Urban Local Climate Zone classification through deep learning using spatio-temporal thermal imagery

Michaja van Capel
4657071

1st supervisor: Azarakhsh Rafiee
2nd supervisor: Roderik Lindenbergh

June, 2023

Contents

1 Introduction

In recent decades, urbanization has been rapidly transforming the global landscape, with a significant proportion of the world's population now living in urban areas (Zhang et al. 2022). This urban growth has given rise to numerous challenges, including the negative effects on urban micro-climates and the well-being of urban inhabitants. Understanding and characterizing the urban climate is crucial for mitigating these challenges and creating sustainable urban environments. To this end, the concept of Local Climate Zones (LCZs) has emerged as a valuable framework for urban climate classification and analysis.

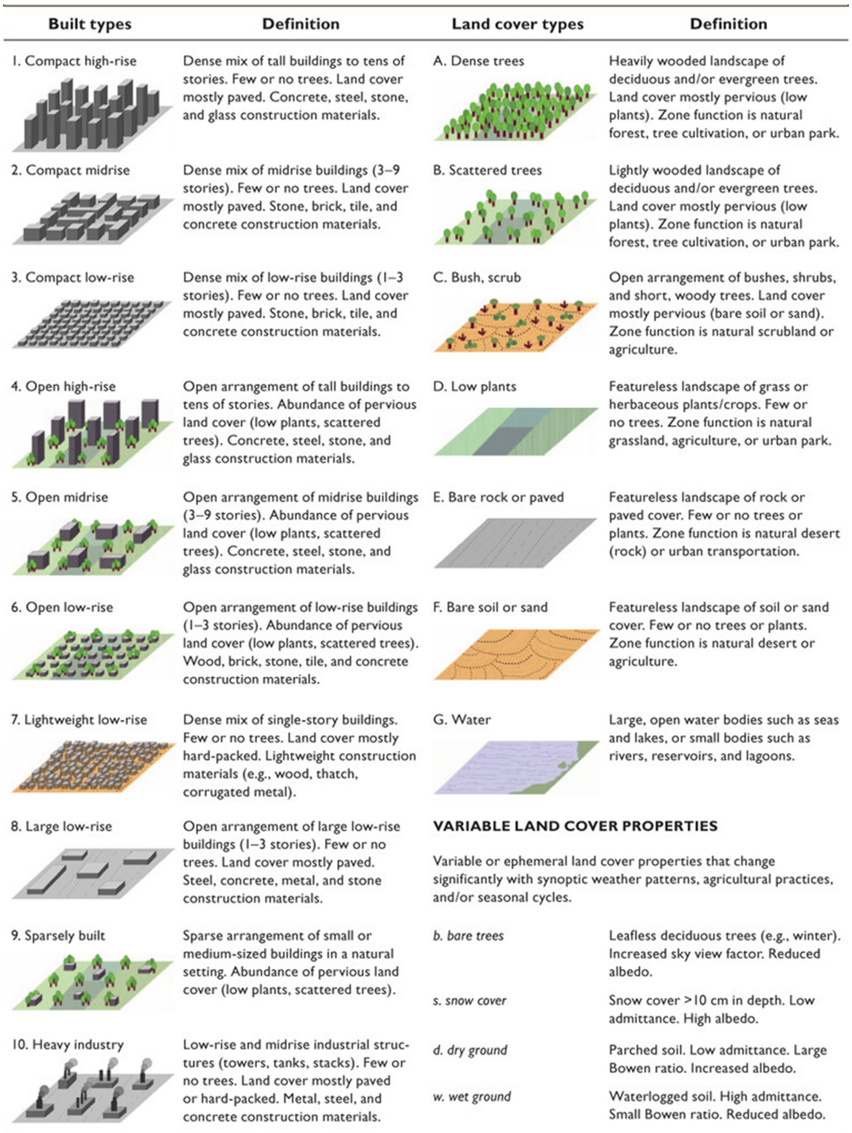
In recent years, the rapid advancement of remote sensing technologies, together with the power and development of deep learning algorithms, has provided promising opportunities for automated and data-driven LCZ classification. Among the various remote sensing modalities, thermal imagery stands out as a rich source of information for capturing the spatio-temporal thermal dynamics of urban areas. Thermal imagery enables the observation of thermal patterns within different LCZs. To harness the potential of thermal imagery for LCZ classification, this research aims to explore the suitability of Convolutional Long Short-Term Memory (ConvLSTM) networks. ConvLSTM is a variant of Recurrent Neural Networks that combines Convolutional and Long Short-Term Memory (LSTM) layers to effectively capture both spatial and temporal dependencies in sequential data (Xavier, 2021). With the spatio-temporal characteristics inherent of thermal imagery, ConvLSTM has the potential to overcome the limitations of traditional LCZ classification methods and achieve higher accuracy in urban climate zoning.

