BISCUTE

Breed Identification & Specific Care Utility Tool Engine

Member:

1. Ferdi Fadillah
2. Meghana Gudamsetty
3. Raquel Brown
4. Vincent G. Capone

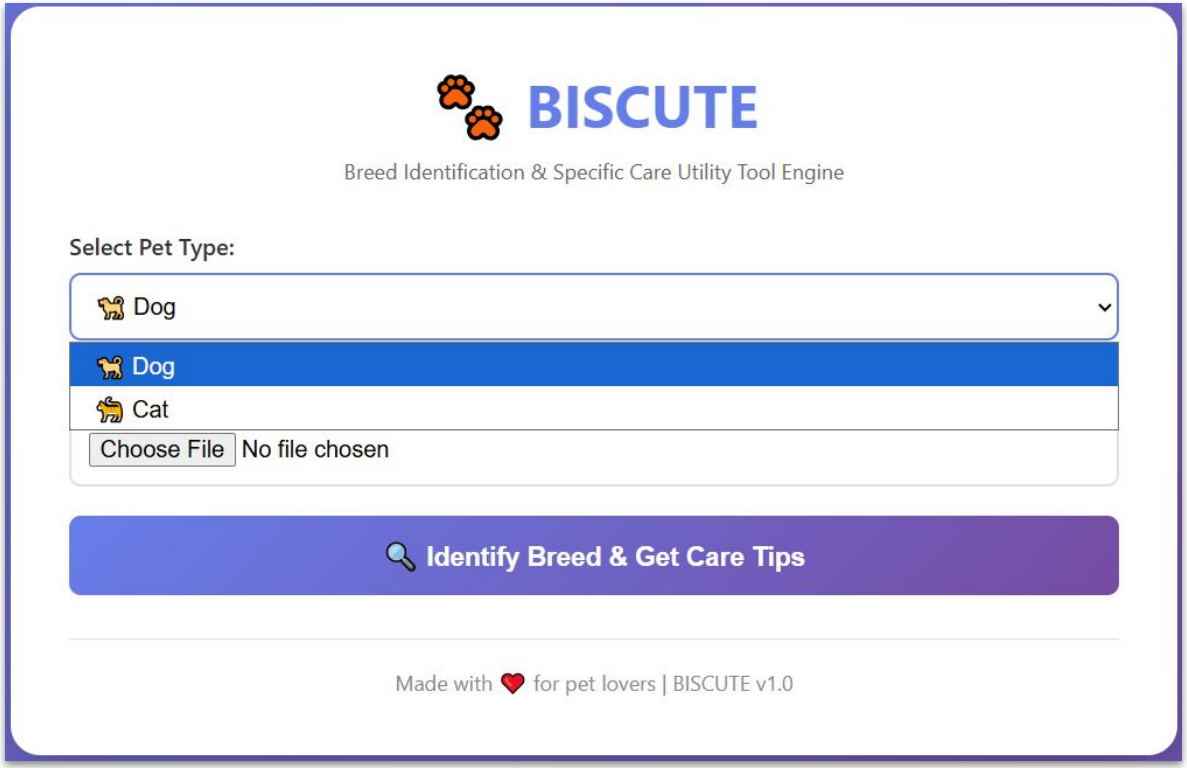
Background

We want to make and engine that can give recommendation care to our pet (dog and cat), but we need the classifier. Nowadays, a good classifiers is already available, especially the dog one. However, the cat one does not cover some niche cat breeds.

Hence, we build the website for this case, BISCUTE (Breed Identification & Specific Care Utility Tool Engine). We still use the dog classifier model from internet, but for the cat, we do fine-tuning by ourself.

Info:

- BISCUTE stands for Breed Identification & Specific Care Utility Tool Engine
- You can first select any image of your choice of a dog or cat.
- This will allow you to identify the breed and provide some helpful tips depending on each breed.

