Internship topic proposals

Tracking and re-identification of players in broadcast sport videos

Contact: o.barnich@evs.com – t.morael@evs.com

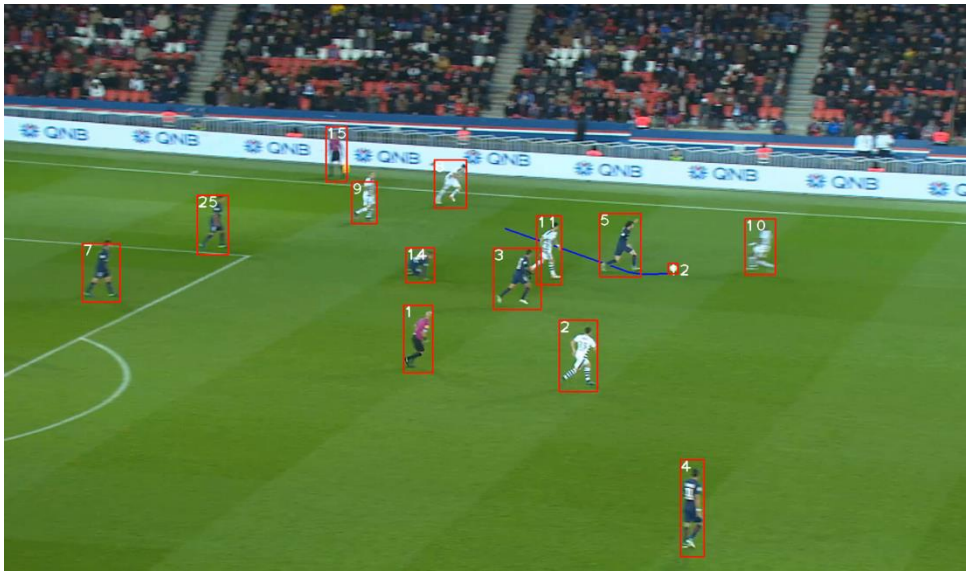
Sport plays a significant part in the entertainment industry. Therefore, it provides many opportunities for innovation and new technologies. One can think of several applications to keep sport broadcast engaging to fans: provide player statistics, add visual effects, make predictions, etc. A lot of these applications require to solve some fundamental computer vision problems such as camera calibration, ball detection, and player tracking. This internship is about the latter problem.

Multi-object tracking (MOT) is a well-studied task in the scientific literature. It consists in identifying each object in a video and representing them as a set of trajectories. When tracking of an object is lost, MOT often includes a re-identification task to identify the lost object in future video frames.

In the context of sport broadcast, typical approaches of the literature are not optimal. Indeed, typical MOT algorithms are mostly trained and evaluated on data recorded from static surveillance cameras filming a crowd of people that are walking and wearing distinct clothes. In contrast, broadcast cameras follow the action of players that are running, changing direction abruptly, and wearing their team jersey.

Fortunately, sport games and sport broadcast have specific elements that we can take advantage of. For instance, we know which player is playing in the game: we can thus use face recognition algorithms to improve our tracking. Although team players wear the same team jersey, they each have a dedicated number which is printed several times on said jersey. Finally, multiple broadcast cameras record the action from different viewpoints, which would increase the duration a player can be tracked.

This internship will take place under the supervision of a computer vision engineer in the innovation team at EVS. You will be tasked to implement and evaluate various strategies (like those mentioned above) to track players in a robust manner.



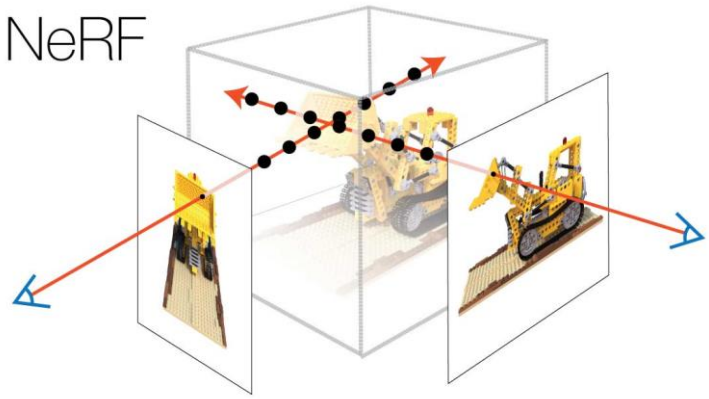
Internship topic proposals

Novel viewpoint synthesis of sport scenes using broadcast images

Contact: o.barnich@evs.com – t.morael@evs.com

Sports play a significant part in the entertainment industry, and as such provide many opportunities for innovation and recent technologies. One can think of several applications to keep sport broadcast engaging to fans: generate player statistics, add visual effects, make predictions, to name a few. One of the most striking visual effects proposed in recent years has been the generation of novel photorealistic viewpoints of sport scenes. This effect is usually presented as the video capture of a virtual camera that would have been freely moving and filming the action alongside broadcast cameras. Existing commercial systems typically discard any video data coming from broadcast cameras, and instead use methods that involve the setup and the cost of tens of proprietary cameras to deliver the novel viewpoint synthesis effect. This internship is about investigating and implementing methods that would only use broadcast images to generate novel viewpoints of a sport scene.

