



CrowdSenseAI

Avert Crowd Crushing with Deep Learning

TABLE OF CONTENTS

01

PROBLEM
STATEMENT

02

EDA

03

MODELLING

04

MODEL
EVALUATION

05

LIMITATIONS &
FUTURE WORKS

06

CONCLUSIONS

SEOUL HALLOWEEN CROWD CRUSH 2022



The Telegraph

HOW THE TRAGEDY UNFOLDED

CELEBRATION

First virtually unrestricted Halloween festivities in three years that attracted 100,000 crowd on the street

CONTROL

Witnesses said police had trouble controlling the crowds.

DENSITY

Bodies are jammed together so tightly that they can no longer choose where they go and they begin to behave like a fluid.

Source:

<https://www.theguardian.com/world/2022/oct/31/how-did-the-seoul-itaewon-halloween-crowd-crush-happen-unfolded-a-visual-guide#:~:text=3-,A%20densely%20packed%20group%20exert%20pressure%20on%20each%20other%2C%20but,effect%20of%20similar%20holes%20elsewhere.>

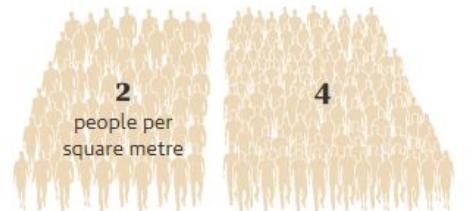
HOW THE TRAGEDY UNFOLDED

Crowd Crushing
Risk increases
when crowd
density
increases.

Risk of accidents in a moving crowd

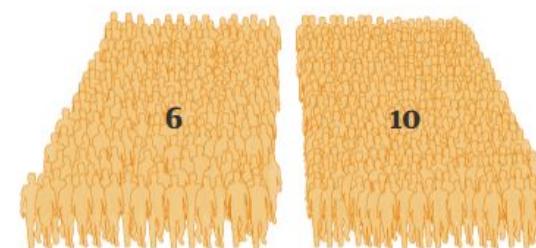
Low risk

Moving crowds with a density of up to four people per square meter are safe, because people have enough room to make decisions and move accordingly



High risk

When the density is higher - at about six people per square metre - bodies are jammed together so tightly that they can no longer choose where they go and they begin to behave like a fluid. Pressure waves can travel through them and they lose control



Guardian graphic. Source: G Keith Still; University of Suffolk

Source:

<https://www.theguardian.com/world/2022/oct/31/how-did-the-seoul-itaewon-halloween-crowd-crush-happen-unfolded-a-visual-guide#:~:text=3-,A%20densely%20packed%20group%20exert%20pressure%20on%20each%20other%2C%20but,effect%20of%20similar%20holes%20elsewhere.>

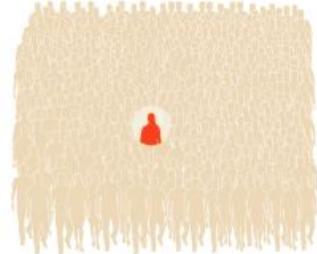
HOW THE TRAGEDY UNFOLDED

Densely packed crowd behaving like fluid.

How the progressive crowd collapse happens

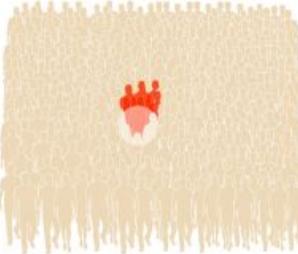
1

A densely packed group exert pressure on each other, but are kept in place by pressure from the other side. A shockwave through the crowd or a slip can cause a fall.



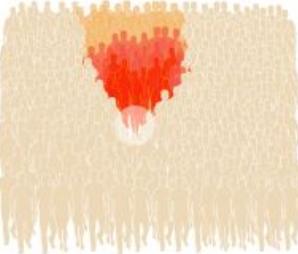
2

The fall of one person creates a sudden gap in the crowd and removes the opposing force that kept the crowd in equilibrium.



3

Others fall into the new space, and more follow them until the pressure eases. That process can create a ripple effect of similar holes elsewhere.



