# AniNex – The 6th Workshop on Next Generation Computer Animation Techniques

In conjunction with the 38th International Conference on Computer Animation and Social Agents (CASA 2025)

2–4 June 2025
Strasbourg, France

# Accepted Posters

Jian Chang, Bournemouth University, UK
Anil Bas, Bournemouth University, UK
Shihui Guo, Xiamen University, China
Jian Jun Zhang, Bournemouth University, UK

This workshop is partially funded by the EU H2020 Marie Skłodowska-Curie COFUND Scheme (CfACTs 900025) and the Bournemouth University Research Capacity Transformation Fund.

# Immersive Video Game Experience through a Sequential Response to In-Game Context NPC Conversation

Michael Adjeisah, Eshani Kawshika Fernando, and Jian Chang*
National Centre for Computer Animation, Bournemouth University, UK
jchang@bournemouth.ac.uk

## Abstract

We present a game companion that dynamically personalizes its responses based on players' emotional shifts in real-time, enhancing user immersion in games. The model integrates sequential data, taking advantage of the dynamic nature of streams to benefit from player information, spoken lines, and the context in the game. We engage LLMs' tool-calling capabilities to extract vital information from memory and recognize potential constraints for accurate reasoning, for NPC conversation.

## 1 Introduction

Non-player characters (NPCs) perform various roles and functions, such as quest-givers, vendors, shopkeepers, allies and companions, enemies, adversaries, and supporting characters [1]. They hold critical information about games to enrich the player's experience by adding to the atmosphere of the game world [2]. Developing a virtual agent that engages in meaningful and contextually appropriate dialogue with players throughout gameplay can provide essential information, plot developments, and mythology that enrich the game world and drive player immersion. This work presents a game companion capable of personalizing its responses in real-time based on players' emotional shifts to enhance immersion within broader gameplay interactions. The approach employs a novel sentinel mechanism to foster emotional connection and manage contextual aspects of the game for a seamless narrative flow.

**Research Gap**. Most studies present various applications of LLMs *in* and *for* within the broader ecosystem of games, as well as the different roles they can play within a game [3]. Their theoretical frameworks are extensively based on emulating human behavior to play games at a human or near-human level [4]. However, an interactive companion intended to enrich or guide the player experience without competing with or altering the game mechanics [5] remains relatively less explored.

## 2 Framework

We detail the pipeline process and illustrate the core modules in Figure 1. We divided the task into three distinct modules: the sentinel mechanism (*SenGent*), a memory capability agent, and a chat planning agent. We leverage the advanced capabilities of the WhisperX[1] model for speech-to-text transcription [6].

## 3 Experiment

As displayed in Table 1, the companion benefits from the supervised fine-tuning for its overall performance compared to the baselines with Few-Shot prompting and role-play models. Specifically, the game companion surpasses the

*Corresponding author

[1] https://github.com/m-bain/whisperX

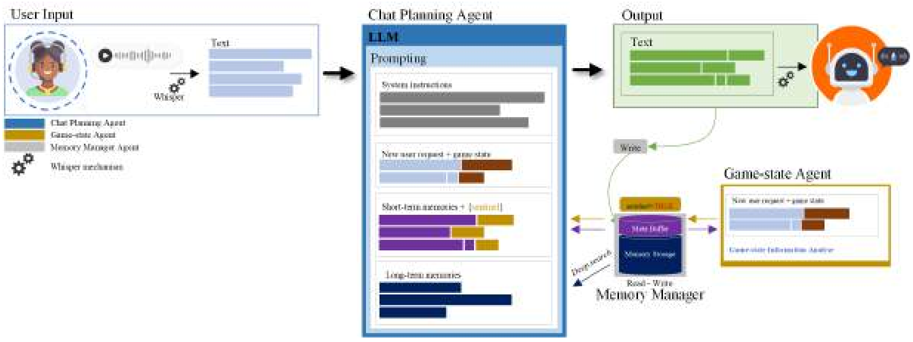



Figure 1: Conceptual framework of the game companion.

| | Model | R-L | R-WE |
|---|---|---|---|
| Base | Llama 2 7B | 0.327 | 0.465 |
| | Llama 2 13B | 0.383 | 0.48 |
| | Llama 2 70B | 0.516 | 0.599 |
| | Llama 3 8B | 0.391 | 0.489 |
| | Llama 3 70B | 0.523 | 0.621 |
| Roleplay | Llama-3-8B | 0.553 | 0.682 |
| | Mythalion 13B | 0.51 | 0.629 |
| Fine-tuned | Llama 2 70B | 0.783 | 0.855 |
| | Llama 3 70B | **0.807** | **0.883** |

Table 1: The overall model performance is based on Llama base, roleplay, and fine-tuned models. The companion response is scored against the human-in-the-loop (gold) responses. We use the ROUGE-L and ROUGE-WE matrices.

base and role-play models by 26.5% and 20.1% on ROUGE-WE scores, respectively.

