Heat Risk Prediction in Philadelphia

MUSA 6950
Jingmiao Fei

Why Heat Risk?

Extreme heat is one of the deadliest weather-related hazards. With ongoing climate change and urbanization, cities are becoming 'heat islands'.

Predicting heat risk accurately is critical for early warning systems, public health planning, and building resilient cities.

Research Goal

Use machine learning methods to build a **heat risk level model** based on the satellite images and LULC index to provide scientific decision-making assistance for urban planning.

Data

Landsat 8 / C02 / T1_L2

- 2020 ~ 2024, June ~ August
- Cloud cover < 20%
- Median

Philadelphia Administrative Boundary

Data

Y

- Band 10 (Thermal Infrared, 100m resolution) → **Land Surface Temperature (LST)** → Resample to 30m

X

- Band 4,3,2 (RGB), Band 5 (NIR),
- NDVI, NDBI, MNDWI, SAVI

Model

Categorize LST to heat risk

- ≤ 35℃ low
- 35 ~45℃ Medium
- ≥ 45℃ High

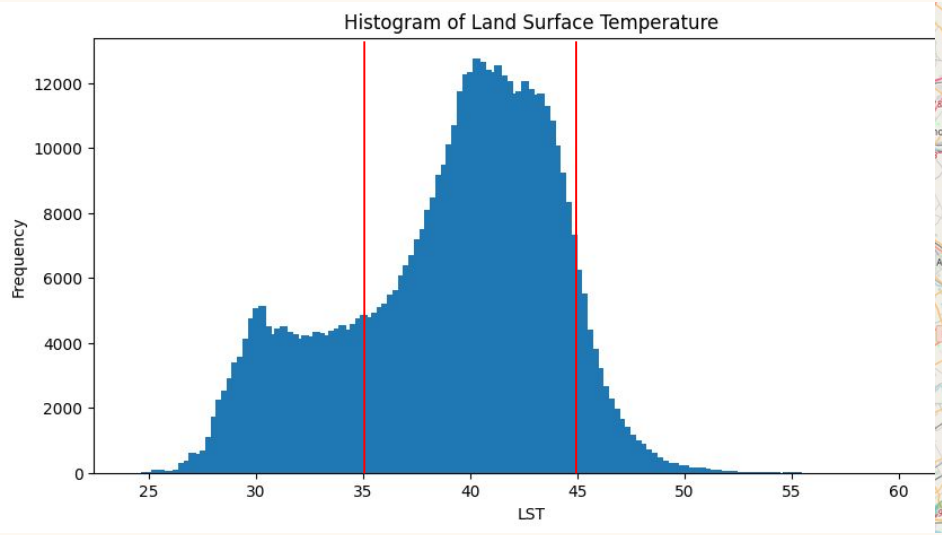
Stratified sampling

- 5000 samples in total

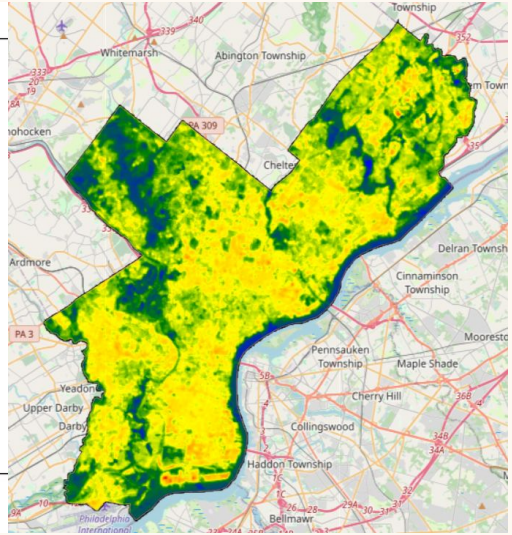
train (70%) **- test** (30%) **split**

Model

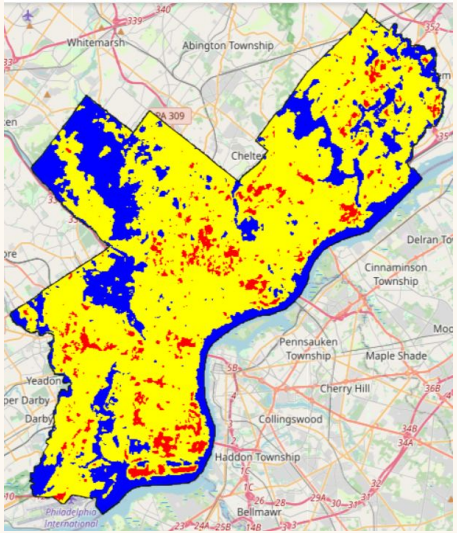
Histogram of LST



LST



Heat Risk



Model

Model 1 - CNN (Adapted from Fu et al. (2024))

