Bird Scouts

- Aditya
- Krishna
- Nagasai
- Sanyat
- Sathvik
- Shankaradithyaa

- Team 6

- Project Objective
- Data Collection
- Audio Pipeline
- Image Pipeline
- Chatbot & other features

Project Objective

To develop a user-friendly web application that leverages machine learning to accurately identify bird species from audio recordings and images. Bird Scouts aims to:

- Facilitate quick and reliable bird identification for researchers, birdwatchers, and nature enthusiasts.
- Enhance understanding of bird species through automated analysis of vocalizations and visual features.

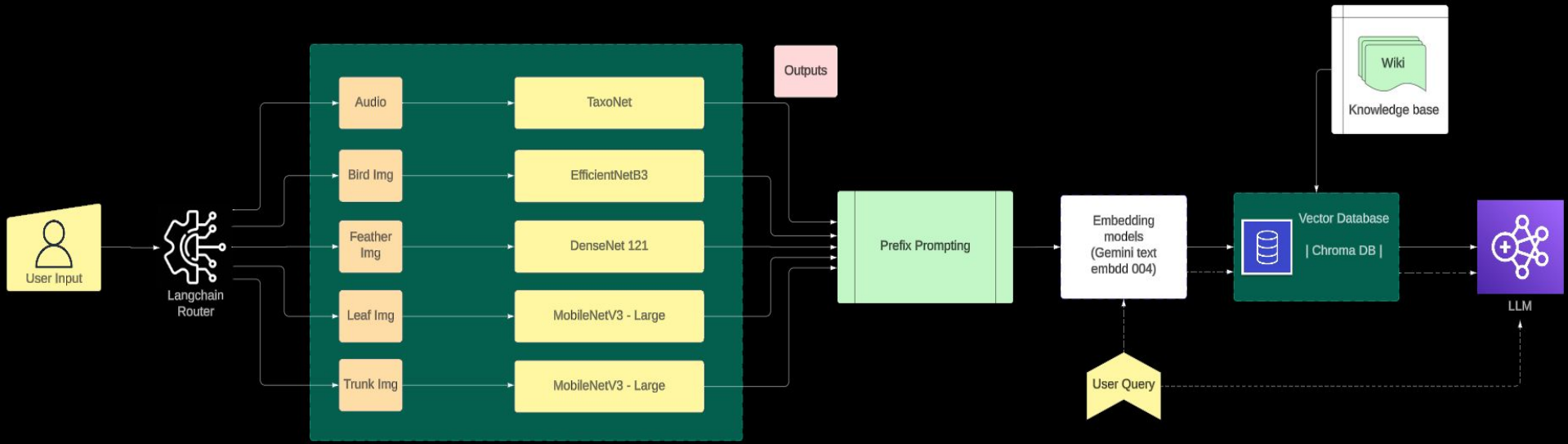
Customer Needs Assessment

- Conducted feedback sessions with fellow birdwatching enthusiasts in our batch.
- Key desired features for an AI-assisted bird identification platform:
  - Species detection from both bird images and audio recordings.
  - Ability to identify bird species using bird feathers found on the ground to help determine the bird population in specific areas.
  - Optional feature to identify trees based on leaves or trunk, enhancing knowledge of the birds' habitat.

DEMO

Repo Github link

Inference Pipeline



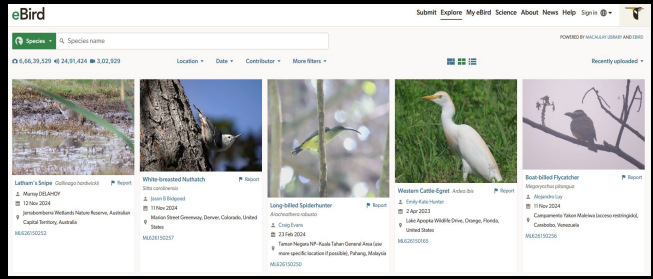
Data Collection

01 BirdCLEF 2024 competition dataset (aud)

02 eBird library by Cornell lab of Ornithology

03 25 Indian Bird Species (img)

04 Feather V1

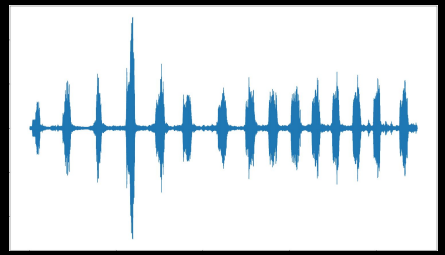


Audio pipeline

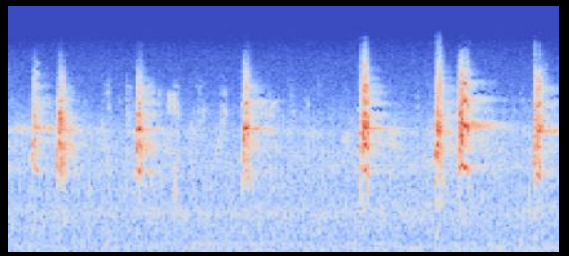
- Input to ML model ?
- Data Augmentation
- Species Identification
- Type of Call Identification

Mel Spectrogram

- '***Mel Spectrogram***' is a 2D representation (freq vs time) of audio signals in the frequency domain using '**mel scale**' to emphasize frequencies that are more perceptible to human hearing.
- It's created by applying ***short-time Fourier transform (STFT)*** on an audio signal and mapping the frequency components to the Mel scale.



