COBRA - Snakes on the Moon

Vanshika Jain[1]
jain.van@northeastern.edu

Keshav Bharadwaj Vaidyanathan[1]
bharadwajvaidyanat.k@northeastern.edu

Vasu Kumar Reddy Sirigiri[1]
sirigiri.v@northeastern.edu

Ruthwik Hemmanahalli Manjunath[1]
hemmanahallimanjun.r@northeastern.edu

Dipak Sairamesh[1]
sairamesh.d@northeastern.edu

[1]**Northeastern University, Boston, MA**

December 16, 2023

Abstract

COBRA (Crater Observing Bio-inspired Rolling Articulator) is an innovative snake-inspired rover designed to explore and analyze lunar craters. The essential components of this system encompass a rolling articulator chassis designed for navigating challenging terrains, gaits inspired by biological sidewinding and tumbling movements, a perception module utilizing semantic segmentation of crater images, and modular elements featuring an actuated transformer mechanism. The perception module leverages deep learning techniques to automate environmental comprehension, thereby augmenting the overall navigation capabilities of the system. The training data included manually annotated samples, which were utilized to retrain a pixel-wise terrain classification model. Through testing on analog crater images without ground truth, the model demonstrated consistent identification of ground surfaces, motionless objects, and moving objects. Prospective improvements in autonomy and the incorporation of advanced scientific payloads will facilitate the exploration of permanently shadowed polar craters to study water ice deposits. This paper delves into the mechatronics, software architecture, innovative locomotion strategies, computer vision sensing, and the systematic validation process employed in crafting the COBRA. The outcomes highlight enhanced mobility achieved through a biologically inspired design, enabling exploration of extreme lunar environments that surpass the capabilities of current planetary rovers. Code for semantic segmentation is available at https://github.com/keshavbharadwaj/COBRA.

1. Introduction

COBRA Fig.1, a rover inspired by snake locomotion, is specifically designed for the exploration of lunar craters. It incorporates a unique multi-modal mobility system that combines sidewinding and hexagonal tumbling, enabling it to navigate steep crater slopes and loose regolith. The rover is composed of 11 identical body modules, including a head and tail, measuring 1.59 meters in length, 10 centimeters in diameter, and weighing 7.11 kilograms. The innovative mobility gaits of COBRA facilitate its descent into the Shackleton crater, where it conducts scientific measurements and transmits valuable data. Key advancements include bio-inspired gaits, a method for the center of mass manipulation to control tumbling, and a reversible transformation mechanism facilitating a switch between mobility modes. Testing of a 1/3 scale model, named COBRA Mini, validated the performance of sidewinding and tumbling. Future developments involve the integration of sensors and active control for autonomous crater exploration. COBRA introduces a pioneering approach to exploring challenging terrains on the Moon and other celestial bodies.

The bio-inspired design leverages principles of snake locomotion for mobility. Some of the key features include:

- **Rolling articulation:** The utilization of rolling articulation serves as a pivotal mechanism for adeptly navigating the irregular surfaces of lunar craters. The incorporation of modular segments interconnected by joints allows for flexibility and dynamic shape adaptation, enhancing the rover's ability to traverse diverse terrains.
- **Perception module:** The inclusion of a perception module equipped with computer vision capabilities significantly contributes to the rover's environmental

Figure 1. COBRA: Crater Observing Bio-inspired Rolling Articulator

sensing capabilities. This module empowers the rover to intelligently perceive and respond to its surroundings, ensuring efficient and informed decision-making during exploration.

- **Sidewinding and tumbling gaits:** The implementation of sidewinding and tumbling gaits represents an innovative approach to crater traversal, optimizing the rover's movement efficiency. These biologically inspired gaits enhance the rover's adaptability to challenging lunar landscapes, enabling it to overcome obstacles with greater ease.
- **Custom rover head:** A custom-designed rover head constitutes a comprehensive integration of cameras, electronics, and battery components. This tailored rover head enhances the overall functionality of the rover, providing a cohesive and efficient platform for data acquisition, processing, and energy management during exploration missions.

2. Problem Definition

2.1. Overview

The COBRA concept integrates articulated mobility and advanced autonomy features to facilitate the exploration and investigation of challenging lunar terrains. Initial testing efforts concentrate on refining terrain perception and semantic segmentation models to enhance the rover's adaptability and decision-making capabilities. Subsequent plans involve ongoing rover upgrades and field testing, particularly in lunar crater analog environments. The multi-functional nature of COBRA is designed to surpass the constraints of current planetary rovers, leveraging its innovative biologically-inspired design to enable unprecedented access to and study of extreme lunar landscapes.

