Wildfire Image Analysis

M. Tzimas, D. Fotiou, P. Dinopoulos, Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@csd.auth.gr
www.aiia.csd.auth.gr
Version 3.1

Artificial Intelligence &
Information Analysis Lab

Wildfire Image Analysis

- **Computer Vision Tasks**
- Image Region Segmentation
- Semantic Image Segmentation
- Fire Detection
- Fire Segmentation

Computer Vision Tasks

Computer vision

- **Classification** → Classify an image into a specific class
- **Object detection** → Classify and localize objects within an image
- **Semantic segmentation** → Classify each pixel in an image into a specific class

Computer Vision Tasks

Image Classification

Fire / No Fire

Smoke / No Smoke

Burnt area / No Burnt area

Artificial Intelligence & Information Analysis Lab

Computer Vision Tasks

Object Detection

Fire detection

Smoke Detection

Computer Vision Tasks

Object Detection = classification + localization

•Find **what** is in a picture as well as **where** it is.

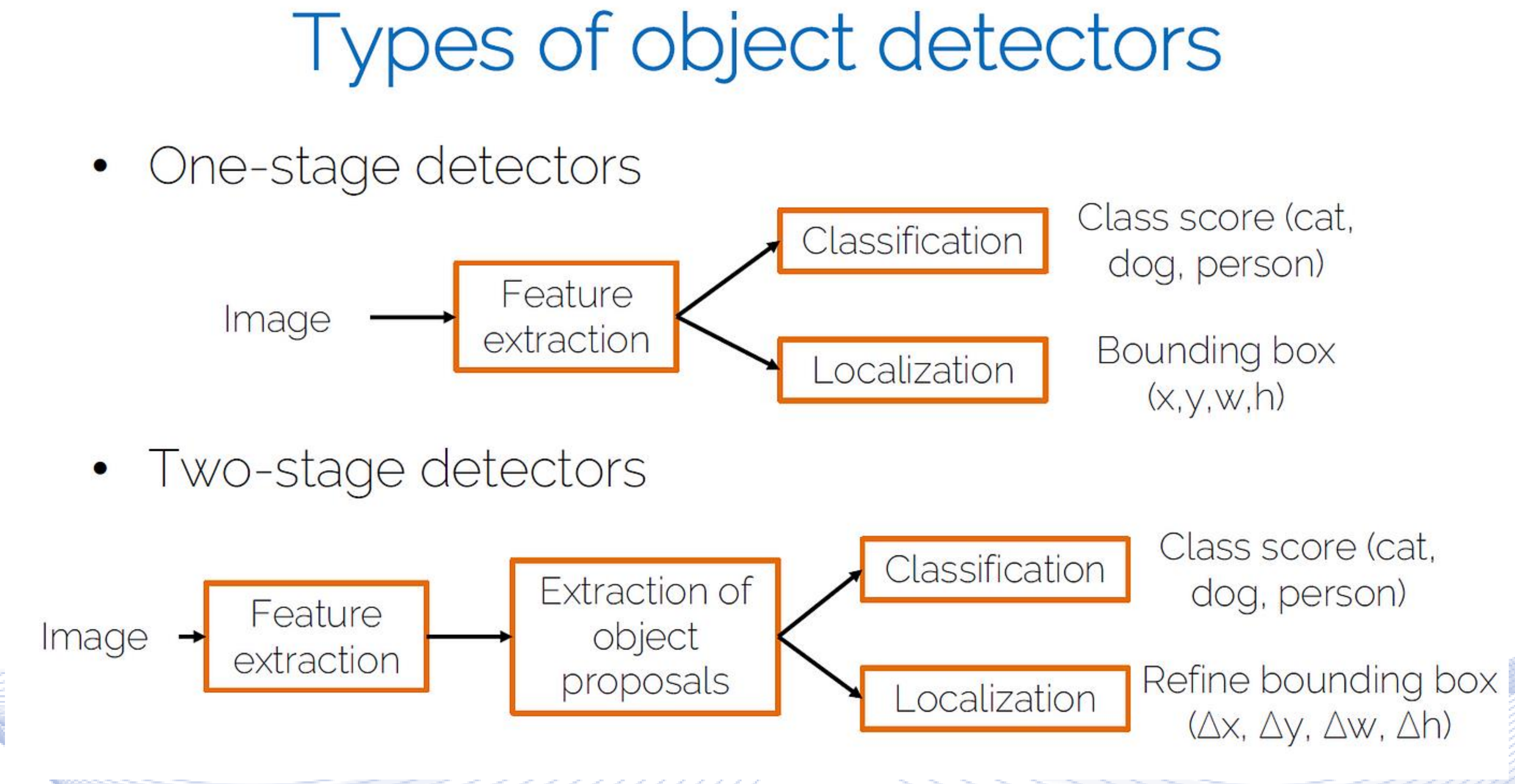
Artificial Intelligence &
Information Analysis Lab

Computer Vision Tasks

Classification – Regression

- Given a training set of **images annotated with bounding boxes** (coordinates and class per depicted object)
  - ***Classification***: predict probabilities that each box belongs to each of the classes present in the dataset
  - ***Regression***: for each depicted object predict bounding box coordinates in some predefined format, e.g., coordinates of the bounding box center along with its width and height (x, y, w, h).

Computer Vision Tasks



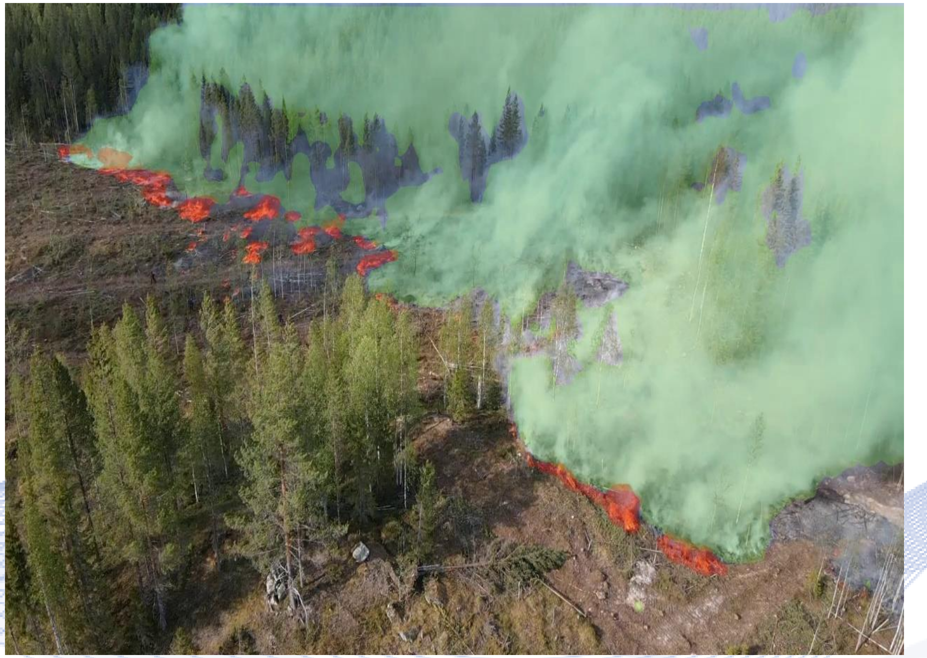
One stage vs Two stage object detection architectures [THA2023].

Computer Vision Tasks

Semantic Image Region Segmentation

- Assign class labels to each image pixel.

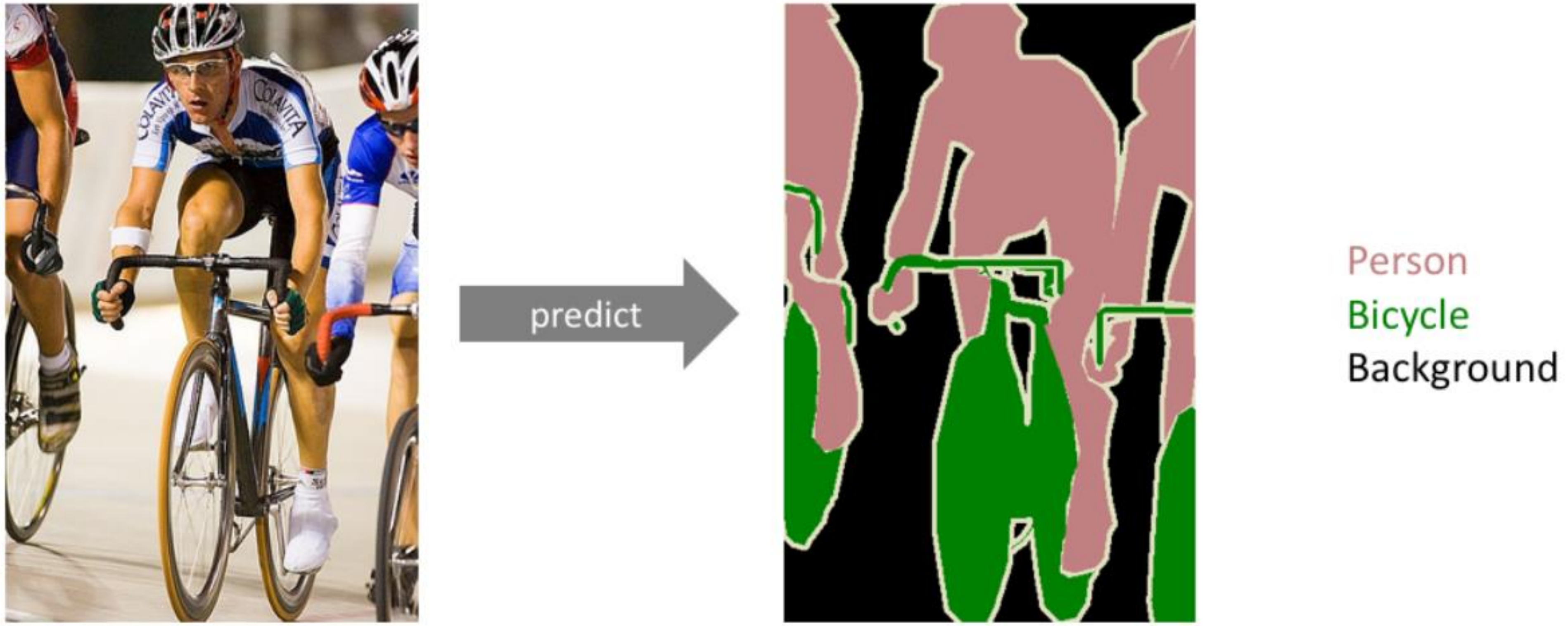
Fire Segmentation



Fire/Smoke Segmentation

Computer Vision Tasks

Semantic Image Region Segmentation



Semantic image segmentation of a sports event [EVE2011].

