Enhancing Fan Engagement in a 5G Stadium With AI-Based Technologies and Live Streaming

Cheng-Wen Wu, *Fellow, IEEE*, Ming-Der Shieh, *Member, IEEE*, Jenn-Jier James Lien, *Member, IEEE*, Jar-Ferr Yang, *Life Fellow, IEEE*, Wei-Ta Chu, *Member, IEEE*, Tsang-Hai Huang, Han-Chuan Hsieh, Hung-Ta Chiu, Kuo-Cheng Tu, Yen-Ting Chen, Shian-Yu Lin, Jia-Jun Hu, Chen-Huan Lin, and Cheng-Siang Jheng

***Abstract*—We participated in Taiwan's National Intercollegiate Athletic Games (NIAG) in early May 2021, and formed a Sport Technology Team of more than 30 scholars, students, and engineers to provide novel systems and solutions that make the athletic games rich in sport technologies. Some of the features could be the first time shown to the Internet audience for large-scale athletic games. The technologies involved include table tennis ball trajectory and bounce distribution, badminton shuttlecock tracking and trajectory, *augmented-reality* enriched video streaming on social networks, real-time three-dimensional broadcasting with free view-angle, in-stadium video stream pushing by a private fifth-generation (5G) network with *multiaccess edge computing*, AI-based sport data analytics during live streaming, etc. All the technologies and applications are integrated in a novel *technology-enhanced broadcasting system* (TEBS) that is dedicated to sport events. This article introduces the respective technologies that we have developed, deployed, and demonstrated in the 2021 NIAG. We stress the layered architecture design and integration of the TEBS, as well as experimental results from real games in the smart stadium and swimming pool. We also discuss the technical challenges and our approaches to tackle them, as well as lessons learned.**

***Index Terms*—Artificial intelligence, fan engagement, fifth-generation (5G) mobile communication, multiaccess edge computing (MEC), operations and management, sports equipment, system of systems, video signal processing, visualization.**

Manuscript received 26 July 2021; revised 26 January 2022 and 21 March 2022; accepted 17 April 2022. Date of publication 13 May 2022; date of current version 9 December 2022. This work was supported in part by the National Cheng Kung University, and in part by the Industrial Technology Research Institute. *(Corresponding author: Cheng-Wen Wu.)*

Cheng-Wen Wu is with the Department of Electrical Engineering, National Tsing Hua University, Hsinchu 300044, Taiwan, with the Miin Wu School of Computing, National Cheng Kung University, Tainan 70101, Taiwan, and also with the Industrial Technology Research Institute, Hsinchu 310401, Taiwan (e-mail: cww@ee.nthu.edu.tw).

Ming-Der Shieh, Jar-Ferr Yang, Yen-Ting Chen, Shian-Yu Lin, Jia-Jun Hu, and Cheng-Siang Jheng are with the Department of Electrical Engineering and Miin Wu School of Computing, National Cheng Kung University, Tainan 70101, Taiwan (e-mail: shiehm@mail.ncku.edu.tw; jfyang@ee.ncku.edu.tw; qaz26845qaz26845@gmail.com; n26090067@gs.ncku.edu.tw; n26094809@gs.ncku.edu.tw; n26094891@gs.ncku.edu.tw).

Jenn-Jier James Lien and Wei-Ta Chu are with the Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan 70101, Taiwan (e-mail: jjlien@csie.ncku.edu.tw; wtchu@gs.ncku.edu.tw).

Tsang-Hai Huang, Hung-Ta Chiu, and Kuo-Cheng Tu are with the Institute of Physical Education, Health and Leisure Studies, National Cheng Kung University, Tainan 70101, Taiwan (e-mail: tsanghai@mail.ncku.edu.tw; htchiu@mail.ncku.edu.tw; phila@mail.ncku.edu.tw).

Han-Chuan Hsieh is with the Industrial Technology Research Institute, Hsinchu 310401, Taiwan (e-mail: hchsieh@itri.org.tw).

Chen-Huan Lin is with the Department of Industrial Design, National Cheng Kung University, Tainan 70101, Taiwan (e-mail: andylin86627@gmail.com).

Digital Object Identifier 10.1109/JSYST.2022.3169553

I. Introduction

In the past few years, even during the COVID-19 pandemic, the sports industry has been busy introducing high-tech solutions into stadiums, arenas, fields, domes, parks, education, and training facilities in many parts of the world [1]–[5]. In addition to the fifth-generation (5G) mobile communication, the engagement of AI-based data analytics and the *Internet of Things* (IoT) in sports also is triggering growing demand for advanced sensor technologies, e.g., microchip Doppler radar, high-speed and high-resolution cameras, camera arrays backed by video stitching/fusion, wearable devices, and RF-ID tags and sensors that monitor athletes and other moving objects in the games [6]–[13]. The huge amount of data generated by all the sensors during training or in the games will need to be processed online and/or offline. For that reason, AI-enhanced data analytics seem inevitable [6]. The technology trends for sports are becoming clear—fan engagement, smart stadium, immersive media, quantified athlete, next-generation sponsorship, e-sport, etc. Among them, *fan engagement* has been under the spotlight recently as fans represent the most significant part in the sports industry value chain. As a result, almost all major 5G players and cloud/AI companies had planned to showcase their products and services in the 2020 Tokyo Olympics [6], though unfortunately as we have already witnessed, most of the plans had been negatively affected by the COVID-19 pandemic.

In Taiwan, thanks to the efforts of the government and the entire society, the pandemic was contained to a level that we could hold the 2021 National Intercollegiate Athletic Games (NIAG 2021) in early May [14]. For exploring and expanding fan engagement, we decided to develop a *system of systems* for fan engagement in sport events. At around the end of 2019, we formed a Sport Technology Team of more than 30 scholars, students, and engineers from the academia and industry to provide novel systems and solutions that make the athletic games rich in sport technologies never seen before in Taiwan, and some of the features could even be the first time shown to the internet viewers during large-scale athletic games. We then integrated the novel systems and solutions into a *technology-enhanced broadcasting system* (TEBS), as a system of systems, and showcased it in the 2021 NIAG [14].

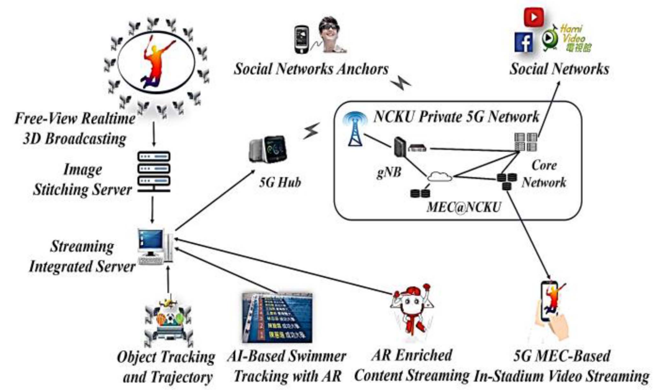
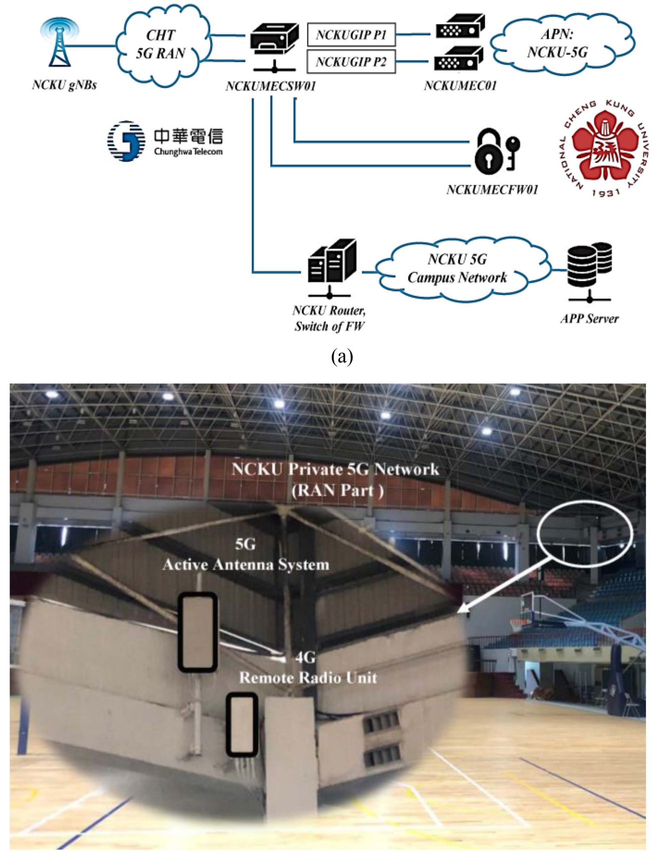


