

A cheetah cub is peeking through tall, dry, golden-brown grass. The cub's face is visible, showing its characteristic spots and a focused expression. The background is a dense field of similar grass, creating a natural, savanna-like setting.

Safari Stickerbook

Unleashing Creativity, One Sticker at a Time!

Dipta Roy - Capstone



Problem

- No easy-to-use apps available to create an animal stickerbook.
- As someone with experience working with kids, a creative tool like this can spark imagination and fun.
- Only software that are available are photoshop or other tools that are too intricate and not tailored towards ease of use.

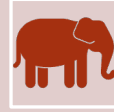
Solution

A web app to identify animals in a given picture and collect them in a virtual stickerbook.

Target Audience and Users



Schools: it can teach students/kids about animals and let them make a stickerbook of their favorite animals.



Zoos: Zoos can have an interactive way to get their zoogoers' involved.



Graphic Designers/Photographers: They can use to crop out animals and put it on another picture of their choice.



No photoshop skills required!

How does it work?



- A picture will be uploaded



- Uploaded picture will be fed to the classification model (YoloV8 architecture)



- Once identified a bounding box of the identified animal will be sent to Segment Anything Model to retrieve a mask



- Mask will be isolated to show a segmented image



App Roadmap



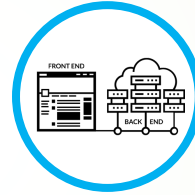
Model Development and Testing

Using YOLOv8 trained on COCO dataset and Segment Anything Model by Meta



Creating Front End

Front end creation through HTML, Javascript, and CSS



Connecting Front End and Back End

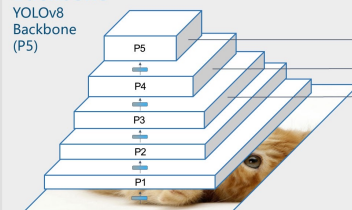
Using Flask and python



Deploying to the web

Deployed through the use of Google Cloud Run

Backbone

YOLOv8
Backbone
(P5)

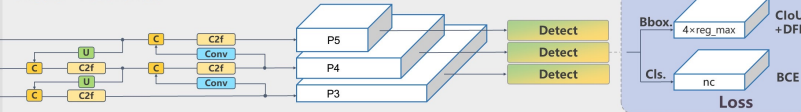
- Model created by the Ultralytics team based on prior iterations of YOLOv5

- YOLO divides an image into a grid and predicts bounding boxes & class probabilities for each grid cell.

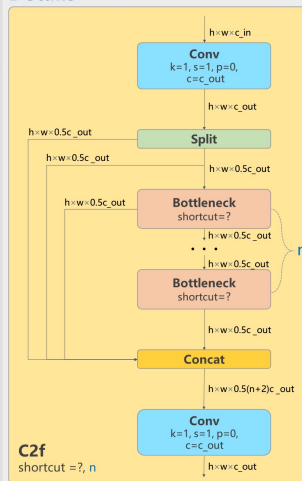
- YOLOv8 uses a convolutional neural network split into two main parts: the backbone and the head.

- The backbone is based on the CSPDarknet53 architecture with 53 convolutional layers.

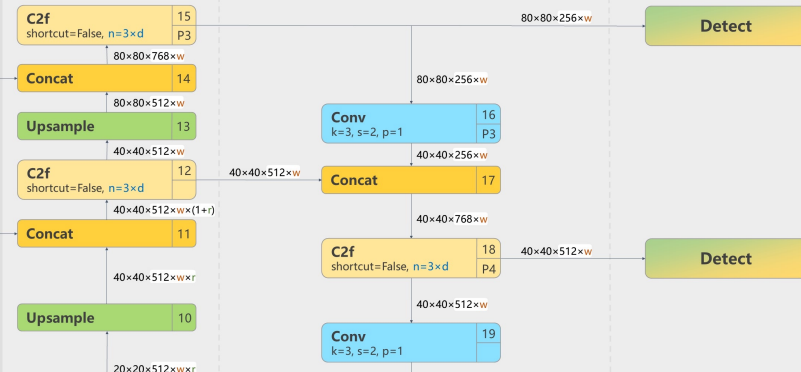
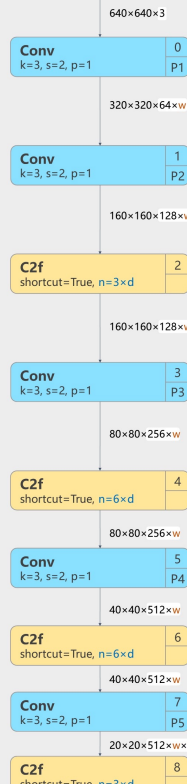
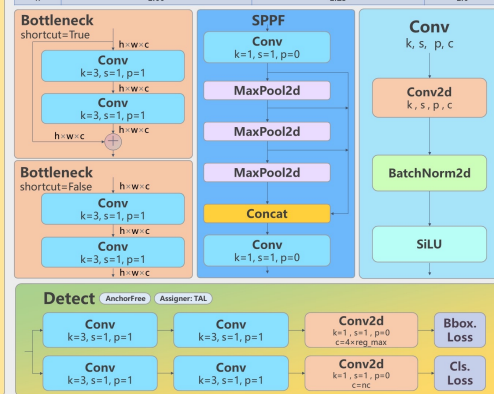
- The head focuses on predicting bounding boxes, objectness scores, and class probabilities.