Computer Vision Tasks

Semantic Image Region Segmentation

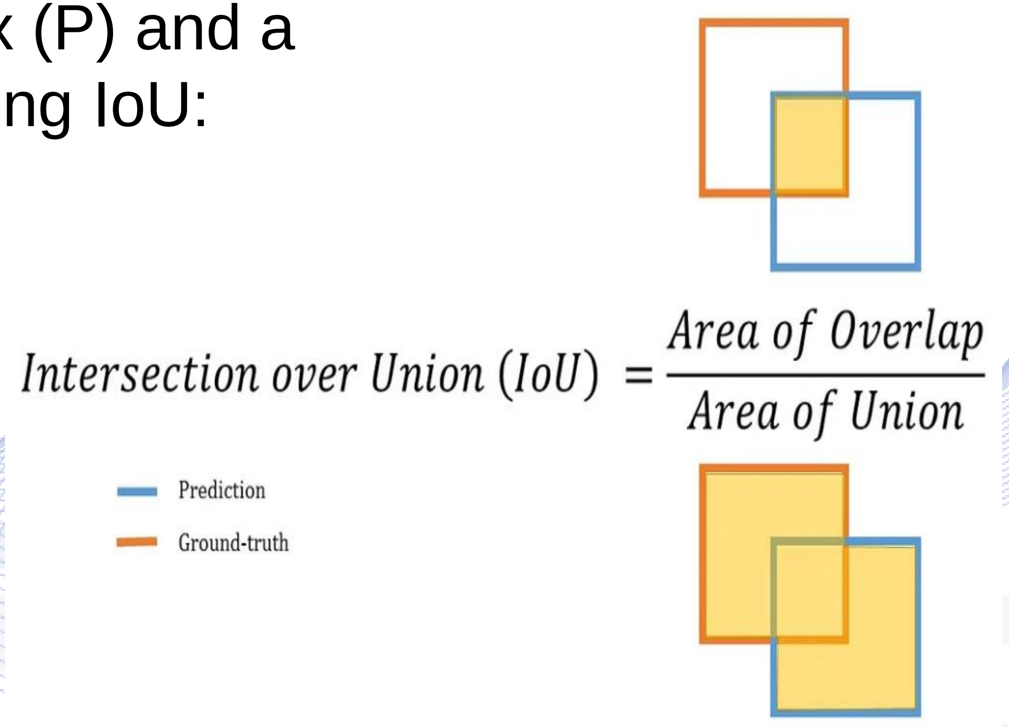
Semantic image segmentation for autonomous driving [COR2016].

Evaluation Measures

Intersection Over Union (IoU)

The overlap between a predicted bounding box (P) and a ground truth bounding box (G) is measured using IoU:

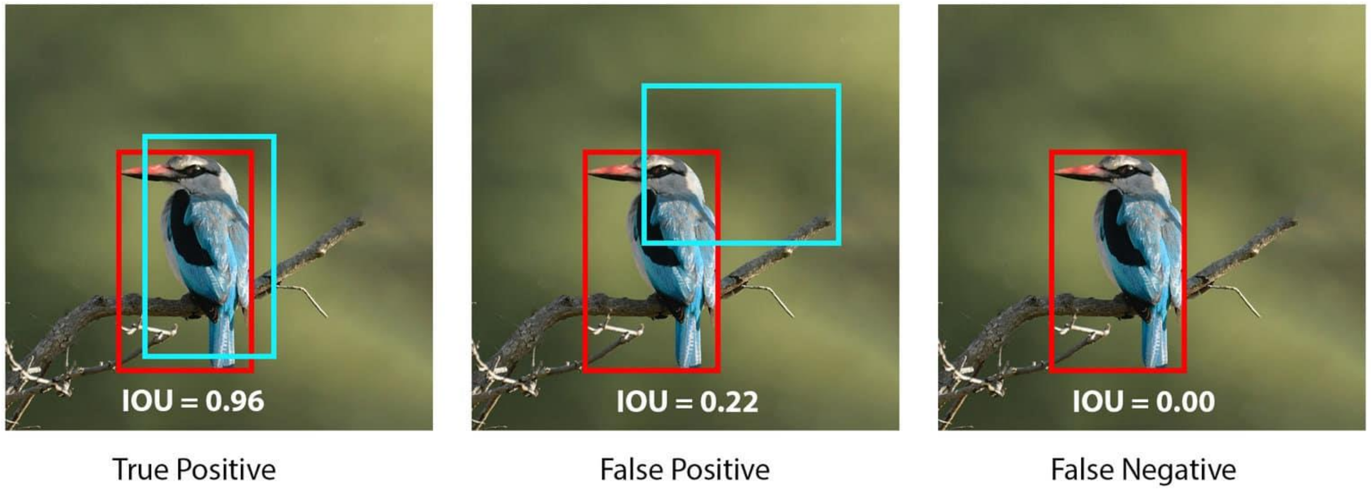
$$IoU(P,G) = \frac{|P \cap G|}{|P \cup G|}.$$



Evaluation Measures

Predicted Bound box evaluation :

TP : IOU > 0.5
FP : IOU < 0.5



Bounding box evaluation [KUK2023].

Evaluation Measures

Mean Average precision (mAP) for object detection:

- Combines precision and recall to evaluate detection accuracy.
- Uses Intersection over Union (IoU) to match predicted and ground-truth bounding boxes.
- Calculates the average precision for each class and averages across all classes.
- Rewards precise alignment and penalizes missing or incorrect predictions.

Mean IoU (mIoU) for Segmentation:

- Averages IoU across all pixels and classes to assess segmentation quality.

Wildfire Image Analysis

- Computer Vision Tasks
- **Image Region Segmentation**
- Semantic Image Segmentation
- Fire Detection
- Fire Segmentation

Artificial Intelligence & Information Analysis Lab

Image Region Segmentation

Definitions

- An image domain $\mathcal{X}$ must be segmented in N different regions $R_1, \ldots, R_N$.
- The segmentation rule is a logical predicate of the form $P(\mathcal{R})$
- Image segmentation partitions the set $\mathcal{X}$ into the subsets R_i, $i = 1, \ldots, N$, having the following properties:

$$\mathcal{X} = \cup_{i=1}^{N} R_i,$$
$$R_i \cap R_j = \emptyset, \qquad i \neq j,$$
$$P(R_i) = TRUE, \qquad i = 1, \ldots, N,$$
$$P\left(R_i \cup R_j\right) = FALSE, \qquad i \neq j,$$

Image Region Segmentation


Image thresholding.

Image Region Segmentation

Image thresholding

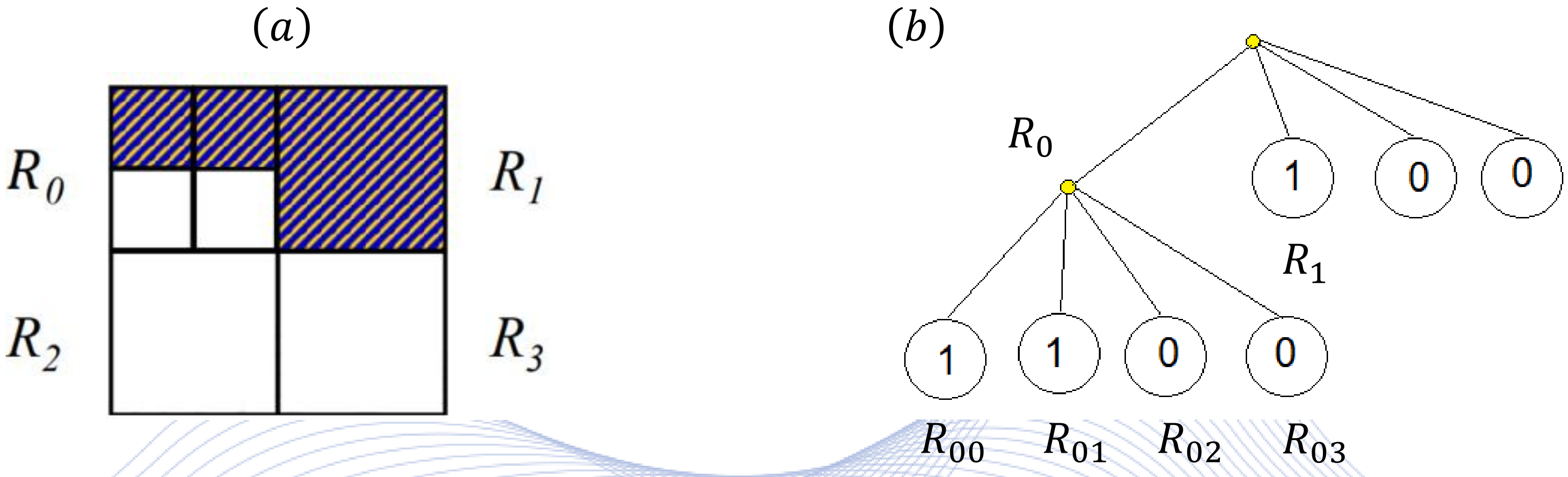
(a)

(b)

a) Original image; b) Image thresholding in four equirange regions.

Image Region Segmentation

Split/merge segmentation algorithm



a) Image segmentation by region splitting; b) Quadtree.

Wildfire Image Analysis

- Computer Vision Tasks
- Image Region Segmentation
- **Semantic Image Segmentation**
- Fire Detection
- Fire Segmentation

Semantic Image Segmentation

CNN Image Classification

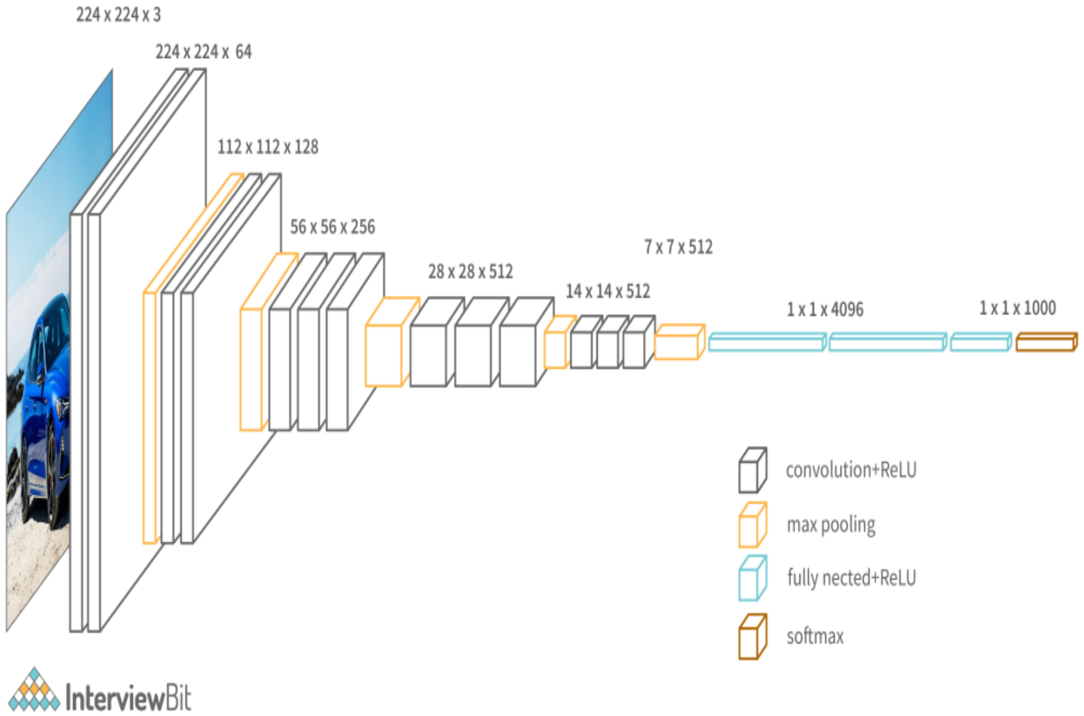
Convolution layers produce feature maps.