Fig. 1. Integrated applications in the TEBS that demonstrates AI-based sport technologies and live 4K video streaming.



(b)

Fig. 2. NCKU private 5G network (a) deployed in the 5G Stadium (b).

In this article, we will introduce the technologies integrated in the TEBS, including table tennis ball trajectory and touch-down (bounce) distribution, badminton shuttlecock tracking and trajectory, *augmented-reality* (AR) enriched content for real-time video streaming on social networks, real-time three-dimensional (3-D) broadcasting with wide-field free view-angle, in-stadium video stream pushing by a private 5G network with *multiaccess edge computing* (MEC), real-time AI-based sport data analytics, etc. We will also discuss the design and integration of the TEBS, and show the experimental results from large-scale athletic games in the smart stadiums that we have established, i.e., the 2021 NIAG, Taiwan [14]. The end-to-end (E2E) latency in the 5G stadium is shown to be less than 30 ms as measured in the drone racing game. Note that we have demonstrated, the first time, the technologies in an integrated system, i.e., the TEBS, with live streaming to the viewers on social networks and on top of a private 5G network.

II. 5G Stadium With AI-Based Sport Technologies

The integrated applications in the TEBS is depicted in Fig. 1, which shows the main application systems deployed on top of the 5G network, including 1) *object tracking and trajectory* (covering table tennis ball trajectory and bounce distribution, and badminton shuttlecock tracking and trajectory), 2) *AI-based swimmer tracking with AR*, 3) *AR-enriched content streaming*, and 4) *free-view realtime 3-D broadcasting*. For real-time social-network streaming of sport games, we also have developed the *5G MEC-based in-stadium video streaming* system. We will discuss all the components in detail, beginning with the private 5G network.

A. NCKU Private 5G Network

The National Cheng Kung University (NCKU) Private 5G Network located in the NCKU Stadium was built in early 2021 by Chunghwa Telecom (CHT) Co., Ltd. and NCKU, which consists of a set of *next generation NodeB* base stations that are connected to the MEC servers and routers via the CHT *radio access network* (RAN) and on-campus backhaul fibers, as depicted in Fig. 2(a). The deployed network devices are compliant with the 3GPP 5G NR Release-15 nonstand-alone (NSA) specifications, including the E2E network architecture, massive antenna array, MEC, and vertical application systems developed at NCKU that reside on the *firewall* protected *APP servers*. Advanced features, including the *enhanced mobile broadband* and *ultra-reliable low latency communications*, are implemented and supported by the settings in the 5G stadium. A photo of the 5G stadium highlighting the RAN part is shown in Fig. 2(b), where we can see the mounted 5G active antenna system and the 4G remote radio unit.

Table I briefly summarizes the elements and specifications of the NCKU 5G network, including the 5G RAN and MEC. The measured maximum uplink and downlink data rates are higher than 1.35 Gbps and 153 Mbps, respectively. The bandwidth is enough for the target real-time video streaming applications. In addition to high bandwidth, the low-latency feature of the 5G technology is quite important to real-time streaming, especially for onsite spectators. The measured E2E latency during real streaming in the games is lower than 30 ms, including the sport-tech applications running on local and MEC servers, such as the 4K/8K video streaming, AR-enhanced video content, onsite ball tracking/trajectory, etc.

B. Technical Challenges of the TEBS

The integrated TEBS provides in-stadium spectators as well as internet viewers (online fans) with exciting new experiences. As the NIAG 2021 has attracted top athletes from all over Taiwan,

TABLE I
NETWORK ELEMENT LIST AND SPECIFICATIONS

Network Name	Network Part	Element or Feature	Model/Spec
NCKU Private 5G Network	RAN	Remote Radio Head	Nokia AEQZ
			Nokia FHEL + FRHG
		Frequency	3.5GHz (N78), 1.8GHz, 2.6GHz
		Throughput	Downlink (Peak) > 1.35Gbps Uplink (Peak) > 153Mbps
		Latency	E2E < 30ms
	MEC	Server	Nokia CMG A2
		L3 Switch	Juniper EX3400-24T
		Firewall	Juniper SRX345

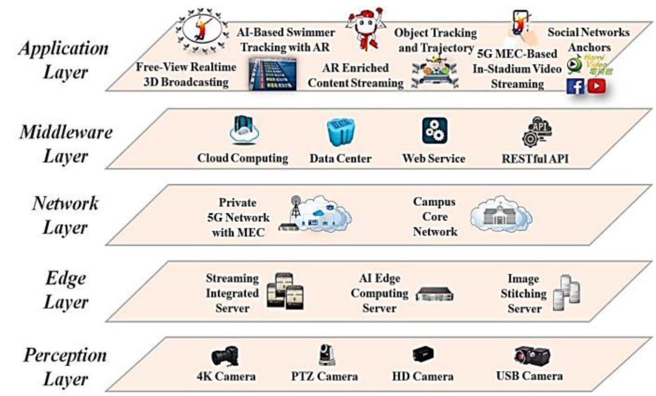


Fig. 3. Layered architecture of the TEBS.

some even top-ranked worldwide, this is an important turning point for fan engagement enabled by emerging technologies, e.g., AI, 5G, IoT, immersive media, video stitching/fusion, 3-D free view-angle replay, AR/*mixed reality* (MR), etc. However, to build the application systems as depicted in Fig. 1 on top of the 5G network, we do face some technical challenges. The major challenges and the proposed approaches to tackle them are as follows:

1) To guarantee the quality of experience (QoE) of the onsite spectators and online viewers is quite important for fan engagement. Integration of all the applications on the 5G network that are developed and operated by different teams, given their various hardware equipment and software tools, is a very complicated task, so we develop a layered architecture of the TEBS for better management of the system of systems, which can quickly connect individual applications and optimize parallel development requirements.
2) Innovative in-stadium spectator service is critical to fan engagement, especially in the 5G age. The local computing power and short 5G latency are key. We have developed software for integrating the MEC servers with the 5G small cells, i.e., the novel *MEC-based in-stadium video stream pushing platform*, so even the onsite spectators can enjoy real-time broadcasting as well, with similar services and QoE for online viewers.
3) Tracking the dynamic and high-speed objects in the badminton and table tennis games requires improved AI models, which will be described later. Also, training of the models requires taking videos from expert players, and labeling the training videos also requires experts from the respective sports, which are performed by coaches and players of the NCKU Athletics Department under the guidance of the EECS professors, with the software tools developed by the respective teams.
4) Tracking the swimmers in all the lanes to create real-time AR content in the outdoor swimming pool is quite challenging as the sunlight keeps changing its strength and direction during the day, and weather condition is dynamic as well. Furthermore, swimmers with part of their bodies in the water and the splashes of the water have been the major obstacle for detection and tracking. Again, we have developed improved AI models that will be discussed later.

