# Combating Human-Trafficking: Using Scene Recognition to Identify Hotels

Shubham Shetty
shubhamshett@cs.umass.edu

Adarsh Kolya
akolya@cs.umass.edu

## 1. Motivation

Human trafficking is estimated to have trapped 24.9 million victims.[1]. The pictures of these victims are often clicked in hotel rooms. Identifying these hotels is vital in the combat of human trafficking. We are trying to build a model that identifies the hotel in which the pictures of the victims had been clicked.

The problem was presented at the FGVC8 workshop [6]. This is more complex than a simple image recognition task, as there are a huge number of classification labels which is the number of hotels in the dataset. Also rooms within the same hotel may not look the same and rooms in different hotels (part of the same chain) may look very similar. Hence we may have high intraclass variance and low interclass variance. It becomes imperative to minimize the search space efficiently in order to identify possible matches within a reasonable margin of error.

## 2. Literature Review

A baseline approach for this challenge is already laid out by the paper that presented this challenge [4]. We will be reviewing ResNet [2] and GoogleNet [8] which have shown high accuracy in static scene recognition tasks. We also intend to experiment with EfficientNetv2 [9] and hence will be reviewing this too. In addition to these, we will learn about Selective contrast triplet loss [10], the loss function used by the state of the art models on the Hotels-50k data set [7].

## 3. Data

Our primary training dataset will be the 2021 Hotel-ID to Combat Human Trafficking Competition Dataset [4]. This dataset consists of photos of hotel room interiors without any people present. It consists of 97000+ images from 7700 hotels worldwide. Training and test data is hosted on Kaggle [3].

Our data can be augmented using the Hotels-50k dataset [7]. Additional model training can be done using this dataset, which contains around a million annotated hotel images and was also built for combating human trafficking.

## 4. Approach

We aim to use recent state of the art approaches in scene recognition to model our solution. Recent advances in deep learning for visual recognition and ensemble learning techniques [5] have achieved high performance in static scene recognition. VGG, ResNet, and GoogleNet have all shown high accuracy for this task. We also intend to experiment with EfficientNetv2 [9] and check its performance versus other models.

## 5. Evaluation

Average multi-class log loss and top-K classification accuracy will be used for evaluation. A baseline is provided in the original paper for comparison [4].

## 6. Group Members

This project will be undertaken by Adarsh Kolya (akolya@cs.umass.edu) and Shubham Shetty (shubhamshett@cs.umass.edu) with tasks equally divided amongst ourselves. Some broad tasks include, but are not limited to -

1. Literature review
2. Data preparation and pre-processing
3. Modeling and evaluation
4. Ablation study
5. Report preparation

In the modelling and evaluation phase we plan to train and evaluate different models. One of us will use ResNet [2] while the other member will use GoogleNet [8]. Literature review will also be split accordingly.

## References

[1] Human Rights First. Human trafficking by the numbers. https://www.humanrightsfirst.org/resource/human-trafficking-numbers. 1

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016. 1

[3] Kaggle. Hotel-id to combat human trafficking 2021 - fgvc8. https://www.kaggle.com/c/hotel-id-2021-fgvc8/, 2021. 1

[4] Rashmi Kamath, Gregory Rolwes, Samuel Black, and Abby Stylianou. The 2021 hotel-id to combat human trafficking competition dataset. 2021. 1

[5] Alina Matei, Andreea Glavan, and Estefania Talavera. Deep learning for scene recognition from visual data: a survey, 2020. 1

[6] Abby Stylianou, Rashmi Kamath, Richard Souvenir, and Robert Pless. Hotel-id 2021. https://sites.google.com/view/fgvc8/competitions/hotel-id2021, 2021. 1

[7] Abby Stylianou, Hong Xuan, Maya Shende, Jonathan Brandt, Richard Souvenir, and Robert Pless. Hotels-50k: A global hotel recognition dataset. In *AAAI*, 2019. 1

[8] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions, 2014. 1

[9] Mingxing Tan and Quoc V. Le. Efficientnetv2: Smaller models and faster training, 2021. 1

[10] Hong Xuan, Abby Stylianou, Xiaotong Liu, and Robert Pless. Hard negative examples are hard, but useful, 2021. 1