This internship will take place under the supervision of a computer vision engineer in the innovation team at EVS, with frequent follow-ups and opportunities for brainstorming with the entire team. The intern will first be asked to cover the relevant recent literature on the topic, including Neural Radiance Fields (NeRF) and other neural network-based data-driven approaches. The intern will then test the most promising methods on EVS broadcast image datasets, and adapt those methods where needed to produce a convincing proof-of-concept demonstrator. Of course, the intern will be granted access to the entire EVS infrastructure for the duration of his work, including EVS' farm of GPU servers.



Internship topic proposals

Deblurring of sport videos using machine-learning

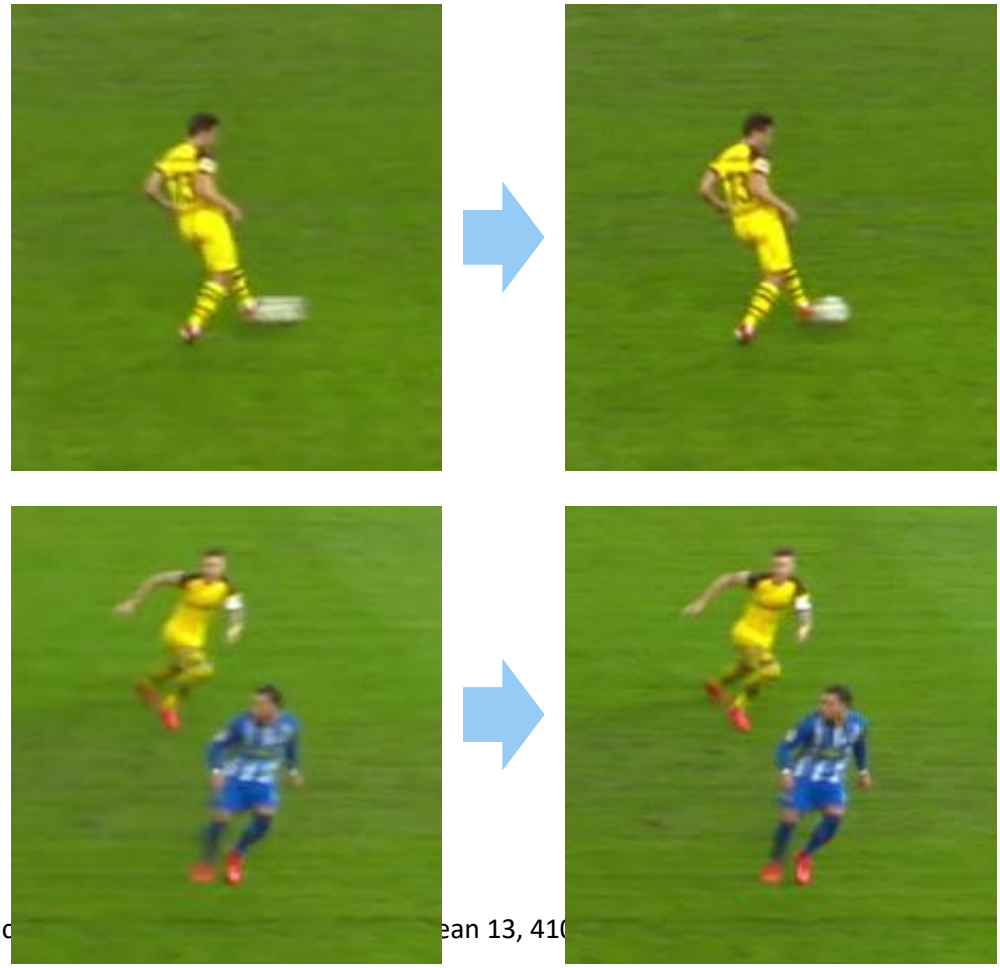
Contact: o.barnich@evs.com – t.morael@evs.com

In the context of live broadcasting, camera operators are under pressure to produce the most compelling images. However, unlike movie productions, they get only one shot to produce content that is interesting, beautiful, does not present motion blur nor noise, and finally has adequate color settings and white balance under any weather conditions. Given the live requirement, it is thus unavoidable that content quality is not always optimal.

