

AI-Enhanced Equestrian Jump Analysis System

Siavash Mortaz Hejri
Ammar Ali
Javeria Munir



British Equestrian

British Equestrian is the National Governing Body for horse sports in the UK, affiliated to the Federation Equestre Internationale (FEI), the international governing body of equestrian sports. It exists to provide leadership, vision and purpose in steering the direction of equestrianism. Our patron is HRH The Queen.

British Equestrian is an umbrella organisation representing the interests of 3 million riders, vaulters and carriage drivers in Great Britain via 19 independent member bodies (14 members and five associates). Established in 1972, we are the largest representative body within the equestrian industry.





ABOUT THE PROJECT

When horses jump over obstacles, there are many important parameters that play a crucial role in scoring during a match. These include the number of steps they take between obstacles, the positioning of their legs (front, rear, right, and left) during takeoff and landing, and whether they touch the poles on the obstacles. A data analyst within a team is in charge of meticulously observing matches to explore and document this kind of information. However, this task is extremely time-consuming and prone to errors due to human limitations in vision and energy. Additionally, it represents a high cost for team leaders.



OUR SOLUTIONS AND CHALLENGES

To address these challenges, we aim to apply appropriate AI models to analyze every detail of the actions of horses in the field during equestrian matches. With the assistance of this system, team leaders can improve their analysis and consequently achieve better scores, while also making the process more affordable and efficient in the long term.



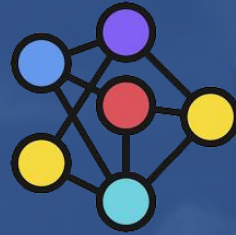
Jump is clear or not?

Is the jump clear? A clear jump is defined as one where no poles are knocked down. We utilize an AI model to detect horses during the play round, and subsequently pass the detected horse data, including detailed action information, through another AI model to classify whether the jump is clear or a failure. To achieve this, we begin by collecting data, including videos and a file containing useful information about the horses and their jumps. The video files encompass the entire round of play, prompting us to meticulously analyze each frame to detect horses and identify cleared jumps.



DETECTION

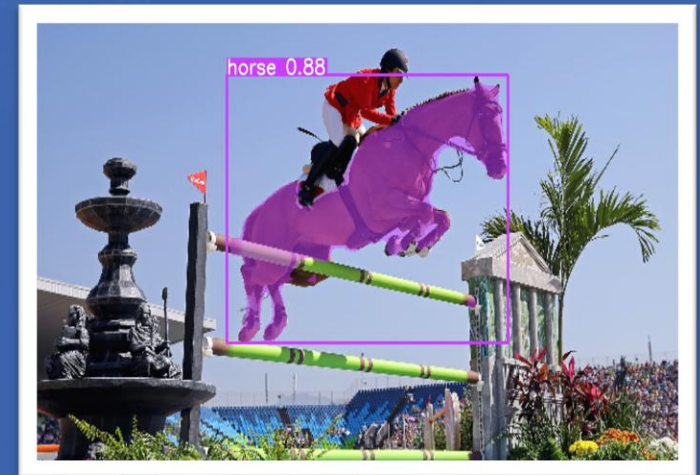
Video Frames



AI MODEL

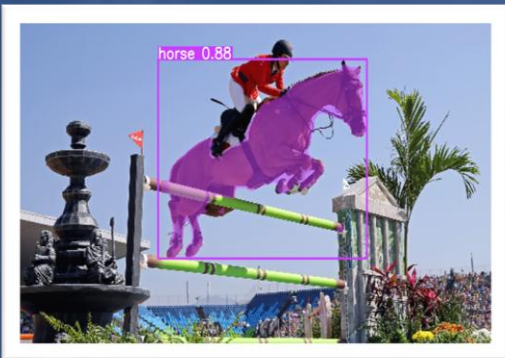


Horse Detection



CLASSIFICATION

Detected Horse



Classification AI Model

input

1
2
3
n

Hidden Layer

output

1
2

Labels

- 'Clear Jump'
- 'Fail Jump'