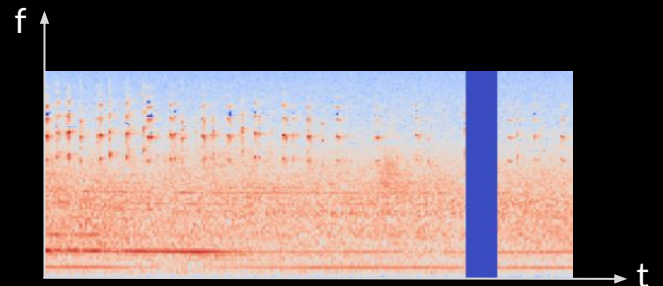
Raw audio



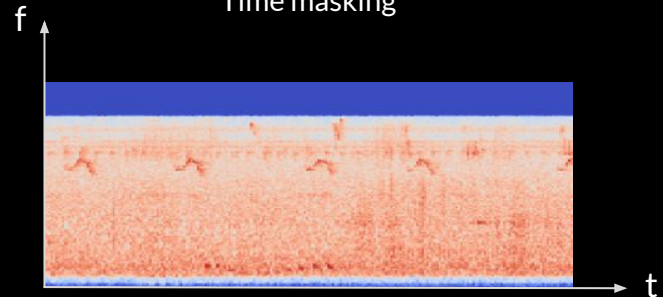
Mel spectrogram

Data Augmentation

- Time stretch
  - Alters speed of audio without changing the pitch
- Time masking
  - Randomly masks a short time segment in the audio
- Frequency masking
  - Randomly masks a range of frequencies



Time masking



Frequency masking

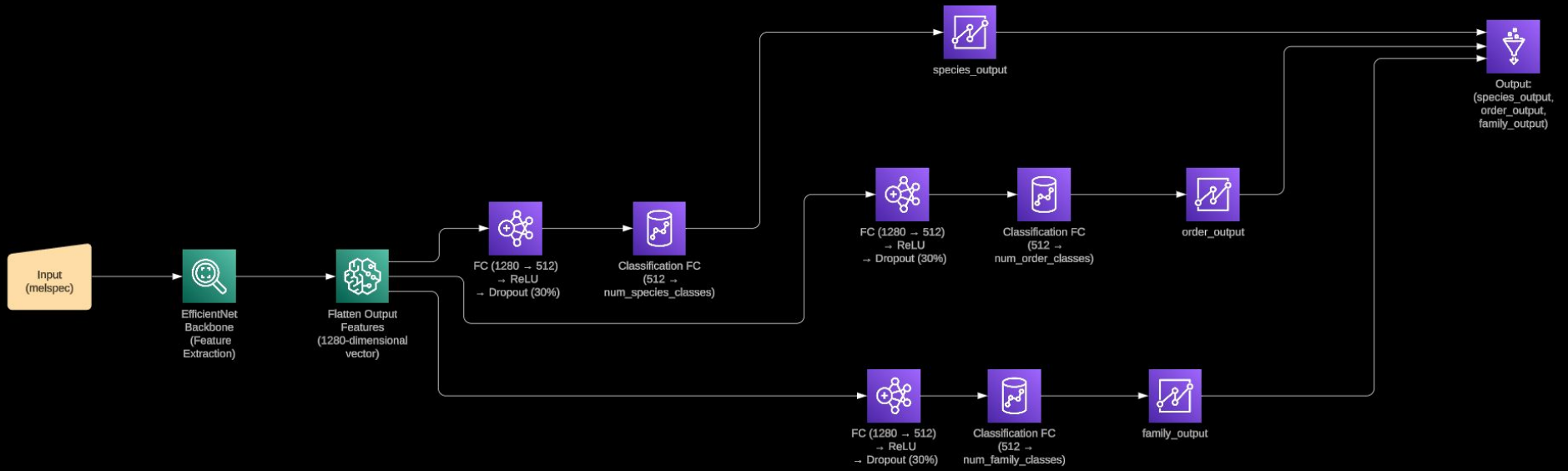
Species Identification

➢ Simple model . . .

- Using ***EfficientNet_v2_s*** as the backbone for feature extraction, with gradual unfreezing of layers implemented to prevent '***Catastrophic Unlearning***'.

Multitask Learning

- We have labels of the taxonomy (order, family, species) in our dataset
- As the 3 tasks are related, why not predict all the 3 labels using separate fully connected layers at the end for each task . . . ?

