CS4243 WEAPON DETECTION

Tan Ying Hui, Gan Jing Wen, Ekam Singh Pandher, Mathan Mahendran
TEAM ID: 34

OVERVIEW

This project aims to identify the threat levels of individual images with one person, **classifying between 'Threat', 'Carrying' and 'Normal' labels.** It uses a pre-trained ANN (Resnet18) as baseline which we seek to improve through **data cleaning** and **various image processing techniques.** Lastly, we combined these techniques into a final model, which achieved a **weighted-F1 score of {insert final score here.}** Our analysis of the results are as provided in this poster.

OUR APPROACH

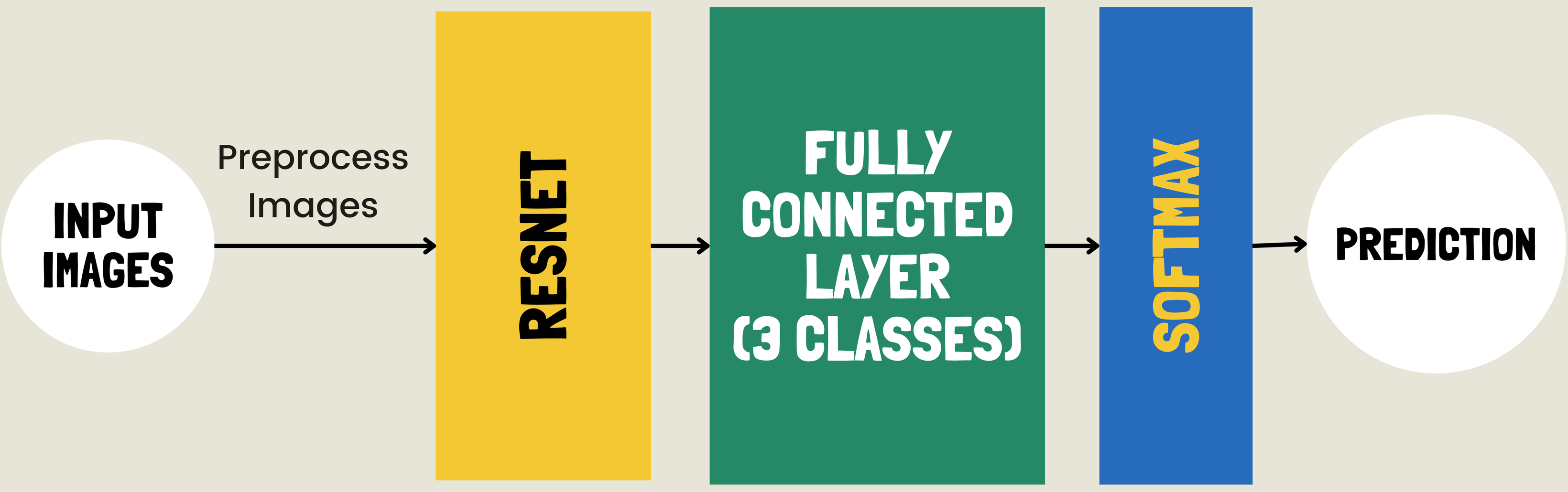
We approached this as an Image Classification Problem. We conducted data cleaning to improve dataset quality and data augmentation methods for data generation. We also experimented introduced noises to the images to mitigate overfitting.

MODEL AND HYPERPARAMETERS USED

Transfer Learning using PyTorch **ResNet18**, a deep residual network pretrained on ImageNet Database.

Optimiser: ADAM
- Learning rate: 1 x 10^-3
- Weight decay: 1 x 10^-5

Loss Function: Cross Entropy Loss
Batch Size: 256, Epochs: 10



DATA PREPARATION

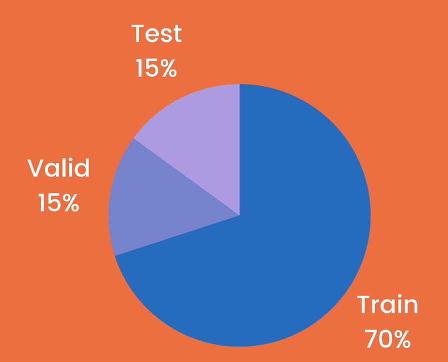
DATA CLEANING

Yolov3 is used to filter away images with zero or more humans.
Cleaned data sample contains only one human.