The aim of this thesis is to optimize an urban Local Climate Zone classification using temporal thermal imagery. The goal for the classification method is to be versatile and therefore to provide desirable classification results for any city.

2 Related work

2.1 Local Climate Zones

In 2012, the Local Climate Zone (LCZ) classification system was introduced by Stewart and Oke. Climate classifications were typically formulated to describe climate zones at larger scales, making them ineffective when applied to smaller, micro-scale areas. Sites in cities with very different physical and climatological features were usually described only as "urban" or "rural". The LCZ system aims to overcome this (Aslam & Rana, 2022). Within this system, there are 17 distinct zones, each characterized by its unique combination of surface structure, cover, and human activity. By considering these factors, the LCZ system provides a more accurate and detailed representation of the climate within specific areas. The different zones and their definition are shown in Figure 1. LCZ 1-10 are different built-up classes, and LCZ A-G different land cover types (Stewart & Oke, 2012).



Built types	Definition	Land cover types	Definition
1. Compact high-rise	Dense mix of tall buildings to tens of stories. Few or no trees. Land cover mostly paved. Concrete, steel, stone, and glass construction materials.	A. Dense trees	Heavily wooded landscape of deciduous and/or evergreen trees. Land cover mostly pervious (low plants). Zone function is natural forest, tree cultivation, or urban park.
2. Compact midrise	Dense mix of midrise buildings (3–9 stories). Few or no trees. Land cover mostly paved. Stone, brick, tile, and concrete construction materials.	B. Scattered trees	Lightly wooded landscape of deciduous and/or evergreen trees. Land cover mostly pervious (low plants). Zone function is natural forest, tree cultivation, or urban park.
3. Compact low-rise	Dense mix of low-rise buildings (1–3 stories). Few or no trees. Land cover mostly paved. Stone, brick, tile, and concrete construction materials.	C. Bush, scrub	Open arrangement of bushes, shrubs, and short, woody trees. Land cover mostly pervious (bare soil or sand). Zone function is natural scrubland or agriculture.
4. Open high-rise	Open arrangement of tall buildings to tens of stories. Abundance of pervious land cover (low plants, scattered trees). Concrete, steel, stone, and glass construction materials.	D. Low plants	Featureless landscape of grass or herbaceous plants/crops. Few or no trees. Zone function is natural grassland, agriculture, or urban park.
5. Open midrise	Open arrangement of midrise buildings (3–9 stories). Abundance of pervious land cover (low plants, scattered trees). Concrete, steel, stone, and glass construction materials.	E. Bare rock or paved	Featureless landscape of rock or paved cover. Few or no trees or plants. Zone function is natural desert (rock) or urban transportation.
6. Open low-rise	Open arrangement of low-rise buildings (1–3 stories). Abundance of pervious land cover (low plants, scattered trees). Wood, brick, stone, tile, and concrete construction materials.	F. Bare soil or sand	Featureless landscape of soil or sand cover. Few or no trees or plants. Zone function is natural desert or agriculture.
7. Lightweight low-rise	Dense mix of single-story buildings. Few or no trees. Land cover mostly hard-packed. Lightweight construction materials (e.g., wood, thatch, corrugated metal).	G. Water	Large, open water bodies such as seas and lakes, or small bodies such as rivers, reservoirs, and lagoons.
8. Large low-rise	Open arrangement of large low-rise buildings (1–3 stories). Few or no trees. Land cover mostly paved. Steel, concrete, metal, and stone construction materials.	**VARIABLE LAND COVER PROPERTIES**	Variable or ephemeral land cover properties that change significantly with synoptic weather patterns, agricultural practices, and/or seasonal cycles.
9. Sparsely built	Sparse arrangement of small or medium-sized buildings in a natural setting. Abundance of pervious land cover (low plants, scattered trees).	*b. bare trees*	Leafless deciduous trees (e.g., winter). Increased sky view factor. Reduced albedo.
		s. snow cover	Snow cover >10 cm in depth. Low admittance. High albedo.
10. Heavy industry	Low-rise and midrise industrial structures (towers, tanks, stacks). Few or no trees. Land cover mostly paved or hard-packed. Metal, steel, and concrete construction materials.	*d. dry ground*	Parched soil. Low admittance. Large Bowen ratio. Increased albedo.
		w. wet ground	Waterlogged soil. High admittance. Small Bowen ratio. Reduced albedo.