## 4 Conclusion

The game companion sequentially captures players' emotional shifts over time; responses are more naturalistic and human-like with appropriate conversational cues, pauses, and sighs; and utterances remain faithful to the dynamics of in-game context, supporting narrative continuity.

## Acknowledgements

This work was supported in part by the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie, grant agreement No. 900025.

## References

[1] Roberto Gallotta, Graham Todd, Marvin Zammit, Sam Earle, Antonios Liapis, Julian Togelius, and Georgios N Yannakakis. Large language models and games: A survey and roadmap. *arXiv preprint arXiv:2402.18659*, 2024.

[2] Maximilian Croissant, Guy Schofield, and Cade McCall. Emotion design for video games: A framework for affective interactivity. *ACM Games: Research and Practice*, 1(3):1–24, 2023.

[3] Penny Sweetser. Large language models and video games: A preliminary scoping review. In *Proceedings of the 6th ACM Conference on Conversational User Interfaces*, pages 1–8, 2024.

[4] Austen Liao, Nicholas Tomlin, and Dan Klein. Efficacy of language model self-play in non-zero-sum games. *arXiv preprint arXiv:2406.18872*, 2024.

[5] Yuqian Sun, Zhouyi Li, Ke Fang, Chang Hee Lee, and Ali Asadipour. Language as reality: a co-creative storytelling game experience in 1001 nights using generative ai. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, volume 19, pages 425–434, 2023.

[6] Max Bain, Jaesung Huh, Tengda Han, and Andrew Zisserman. Whisperx: Time-accurate speech transcription of long-form audio. *INTERSPEECH 2023*, 2023.

# Photorealistic 3D Head Reconstruction via 2D Gaussians

Anil Bas
Bournemouth University
abas@bournemouth.ac.uk

Oleg Fryazinov
Bournemouth University
ofryazinov@bournemouth.ac.uk

Xiaosong Yang
Bournemouth University
xyang@bournemouth.ac.uk

Callum Rex Reid
Visualskies
cal@visualskies.com

## Abstract

Radiance fields have significantly enhanced novel view synthesis and 3D reconstruction techniques. Recently, 3D Gaussian Splatting [1] has proven to be a milestone, offering compact and differentiable volumetric primitives that enable photorealistic rendering with efficient training. However, while 3DGS excels at view synthesis, extracting high-quality surface meshes remains a challenge due to the loose geometric alignment of ellipsoidal Gaussians [2]. This issue is particularly pronounced in high-detail domains like human head reconstruction, where capturing subtle anatomical features and precise surface geometry is essential.

Following [3, 4], we represent 3D Gaussians as 2D ellipses (disks) to address surface misalignment, which improves reconstruction quality. Furthermore, our approach allows for efficient optimisation through multiple loss functions and uses Poisson reconstruction [5] to extract photorealistic meshes. We apply our method to the NeRSemble dataset [6], preprocessing data similar to GaussianAvatars [7]. We provide known camera parameters, background masks, and normal maps to support optimisation and accurate surface alignment. Despite using only 16 input images per subject, our method successfully reconstructs high-fidelity, textured meshes, as shown in Figure 1. Unlike previous studies, our method is tailored for mesh extraction rather than view synthesis only, and does not require morphable face models (e.g. FLAME or BFM) or facial landmark detection, tracking, and annotation processes.

Figure 1: Comparison of mesh reconstruction results. Rows 1 & 3: 2DGS [4] and Rows 2 & 4: Ours, which provides finer detail and richer surfaces.

**Keywords:** Gaussian splatting, radiance fields, 3D head reconstruction, mesh extraction

## Acknowledgements

The results incorporated in this work have received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 900025.

## References


[1] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3D Gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 2023.

