EVS Internships 2026-2027

Generic correspondences matching between soccer images for VOL

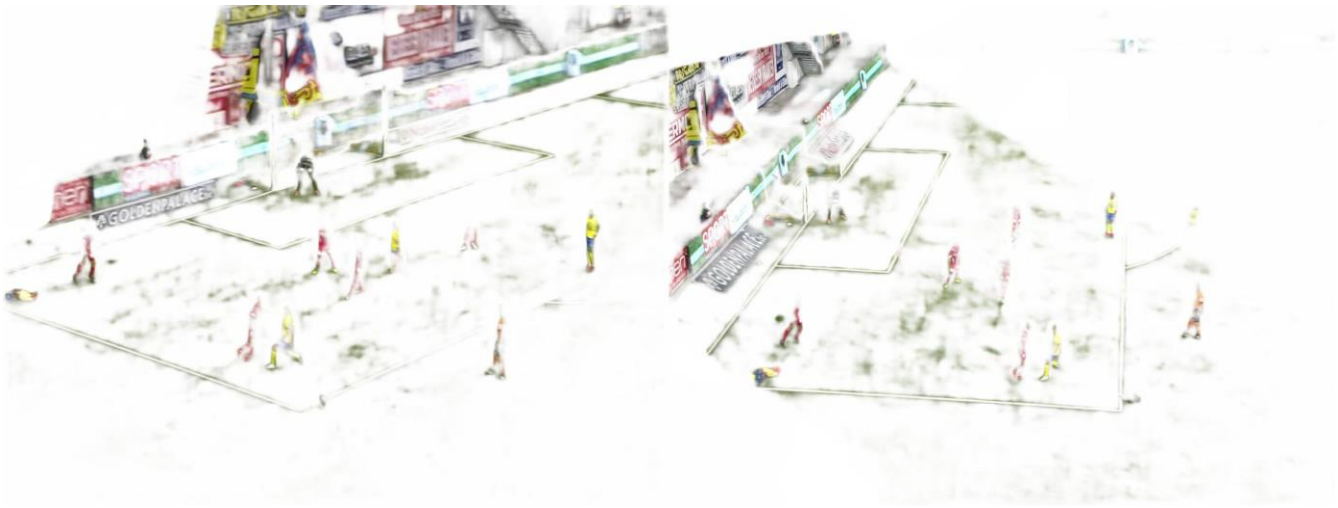
Keywords: VOL, camera calibration, deep learning, feature matching, keypoint detectors

Xeebra is the virtual assistant referee tool developed by EVS. This product includes a Virtual Offside Line (VOL) functionality that is essential for soccer games nowadays. In order to ease the drawing of the offside line in augmented reality, this VOL feature contains sophisticated deep learning and computer vision algorithms.

Feature matching is the task of establishing (point) correspondences between pairs of images. This task is an essential step in stereo vision, which allows relative calibration of cameras. With the current state of dense feature matching detectors, the VOL functionality could benefit from generic detectors to calibrate synchronized cameras with viewpoint overlap.

The goal of this internship is to refine existing state-of-the-art feature matchers to support the existing camera calibration algorithms, enabling calibration of new images, or the improvement of the precision of the camera parameters when the existing primitives used (players, line markings) are scarce.

The ideal candidate for this internship has a good background in deep learning and computer vision, especially on camera calibration, and some experience with python, PyTorch, and Linux.



Results of RoMa[1] on main and 16 meters camera.