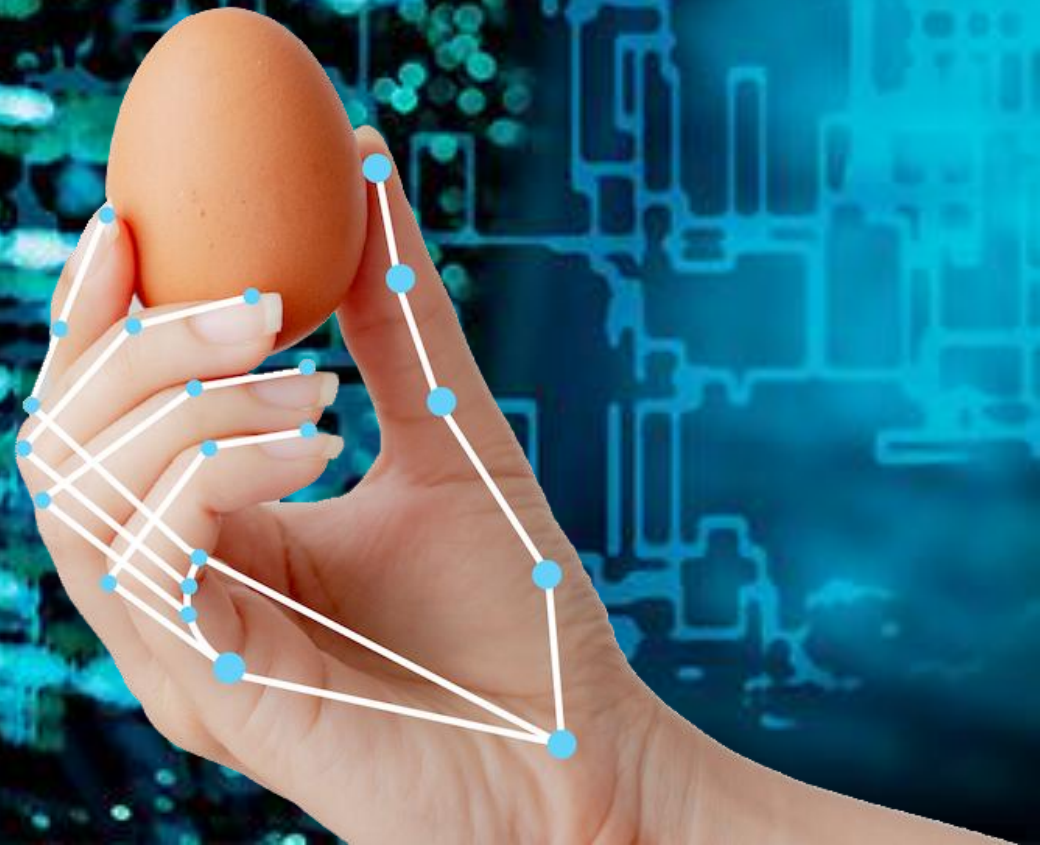
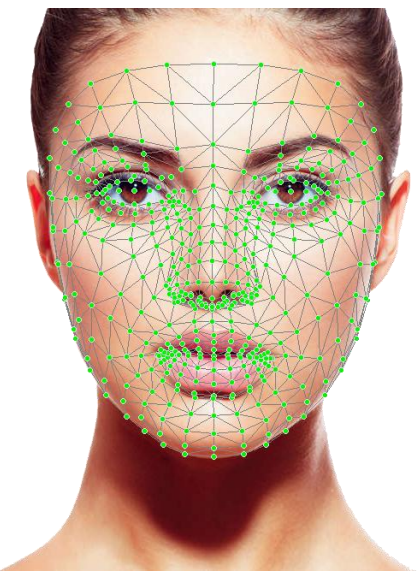


AI MODELS EXPLORED FOR IMPLEMENTATION

- ❖ MediaPipe
- ❖ Yolo v8
- ❖ CNN+RNN
- ❖ 3DCNN
- ❖ OmniMotion

MediaPipe

MediaPipe is a toolbox for developers who want to add cool features to their apps or projects, especially ones involving images or videos. It helps them do things like track hands or faces, detect objects, recognize gestures, and much more.