[2] Antoine Guédon and Vincent Lepetit. SuGaR: Surface-aligned Gaussian splatting for efficient 3D mesh reconstruction and high-quality mesh rendering. In *Proc. ECCV*, 2024.

[3] Pinxuan Dai, Jiamin Xu, Wenxiang Xie, Xinguo Liu, Huamin Wang, and Weiwei Xu. High-quality surface reconstruction using Gaussian surfels. In *Proc. SIGGRAPH*, 2024.

[4] Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2D Gaussian splatting for geometrically accurate radiance fields. In *Proc. SIGGRAPH*, 2024.

[5] Michael Kazhdan and Hugues Hoppe. Screened Poisson surface reconstruction. *ACM Trans. Graph.*, 2013.

[6] Tobias Kirschstein, Shenhan Qian, Simon Giebenhain, Tim Walter, and Matthias Nießner. NeRSemble: Multi-view radiance field reconstruction of human heads. *ACM Trans. Graph.*, 2023.

[7] Shenhan Qian, Tobias Kirschstein, Liam Schoneveld, Davide Davoli, Simon Giebenhain, and Matthias Nießner. GaussianAvatars: Photorealistic head avatars with rigged 3D Gaussians. In *Proc. CVPR*, 2024.

# LLM-Powered VR Nursing Training for Dynamic Risk Assessment

Ehtzaz Chaudhry[1], Cliff Kilgore[2], Michele Board[2], Jian Chang[1], Michael Board[1] and Jian Jun Zhang[1]

[1]National Centre for Computer Animation, Bournemouth University, Fern Barrow, Poole, Dorset, BH12 5BB, United Kingdom
[2]Department For Nursing Science, Bournemouth University, Bournemouth Gateway Building, St Pauls Lane, Bournemouth, BH8 8GP, United Kingdom
echaudhry@bournemouth.ac.uk

## Abstract

This project presents a Virtual Reality (VR) training simulation system powered by Large Language Models (LLMs), aimed at improving the skills of nurses by enabling them to practice various procedures through realistic patient interactions and enhance risk management skills with in patient home environments. In this poster, we present our work focused on developing a VR-based nurse training simulation system integrated with Open AI. The LLM architecture enables learners to interact with realistic patient scenarios, ask questions, make decisions, receive context-relevant responses and accompanied by appropriate facial expressions. Our system lays the foundation for the future development of LLM-powered Non-Player Characters (NPCs) in VR.

**Keywords:** Virtual Reality, Simulation, Large Language Models, Nurse Training, Risk assessment.

## 1. Introduction

The integration of LLMs enables dynamic, natural conversations with virtual patients or NPCs, allowing simulations to incorporate complex dynamic risk assessments (DRA) and realistic emotional responses [1]. VR-based training simulations (LLMs), have significantly enhanced the ability of nurses and carers to practice a wide variety of procedures through realistic patient interactions [2]. Most commercial VR-based training simulations primarily focus on clinical and hospital settings, lacking sufficient training for nurses in conducting DRA within home care environments presents unique challenges for nurses, as the setting is often unpredictable and varies significantly from patient to patient [3]. Unlike controlled hospital environments, patient homes may have hazards such as poor lighting, cluttered spaces, pets, inadequate hygiene, or lack of proper medical equipment. Nurses must quickly evaluate these factors while also managing patient needs, ensuring infection control, and maintaining their own safety.

## 2. Integrating LLMs into VR-Based Nurse Training Systems

Our work focuses on developing a VR-based nurse training simulation system enhanced by LLMs. We have integrated Open AI in our VR system which embedded to support dynamic, interactive scenarios, enabling more realistic communication, decision-making, and adaptive learning for nurses in training. Our framework for the VR Nurse Training system includes a VR application developed in Unity for Oculus Quest 3 headsets. The VR app interacts with the Amazon Polly (AWS) API for text-to-speech and speech-to-text processing. Transcribed text from user is passed to Open AI language model as a text prompt. The response, enhanced through Retrieval

Augmented Generation (RAG), is then converted back to audio using Amazon Polly's text-to-speech service as shown in figure 1.