[1] *"RoMa: Robust Dense Feature Matching", J. Edstedt et al., CVPR 2024*

Interested? Send your CV and motivation letter to f.magera@evs.com, t.michel@evs.com, v.botta@evs.com *and* o.barnich@evs.com

Academic supervision provided by one of two supervisors: anthony.cioppa@uliege.be *or* g.louppe@uliege.be

EVS Internships 2026-2027

Multi-View Feed-Forward Networks for 3D Scene Reconstruction

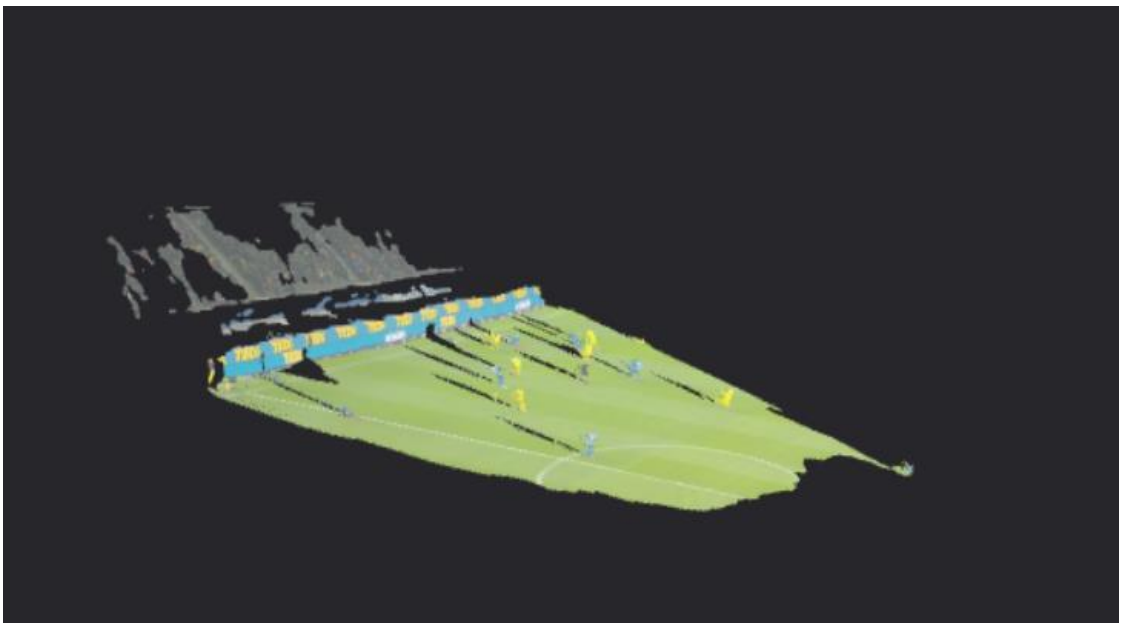
Keywords: computer vision, deep learning, transformers

The past years have seen several breakthroughs in terms of 3D reconstruction from images. Until recently, the best methods for 3D reconstruction were classical SfM methods such as COLMAP[2], which performs feature detection and matching, camera calibration, triangulation as sequential steps in an iterative setup.

The latest developments from DUSt3R[3], and alikes as well as VGGT[4], DepthAnythingV3[5] suggest that big transformer models are able to capture the scene geometry from very few images as demonstrated in the figure below, in a single step inference.

The purpose of this internship is to investigate the relevance of such transformers for our applications. How accurate are the predicted point clouds from the limited and sparse viewpoints available through broadcast cameras? What level of precision can we expect from the predicted camera parameters?

The ideal candidate for this internship has a good background in deep learning and computer vision, especially on camera calibration, and some experience with python, PyTorch, and Linux.



[2] *"Structure-from-Motion Revisited", J. Schönberger et al., CVPR, 2016*
[3] *"DUSt3R: Geometric 3D Vision Made Easy", S. Wang et al., CVPR 2024*
[4] *"VGGT : Visual Geometry Grounded Transformer", J. Wang et al., CVPR 2025*
[5] *"Depth Anything 3: Recovering the Visual Space from Any Views", Haoting Lin et al, arXiv 2025.*

Interested? Send your CV and motivation letter to f.magera@evs.com, t.michel@evs.com, v.botta@evs.com and o.barnich@evs.com
Academic supervision provided by one of two supervisors: anthony.cioppa@uliege.be or g.louppe@uliege.be

EVS Internships 2026-2027

Point cloud results of VGGT from a single soccer image.

