AP157 Capstone Project

Jed Jerrel K. Escaran
BS3 Applied Physics

May 21, 2024

Racial profiling is the practice of targeting individuals for suspicion of crime based on their race, ethnicity, religion, or national origin. [a]

- Reasons Why Racial Profiling is Bad: [a]
  1. Violation of Human Rights
  2. Ineffectiveness in Law Enforcement
  3. Erosion of Trust
  4. Psychological and Social Harm
  5. Legal and Ethical Concerns

- Instances of Wrongful Incrimination:
  1. Traffic Stops and Searches
  2. Stop-and-Frisk Practices [b]
  3. Post-9/11 Policies [c]
  4. Cases of Wrongful Arrest [d]

[a] https://www.aclu.org/issues/racial-justice/race-and-criminal-justice/racial-profiling
[b] https://www.nyclu.org/data/stop-and-frisk-data
[c] https://www.hrw.org/reports/2002/usahate/usa1102-04.htm
[d] https://www.npr.org/sections/thetwo-way/2018/05/03/607973546

This study aims to spread awareness and educate that no matter the situation, one must never judge another based on their color.

“never judge a book by its cover”

-Aristotle (probably)


Wild ???
appeared!

???
???/???

Scyther
Bug/Flying

Togekiss
Fairy/Flying

Mamoswine
Ice/Ground

Gotta catch 'em all!

Type profiling Pokémons based on their color using image processing and machine learning

Jed Jerrel K. Escaran
BS3 Applied Physics
AP157 Capstone Project

May 21, 2024

Dataset

- Datasets were retrieved from Kaggle.com
- Dataset by Vishal Subbiah was used in this study [1]

	Name	Type1	Type2	Evolution
0	bulbasaur	Grass	Poison	ivysaur
1	ivysaur	Grass	Poison	venusaur
2	venusaur	Grass	Poison	NaN
3	charmander	Fire	NaN	charmeleon
4	charmeleon	Fire	NaN	charizard
...	...	...	...	...
804	stakataka	Rock	Steel	NaN
805	blacephalon	Fire	Ghost	NaN
806	zeraora	Electric	NaN	NaN
807	meltan	Steel	NaN	NaN
808	melmetal	Steel	NaN	NaN