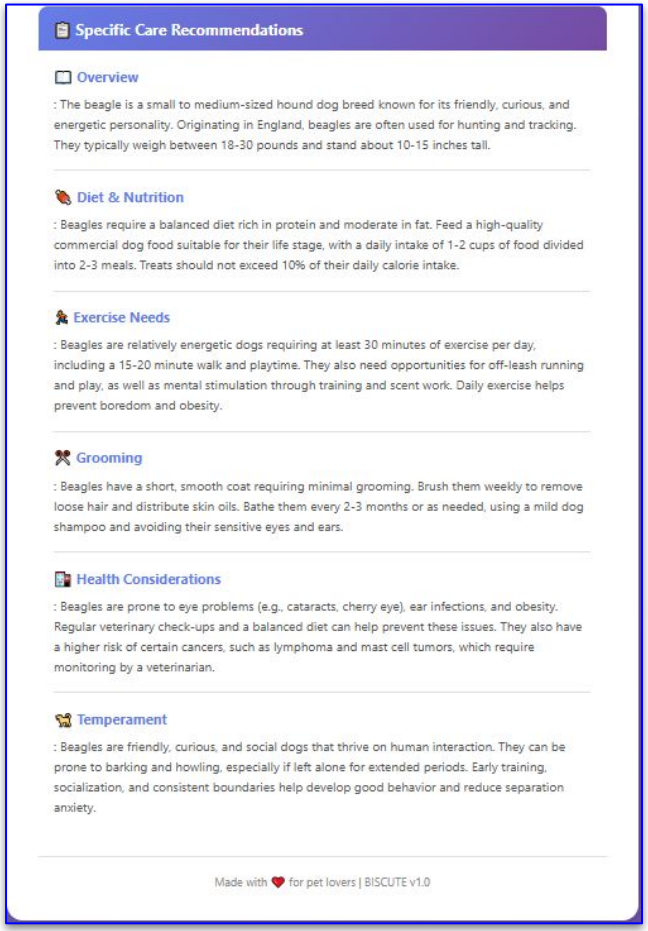


Breed Specific Specification

Beagle

Confidence: 18.47%

Helpful Care Tips



Specific Care Recommendations

Overview

: The beagle is a small to medium-sized hound dog breed known for its friendly, curious, and energetic personality. Originating in England, beagles are often used for hunting and tracking. They typically weigh between 18-30 pounds and stand about 10-15 inches tall.

Diet & Nutrition

: Beagles require a balanced diet rich in protein and moderate in fat. Feed a high-quality commercial dog food suitable for their life stage, with a daily intake of 1-2 cups of food divided into 2-3 meals. Treats should not exceed 10% of their daily calorie intake.

Exercise Needs

: Beagles are relatively energetic dogs requiring at least 30 minutes of exercise per day, including a 15-20 minute walk and playtime. They also need opportunities for off-leash running and play, as well as mental stimulation through training and scent work. Daily exercise helps prevent boredom and obesity.

Grooming

: Beagles have a short, smooth coat requiring minimal grooming. Brush them weekly to remove loose hair and distribute skin oils. Bathe them every 2-3 months or as needed, using a mild dog shampoo and avoiding their sensitive eyes and ears.

Health Considerations

: Beagles are prone to eye problems (e.g., cataracts, cherry eye), ear infections, and obesity. Regular veterinary check-ups and a balanced diet can help prevent these issues. They also have a higher risk of certain cancers, such as lymphoma and mast cell tumors, which require monitoring by a veterinarian.

Temperament

: Beagles are friendly, curious, and social dogs that thrive on human interaction. They can be prone to barking and howling, especially if left alone for extended periods. Early training, socialization, and consistent boundaries help develop good behavior and reduce separation anxiety.