Computer Graphics to Photorealistic Style Transfer for Sports Broadcasting

Keywords: Computer Vision, Deep Neural Networks, Diffusion Models, Domain Adaptation, Computer Graphics

Acquiring annotated real-world data for deep learning is often labor-intensive and technically demanding. Generating ground-truth depth maps or 3D environment representations frequently requires specialized hardware and remains prone to noise. Consequently, synthetic data produced via 3D rendering pipelines like Blender or Unreal Engine is invaluable, as it allows for the generation of infinite image pairs with precise ground truth from any perspective. However, a significant "sim-to-real" gap persists; if synthetic renderings lack photorealism, models trained on them often fail to generalize to real-world scenarios.

We currently utilize a 3D soccer dataset, the same featured in the 2026 SoccerNet NVS challenge, but its lack of photorealism currently limits its effectiveness for real-world depth transfer. In this internship, you will investigate state-of-the-art methods, such as diffusion models (for example NeuralRemaster[6]), to enhance the realism of these 3D rendered assets. A primary focus will be maintaining structural integrity while ensuring cross-view and temporal consistency, so that players and environments remain identical across frames.

As part of our Innovation Department, you will collaborate with engineers to develop style transfer techniques that enhance the photorealism of our synthetic data. You will then validate these improvements by training a metric depth estimation model and evaluating its performance.

[6] "*NeuralRemaster: Phase-Preserving Diffusion for Structure-Aligned Generation*", *Y. Zeng et al., arXiv, 2025*

Interested? Send your CV and motivation letter to f.magera@evs.com, t.michel@evs.com, v.botta@evs.com and o.barnich@evs.com

Academic supervision provided by one of two supervisors: anthony.cioppa@uliege.be or g.louppe@uliege.be

EVS Internships 2026-2027

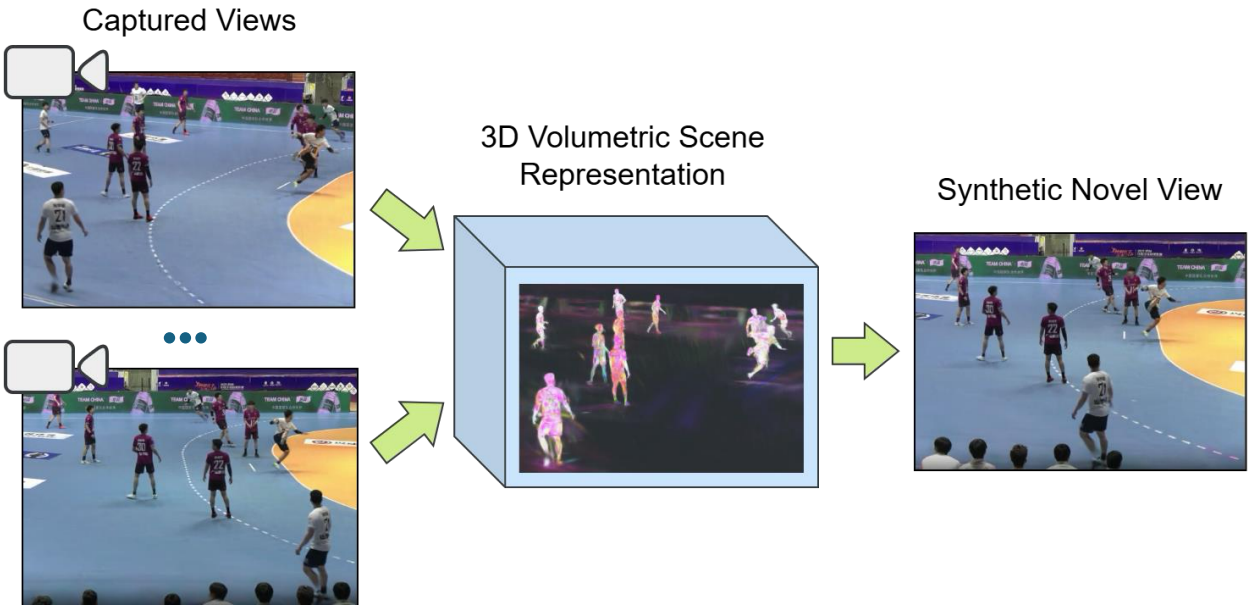
Novel View Synthesis for Immersive Sports Broadcasting

