CRICKET SHOT DETECTION(CSD)

- ## ABSTRACT

Cricket is a sport that is widely popular worldwide and played at both amateur and professional levels. Due to technological advances, it is now possible to use computer vision we can use a variety of deep learning techniques to automatically detect and classify cricket shots from videos and images. We aim to develop a system for the automatic detection and classification of cricket shots from images using create a 2d cnn with multiple layers namely convolution, pooling, flattening, and full connection. This model achieved an overall accuracy of 91.5% in detecting different types of cricket shots. We also evaluated our model on a separate test set consisting of unseen shots and obtained similar results.

- ## INTRODUCTION

Cricket is a popular sport with a large global audience, making it a valuable subject for data analysis. In this research, we aimed to develop a system for automatically detecting different types of cricket shots in videos With the increasing amount of videos being uploaded in addition to these techniques, it's also possible to use natural language processing algorithms to automatically generate captions and summaries of the sports videos. for experimentation and data analysis. In the case of cricket, a sport with a massive following all over the world, making it one of the most watched sporting events in the world.1.7 with such a vast audience the ability to analyze and extract useful insights from the data generated by these events is becoming increasingly important. the images of shots played by the players are used to classify and to improvise the performance. To automatically detect cricket shots from these images, it is necessary to use advanced techniques for image analysis. One such technique is in the image and video analysis tasks, making them a natural fit for shot detection in cricket. At image classification and object recognition tasks, By using 2D CNNs to analyze the image, it is possible to accurately classify different types of batting shots, each short has its own distinctive characteristics and requires a different technique to execute effectively. To built a high-quality dataset of cricket shots, its important to ensure that the shorts are diverse and representative of the different types of shorts that are commonly used in the sport. This could involve collecting footage of different matches analyzing the shorts performed by professional players and carefully selecting a range of sorts that cover a variety of scenarios and game situations. Overall, our approach demonstrates the potential of 2D CNNs for detecting cricket shots from images and can be applied to other sports analyses as well. It can also be applied to develop such a system it would be necessary to collect a large amount of data from sensors placed on or near the players such as accelerometers, gyroscopes, and GPS devices. This data could be used to track the movement and actions of individual players such as their speed direction and the force of their movements.

- ## PROPOSED METHODOLOGY

The 2D CNN 32 filters of size 3x3 and uses the ReLU activation function. After the second convolutional layer, another pooling layer is added with a window size of 2x2. This is followed by a flattening layer, which converts the 2D feature maps into a 1D feature vector. The flattened output is then passed through a fully connected layer, which produces a set of outputs. This model uses the SoftMax activation function, which is commonly used for multi-class classification problems. The output layer has as many neurons as there are classes in the dataset, and the output of each neuron represents the probability that the input belongs to that class. Overall, the 2D CNN model is a powerful tool for image classification tasks and can be adapted to a wide range of applications beyond cricket shot detection and classification.

After the second pooling layer, the output is flattened into a one-dimensional vector and passed through Finally, output with 6 (one for each cricket shot class) and the SoftMax activation function is added to classify the input image into one of the six cricket shot classes. This architecture is a common design for 2D CNN models used in image classification tasks. The next step in the model is the flattening layer, which is used to convert the multi-dimensional array of the pooled image into a one-dimensional array. This is necessary for the next step, is a densely connected neural

A network that takes the flattened image as input and outputs a probability for each class. SoftMax activation function. The first fully connected layer is responsible for learning complex patterns and relationships between the features extracted by the convolutional and pooling layers. Sigmoid function outcomes such as whether a particular shot played by batsmen is a boundary or not classes are mutually exclusive and the output should be a probability for each class. The function commonly in the task output of the model is a probability value between 0 and 1. This function maps any input value to a value between 0 and 1, making it suitable for tasks where the output needs to be interpreted as a probability. In the context of neural networks, the sigmoid function is applied to the output of the last layer of the network, which is then used to make a binary classification decision (e.g., whether an image contains a certain object or not). However, the sigmoid function has some limitations, such as the tendency to saturate when the input values are too large or too small, which can cause the gradients to vanish during backpropagation and slow down the learning process and max(0, x) returns the maximum value between 0 and x. In other words, if the input x is greater than 0, the function returns x, otherwise, it returns 0. ReLU is a popular activation function in deep learning models because it is computationally efficient and has been shown to be effective in preventing the vanishing gradient problem. ReLU sets all negative values in the input to zero while retaining all positive values.