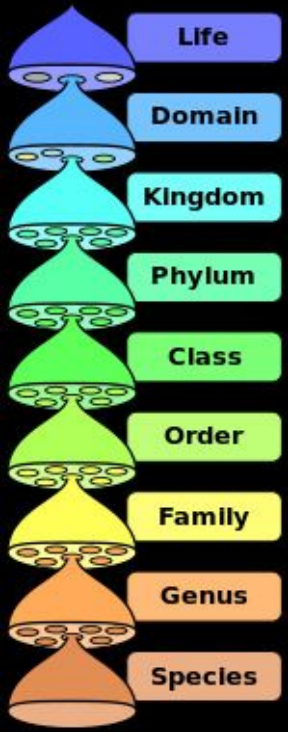


Conflicting Gradients . . . Under Optimization . . .

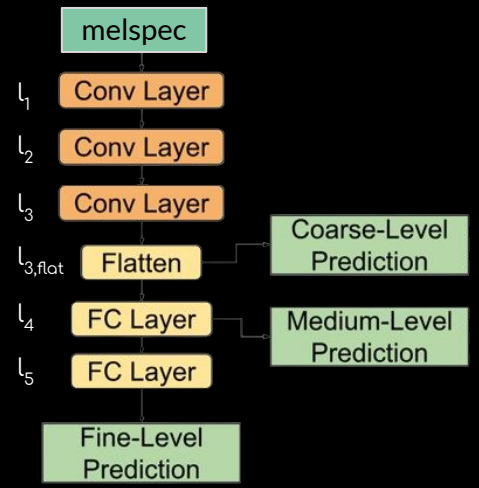
- We use ***Fast Adaptive Multitask Optimization (FAMO)***
- It adjusts the importance (by adjusting weights) of each task during training depending on how well each task's loss is progressing.

Hierarchy . . .

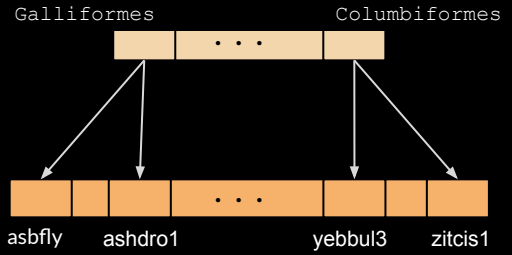
- The hierarchy is the *order, family and species* taxa
- Our taxonomy:
  - Coarse level - Order
  - Medium level - Family
  - Fine level - Species



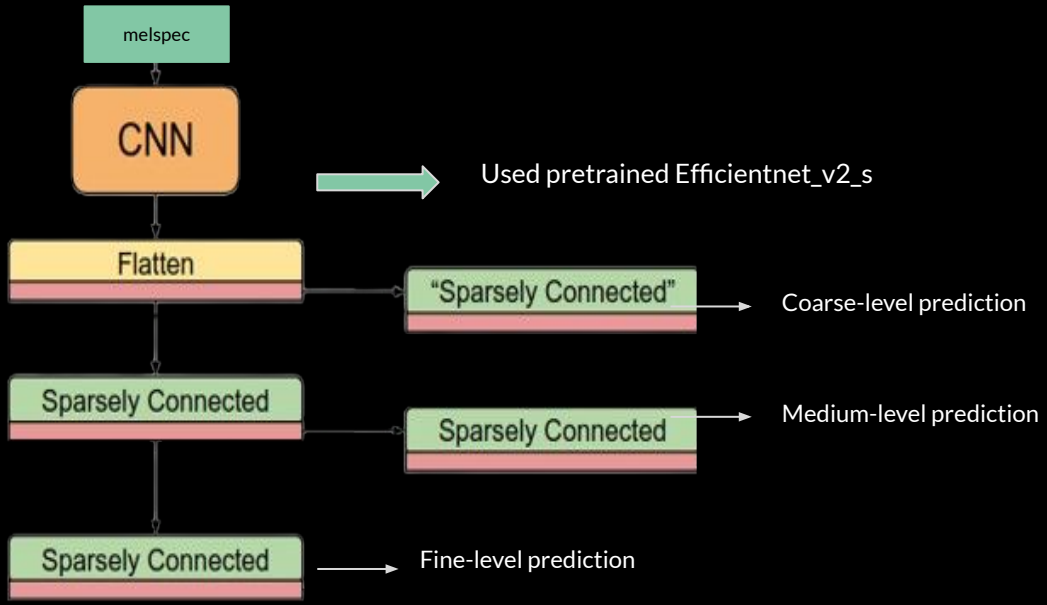
➢ TaxoNet - *Hierarchical Multitask learning*



- In single-task models, only l_5 contains output nodes. But in this Hierarchical Multitask learning model, $l_{3,flat}$ is associated with coarse-level prediction(order), l_4 to the medium-level(family) and l_5 for the fine-level prediction(species).
- Nodes at each layer are partitioned, with each partition mapping to a member of the current taxonomic level
- Partitions for a taxa member in the current layer only connect to partitions corresponding to descendents in the next taxonomic level, mirroring the structure of the taxonomy



In our case . . .


melspec
CNN
Used pretrained Efficientnet_v2_s
Flatten
"Sparsely Connected"
Coarse-level prediction
Sparsely Connected
Sparsely Connected
Medium-level prediction
Sparsely Connected
Fine-level prediction