Keywords: Novel View Synthesis, Volumetric Rendering, Gaussian Splatting

Novel view synthesis has the potential to allow viewers to experience sports events from perspectives not captured by physical cameras. Recent breakthroughs in this field, such as 3D/4D Gaussian Splatting, have revolutionized volumetric rendering by offering unprecedented quality and speed compared to earlier neural radiance field (NeRF) approaches. As a teaser, check out https://zju3dv.github.io/longvolcap/ for an example of volumetric video rendering.

However, in real-world scenarios, only a limited number of cameras are often available, making it challenging for these methods to function effectively. The internship will therefore focus on exploring novel view synthesis methods capable of rendering real-time dynamic sports scenes from a limited set of cameras. You will evaluate these methods on multi-view sports datasets, paying particular attention to rendering quality, temporal consistency, and computational efficiency.

Working within our innovation department, you will collaborate with engineers and researchers to develop proof-of-concept applications that showcase the potential of these technologies for enhancing sports broadcasts. This internship offers a unique opportunity to work at the cutting edge of computer graphics and computer vision, with direct applications to next-generation broadcast technology.



Interested? Send your CV and motivation letter to f.magera@evs.com, t.michel@evs.com, v.botta@evs.com and o.barnich@evs.com
Academic supervision provided by one of two supervisors: anthony.cioppa@uliege.be or g.louppe@uliege.be

EVS Internships 2026-2027

Multimodal Embeddings for Enhanced Video Content Search and Captioning

Keywords: CLIP, Vision-Language Models, Embeddings, Video Captioning, Semantic Video Search

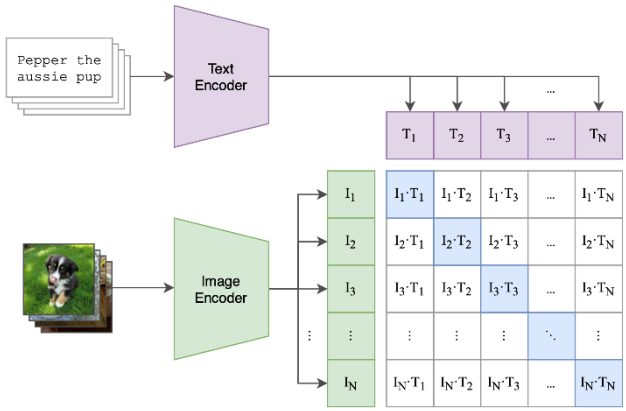
Vision-Language Models (VLMs) like CLIP (Contrastive Language-Image Pre-training) have revolutionized the way machines understand visual content by creating a shared embedding space between images and text. These models enable powerful semantic search capabilities that go beyond traditional metadata-based approaches, allowing for intuitive content discovery based on natural language descriptions.

In the broadcast industry, the ability to efficiently search, categorize, and caption vast video archives is crucial for content monetization and accessibility. Current methods often rely on manual tagging or basic computer vision techniques that fail to capture the rich semantic content of video material.

This internship focuses on leveraging CLIP-like architectures and other VLMs to develop an advanced video search and automatic captioning system specifically tailored for broadcast applications. You will explore how these models can be fine-tuned on domain-specific content to improve search accuracy and generate descriptive captions that capture the nuances of broadcast footage.

Your responsibilities will include implementing and evaluating different embedding approaches, developing efficient indexing methods for large-scale video libraries, and creating a prototype system that demonstrates improved search capabilities and automatic captioning. You will also investigate how these technologies can be integrated into existing broadcast workflows to enhance content discovery and accessibility.

This internship offers a unique opportunity to work with EVS's innovation team, applying cutting-edge AI research to solve real-world challenges in the broadcast industry. Your work will be conducted under the guidance of experienced computer vision engineers and will contribute directly to next-generation media management solutions.