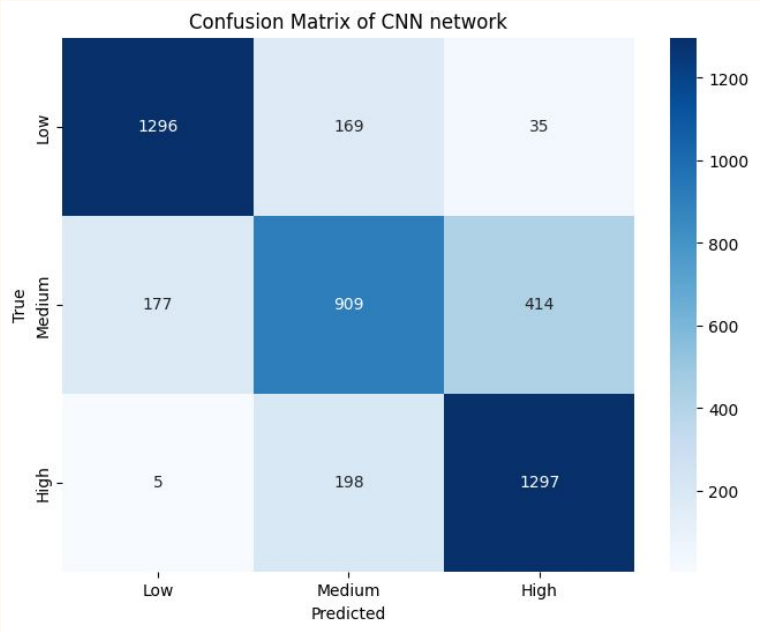
Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 1, 8, 16)	80
conv2d_1 (Conv2D)	(None, 1, 8, 32)	544
conv2d_2 (Conv2D)	(None, 1, 8, 32)	1,056
conv2d_3 (Conv2D)	(None, 1, 8, 64)	2,112
flatten (Flatten)	(None, 512)	0
dense (Dense)	(None, 128)	65,664
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 64)	8,256
dense_2 (Dense)	(None, 3)	195

Model 2 - Random Forest

Fu, S., Wang, L., Khalil, U., Cheema, A. H., Ullah, I., Aslam, B., Tariq, A., Aslam, M., & Alarifi, S. S. (2024). *Prediction of surface urban heat island based on predicted consequences of urban sprawl using deep learning: A way forward for a sustainable environment. Physics and Chemistry of the Earth, Parts A/B/C, 135*, 103682.