C. Layered Architecture of the TEBS

In the 5G private-network business, the key success factors would not simply be higher bandwidth and shorter latency, but also the ecosystem and the provision of customized services in specific domains that 4G cannot easily achieve. In recent years, the 5G technology and applications have been showcased at major international information and communication exhibitions, where global trends in mainstream technological applications are mostly focused on *business to consumer* audio-visual entertainment and *business to business* electronic commerce private-network applications [15]. Today, the 5G developments are no longer oriented toward technological research, but toward applications associated with AR and *virtual reality*, streaming of 4K ultra-HD video and audio, etc. [16]. We have observed that *justification of the value of 5G is a complicated and tedious collaboration task*, which will have to be jointly demonstrated by the 5G network operators, stadium (private 5G network) owner, vertical application developers, equipment suppliers, commercial software providers, system architects and integrators, as well as service providers. Most important of all, we must *ensure that all stakeholders benefit from the integrated task*, which mainly lies in the *high potential of business opportunities from fan engagement*. For that purpose, we propose the TEBS for fan engagement in sport events, as shown in Fig. 3.

In the figure, we show the *layered architecture* of the TEBS, which combines vision-based sports technology applications and live video streaming supported by the NCKU private 5G network. The top layer is the *application layer*, which exhibits the applications and services we have demonstrated in NIAG 2021. The applications will be described one-by-one later. Each application can be considered as an independent system with its own development and operation team. The applications, however, share the *middleware layer* and the infrastructure (*network layer* and *edge layer*), as shown in Fig. 3.

Regarding the edge layer, we adopt the decentralized MEC approach for image processing and sport data preanalysis, taking advantage of its short setup time to achieve higher flexibility in system computing management. The AI edge computing servers are connected directly to the cameras, which are considered as an analysis module at the edge layer for fast front-end AI computation (image processing, AR, object tracking, trajectory analysis, bounce distribution, etc.). They and the stitching servers feed the streaming server, which in turn send the video streams to the 5G hub (Fig. 1). In short, the front-end images and other data are processed at the edge layer, and the stitched images and data analysis results are offloaded through the MEC at the network layer to minimize application latency and improve network efficiency. The approach greatly enhances image processing and sport data analysis performance in the 5G sport event streaming framework.

Regarding the middleware layer, it provides the application data interface for exchange services in accordance with the open system interconnection (OSI) model. Considering that the applications are developed and operated by different teams, guaranteeing their interoperability and maintaining system availability and efficiency are difficult but important. Although they could all be migrated to the MEC servers, the engineering effort would be quite high and risky. Therefore, we propose the layered TEBS architecture, based on which we can quickly integrate individual applications and optimize the development tasks in parallel under a tight schedule. We specify the application programming interfaces (APIs) of all applications in the same format, so system integration and configuration are more efficient and flexible.

During NIAG 2021, online fans have enjoyed the real-time streamed games on social networks, provided by the Facebook, YouTube, and CHT Hami Video platforms. The live streaming of the games was done by six teams of students and teaching assistants from 18 different departments at NCKU, where each team consists of students responsible for, respectively, planning and marketing, network and equipment, live broadcasting, content and software management, etc. Note that the applications exhibited include not just the official games in NIAG 2021, but also a drone racing game that is demonstrated the first time in a 5G stadium using 5G-equipped drones and MR goggles. The drone racing game was shown to some selected onsite student spectators, and open to the fans on the internet by live streaming on May 1, 2021. The recorded short and full videos can be checked out at the following URLs, respectively: 1) NIAG 2021 Sport Technology Debut[1], and 2) 5G Drone Racing Show and AR Opening[2].

Finally, the *perception layer* contains various sensing devices for sport events (so far, they are mainly the camera modules), which are connected directly to the video processing and AI computing servers in the edge layer. Note that different applications may require different sensors, and they can share some of them, depending on the user requirements. In the 5G stadium, on top of all these layers, we have consolidated and integrated cross-domain solutions to facilitate different applications in the application layer.

[1]https://www.youtube.com/watch?v=8zhLIj53cyo
[2]https://www.youtube.com/watch?v=QASDiljBWM0

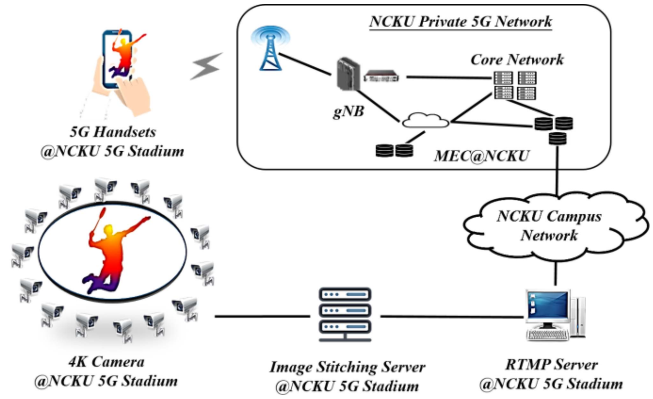


Fig. 4. MEC-based in-stadium video stream pushing platform.

D. In-Stadium Video Pushing by a Private 5G Network With Multiaccess Edge Computing

The selected NIAG 2021 sport events were broadcasted with the help from CHT, which provides the NSA-compliant 5G private network, together with our MEC servers that are installed on campus. The MEC servers are responsible for cloud datacenter offloading as well as latency reduction of the 5G services that support the TEBS applications. To perform packet processing and traffic aggregation closer to the network edge, e.g., they analyze the user packet destination to determine the service location and direct local traffic to the dedicated user plane function (UPF) for offloading, i.e., to ensure that local broadcasting data will be handled on campus. Locally processing the data also helps to meet the 5G low-latency specifications. From Fig. 4, we can see that in the NCKU private 5G network, the MEC servers are also connected to the CHT core network, which implements the 5G control and user plane separation architecture that allows the session management functions (SMFs) to be decoupled from the UPFs. As the UPFs are deployed on the MEC servers, they are close to the 5G mobile users, i.e., the onsite spectators for NIAG 2021 sport events. The SMFs that handle the control plane, on the other hand, are deployed at the core network to manage other UPFs for traffic routing.

Before the opening of NIAG 2021, we had installed 44 4K cameras in the 5G stadium. The first event to showcase the sports technologies is the drone racing game, which is a demonstration game instead of an official game for NIAG 2021. It is the first time that drone racing competition is held in a 5G stadium using *5G-enabled first-person-view* (FPV) drones and MR goggles. With the 5G technology and a commercial 4K live streaming system, onsite and offsite fans can use 5G mobile phones and pads to access real-time content produced on the application platforms in the 5G smart stadium. They can also switch between 3-D viewing angles or instantly playback to experience the racing drones from various viewing angles. Fig. 4 depicts the proposed MEC-based in-stadium video stream pushing platform, which is designed to support two options. It can determine the service packets to be processed by the local

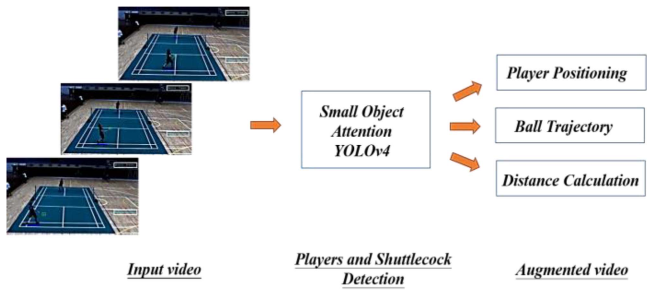


Fig. 5. Block diagram of the smart badminton analysis system.

