

PROJECT REPORT

T-REX DINO USING CNN

Team Members – Lakshmi Prasanna, Vyshnavi

INTRODUCTION

Author –Vyshnavi

What AI system are you interested in investigating further?

The AI system we are investigating is **T-Rex Dino game using CNN** (Convolutional Neural Network). This game is a browser game developed by Google and built into the Google Chrome web browser. The player guides a pixelated T-rex across a side-scrolling landscape, avoiding obstacles to achieve a higher score. The obstacle can be a cactus or a bird. If the obstacle is a cactus, then we must press up arrow to jump and if the obstacle is a bird, then we must press down arrow to duck.

Author – Vyshnavi

Why is it of interest to you?

Chrome users find themselves without Internet connection occasionally, and when this happens, they enjoy playing the popular T-Rex, Run! Game.

Author – Vyshnavi

What specifically your project will focus on? What's the focus of your investigation?

In this project, we want to focus on predicting the next move of the dino based on the data provided by playing the game manually.

We want to focus on how to collect the data, train a CNN model using keras and sklearn and how to make the trained model play the game. We want to visualize the trained model play the game using tensorflow-gpu.

THEORETICAL BACKGROUND

Author - Prasanna

What area or field of AI provides the theoretical framework for the chosen AI system?

T-Rex-Runner uses a **Convolutional Neural Network (CNN) model with Keras** that takes an input of the game screen and outputs whether it should or should not jump at that specific moment.

It can capture the screen, evaluate it, and output its calculations in real time to play the game.

Convolutional Neural Network Architecture

First we need to create a new neural network object in Keras. To do that, we create a new sequential object. So we say `model = sequential()`. The sequential api lets us create a neural network by adding new layers to it one at a time.

The core features of the model are as follows :

- Input shape consists of (WIDTH, HEIGHT, 1) as values.
- First layer, **Conv2D** consists of 32 filters and 'relu' activation function with kernel size, (3,3).
- Second layer, **Conv2D** consists of 64 filters and 'relu' activation function with kernel size, (3,3).
- Third layer, **MaxPooling** has pool size of (2, 2). Max pooling is where we scale down the output of the convolutional layers by keeping only the largest values and throwing away the smaller ones.
- Fourth layer, **Dropout** has 0.25 as its value.
- Fifth layer, **Flatten** is used to flatten all its input into single dimension.
- Sixth layer, **Dense** consists of 128 neurons and 'relu' activation function.
- Seventh layer, **Dropout** has 0.4 as its value.
- Eighth layer consists of 3 neurons and 'softmax' activation function.

References:

- [1]. Valeti Dilip, 2021 "Convolutional Neural Network Using Keras".
<https://medium.com/@dilip.voleti/convolutional-neural-network-using-keras-e8c5f4ff8a37>
- [2]. Dabakoglu Caner, 2018 "What Is Convolutional Neural Network (CNN) ? — With Keras".
<https://medium.com/@cdabakoglu/what-is-convolutional-neural-network-cnn-with-keras-cab447ad204c>.

Author - Prasanna

What are the key concepts and techniques that the AI application is based on?

T-Rex-Runner uses a Convolutional Neural Network (CNN) to implement the CNN model, we will do these in order:

- Firstly, we are going to create a dataset by playing the game for a while.
- To create a dataset, we are going to grab the screen while playing the game and record which key was pressed at that moment.
- After the dataset is created, we are going to train a classification model which is prepared using CNN. We can train the model by calling **model.fit** and pass in the training data and the expected output. Keras will run the training process and print out the progress to the console.

Keras in CNN first creates a new instance of a model object and then add layers to it one after the another. It is called a sequential model API. We can add layers to the neural network just by calling `model.add` and passing in the type of layer we want to add.

We need to specify the **ROI** (Region of Interest) according to screen resolution. We need to change limits (width and height) for the specified ROI.

When each press direction buttons, it saves ROI pixels by taking the screenshot using **mss** (multiple screenshot module).