Made with ❤ for pet lovers | BISCUTE v1.0

Machine Learning Process

1. Loading Dataset and Pretraining Model
2. Image Pre-Processing and Background Removal
3. Data Loading & Sampler
4. Fine-tuning Cat Breed Model with New Breeds

1. Loading Dataset and Pretraining Model

- Imports and organizes the cat breed image dataset from storage:
  - The original breed dataset (for rehearsal strategy)
    - https://www.kaggle.com/datasets/ma7555/cat-breeds-dataset
    - https://www.kaggle.com/datasets/doctrinek/catbreedsrefined-7k
  - The new dataset, we only take 12 new breeds, and delete the duplicates
    - https://www.kaggle.com/datasets/almanaqibmahmooddar/37-cats-breeds-dataset
- Loads a pre-trained model with learned features

New Breeds

```
NEW_BREEDS = [
    "Chartreux",
    "Chausie",
    "Khao Manee",
    "Kurilian Bobtail",
    "LaPerm",
    "Neva Masquerade",
    "Ocicat",
    "Peterbald",
    "Savannah",
    "Selkirk Rex",
    "Serengeti",
    "Singapura",
]
```

```
label     size
Savannah           128
Ocicat             98
Singapura          95
Neva Masquerade    92
LaPerm             91
Peterbald          90
Chartreux          86
Chausie            80
Selkirk Rex        79
Serengeti          78
Khao Manee         69
Kurilian Bobtail   50
Name: count, dtype: int64
```

2. Image Pre-Processing and Background Removal

- Image processing steps:
  a. Background removal with gray background (handle multi-color of cat)
  b. Cropping (to make cat is full in the picture with small padding)
  c. Resize to 224x224 and center it

3. Data Loading & Sampler

- Do undersampling and oversampling data to 5645
- Load new breed cat data (for training and evaluation)
- Load original breed cat data (for rehearsal strategy and evaluation)

4. Fine-tuning Cat Breed Model with New Breeds

Now that our base model understands the common cat features, we can efficiently teach it to recognize new breeds. This can be done by:

- Expanding it to 12 new breeds (from 48 breeds to 60 breeds)
- Train test split (60% training, 40% test)
- The Rehearsal Strategy ratio is of 70% (to prevent catastrophic forgetting)
- Freeze the base model and unfreeze the classifier head (for fine-tuning)
- Augmentation for training
- Adjusting the class weight to handle imbalanced dataset (needed for rehearsal strategy)
- Train (3 epochs) and Perform Evaluation on the model

As a result, this makes our model more efficient and produces better accuracy.

Results for Accuracy, Training and Validation Loss

[3240/3240 1:36:51, Epoch 3/3]

Epoch	Training Loss	Validation Loss	Accuracy
1	1.115900	0.976417	0.723502
2	0.960700	0.908906	0.748280
3	0.966600	0.899622	0.749897

New Breeds Accuracy Outcomes

```
============================================================
PER-BREED ACCURACY (New Breeds Only)
============================================================
Chartreux          :  94.51% (2134/2258)
Chausie            :  83.08% (1876/2258)
Khao Manee         :  85.12% (1922/2258)
Kurilian Bobtail   :  86.05% (1943/2258)
LaPerm             :  64.57% (1458/2258)
Neva Masquerade    :  91.90% (2075/2258)
Ocicat             :  82.95% (1873/2258)
Peterbald          :  76.97% (1738/2258)
Savannah           :  77.41% (1748/2258)
Selkirk Rex        :  88.09% (1989/2258)
Serengeti          :  64.13% (1448/2258)
Singapura          :  78.17% (1765/2258)
============================================================
Average accuracy (new breeds): 81.08%
Min accuracy: 64.13%
Max accuracy: 94.51%
```

