SeatScore: A Dynamic Deep Learning Approach to Determine How Much You Deserve a Seat

Kushagra Karar (2024CSY7554); Saurabh Barthwal (2024CSY7562); Vedant Sharma (2024CSY7553)

Abstract:

In crowded public transport, the decision about who should get a seat is usually based on personal judgment. This can lead to unfair situations: someone who looks tired might still be standing while others stay seated comfortably. Our project addresses this problem by building a smart system that gives each passenger a "seat score" based on how tired they are, their physical condition and other important details. This makes seat sharing fairer and more thoughtful.

1 Introduction

Public transport often depends on social habits, not clear rules to decide who gets a seat. This can result in unfair seat sharing. Our goal is to build a system that gives people a "seat-deservingness" score using different inputs such as fatigue, emotions, age, and gender.

For fatigue detection, we combine existing EAR (Eye Aspect Ratio) solutions [1] (20% weightage) with our custom YOLOv11 medium model [2] trained on COCO dataset [3] (80% weightage). Section 2.2.1 elaborates on our custom model development and weighting rationale.

After evaluating multiple approaches for determining age and gender, we selected DeepFace [4] for optimal accuracy. Section 2.2.2 details the DeepFace architecture and implementation.

All extracted features are normalized into feature vectors paired with the corresponding seat score labels, as explained in Section 2.3. The complete dataset is then trained using a regression tree model to produce informed seating priority decisions based on physical and contextual data.

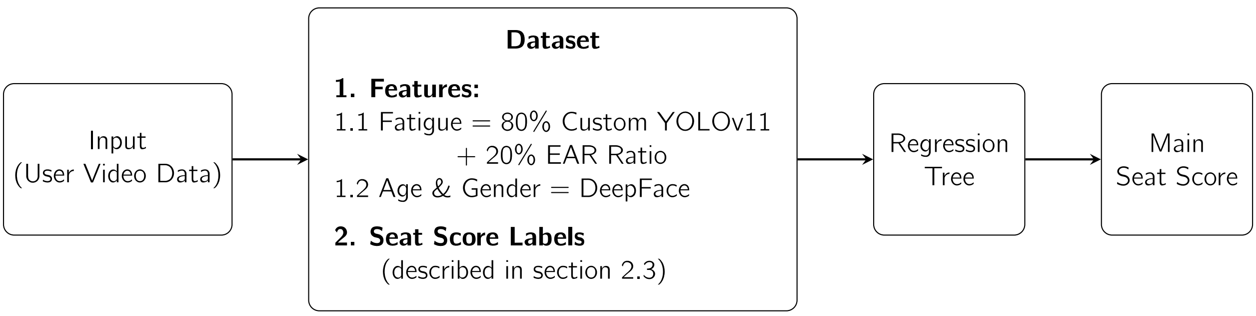

Figure 1: Methodology Overview and Pipeline

2 Dataset – Feature Extraction and Label Assignment

The most crucial step of the project is the creation of the dataset because choosing relevant features, extracting those features from the user video data and assigning proper accurate labels to it is the most crucial and most challenging task involved in the project. After we have a dataset ready, all we need to do is train a simple regression tree. The various steps in the creation of the dataset are described below.

2.1 Data points: Synthetic and Real World

Since we need a significant amount of video data to train the model, getting all real life video data is an impractical task within the time frame. So, we took **video data of 15 volunteers**, and these served as real life data points, and rest of the data points are synthetic in nature. Note that a data point here refers to the set of features and the corresponding seat score label.

2.2 Feature Extraction

Our seat score calculation uses several important features from passenger images, each chosen for specific reasons:

- **Fatigue Detection:** How tired a passenger looks helps determine their seating needs. A tired passenger might need different seat adjustments than an alert one.
- **Age and Gender:** These factors affect physical comfort since people of different ages and genders have varying body structures and comfort preferences.
- **Emotion:** How a passenger feels can indicate their current comfort level and satisfaction with their seat. It should be noted that as we continued with the project, we came to a consensus that emotion is not an important criterion for public transport seating, so it was scrapped as a feature in the final feature set.