- The final feature map is **flattened** into a 1D vector.

Fully Connected Layers:

• Reduce dimensionality to match the number of classes in the dataset.

• Perform the final classification by mapping features to class probabilities.



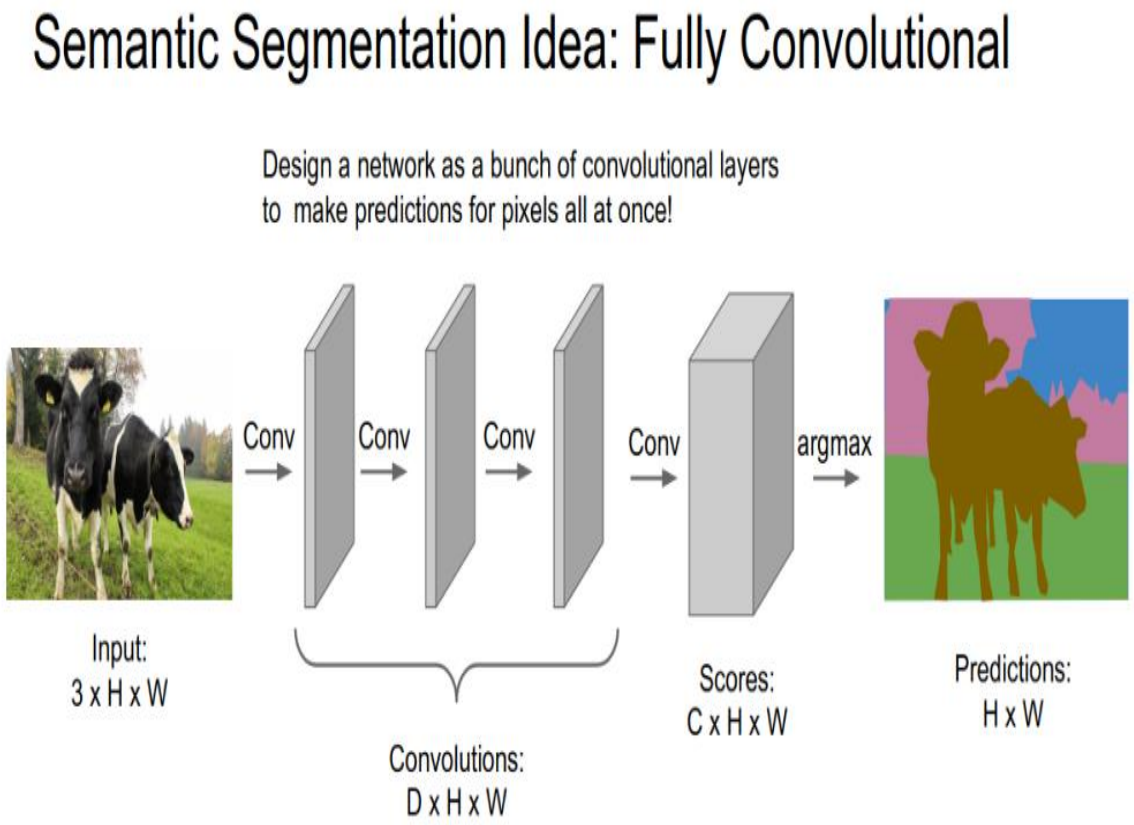
[BIT2024]

Semantic Image Segmentation

Semantic image segmentation

In contrast to Image classification, in segmentation C segmention maps of dimensions $H \times W$ are produced:

- C: Number of classes in the dataset.
- H, W: Height and width of the image.

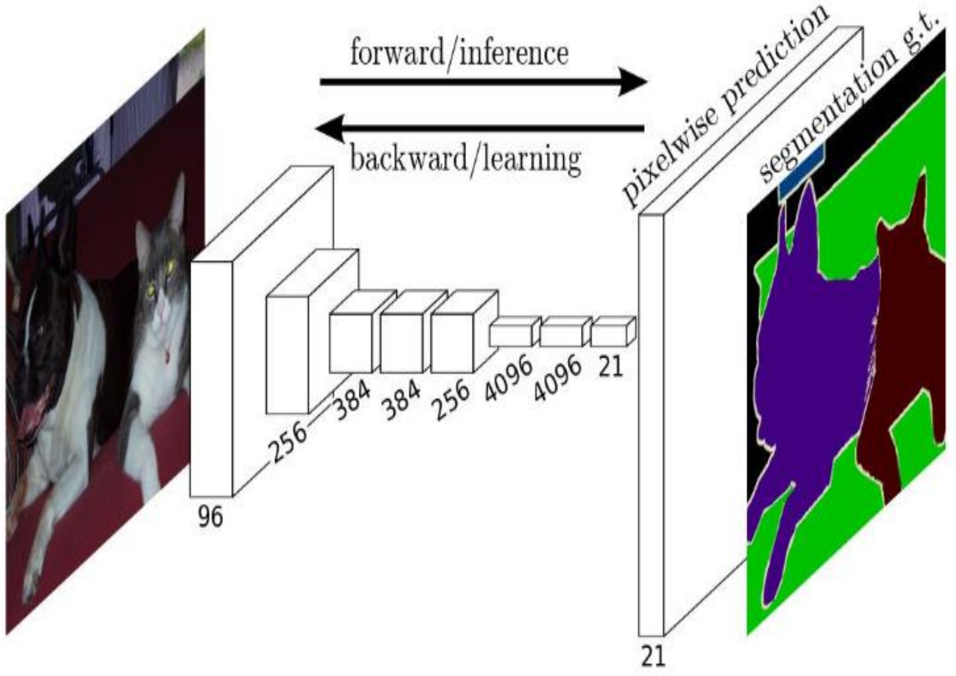


[SUP2024]

Semantic Image Segmentation

Semantic image segmentation

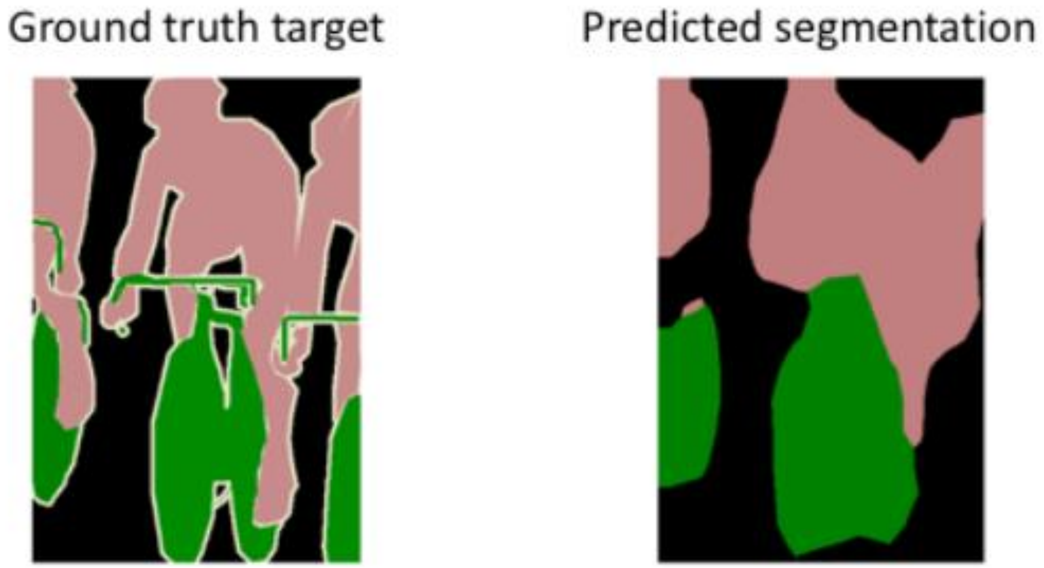
- Fully convolutional network for semantic segmentation.
- Usually, the final feature map is upsampled to match the resolution of the input image.



End-to-end CNN training for semantic image segmentation [LON2015].

Semantic Image Segmentation

- As the CNN model radically reduces the resolution of the input image, it fails to produce fine-grained segmentations.



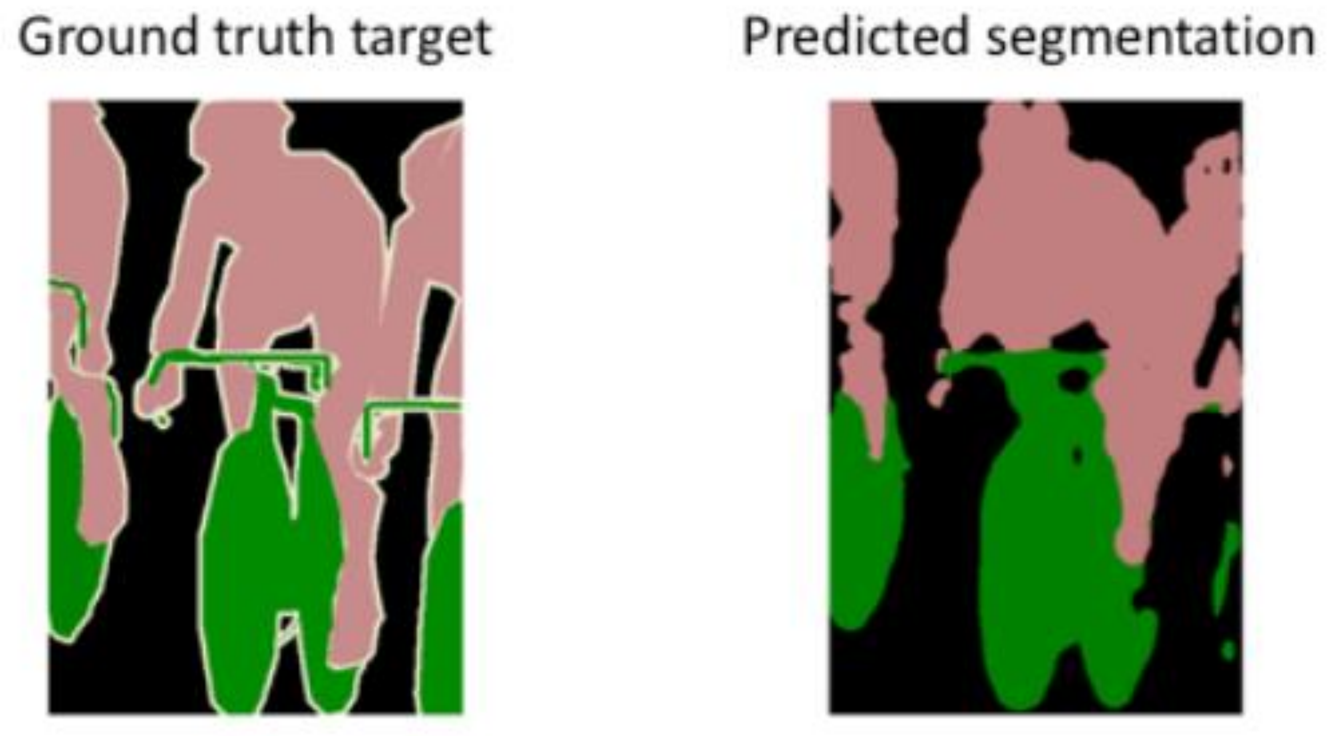
Coarse image segmentation [LON2015].