(edge) computing platform directly, and those need to go through the backbone and core network to access the cloud computing services. Note that the live streaming and playback contents are to be broadcasted locally as well, using the 5G private network. This is necessary in the drone racing game and many other sport events, where the low-latency requirement is a must.

III. Sport Technologies Demonstrated

A. Object Tracking and Player Positioning

Object detection and tracking is a key technology in computer vision. For both table tennis and badminton games, we have adopted the state-of-the-art object detection AI model, i.e., YOLOv4 [17], to improve object detection accuracy based on *convolutional neural network* models. In the following, we will give an overview of the adopted technical schemes and address the main challenges and refinements when applying those to our target applications. More detailed information will be described in Sections III-B–E.

In the badminton games, we use YOLOv4 to track the movement of badminton shuttlecock and the players as well. To detect the small and fast-moving object, we have added the *context attention* [19] and the *residual attention* modules [20] to YOLOv4, as shown in Fig. 5.

The context and residual attention modules that we propose for the modified YOLOv4 can detect and track the players and the shuttlecock simultaneously. With the residual attention module, the network can retrieve the high-level information that helps semantic understanding, so we can focus on detecting small objects. Experimental results show that the *mean average precision* of small object detection is improved by 18% as compared with the original YOLOv4, including the badminton shuttlecock and rackets. After a series of detection and tracking, we then use a postprocessor to overlap the shuttlecocks and the player positions to draw their trajectories. Both trajectories and player positions are exhibited by AR, i.e., merged with the original image, providing exciting new information and intelligence that improves the audience's experience when watching the game online. It also helps a lot for the live broadcasters, anchors, and commentators to deliver colorful and interesting messages to the fans during live streaming.

For table tennis, real-time object detection in high accuracy also can be achieved after we fine-tune the network parameters of YOLOv4 with manually collected ball samples. For table tennis matches in NIAG 2021, we have manually collected 480 ball samples captured in the stadium in advance, and fine-tuned the YOLOv4 model that was originally trained based on datasets of generic objects like chairs, horses, and persons. Overall, the fine-tuned YOLOv4 model can detect the table tennis ball in real time with an accuracy higher than 94%. More details of the ball detection algorithm can be found in [21]. When deploying the model in our 5G stadium to avoid the influence of the noisy background, we have manually defined an *ignore mask* before we start the detection process. Moreover, based on the captured image, we can identify the four corners of a *quadrilateral* that defines the ignore mask. Pixels in this quadrilateral will not be examined by YOLOv4. This simple process not only reduces the influence of the background, but also boosts the efficiency of object detection.

We also have applied the AR view to swimming competition in NIAG 2021. For tracking the swimmers, the *Siamese region proposal networks* (SiamRPN++) as proposed in [18] is a potential tracking algorithm that can be used, which has evolved from the SiamRPN model by the same group. In our use case, we modified the SiamRPN++ to meet our application requirements—real-time detection and tracking of eight swimmers simultaneously in the swimming competitions. We did not change the structure of the model, but trimmed it to remove unnecessary modules and reduce its complexity. To implement this tracking algorithm for swimming competition, we need to overcome several problems. First, there are space, network, and equipment constraints for installing cameras, so the view-angles of the cameras can be limited. Because the camera needs to have long *depth of field* to cover the entire swimming pool, which is around 28 m $\times$ 56 m in physical rectangular dimensions, the best perspective view of the two lateral cameras that we can install is still unsatisfactory, as nearer swimmers can look much bigger than farther swimmers. The second problem is that the swimming pool is located outdoor, and the sunlight reflects from the waving water surface, which generates noisy signals. This reflection can cause overexposure on the cameras. Another issue is that the swimmers are always occluded by the continuously changing water splashes, and sometimes they swim underwater. Finally, there are basically four different swimming styles, i.e., the freestyle stroke, breaststroke, backstroke, and butterfly stroke. Different strokes can cause different intensities of water splash. We, of course, need to solve these issues to maintain the accuracy of our tracking and analysis.

B. Badminton Shuttlecock and Player Tracking

The tracking of badminton shuttlecock is challenging, as the ball is small and fast, and it could be stretched and deformed during flight. Again, we choose YOLOv4 as our base algorithm. For the small and fast-moving object, the customized YOLOv4 achieves a good precision-speed tradeoff among all popular object detection models, which help us to accurately detect and track the shuttlecock during the entire badminton match. Of course, we need to use labeled data to train the *neural network* (NN) model. As there are stretched badminton shuttlecock images in real games, we have trained the NN by using the stretched labeling data, and the output detection result

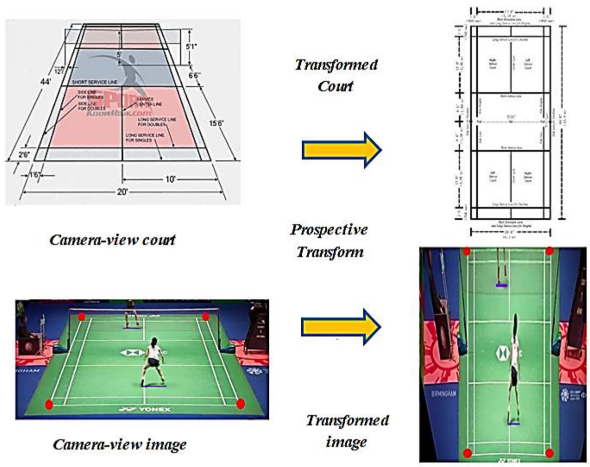


Fig. 6. Perspective transform to calibrate the badminton court.

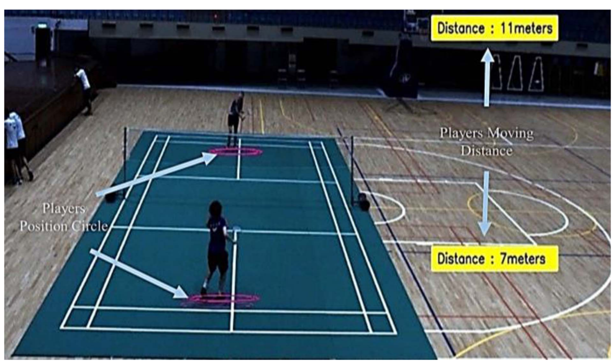


Fig. 7. Snapshot of the detection and analysis results.

has been satisfactory. For calculating the moving distances of the players, we use a *perspective transform matrix* to calibrate the distorted badminton court into a rectangular court by using only a single camera, as shown in Fig. 6. The low-cost and fast solution proves satisfactory in real badminton games. In the figure, the court is an international standard badminton court. After the perspective transform, we translate the bottom center of the original bounding box in the camera-view to the bottom center of the transformed badminton court for correct alignment of the badminton court. Finally, we calculate the distance of the same player in two adjacent frames. By combining all moves collected by adjacent frames, we come up with the total moving distance and the average moving speed for each of the two badminton players. With this mechanism, we can calculate the moving distance of a player in any shooting angle of the camera with only the coordinates of the four corners. Note that the moving distance is an important data item for competitiveness analysis, fatigue evaluation, injury assessment, tactics analysis, etc. of the badminton players during the match.