Comparison

Model	**Validation accuracies** *(species)*
ConvModel	68%
MTL	70%
TaxoNet	73%

Type of sound identification

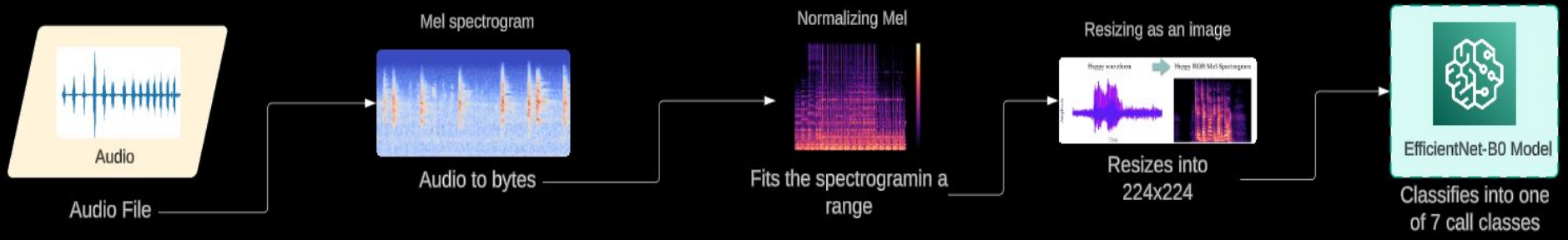
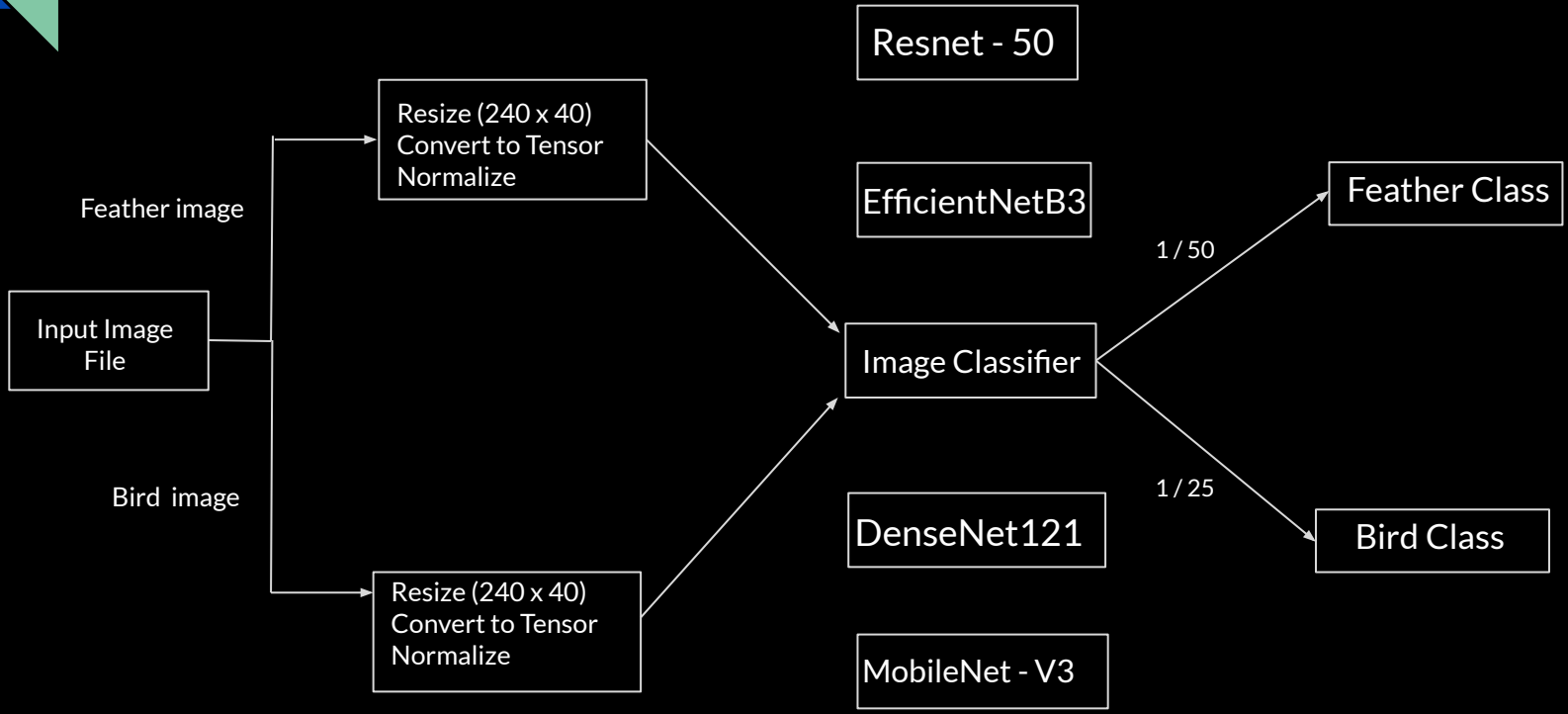
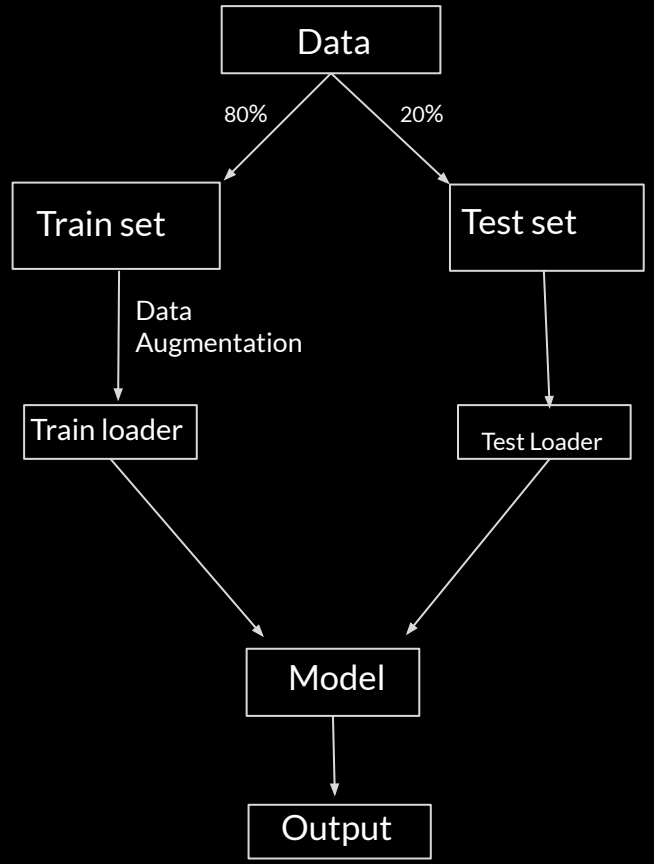