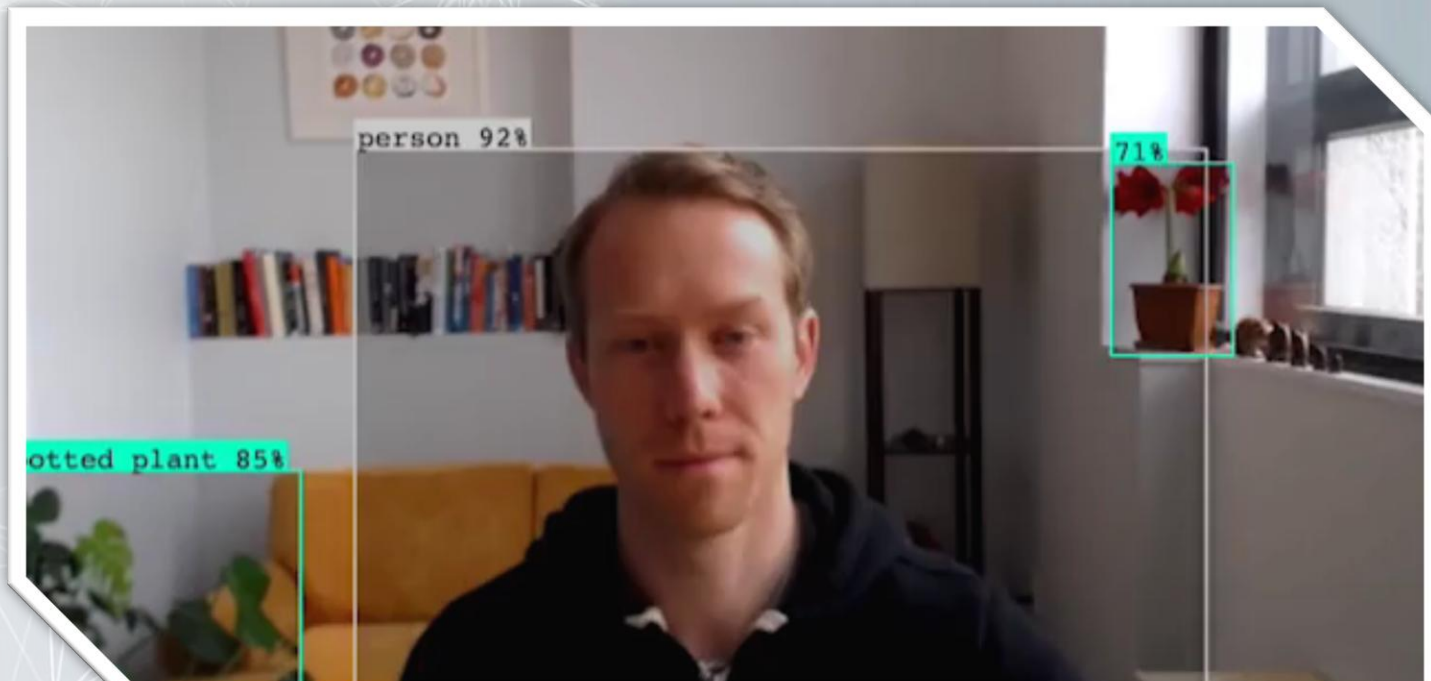
MediaPipe

Although this tool worked well for spotting horse legs and joints, it struggled when there were more than one horse or a person in the video, leading to mistakes in what it detected.





A type of computer program that's really good at spotting objects in pictures or videos. It's like having a smart helper that can quickly figure out what's in a picture or video, even if there's a lot going on. It's especially handy because it can do this detection super fast, making it useful for all sorts of things like self-driving cars, security cameras, or even just organizing your photo collection!



HORSE DETECTION



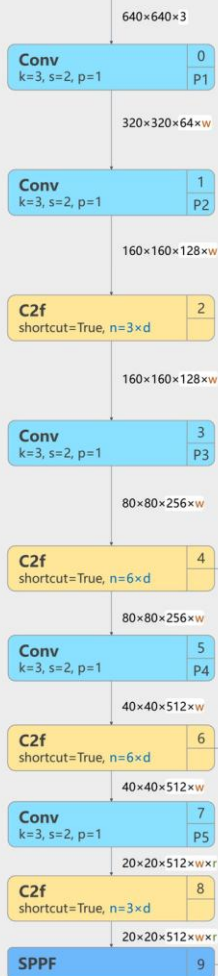
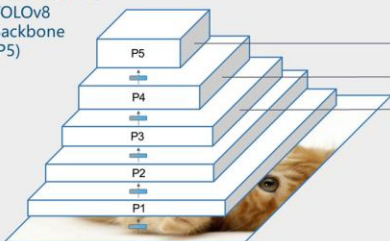


HORSE SEGMENTATION

YOLOv8

Backbone

YOLOv8
Backbone
(P5)