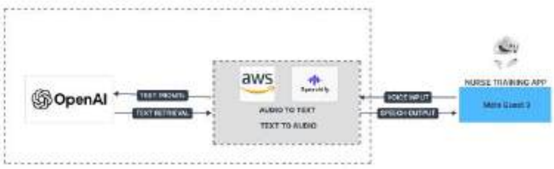



Figure 1: Framework of our VR Nurse Training System

We have developed several features to the VR nurse training system to improve realism. Both player-controlled (PC) and NPCs can engage in interactions through a dialogue system that displays text on-screen during conversations. NPC character have a lip-syncing and eye-blinking animations, while the VR environment allows players to interact by grasping objects and focusing their gaze on them. Additionally, NPC character also feature eye and head tracking, which respond dynamically to player movements, providing an immersive and highly realistic training experience.

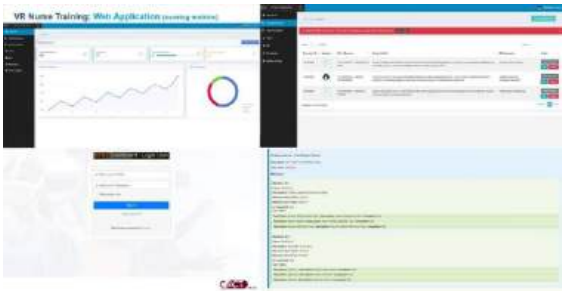



Figure 2: VR nurse training web application and user dashboard interface

## 3. Results & Conclusions

Our web-based Nurse Training application offers a comprehensive platform for both users and administrators to analyze performance through a structured scoring matrix. It features secure login, intuitive signup, and efficient password management. The "My Dashboard" presents a clear overview of training progress, including scenario completion, time spent in VR simulations, and performance metrics visualized through interactive charts (see Figure 2). The "Scenarios" section enables users to browse, filter, and search for training modules, each accompanied by briefings, milestones, objectives, and expected outcomes. Additionally, the "Session/Timeline" feature generates comprehensive logs and visual reports, offering detailed insights into user activity and performance metrics upon completing each VR scenario. We are conducting an evaluation to assess the effectiveness of our VR-based nurse training platform, along with the underlying LLM architecture. In subsequent phases, the evaluation will focus on (a) gathering surveys and feedback from nurses and carers to assess the simulation's realism, usability, and educational impact; and (b) analyzing LLM interactions, comparing the conversational capabilities of various LLM models, such as Gemini and Meta's LLaMA 4.

## Acknowledgments

This research is supported by Bournemouth University, UK as part of "The Centre for Applied Creative Technologies PLUS (CfACTs+)" project[1]. The VR scenarios are being developed as part of the BU Simulation Team's contribution, under the leadership of Charlotte Mutton.

## References

[1] Ferri-Molla, I., Linares-Pellicer, J. and Izquierdo-Domenech, J., 2024. Virtual reality and language models, a new frontier in learning.

[2] Woo, B., Huynh, T., Tang, A., Bui, N., Nguyen, G. and Tam, W., 2024. Transforming nursing with large language models: from concept to practice. European journal of cardiovascular nursing, 23(5), pp.549-552.

[3] Chaudhry, E., Chang, J. and Zhang, J.J., 2024, December. Virtual Reality Training System for Nurses: Evaluating Risks in Patients Home Environments. In The 19th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry (pp. 1-4).

[1] https://www.bournemouth.ac.uk/research/centres-institutes/centre-applied-creative-technologies/centre-applied-creative-technologies-plus-cfacts

# Using Large Language Models for Evaluation of Radiological Textual Reports

**Nicolay Rusnachenko**[1] **James Franklin**[2] **Theophilus Akudjedu**[2] **Neel Doshi**[3]
**Michael Board**[1] **Jian Chang**[1] **Jian Jun Zhang**[1]

[1] Bournemouth University, Centre for Applied Creative Technologies (CFACT+), Bournemouth, United Kingdom
[2] Bournemouth University, Institute of Medical Imaging and Visualisation, Bournemouth, United Kingdom
[3] University Hospital Southampton, Southampton, United Kingdom

## Abstract

The accurate assessment of practitioners' knowledge in the field of radiology represents a critical task that heavily relies on the expertise and availability of experts. Textual narratives represent a common approach of reporting patient conditions. To reduce the need for manual annotation by experts, the development of autonomous systems in radiology is essential. At present, Large Language Models (LLMs) represent a promising yet effective framework for developing autonomous explainable systems. Models of this type demonstrate a promising solution across various fields of Natural Language Processing (NLP), including the Information Retrieval (IR) domain. In this poster we propose methodology of adopting LLMs in radiological textual reports evaluation by leveraging IR capabilities of facts. We use $\mathrm{SN}_{\text{HCC-TCIA}}$[1] dataset of structured MR/CT image textual narratives on image acquisitions in liver cancer imaging. Our evaluation results provide insights into the relationship between model size and performance. Practical applications of our experimental findings are demonstrated through a web application which evaluates practitioner response on scoring a patient's liver condition.