2.2. Hardware Design

2.2.1 Mechanical Design

Fig.2 depicts the previous design of the COBRA head and its positioning on the COBRA system. A major component of space exploration missions is investigating lunar topography for potential water sources. Northeastern University's COBRA head, which navigates difficult crater environments, is essential to this effort. However, limitations exist with the current design. Protruding cables, insufficient fan protection, lack of a vision camera, and inadequate battery access present challenges for effective operation.

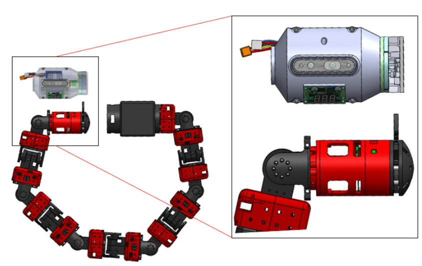


Figure 2. COBRA head and its positioning on the COBRA system

Problems with old COBRA

1. The outdated COBRA head lacks a vision camera, limiting autonomous hazard detection and navigation.

2. The absence of battery replacement options poses a risk of disrupted operations due to depleted power.

3. Exposed cables protruding from the head pose reliability risks and mobility limitations, diminishing the overall robustness of the system.

4. The Jetson Orin fan lacks protective mesh, leaving internal components vulnerable to harm from particles in the harsh lunar terrain, compromising dependability.

5. Inadequate shielding for vital parts, such as the fan, undermines resilience, especially considering the presence of abrasive dust in the extreme space environment.

6. The existing COBRA head design hinders the effective assessment of lunar ice water deposits within craters, limiting the robot's functionality to characterize frozen water resources.

Addressing these deficiencies through a SolidWorks-based redesign is crucial for persistent operation on the moon's surface and improved scientific contributions to lunar exploration objectives, particularly in supporting the feasibility of future colonization efforts.

2.2.2 COBRA Head Redesign

Fig.3 displays the updated design of the COBRA head. To address key deficiencies, an enhanced COBRA head is proposed with several new capabilities:

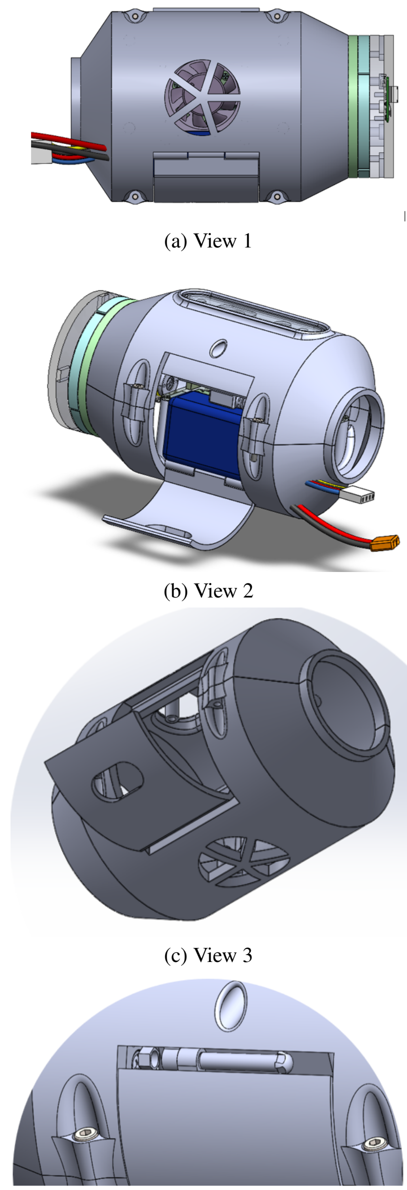
1. Visual Perception Enhancement: Integration of dedicated space and mounts for cameras to enhance visual perception capabilities. Enables efficient obstacle identification through the deployment of neural network algorithms onboard.

2. Battery Accessibility Optimization: Incorporation of sliding glass panels to facilitate easy access to batteries, ensuring swift and hassle-free exchange. Introduces a quick slide-in/slide-out method for rapid battery replacement in under 60 seconds, minimizing downtime during operations.

3. Durable Internal Cable Management: Implementation of internal cable routing through integrated conduits to enhance overall durability. Utilization of plastic ducts to safeguard wires against abrasion, and crimping, and reduce snag hazards, ensuring sustained performance in challenging lunar conditions.

