Identify Potential Habitat for the South China Tigers

Yuchen Liu, Haojun Zhang, Lirong Wu

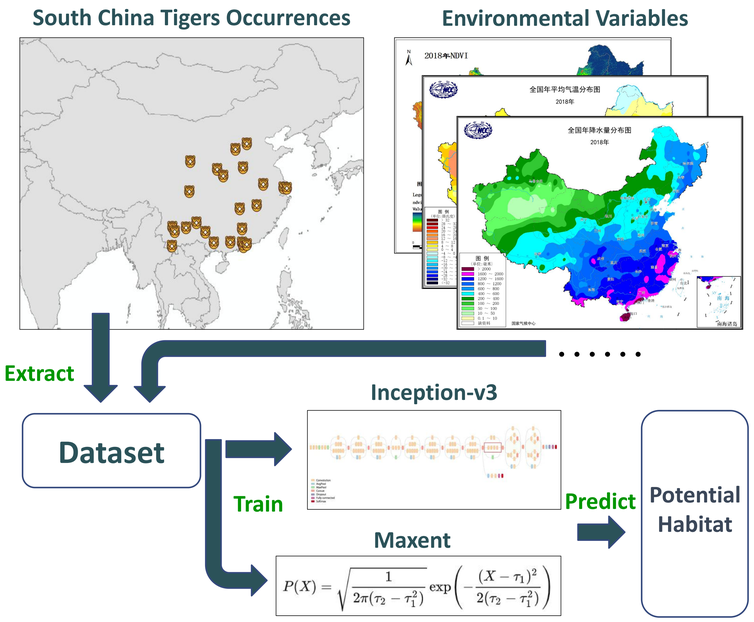
SIGS, Tsinghua University, China

Abstract

- **Construct a presence-only dataset for the tigers based on the consideration of 20 environmental raster variables.**
- **Implement and improve the Inception-v3 model to specifically predict the categorical probability of the South China tigers.**
- **Combine CNN with the Maxent model to form a potential habitat suitability map.**

Introduction

- **Motivation**: Since South China tigers have come to be the most endangered tiger subspecies, identifying the potential habitats is important for their reintroduction.

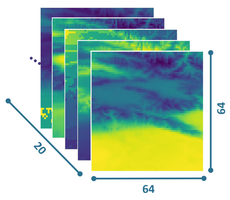


Problems to be solved:

- **Model Transfer**: How to implement prediction by CNN?
- **Accuracy Improvement**: Which parts can be optimized?

Dataset

- Obtain 68 occurence points of tiger from GBIF and INAT,including 36 items of SC tiger.

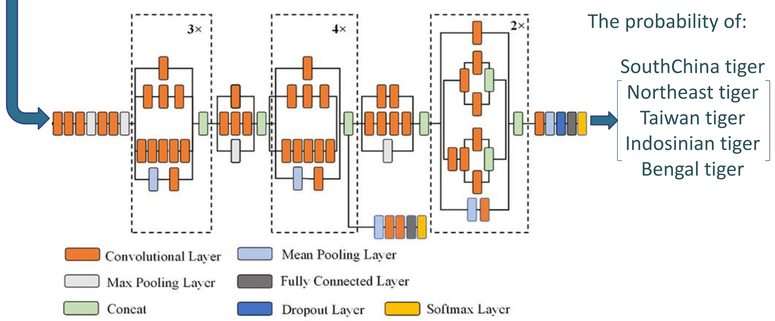


- Collect 20 environmental factors as variable
- To train CNN combine 20 environment variables, each with 64×64 pixels, to form a 3D tensor (64×64×20) for each occurrence point.

Fig1: example of environmental tensor

Method

INCEPTION-V3:



- Use inception-v3 model,whose output takes the form of a categorical probability over all possible tiger species conditionally to the fact that a tiger has been observed at a given location.

MAXENT:

$$\max_{p} = \sum_{x,y} \overline{P}(x)P(y \mid x)\log P(y \mid x)$$

$$s.t\sum_{x,y} \overline{P}(x,y)f(x,y) = \sum_{x,y} \overline{P}(x)P(y \mid x)f(x,y)$$

$$s.t\sum_{y} P(y \mid x) = 1$$

- The probability of tiger in a certain location was obtained by using the Lagrange multiplier to solve the above optimization problems.
- The product of the results of Maxent and inception-v3 model is the probability of South China tiger at a position.

Result

Occurrence probability of the tigers - MAXENT:

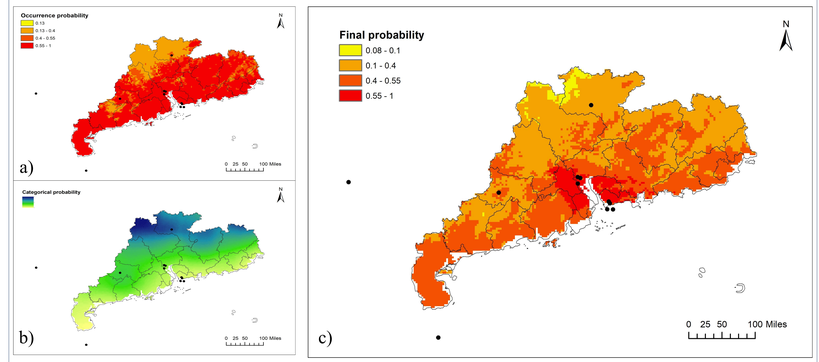
- Maxent model for the tigers performs better than random. The average test AUC for the 5 replicate runs is 0.885, and the standard deviation is 0.042.

categorical probability of the south China tigers - CNN:

- high Top-k accuracy (>0.85) of the model and by species of the model on the test set

Habitat Suitability (HS) map:

- We take Guangdong Province as an example, and plot the Habitat suitability map in the whole field.
- a) shows the occurrence probability of the tigers, b) shows the categorical probability of the SC tigers, and c) indicates the potential habitat suitability of the SC tigers.
- The high suitable area is mainly distributed in Shenzhen, Dongguan, Foshan, Zhongshan and southern Guangzhou.



Conclusion

- This work combined CNN with Maxent to predict the potential suitable habitat for a rare species in China, south China tiger, so as to guide the reintroduction of this species. The application of machine learning on environmental science is still a topic worth researching on.

References

Deneu, B. et al. (2021) 'Convolutional neural networks improve species distribution modelling by capturing the spatial structure of the environment', PLOS Computational Biology, 17(4), p. e1008856. Available at: https://doi.org/10.1371/journal.pcbi.1008856.