Interested? Send your CV and motivation letter to f.magera@evs.com, t.michel@evs.com, v.botta@evs.com and o.barnich@evs.com
Academic supervision provided by one of two supervisors: anthony.cioppa@uliege.be or g.louppe@uliege.be

EVS Internships 2026-2027

Automated Sports Commentary Generation with Expressive Speech Synthesis

Keywords: large language models, text-to-speech, voice cloning, speech synthesis, emotion control, sports commentary, natural language generation

Live sports commentary is a craft that blends factual accuracy, narrative flair, and real-time awareness. Traditionally, the domain of experienced broadcasters, commentary plays a central role in shaping the viewer's experience, providing context, building tension, and making sense of fast-moving events. As demand grows for multilingual, multi-platform, and personalized sports content, automating the generation of high-quality, expressive spoken commentary represents a compelling frontier for AI research.

Recent breakthroughs in large language models and neural text-to-speech synthesis have demonstrated remarkable capabilities in natural language generation and expressive voice rendering. Modern speech synthesis systems are increasingly able to produce natural-sounding speech with fine-grained control over prosody, tone, and emotional delivery. Combined with voice cloning techniques, these technologies open up exciting possibilities for broadcast applications, from automatic multilingual dubbing to generating adapted voice-overs for shortened video versions.

This internship will take place at EVS headquarters in Liège, under the supervision of engineers from the Innovation team. The objective is to investigate how to generate natural, expressive spoken commentary by combining LLM-based text generation with advanced speech synthesis techniques. You will explore methods to control and modulate emotions during speech synthesis, evaluate voice cloning approaches for automatic translation and content adaptation, and study how synthesized voices can be seamlessly integrated into edited video outputs, for instance, generating a new voice-over when producing a condensed version of a broadcast. Your work will contribute to a broader vision of AI-assisted, multilingual, and personalized sports production at EVS.

Interested? Send your CV and motivation letter to f.magera@evs.com, t.michel@evs.com, v.botta@evs.com and o.barnich@evs.com
Academic supervision provided by one of two supervisors: anthony.cioppa@uliege.be or g.louppe@uliege.be

EVS Internships 2026-2027

Advanced Action Spotting for Football Broadcasts: Enhancing Viewer Experience Through Real-Time Event Detection

Keywords: Action Spotting, Sports Analytics, Computer Vision, Deep Learning, Football/Soccer

Action spotting in soccer has emerged as a critical technology for enhancing broadcast experiences, enabling automatic highlight generation, advanced statistics, and interactive viewing features. The ability to automatically detect key events such as goals, cards, substitutions, and corner kicks in real-time creates opportunities for more engaging and informative sports broadcasts.

Recent advances in computer vision and deep learning have significantly improved the accuracy and efficiency of action spotting systems. However, challenges remain in handling the variability of camera angles, broadcast styles, and the subtle nature of some important game events.

This internship focuses on reviewing and implementing state-of-the-art methods for action spotting in football broadcasts. You will explore various approaches including temporal convolutional networks, attention-based models, and multi-modal architectures that combine visual and audio cues for more robust event detection.

Your responsibilities will include conducting a comprehensive literature review of current techniques, implementing selected methods on standard datasets (such as SoccerNet), and developing evaluation frameworks to assess performance in real-world broadcast scenarios. You will also investigate how these technologies can be optimized for low-latency applications required in live production environments.

Working within EVS's innovation team, you will have the opportunity to test your implementations on professional broadcast footage and receive guidance from experts in broadcast technology. This internship offers valuable experience at the intersection of computer vision, sports analytics, and broadcast engineering, with potential contributions to next-generation sports production tools.

Interested? Send your CV and motivation letter to f.magera@evs.com, t.michel@evs.com, v.botta@evs.com and o.barnich@evs.com
Academic supervision provided by one of two supervisors: anthony.cioppa@uliege.be or g.louppe@uliege.be

EVS Internships 2026-2027

Replay Grounding for Football Broadcasts: Matching Replays to Their Original Match Time

Keywords: Replay Grounding, Temporal Alignment, Computer Vision, Deep Learning, Football/Soccer

Modern sports broadcasting generates massive amounts of video, blending continuous live feeds with various replay shots to enhance the viewer experience. Replay grounding consists of retrieving the exact timestamp of an action shown in a given replay shot within the complete broadcast game. Linking replays to their corresponding live action automatically identifies the most crucial moments of a match, enabling powerful applications like automatic highlight generation.