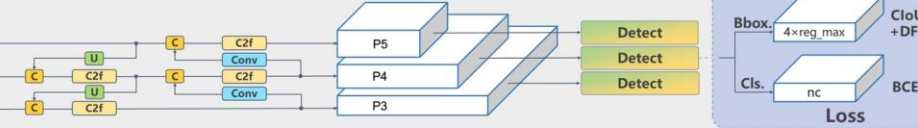


Note:
height*width*channel

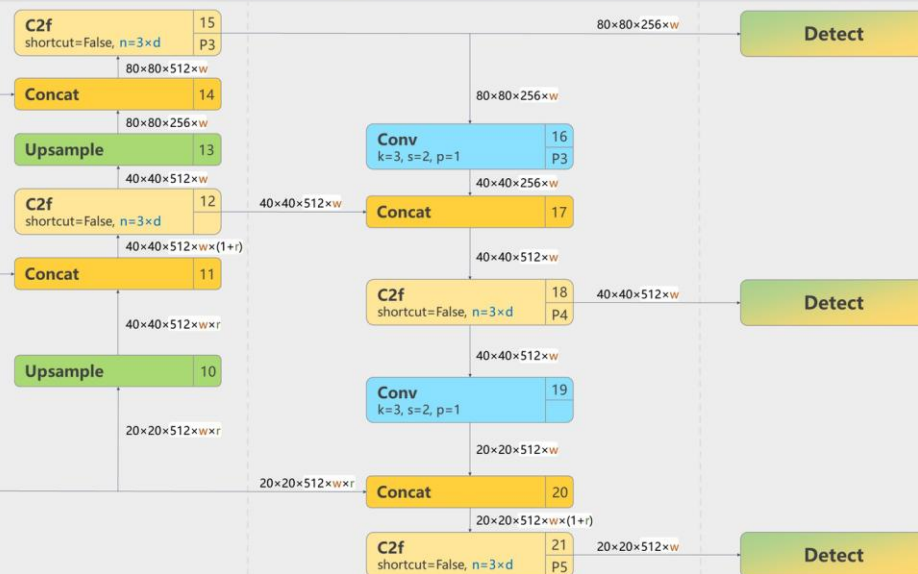
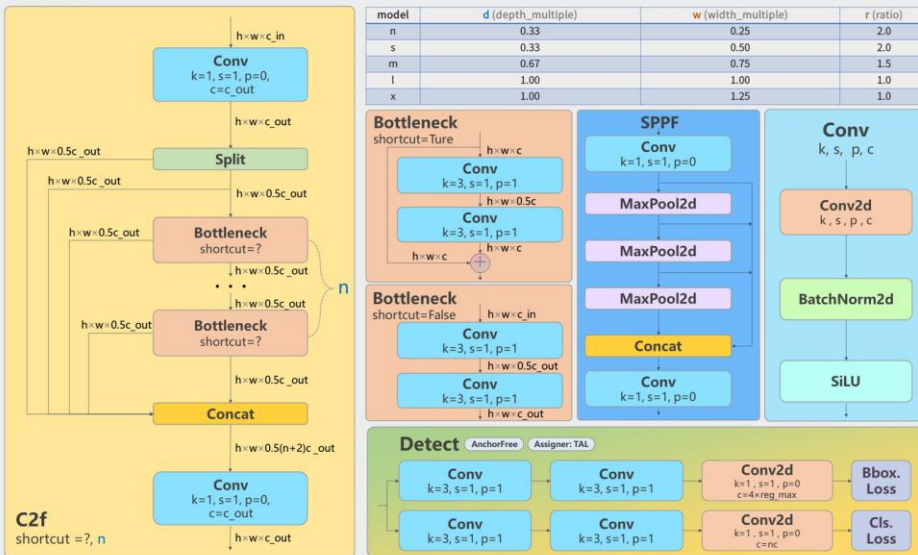
Backbone

Head

YOLOv8Head



Details



Head

While Yolo excels in detecting, segmenting, and classifying images, it struggles with action classification in videos. This is because Yolo models are designed to work with individual images rather than sequences of frames found in videos, which are typically labeled differently for classification purposes.



COMBINATION OF CNN AND RNN

A Convolutional Neural Network (CNN) is like a smart brain inspired by how our eyes work. It's great at looking at pictures and figuring out what's in them. Imagine it's like a puzzle solver: it breaks the picture into smaller pieces, analyzes each piece to understand its features, and then puts everything back together to tell you what it sees. It's used in lots of things like recognizing faces, identifying objects, and even helping cars drive themselves safely!