Guardian graphic. Source: G Keith Still; University of Suffolk

Source:

<https://www.theguardian.com/world/2022/oct/31/how-did-the-seoul-itaewon-halloween-crowd-crush-happen-unfolded-a-visual-guide#:~:text=3-,A%20densely%20packed%20group%20exert%20pressure%20on%20each%20other%2C%20but,effect%20of%20similar%20holes%20elsewhere.>

17,512

Total Deaths from 164 Crowd Crushing Incident

01

PROBLEM STATEMENT

PROBLEM STATEMENT

SAFETY



Address the critical issue of crowd safety during public gatherings and events.

CONTROL



Prevention of crowd crushing disasters caused by overcrowding and mismanagement of crowded spaces.

DETECTION



Develop sophisticated deep learning model for real-time crowd presence detection



MITIGATION

Early warnings to authorities in hazardous conditions to protect lives

METHODOLOGY

DATA

Source for the right data to solve the problem



PROBLEM

Identifying the problem statement



EDA

Explore & Analysis the Data

MODELING

Develop Deep Learning Model for prediction



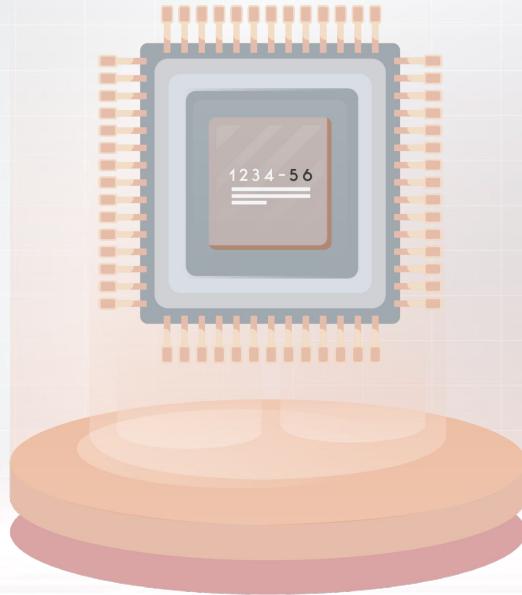
EVALUATION

Evaluate the performance of the models

DEPLOYMENT

Test & deploy the final product





02

EXPLORATORY DATA ANALYSIS

DATASET

Crowd dataset
with **4k images**
for model training.



GROUND-TRUTH LABELS

Ground-Truth Labels for
training the model in the
format:

(c, x, y, w, h)

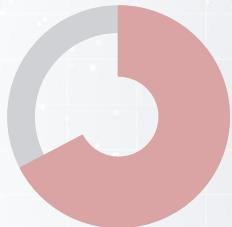
C: Class

X & Y: Center of Bounding Box

W & H: Width & Height of Bounding Box

```
0 0.121875 0.27421875 0.0328125 0.05625
0 0.40703125 0.28046875 0.01875 0.0421875
0 0.221875 0.28203125 0.0234375 0.0421875
0 0.32109375 0.28359375 0.0234375 0.0421875
0 0.51328125 0.2890625 0.0234375 0.0375
0 0.2703125 0.2875 0.0203125 0.046875
0 0.859375 0.309375 0.0234375 0.028125
0 0.90625 0.33125 0.0234375 0.0421875
0 0.73203125 0.34453125 0.01875 0.0328125
0 0.98359375 0.3515625 0.0203125 0.0515625
0 0.74296875 0.3671875 0.01875 0.0328125
0 0.13671875 0.35703125 0.0203125 0.0515625
0 0.62265625 0.37265625 0.025 0.046875
0 0.30078125 0.37421875 0.0203125 0.0609375
0 0.70390625 0.38125 0.01875 0.0421875
0 0.3796875 0.38203125 0.01875 0.0234375
0 0.96171875 0.046875 0.0140625 0.0421875
0 0.090625 0.38203125 0.025 0.05625
0 0.328125 0.38828125 0.0296875 0.05625
```