809 rows x 4 columns

Figure 1. Pokémon Dataset

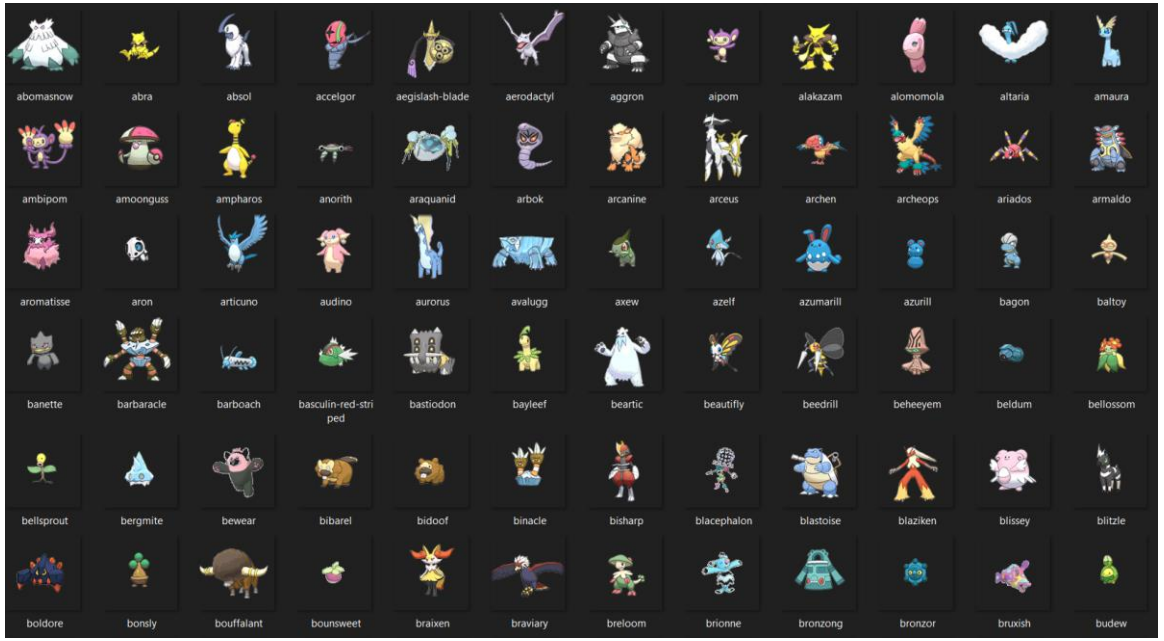


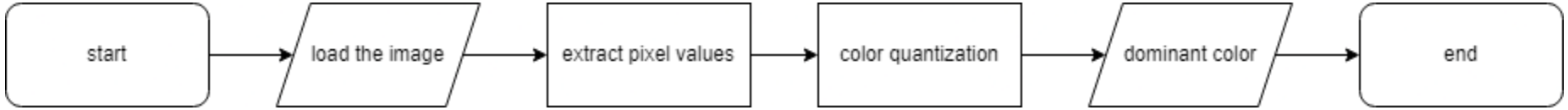
Figure 2. Pokémon Images

[1] https://www.kaggle.com/datasets/vishalsubbiah/pokemon-images-and-types

Image processing

- Color Thief [2]
  - a Python library that extracts the dominant color or a palette of colors from an image
  - utilizes the k-means clustering algorithm to analyze colors and identify the most prominent colors

Figure 3. Color Thief process flowchart



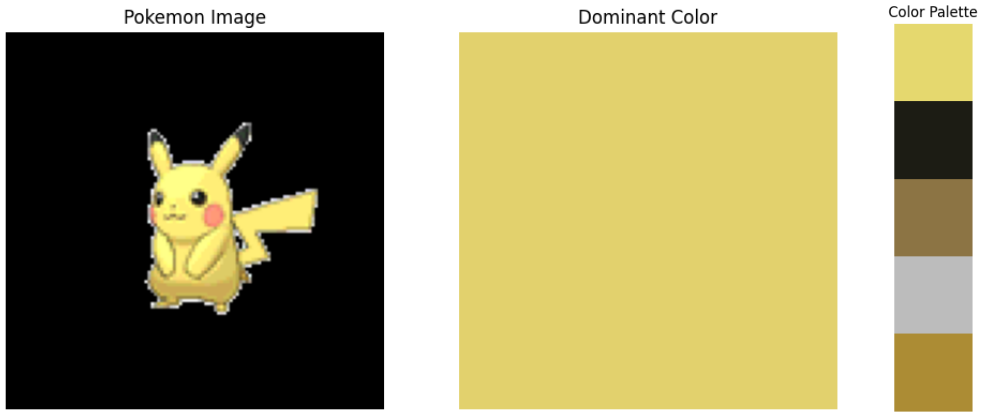


Figure 4. Pikachu dominant color and palette

Figure 5. Eevee dominant color and palette

[2] https://github.com/fengsp/color-thief-py

Machine learning

- Scikit-learn [3]
  - Model selection: train_test_split was used to split dataset (names, types, colors) into train and test subsets
  - Model training: random forest classifier was used to train the model