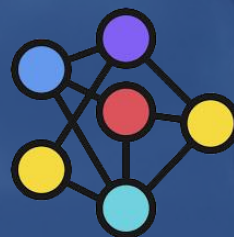
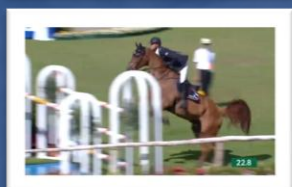
With the aid of these models in our project we extract features from videos frame by frame.

COMBINATION OF CNN AND RNN

A Recurrent Neural Network (RNN) is like a smart friend who remembers what you've told them before. It's great for understanding sequences of things, like words in a sentence or steps in a recipe. It looks at each piece of information one by one and keeps track of what it's seen so far. This helps it make predictions or decisions based on what it's learned from the past. RNNs are used in things like predicting the next word in a text message or understanding the context of a conversation.

Eventually, we pass sequence of features that comes from CNN model through these RNN models and training them for classification the jumps.

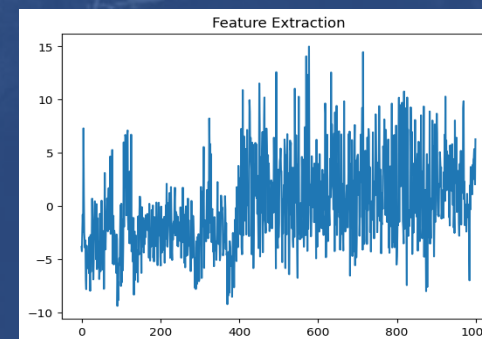
Jumping Frames



CNN Model

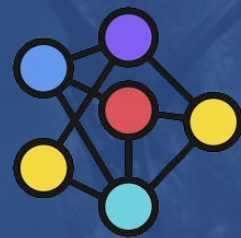
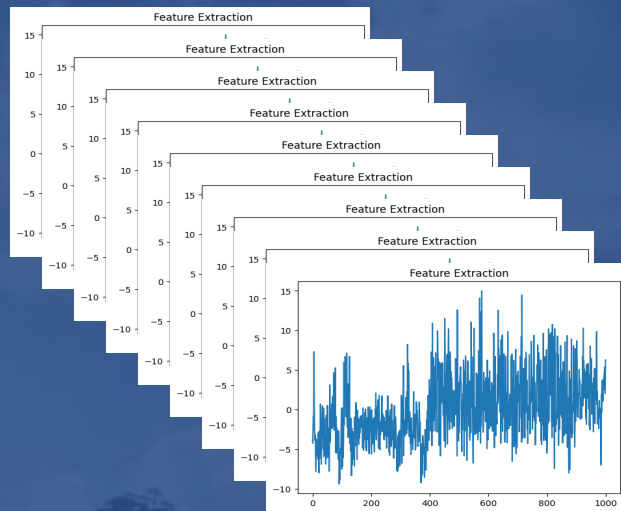


Extract Features



**COMBINATION OF
CNN AND RNN**

Sequence of Features



RNN Model



**Classification
Jump videos**



- 'Clear Jump'



- 'Fail Jump'

**COMBINATION OF
CNN AND RNN**

FUTURE APPROACH



3D Convolutional Neural Network, is like a super-smart brain that can understand both images and videos. It's great at recognizing patterns and features in three-dimensional space, which helps it understand moving images like videos. Just like how we see and understand things in the real world, 3DCNNs can analyze objects and movements in videos to recognize and classify them.



FUTURE APPROACH

OmniMotion “Tracking Everything Everywhere All at Once”