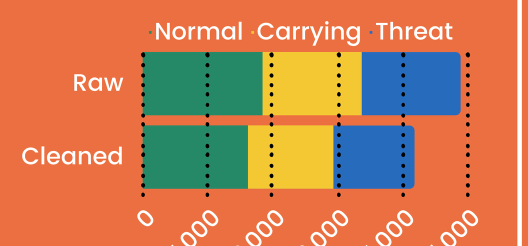
NORMALISATION

Images are normalised before feeding into ResNet model

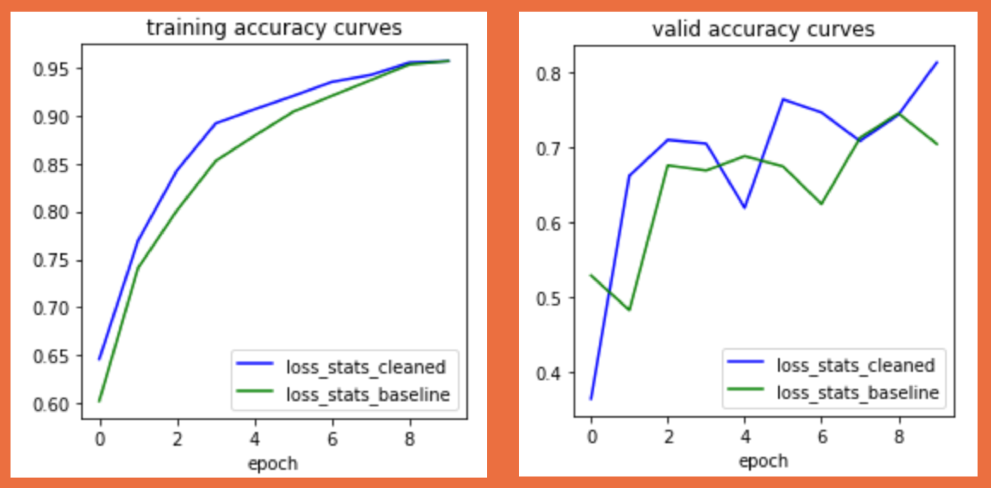
SPLITTING DATASET



PROPORTION OF DATA



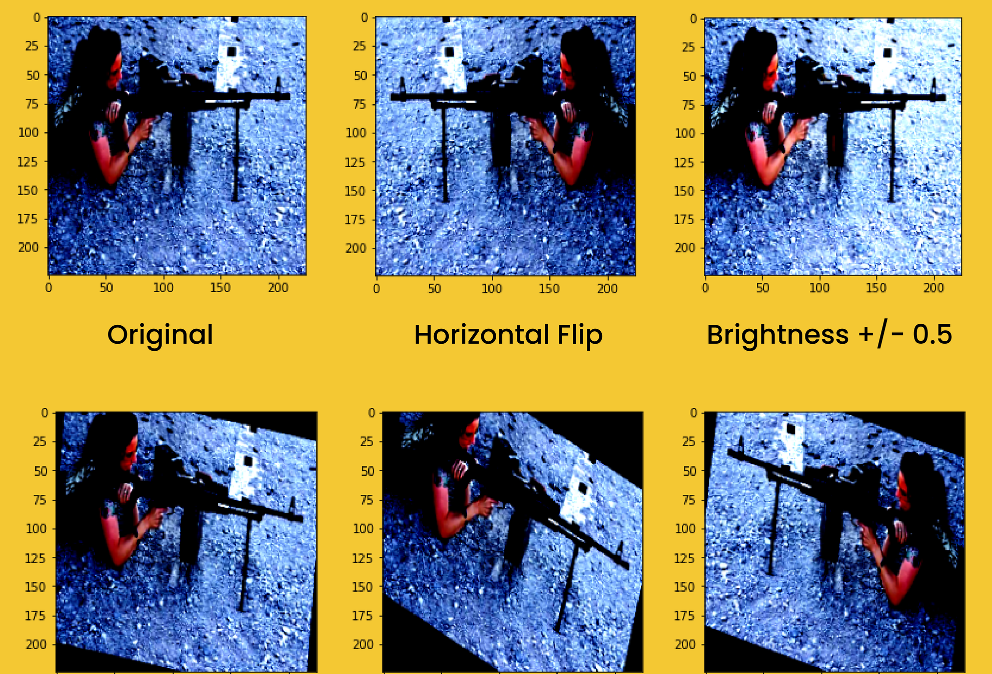
COMPARISON OF RAW VS CLEANED DATA PERFORMANCE



Cleaned data has shown to yield better training and validation accuracies across 10 epochs.