Image Pipeline :- Part 1

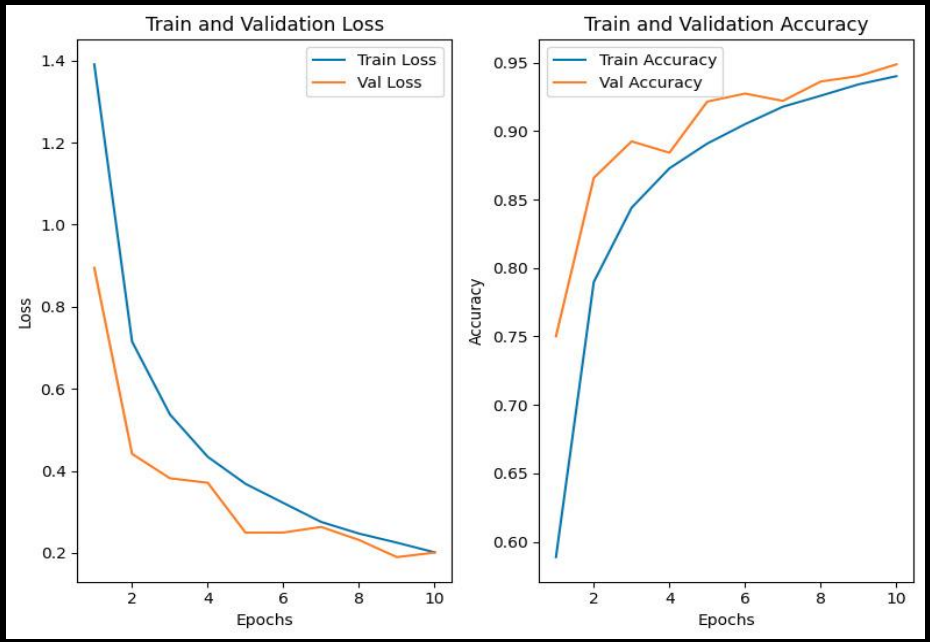


Training

Data Augmentations
`Random Horizontal Flip`
`Random Rotation upto 10*`
`Random Resized Crop`
`Color Jitter`