2.2.1 Fatigue Detection – Custom YOLO Model + Pre-Trained EAR Models

Most existing fatigue detection systems are designed for driving and focus on limited features such as eye closure or yawning. While effective in some contexts, these can produce false negatives in public transport scenarios (e.g., speaking misclassified as yawning). Despite the limitations, we retain the Eye Aspect Ratio (EAR) method for its precision in detecting eye openness and assign it a 20% weight. It uses Dlib's 68-point facial landmark detector [5, 6] to compute EAR, flagging fatigue when the ratio drops below a threshold.

To address EAR's limitations, we trained a **custom fatigue detection model** using the YOLOv11m architecture. Initially trained on COCO [3], the model was fine-tuned to classify faces as `Awake` or `Drowsy`.

Custom model creation:

- 15 volunteers recorded two 60-second videos: one in an awake state and another in a drowsy state.
- We annotated facial bounding boxes for compatibility with the YOLO format: [class_id, center-x, center-y, width, height].
- Initially, annotations were manual. Later, we automated face detection using FaceNet [7].
- These labeled images were used to fine-tune the YOLOv11m model on fatigue detection.

The final fatigue score is a weighted combination:

$$\text{Fatigue Score} = 0.2 \times \text{EAR} + 0.8 \times \text{YOLO}$$

This hybrid approach balances eye-specific precision with general facial fatigue features, improving seat score reliability.

2.2.2 Age and Gender Detection

We experimented with multiple models for age and gender detection, including AgeNet[8] and SSR-Net [9]. Some of these models were pre-trained, while others we trained on public datasets. After thorough testing, we found that **DeepFace** provided the most accurate results for our specific subjects.

DeepFace uses a model based VGG-Face model trained on a combination of datasets including IMDB-WIKI [10]. This model, often referred to as DeepFace's Age and Gender Model, was adapted from the famous research paper [1]. The model's robust performance across diverse facial characteristics made it the optimal choice for our seat scoring system.

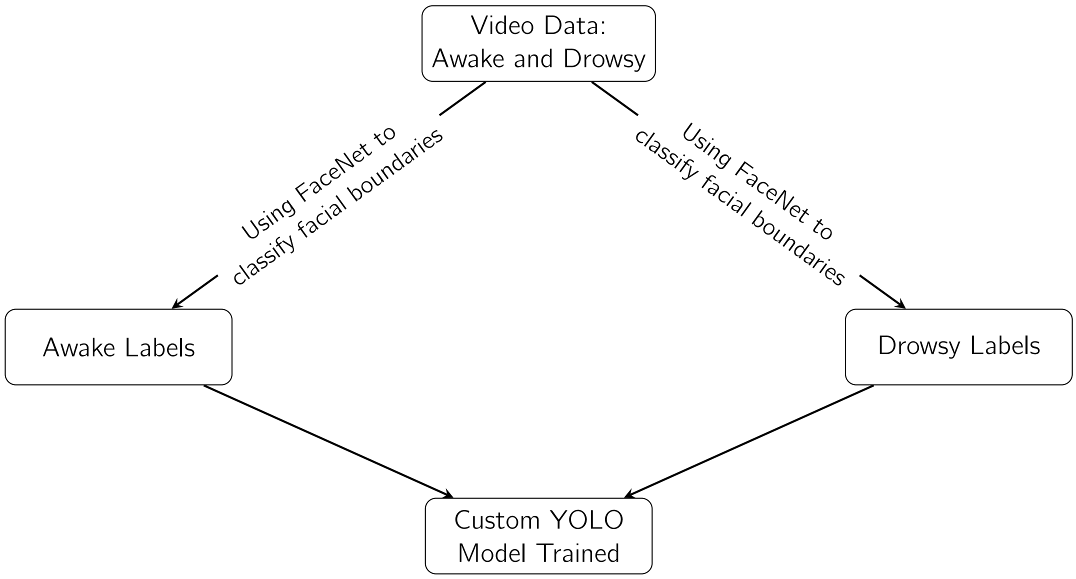

Figure 2: Training Process for Custom YOLO-based Fatigue Detection Model

2.2.3 Processing the Features

After normalising the features on a scale of 0-100, our final dataset has this structure:

$$[\text{fatigue}_{norm}, \text{age}_{norm}, \text{gender}_{norm},] \tag{1}$$

Each set of features is paired with a corresponding seat score label.

