

KAGGLE SYSTEMS DESIGNS – NFL CONTACT DETECTION

ANDERSON DAVID ARENAS GUTIERREZ - 20231020030
NICOLAS FELIPE PULIDO SUAREZ – 20231020045
NAHIN JOSE PEÑARANDA MEJIA – 20231020032



UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

INTRODUCTION

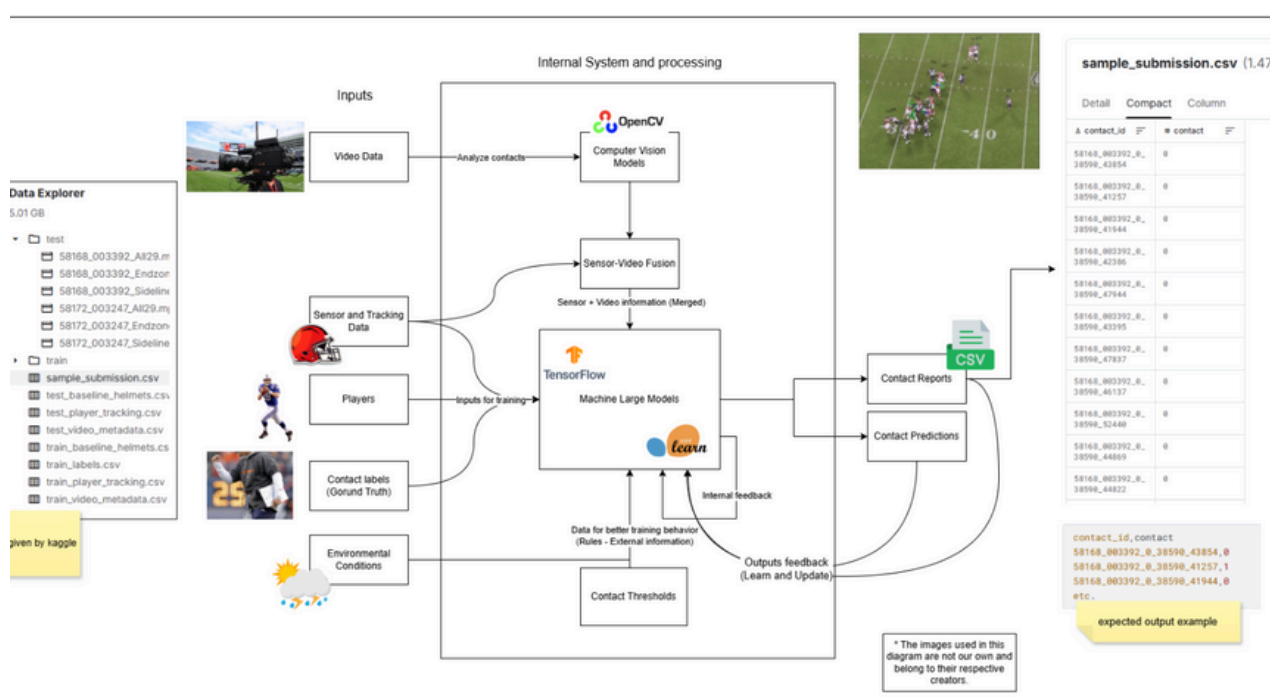
Detecting player contact events in NFL games is a complex challenge due to the combination of video and sensor data, often plagued by low video resolution, synchronization issues, and label inaccuracies. Existing solutions struggle with real-time demands, imbalanced datasets, and environmental variability. Prior approaches, like those from winning Kaggle teams, use GPU-heavy architectures that limit generalizability. Our challenge was to design a lightweight, CPU-friendly system that maintains precision and interpretability without requiring high-end resources

GOAL

Research Question: How can we build a lightweight, real-time NFL player contact detection system using CPU-optimized tools that maintain accuracy and scalability?

Expected Product: A modular, CPU-efficient pipeline capable of ingesting, processing, and classifying contact events from multimodal game data.

PROPOSED SOLUTION



The architecture is modular, enabling flexibility and maintainability. It uses Python-based tools such as TensorFlow (MobileNet SSD), OpenCV, and Scikit-learn. Data from sensors and video streams is synchronized, features are extracted, and contact events are predicted using a RandomForest model. Monitoring tools feed performance metrics back into the model pipeline for potential retraining.

EXPERIMENTS

We will conduct experiments on a small subset of the NFL contact detection dataset.

- Setup: The data will be downsampled (1 FPS, 640×480 resolution), processed in batches, and executed using Jupyter Notebooks on CPU.
- Model: We will use MobileNet SSD for detection, OpenCV CSRT for tracking, and RandomForestClassifier for contact prediction.
- Validation: A 5-fold cross-validation will be used to evaluate model robustness.

RESULTS

In process.

CONCLUSIONS

Our approach focuses on building a modular, efficient, and flexible system capable of handling the complexity of real-game scenarios. Systems thinking has played a crucial role in shaping the architecture and informing key decisions, helping us pinpoint important challenges and areas for improvement. Moving forward, we plan to expand testing, enhance performance, and investigate deployment options for both real-time and batch processing in sports analytics.

BIBLIOGRAPHY

- <https://www.kaggle.com/competitions/nfl-player-contact-detection>
- <https://github.com/ItzNxhin/SAD---Nahin-Nicolas-and-Anderson>