**Keywords:** medical reports, textual narratives, natural language processing, education, demo

## 1 Introduction

Texts or textual narratives represent the most accessible format of reporting patient conditions by practitioners and radiological experts. Such a principle is suitable for evaluating practitioner knowledge about patient conditions with respect to the ground truth knowledge and facts (Figure 1, left). *Information Retrieval* (IR) refers to the domain of knowledge on finding and accessing relevant information from resources of various types (image, texts, etc.). In the context of reports, retrieving facts is important for reducing the influence of subjectivity in their assessment (*IR Module*, Figure 1). Adapting LLM for IR promises scalable solutions that become majorly distributed in radiology field and patient-centered systems [1, 2, 3]. However, the gap in IR application *for practitioner-oriented systems* is still poorly covered.

## 2 Methodology

In this poster we propose a methodology for textual reports evaluation for the *task* ($T$) provided in description form and set of the expected closed-end answers ($C$). Given task description $T$, set of classes $C$, patient info $I$, practitioner textual narrative $A_p$, ground truth answer $A_g \in C$, list of explanation categories $S$. Our methodology exploits the chain-of-thought concept (CoT) for implicit IR [4] and facts extraction [5]. Our methodology relies on the single instance of the LLM model ($M$). We illustrate each step as an inference function $\mathbf{I}_M(*) \to O$, in which "$*$" refers to input parameters that forms prompt, and $O$ is a textual output[2]:

$$\mathbf{I}_M(T, A_p, C) \to A'_p \quad (1)$$

$$\mathbf{I}_M(T, A_p, A'_p) \to F_p \quad (2)$$

$$\mathbf{I}_M(T, I, A_g) \to F_g \quad (3)$$

$$\mathbf{I}_M(T, A'_p, A_g, F_p, F_g, S) \to E \quad (4)$$

Where, steps 1-4 denote: answer interpretation ($A'_p \in C$) of the practitioner narrative, facts extraction ($F_p$, $F_g$) for practitioner narrative and ground truth, and the result *explanation* ($E$) by relying on categories $S$.

[1] `https://github.com/nicolay-r/Reasoning-for-Radiology-Report-Evaluation`

[2] Due to confidentiality agreements, the specific formulation details cannot be disclosed.

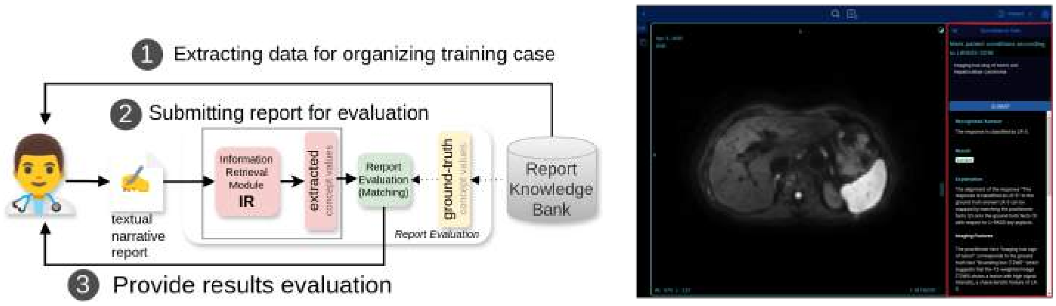



Figure 1: *left*: The workflow of the training session organization for radiologists; *right*: NLP component (highlighted red box) of the demo system showcasing a task for scoring patient's liver condition (scale 1-5) according to LiRADS-2018 framework