Semantic Image Segmentation

- To address this problem, ***skip network connections*** are added in fully convolutional network that combine the final prediction layer with previous fine-grained layers.

- Combining fine layers and coarse layers allows the model to make local predictions that respect global structure.

Semantic Image Segmentation



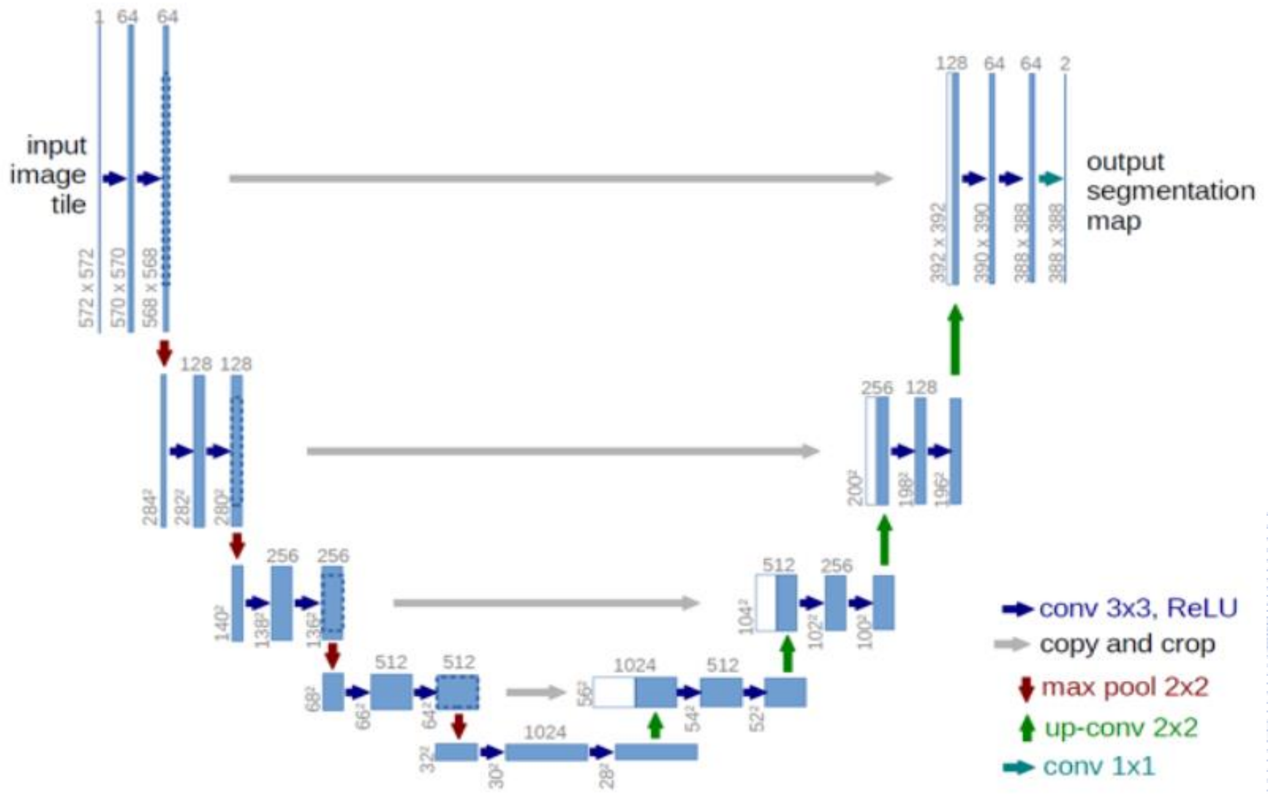
Improved segmentation results with skip connections [LON2015].

Deep Semantic Image segmentation

U-Net architecture

- More advanced semantic segmentation network architectures have emerged.
- The capacity of the decoder was expanded by using a ***U-shaped network*** architecture (***U-Net***).
- Consists of a ***contracting path*** to capture context and a ***symmetric expanding path*** that enables precise localization.

Semantic Image Segmentation



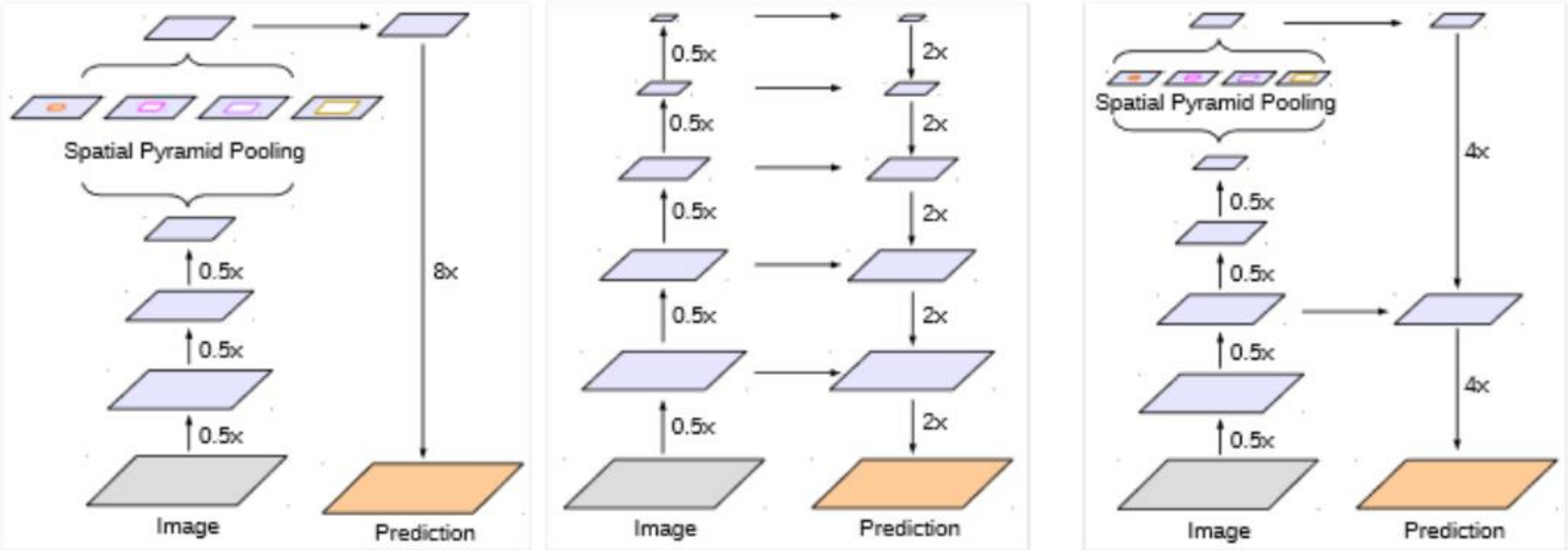
U-Net network architecture [RON 2015].

Semantic Image Segmentation

Spatial Pyramid Pooling

- Semantic image segmentation performance was also increased by combining the advantages of a **Spatial Pyramid Pooling (SPP)** [ZHA2017] module and the encoder-decoder architecture.
- SPP module can encode multi-scale contextual information, by probing the incoming features with filters or pooling operations at multiple rates and multiple effective fields-of-view.

Semantic Image Segmentation



Spatial Pyramid Pooling. Encoder-Decoder. Combined approach [CHE2018].

Semantic Image Segmentation

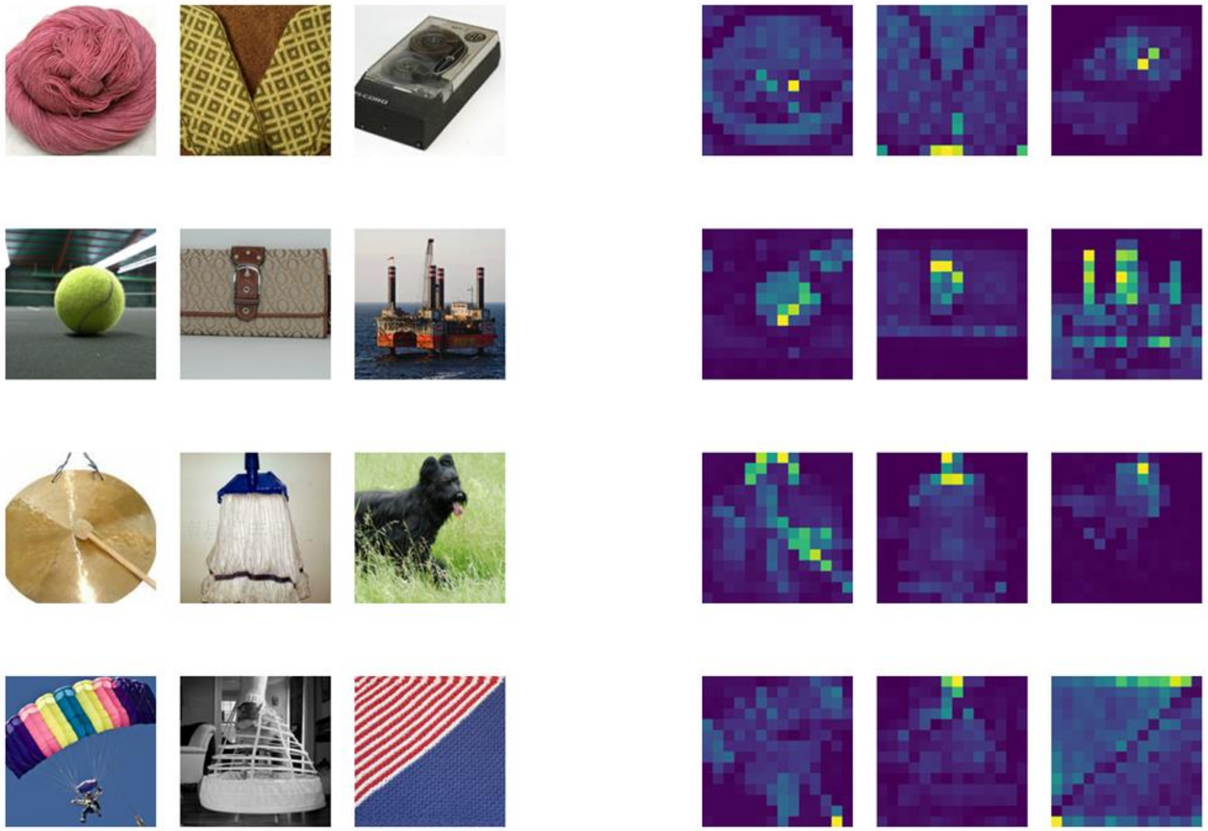
Vision Transformer (ViT) [DOS2020].