TRAIN, VALIDATION & TEST



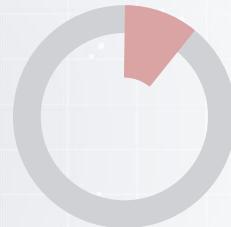
70%

Train Dataset
2.9k images



20%

Validation Dataset
836 images



10%

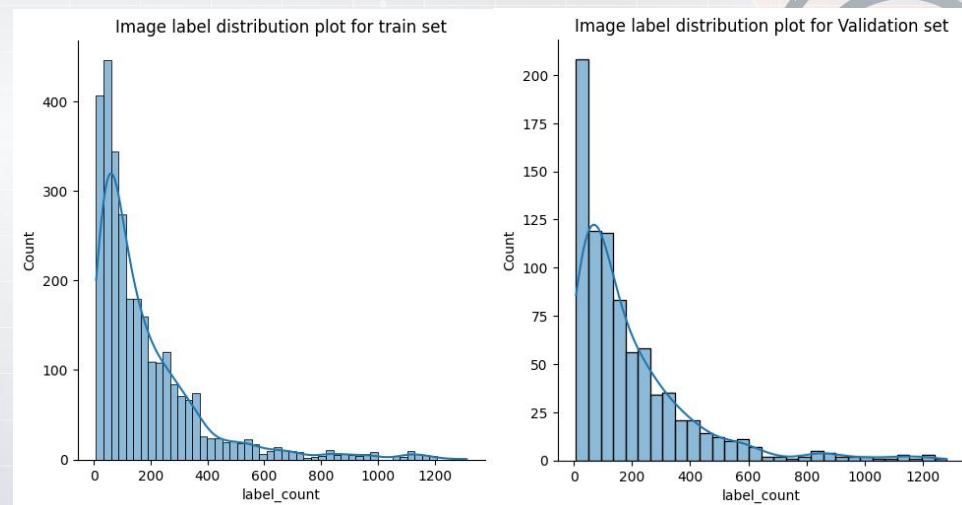
Test Dataset
415 images

PROPERTIES

SIZE



Image Size:
640 x 640



LABELS

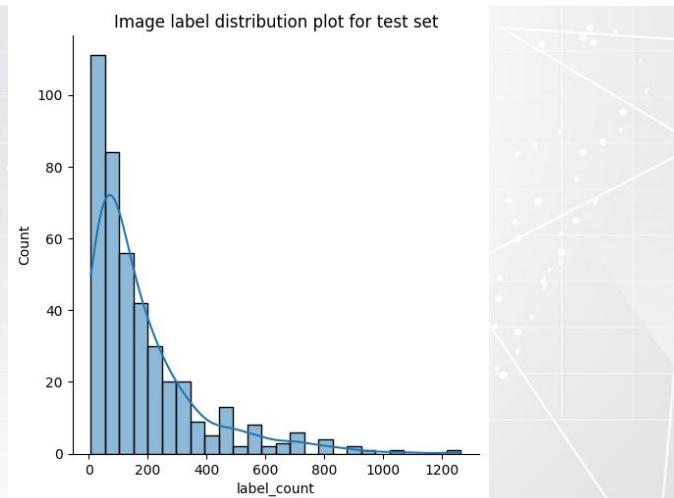


Ground-Truth Labels majority
below 200 per image

AUGMENTATION



Images were randomly tilted
with mixture of colored &
gray-scaled images



03

MODELING

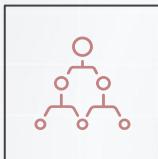


Model selection

- Project utilized two models: YOLOv5 and YOLOv8.
- YOLO stands for "You Only Look Once."
- YOLO's algorithm is real-time object detection, identifies and locates multiple elements in an image.
- Uses convolutional neural networks (CNN) for instant recognition.
- YOLO performs predictions for the entire image simultaneously.
- Offers fast and accurate object detection for various applications.



THE PREPARATION



MODELS

YOLOv5 & YOLOv8



MACHINE

Modeling performed on Google Collab
GPU: NVIDIA A100-SXM4-40GB

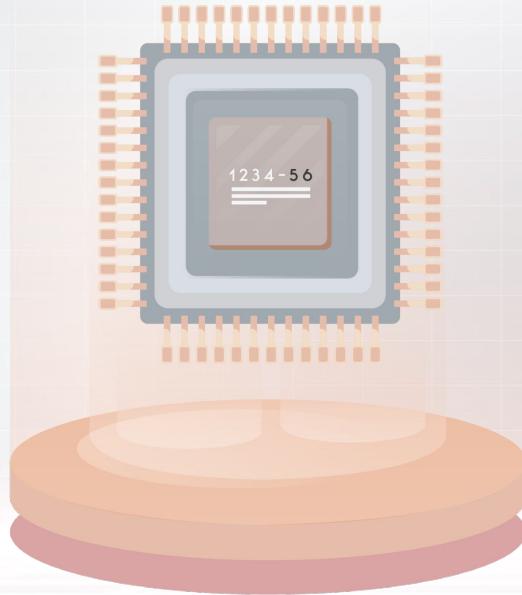


PARAMETERS

Epoch: 100 & 200
Batch: 8 & 16

Modeling

| MODEL | EPOCH | BATCH | REMARKS |
|--------|-------|-------|--|
| YOLOv5 | 100 | 16 | Training: train dataset Validation: All 3 dataset (train, valid, test). |
| YOLOv8 | 100 | 8 | Training: train dataset Validation: All 3 dataset (train, valid, test). Batch: 8 due GPU out of memory error |
| YOLOv8 | 200 | 8 | Training: train dataset Validation: All 3 dataset (train, valid, test). Batch: 8 due GPU out of memory error |