DATA AUGMENTATION – IMAGE TRANSFORM

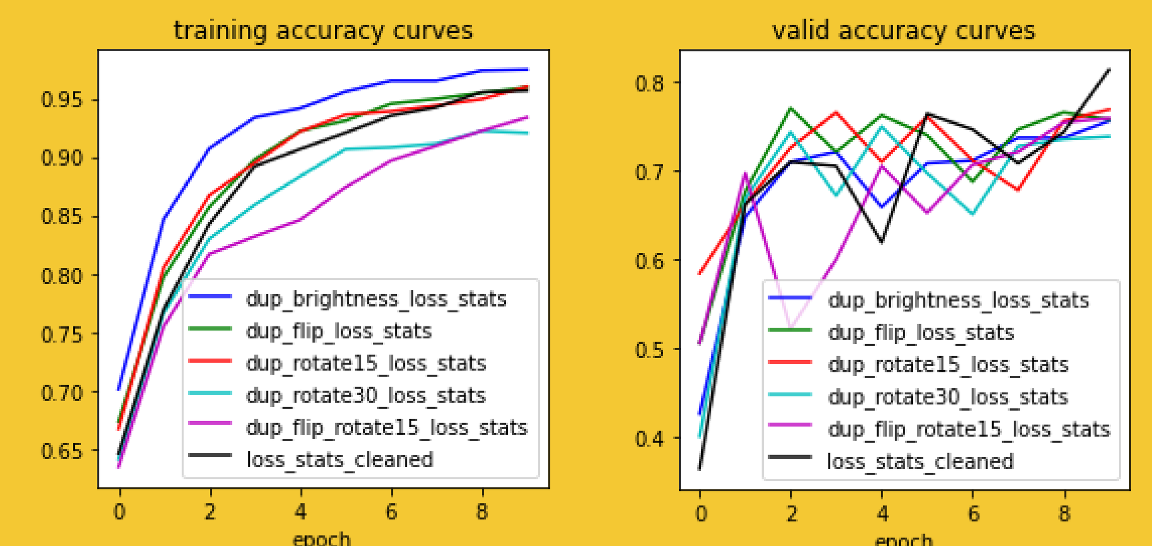
Wang et al. (2017) shows that adding traditional image augmentation improves performance as compared to no augmentation used.



Original | Horizontal Flip | Brightness +/- 0.5

Rotate +/- 15 deg | Rotate +/- 30 deg | Horizontal Flip & Rotate +/- 15 deg

Comparison of Different Image Augmentation Methods

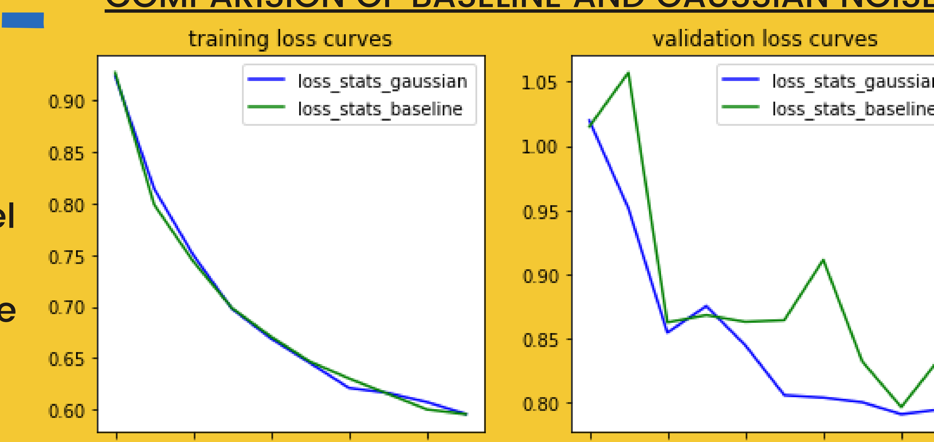


From the two Accuracy Curves, the model with 15 deg rotation and the model with horizontal flipping performs slightly better than baseline (cleaned data). We also attempt to combine the 2 augmentations to see if it improves performance.