Figure 1: LCZ system by Stewart & Oke (2012)

2.2 LCZ and LST

Because of the differences in (combinations of) built-land cover types, the LCZ and Land Surface Temperature (LST) have correlations. Understanding these correlations has important potential for managing urban heat islands, improving urban micro-climates, addressing climate change impacts, and assessing environmental and ecological aspects of urban areas, and has therefore been the subject of numerous studies. The research by Ünal & Çilek (2021) shows differences in the mean LST values and differences per LCZ. Another study by Zhao et al. (2021) shows differences in seasonal LST variabilities per LCZ (Zhao et al. 2021). A general trend found is that the diurnal LST variation increases with the urbanization index (Chen et al. 2017). However, despite the demonstrated differences and trends, all these studies have concluded that more information and detailed investigations are needed to eliminate the gap between the LCZ and LST relationships.

2.3 LCZ classification methods

There are three main methods of creating a LCZ classification: manual sampling, GIS, and remote sensing. Each of these methods has its own advantages and limitations. Manual sampling involves collecting data on LCZ characteristics through on-site observations or surveys. However, this method is time-consuming and can be prone to biased results due to variations between different inputs. GIS uses spatial analysis techniques to integrate multiple data sources and generate LCZ classification maps. This method is more data-intensive than the other methods, but it can take into account numerous characteristics of LCZs. Ünal & Çilek (2021) classified LCZs in Adana city, Turkey using five parameters including building height, building surface fraction, aspect ratio, pervious and impervious surface fraction. Another study by Zheng et al. (2018) created a LCZ classification of Hong Kong with the same parameters, but also using areal mean SVF of non-building areas of the sample site and mean street width.

At a bigger scale, classification using remote sensing can be used. Demuzere et al. (2019) mapped Europe into LCZs using tools and techniques developed as part of the World Urban Database and Access Portal Tools (WUDAPT) project. The supervised pixel-based method enables LCZ classification using freely accessible Landsat imagery supported by NASA. LCZ classification using remote sensing is fast and cost-effective.

By incorporating temporal thermal heat imagery, LCZ classification methods can potentially achieve higher accuracy and objectivity in defining Local Climate Zones. Thermal heat imagery can capture fine-scale temperature variations within an urban area (Zhao et al., 2021). LCZ classification methods based on multi-spectral satellite imagery may not capture the micro-climate variations accurately.

2.4 Spatio-temporal deep learning architecture

2.4.1 ConvLSTM

ConvLSTM is a variant of a Long-Short Term Memory (LSTM) containing a convolution operation inside the LSTM cell. LSTM networks are well-suited to make predictions on time series datasets, but use one-dimensional input data. The LSTM cell contains a matrix multiplication instead of a convolution operation, which makes them not suitable for spatial sequence data. This convolution process allows the ConvLSTM multi-dimensional input data and to capture spatial features additional to the temporal features.

A ConvLSTM cell is shown in Figure 2. Each cell takes as input the hidden state from the previous time step (if any), along with the current input data, and produces an output along with a new hidden state. The convolutional operations are performed within the cell, replacing the (for LSTM) usual matrix multiplications with convolutions (Xavier, 2021).

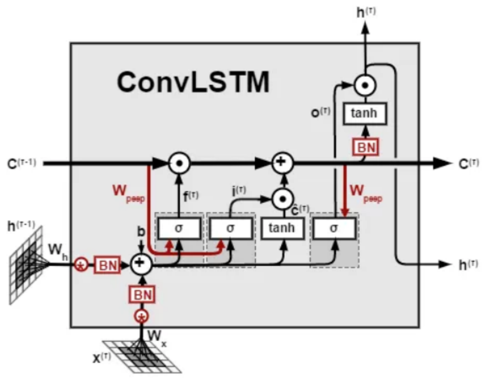


Figure 2: ConvLSTM cell (Xavier, 2021)