2.3 Generating Ground Truth Labels

Generating accurate ground truth labels is the most challenging part of our project. Unlike other machine learning problems, there is no standard way to create seat score labels from our features. This section explains our approach to this challenge.

2.3.1 Our Step-by-Step Approach

We developed a practical method that starts with basic labels and improves them over time:

- **Starting with Edge Cases:** We identified key scenarios that represent different combinations of features. These edge cases help us establish initial scoring guidelines. Few of the interesting edge cases to be considered are (For Gender, Male is 1 and Female is 2). The seat score reflects a blend

Fatigue	Age	Gender	Seat Score
10	60	1	76
85	40	1	75
20	35	2	44
87	57	1	90
60	22	2	63

Table 1: Synthetic Edge Cases: Feature Values and Corresponding Seat Scores

of fatigue, age, and gender. For example, a highly fatigued younger person (Fatigue = 85, Age = 40) receives a score similar to an older, less fatigued person (Fatigue = 10, Age = 60), showing how multiple factors balance each other to determine seat priority.

- **Estimate Based Method for Seat Score Computation**

  To compute seat scores, we developed an estimate based approach that uses our team's knowledge and judgment. Instead of randomly assigning importance to different features, we carefully analyzed specific examples that represent extreme or notable cases.

For these selected cases, we determined the most appropriate seat scores by considering how fatigue levels, age, gender, and other features work together to affect seating comfort. We chose these examples to cover a wide range of passenger types and conditions.

With these carefully scored examples as guides, our model could then learn the patterns in our scoring logic and apply similar reasoning to new data.

- **Continuous Improvement:** As we train the model with these initial labels, we regularly check its performance and adjust our approach based on feedback.

2.3.2 Ensuring Fair Labels

To avoid personal bias and keep the process transparent, we set up a team-based review system:

- **Shared Responsibility:** For each test case, two team members independently create labels using our criteria.
- **Independent Review:** A third member of the team reviews both proposed labels and makes the final decision if the absolute difference between the labels of the first two team members is greater than 15. Otherwise, we simply take the average of the labels by the two team members. This helps catch and correct for any inconsistencies or biases.

2.3.3 Practical Implementation

Our approach recognizes that label generation for this problem is inherently based on human judgment rather than absolute truth. By starting with carefully selected examples and using a team-based review process, we create the most reliable labels possible. This peer review system ensures our ground truth labels represent team consensus rather than individual opinion, providing a solid foundation for our seat scoring model despite the lack of established standards.

3 Model Training and Results

After dataset creation, we train a regression tree (DecisionTreeRegressor) to learn the mapping from features to seat scores.

3.1 Training Procedure

The normalized feature vectors and heuristic labels are split into training and validation sets. We fit a regression tree, tuning its depth and leaf parameters through cross-validation to minimize the mean squared error.

3.2 Inference Pipeline

1. **Fatigue Computation:** YOLO-based drowsiness detection and eye-aspect-ratio (EAR) tracking to obtain a final fatigue percentage by taking the percentage of frames that are higher than threshold for the last 60s. For eg, if 70% of the frames in last 60s are drowsy, the final YOLO fatigue would be 0.7
2. **Face Analysis:** DeepFace estimates age and gender periodically.
3. **Score Prediction:** The function `predict_seatscore(age, gender, fatigue)` applies the trained regression tree to output the seat score.
4. **Display:** Results (YOLO fatigue, EAR fatigue, age, gender, seat score) are overlaid in the video feed.