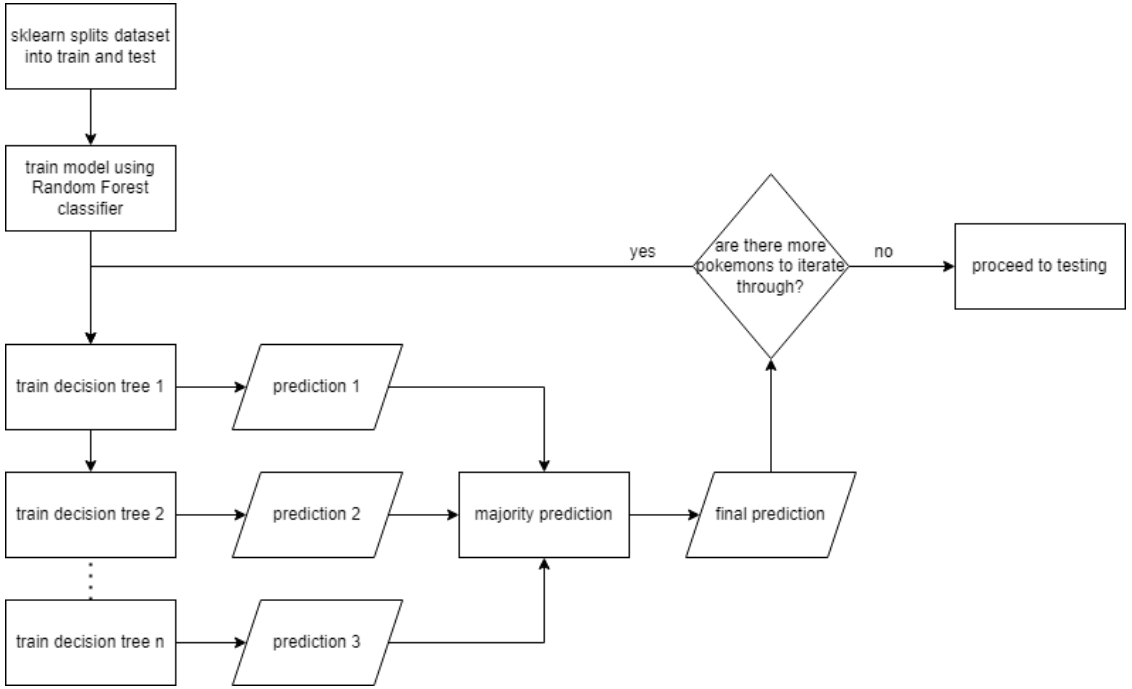


Figure 6. Model training flowchart using Random Forest classifier

[3] https://scikit-learn.org/stable/

Machine learning

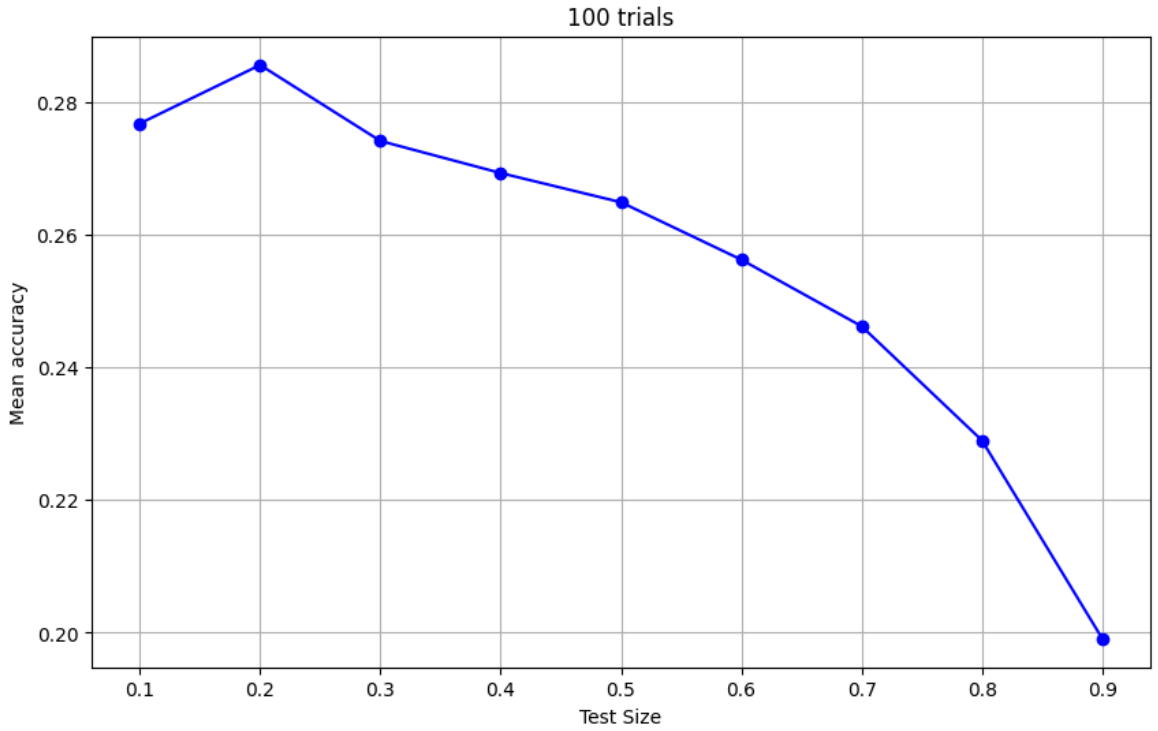


Figure 7. Mean accuracy of varying test sizes over 100 trials

[3] https://scikit-learn.org/stable/

Results

- First test: at least one predicted type matches

```
swinub
predicted: Normal
real type: Ice and Ground
Failed

cyndaquil
predicted: Fire
real type: Fire
SUCCESSFUL PREDICTIONS = 13

wartortle
predicted: Psychic
real type: Water
Failed

lombre
predicted: Grass
real type: Water and Grass
SUCCESSFUL PREDICTIONS = 14

haunter
predicted: Poison
real type: Ghost and Poison
SUCCESSFUL PREDICTIONS = 15
```

```
Accuracy: 26.9%
```

Figure 8. First test sample results and accuracy

- Second Test: predicted type must match primary type

```
huntail
predicted: Rock
real type: Water
FAILED

pelipper
predicted: Water
real type: Water and Flying
SUCCESS

heatmor
predicted: Fire
real type: Fire
SUCCESS

lotad
predicted: Grass
real type: Water and Grass
FAILED

marill
predicted: Normal
real type: Water and Fairy
FAILED
```

```
Accuracy: 25.2%
```