4. Protective Fan Mechanism: Integration of a protective mesh for the fan to prevent particle damage without disrupting airflow. The inclusion of a filter to prevent the ingestion of moon dust into the CPU heatsink and fan mechanism enhances the longevity of critical components.

These design improvements collectively contribute to a more resilient and efficient lunar rover, addressing key aspects such as visual perception, battery accessibility, cable management, and fan protection for enhanced performance in extreme space environments.



(a) View 1

(b) View 2

(c) View 3

(d) View 4

Figure 3. COBRA new head design

2.2.3 3D printing

Table.1 shows the 3D printing parameters used to print latching door and head housing module

2.3. Perception Module

A major addition to COBRA is the incorporation of a perception module to enable computer vision capabilities.

Feature	Latching Door	Head Housing Module
3D Printing Process	Fused Deposition Modeling (FDM)	Fused Deposition Modeling (FDM)
Slicing Software	Ultimaker Cura	Ultimaker Cura
Material	Generic PLA (Polylactic Acid)	Generic PLA (Polylactic Acid)
Layer Resolution	0.15 mm	0.15 mm
Orientation	Straight up on the 3D printer bed	Straight up on the 3D printer bed
Infill Pattern	Triangular	Triangular
Shell Thickness	0.8 mm and 1 mm	0.8 mm and 1 mm
Support	Normal support	Normal support
Weight	78 grams	152 grams
Printing Time	2.5 hours	20 hours

Table 1. 3D printing parameters

This allows the robot to actively sense and interpret its surroundings to make informed decisions. The module hardware consists of:

- Nvidia Jetson Orin: Powerful edge AI platform that runs the deep learning models.
- Intel RealSense Camera: Provides RGB and depth imaging for model training.

The software techniques used include:

- Object Detection: Identifies and localizes objects in the scene.
- Semantic Segmentation: Classifies image pixels into semantic classes to understand the layout.

Together, these algorithms allow robust scene understanding:

1. Robust Object Detection: Implementation of advanced algorithms for precise identification and localization of objects within the rover's operational environment. Enhances the rover's ability to navigate and interact with its surroundings by providing real-time information about objects in the scene.

2. Semantic Segmentation for Environmental Analysis: Utilization of semantic segmentation algorithms to classify image pixels into semantic classes, contributing to a comprehensive understanding of the terrain's layout. Enables the rover to differentiate between various elements in its surroundings, fostering informed decision-making during exploration missions.

3. Scene Understanding for Informed Navigation: Integration of object detection and semantic segmentation collectively contributes to robust scene understanding. Empowers the rover with the capability to navigate through complex terrains by comprehensively analyzing the environment, identifying obstacles, and understanding the spatial layout.

4. Adaptive Decision-Making: The combination of object detection and semantic segmentation supports adaptive decision-making by providing the rover with a holistic understanding of its surroundings. Enables the rover to respond dynamically to diverse environmental challenges, ensuring a more versatile and resilient exploration capability.

These integrated algorithms not only facilitate object detection and semantic segmentation but also synergize to provide the rover with a sophisticated level of scene understanding, ultimately enhancing its overall autonomy and navigation capabilities.

2.3.1 Semantic Segmentation

This module incorporates semantic segmentation algorithms to analyze camera images and better comprehend the surrounding terrain. Semantic segmentation refers to assigning a classification label to each pixel in an image, such as ground, rocks, craters, etc. This provides much more detailed scene understanding compared to basic object detection.

The COBRA team collected a novel dataset of images from inside and outside their lab facilities. These images were manually annotated with semantic labels to generate ground truth data for algorithm training. Four main terrain classes were defined: **moving objects, stationary objects, ground areas, and other regions**. Preprocessing steps included analyzing class imbalance and applying data augmentation techniques.

A deep neural network model called MIT ADE20K, originally designed for autonomous driving datasets, was adapted and retrained for this application. The model outputs a segmented image with each pixel classified according to the trained labels. Evaluation using Intersection over Union (IoU) metrics shows reliable performance in detecting ground and stationary objects. Further optimization of the perception module and inference speed is ongoing, with the goal of eventual integration on the COBRA rover prototype.

The addition of semantic segmentation gives COBRA a vision-based understanding of the lunar landscape to safely navigate and select optimal locations to gather scientific data during its crater exploration mission. This also serves as a pilot application and benchmark for deploying deep learning algorithms on lunar rovers.