Recent advances in computer vision and deep learning have significantly improved the accuracy and efficiency of replay grounding models. However, matching a potentially slow-motion, highly zoomed replay from a unique camera angle back to the wide-angle live broadcast remains a complex computer vision challenge.

This internship focuses on reviewing and implementing state-of-the-art methods for replay grounding in football broadcasts. You will explore various approaches including temporal convolutional networks, attention-based models, and multi-modal architectures that combine visual and audio cues for highly precise temporal localization.

Your responsibilities will include conducting a comprehensive literature review of current techniques, implementing selected methods on standard datasets (such as SoccerNet), and developing evaluation frameworks to assess performance in real-world broadcast scenarios. You will also investigate how these technologies can be optimized for low-latency applications required in live production environments.

Working within EVS's innovation team, you will have the opportunity to test your implementations on professional broadcast footage and receive guidance from experts in broadcast technology. This internship offers valuable experience at the intersection of computer vision, sports analytics, and broadcast engineering, with potential contributions to next-generation sports production tools.

Interested? Send your CV and motivation letter to f.magera@evs.com, t.michel@evs.com, v.botta@evs.com and o.barnich@evs.com
Academic supervision provided by one of two supervisors: anthony.cioppa@uliege.be or g.louppe@uliege.be

EVS Internships 2026-2027

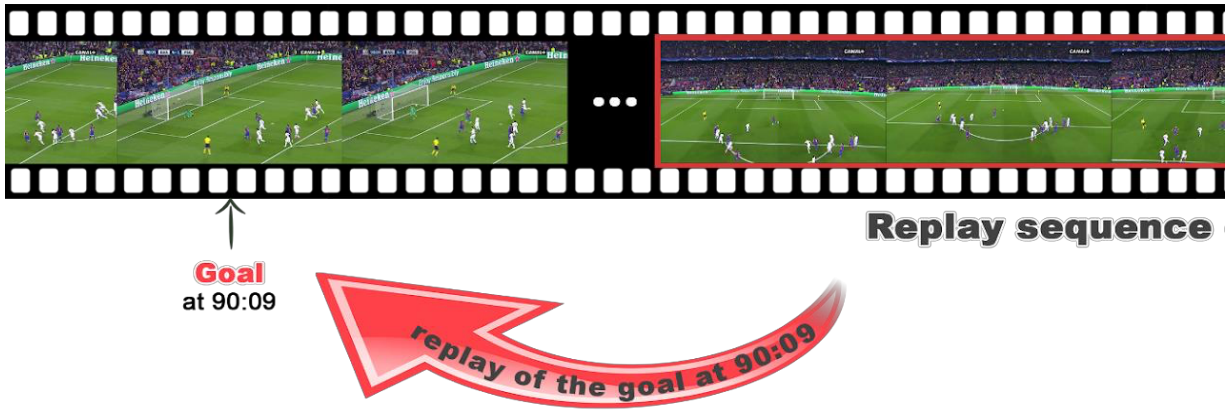


Illustration taken from SoccerNet[7].

Automatic Highlights Generation for Sport Broadcasting

Keywords: content summarization, speech-to-text, video captioning, natural language processing.

In sports broadcasting, a highlight is a curated segment of an event that captures its most important, exciting, or impactful moments. These summaries typically range from quick 3–5 minute recaps to more in-depth 10–12 minute overviews, catering to audiences from casual viewers to dedicated fans. With growing demand for personalized and on-demand sports content, broadcasters are increasingly seeking efficient, scalable methods to produce multiple highlight versions tailored to various platforms and user preferences.