The same technology apparently can be applied to other sports as well. We use a relatively low-cost GTX 1080 as our training and inference GPU. Also, we use the FLIR Grasshopper 3 digital camera to capture the real-time event. A snapshot of the object detection and data analysis results during the match is shown in Fig. 7, where the pink ellipses are the positions of the players

Fig. 8. Snapshot of a table tennis match with real-time ball tracking and the bounce distribution analysis result.

and the yellow boxes show the respective total moving distances of the two players so far. The data can be turned into AR content to be embedded in the video for broadcasting. With a resolution of 1920 $\times$ 1080, the proposed system can achieve 5 fps on the Nvidia GTX1080 GPU, which is good enough for the broadcaster and commentator to show the analysis results and explain the summary at the end of each match and during the breaks. In real practice, when the commentator hits the button to show the replay video, the system will start to play the 3-D replay video, which is from the stitched images captured by a separate array of 12 cameras shooting the same event from 12 different angles. Meanwhile, the GPU will finish the analysis and the AR results will be embedded in the video from the final camera. If the application requirement asks for a higher frame rate, such as real-time applications using AR glasses, we can use, e.g., the GTX 3080Ti GPU to improve the performance to 24 fps. We also believe that an algorithmic approach (e.g., an interpolation-based one) can be used to improve the frame rate as well.

C. Table Tennis Ball Tracking and Trajectory for Bounce Distribution Analysis

When deploying the proposed table tennis analysis system in the 5G smart stadium, we use one SONY FDR-AX700 digital camera to capture the table tennis match. Fig. 8 shows a snapshot from the camera with AR results displayed as well. Based on the captured image, we can label four corners of a quadrilateral, which defines an *ignore mask*. The captured signal is transmitted to the server for analysis, via a 20-m HDMI 2.0 wire. In the analysis machine, one AVerMedia GC570D video capture card is equipped to capture the signal into memory, which largely eases the computation load of the CPU. The resolution of the captured video is 1920 $\times$ 1080 and the max speed of this card is 60 fps. Of course, a more advanced card can be used, if necessary, for the system to provide analysis results at a higher frame rate. Practices in NIAG 2021 have shown that real-time ball tracking and bounce distribution analysis result can be successfully delivered to the fans on the internet during live streaming.

The main target of table tennis live streaming is showing the ball trajectory and statistics of bounces on the table in real time. Bounce (ball touch-down on table) detection relies on analysis of temporally consecutive frames. Based on the fine-tuned YOLOv4 model, basically only one ball is detected at each frame. Assume now the match has proceeded to frame $t + 1$, and positions of the balls in these $t+1$ frames are $\boldsymbol{a}_1 = (x_1, y_1)$, $\boldsymbol{a}_2 = (x_2, y_2)$, ..., $\boldsymbol{a}_{t+1} = (x_{t+1}, y_{t+1})$. Starting from frame $t + 1$ backward to frame $t - K$, we calculate the spatial displacement between balls in every two temporally consecutive frames in the x- axis and y-axis, respectively. The displacements in the x axis are $\Delta\, x_t = x_{t+1} \; - x_t$, $\Delta\, x_{t-1} = x_t \; - x_{t-1}$, ..., and $\Delta\, x_{t-K} = x_{t-K+1} \; - x_{t-K}$; the displacements in the y axis are $\Delta\, y_t = y_{t+1} \; - y_t$, $\Delta\, y_{t-1} = y_t \; - y_{t-1}$, ..., and $\Delta\, y_{t-K} = y_{t-K+1} \; - y_{t-K}$. To determine whether the ball at the tth frame is a bounce on the table, two criteria are checked as follows:

1) Δx should be larger than a predefined threshold τ_x. Before a serve, players usually throw the ball on the table and the ball bounces vertically multiple times. To ensure the bounces before serving not to be confused with the bounces in play, we check the horizontal displacement between two consecutive frames. If the displacement Δx_t is not larger than the threshold τ_x , the ball at the tth frame cannot be a bounce in play.
2) A bounce at frame t on a table happens when the ball keeps falling from frame $t - K$ to frame t , and it reverses at frame $t + 1$. Specifically, we obtain the sign of the product of two consecutive vertical displacements $\Delta y_{i-1}$and Δy_i, $i = t - K, \ldots, t, t+1$, i.e., $s_{i-1} = sign(\Delta y_{i-2} \times \Delta y_{i-1})$. If $s_t < 0$, $\Delta y_{t-1} < 0$, and $\boldsymbol{a}_t = (x_t, y_t)$ is in the *table mask*, then we say there is a bounce on the table at the tth frame.

In real implementation, only the cases passing the first criterion will be examined by the second criterion. In this way, we can filter out many false positives to reduce complexity, and still be able to obtain accurate bounce detection results.

Once the bounce positions are detected, we project the positions onto a bird-view table (shown on the right of Fig. 8). Based on the positions of the four corners of the table mask and the positions of the four corners of the bird-view table, we can estimate the homography matrix $\boldsymbol{H}$ representing the projective transformation [22] between the two views. For a set of N bounce coordinates $\boldsymbol{B} = (\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_N)$, where $\boldsymbol{b}_i = (x_i, y_i)$, the corresponding coordinates on the bird-view table are obtained as $\hat{\boldsymbol{B}} = \boldsymbol{H}\boldsymbol{B}$, where $\hat{\boldsymbol{B}} = (\hat{\boldsymbol{b}}_1, \hat{\boldsymbol{b}}_2, \ldots, \hat{\boldsymbol{b}}_N)$ and $\hat{\boldsymbol{b}}_i = (\hat{x}_i, \hat{y}_i)$. The right part of Fig. 9 shows the statistics of bounces $\{\hat{\boldsymbol{b}}_i\}$ on the table. We can divide each half of the table into several regions and calculate the probability of bounces in each region. These visualized statistics not only can enrich the audience's viewing experience, but also can be used to analyze the players' performance or even team tactics.

D. AI-Based Multiswimmer Tracking for AR Flag Generation

Swimming competition is another story, as tracking the swimmers in all the eight lanes and create real-time AR content in an outdoor swimming pool are very challenging. The issue is that

Fig. 9. Example of live bounce statistics on the right-hand side.

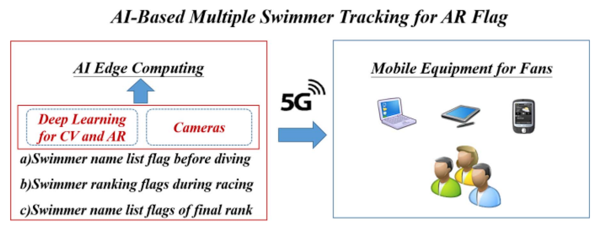


Fig. 10. Video streaming of the NIAG swimming competitions with AI-enhanced AR flags is first-time seen by 5G users and fans on the internet.

the sunlight keeps changing its strength and direction during the day, and weather condition is dynamic as well. Also, swimmers with part of their bodies in the water and the splashes of the water have been the major obstacle for detection and tracking. Therefore, improved AI models need to be developed.

In our outdoor swimming pool, we install only three regular Point Grey GS3-U3-23S6C cameras with 1920 × 1200 resolution. The AI-based system runs at 10 fps under the GTX 3080 accelerator on the edge computing server, as depicted in Fig. 10. The input RGB-channel values of the template and search region are preprocessed to reduce the sunlight effects. We use the ResNet-50 neural net model as the baseline subnetwork, but the number of layers is reduced and several shortcuts between stages are added to reduce the number of parameters and increase the detailed features, respectively. For the depthwise separable convolution approach adopted by the SiamRPN, all the eight swimmer templates can be processed in parallel to reduce the running time. We thus use a modified version of the SiamRPN++ algorithm [18] to track all the eight swimmers in the pool, individually and simultaneously. Note that the swimmers can be occluded by water splashes anytime, so we update the swimmer templates as soon as the tracking confidence is lower than a predefined threshold.