Table 1: $F1$-based evaluation results on $\mathrm{SN}_{\text{HCC-TCIA}}$ in zero-shot learning mode

| Method | WEIGHTING | TIMING | PLANE |
|---|---|---|---|
| ChatGPT-4 (1.76T) | 0.90 | 0.99 | 0.98 |
| ChatGPT-3.5 (175B) | 0.93 | 0.97 | 0.99 |
| LLaMA-3-70B | 0.93 | 0.98 | 0.99 |
| LLaMA-3-8B | 0.85 | 0.65 | 0.97 |
| LLaMA-3.2-3B | 0.68 | 0.72 | 0.36 |
| LLaMA-3.2-1B | 0.32 | 0.27 | 0.18 |
| RANDOM | 0.19 | 0.26 | 0.29 |

# 3 Experiments

We consider the simplified problem of *facts extraction*: given text ($t$) and mentioned fact ($f$), extract the value of this fact with respect to the set of classes ($C$). We use $\mathrm{SN}_{\text{HCC-TCIA}}$ resource[1], which yields short-text *studies* and *series descriptions* narratives of liver-related cancer-targeted resources. $\mathrm{SN}_{\text{HCC-TCIA}}$ provides annotation of image series facts of the textual narratives for the liver. The following set of facts considered: image WEIGHTING type ($C_{weighting} = \{$T1, T2, DWI, ADC$\}$), contrast injection TIMING ($C_{timing} = \{$pre, art, port, del$\}$), and PLANE type ($C_{plane} = \{$axial, sagittal, coronal$\}$). Statistics of the amount of series in $\mathrm{SN}_{\text{HCC-TCIA}}$ is as follows: 467 (WEIGHTING), 421 (TIMING), 973 (PLANE).

In the context of selected models ($M$), we experiment with evaluation of LLM models in zero-shot mode[3]. We use top-tier LLMs of *closed transparency* type (ChatGPT series) and *open transparency* (LLaMA-3 series [6]) of different sizes: 1B, 3B, 7B and 70B. Table 1 illustrates the obtained results (F1-measure) for each selected model, including baseline (RANDOM). Our observations are two fold:

**Finding 1.** The small versions of the open-access LLaMA-3.2 models, with 1B and 3B parameters, represent the least preferable scale for handling textual narratives. It is suggested to strengthen reasoning capabilities of models at that scale by applying more advanced techniques that are beyond the scope of the present studies.

**Finding 2.** The out-of-the-box 8B and 70B scaled LLaMA-3 models represent a decent solution for automatic handling textual narratives using the proposed methodology in Section 2.

[3]**Utilized Prompt:** Given the study and series description $t$ in the domain of liver surveillance, to which fact $f$ the series description is related to? Choose one from: $C$

# 4 Demo

Figure 1 (right) represent a screenshot of the NLP component[4] for scoring patient's liver condition according to LiRADS-2018 framework. This demo illustrates application of the Methodology (Section 2) towards submitted textual narrative. In this example practitioner submit the response "Imaging has sing of tumor and hepatocelluar carcinoma", which been classified as LR-5 (definite case of cancer).

# References

[1] Davide Caffagni et. al. The (r)evolution of multimodal large language models: A survey, 2024.

[2] Khaled Saab et. al. Capabilities of gemini models in medicine, 2024.

[3] Sucheng Ren et. al. Medical vision generalist: Unifying medical imaging tasks in context, 2024.

[4] Hao Fei, Bobo Li, Qian Liu, Lidong Bing, Fei Li, and Tat-Seng Chua. Reasoning implicit sentiment with chain-of-thought prompting, 2023.

[5] Stella Li and et. al. Mediq: Question-asking llms and a benchmark for reliable interactive clinical reasoning. *Advances in Neural Information Processing Systems*, 37:28858–28888, 2024.

[6] Abhimanyu Dubey et al. The llama 3 herd of models, 2024.

[4]Powered by LLaMA-3-70B model.

# AssetMask: Mask R-CNN-based approach for Asset detection in railroad track health monitoring

Aradhya Saini1, Hammadi Nait-Charif1, Jian Chang1, Tolga Varol2

[1]Bournemouth University, Fern Barrow, Poole, Dorset, BH12 5BB, United Kingdom
[2]DriveTrust: RailEyes

## Abstract

Railroad track safety is of utmost importance as the negligence may lead to loss of life and property. Track assets are inspected for the smooth functioning of the railroad track environment. Mask R-CNN-based framework is considered useful for the detection of track assets namely ‘Train’, ’Track’, ’Vegetation’,’ Sign’, ’Person’. The monitoring and safety of these aforementioned assets is useful for implementing the track safety in the railroad environment. The asset management framework is developed based on this purpose.