Recent advances in artificial intelligence—particularly in computer vision, natural language processing, and multimodal learning—have paved the way for increasingly sophisticated and high-quality automated highlight generation systems. For instance, speech-to-text technology now delivers high-accuracy performance, while large language models (LLMs) offer powerful capabilities for identifying key moments and filtering out irrelevant content.

This internship will take place at EVS headquarters in Liège, under the supervision of engineers from the Innovation team. The objective is to explore how to most effectively identify video segments that encapsulate the most relevant moments by fusing a variety of modalities, including live speech-to-text commentaries, scripted live commentaries, video captioning, and human-labeled metadata.

[7] https://www.soccer-net.org/tasks/replay-grounding

Interested? Send your CV and motivation letter to f.magera@evs.com, t.michel@evs.com, v.botta@evs.com and o.barnich@evs.com

Academic supervision provided by one of two supervisors: anthony.cioppa@uliege.be or g.louppe@uliege.be

EVS Internships 2026-2027

Automatic video summarization for Broadcast News and Documentaries

Keywords: video summarization, content condensation, natural language processing, multimodal learning, news production

In news and documentary broadcasting, the ability to produce concise, compelling summaries of long-form content is essential for reaching audiences across multiple platforms and formats. A well-crafted summary distills hours of footage into a coherent narrative — preserving the most informative, impactful moments while remaining accessible to a broad audience. With the explosion of on-demand content consumption, broadcasters face increasing pressure to deliver tailored versions of their content efficiently and at scale.

Recent advances in artificial intelligence — particularly in natural language processing, computer vision, and multimodal learning — have opened new avenues for automating this editorial process. News and documentary summarization requires a deep understanding of narrative structure, journalistic relevance, and contextual coherence, making it a uniquely challenging and intellectually rich problem.

This internship will take place at EVS headquarters in Liège, under the supervision of engineers from the Innovation team. The objective is to explore how to most effectively identify and assemble video segments that capture the essential narrative of news reports and documentaries, by fusing multiple modalities including speech-to-text transcripts, video captioning, on-screen text recognition, and editorial metadata. You will investigate how large language models can be leveraged to assess journalistic relevance and structure coherent summaries, and how these approaches can be integrated into broadcast-grade workflows.

Interested? Send your CV and motivation letter to f.magera@evs.com, t.michel@evs.com, v.botta@evs.com and o.barnich@evs.com
Academic supervision provided by one of two supervisors: anthony.cioppa@uliege.be or g.louppe@uliege.be

EVS Internships 2026-2027

Video Stabilization

Keywords: Computer Vision, Motion Estimation, Deep Learning

With the recent integration of advanced robotic camera systems, drones, and cable cams into live sports broadcasting, the ability to capture dynamic, high-impact shots has never been greater. However, the complex physical movements of this specialized equipment can sometimes introduce unwanted camera shaking. Video stabilization aims at computationally removing this jitter to produce smooth, steady footage, making it a critical requirement for maintaining broadcast-quality output.

Traditional video stabilization relies on 2D methods, which estimate the global motion between frames in two dimensions, smooth the trajectory and then warp the images to obtain the new frames. While these methods are generally faster and simpler to implement, they often fail to accurately represent the true geometry of the scene, resulting in artifacts when dealing with perspective changes or parallax.

More recently, 2.5D and 3D stabilization methods have emerged to address these geometrical limitations. 3D methods attempt to explicitly reconstruct the scene geometry and the 3D camera path. When they work well, they provide highly accurate and realistic stabilization. However, estimating a precise 3D path from a shaky video is notoriously difficult and computationally expensive. 2.5D methods offer a promising alternative, aiming to combine the computational efficiency and robustness of 2D techniques with the geometrical accuracy of 3D approaches.

The goal of this internship is to investigate, implement and evaluate state-of-the-art video stabilization algorithms tailored for dynamic camera movements. Working within the EVS Innovation team, you will explore how modern deep learning architectures can improve motion estimation and frame warping to effectively smooth complex trajectories.

Illustration taken from GaVS[8].