The AR flag view of a swimming competition video clip consists of three intervals: 1) name list flags before diving, 2) ranking flags during competition, and 3) name list flags of the final ranking, as shown in Fig. 11. Specifically, the AR view of both name list flag intervals is based on the *perspective projection mapping* from the virtual rectangular swimming pool to the real swimming pool. When we detect the eight swimmers, who stand on the diving boards or position at the starting line, the "name list flags before diving" process is triggered. An

(a)

(b)

(c)

Fig. 11. Three intervals of the AR view of a swimming competition: (a) name list flags before diving, (b) ranking flags during competition, and (c) name list flags of the final ranking.

AR flag is rendered for each lane at the water surface for 6 s, which contains the lane number, name of the swimmer, and his/her affiliated university or college, as shown in Fig. 11(a). After the match starts, the eight AR ranking flags will follow the eight swimmers and be rendered on the surface of the respective lanes, as shown in Fig. 11(b). The ranking flags are updated dynamically to show the correct ranking of the swimmers. Note that in Fig. 11(b), there are two perpendicular lines displayed on the pool surface, i.e., the national record line (yellow) and the game record line (red), which are also dynamically updated (moving). When each swimmer touches the finishing line, his/her name list flag is shown again on the water surface of the corresponding lane, which also contains the final ranking, swimmer and school names, and the elapsed time, as shown in Fig. 11(c). This proves very convenient for

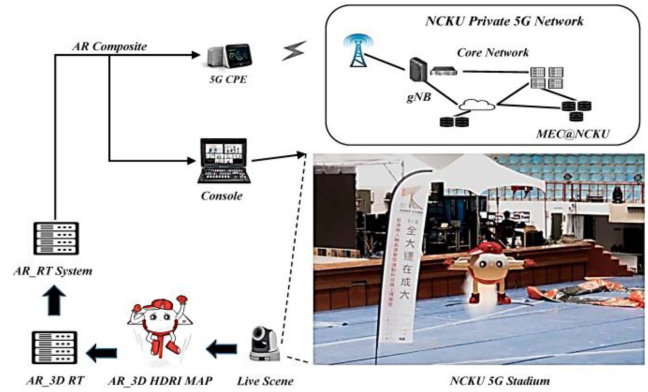


Fig. 12. AR-enriched content platform in the 5G stadium.

not just the online viewers, but also the onsite spectators with 5G viewing devices (cell phones, pads, laptops, etc.). Again, every swimmer is tracked concurrently and continuously, and the dynamic ranking list of all the eight swimmers is updated based on the relative locations of the center points of all the tracking bounding boxes (each representing a swimmer) along the y-axis (which is in parallel with the lanes). The AR flags of the ranking are mapped to the corresponding lanes on the water surface for the swimmers, as shown in Fig. 11(c).

E. Platform for AR-Enriched Content

Before the AR contents are integrated with the NIAG games in the 5G smart stadium and the swimming pool, we had practiced the AR-enriched content during the drone racing game. The AR technology allows virtual objects to interact with real objects and humans onsite in the video, during live streaming to the internet, as well as the on-site big screens and 5G users. This is done by predesigned scenarios for the virtual objects to be located at positions in the video of the real world, with real objects and humans onsite, through simultaneous localization and mapping of camera images. To make the overall performance more vivid and exciting, we have used an AR platform to build the NIAG 2021 mascot (see Fig. 12, in red). We turn the original 2-D image to an animated 3-D image that matches the characteristics of the mascot role, such as looking around the stadium and waving hands to the onsite spectators, and showing jet flames and smoke during flight.

Once the onsite live scenes are captured through the positioning 4K cameras, the 3-D animated mascot will automatically be embedded in the videos as the source of the *high dynamic range imaging* environmental light, and the shadow effect is automatically rendered. The instant rendering by the content platform and the 3-D-AR animation has matched the real scenes, which are then precisely positioned to combine with the virtual 3-D items in the stadium scenes. Finally, a complete animation performance was presented to show the characteristics and features of the mascot, as shown in Fig. 12. The result of the animation performance was also directed to a 250-inch onsite screen, so fans onsite knew what

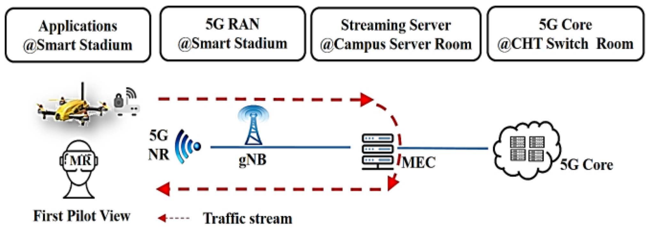


Fig. 13. FPV drone racing scenario for measuring the latency.

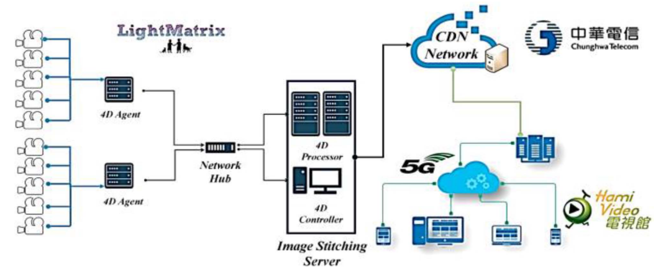


Fig. 14. Broadcasting platform supporting 3-D free view-angle replay.

was going on, and they were extremely excited about the technology.

F. Drone Racing and E2E Latency Measurement

The drone racers (pilots) wear MR goggles and handle remote controls that are both 5G enabled, which is a technology developed by the Industrial Technology Research Institute (ITRI)—through a customized 5G communication module and the private 5G network. The drone racing system is integrated together by CHT and ITRI, which fulfills the bandwidth requirement for real-time stitching of the video streams from all the 44 cameras mounted in the stadium. The 5G network provides a high-speed infrastructure required for the instant free view-angle playback application, and guarantees stable and smooth live transmission to the core network. Moreover, to enhance the experience of the instant free-view playback of the racing drones, highlights of the flight path are captured by dedicated cameras with more than 120 fps installed at different viewing angles of the stadium.

In drone racing games, the traditional FPV images are transmitted in analog form, whose quality is low—there are image flickers and jitters that lead to intermittent feeds. In the proposed TEBS, the FPV image frames from the drone camera are converted to compressed digital video stream for transmission to the MEC server via the 5G network, as shown in Fig. 13. Note that the drone is equipped with a 5G communication module developed by one of our industry partners. Also, by using an MR goggle equipped with the same 5G communication module, operating on the private 5G network, for receiving FPV image frames and sending pilot commands in real-time during the drone racing game, the image frames are very clear. The FPV image frames from the drones are streamed not just to the MR goggles of the respective competing pilots, but also the big screens in the stadium and the social network, so the onsite spectators and online viewers also enjoy the exciting competitions much the same as the pilots.

As shown in Fig. 13, the real-time FPV video is streamed to the MEC server on campus, through the private 5G network. Although the overall latency on the entire traffic path (shown as the red-dashed arcs in the figure) can be high, most of the critical computations are done at the edge layer, so it does not affect the application performance. The E2E latency measurement starts from the drone-side to the MEC server and then to the player-side. The FPV videos are in *high-efficiency video coding* form, with an advanced specification of 1080p/60fps/8Mbps. The latency along the traffic path is measured, and the average number for multiple sessions of drone flights is calculated, which is less than 30 ms.