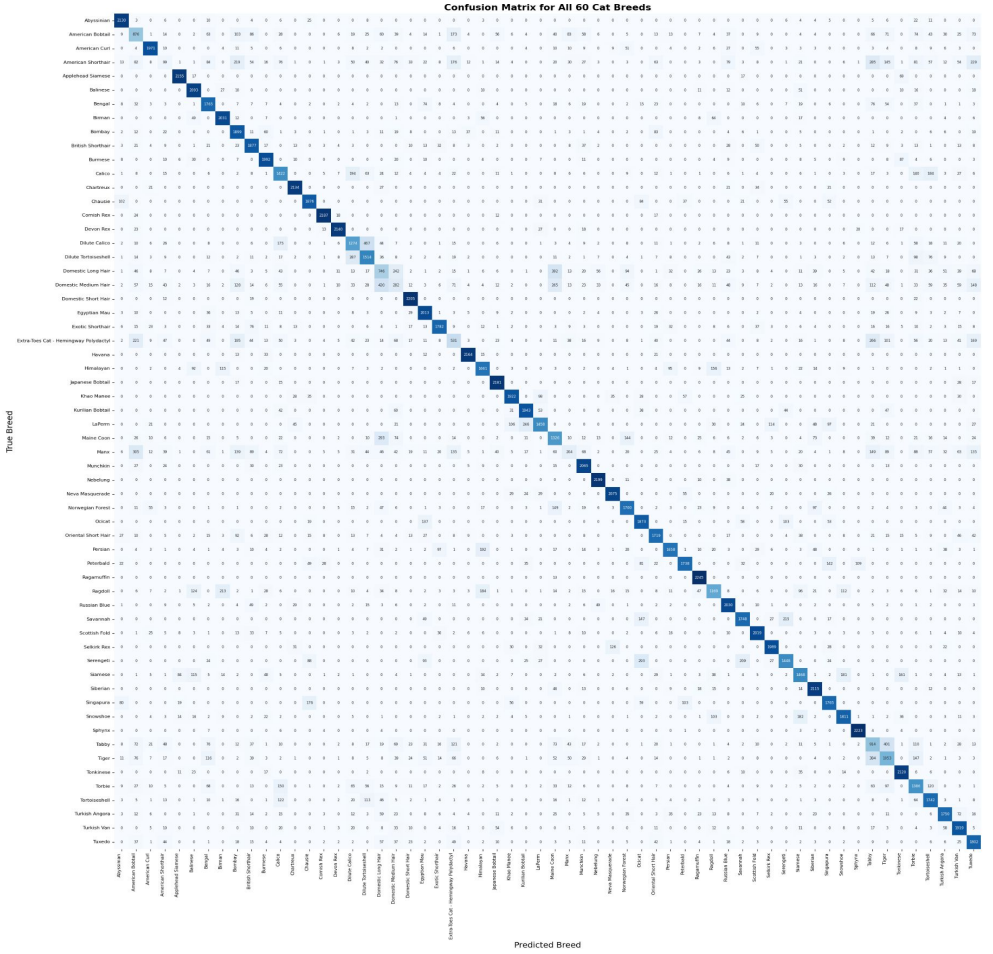
Original Breeds

Abyssinian : 94.33% (2130/2258)
American Bobtail : 38.80% (876/2258)
American Curl : 87.29% (1971/2258)
American Shorthair : 4.38% (99/2258)
Applehead Siamese : 95.44% (2155/2258)
Balinese : 92.69% (2093/2258)
Bengal : 78.17% (1765/2258)
Birman : 89.95% (2031/2258)
Bombay : 84.10% (1899/2258)
British Shorthair : 83.13% (1877/2258)
Burmese : 88.22% (1992/2258)
Calico : 62.98% (1422/2258)
Cornish Rex : 96.86% (2187/2258)
Devon Rex : 94.77% (2140/2258)
Dilute Calico : 56.42% (1274/2258)
Dilute Tortoiseshell : 67.05% (1514/2258)

Domestic Long Hair : 33.04% (746/2258)
Domestic Medium Hair : 12.49% (282/2258)
Domestic Short Hair : 97.65% (2205/2258)
Egyptian Mau : 89.15% (2013/2258)
Exotic Shorthair : 78.92% (1782/2258)
Extra-Toes Cat - Hemingway Polydactyl: 23.52% (531/2258)
Havana : 95.84% (2164/2258)
Himalayan : 73.56% (1661/2258)
Japanese Bobtail : 96.59% (2181/2258)
Maine Coon : 58.72% (1326/2258)
Manx : 9.03% (204/2258)
Munchkin : 91.45% (2065/2258)
Nebelung : 97.39% (2199/2258)
Norwegian Forest : 75.29% (1700/2258)
Oriental Short Hair : 76.13% (1719/2258)
Persian : 73.43% (1658/2258)