Accuracies and Plots

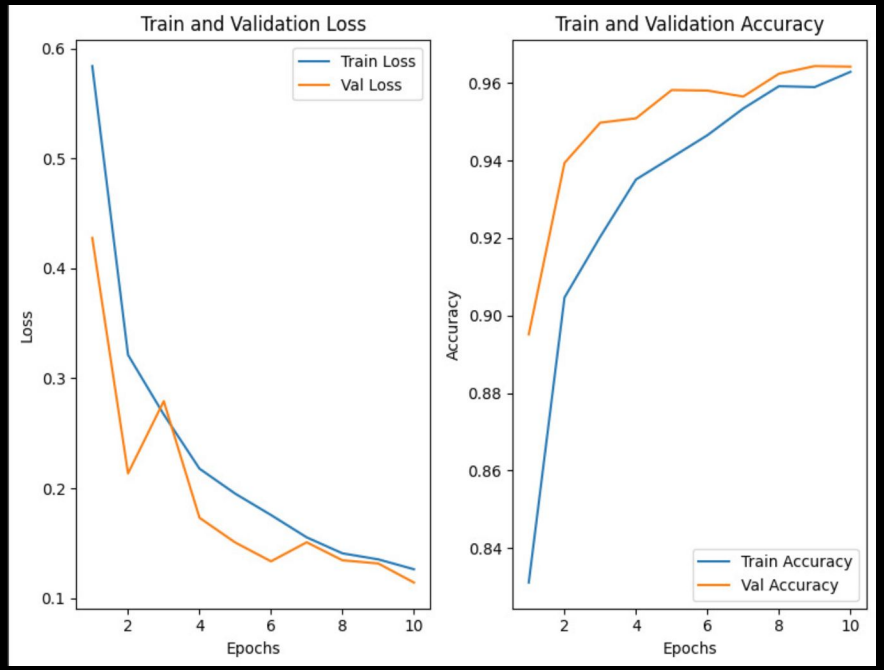


EfficientNetB3 on 25 Indian Bird Species

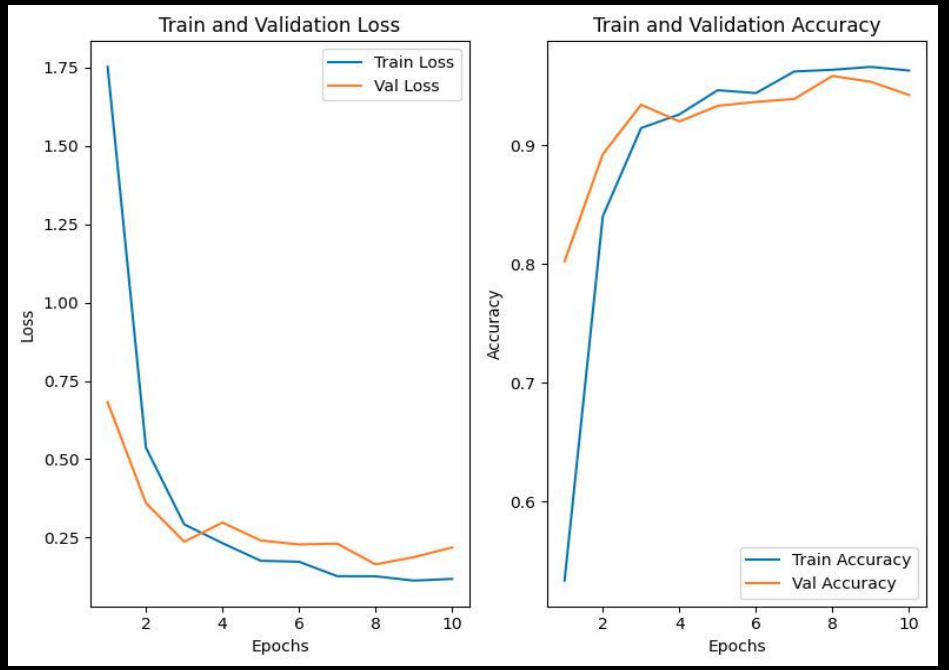
Model	Accuracy on Test-Set	Hyper-Parameters
EfficientNetB3	96.4 95.88(FeatherV1)	Batch Size = 8 Lr = 0.001 Epochs = 10
DenseNet 121	94.54	Batch Size = 8 Lr = 0.001 Epochs = 10
Resnet 50	87	Batch Size = 32 Lr = 0.001 Epochs = 15
MobileNet V3	83	Batch Size = 32 Lr = 0.001 Epochs = 10

* Accuracies are for Indian 25 bird Species Dataset, unless mentioned otherwise .