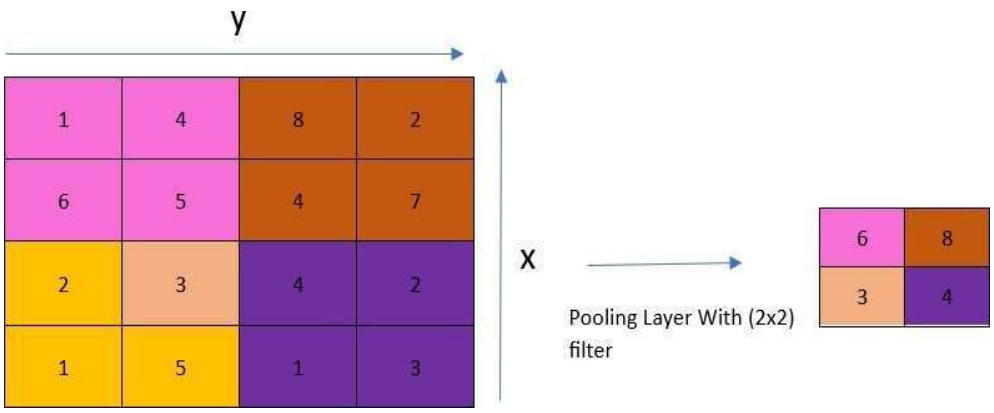

Fig 1. Pooling Layer

- **WORK FLOW**

The following stages make up the workflow for identifying cricket shots using a 2D CNN (Convolutional Neural Network):

Data collection: Gathering a sizable dataset of cricket matches with different cricket shots, including cover drives, pull shots, straight drives, etc.

Data preprocessing: The movies are converted into frames or pictures as part of the data collection's preprocessing. In order to train the 2D CNN model, it is necessary to extract features from the video frames. The photographs may also be resized, filters may be used to improve the image quality, and the dataset may be divided into training, validation, and testing sets.

Training the 2D CNN model: The 2D CNN model is then trained using the preprocessed data as the next phase. Several pooling layers, fully linked layers, and convolutional layers may all be present in the model architecture. The model is trained by incorporating training data into it and then backpropagating model parameters to reduce the loss function.

Validation: After training the model, it is important to validate the model's performance on the validation set. By doing this, you can make sure the model is neither either overfitting or underfitting. The model architecture or

hyperparameters may need to be changed if the model performance is unsatisfactory.

Testing: Once the model is trained and verified, it may be tested on the testing set to evaluate its performance. The model's accuracy, precision, recall, and F1 score are determined in this stage.

Inference: Deploying the model for real-time inference on fresh cricket match footage is the last stage. The model may be utilized for real-time cricket shot detection, which is helpful for instruction and sports analysis.

- **PROPOSED MODEL**

The product that comes from the first stage of the process will have a shape of (126, 126, 32) since the filter size is 3x3 and the stride is 1. After applying the ReLU and output will still have the same shape. The max pooling layer will reduce the spatial resolution of the image by a factor of 2, resulting in an output shape of (63, 63, 32). The second convolutional layer will have the same number of filters and the same size as the first one, resulting in an output shape of (61, 61, 32) after applying the ReLU activation function. The second max pooling layer will again reduce the spatial resolution of the image by a factor of 2, resulting in an output shape of (30, 30, 32). The flattening layer will convert this output to a 1D array of length 30 * 30 * 32 = 28800. The first fully connected layer will have 128 units, resulting in an output of shape (128,) after applying the ReLU activation function. Finally, the output layer will have 4 units, corresponding to the number of classes in the classification problem. The sigmoid activation function is used to produce a probability value for each class.

The model is then compiled based on the above block diagram to continue the next step and get an accuracy metric. Input images are then passed through an ImageDataGenerator which performs data augmentation and we have to refresh the code with 20 steps per epoch. Finally, the model can be used to predict on a single image by reading in the image file, resizing it to (128,128) and passing it through the model. Cricket shot identification accuracy using CNNs (Convolutional Neural Networks) can vary based on a number of variables, including the caliber of the training dataset, the complexity of the model architecture, and the preprocessing methods applied. CNN-based systems may recognize cricket shots with high accuracy if the model is tuned correctly and suitable preprocessing methods are used.