DATA AUGMENTATION – NOISE

During training, we noticed that the model was showing indications of overfitting. Adding Gaussian Blur was one method we experimented to mitigate this issue.

COMPARISION OF BASELINE AND GAUSSIAN NOISE



KEY FINDING

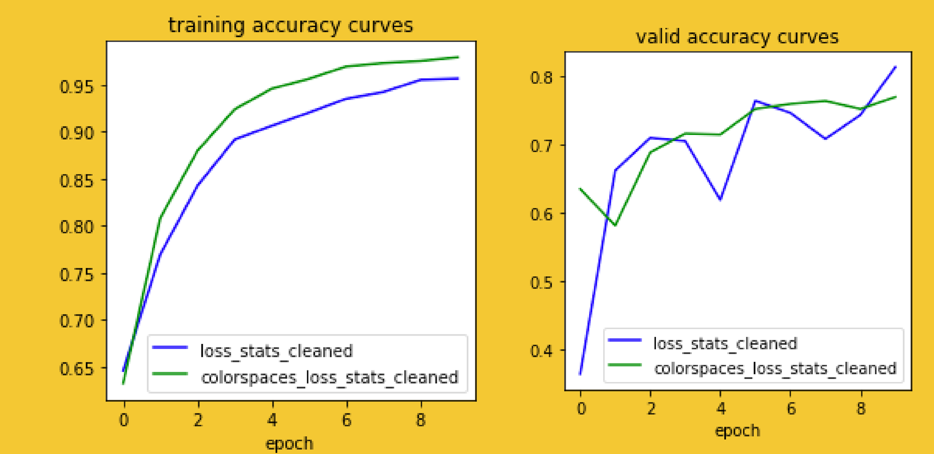
A 3x3 and a 5x5 kernal was experimented with. We found the 5x5 kernal performed better, likely due to a stronger blurring effect

DATA AUGMENTATION – COLOUR SPACE

Gowda et al. (2019) has shown that using different colour spaces can improve image classification performance. We use this as inspiration in investigating the effects of colour space for our problem.

RGB Color Space | HSV Color Space | LAB Color Space

COMPARISION OF USING COLOUR SPACES VS ONLY CLEANED DATA (BASELINE)

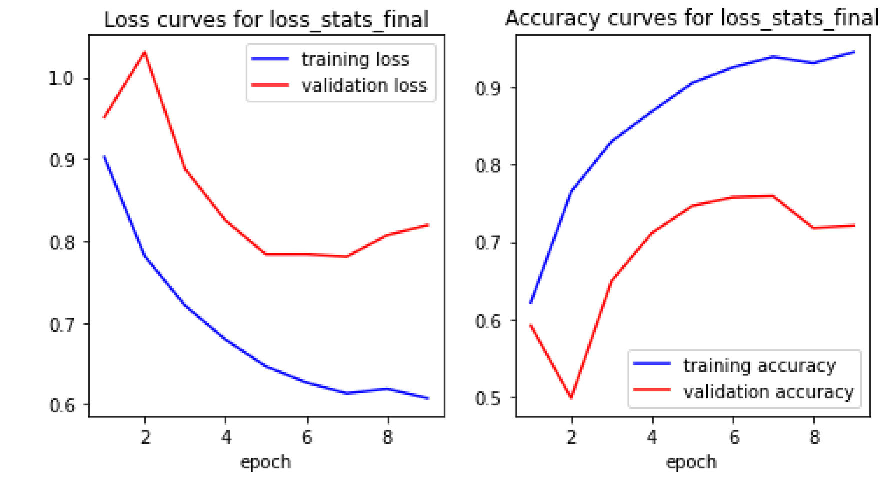


Using color spaces has shown better training accuracy and more consistent validation accuracy rise without fluctuating - but the final validation accuracy was slightly lower

FINAL MODEL

Model configuration: Trained on cleaned dataset. Train dataset was duplicated, where the duplicate image were randomly augmented with Rotations, Horizontal Flips, Brightness Jitter and Gaussian blur.