IV. Real-Time Broadcasting With Free-View Replay

The high-speed 4K camera arrays are installed in the stadium to capture image sequences of the drone racing and sport games (table tennis and badminton). They capture 3-D images and generate a large amount of data that are processed on local servers prepared by our industrial partner, LightMatrix. The real-time broadcasting platform supporting 3-D free view-angle replay is shown in Fig. 14, which is composed of three major components. The first are the *4-D agents* that control and calibrate the cameras (shown to the left of the figure in two arrays), and provide the decoding/encoding media storage. Note that “4-D” means exact positioning of the cameras in the 3-D space, plus synchronization in the time domain. Each 4-D agent can handle up to 12 cameras, limited by the network bandwidth. The cameras capture videos of the same event from 12 different view-angles, which are to be stitched together to provide free views from different angles by the application or directly by the viewer on the end-point device. The second major component is the *4-D processor* that creates the integrated video stream to be forwarded to the *content delivery network* (CDN) operated by CHT. Finally, the *4-D controller* is mainly for converting the captured video streams to the standard *serial digital interface* format and/or other popular internet formats, depending on requirements of the streaming platforms.

Note that the 4-D processor and 4-D controller, together, are considered as the *image stitching server* (ISS). There is another component called the *network hub*, which connects all the major components together, as shown in the figure between the 4-D agents and the ISS. The live streaming experiments had been successfully carried out during NIAG 2021 through the private 5G network that meets the bandwidth requirement of the 3-D free view-angle playback application [14]. The CDN is based on CHT’s *video-on-demand* system that delivers 4K videos to their subscribers. This design ensures smooth, stable live transmission of the streaming service. Both online and offline users can control the playing speed and the forward/rewind play freely through a decoder-ready APP on the 5G endpoint devices. In addition, the video can be viewed in different angles, aiming at 360° playback and frozen images. The viewers can also zoom in/out with their fingers at will through the APP on their devices. Slow motion replay and the ball trajectory can be shown as well. The feature-rich system, thus, is reshaping the way fans participate in the sport events. Particularly, for onsite fans, the seamless and realistic display of the AR-embedded content can deliver an immersive experience they have never enjoyed before.

V. Discussions

A. Value-Added Game Streaming on Social Networks

With the rise of *social network services* (SNS), spectators and viewers have more options and ways to watch a sport game on different online platforms. The traditional logic of TV broadcasting does not meet the needs of the users of SNS, which touts online interaction and e-commerce for value creation. In addition to video streaming via our private 5G network, we have also set up a backup solution to transmit the live video content through the existing network infrastructure in the stadium, i.e., the WiFi/4G network, though it was never used during NIAG 2021. Note that the conventional broadcasting of sport events through TV channels are losing viewers fast as the viewers can only passively receive the content without interaction with either the platform, the content source, or other viewers. Also, with the increasing usage of mobile devices [23] and social networks, the online interaction between the live broadcaster (YouTuber) and the fans is the key of generating values on top of the new social media, which is quite different from the traditional one. Online viewers are keen to have more connections with the real world and meanwhile be immersed in the games that are full of features from the virtual world. Sports games and other types of live video streams that provide interactive service, therefore, is becoming essential to the young generations [24].

To resolve all kinds of issues associated with live streaming of high-quality sport games, we should, of course, take advantage of high-tech solutions. With the proposed TEBS, the free view-angle and real-time 3-D playback creates flexibility for the spectators and viewers to choose their preference angle by themselves, and they can even watch playback onsite when they have missed an event just happened or when they have any doubt about it. Besides, all kinds of sport data can be analyzed for different purposes, and objects (such as balls, rackets, and players) can be tracked, new applications and services can be integrated into the virtual cloud system that is basically without limit. Altogether, the TEBS increases fan engagement and raises the value of sport events through live streaming on social networks. The high-tech-enabled solutions also provide many options for the viewers to enjoy the same sport event by choosing from different live streaming sources and services.

Note that conventional sport events broadcast through TV channels are a one-way service in prescheduled time slots. However, though the content of the games in the proposed TEBS can also be broadcast through TV channels, we chose to broadcast the content through social network platforms by AI-enhanced video streaming (Fig. 1), which can be combined with all other SNS, especially online interaction and e-commerce for value creation. Therefore, direct fan engagement is not just for onsite spectators, but also online fans anywhere in the world, as interaction between the broadcasters and online fans becomes possible. Also, all other activities associated with the games, athletes, sponsors, etc. can be done online in real time and any time before/after the games.

TABLE II
On-Line Views and Fans Reached

Sport	Date	Views	Reach
Table Tennis	5/10/2021	3,000	6,800
	5/11/2021	4,400	10,600
Swimming	5/10/2021	7,200	10,900
	5/11/2021	5,200	7,800
	5/12/2021	3,800	5,500
	5/13/2021	6,200	9,500
	5/14/2021	3,500	5,400

Experimental results from NIAG 2021 about online views and potential fans reached for table tennis and swimming during the game time slots, respectively, are listed in Table II. During the period, each live streaming has attracted thousands of online viewers. Compared with other sports that have only TV viewers watching the games without AI-enhanced content, the table tennis and swimming competitions have reached far more viewers through our social-network broadcasting, engaging a record number of fans online. We believe this will gradually become the norm in future sport events and athletic games.

B. Contributions and Lessons Learned

What we have contributed in this article is to design and demonstrate the integrated system of systems, i.e., the TEBS, through the deployment at NIAG 2021, and show that the systems and applications that we have developed do work. One of the key characteristics of our system is that any consumer-level cameras with acceptable frame rate can be adopted. As opposed to existing commercial systems that need high-end industrial cameras, our solution is affordable, generic, and easy to deploy. Without binding to APIs specialized to industrial cameras, our software is portable to generic consumer-level cameras with acceptable automatic focus and exposure, reliability, and other quality control characteristics that we need.

Specifically, some tough issues that we have faced and resolved include the following conditions:

1) Managing the heterogeneous interfaces between parties and organizations (research teams, engineering teams, technology companies, service providers, students, etc.).
2) Establishing the private 5G network on university campus, facing privacy and safety challenges, as well as network integration issues.
3) Allocating university and commercial bandwidth among the supporting subsystems and target applications.
4) Incorporating a backup solution with the existing WiFi/4G infrastructure.
5) Developing feasible business models for all parties from fan engagement (which is an ongoing task).

We have tried to justify the value of 5G, which is a complicated and tedious collaboration task. At the end, the value is jointly demonstrated by the 5G network operator (CHT), 5G smart stadium and vertical applications owner (NCKU), equipment suppliers, software providers, system architects (the team leaders) and integrators (ITRI and industry partners), as well as service providers. Some quotable experiences and lessons learned in the development and deployment are listed as follows:

1) To be successful, *we must ensure that all stakeholders benefit from the integrated task*, which mainly lies in the high potential of business opportunities from fan engagement. That is why all parties have jointly developed the TEBS, and make the *availability of a test bed/field* a common goal, allowing us to really deploy the TEBS and test all the systems. Integration and operation of such a complicated system requires interdisciplinary talents, and we found that *interdisciplinary talents can be trained, mostly from real deployment.* The research teams, engineering teams, service teams, etc., all learned from each other and become more experienced in system integration.
2) In the table tennis system, the main challenge is not the video analysis, but the environment setup and system integration issues. As an example, the broadcasting process begins with video capture and sport data computing, which can be done easily by the research team on the fly using a local AI server in the lab setup. However, integrating a commercial camera in the stadium using HDMI cables with the AI server that was located 20-m away was a different story. It was eventually done by an experienced engineer to avoid video quality degradation, who connected two 10-m HDMI cables with an amplifier in between. Another example is that the camera can run at 120 fps, but we found no hard disk that can reliably store 120 frames of 1080p images per second. The interdisciplinary teams worked together to solve the problem. They found a way to conduct video analysis directly on the GPU card of the server that is equipped with 24 GB memory, successfully achieving an overall throughput of 55–59 fps.
3) In swimming, existing solutions for indoor pools did not work well. It took us months to work around the solution for tracking swimmers correctly in the outdoor swimming pool, where the sunlight keeps changing its strength and direction during the day, and weather condition is dynamic as well.
4) We have also observed that a successful landing site of technology, whether it is a sport field or stadium, should be one that can house and link all people's minds together, not just for the integration of the sport activities and technologies. The key success factor, therefore, is not just the integration of modules or fusion of images, but also *the fusion of people, their domain skills and knowledge, intelligence, experiences, social networks, and, most important of all, their minds*.