Figure 9. First test sample results and accuracy

Results

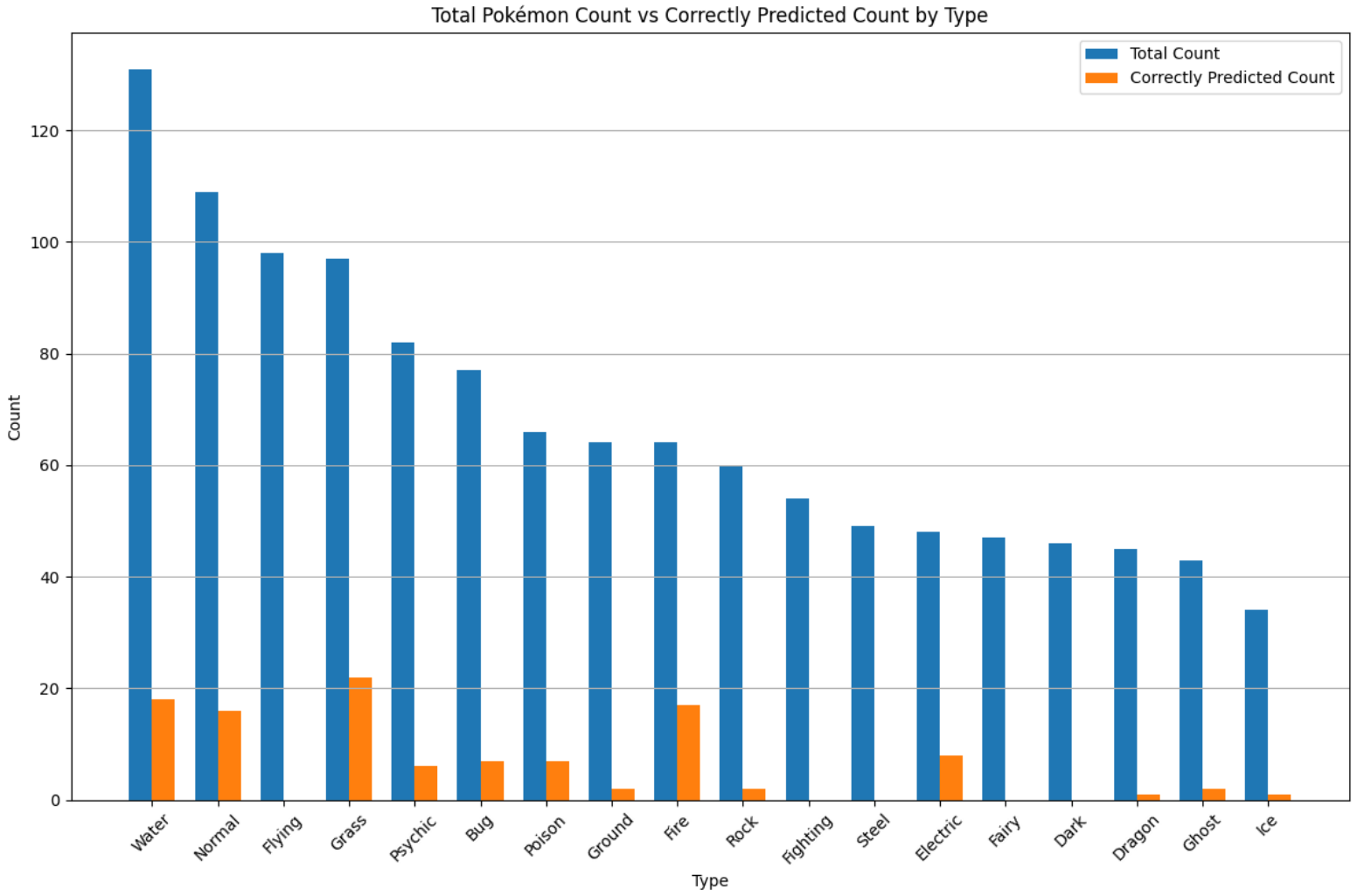


Figure 10. Total Pokémon count with the number of correctly predicted count for each type