Where the Convolutional Neural Network (CNN) has proven to be efficient for image analysis (Ajit et al. 2020), ConvLSTM has been used for series of images or video analysis. ConvLSTMs have been successfully applied in various domains. A recent study by Huang et al. (2022) showcased the utilization of ConvLSTM for the prediction of vehicle driving intentions. In a similar way, big datasets of spatio-temporal data have been used to predict traffic accidents with a Hetero-ConvLSTM framework (Yuan et al., 2018). Furthermore, ConvLSTM has also been used for video saliency detection, specifically finding the most interesting segments in every video frame that attract human attention (Song et al., 2018). When it comes to spatio-temporal satellite data, the ConvLSTM framework has also been applied for soil moisture prediction. This was done using NDVI and NSMI satellite data (Habiboullah & Louly, 2023). Durrani et al. (2023) have applied a ConvLSTM for crop classification, by optimizing the hyper-parameters of the network. With the optimal parameters, the network led to an overall validation accuracy of 97.71%.

2.4.2 Hyper-parameters

The results of deep learning algorithms are highly dependent on the used hyper-parameters. This is why the hyper-parameters for the ConvLSTM network in this thesis will need to be optimized. Three chosen hyper-parameters and their role are explained.

- Batch Size: Batch size refers to the number of samples fed into the network during each training iteration. A larger batch size can potentially lead to more stable gradient estimates but requires more memory. Choosing an appropriate batch size involves balancing computational efficiency and model convergence.
- Filter Size: Filter size determines the spatial extent of the filters applied by the ConvLSTM network. Larger filter sizes capture more complex spatial patterns but increase the number of parameters in the model. Smaller filter sizes may capture more local details but could potentially miss larger-scale spatial dependencies. Selecting an optimal filter size depends on the specific characteristics of the dataset and the desired level of spatial granularity.

- Number of Layers: The number of layers in a ConvLSTM network represents its depth or the number of stacked ConvLSTM units. Increasing the number of layers allows the network to learn hierarchical representations of the input data, capturing both low-level and high-level features. However, deeper networks require more computational resources and may be prone to overfitting if not properly regularized.

The values of hyper-parameters are typically determined through experimentation and model performance evaluation (Durrani, et al. 2023).

3 Research questions

The main research question for this thesis is:

To what extent is a ConvLSTM using spatio-temporal thermal imagery suitable for the classification of urban Local Climate Zones?

In order to assess the performance of the ConvLSTM network, the different contributing factors need to be considered. The training data and architecture of the deep learning network can significantly influence the accuracy of the model. Therefore the following research sub-questions need to be addressed to provide an exploration of the main research question:

- How can a representable training data set be collected?
- When it comes to the architecture of ConvLSTM, what values for the hyper-parameters of the deep learning network lead to the best classification result?
- Which time frequency (day-night, seasonal) leads to the best classification result?
- How and why is the resulting LCZ classification map different from existing LCZ maps?

4 Methodology

The methodology of this thesis can be divided into three main parts, namely (1) data collection and pre-processing, (2) training and application of the ConvLSTM model, and (3) the evaluation of the results. The steps are shown in Figure 3. The steps will not be chronological and treated completely separately, since unsatisfactory results in part (3) will need review and adjustments in part (1) and/or part (2).