**Keywords:** Assets, Railroad track, Mask R-CNN, monitoring.

## 1. Introduction

The advancements in the technology in the 21st century has led to object detection garnering a lot of attention in the field of machine learning and computer vision. Railroad track safety is of utmost importance as derailments can lead to loss of life and property. For the purpose of asset monitoring in the railroad environment the classes in which the images are categorized are listed as ‘Train’, ‘Track’, ‘Vegetation’, ’Sign’, ’Person’. These assets have been as shown in Figure 1.

Mask R-CNN method has been developed for the instance segmentation purpose. The masks are applied in varied colors on different aforementioned asset classes. Multiple instances, considerably tiny, are also detected in the images captured from afar. The detections are observed to be detected in out of distribution images also.

**Figure 1Assets in a railroad track environment.**

## 2. Literature Review

Various assets such as ‘ Construction’, ‘Power Junction’, ‘ Cement slab’ have been identified using YOLO [1]. A risk informed asset management framework has been developed in [2]. Asset management model studying degradation is discussed in [3]. It is observed that the Mask R-CNN method has been observed to detect multiple instances smoothly and hence is a candidate for consideration during development of framework.

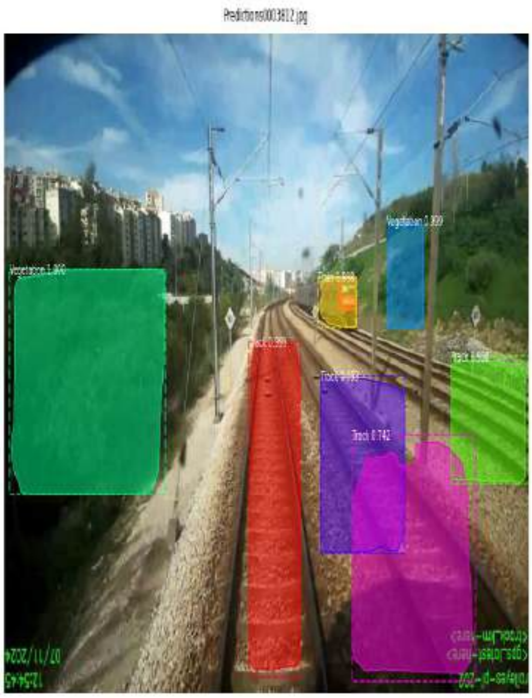
## 3. Proposed Approach

As per the literature review in Section 2, the Mask R-CNN method is considered appropriate for object detection. The frames are extracted from the video with frame width and height of 1280* 720 and 1920*1088 respectively. The loss function is considered as in Eq. 1.

loss = w1 × rpn_class_loss + w2 × rpn_bbox_loss + w3 × mrcnn_class_loss + w4

**(a)**

**(b)**



**(c)**

**(d)**

**Figure 2 Assets in a railroad environment**

× mrcnn_bbox_loss + w5 × mrcnn_mask_loss (1)

## 4. Conclusion

The detections are well observed in the railroad environment in the train as well as the test images. All the classes with most of the multiple instances can be observed in the test images. Mask R-CNN can concluded as a successful multiple instance detection technique.


## Acknowledgements

This research work is supported by the project which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 900025.


## References


[1] A. Saini, K. KG, K. Sriram, D. Singh and K. P. Singh, "Machine Learning Approach for Detection of Track Assets for Railroad Health Monitoring with Drone Images," IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 2022, pp. 4891-4894, doi: 10.1109/IGARSS46834.2022.9883490.

[2]Sasidharan, Manu, Michael Peter Nicholas Burrow, Gurmel Singh Ghataora, and Rishi Marathu. "A risk-informed decision support tool for the strategic asset management of railway track infrastructure." *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 236, no. 2 (2022): 183-197.

[3] Clarke, Steve. "Railway track asset management modelling." PhD diss., University of Nottingham, 2021.

[4] Gbadamosi, Abdul-Quayyum, Lukumon O. Oyedele, Juan Manuel Davila Delgado, Habeeb Kusimo, Lukman Akanbi, Oladimeji Olawale, and Naimah Muhammed-Yakubu. "IoT for predictive assets monitoring and maintenance: An implementation strategy for the UK rail industry." *Automation in Construction* 122 (2021): 103486.