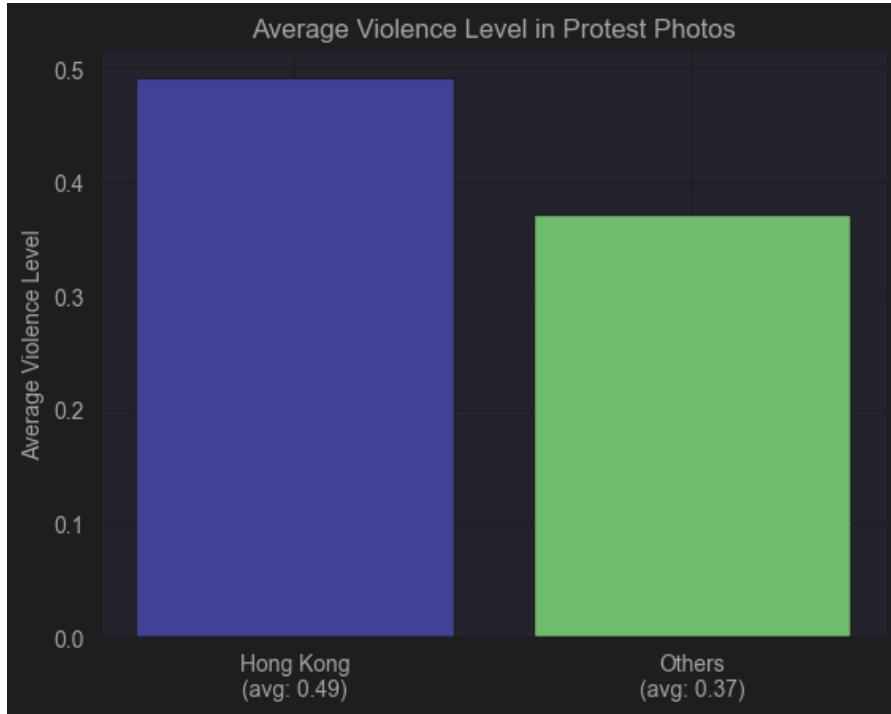
In summary, the statistical evidence is compelling and supports the hypothesis that Hong Kong protest photos contain higher levels of perceived violence compared to a broader category of

protest images from social media. This finding is not only statistically significant but also suggests that the methodology used to evaluate these images, likely involving a convolutional neural network (CNN) model, is both sensitive and specific in assessing the degree of violence depicted in protest imagery.



This box plot compares the spread and outliers of the violence levels between Hong Kong protests and general protests. The Hong Kong box plot shows a higher median, and the data are spread out indicating more variation in the violence levels within the Hong Kong photos. The 'General Protest' plot has a lower median, and while it has a wider interquartile range (indicating that the middle 50% of the data is more spread out), the overall levels of violence are lower. Outliers are present in both categories, which are individual points that fall well outside the general spread of the data.

To refine our understanding of perceived violence in protest imagery, we analyzed images with a protest score of 0.4 and above. This focused approach provides insight into the higher end of the violence spectrum.



The bar graph illustrates that among images with notable protest characteristics, those from Hong Kong protests have an average perceived violence level of 0.49, which is significantly higher than the 0.37 average for other protests. This reinforces our hypothesis, indicating a consistent perception of greater violence in Hong Kong protest imagery.

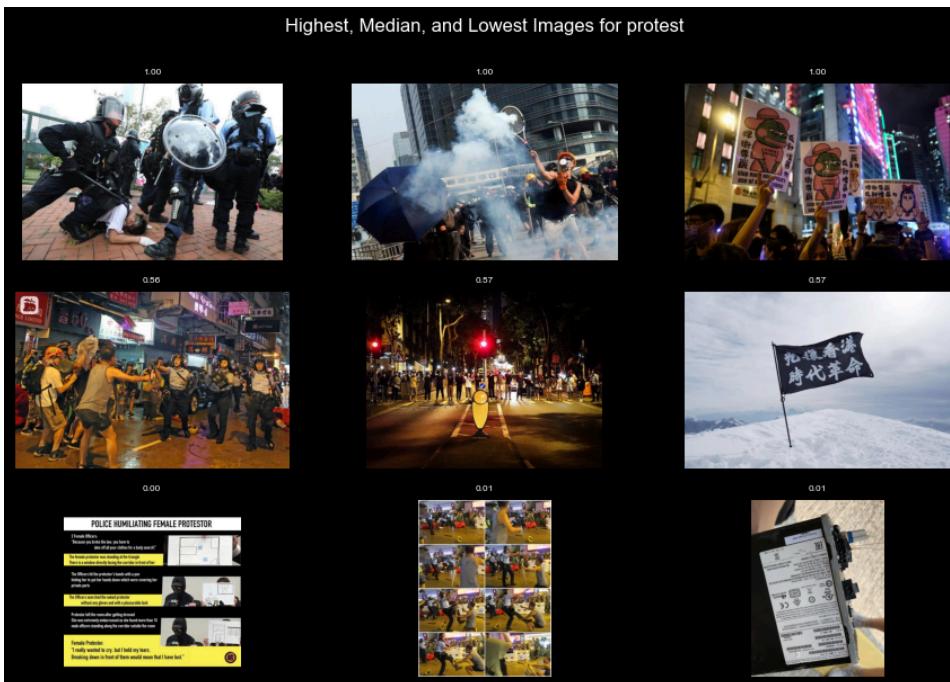
The statistical analysis through a t-test confirms the visual data:

- The t-statistic stands at 4.9, suggesting a robust difference between the Hong Kong protests and other protest images.
- The p-value is 9.63e-17, which is extremely low, reinforcing the likelihood that this difference is not due to chance, but a real disparity in the data.

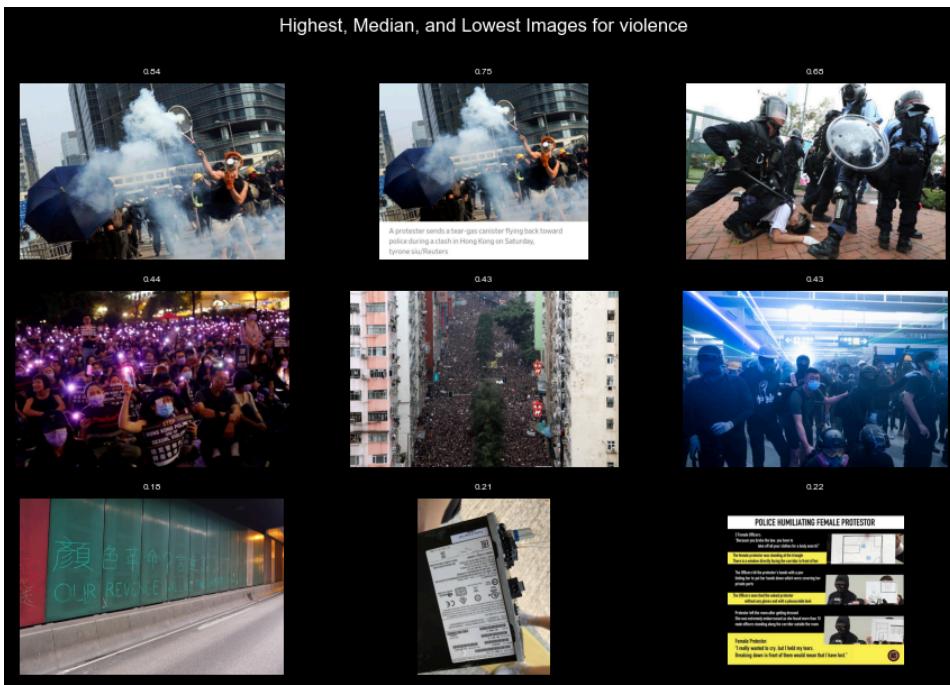
This analysis confirms with statistical significance that Hong Kong protest photos are perceived to have higher levels of violence compared to those from other protests across social media.

Informal Evaluation of Trained Model on Tested Hong Kong Protest Images

Highest, Median, and Lowest Images for protest



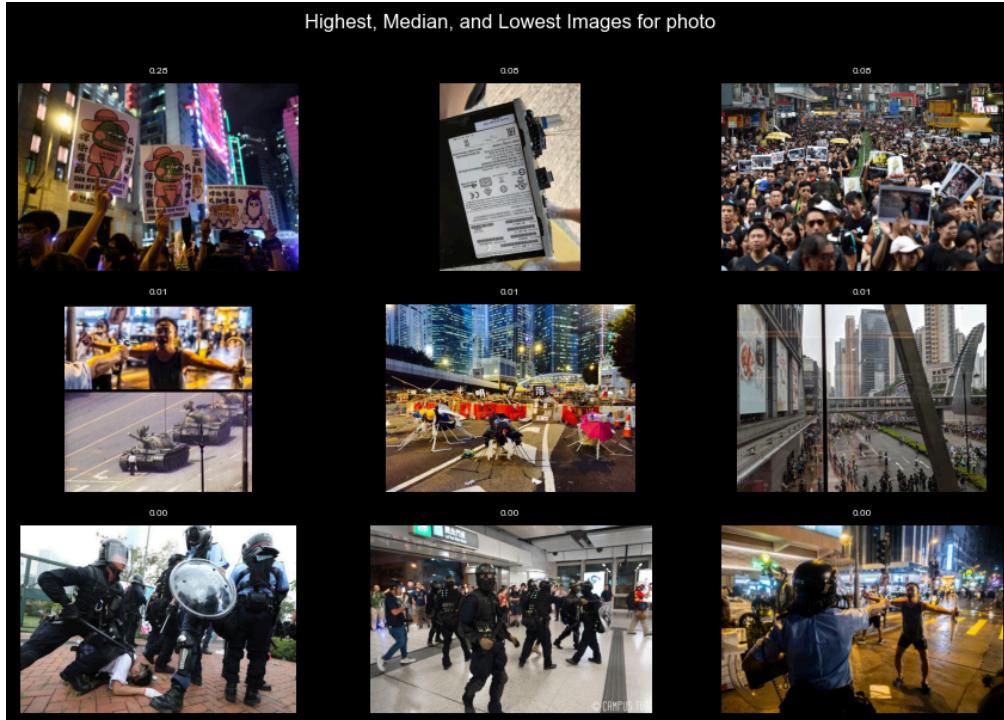
Highest, Median, and Lowest Images for violence