- Implementation of transformer architecture in Computer Vision.
- A pure transformer applied directly to sequences of image patches works exceptionally well on image classification, segmentation and object detection tasks.
- Uses self-attention mechanisms to process images

Semantic Image Segmentation

Self-Attention

- A mechanism which computes a weighted sum of the input image data vectors (block pixels or their CNN feaures).
- The weight are computed based on the similarity between the input data vectors.



Wildfire Image Analysis

- Computer Vision Tasks
- Image Region Segmentation
- Semantic Image Segmentation
- **Fire Detection**
- Fire Segmentation

Fire Detection Evaluation

3 TP – 0 FP

1TP – 2 FP

0 TP – 3 FP

BASED ON IOU

Raw image

Possible ground-truth annotation strategies

Predictions

Fire Detection Evaluation

Challenges with mAP for Fire Detection

- Unlike most objects, fires consist of "children" objects (flames) that belong to the same class as the "parent" object (fire), making it uncertain how many bounding boxes are needed for accurate representation.

Limitations of mAP: Inconsistent annotation styles for fire objects can misalign with predicted bounding boxes, leading to mAP scores that do not accurately reflect model performance.

Fire Detection Evaluation

Image-level mean Average Precision (ImAP) [TZI2023]:

- It evaluates fire detection models based on their ability to predict bounding boxes for the entire image rather than individual boxes.

- ImAP demonstrates greater suitability than mAP for evaluating object detectors in fire detection tasks, addressing the unique properties of fire entities.

Fire Detection Evaluation

ImAP utilize Image Level Intersection Over Union (**ImIOU**) instead of **IOU** in order to evaluate fire detection in the entire image

ImIOU: Intersection over Union between all predictions and all ground truth bounding boxes of the same image:

$$ImIoU\,(\mathcal{P},\mathcal{G}) = \frac{\left|\left(\bigcup_{i=1}^{|\mathcal{P}|} P_i\right) \cap \left(\bigcup_{i=1}^{|\mathcal{G}|} G_i\right)\right|}{\left|\left(\bigcup_{i=1}^{|\mathcal{P}|} P_i\right) \cup \left(\bigcup_{i=1}^{|\mathcal{G}|} G_i\right)\right|}.$$

- $\mathcal{G} = \{G_i\}_{i=1,\dots,N}$: bounding box ground truths.
- $\mathcal{P} = \{P_i\}_{i=1,\dots,M}$: predictions.

Fire Detection Evaluation

BASED ON IOU 3 TP – 0 FP 1TP – 2 FP 0 TP – 3 FP

BASED ON IMIOU IMIOU=0.78 -> TP IMIOU=0.6 -> TP IMIOU=0.51 -> TP

Artificial Intelligence & Information Analysis Lab

Fire Detection Localization Loss

Regression losses for the localization task of the object detection:

- L1
- IoU-based.

The state of the art object detection model RTDETR [ZHA2024] combines the L_1 loss with L_{IoU} to improve the detection of object of interest:

$$L_{loc} = \lambda_1 \cdot L_1 + \lambda_{IoU} \cdot L_{IoU}\,.$$

Fire Detection Localization Loss

L_1 and L_{IoU} formulae for an image with N bounding boxes:

$$L_1(\mathcal{P},\mathcal{G}) = \frac{1}{N}\sum_{i=1}^{N}\left(\sum_{j\in\{x,y,w,h\}}\left|P_{i,j} - G_{i,j}\right|\right)$$

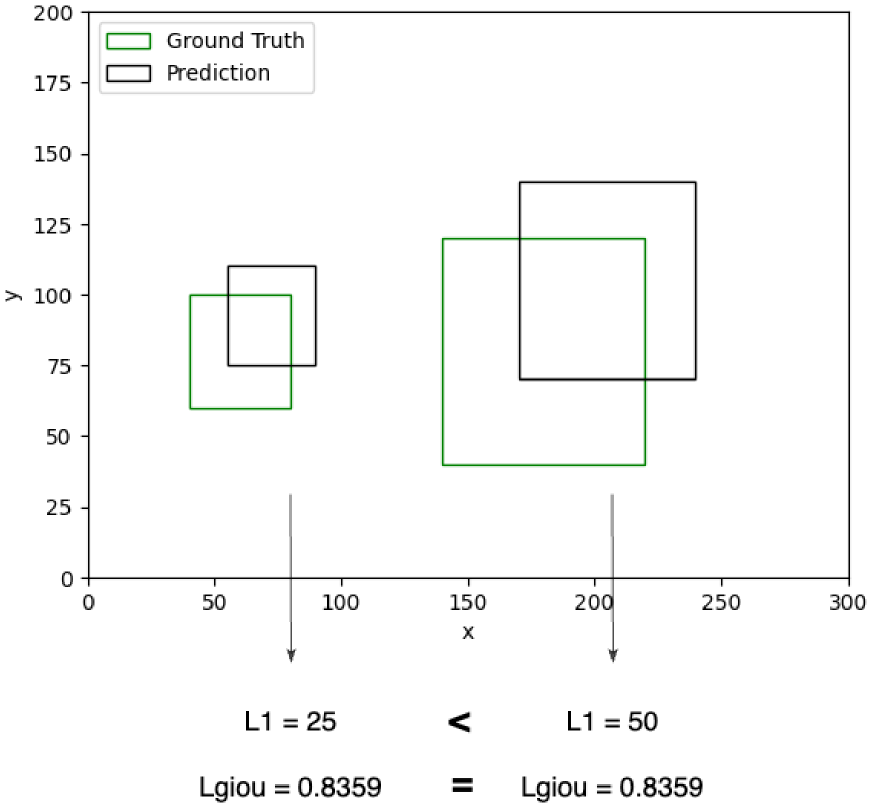
$$L_{IoU}(\mathcal{P},\mathcal{G}) = \frac{1}{N}\sum_{i=1}^{N}\big(1 - IoU(P_i, G_i)\big)$$

- $\mathcal{G} = \{G_i\}_{i=1,\ldots,N}$: bounding box ground truths
- $\mathcal{P} = \{P_i\}_{i=1,\ldots,N}$: predictions.

Fire Detection Localization Loss

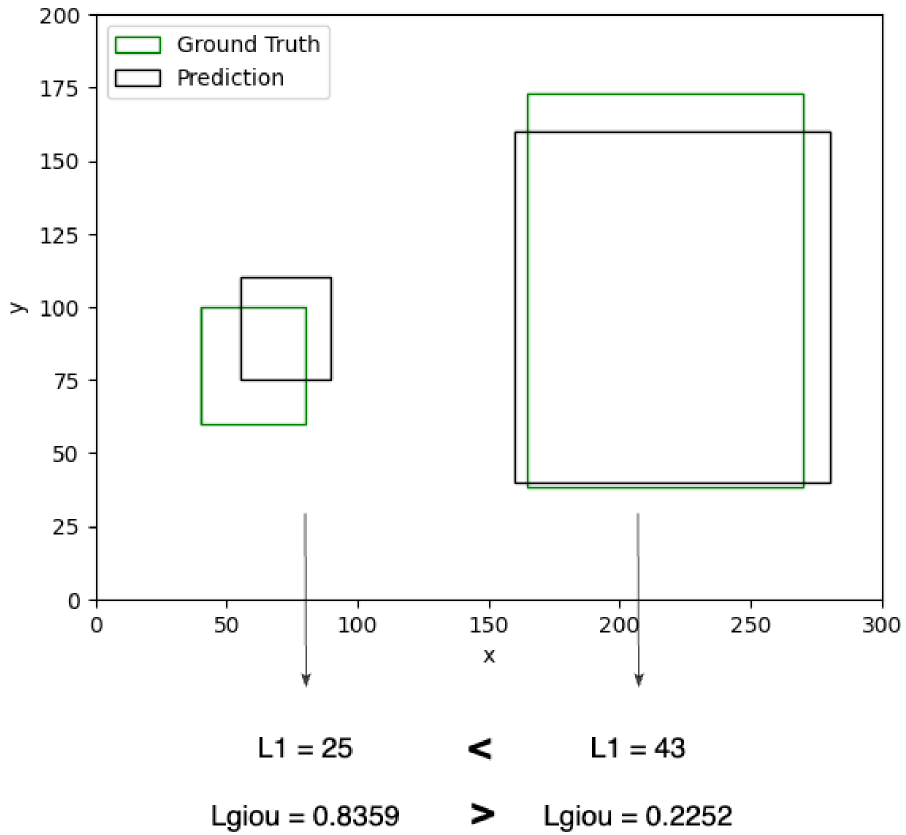
In fire detection, there are many scenarios where small and large flame regions can appear in the same image.

- In this case the larger prediction boxes have larger error with their corresponding due to the L_1 loss.
- L_{IoU} loss is invariant of the bounding box sizes.



Fire Detection Localization Loss

In many cases, there may be disagreements between the L_1 and IoU losses, which can affect training, as the two losses may not share the same local minimums.



Artificial Intelligence & Information Analysis Lab

Fire Detection Localization Loss

Size balanced loss L_{SB} is an L_1 loss variant adding a weighting mechanism on the L1 loss based on the ground-truth bounding box size.

$$L_{SB}(\mathcal{P},\mathcal{G}) = \sum_{i=1}^{N} W_i \left(\sum_{j\in\{x,y,w,h\}} \left|P_{i,j} - G_{i,j}\right| \right).$$

Experiments on fire detection datasets demonstrate +2% improvement over mAP and ImAP using L_{SB}.

Wildfire Image Analysis

- Computer Vision Tasks
- Image Region Segmentation
- Semantic Image Segmentation
- Fire Detection
- **Fire Segmentation**

Fire Segmentation

Semantic segmentation Deep Neural Networks architectures were trained on the FLAME dataset:

- BiSeNet (backbones: ResNet18, ResNet101) [CYO2018]
- I2I-CNN (backbones: ResNet18) [PAP2021]
- PIDNet (backbone: ResNet18) [JXU2023]