3. Experiments

3.1. Data Collection

Data collection is essential to create a dataset to train machine learning models. 10 videos were recorded in an indoor lab and outdoor environment using an Intel Realsense camera. This results in a good diversity of scenes. The frame rate was around 30 fps and the resolution of each frame was 720 x 1280 providing detailed RGB images. The data collection process consisted of:

- Location: Indoors and outdoors in the Robotics Systems Lab
- Details: 10 videos captured, with 30 FPS totalling 10,000 frames
- Resolution: 720 x 1280 px RGB images

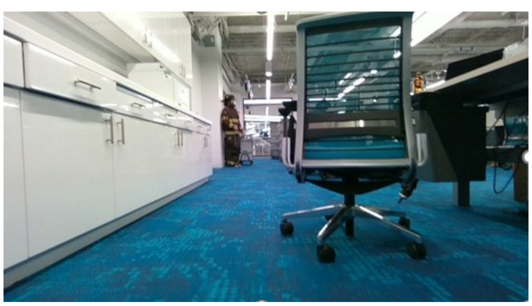
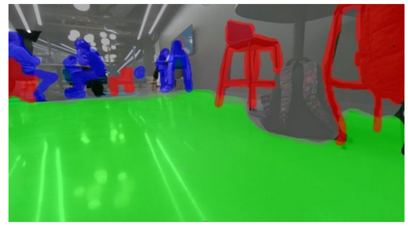
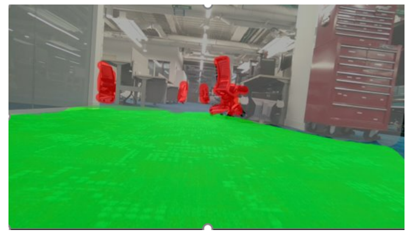


Figure 4. Example of raw image frame

A set of 100 frames was chosen randomly, and each frame underwent manual labeling with pixel-wise annotations. The annotations categorized the pixels into four classes: moving objects, non-moving objects, ground, and miscellaneous background regions. The annotation process encountered challenges, particularly in dealing with lighting reflections and accurately labeling intricate shapes such as chairs. Despite these challenges, the hand-labeling process aimed to establish a comprehensive dataset for subsequent analysis and model training.



(a) Example 1



(b) Example 2

Figure 5. Example of images Labelled using Label Studio. Green represents ground,Gray represents Immovable Objects, Red represents Movable Objects and Blue represents Misc

3.2. Data Labelling

Manual labeling was employed to create ground truth data, serving as the basis for evaluating the accuracy of the model. The open-source tool, Label Studio, was chosen for its accessibility and used in the labeling process. Four distinct classes were designated to represent various scenes encountered by the COBRA during its operation.

1) Ground: Represents the path where COBRA can navigate freely.

2) Immovable Objects: Encompasses obstacles, such as walls and ceilings, that cannot be shifted.

3) Movable Objects: Includes obstacles like chairs that could be shifted or moved.

4) Misc: Encompasses objects or structures of interest, such as humans. Fig .5 shows examples of labelling done using Label Studio. Acknowledging the potential for enhanced granularity, the addition of more labels could further refine the dataset and improve the model's understanding of diverse scenarios. Ten images from each video were randomly selected to form the test set, which was then labeled to assess model performance. While the dataset consisted of a modest 100 labeled frames, it is considered reasonable for initial analysis, providing a representative glimpse of dataset variability.

This labeling approach, utilizing manual labeling and an accessible tool like Label Studio, ensures the creation of a reliable ground truth dataset for evaluating the model's accuracy and performance across diverse environmental scenarios encountered by the COBRA.

3.3. Data Preprocessing

Preprocessing techniques were implemented to improve the consistency of the data and enhance the stability of model training. The application of standard techniques contributed to faster convergence when compared to training on raw, unprocessed data.

1. **Normalization:** The dataset underwent normalization to rescale data and address variability. The standard scalar method was employed to normalize pixel intensities, ensuring a more uniform and manageable input for the model.

2. **Cropping:** Cropping was applied to minimize background noise, focusing on refining annotations by cropping them to tight object boundaries. This step aimed to enhance the clarity and relevance of training data.

Fig.7 example of different ground truth masks for a frame.

Examination of class distribution revealed significant imbalances. Moving objects and miscellaneous classes exhibited considerably fewer samples compared to ground and background classes.