Highest, Median, and Lowest Images for sign



Highest, Median, and Lowest Images for photo

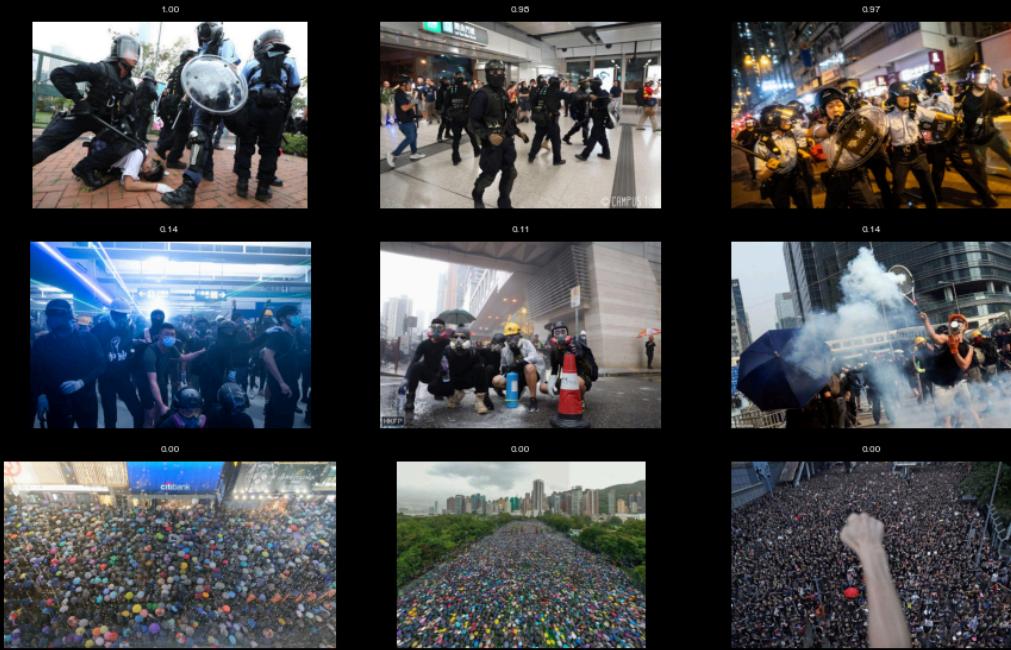


Katherine Driebe (Leader), Lasata Tuladhar, Ramya Tangirala

DS 4002

4/22/24

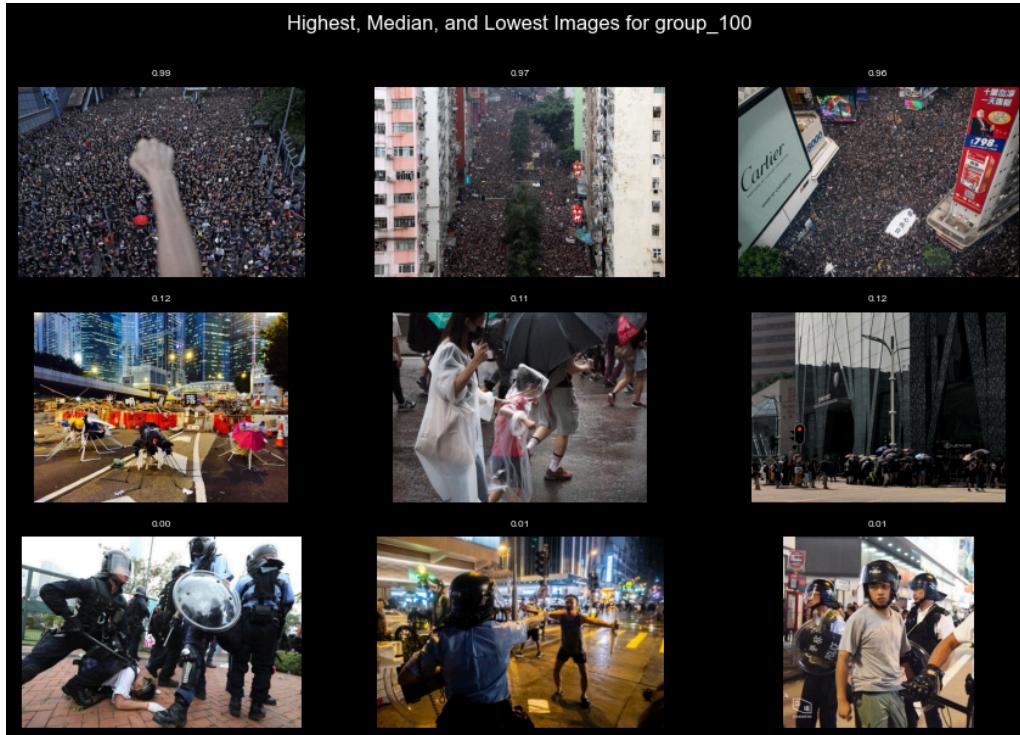
Highest, Median, and Lowest Images for police



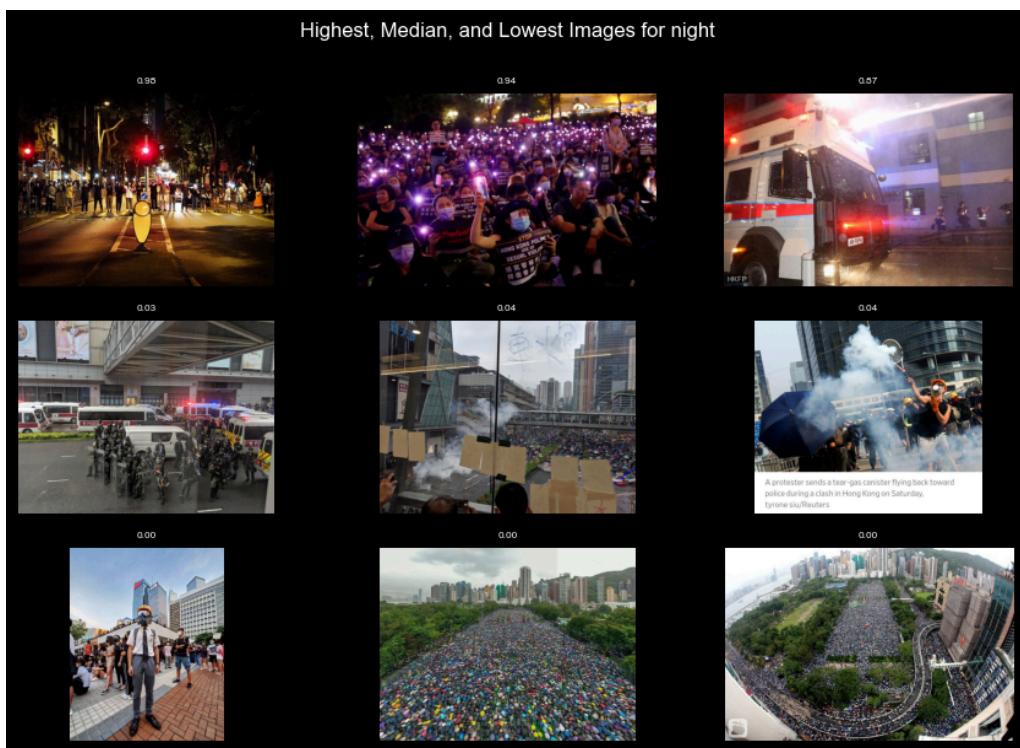
Highest, Median, and Lowest Images for group_20



Highest, Median, and Lowest Images for group_100



Highest, Median, and Lowest Images for night



Each plot displays the highest, median, and lowest three images scored for each variable. Based on an informal evaluation, the model demonstrates considerable accuracy in its scoring, with the main inaccuracies arising from photos not related to protests.

References

- [1]D. Won, Z. C. Steinert-Threlkeld, and J. Joo, "Protest Activity Detection and Perceived Violence Estimation from Social Media Images," in Proceedings of the 25th ACM International Conference on Multimedia, 2017.
- [2] "HongKongProtests2019_gallery_001," retrieved from Archive.org, [Online]. Available: https://archive.org/details/HongKongProtests2019_gallery_001/EC0SY18UEAA4HnU.jpg.