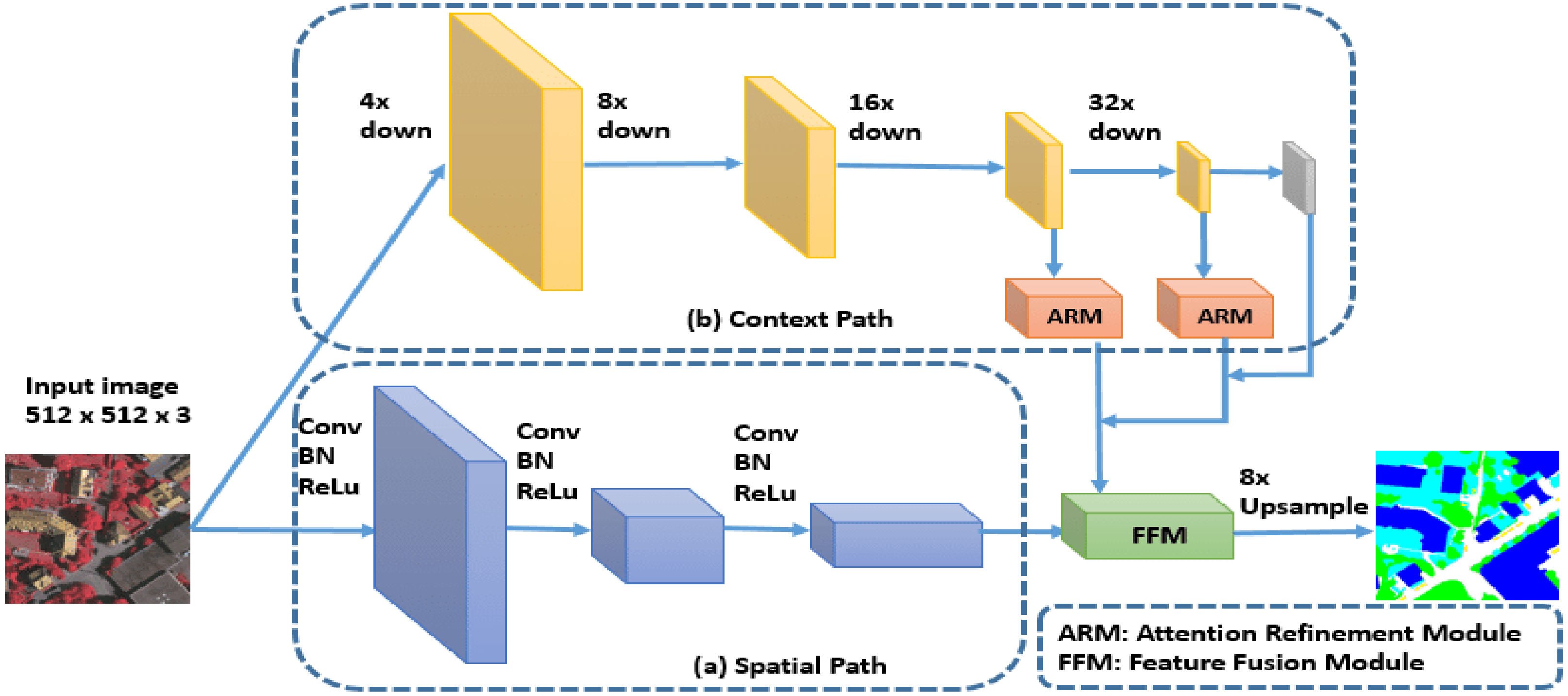
FLAME dataset [SHA2021]

Fire Segmentation

BiseNet architecture

- **Two-Stream Network:** Combines spatial and contextual information for high accuracy in segmentation.
- **Efficient and Fast:** Designed for real-time performance with lightweight structure, ideal for real-time applications like fire detection.
- **Context Path:** Captures large-scale features for better scene understanding.
- **Spatial Path:** Retains high-resolution details for precise boundary segmentation.

Fire Segmentation



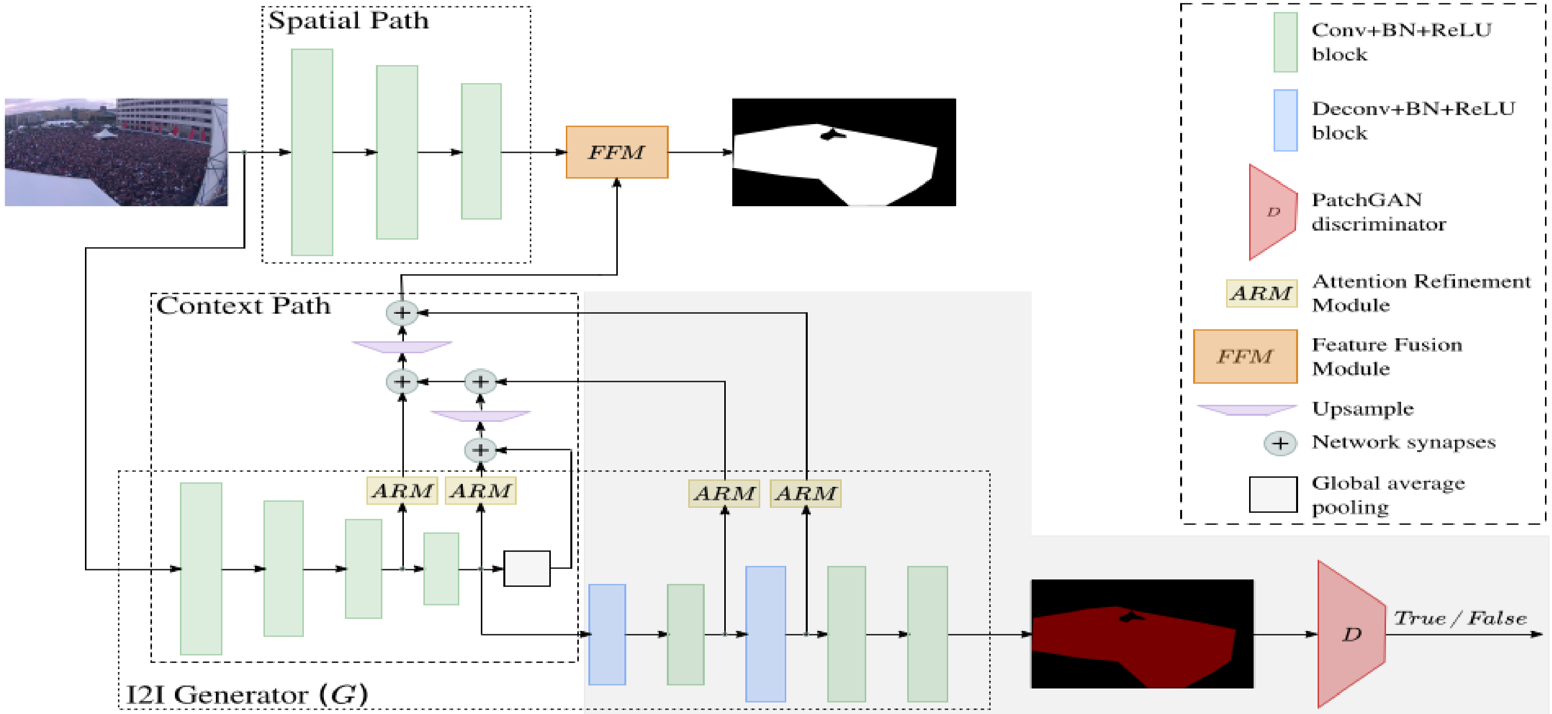
BiSeNet architecture [CYO2018]

Fire Segmentation

I2I-CNN architecture

- **Dual-Branch Design:** Adds an auxiliary neural branch to the BiseNet branch for enhanced semantic accuracy without slowing down execution.
- **GAN-Based Auxiliary Branch:** Trained using a Generative Adversarial Network (GAN) to generate RGB-like segmentation maps, capturing additional semantic information.
- **Adversarial Training with Discriminator:** The auxiliary branch learns through adversarial loss, where a Discriminator validates its output for improved semantic feature extraction.
- **Lightweight and Fast:** This network has the same inference speed as Bisenet.

Fire Segmentation



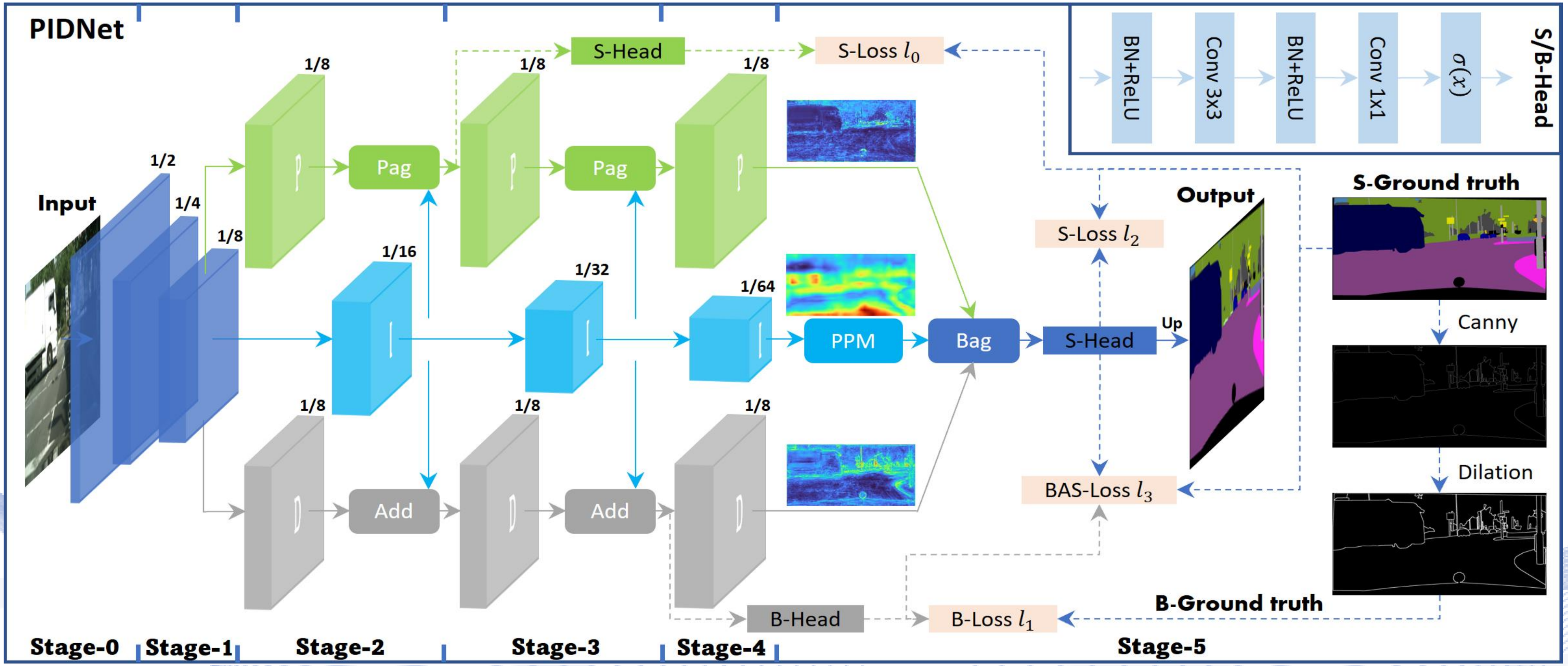
I2I-CNN architecture [PAP2021].

Fire Segmentation

PIDNet architecture

- **Triple-Branch Design:** Uses three branches—Proportional (P), integral (I), and derivative (D)—to balance accuracy and efficiency.
- **Real-Time Performance:** Optimized for real-time applications, making it suitable for tasks like fire detection in edge environments.
- **High Precision in Edge Detection:** The Detail branch captures fine edges, crucial for accurately outlining objects in segmentation.
- **Competitive Accuracy:** Delivers performance close to more complex models, but with much faster inference speeds.

Fire Segmentation



PIDNet architecture [JXU2023].