These preprocessing steps, encompassing normalization and cropping, play a vital role in improving the overall quality of the dataset for model training. Additionally, addressing class imbalances is crucial for ensuring that the model is trained on a representative and diverse set of data, contributing to its robust performance across different classes. Fig. 6 shows ground truth distribution for test data.

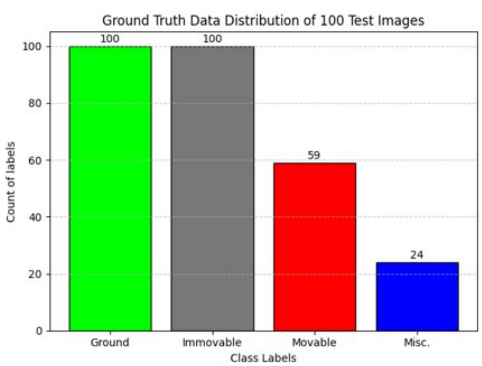
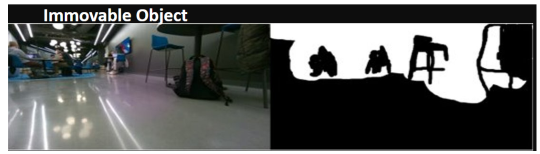


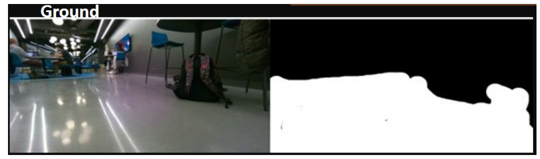
Figure 6. Ground Truth Distribution

3.4. Model Architecture and Training

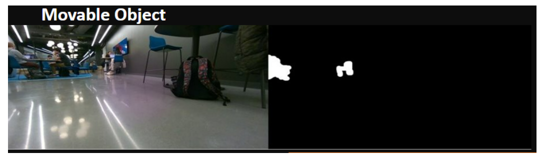
The base model architecture used was the open-source MIT ADE20K semantic segmentation model. This deep CNN extracts multi-scale features which are aggregated and upsampled to pixel-wise predictions. The autoencoder leverages a pre-trained encoder for efficient feature extraction paired with a semantic segmentation decoder to classify image regions. Fine-tuning adapts the layers to a novel dataset. Fig. 9 shows the entire software pipeline from raw frame to semantic segmented and smoothed image output.



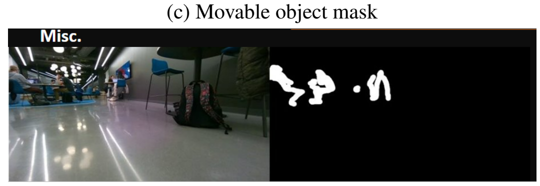
(a) Immovable object mask



(b) Ground Mask



(c) Movable object mask



(d) Misc mask

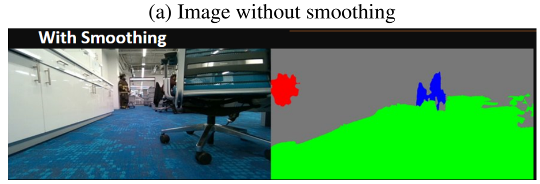
Figure 7. Ground Truth Masks

The model leverages state-of-the-art techniques for the task. Training on large datasets like ADE20K enables reasonable performance even with limited labeled frames.

3.4.1 Smoothing

A smoothing mechanism based on area is employed to mitigate noisy labels. Contours with an area below a specified threshold (hyperparameter) are either eliminated or smoothed. Lowering the threshold reduces the model's sensitivity to smaller objects, as they are more likely to be smoothed out. Fig.8 shows Effects of smoothing