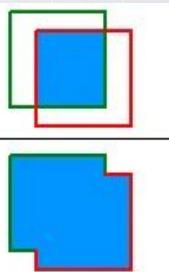


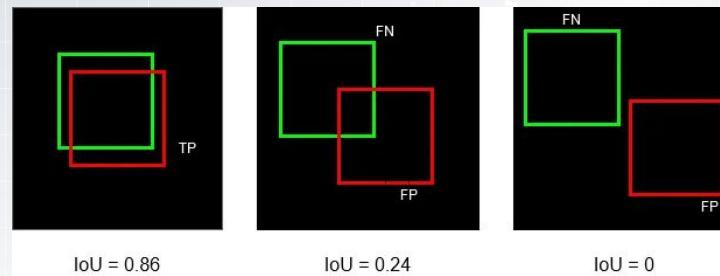
04

MODEL EVALUATION

Intersection over Union

- Intersection over Union measures the extent of overlap between the ground truth (gt) and the prediction (pd) bounding boxes.
- IoU is used with a threshold (α) to determine correct detections.
- Example on the right shows the output when α is set to 0.5 (TP: $\text{IoU} \geq \alpha$, FP & FN: $\text{IoU} < \alpha$)

$$IOU = \frac{\text{area of overlap}}{\text{area of union}} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$




Definition of terms:

- True Positive (TP) — Correct detection
- False Positive (FP) — Incorrect detection
- False Negative (FN) — A Ground-truth missed (not detected)
- True Negative (TN) — Not used in object detection as it pertains to background regions that are correctly not detected by the model.

METRICS

mAP

mean Average Precision (mAP) assess the model's overall detection accuracy across all classes.

PRECISION

Measures the model's accuracy in identifying relevant objects by calculating the ratio of True Positives (TPs) to all detections

RECALL

Evaluates the model's ability to detect all ground truth objects, indicating the proportion of True Positives (TPs) among all actual ground truth instances.