More plots



DenseNet121 on 25 Indian Bird Species



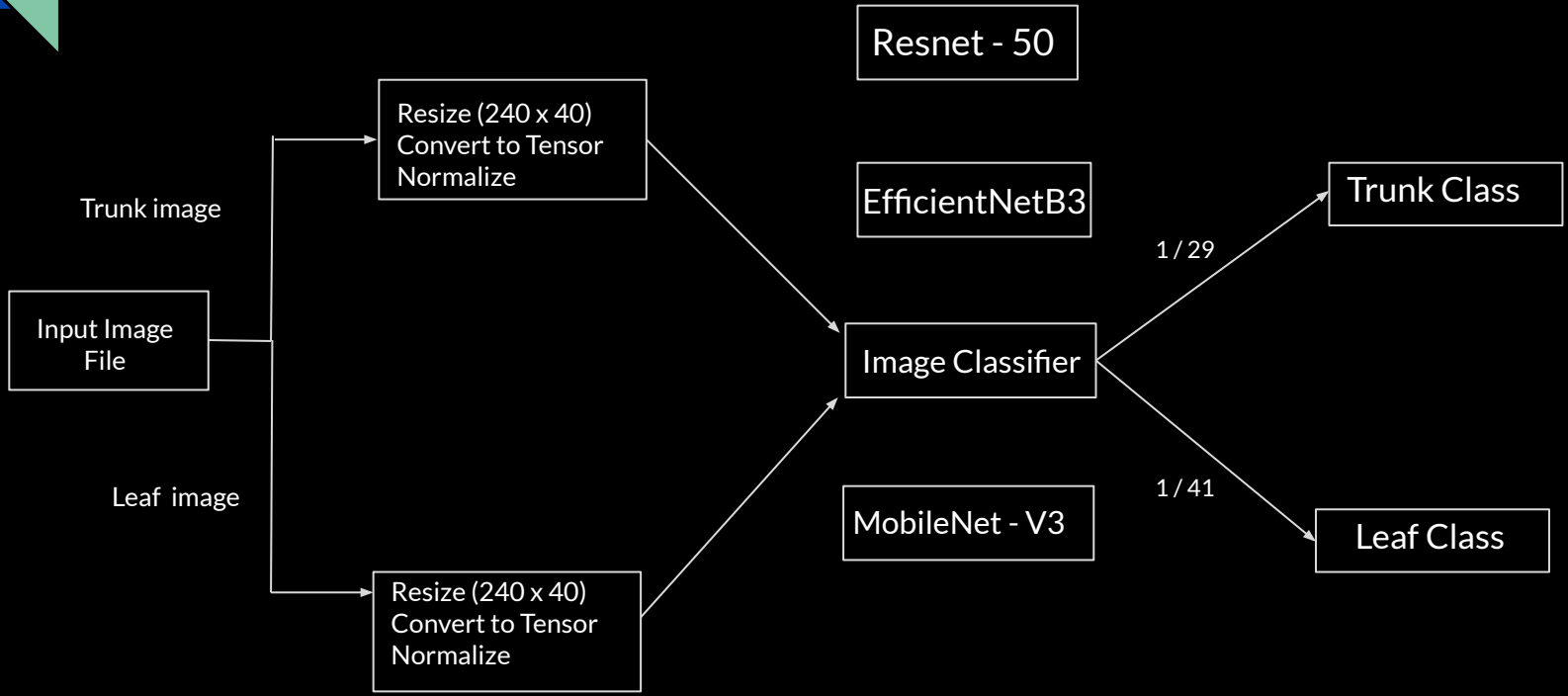
EfficientNetB3 on FeatherV1

Vegetation

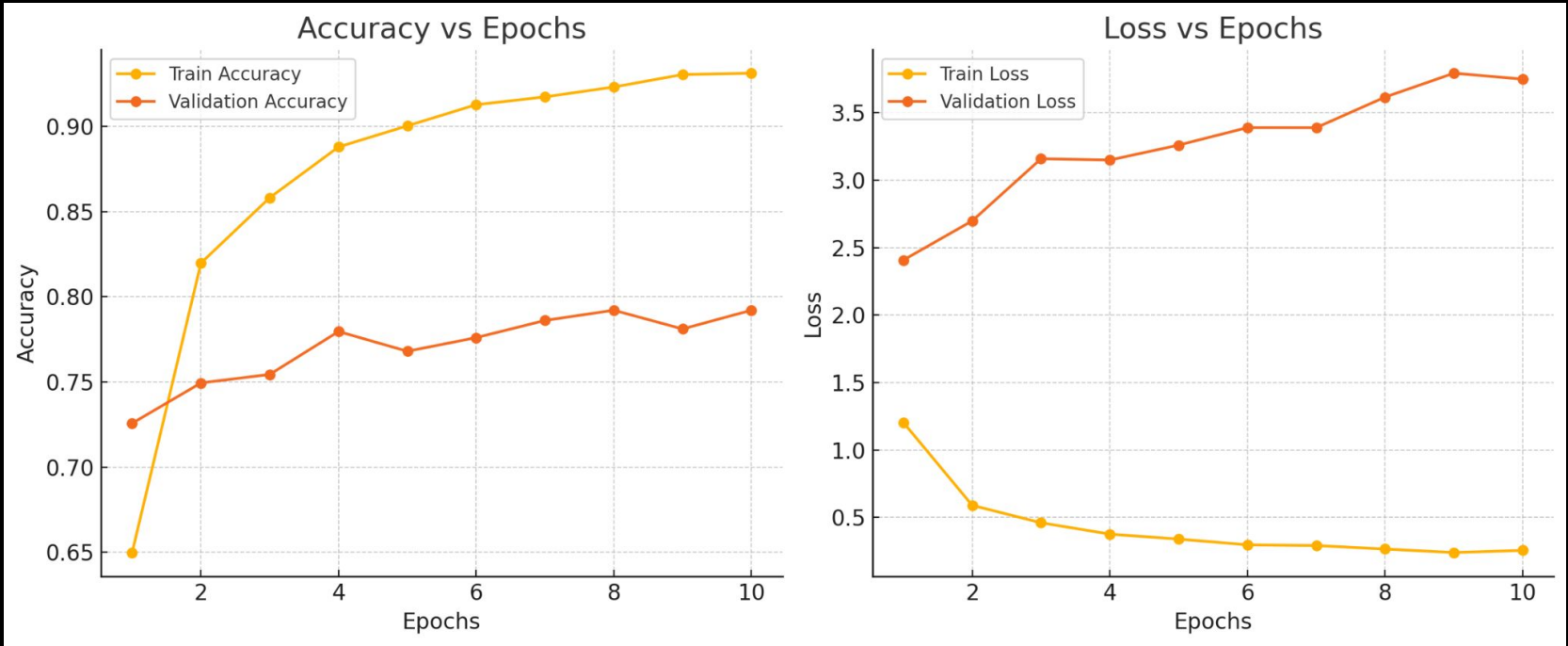
- Species of tree identified from the input image of leaves/trunk using a MobileNet-v3 model.
- We use the RAG model to query information like:
  - Family, Genus, Climate preference, Native Region and Associated birds
- Knowing the Genus helps us identify the type of forest (Deciduous/Coniferous/etc) the tree might be surrounded by.
- Depending upon this we can figure out the average climate the region endures
- Many birds prefer only specific trees for nesting (eg: Woodpecker prefer Aspen wood or similar softwood trees).