Ragamuffin : 99.42% (2245/2258)
Ragdoll : 51.77% (1169/2258)
Russian Blue : 89.90% (2030/2258)
Scottish Fold : 89.42% (2019/2258)
Siamese : 65.01% (1468/2258)
Siberian : 93.67% (2115/2258)
Snowshoe : 80.20% (1811/2258)
Sphynx : 98.45% (2223/2258)
Tabby : 40.48% (914/2258)
Tiger : 46.63% (1053/2258)
Tonkinese : 93.89% (2120/2258)
Torbie : 61.38% (1386/2258)
Tortoiseshell : 77.15% (1742/2258)
Turkish Angora : 77.50% (1750/2258)
Turkish Van : 84.99% (1919/2258)
Tuxedo : 79.81% (1802/2258)

Confusion Matrix



BREEDS WITH POOR PERFORMANCE (< 20% accuracy)

American Shorthair : 4.38% [ORIGINAL] (99/2258)
Domestic Medium Hair : 12.49% [ORIGINAL] (282/2258)
Manx : 9.03% [ORIGINAL] (204/2258)

Total: 3 breeds (3 original, 0 new)

CONFUSION ANALYSIS FOR BREEDS With < 20% ACCURACY

```
American Shorthair (ORIGINAL) - Top-1 Accuracy: 4.38%
Most common TOP-1 predictions (what model thinks it is):
  Tabby                          [ORIGINAL]:  285/2258 ( 12.6%) ✗ WRONG
  Tuxedo                         [ORIGINAL]:  229/2258 ( 10.1%) ✗ WRONG
  Bombay                         [ORIGINAL]:  219/2258 (  9.7%) ✗ WRONG

Domestic Medium Hair (ORIGINAL) - Top-1 Accuracy: 12.49%
Most common TOP-1 predictions (what model thinks it is):
  Domestic Long Hair             [ORIGINAL]:  420/2258 ( 18.6%) ✗ WRONG
  Domestic Medium Hair           [ORIGINAL]:  282/2258 ( 12.5%) ✓ CORRECT
  Maine Coon                     [ORIGINAL]:  265/2258 ( 11.7%) ✗ WRONG

Manx (ORIGINAL) - Top-1 Accuracy: 9.03%
Most common TOP-1 predictions (what model thinks it is):
  American Bobtail               [ORIGINAL]:  305/2258 ( 13.5%) ✗ WRONG
  Manx                           [ORIGINAL]:  204/2258 (  9.0%) ✓ CORRECT
  Tabby                          [ORIGINAL]:  149/2258 (  6.6%) ✗ WRONG
```

Original's and New Breeds' Statistics

==========
SUMMARY STATISTICS
==========

ORIGINAL BREEDS (48 breeds):
  Average accuracy: **73.47%**
  Min accuracy: 4.38%
  Max accuracy: 99.42%

NEW BREEDS (12 breeds):
  Average accuracy: **81.08%**
  Min accuracy: 64.13%
  Max accuracy: 94.51%

OVERALL (All 60 breeds):
  Average accuracy: **74.99%**
  Min accuracy: 4.38%
  Max accuracy: 99.42%

In Relation to Our Class

There has been one lab from this class that was useful to this project:

- unnit10_cnn/Lab_weather_partial
- Helped us with fine tuning our model for the classification of the dog and cat breeds

Sources:

Dataset:

https://www.kaggle.com/datasets/ma7555/cat-breeds-dataset

https://www.kaggle.com/datasets/doctrinek/catbreedsrefined-7k

https://www.kaggle.com/datasets/almanaqibmahmooddar/37-cats-breeds-dataset

Pretrained-Model:

https://huggingface.co/dima806/cat_breed_image_detection

https://huggingface.co/wesleyacheng/dog-breeds-multiclass-image-classification-with-vit

Notebook:

https://www.kaggle.com/code/dima806/cat-breed-image-detection-vit


THANK
YOU!