Interestingly, a lot of research has tackled the problem of image and video enhancement: deblurring, denoising, artefact removal, inpainting, tone mapping… The latest techniques based on Deep Learning have shown promising results.

Deblurring in particular is among the most interesting image enhancement tasks for EVS. Indeed, motion blur can often be present in the images. Moreover, for producing beautiful slow-motion content, images need to be even sharper than at normal playback speed. Deblurring algorithms would thus be a very useful preprocessing stage for our XtraMotion algorithm that interpolates 50 fps videos to 150 fps to make them smooth when played back at 33% speed.

The intern will thus be responsible for reviewing the deblurring literature, testing available codes, implement the most promising methods in PyTorch and train them on EVS data. A dataset of synthetically blurred videos with their sharp counterpart will be provided to explore the adaptability of these methods on sports content. Moreover, the intern will compare the different methods both quantitatively on public and EVS benchmarks and qualitatively on relevant videos in order to choose the best method and hyperparameters for EVS. The intern will need to be able to work quite autonomously, while receiving regular guidance and relevant help from members of the Innovation team.



Internship topic proposals

Automatic Captioning of Broadcast Videos

Contact: o.barnich@evs.com – t.morael@evs.com

The television broadcast industry generates huge amounts of video. Between the camera feeds and the edited production, only a small part of video material is used for a broadcasted event. However, the data is still available and might be used for other production purposes. In order to ease the use of old video content, tags are sometimes associated to videos in our client's database, allowing for easy search in the database. Unfortunately, these tags are often a scarce resource, and only top productions spend money on the annotation of tags.

With recent advances in computer vision and natural language processing, the task of automatically generating a video description seems feasible. Therefore, automatic generation of video tags or descriptions seems to be promising to help EVS customers to fully leverage their video resources and to enrich their databases with relevant metadata.

The goal of this internship is to explore current video captioning techniques, identify those that seem adequate for EVS data, and provide an implementation of a novel method that works on EVS data. The intern will provide qualitative and quantitative evaluations for this new algorithm. EVS has a large amount of soccer clips that have been tagged. The tags do not provide a complete description as expected at the end of this internship but will still allow us to evaluate the quality of the descriptions provided.

Automatic caption: "10 soccer players standing in front of a goal cage"

Internship topic proposals

Image/Sequence Similarity Engine for Sport Broadcast Industry

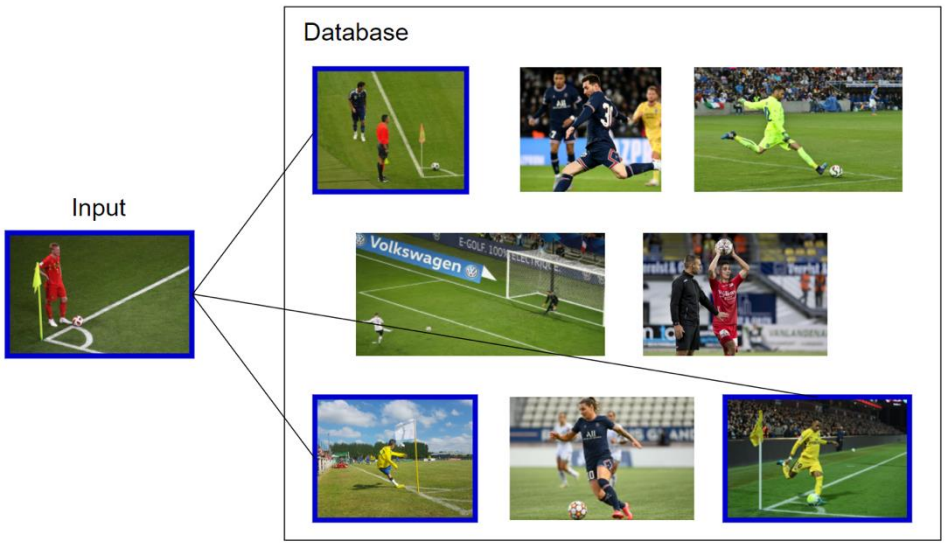
Contact: o.barnich@evs.com – t.morael@evs.com

Sport plays a significant part in the entertainment industry. Therefore, it provides many opportunities for innovation and new technologies. One can think of several applications to ease EVS operators' work and improve live production workflows. Since off-the-shelf methods are not always well suited for sports content, these applications require specializing and solving fundamental computer vision problems such as camera calibration, player detection, player tracking, complex scene understanding, action detection, ...

Common image retrieval systems aim to find similar images based on a query image among a large image dataset. They rely on metadata, keywords, title, and description that are easily collected from the indexation of internet web pages for instance but harder to find for sport video broadcast content.