Birdwatchers actively look for cues like these in their surroundings!

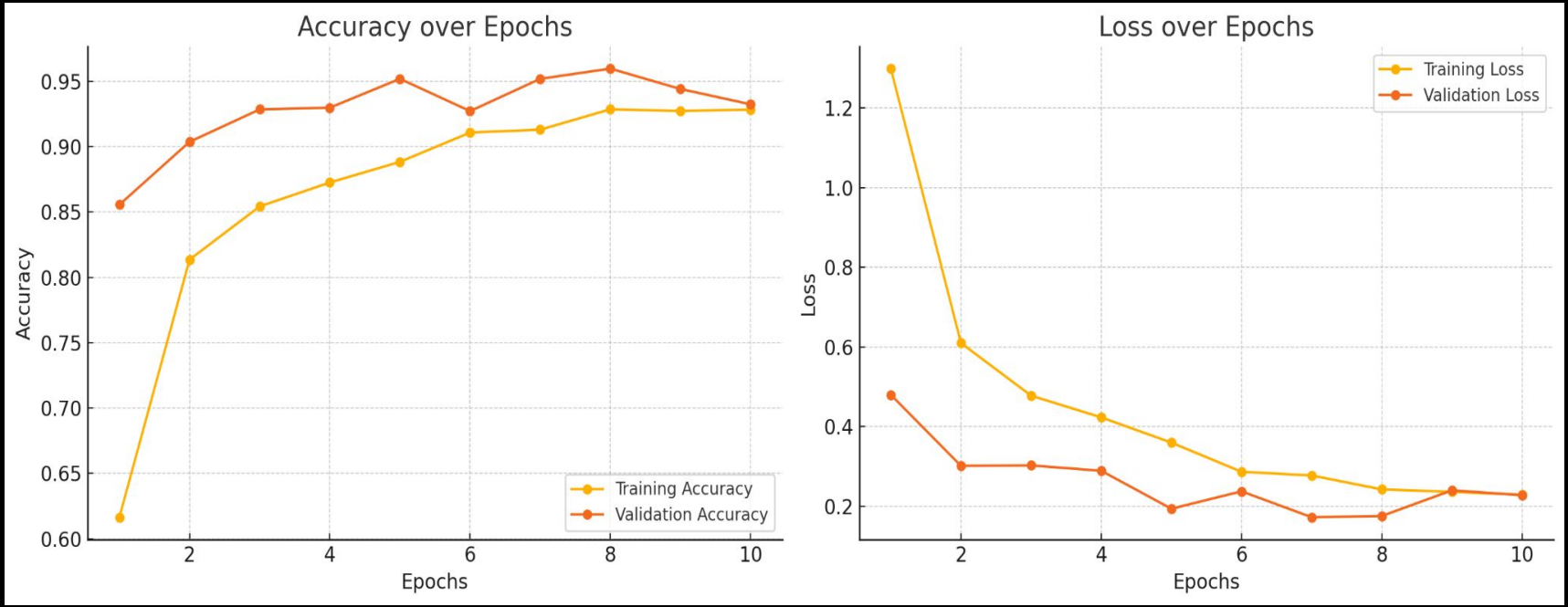
Image Pipeline :- Part 2



Accuracy plots:
MobileNetv3 on Leaf dataset



MobileNetv3 on Trunk dataset



Ideations for the Future

- We will create a more robust login system
- We will make location pinning of birds more user friendly
- We will create a community of all the people who use the app
- We will encourage users to enhance online databases through our app
- We will create a dataset on bird excreta to further develop ml models
- In future we will use a more standard tech stack

Contributions

Member	Contribution
Nagasai	Backend, Website-Logic, LLM, RAG
Sathvik	Species identifier with audio, Flowcharts
Shankar	Type of sound identifier, Website-UI, User-auth
Krishna	Dataset handling, Web Scraping
Aditya	Bird species identifier with bird image/ feather image
Sanyat	Tree identification with trunk/leaf images, BBox detection(RCNN)