(a) Image without smoothing



(b) Image after smoothing

Figure 8. Effect of smoothing

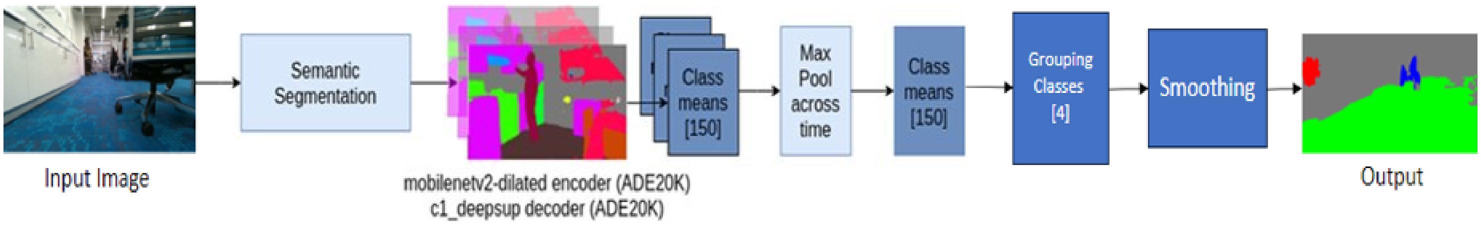


Figure 9. Software Pipeline

3.4.2 Evaluation

Model evaluation was done using mean Intersection-over-Union (mIOU) per class. This measures overlap between prediction and ground truth. Results showed: Strong mIOU (>75%) for ground and non-moving objects owing to more training data Weak mIOU (<30%) for other classes with fewer samples Thus, the priority was tuning performance for ground and background.

IOU Results: IOU evaluates semantic segmentation accuracy by comparing pixel label assignments between predictions and truth. Scores near 1 indicate strong consistency. The IOU scores demonstrate accurate classification, especially for ground and static objects. The misc. category understandably suffers from high variability as can be seen in Table.2. Smoothing improves the consistency of predictions.

Label Name	IOU without Smoothing	IOU with Smoothing
Ground	0.77	0.78
Immovable Objects	0.75	0.76
Movable Objects	0.05	0.03
Misc.	0.31	0.26

Table 2. IOU performance metric with and without smoothing

3.4.3 Optimization

To optimise the model for real-time usage on Jetson Orin, OpenVino was used: Quantization: Reduces model precision for a smaller memory footprint. This improves inference speed by 2-3x as can be seen in Fig. 10 with negligible accuracy loss as can be seen in Table.3. The memory footprint of the model was reduced from 8.72mb to 3.88mb 2x reduction in size.

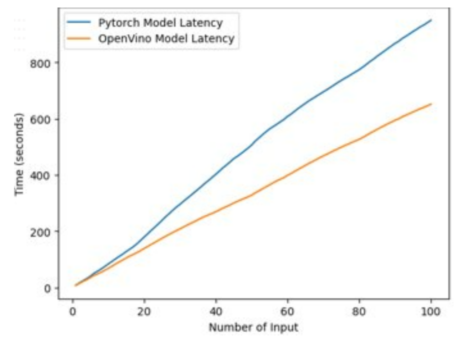


Figure 10. Latency of OpenVino optimized model and pytorch model

Label Name	Model	Optimized model
Ground	0.78	0.78
Immovable Objects	0.76	0.75
Movable Objects	0.03	0.03
Misc.	0.26	0.26

Table 3. IOU performance comparison of Optimised model

4. Conclusions

In summary, the proposed redesign of the COBRA head encompasses enhancements in vision integration, battery replacement mechanisms, cabling infrastructure, and fan protection. This comprehensive redesign aims to address the limitations inherent in the current model, ultimately boosting performance, dependability, and overall functionality. These improvements are strategically aligned with the goal of optimizing lunar navigation capabilities and advancing the rover's capacity for detecting water deposits within lunar terrains. The proposed changes are anticipated to significantly elevate the rover's effectiveness in achieving its exploration objectives on the moon.

Acknowledgement

This work was conducted at Silicon Synapse Lab, Northeastern University , Boston as part of the Assistive Robotics course. We'd like to thank Prof Alireza Ramezani for his guidance as well as Adarsh Salagame and Darshan Krishnamoorthy for their support throughout the course.

References

[1] Long, J., Shelhamer, E. & Darrell, T. "Fully Convolutional Networks for Semantic Segmentation," *IEEE Conference on Computer Vision and Pattern Recognition*, doi: https://doi.org/10.48550/arXiv.1411.4038.

[2] Chen, L., Papandreou, G., Kokkinos, I., Murphy, K. & Yuille, A. "Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs," *IEEE Computer Vision and Pattern Recognition* doi: https://doi.org/10.48550/arXiv.1412.7062.

[3] Nie, D., Xue, J., Ren, X. "Bidirectional Pyramid Networks for Semantic Segmentation". *Computer Vision – ACCV 2020* doi:https://doi.org/10.1007/978-3-030-69525-5_39

[4] Ronneberger, O., Fischer, P., Brox, T., "U-Net: convolutional networks for biomedical image segmentation," *MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015)* doi: https://doi.org/10.1007/978-3-319-24574-4_28

[5] Sun, K., et al, "High-resolution representations for labeling pixels and regions," *IEEE Computer Vision and Pattern Recognition* doi: arXiv:1904.04514.