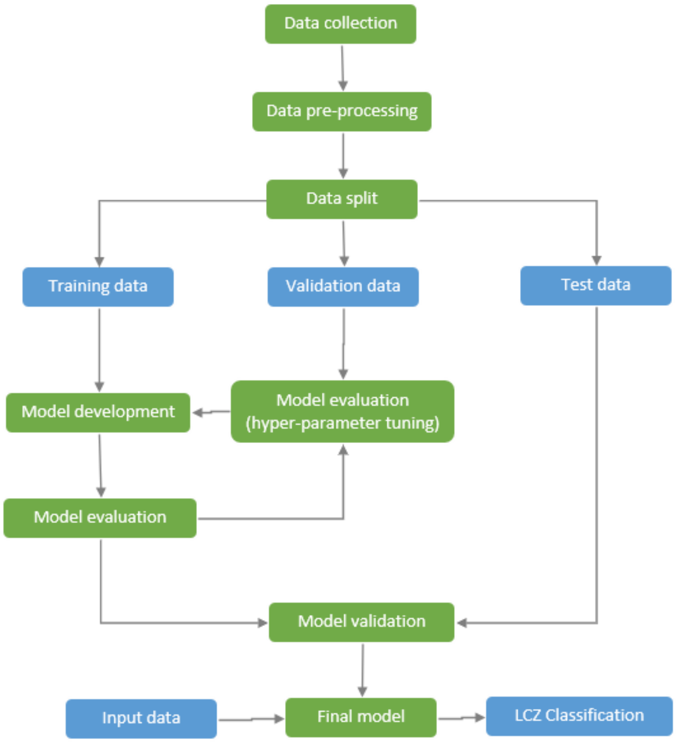


Figure 3: Workflow

4.1 Approach

Two cities in the Netherlands with different LCZ distributions were selected as study area for data collection, Rotterdam and Utrecht specifically. This was decided because the goal for the classification method is to be versatile and therefore to provide desirable classification results for any city.

Collecting representative and reliable training data for LCZ classification poses several challenges. Some of these challenges include:

- Spatial and temporal variability: LCZs exhibit significant spatial and temporal variability within urban areas. To train an accurate ConvLSTM model, it is crucial to collect training data that adequately captures this variability. This requires a diverse and comprehensive dataset that represents different LCZs and their temporal dynamics.
- Data availability and accessibility: Acquiring thermal imagery data at the desired spatial and temporal resolutions can be challenging.
- Class imbalance and data bias: Urban areas often exhibit class imbalance in terms

of the distribution of different LCZs. Some LCZs may be more prevalent or occur in larger spatial extents than others. This class imbalance can lead to biases in the training data, potentially affecting the performance of the ConvLSTM model.

- Different behavior of labeled data: For the classification labeled data will be used. LCZ classification maps exist but are created with different input data, models and requirements. Therefore, one LCZ on a classification map may have very different spatio-temporal thermal characteristics than the same LCZ on the same map. This will confuse the deep learning model.

To overcome the last challenge, the training data will require careful and accurate labeling. To analyze the thermal behaviour over time for different LCZs, an LCZ map of Europe created by Demuzere et al. (2019) will be used. Polygons will be drawn in each of the available LCZs in or close to the city of Rotterdam. They are shown in Figure 4. The available LST values inside the polygons over a time span of two months will be plotted and analyzed. The results will gain insight into the behavior of LST in different LCZs and help with the data labeling. It is possible that different LCZs show similar behavior, or that the same LCZ shows different behavior within the LCZ. If this hypothesis holds true, this means that not exactly the same classes as the map by Demuzere et al. (2019) will be used for the data labelling.

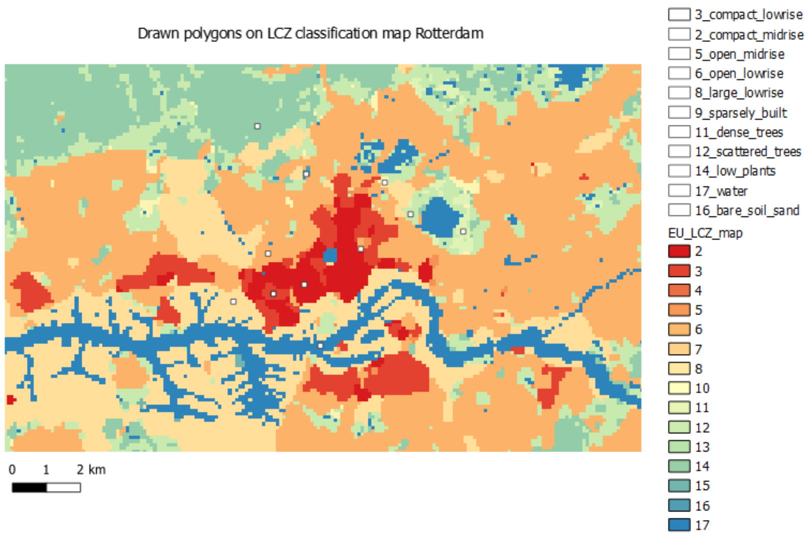


Figure 4: Drawn polygons on LCZ classication map Rotterdam

4.2 Tools and datasets

The tools and datasets needed for this thesis are listed below.

4.2.1 Tools

For the ConvLSTM model development (training, testing, and validating) code will be programmed in Python. Python offers a wide range of libraries for data analysis, machine learning, and image processing.