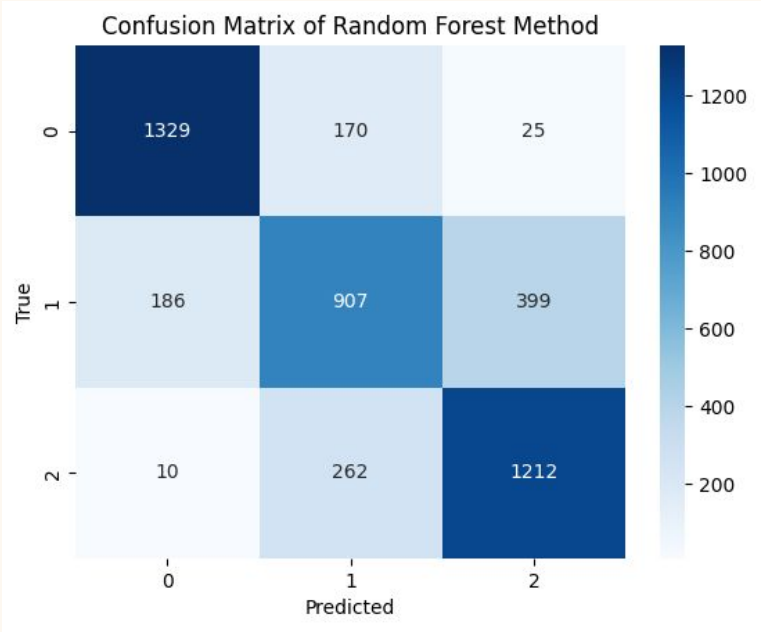
Result

Model 1 - CNN



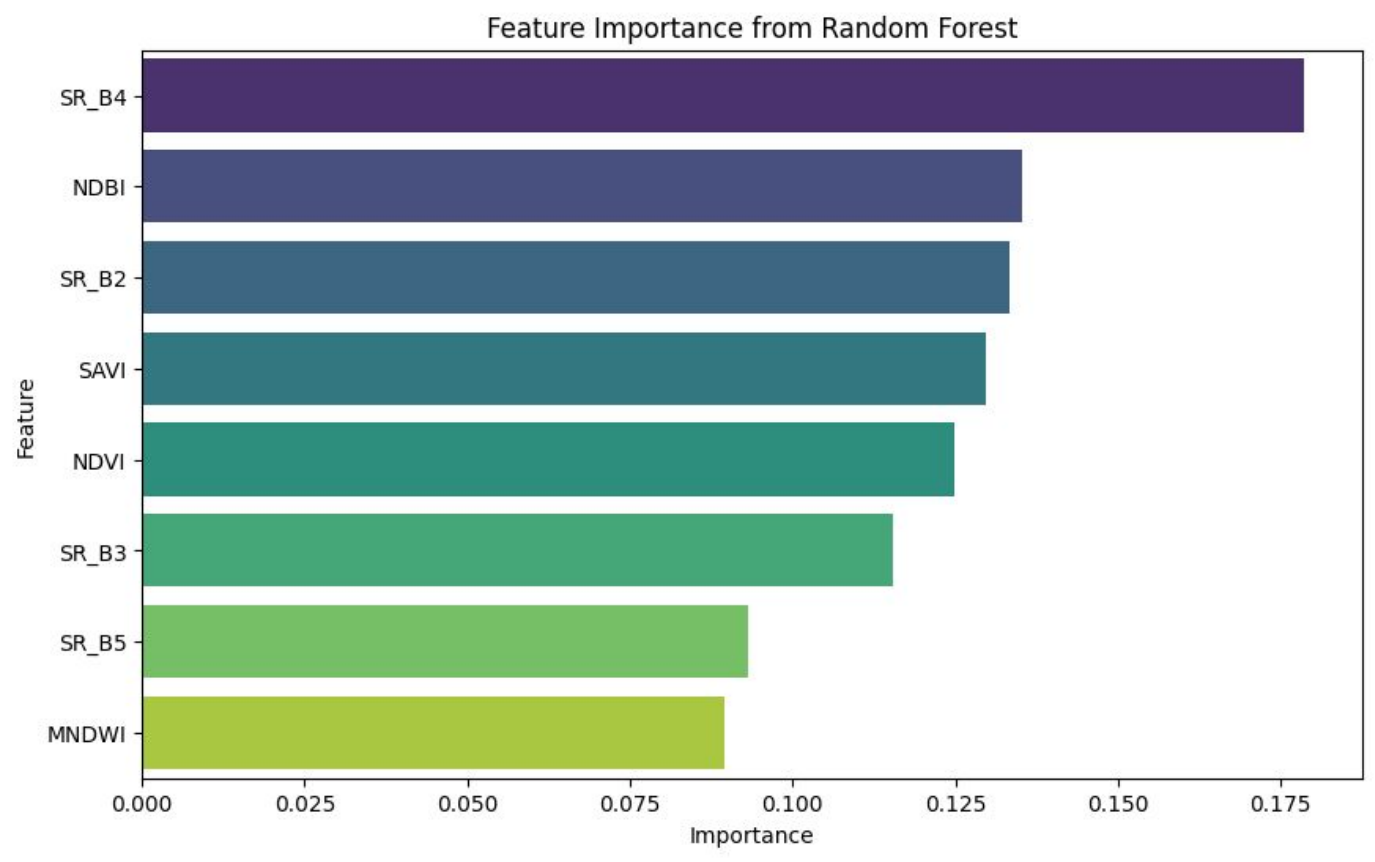
Overall Accuracy = 0.78

Model 2 - Random Forest



Overall Accuracy = 0.77

Result


Feature Importance from Random Forest
SR_B4
NDBI
SR_B2
SAVI
NDVI
SR_B3
SR_B5
MNDWI
Feature
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175
Importance

Future Approach

- Test generalizability across time and space
- Feature selection based on feature importance rank
- Use **air temperature** instead of LST, which better measures the heat risk. (Using deep learning and interpolation methods based on dispersed meteorological station data to construct a continuous raster dataset of air temperature.)