TRAINING AND VALIDATION STATISTICS



Diverging accuracy curves indicate possible overfitting.
Diverging loss curves indicate possible unrepresentative training dataset - training dataset does not provide sufficient information to learn the problem.

FINAL MODEL STATISTICS

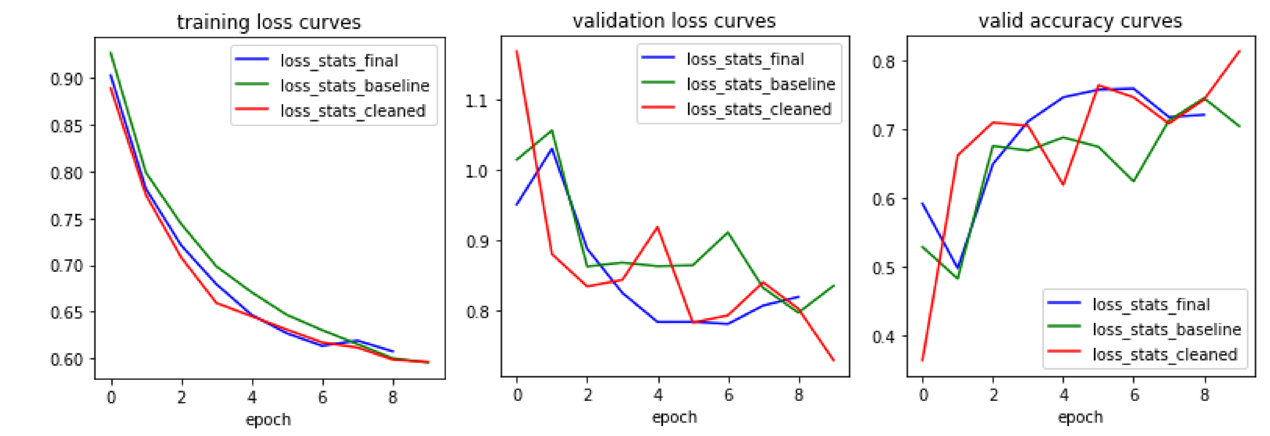
```
              precision    recall  f1-score   support

           0       0.83      0.90      0.86       243
           1       0.84      0.75      0.79       197
           2       0.78      0.79      0.79       188

    accuracy                           0.82       628   [[218   9  16]
   macro avg       0.82      0.81      0.81       628    [ 25 147  25]
weighted avg       0.82      0.82      0.82       628    [ 19  20 149]]
```

RESULTS

Model	Accuracy	Precision	Recall	Weighted-F1 score
Baseline (uncleaned)	73	77	73	73
Baseline (cleaned)	77	77	77	77
Final model	82	82	82	82

LEARNING CURVE FOR ALL

TRAINING AND VALIDATION STATISTICS COMPARING BASELINE, CLEANED AND FINAL MODELS



From the learning curves, the final model gives a more consistent validation curve as compared to the fluctuating ones in baseline models.

ANALYSIS

1. The results show improvements on the cleaned data with a better weighted f-1 score.
2. Augmentations expanded our limited dataset.
3. Augmentations helped to identify different orientations of images.
4. Gaussian Blur helps to upsample and mitigate overfitting to some extents.

We observe most of our mislabelled images are from the "Threat" abd "Carrying" classes. They could be the Hardest to detect since we are observing posture which may be unclear.

CONCLUSION

WHAT COULD BE IMPROVED?

- Experiment with **hyperparameter tuning**, use **regularisation** or **dropout** to further reduce overfitting
- Further cleaning data in 'carrying'/ 'threat' using a model to detect presence of guns and knives, and remove those without

OTHER IDEAS TO EXPLORE

- Incorporate **Human Pose Detection** as a feature extractor - Different poses when not carrying vs carrying vs using a weapon
- Generation of artificial images using **GANs**
- **Clustering algorithm** to handle mislabelled data

Note:

The team also experimented with processing images using Sobel Edge Detection, S&P Noise and CLAHE. However, the results were not notable, and excluded from the poster.