[8] *"GaVS: 3D-Grounded Video Stabilization via Temporally-Consistent Local Reconstruction and Rendering", Z. You et al., SIGGRAPH 2025*

Interested? Send your CV and motivation letter to f.magera@evs.com, t.michel@evs.com, v.botta@evs.com and o.barnich@evs.com

Academic supervision provided by one of two supervisors: anthony.cioppa@uliege.be or g.louppe@uliege.be

EVS Internships 2026-2027

Ball Action Spotting from Multi-View Broadcast Videos

Keywords: Computer Vision, Transformers, Ball Trajectory Estimation, Football/Soccer.

The soccer ball serves as the fundamental epicenter of gameplay, yet it remains one of the most elusive objects to track accurately within a broadcast environment. Its detection is perennially challenged by a combination of its minimal pixel footprint, high-velocity motion blur, and frequent occlusions by players or field equipment.

However, overcoming these technical hurdles is essential, as precise ball localization and state analysis act as the primary data engine for modern sports analytics. By successfully capturing the ball's trajectory and interaction points, the system can automate high-stakes officiating tasks like kickpoint detection for offside lines, while simultaneously generating granular performance metrics such as possession percentages, pass completion rates, and tackle success. Furthermore, identifying specific physical events (such as hand-ball infractions or floor-bounce patterns) enables a deeper, context-aware understanding of the match that is invaluable for both broadcast augmentation and tactical coaching.

This internship will take place at EVS headquarters in Liège, under the supervision of engineers from the Innovation team. The objective is to explore how to leverage deep learning methods to detect and identify a wide array of ball interactions based on temporal and multi-view observations from broadcast cameras.

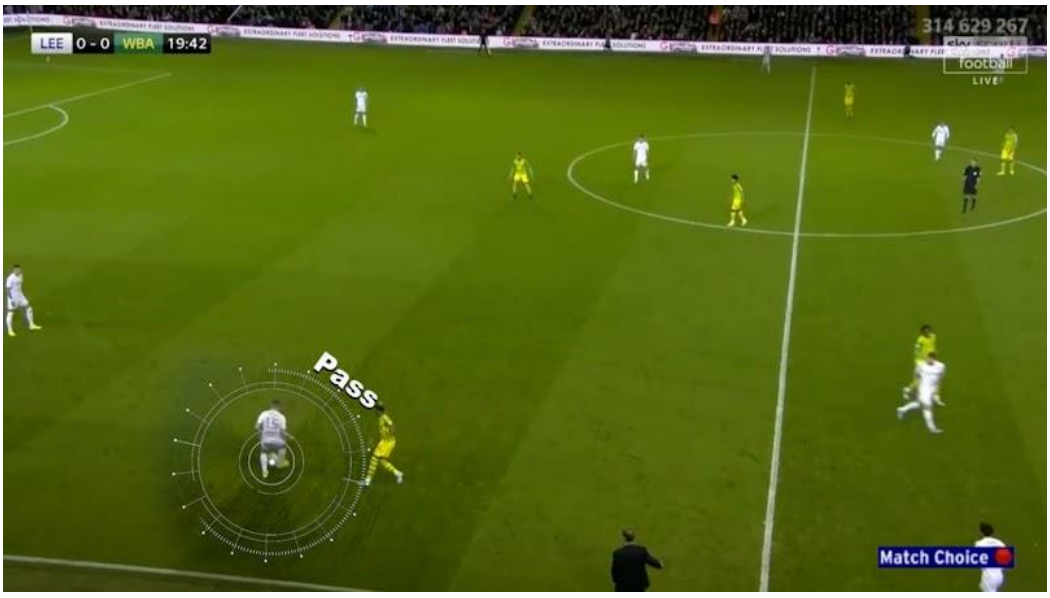


Illustration from https://www.soccer-net.org/tasks/ball-action-spotting

Interested? Send your CV and motivation letter to f.magera@evs.com, t.michel@evs.com, v.botta@evs.com and o.barnich@evs.com
Academic supervision provided by one of two supervisors: anthony.cioppa@uliege.be or g.louppe@uliege.be