Results

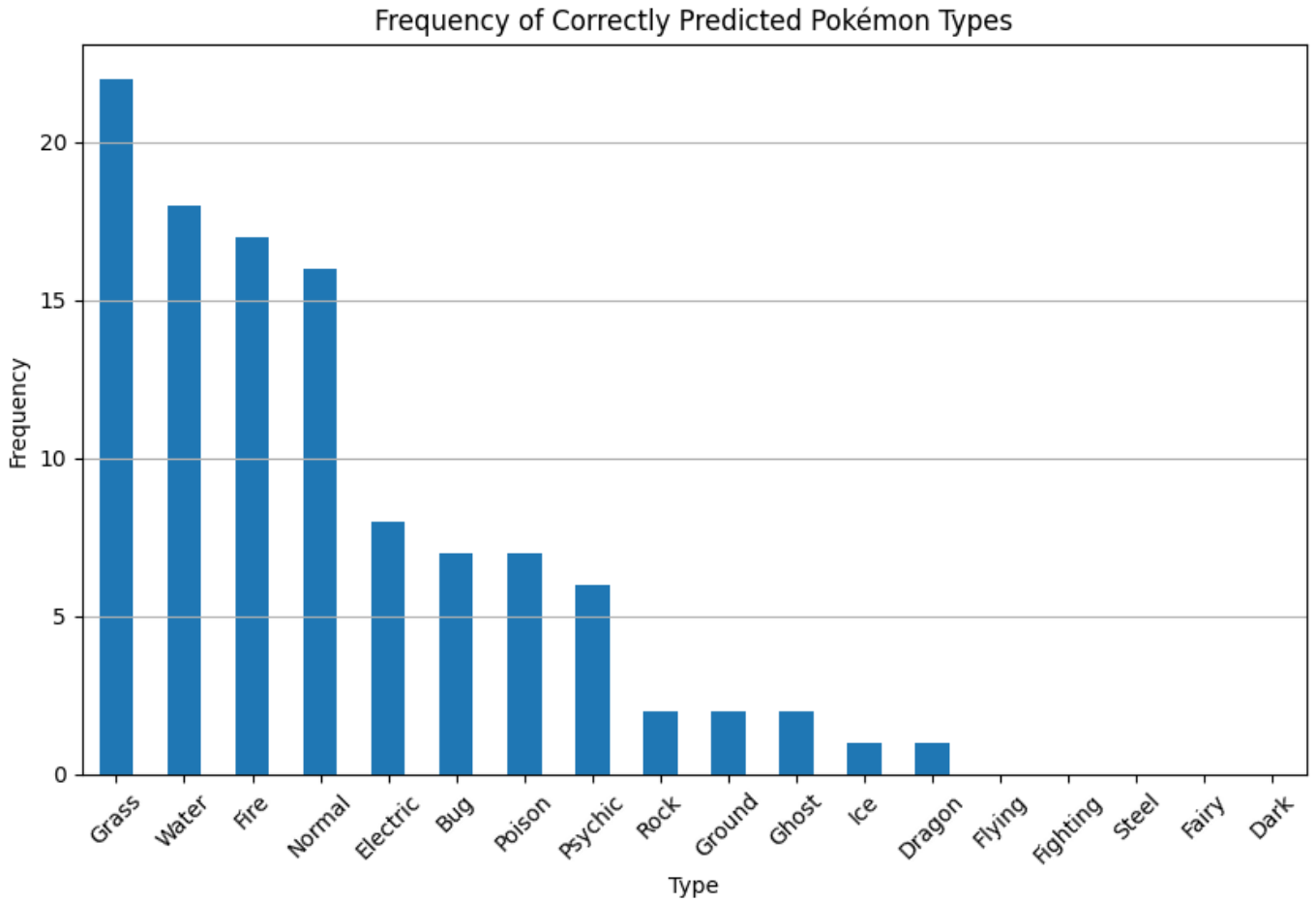


Figure 11. Frequency of successful predictions by type

Figure 12. Grass, Water, Fire, and Normal type sample Pokémons and their dominant color