References:

[1]. Pramoditha Rukshan, 2022 “Convolutional Neural Network (CNN) Architecture Explained in Plain English Using Simple Diagrams.” <https://towardsdatascience.com/convolutional-neural-network-cnn-architecture-explained-in-plain-english-using-simple-diagrams-e5de17eacc8f>.

DEVELOPMENT APPROACH

Author – Vyshnavi

What tools, platforms, and APIs are available and needed to develop the AI application?

Image from PIL - The Image module provides a class with the same name which is used to represent a PIL image. The module also provides a number of factory functions, including functions to load images from files, and to create new images.

MSS(Multiple Screenshots) - An ultra-fast cross-platform multiple screenshots module in pure python using ctypes.

Keyboard - Python provides a library named keyboard which is used to get full control of the keyboard. It can hook global events, register hotkeys and simulate key presses.

uuid - helps in generating random objects of 128 bits as ids. It provides the uniqueness as it generates ids on the basis of time, Computer hardware.

os - The OS module in Python provides functions for creating and removing a directory, fetching its contents, changing and identifying the current directory.

glob - In Python, the glob module is used to retrieve files/pathname matching a specified pattern.

Numpy - It is a Python library that provides a multidimensional array object, various derived objects, and an assortment of routines for fast operations on arrays, including mathematical, logical, shape manipulation, sorting, basic linear algebra and basic statistical operations.

Keras - Keras is a high-level, deep learning API developed by Google for implementing neural networks. It is written in Python and is used to make the implementation of neural networks easy. It also supports multiple backend neural network computation.

Sklearn - It provides a selection of efficient tools for machine learning and statistical modeling including classification, regression, clustering and dimensionality reduction via a consistency interface in Python.

Tensorflow – It helps to create machine learning models for mobile, web, desktop and cloud.

Author – Vyshnavi

How did you carry out the plan?

Initially all the modules required for the project are imported. Then developed the code to teach t-rex dodge obstacles in 3 stages- collect data, train the model, test the model.

Collecting the data:

- We will run the code and then play the game manually. The keyboard actions will be saved using keyboard library.
- We need to set the co-ordinates in order to cut off the area that we want the model to see only using mss.
- A function should be defined for recording the screen in which the images with area which we want to cut off using mss will be grabbed and saved.
- An exit function will be defined and whenever we want to exit in recording, we will press esc and then exit function will be called.

Training the model:

- Binary values will be generated for the labels (UP, DOWN, RIGHT) using Label Encoder and One Hot Encoder.
- Collected images will be resized and normalized. Then data will be split into 2 parts – one part for training and another for testing.
- A CNN model will be developed using keras models and keras layers and then trained.
- The trained model will be saved for testing.

Testing the model:

- Model will be tested in the game as real-time.
- Labels will be determined again because the model predicts images, we make it choose in this list according to output of the model.
- Set up the sizes for the screen and load the model.
- Apply resize and normalization.
- `argmax()` is used to find the label maximum probability.
- If the result is 0 it means DOWN, using keyboard library, model will press DOWN.
- If the result is 2 it means UP, again using keyboard library model will press UP.

Author – Vyshnavi

How did the team divide the work?

Code development for collecting the data and testing the model was done by Vyshnavi and code development for training the model was done by Prasanna. Dataset was maintained by both.

WORKING EXAMPLE

Author – Vyshnavi

What example illustrated the basic components, processes and outcomes of the chosen AI application?

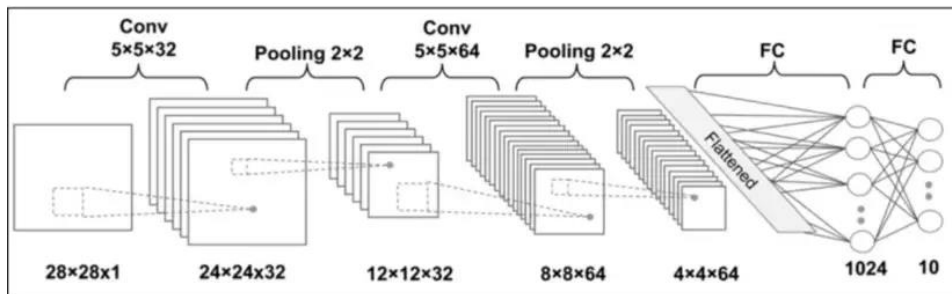
Image classification for CIFAR dataset demonstrates the basic components, processes and outcomes of the AI system T-Rex Dino using CNN.