Looking forward, we believe that sports are basic human needs, which should be considered as a part of human culture. Therefore, it is quite natural that new industries can evolve from such needs. When technologies can be introduced to help innovate the sport events, innovations can migrate to all related industries, so far as human needs are taken care of, and that we understand where the culture in human societies are headed. However, when we introduced technologies into sports, we also found that the deployed systems demand high computing power, so very low-power and environmentally friendly systems will need to be developed in the future.

VI. Conclusion

We have introduced the TEBS for enhancing fan engagement and the NCKU 5G smart stadium that houses a private 5G network and showcases the TEBS. The TEBS contains a series of *application systems* that demonstrate AI-based sports technologies, for which we put together all sorts of commercial systems to enable live 4K video streaming on top of the 5G network. The entire system of systems, i.e., the hardware/network structure is depicted in Fig. 1, and the TEBS is shown in Fig. 3. The technologies integrated include table tennis ball trajectory and touch-down (bounce) distribution, badminton shuttlecock tracking and trajectory, AR-enriched content for real-time video streaming on social networks, real-time 3-D broadcasting with wide-field free view-angle, in-stadium video stream pushing by private 5G network with MEC, AI-based sport data analytics, etc. The tech-enhanced events are streamed to the social networks, e.g., Facebook, YouTube, and CHT Hami Video.

We believe the proposed TEBS approach is not just applicable to the live events we have presented, but the TEBS approach as detailed in the article can also be adapted for other venues where a 5G network and MEC servers can be deployed. Also, given that the technologies presented here are all portable, the live streaming of the AI-based sport contents can be applied to many other live events as well, as we use standard interface protocols in TEBS.

Finally, to enhance fan engagement, professional sports commentators and live broadcasters were invited to participate in the NIAG 2021 events. We have prepared video clips for table tennis and swimming from the full recorded live streaming [14] with English and Japanese subtitles, which are available online.[3]

Acknowledgment

The authors appreciate the collaborating industry partners for their great contribution in all the engineering tasks in this article, including Chunghwa Telecom, LightMatrix, Compal Electronics, DroneFighter, Jorjin Technologies, and Chunghwa System Integration.

References

[1] H.-C. Shih, "A survey of content-aware video analysis for sports," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 28, no. 5, pp. 1212–1231, May 2018.
[2] M. Stein *et al.*, "Bring it to the pitch: Combining video and movement data to enhance team sport analysis," *IEEE Trans. Visual. Comput. Graph.*, vol. 24, no. 1, pp. 13–22, Jan. 2018.
[3] P. Soltani and A. H. P. Morice, "Augmented reality tools for sports education and training," *Comput. Educ.*, vol. 155, May 2020, Art. no. 103923.

[3]https://drive.google.com/drive/folders/1Dl7UOWlB0Qw22wlthbywC_2ThWkRhxoX?usp=sharing

[4] S. A. Alhadad and O. G. Abood, "Enhancing smart sport management based on information technology," *IOSR J. Sports Phys. Educ.*, vol. 5, no. 5, pp. 19–26, Sep./Oct. 2018.
[5] K. Rangasamy, M. A. As'ari, N. A. Rahmad, N. F. Ghazali, and S. Ismail, "Deep learning in sport video analysis: A review," *TELKOMNIKA Telecommun., Comput., Electron. Control*, vol. 18, no. 4, pp. 1926–1933, Aug. 2020.
[6] C.-W. Wu, "Baseball and testing?," *IEEE Des. Test*, vol. 36, no. 6, pp. 88–88, Dec. 2019.
[7] S. Sinha, "Celebrating sports technology awards," *IEEE Potentials*, vol. 38, no. 3, pp. 35–38, May/Jun. 2019.
[8] M. Bhatia, "IoT-inspired framework for athlete performance assessment in smart sport industry," *IEEE Internet Things J.*, vol. 8, no. 12, pp. 9523–9530, Jun. 2021, doi: 10.1109/JIOT.2020.3012440.
[9] J. J. Martin, "Evaluation of Doppler radar ball tracking and its experimental uses," M.S. thesis, Dept. Mech. Mater. Eng., Washington State Univ., Pullman, WA, USA, Dec. 2012.
[10] W. Carroll and B. Hansen, "The new science of bioball: Sensors and software protect baseball pitchers from injury," *IEEE Spectr.*, vol. 56, no. 10, pp. 34–53, Oct. 2019.
[11] D. Calzada, "DeepBall: Modeling expectation and uncertainty with recurrent neural networks," M.S. thesis, Dept. Comp. Sci., Univ. Illinois at Urbana-Champaign, Illinois, USA, Apr. 2018.
[12] M. Gowda *et al.*, "IoT platforms for sports analytics," *GetMobile: Mobile Comput. Commun.*, vol. 21, no. 4, pp. 8–14, Dec. 2017.
[13] A. Nistala and J. Guttag, "Using deep learning to understand patterns of player movement in the NBA," in *Proc. 13th MIT Sloan Sports Anal. Conf.*, Mar. 2019, pp. 1–14.
[14] The 2021 National Intercollegiate Athletic Games (NIAG), May 2021. [Online]. Available: https://110niag.ncku.edu.tw/Module/Home/Index.php
[15] M. A. Lema *et al.*, "Business case and technology analysis for 5G low latency applications," *IEEE Access*, vol. 5, pp. 5917–5935, 2017.
[16] Y. Siriwardhana, P. Porambage, M. Liyanage, and M. Ylianttila, "A survey on mobile augmented reality with 5G mobile edge computing: Architectures, applications, and technical aspects," *IEEE Commun. Surv. Tut.*, vol. 23, no. 2, pp. 1160–1192, Apr.–Jun. 2021.
[17] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," *arXiv:2004.10934*.
[18] B. Li, W. Wu, Q. Wang, F. Zhang, J. Xing, and J. Yan, "Evolution of Siamese visual tracking with very deep networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 4282–4291.
[19] J.-S. Lim, M. Astrid, H.-J. Yoon, and S.-I. Lee, "Small object detection using context and attention," in *Proc. Int. Conf. Artif. Intell. Inf. Commun.*, 2021, pp. 181–186.
[20] F. Wang *et al.*, "Residual attention network for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 3156–3164.
[21] C.-N. Chen and W.-T. Chu, "How it flies and why it flies? Volleyball trajectory segmentation and classification," *IEEE Trans. Circuits Syst. II: Exp. Briefs*, vol. 68, no. 5, pp. 1591–1595, May 2021.
[22] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[23] W. Gantz and N. Lewis, "Sports on traditional and newer digital media: Is there really a fight for fans?," *Telev. New Media*, vol. 15, no. 8, pp. 760–768, 2014.
[24] K. Pires and G. Simon, "YouTube live and twitch: A tour of user-generated live streaming systems," in *Proc. 6th ACM Multimedia Syst. Conf.*, Mar. 2015, pp. 225–230.