Insights and Limitations

- Pokémon aesthetics

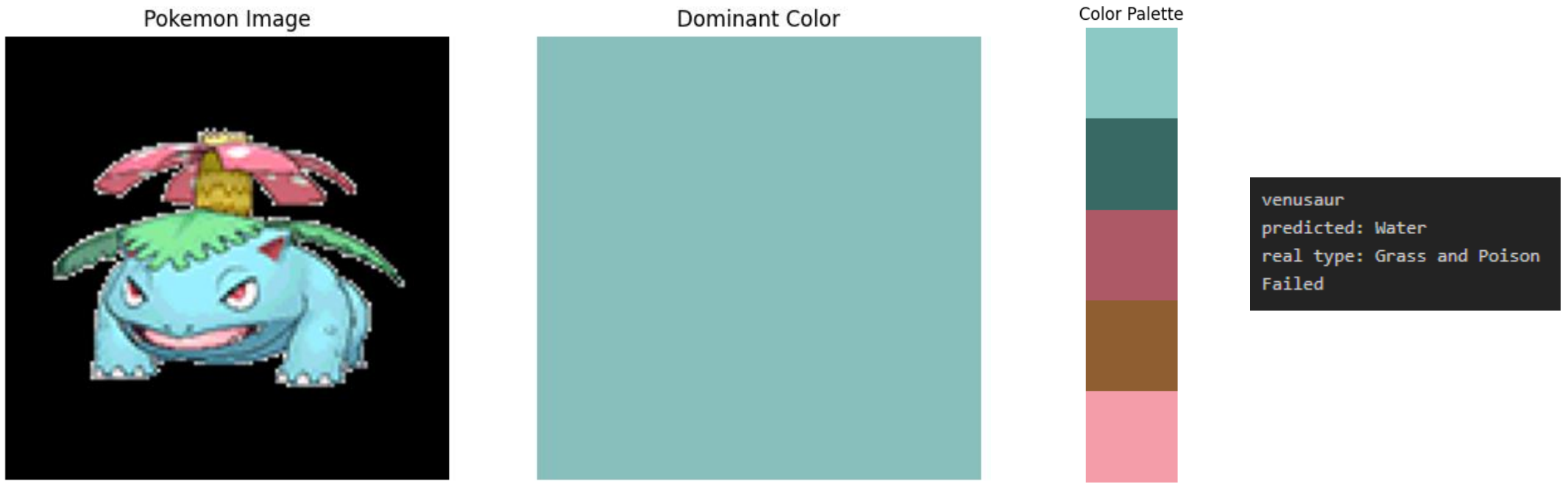


Figure 13. Venusaur, a Grass/Poison type Pokémon, with its dominant color, palette and model-predicted type