Head YOLOv8Head

Details

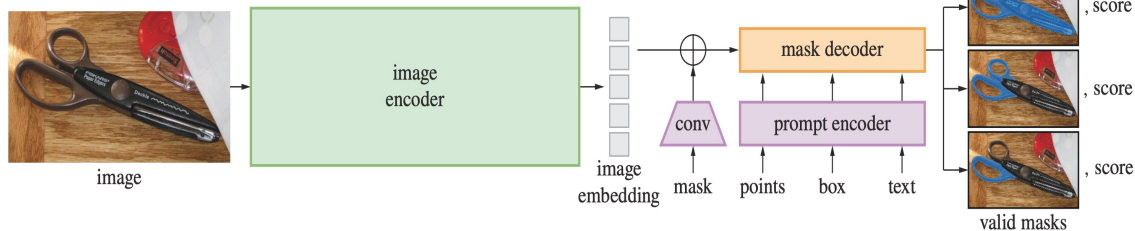


model	d (depth_multiple)	w (width_multiple)	r (ratio)
n	0.33	0.25	2.0
s	0.33	0.50	2.0
m	0.67	0.75	1.5
l	1.00	1.00	1.0
x	1.00	1.25	1.0



A little about Segment Anything Model (SAM)

- Model created by the Facebook AI research team at Meta
- Trained on over 11m high resolution dataset in three different stages using a Vision Transformer (ViT):
 - Manual annotation stage
 - Semi-manual annotation stage
 - Automatic annotation stage
- Images get converted into embeddings that goes through a convolution layer and a mask is decoded given certain prompt





Impact of Project

- Effortlessly craft beautiful canvases without a learning curve.
- Can be incorporated into classroom activities, enhancing learning with fun.
- Encourages cognitive development, fine motor skills, and imaginative play.
- No need to learn or invest in expensive software!

Coursework that enabled this project



Web Engineering:

Developed using Javascript, HTML, and CSS.

Ensuring a seamless, simple and user-friendly interface.



Security:

Safe Interactions:
Implemented CORS for secure cross-origin requests.

Accepted only specific file extensions, ensuring robustness and safety.



Machine Learning:

Grasped the fundamentals of model architectures.

Utilized Python for implementing and managing our models.



Algorithms:

Delved into intricate data structures.

Facilitated smooth data transfer between frontend and backend systems.

Learnings Beyond the Technical



ADAPTABILITY
LIKE NAVIGATING
UNFORESEEN
CHALLENGES
AND ADJUSTING
PROJECT PLANS.



TIME
MANAGEMENT
LIKE BALANCING
IDEAL FEATURES
WITH TIME
CONSTRAINTS.



IMPORTANCE OF
RESEARCHING
WHAT
TECHNOLOGIES
ARE ALREADY
OUT THERE.



KNOWING WHEN
TO PUSH
THROUGH THE
CHALLENGES
AND WHEN TO
REACH OUT FOR
HELP OR EVEN
TAKE A BREAK.



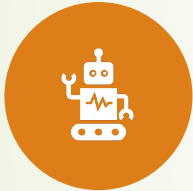
MAKING A
PRODUCT I'D
USE.

Retrospective

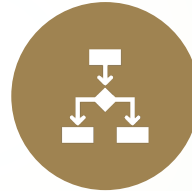
- If I were to do it over again:
 - Prepare a better timeline so I wouldn't be rushing to get things across.
 - Started front end sooner.
 - Trained the YOLOv8 model to be more robust and classify more animals.
 - Trained segmentation model from scratch and compare to the SAM model.



Understanding the life of an engineer



Combination of practical and theoretical.



Balancing design ideas and functionality.



In the age of AI, it's also an ethical duty of engineers to build for good.



A LOT OF RESEARCH!



THANK YOU!