Author – Vyshnavi & Prasanna

Describe the working example.

The example uses CIFAR dataset which consists of thousands of images of different objects. In this example, a CNN model is developed using keras to recognise those objects. Samples of images of each object will be used for training the model. The trained model should be able to recognise the images of objects.

CNN Framework:



The summary of model will look like:

Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 32, 32, 32)	896
conv2d_5 (Conv2D)	(None, 30, 30, 32)	9248
max_pooling2d (MaxPooling2D)	(None, 15, 15, 32)	0
dropout (Dropout)	(None, 15, 15, 32)	0
conv2d_6 (Conv2D)	(None, 15, 15, 64)	18496
conv2d_7 (Conv2D)	(None, 13, 13, 64)	36928
max_pooling2d_1 (MaxPooling2D)	(None, 6, 6, 64)	0
dropout_1 (Dropout)	(None, 6, 6, 64)	0
flatten_1 (Flatten)	(None, 2304)	0
dense_2 (Dense)	(None, 512)	1180160
dropout_2 (Dropout)	(None, 512)	0
dense_3 (Dense)	(None, 10)	5130
Total params: 1,250,858		
Trainable params: 1,250,858		
Non-trainable params: 0		

Author – Vyshnavi

Describe the data the AI application uses.

The dataset contains 60,000 32x32 color images in 10 different kinds of objects, with 6,000 images in each type. The dataset is divided into 50,000 training images and 10,000 testing images. Each image in the dataset includes a matching label so we know what kind of image it is. This dataset will be used to train our neural network to recognize any of these 10 different kinds of object. The dataset is maintained by Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton.

EVALUATION

Author – Prasanna

What are the safety and trustworthy considerations for developing your project's AI application example?

AI Performance

AI system performance may be impaired if the testing of AI models is ineffective or outcomes are unclear. A plan defining AI performance targets should be established at the outset of AI system development, typically with traffic light indications of acceptable ranges, e.g. of precision and recall. Where operational performance levels cannot be established before a model is trained, the training exercise can be viewed as a form of prototyping, with detailed targets, redefined ahead of formal testing.

Privacy

The GDPR requires that a data privacy assessment must be performed on all training datasets. If the datasets contain personal data, they are subject to the same privacy controls as production data.

Replication of outcomes

When an AI system fails, the ability to reproduce the failure in a test setting is an important engineering design consideration, particularly if the reasons why an AI system made a decision or a prediction is opaque. Many of the subsequent policy considerations depend on repeatable processes being in place.

References:

[1]. 2022 “10 Critical Considerations When Developing an AI System | BCS.”
<https://www.bcs.org/articles-opinion-and-research/10-critical-considerations-when-developing-an-ai-system/>.

Author – Prasanna

What are the ethical considerations for developing your project's AI application example?

The legal and ethical issues that confront society due to Artificial Intelligence (AI) include privacy and Transparency, Technical safety and Bias in Data, Training Sets, etc.

Technical Safety

The first question for any technology is whether it works as intended. Will AI systems work as they are promised, or will they fail? When they fail, what will be the results of those failures? And if we are dependent upon them, will we be able to survive without them?

Transparency and Privacy

Once we have determined that the technology functions adequately, can we understand how it works and properly gather data on its functioning? Ethical analysis always depends on getting the facts first-only then can evaluation begin.

Bias in Data, Training Sets, etc.

Algorithmic bias is one of the major concerns in AI right now and will remain so in the future unless we endeavor to make our technological products better than we are.

Automating Ethics

One strength of AI is that it can automate decision-making, thus lowering the burden on humans and speeding up - potentially greatly speeding up—some kinds of decision-making processes. However, this automation of decision-making will present huge problems for society, because if these automated decisions are good, society will benefit, but if they are bad, society will be harmed.