4.2.2 Datasets

The model will be trained with ECOSTRESS Thermal Imagery. The ECOSTRESS mission, operated by NASA, provides high-resolution (70m) thermal imagery data captured by the ECOSTRESS instrument on board the International Space Station (ISS). As already mentioned in section 4.1, the WUDAPT LCZ classification map will also be used for analysis of spatio-temporal thermal behavior.

These are the specific dataset to be used:

- ECOSTRESS Land Surface Temperature and Emissivity Daily L2 Global 70m V001
- ECO1BGEO v001 ECOSTRESS Geolocation Daily L1B Global 70 m
- WUDAPT LCZ map Europe (Demuzere et al. 2019)

5 Time planning

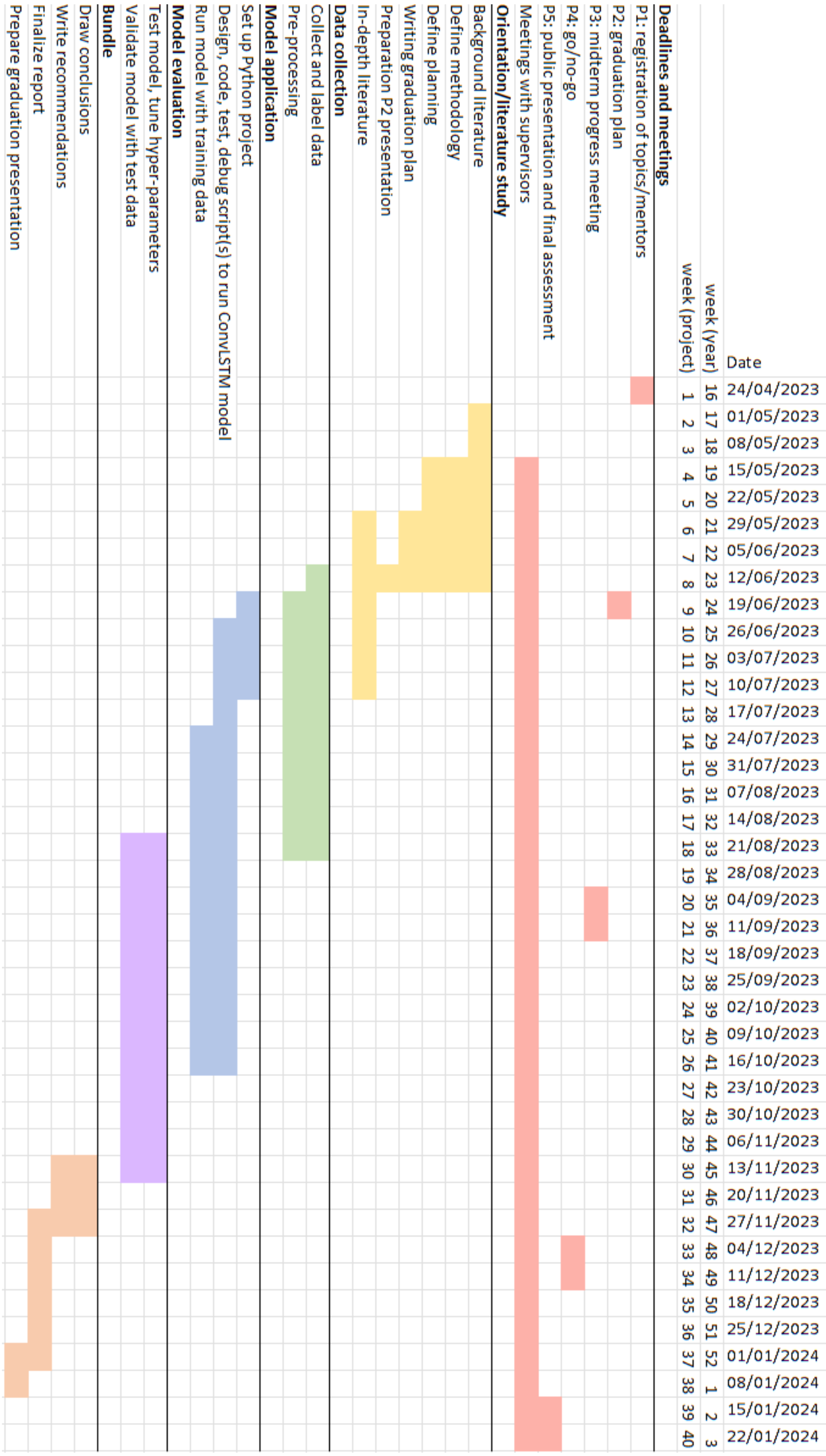


Figure 5: Time planning of the project

Bibliography

Ajit, A., Acharya, K., & Samanta, A. (2020). A Review of Convolutional Neural Networks. *International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, 1-5, https://doi.org/10.1109/ic-ETITE47903.2020.049