$$P = \frac{TP}{TP + FP} = \frac{TP}{\text{all detections}}$$

$$R = \frac{TP}{TP + FN} = \frac{TP}{\text{all ground-truths}}$$

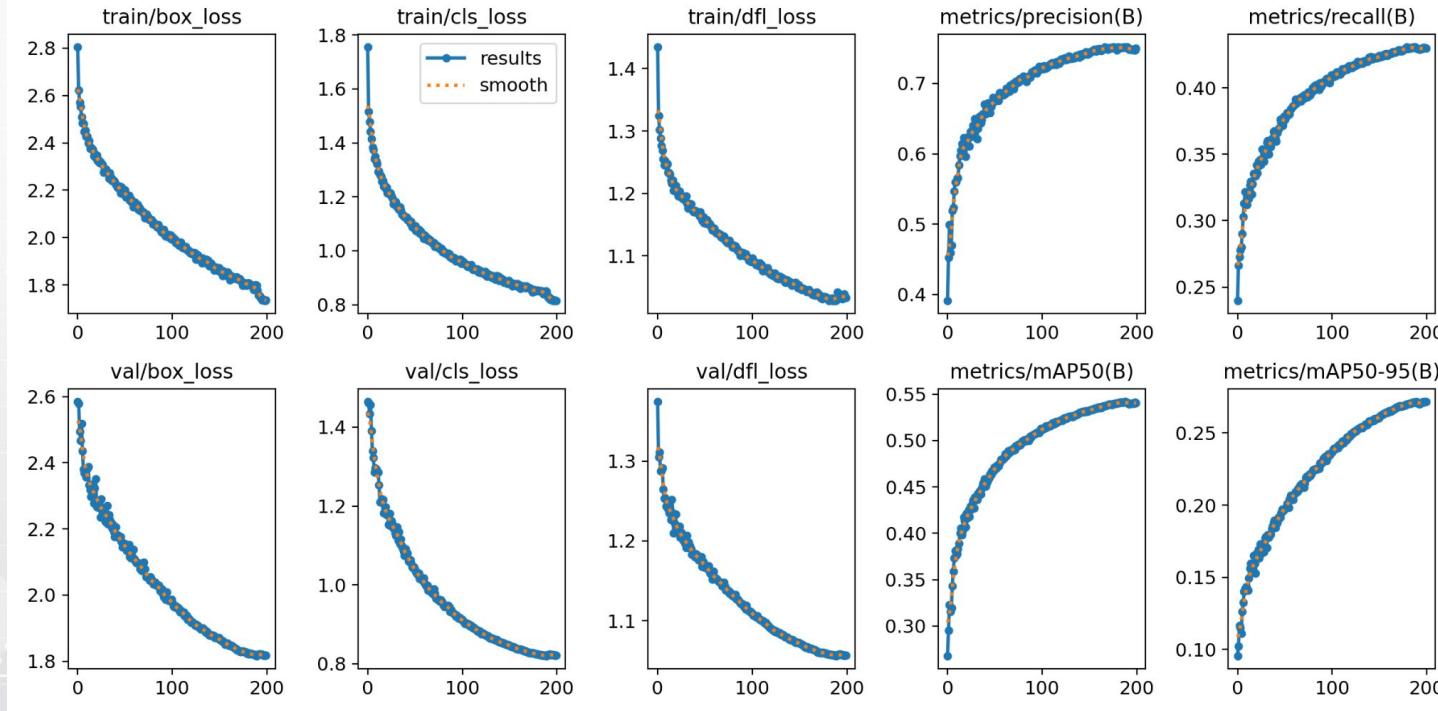
Model Performances

| | mAP50 ₁ | | | Precision | | | Recall | | |
|-----------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| | Train | Val ₂ | Test | Train | Val ₂ | Test | Train | Val ₂ | Test |
| YOLOv5 (Baseline) | 46.7% | 45.6% | 45.8% | 70.7% | 69.6% | 69.0% | 37.3% | 36.5% | 36.9% |
| YOLOv8 (100 Epoch) | 55.1% <i>(+8.4%)</i> | 52.9% <i>(+7.3%)</i> | 54.6% <i>(+8.8%)</i> | 75.6% <i>(+4.9%)</i> | 73.8% <i>(+4.2%)</i> | 73.7% <i>(+4.7%)</i> | 44.4% <i>(+7.1%)</i> | 42.3% <i>(+5.8%)</i> | 44.6% <i>(+7.7%)</i> |
| YOLOv8 (200 Epoch) | 56.6% <i>(+9.9%)</i> | 54.1% <i>(+8.5%)</i> | 55.7% <i>(+9.9%)</i> | 77.6% <i>(+6.9%)</i> | 74.8% <i>(+5.2%)</i> | 74.9% <i>(+5.9%)</i> | 45.3% <i>(+8.0%)</i> | 43.0% <i>(+6.5%)</i> | 45.3% <i>(+8.4%)</i> |

1: Mean Average Precision @ IoU 0.5 2: Validation Score

Model Results

- Metrics performance rate of increase diminishes as epoch approaches the end.
- Validation losses (box, class, distribution focal loss) approaches zero towards the end.



MODEL INFERENCE

Predictions on
the test set
made by the
best model.



PRODUCT DEMONSTRATION



05

LIMITATIONS & FUTURE WORKS

LIMITATIONS & FUTURE WORKS

HARDWARE

High processing power is required for deep learning models. Higher processing GPU required.

DATA

Model is as good as the data gets. More data required to improve model accuracy.

TIME

Besides processing power, more time is required to run models for further improvement.

AREA

Future work: Automated calculation / retrieval from mapping database for the triggering limits

WARNING

Future work: Automated messaging / email system to alert authorities.

06

CONCLUSIONS

CONCLUSIONS



PREEMPTIVE

Closing off high-risk areas during public gatherings and events before they become overcrowded



PROACTIVE

Deployment of security personnel, regulation of crowd control at congregation areas and potential chokepoints



PREDICTIVE

CCTVs and drones with deep learning detection model to monitor the crowd capacity & provide early warning.

Technology alone is not enough!

THANKS!



<https://www.linkedin.com/in/neo-wenxi/>

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon** and infographics & images by **Freepik**