3.3 Model Evaluation Results

- **Fatigue Detection (Custom-trained YOLO model, dataset 1796 images, 10% test) :**
  - Precision: 0.916
  - Recall: 0.931
  - **Augmentations that Improved the Model:**

* Mosaic Augmentation: Enabled (value = 1.0)
* Horizontal Flip (Left-Right): 50% probability (fliplr = 0.5)
* HSV Color Augmentation: hue shift = 0.015, saturation shift = 0.7, value shift = 0.4
* Translation Augmentation: Up to 20% (translate = 0.2)
* Scaling Augmentation: Up to 60% (scale = 0.6)

- **Seat Score Prediction (Regression Tree model, dataset 100 data points 30 real world and 70 synthetic, 10% test):**
  - Mean Absolute Relative Deviation (MARD): 3.8
  - Mean Squared Error (MSE): 9.4

3.4 Example Outputs

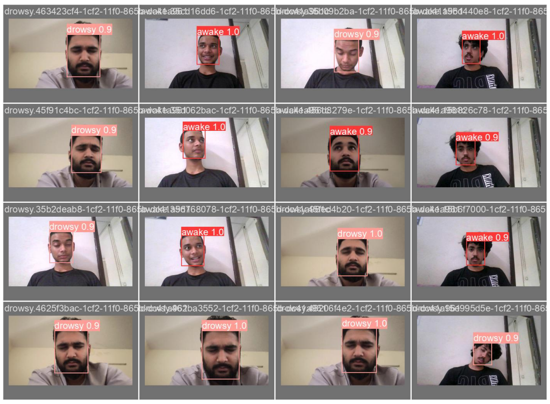

Figure 3: YOLO-based Fatigue Detection

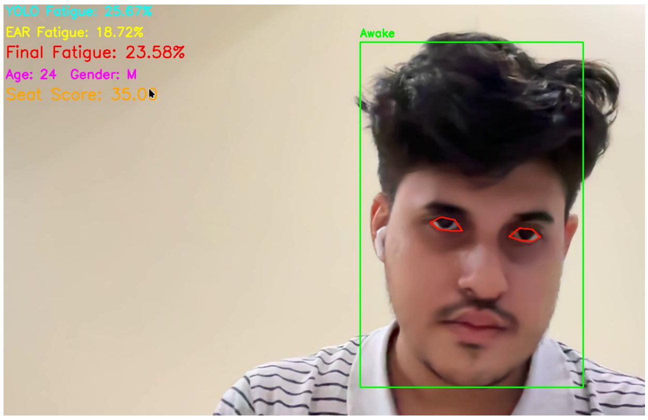

Figure 4: Final SeatScore Display

4 Conclusion

SeatScore is the first intelligent system to decide and help assign seats to people in public transport. It combines measures of fatigue, age, and gender, then uses a simple regression tree to compute a clear seat score for each person. In practice, SeatScore gives sensible and fair results by intelligently mixing these features. Currently, seat scores are based on heuristic labels, which can be improved over time with more data and a refined labeling process to make SeatScore even more accurate and fair.

References

[1] G. Levi and T. Hassner, "Age and gender classification using convolutional neural networks," in *2015 IEEE CVPRW*, Boston, MA, 2015, pp. 34-42.

[2] Ultralytics, "YOLOv11 Medium," 2023. [Online]. Available: `https://github.com/ultralytics/ultralytics`.

[3] T. Lin et al., "Microsoft COCO: Common Objects in Context," in *ECCV*, 2014, pp. 740-755.

[4] S.I. Serengil and A. Ozpinar, "LightFace: A Hybrid Deep Face Recognition Framework," in *ASYU*, 2020, pp. 23-27.

[5] D.E. King, "Dlib-ml: A Machine Learning Toolkit," *JMLR*, vol. 10, pp. 1755-1758, 2009.

[6] C. Sagonas et al., "300 Faces in-the-Wild Challenge," in *ICCVW*, 2013, pp. 397-403.

[7] F. Schroff et al., "FaceNet: A Unified Embedding for Face Recognition and Clustering," in *CVPR*, 2015, pp. 815-823.

[8] Z. Qawaqneh et al., "Deep CNN for Age Estimation based on VGG-Face Model," *arXiv:1709.01664*, 2017.

[9] T.-Y. Yang et al., "SSR-Net: A Compact Soft Stagewise Regression Network for Age Estimation," in *IJCAI*, 2018, pp. 1078-1084.

[10] R. Rothe, R. Timofte, and L. Van Gool, "DEX: Deep EXpectation of apparent age from a single image," in *IEEE International Conference on Computer Vision Workshops (ICCVW)*, 2015, pp. 10-15. [Online]. Available: `https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/`.