In the broadcast industry context, video sequences are first class citizens. For instance, the creation of replay and highlights could easily benefit from finding similar moments from other camera feeds and/or huge video archives for the enrichment and the improvement of storytelling.

In this context, you will be responsible for specializing current state of the art methods to sports and live broadcast content as well as extending image retrieval to sequence retrieval based on computer vision and artificial intelligence. During this internship you will become a part of the innovation team at EVS, and your work will take place under the supervision of a computer vision engineer. You will investigate, implement, and evaluate various strategies.



Internship topic proposals

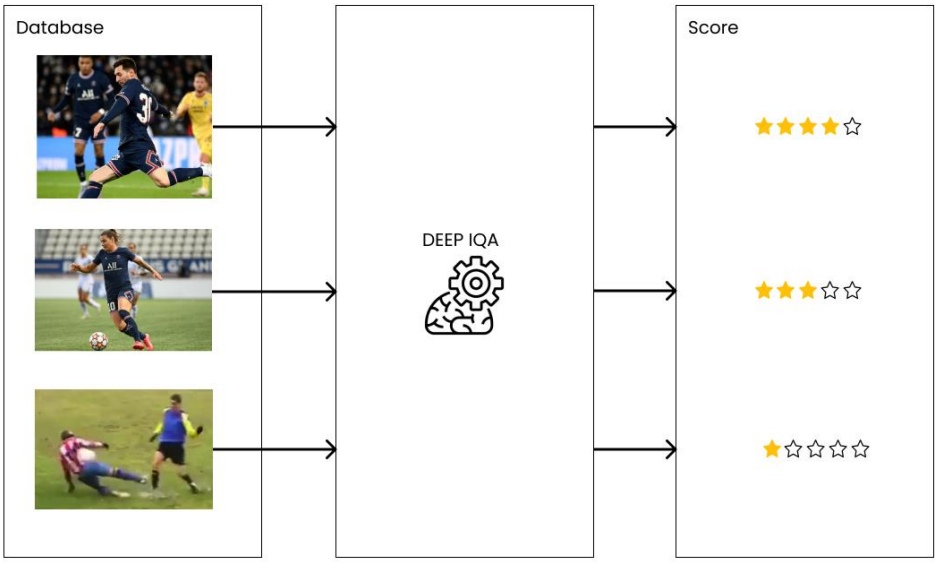
Image quality assessment for Sport Broadcast Industry

Contact: o.barnich@evs.com – t.morael@evs.com

Measuring image quality is of fundamental importance for great content creation and storytelling. Image quality assessment (IQA) aims to determine the level of “suitability” of an image for a given context. It is often based on the subtle combination of objective (sharpness level, contrast, presence of distortion, good exposure, …) and subjective (based on user ratings and publicly available datasets) methods.

Once a large amount of video content is available, the benefits of rating and ranking images based on quality metrics are numerous. For instance, showing top quality clips in first positions to operators would increase the processing times in live broadcast situations where editing needs to be the fastest. Another example is the display of the best previews/thumbnails in asset management UI (User Interface) would improve user experience. As well as the automatic extraction of the best sub-sequence of a clip for social media feeds.

In this context, you will be responsible for investigating current state of the art automatic IQA methods. You will apply and extend the best candidate approach to sports and live broadcast content. During this internship you will become a part of the innovation team at EVS, and your work will take place under the supervision of a computer vision engineer. You will investigate, implement, and evaluate various strategies.



Internship topic proposals

Video Style Transfer for Soccer

Contact: o.barnich@evs.com – t.morael@evs.com

Sport plays a significant part in the entertainment industry. Therefore, it provides many opportunities for innovation and new technologies. One can think of several applications to ease EVS operators' work and improve live production workflows. Since off-the-shelf methods are not always well suited for sports content, these applications require specializing and solving fundamental computer vision problems such as camera calibration, player detection, player tracking, complex scene understanding, action detection, …

Training novel methods, specializing and improving existing ones requires a lot of annotated data. The process of image annotation is cumbersome, it requires a lot of time, efforts and sometimes needs complex tools to remain efficient and impactful.

In this context, with recent advances in deep learning and computer vision, neural style transfer seems to be a great opportunity to generate huge amounts of data from 3D rendered images (such as video games) to turn artificially generated images into photorealistic training samples. This would allow for unlimited data generation, as well as the creation of specific scene that are harder to find in the real world. The goal is to improve the robustness of our computer vision models.

In this context, you will be responsible for investigating current state of the art style transfer methods. You will apply and extend the best candidate approach to transform CGI images into photorealistic images and vice-versa. During this internship you will become a part of the innovation team at EVS, and your work will take place under the supervision of a computer vision engineer. You will investigate, implement, and evaluate various strategies.