Fire Segmentation

DNN input has the 3 following forms:

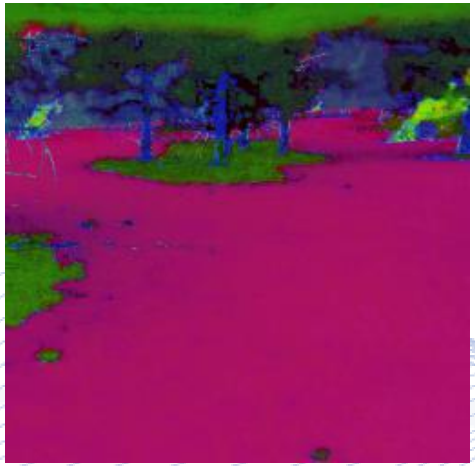
- RGB (3 channels)
- RGB+HSV (6 channels)
- RGBS (4 channels)

- S: processed saturation channel of HSV image transform and is used to suggest potential fire regions. This mask is then concatenated with the RGB image to form a new 4-channel input.

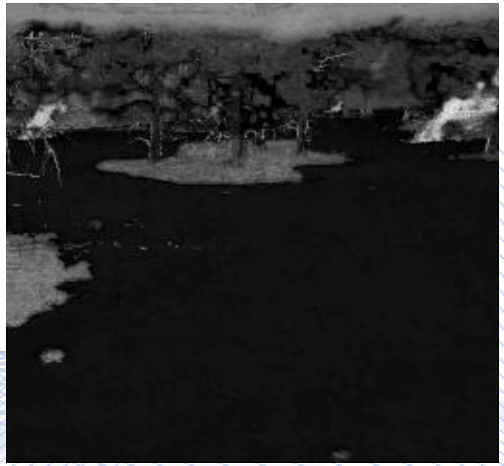
Fire Segmentation

Process of creating the S channel (visualization)

RGB input image



HSV transformation of RGB image



Saturation channel (S)

Thresholding of channel S

Fire Segmentation

BiSeNet, I2I-CNN and PID-Net were evaluated using mIoU and novel fire region segmentation measures:

- Fire region number difference D_N,
- Average fire region area difference D_A,
- Spatial dispersion D_A of fire regions.

These measures distill meaningful information about the extent of a forest fire and target the explainability to the end-user.

- The experiments on the FLAME dataset demonstrate that the PIDNet with RGB+S as input achieve the best mIoU among all the other configurations

Artificial Intelligence & Information Analysis Lab

Fire Segmentation

Fire region number difference D_N:

$$= \frac{1}{|I|} \sum_{i=1}^{|I|} |N_i - N_i'| .$$

Average fire region area difference D_A (in pixels):

$$D_A = \frac{1}{|I|} \sum_{i=1}^{|I|} |A_i - A_i'|.$$

Fire Segmentation

Spatial dispersion D_A of fire regions:

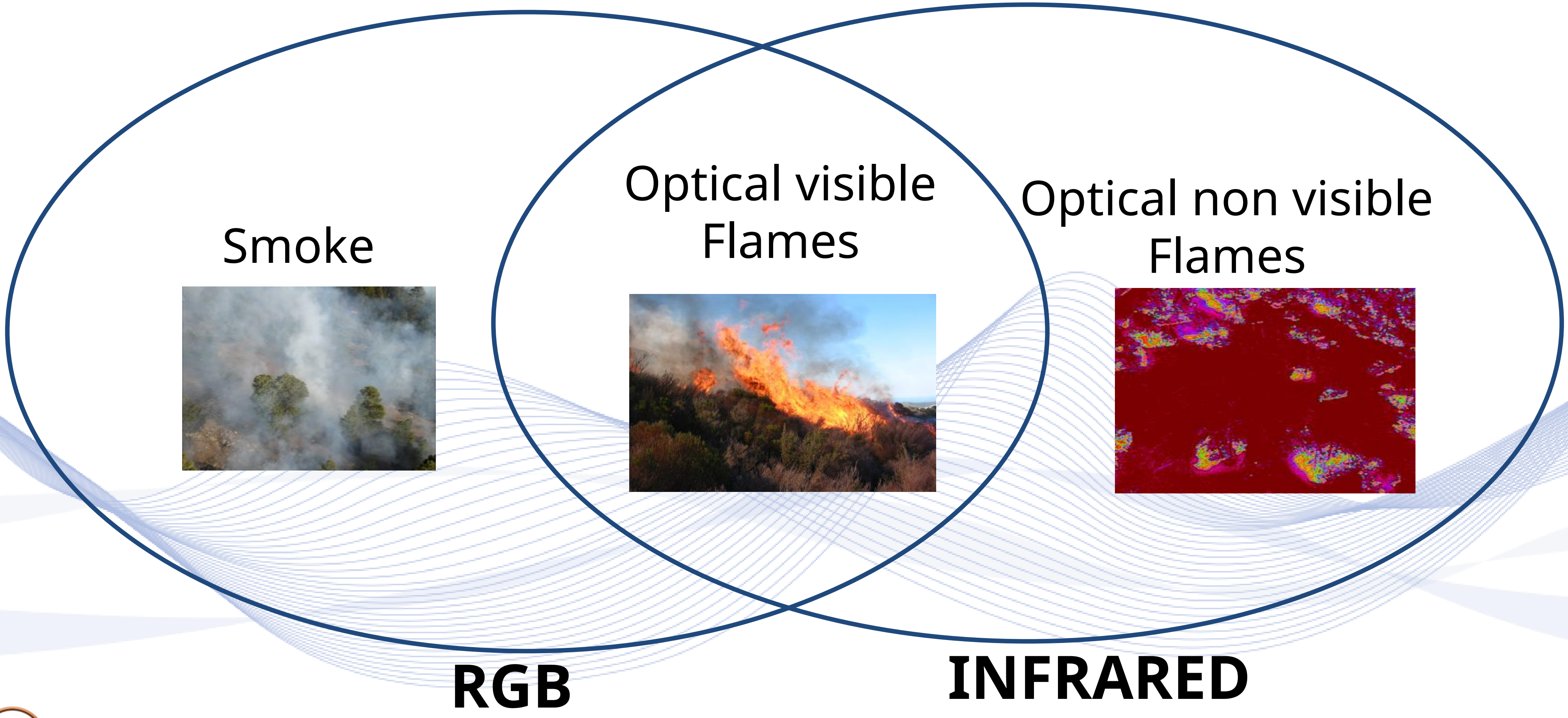
$$D_A = \frac{1}{|I|} \sum_{i=1}^{|I|} |s_i - s_i' |,$$

$$s_i = \frac{1}{|N|} \frac{1}{|N|-1} \sum_{j=1}^{|N|} \sum_{k=1, j \neq k}^{|N|} \|\mathbf{p}_j - \mathbf{p}_k \|_2$$

$\mathbf{p}_i$: centers of fire sources

RGB/IR Fire Segmentation

A Venn Diagram of RGB and IR Capabilities



RGB/IR Fire Segmentation

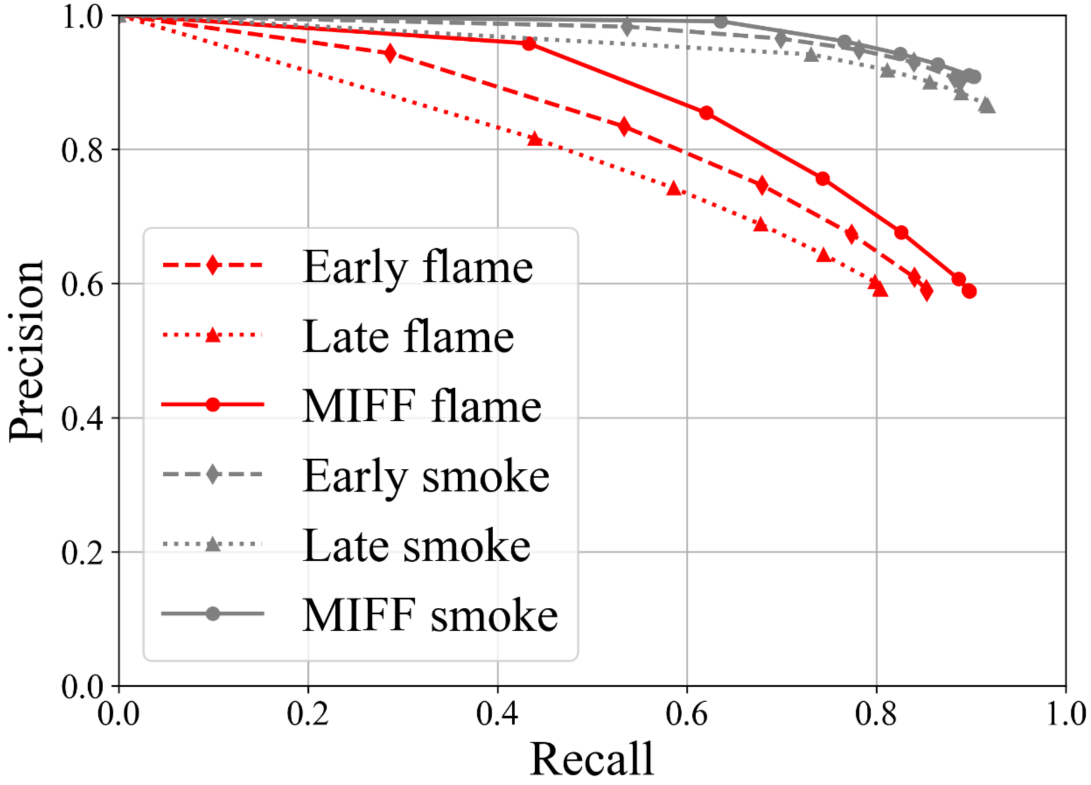
IR and RGB image fusion

Early fusion: Concatenate the three RGB channels with the IR image to create a unified 4D input for the DNN.

Intermediate fusion: Feed the RGB and IR images separately into their respective DNNs, concatenate their intermediate feature maps, and then pass the aggregated map through a common network for further processing.

Late fusion: Process the RGB and IR images separately through their respective DNNs, then concatenate the segmentation results from both networks to obtain the final output.

RGB/IR Fire Segmentation



Flame and smoke segmentation performance.

RGB/IR Fire Segmentation

RGB/IR fire segmentation results.

Artificial Intelligence &
Information Analysis Lab

Wildfire Image Analysis

- Computer Vision
- Classical image segmentation techniques
- Deep semantic image segmentation
- Fire detection
- Fire segmentation
  - RGB/IR Fire segmentation
  - **Prompted fire segmentation**

Prompted Fire Segmentation

An enormous amount of labeled data to train deep learning models to detect objects of interest. Annotating datasets is a time-consuming and expensive task.

Raw images and the corresponding labels.

Prompted Fire Segmentation

Unsupervised semantic segmentation architectures do not rely on labeled datasets.

- Without prior information about the objects of interest, they have poor segmentation accuracy.

Dataset [TOU2017]

Unsupervised segmentation results.