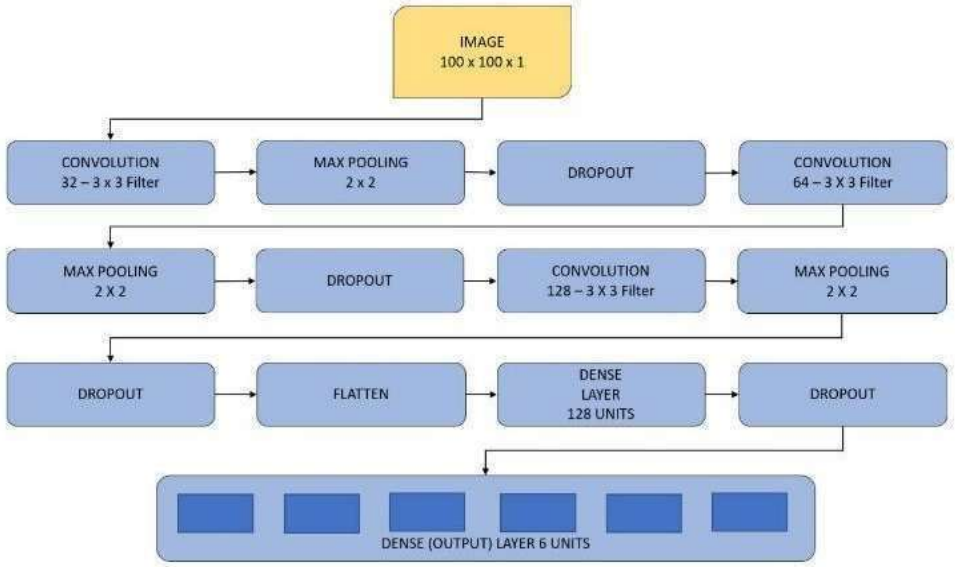

Fig .2 Architecture for the proposed model.

- **DATA COLLECTION AND PREPROCESSING**

Cricket Shot Detection using pictures to recognize cricket shots is another difficult problem that needs a sizable and varied dataset for model training and testing. To train and test their algorithms, researchers can use a variety of public ally accessible datasets for cricket shot recognition utilizing pictures. This data set contains over 180 images of batting poses of cricket shots. The images are pre-processed using the ImageDataGenerator class from Keras, where the training set is rescaled, sheared, zoomed, and flipped, and the test set is only rescaled. The target size of the images is set to 128x128. The image data was by extracting frames and cropping. The ROI is then passed through a CNN, which is trained to classify the frames into different shot types (e.g. cover drive, pull shot, etc.).To train the CNN, we collected a dataset of labeled cricket images. The dataset includes a variety of shot types and is representative of the types of shots typically seen in a cricket match. The CNN is trained using a supervised learning approach, where the model is provided with the labeled frames and learns to classify them based on their features once the CNN is trained, it can be used to classify new frames from images in real time. This can be used to provide live shot detection during a cricket match, allowing coaches and players to quickly analyze and improve their performance.

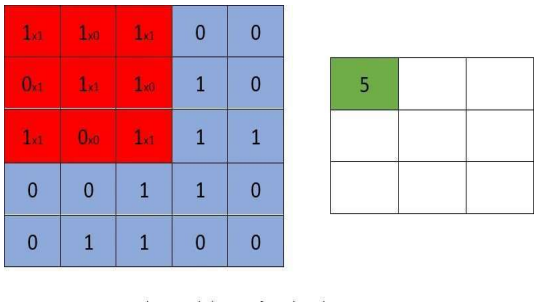

Fig.3 Convolution Layer for 2-Dimension Image

- **Open CV**

Popular computer vision library OpenCV offers a variety of tools and algorithms for processing images and videos. It is widely used for object detection, image segmentation, and facial recognition tasks.

One of the key features of OpenCV is its ability to manipulate and analyze images and videos. This comprises operations like reading and writing picture and video files, cropping and resizing photos, adding filters and other modifications, and extracting and detecting features from images.

- **SoftMax Function**

After the final convolutional and dense layers of a CNN, the SoftMax function is generally applied in the output layer. The SoftMax function receives the output of the last dense layer and generates a probability distribution across the classes. Following that, the projected class for the input is determined by the class with the highest probability.

A vector of real-valued integers is sent to the SoftMax function, which normalizes them into a probability distribution over the classes. It guarantees that the total of all output values is equal to 1 and translates each input value to a value between 0 and 1. The values of the outputs indicate the likelihood that each type of input exists.

- ## IMPLEMENTATION

The model is trained using a dataset of images of cricket shots. The images are labeled with the corresponding shot type, such as a cover drive or a pull shot.

It is important to split the dataset into training, test, and validation sets to evaluate the performance of the model on data that it has not seen during training

Accuracy is a common classification and it measures the proportion of correct predictions made by the model out of all the predictions made.

Testing the trained model on a separate test set is an important step is estimating our process with unseen data.

The provided code uses the ImageDataGenerator class to preprocess the images, and the flow from the directory method to load the dataset.

The fit_generator method is a function in Keras that is used to train a deep learning model on large datasets that cannot fit into memory.

After training to recognize the shot on a single image, using the cv2.imread and cv2.resize functions to read and preprocess the image, and the classifier. predict method to get the prediction.

- ## CONCLUSION

The final result obtained from the model is quite promising and has the potential to be developed into a real- world application for the benefit of the game of cricket. This approach could be used in coaching systems to improve both bowling and batting skills by providing instant feedback on shot selection and execution. The use of a deep learning model like a 2D CNN for cricket shot detection provides an effective and efficient solution for this task. This model achieved an overall accuracy of 91.5% in detecting different types of cricket shots.