Aslam, A., & Rana, I.A. (2022). The use of local climate zones in the urban environment: A systematic review of data sources, methods, and themes. *Urban Climate*, 42, 101120. https://doi.org/10.1016/j.uclim.2022.101120

Demuzere, M., Bechtel, B., Middel, A., & Mills, G. (2019). Mapping Europe into local climate zones. *PLOS ONE, 14*(4), e0214474. https://doi.org/10.1371/journal.pone.0214474

Durrani, A.U.R., Minallah, N., Aziz, N., Frnda, J., Khan, W., & Nedoma, J. (2023). Effect of hyper-parameters on the performance of ConvLSTM based deep neural network in crop classification. *PLOS ONE, 18*(2), e0275653. https://doi.org/10.1371/journal.pone.0275653

Chen, Y., Chiu, H., Su, Y., Wu, Y., & Cheng, K. (2017). Does urbanization increase diurnal land surface temperature variation? Evidence and implications. *Landscape and Urban Planning, 157*, 247–258. https://doi.org/10.1016/j.landurbplan.2016.06.014

Habiboullah, A., & Louly, M.A. (2023). Soil Moisture Prediction Using NDVI and NSMI Satellite Data: ViT-Based Models and ConvLSTM-Based Model. *SN Computer Science, 4*(2), 140. https://doi.org/10.1007/s42979-022-01554-7

Huang, H., Zeng, Z., Yao, D., Pei, X., & Zhang, Y. (2022). Spatial-temporal ConvLSTM for vehicle driving intention prediction. *Tsinghua Science and Technology, 27*,(3), 599-609. https://doi.org/10.26599/TST.2020.9010061

Song, H., Wang, W., Zhao, S., Shen, J., & Lam, K. (2018). Pyramid Dilated Deeper ConvLSTM for Video Salient Object Detection. *Lecture Notes in Computer Science*, 11215. https://doi.org/10.1007/978-3-030-01252-6_44

Stewart, I., & Oke, T.R. (2012). Local Climate Zones for urban temperature studies. *Bulletin of the American Meteorological Society, 93*(12), 1879–1900. https://doi.org/10.1175/bams-d-11-00019.1

Ünal, M., & Çilek, A. (2021). Analyses of land surface temperature (LST) variability among local climate zones (LCZs) comparing Landsat-8 and ENVI-met model data. *Sustainable Cities and Society, 69*, 102877. https://doi.org/10.1016/j.scs.2021.102877

Xavier, A. (2021, December 9). *An introduction to ConvLSTM*. Medium. https://medium.com/neuronio/an-introduction-to-convlstm-55c9025563a7

Yuan, Z., Zhou, X., & Yang, Y. (2018). Hetero-ConvLSTM: A Deep Learning Approach to Traffic Accident Prediction on Heterogeneous Spatio-Temporal Data. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery*, 984–992. https://doi.org/10.1145/3219819.3219922

Zhang, Y., Li, Y., Chen, Y., Liu, S., Yang, Q. (2022). Spatiotemporal Heterogeneity of Urban Land Expansion and Urban Population Growth under New Urbanization: A Case Study of Chongqing. *Int. J. Environ. Res. Public Health 2022, 19*, 7792. https://doi.org/10.3390/ijerph19137792

Zhao Z., Sharifi A., Dong X., Shen L., He B. (2021). Spatial Variability and Temporal Heterogeneity of Surface Urban Heat Island Patterns and the Suitability of Local Climate Zones for Land Surface Temperature Characterization. *Remote Sensing, 13*(21), 4338. https://doi.org/10.3390/rs13214338

Zheng, Y., Ren, C., Xu, Y., Wang, R., Ho, J. C., Lau, K. K., & Ng, E. (2017). GIS-based mapping of Local Climate Zone in the high-density city of Hong Kong. *Urban Climate, 24*, 419–448. https://doi.org/10.1016/j.uclim.2017.05.008