Prompted Fire Segmentation

Raw Images

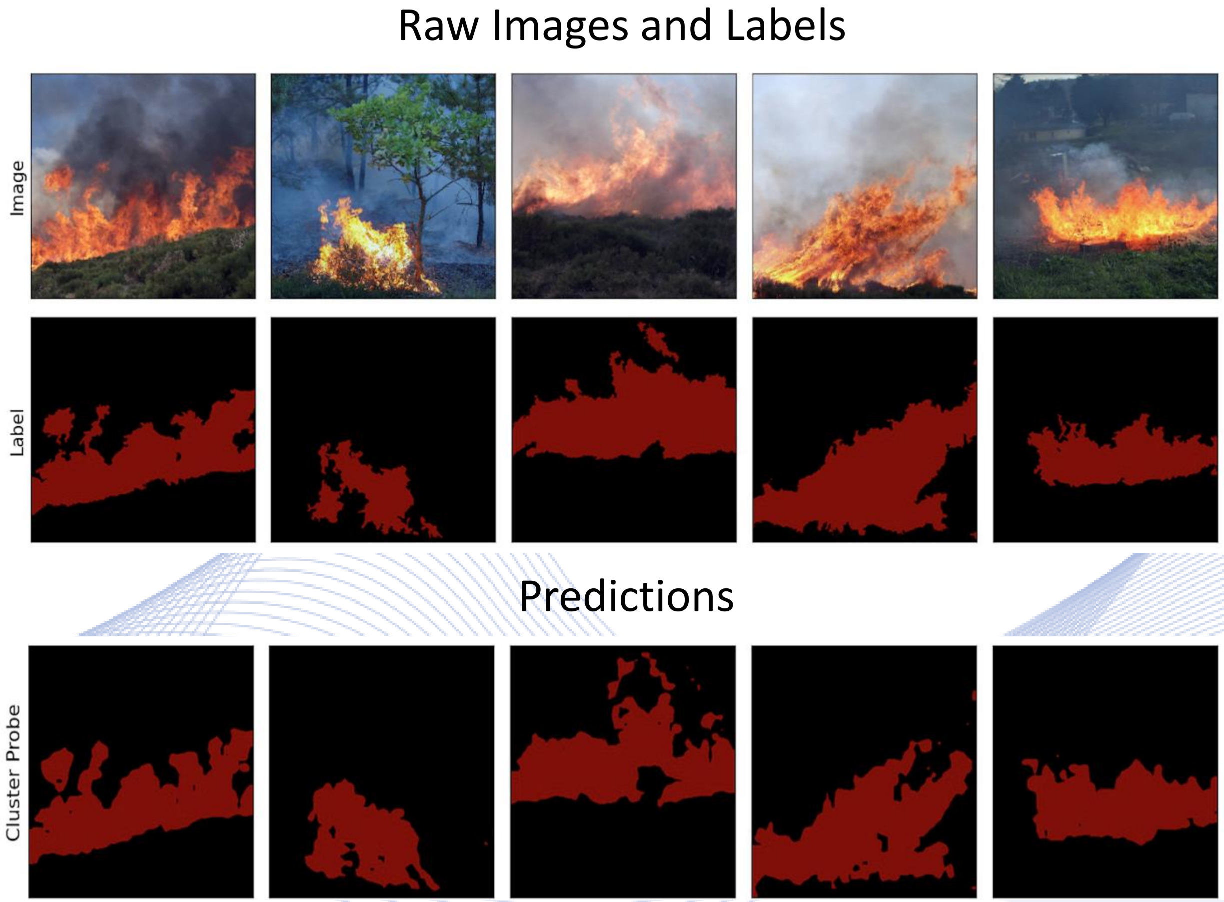
We select a single image from the dataset and specify only one point where our object of interest is located.

1. Combine the raw images with the signal from the annotated point.
2. Push fire representations closer together in the feature space
3. Create a cluster head that separates fire from the background

Artificial Intelligence & Information Analysis Lab

Prompted Fire Segmentation

- Unsupervised segmentation performance: 40 % mIoU.
- Prompted segmentation performance: 80 % mIoU.
- Prompting achieves a 40% increase in mIoU using only a single fire point.



References

[PIT2000] I. Pitas, “Digital image processing algorithms and applications”, Wiley 2000.
[LON2015] J. Long, E. Shelhamer, T. Darrell, “Fully convolutional networks for semantic segmentation”, In Proceedings of the IEEE conference on computer vision and pattern recognition, 2015.
[RON 2015] O. Ronneberger, P. Fischer, T. Brox, “U-net: Convolutional networks for biomedical image segmentation”, In Proceedings of the International Conference on Medical image computing and computer-assisted intervention, Springer, Cham, 2015.
[CHE2014] L.C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, “Semantic image segmentation with deep convolutional nets and fully connected crfs”, arXiv preprint arXiv:1412.7062, 2014.
[TOR2014] O.A.J del Toro, O. Goksel, B.Menze, H. Muller, G. Langs, “VISCERAL–VISual Concept Extraction challenge in RAdioLogy: ISBI 2014 challenge organization”, In Proceedings of the VISCERAL Challenge at ISBI, 2014.
[YU2018] C. Yu, J. Wang, C. Peng, C. Gao, G. Yu, N. Sang, “ Bisenet: Bilateral segmentation network for real-time semantic segmentation”, In Proceedings of the European conference on computer vision (ECCV), 2018.
[ZHU2019] J. Zhuang, J. Yang, L. Gu, N. Dvornek, “ShelfNet for Fast Semantic Segmentation”, In Proceedings of the IEEE International Conference on Computer Vision Workshops, 2019.
[CHE2017] L.C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, “Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs”, IEEE transactions on pattern analysis and machine intelligence (PAMI), 2017

References

[EVE2011] M. Everingham, John Winn, “The PASCAL visual object classes challenge 2012 (VOC2012) development kit”, Pattern Analysis, Statistical Modelling and Computational Learning, Tech. Rep, 2011.
[COR2016] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, B. Schiele, “The cityscapes dataset for semantic urban scene understanding”, In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), 2016.
[CHE2018] L.C. Chen, Y. Zhu, G. Papandreou, F. Schroff, A. Hartwig, “Encoder-decoder with atrous separable convolution for semantic image segmentation”, In Proceedings of the European conference on computer vision (ECCV). 2018.
[MOR2018] G. Morales, G. Kemper, G. Sevillano, D. Arteaga, I. Ortega, J. Telles, “Automatic segmentation of Mauritia flexuosa in unmanned aerial vehicle (UAV) imagery using deep learning.” Forests, 2018.
[YUA2019] Y. Yuan, X. Chen, J. Wang, “Object-contextual representations for semantic segmentation.” arXiv preprint arXiv:1909.11065, 2019.
[TAK2019] T. Takikawa, D. Acuna, V. Jampani, S. Fidler, “Gated-SCNN: Gated shape CNNs for semantic segmentation.” In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2019.
[HUA2019] Z. Huang, X. Wang, L. Huang, C. Huang, Y. Wei, W.Liu, “CCNet: Criss-cross attention for semantic segmentation.” In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2019.
[CHE2018] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, “Encoder-decoder with atrous separable convolution for semantic image segmentation.” In Proceedings of the European Conference on Computer Vision (ECCV), 2018.

References

[MOU2016] A. Mousavian, H. Pirsiavash, J. Kosecḱa, “Joint semantic segmentation and depth estimation with deep convolutional networks.”, In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV). IEEE, 2016.
[LIU2018] J. Liu, Y. Wang, Y. Li, J. Fu, J. Li, H. Lu, “Collaborative deconvolutional neural networks for joint depth estimation and semantic segmentation.”, IEEE transactions on neural networks and learning systems, 2018.
[ALA2020] M. Aladem, S.A. Rawashdeh. “A single-stream segmentation and depth prediction CNN for autonomous driving.”, IEEE Intelligent Systems, 2020.
[CHE2019] P.Y. Chen, A. H Liu, Y.C. Liu, Y.C.F Wang, “Towards scene understanding: Unsupervised monocular depth estimation with semantic-aware representation.”, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019.
[QI2017] X. Qi, R. Liao, J. Jia, S. Fidler, R. Urtasun, “3D graph neural networks for RGBD semantic segmentation.”, In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017.
[KEN2018] A. Kendall, Y. Gal, R. Cipolla, “Multi-task learning using uncertainty to weigh losses for scene geometry and semantics.”, In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), 2018.
[ZHA2017] H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, “Pyramid scene parsing network.”, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
[APOLLO] http://apolloscape.auto/

References

[DOS2020] A. DOSOVITSKIY, An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929. 2020.
[KRI2023] Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, Xiao T, Whitehead S, Berg AC, Lo WY, Dollár P. Segment anything. InProceedings of the IEEE/CVF International Conference on Computer Vision 2023 (pp. 4015-4026).
[SHA2021] Shamsoshoara, Alireza, et al. "Aerial imagery pile burn detection using deep learning: The FLAME dataset." *Computer Networks* 193 (2021): 108001.
[CYU2018] C. Yu, J. Wang, C. Peng, C. Gao, G. Yu, and N. Sang, "Bisenet: Bilateral segmentation network for real-time semantic segmentation," in European Conference on Computer Vision. Springer, 2018, pp. 334–349.
[PAP2021] C. Papaioannidis, I. Mademlis, and I. Pitas, "Autonomous uav safety by visual human crowd detection using multi-task deep neural networks," in 2021 IEEE International Conference on Robotics and Automation (ICRA), 2021, pp. 11 074–11 080.
[JXU2023] J. Xu, Z. Xiong, and S. P. Bhattacharyya, "Pidnet: A real-time semantic segmentation network inspired by pid controllers," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 19 529–19 539.
[BIT2024] *CNN Architecture - Detailed Explanation*. (n.d.). InterviewBit. https://www.interviewbit.com/blog/cnn-architecture/
[SUP2024]SuperAnnotate AI Inc. (n.d.). *Semantic segmentation: Complete guide [Updated 2024] | SuperAnnotate*. SuperAnnotate. https://www.superannotate.com/blog/guide-to-semantic-segmentation#:~:text=Semantic%20segmentation%20is%20simply%20the,Image%20Source

Artificial Intelligence & Information Analysis Lab

References

[TZI2023] M. D. Tzimas, C. Papaioannidis, V. Mygdalis, and I. Pitas, "Evaluating Deep Neural Network-based Fire Detection for Natural Disaster Management," in 2023 IEEE/ACM 16th International Conference on Utility and Cloud Computing (UCC '23), Taormina (Messina), Italy, Dec. 2023.

[ZHA2024] Zhao, Yian, et al. "Detrs beat yolos on real-time object detection." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024.

[THA2023]Thakur, N. (2023, June 2). A detailed introduction to Two Stage Object Detectors. *Medium*. https://namrata-thakur893.medium.com/a-detailed-introduction-to-two-stage-object-detectors-d4ba0c06b14e

[KUK2023] Kukil, & Kukil. (2023, August 4). *Intersection over union IOU in object detection segmentation*. LearnOpenCV – Learn OpenCV, PyTorch, Keras, Tensorflow With Code, & Tutorials. https://learnopencv.com/intersection-over-union-iou-in-object-detection-and-segmentation/

[TOU2017]Toulouse, Tom, et al. "Computer vision for wildfire research: An evolving image dataset for processing and analysis." *Fire Safety Journal* 92 (2017): 188-194.

Q & A

Thank you very much for your attention!

More material in
http://icarus.csd.auth.gr/cvml-web-lecture-series/

Contact: Prof. I. Pitas
pitas@csd.auth.gr