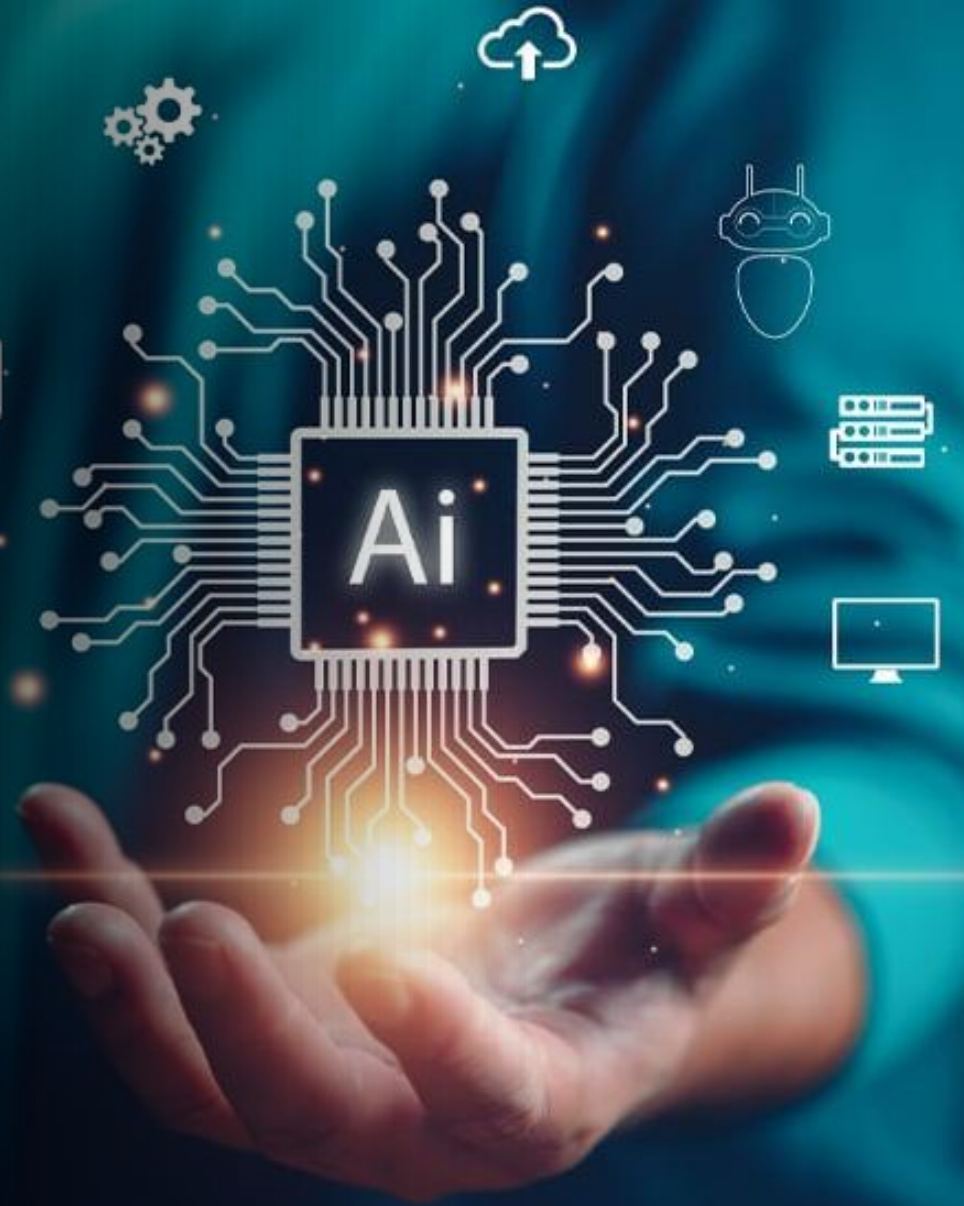
Imagine you're watching a movie with someone running really fast. Normal motion tracking in videos has trouble keeping up with all the movement, especially if the person runs behind a tree.

OmniMotion is like a super-powered motion tracker for videos. It uses a special 3D space to follow everything that's happening, even stuff that gets hidden for a moment.



FUTURE APPROACH

One challenge we face in videos captured by cameras is that during horse jumps, their legs often become obscured behind obstacles, affecting horse detection. Additionally, obstacles or horses may be masked by noise. In such cases, OmniMotion models aid in improving our system for detecting and tracking horses and obstacles.





We've split our tasks into three areas: research and documentation, project management, and programming. Each week, we rotate tasks to ensure thorough coverage and to learn new subjects from each other.

WELCOMING NEW DISCOVERIES IN PROJECT DEVELOPMENT



A photograph of a horse race in progress. Four horses and their jockeys are visible, running from left to right on a grassy track. The jockey on the leftmost horse is wearing a yellow and black outfit with 'OLIVIER' on the back. The horse has a blue saddle cloth with the number '1'. The second horse is brown with a jockey in pink and white, and a blue saddle cloth with the number '6'. The third horse is dark brown with a jockey in blue and white, and a blue saddle cloth with the number '5'. The fourth horse is grey with a jockey in red and white, and a blue saddle cloth with the number '8'. The word 'CHEERS' is overlaid in large, white, distressed capital letters across the middle of the image. A white horizontal line is positioned below the text. The background shows a cloudy sky, trees, and a distant building.

CHEERS