Insights and Limitations

- Lack of representative color

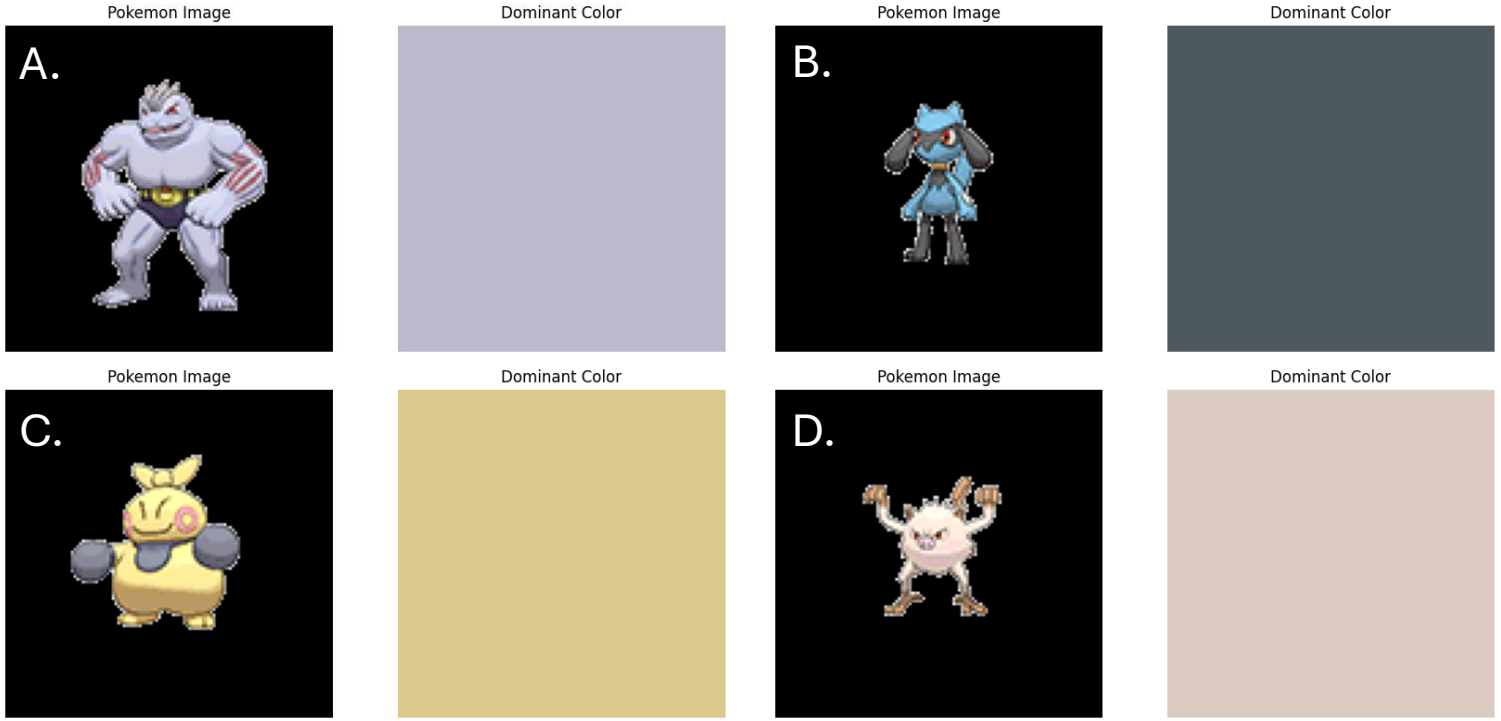


Figure 14. Four pure fighting-type pokemons with varying dominant colors.
A = Machoke. B = Riolu. C = Makuhita. D = Mankey

???
???/???

Scyther
Bug/Flying

Togekiss
Fairy/Flying

Mamoswine
Ice/Ground

???
DARK???/???

Scyther
Bug/Flying

Togekiss
Fairy/Flying

Mamoswine
Ice/Ground

???
DARK???/???

Scyther
Bug/Flying

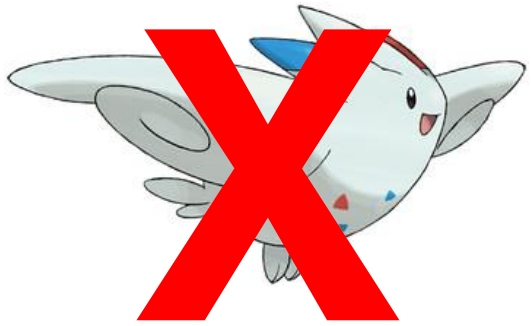
Togekiss
Fairy/Flying

Mamoswine
Ice/Ground

Eelektross
ELECTRIC

Scyther
Bug/**Flying**

Togekiss
Fairy/**Flying**



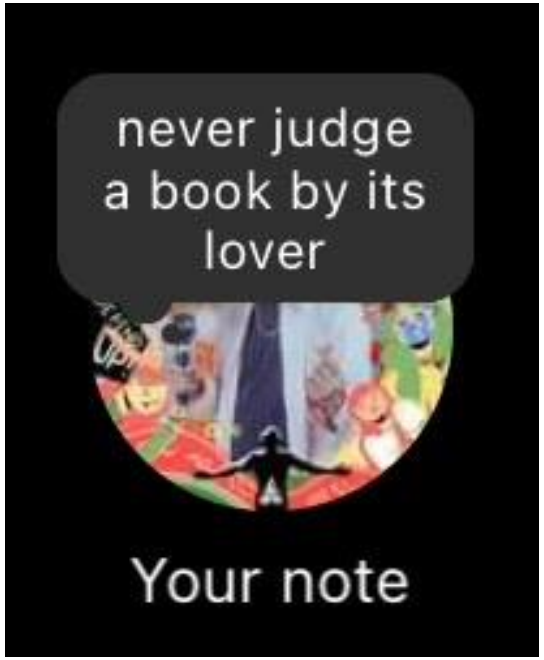
Mamoswine
Ice/**Ground**

Conclusion

- Color of Pokémon is **not** a good indicator for type, with an accuracy of only 26.9% and 25.2% on the first and second test, respectively.

- Usage of CNN deep learning algorithm might be able to achieve higher accuracy.
- Using CNN, double type prediction may be possible.

Conclusion



References

Dataset:

Subbiah, V. (March 2024). Pokemon Image Dataset. Retrieved from https://www.kaggle.com/datasets/vishalsubbiah/pokemon-images-and-types on May 16, 2024

Other images and GIFs:

Pokémon Database (n.d.). Retrieved from https://pokemondb.net/ on May 16, 2024