Training Course on Object Detection and Segmentation with Deep Learning

Professional Development Training Course Outline

Category	Data Science	**Duration**	10 Days
Base Fee	$3,000 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

In the rapidly evolving landscape of Artificial Intelligence and Computer Vision, **Object Detection** and **Image Segmentation** stand as cornerstone technologies driving innovation across diverse industries. This intensive training course delves into the cutting-edge realm of deep learning architectures, specifically focusing on **YOLO (You Only Look Once)**, **Mask R-CNN**, and advanced **Segmentation Networks**. Participants will gain comprehensive theoretical understanding and practical expertise in implementing these powerful models for real-world applications, from **autonomous vehicles** and **medical imaging** to **retail analytics** and **industrial automation**.

Training Course on Object Detection & Segmentation with Deep Learning is designed to equip professionals with the skills to leverage **state-of-the-art deep learning models** for precise object localization and pixel-level understanding. Through hands-on exercises, real-world case studies, and expert-led instruction, attendees will master techniques for **data annotation**, model training, optimization, and deployment. The curriculum emphasizes **real-time processing**, **computational efficiency**, and tackling challenges like **small object detection** and **class imbalance**, ensuring participants are prepared for the demands of modern computer vision projects.

Programme Curriculum

Training Course on Object Detection & Segmentation with Deep Learning

Introduction

In the rapidly evolving landscape of Artificial Intelligence and Computer Vision, **Object Detection** and **Image Segmentation** stand as cornerstone technologies driving innovation across diverse industries. This intensive training course delves into the cutting-edge realm of deep learning architectures, specifically focusing on **YOLO (You Only Look Once)**, **Mask R-CNN**, and advanced **Segmentation Networks**. Participants will gain comprehensive theoretical understanding and practical expertise in implementing these powerful models for real-world applications, from

autonomous vehicles and **medical imaging** to **retail analytics** and **industrial automation**.

Training Course on Object Detection & Segmentation with Deep Learning is designed to equip professionals with the skills to leverage **state-of-the-art deep learning models** for precise object localization and pixel-level understanding. Through hands-on exercises, real-world case studies, and expert-led instruction, attendees will master techniques for **data annotation**, model training, optimization, and deployment. The curriculum emphasizes **real-time processing**, **computational efficiency**, and tackling challenges like **small object detection** and **class imbalance**, ensuring participants are prepared for the demands of modern computer vision projects.

Course Duration

10 days

Course Objectives

1. Master the foundational concepts of Deep Learning and Convolutional Neural Networks (CNNs) for computer vision.
2. Understand the principles and architectures of Object Detection algorithms, including single-stage (YOLO) and two-stage (Faster R-CNN, Mask R-CNN) detectors.
3. Implement and fine-tune YOLO models (YOLOv8, YOLOv11) for high-performance, real-time object detection.
4. Grasp the intricacies of Instance Segmentation using Mask R-CNN for pixel-accurate object mask generation.
5. Explore various Semantic Segmentation networks (e.g., U-Net, FCN) for holistic scene understanding.
6. Learn effective data annotation strategies and techniques for creating high-quality datasets for training.
7. Apply transfer learning and fine-tuning methodologies to adapt pre-trained models to custom datasets.
8. Understand evaluation metrics for object detection and segmentation (mAP, IoU, F1-score) and interpret model performance.
9. Develop robust strategies for model optimization, including quantization, pruning, and edge deployment.
10. Tackle common challenges in deep learning for computer vision, such as class imbalance, overfitting, and small object detection.
11. Gain practical experience with popular deep learning frameworks like PyTorch or TensorFlow.
12. Explore advanced topics like attention mechanisms, transformer-based models (Vision Transformers), and self-supervised learning in the context of object detection and segmentation.
13. Prepare for real-world AI deployment and MLOps best practices for computer vision systems.

Organizational Benefits

- Implement intelligent systems for tasks like automated quality control, inventory management, and surveillance.
- Leverage deep learning models for faster and more precise object identification and analysis, reducing manual errors and processing time.
- Stay at the forefront of AI innovation by adopting cutting-edge computer vision technologies for product development and operational enhancement.
- Automate labor-intensive visual inspection processes, leading to significant savings in operational costs.

- Develop novel AI-powered products and services based on advanced object detection and segmentation capabilities.
- Extract actionable insights from visual data for better business intelligence and strategic planning.
- Empower technical teams with specialized expertise in a high-demand field, fostering internal innovation and reducing reliance on external consultants.

Target Audience

1. AI/ML Engineers.
2. Data Scientists.
3. Computer Vision Researchers
4. Software Developers.
5. Robotics Engineers.
6. Quality Control & Inspection Professionals
7. Medical Imaging Specialists
8. Graduate Students.

Course Outline

Module 1: Introduction to Deep Learning for Computer Vision

- Fundamentals of Deep Learning: Neural networks, activation functions, loss functions, optimization algorithms.
- Convolutional Neural Networks (CNNs): Architecture, convolutional layers, pooling layers, fully connected layers.
- Image Classification Review: Understanding how CNNs classify images.
- Introduction to Computer Vision Tasks: Classification vs. Detection vs. Segmentation.
- Setting up Your Deep Learning Environment: Python, PyTorch/TensorFlow, CUDA.
- **Case Study**: Using a pre-trained ResNet for image classification on a custom dataset.

Module 2: Object Detection Fundamentals

- Defining Object Detection: Bounding boxes, confidence scores, class labels.
- Traditional Object Detection Approaches: HOG, DPM (brief overview for context).
- Evolution of Deep Learning Detectors: Two-stage vs. One-stage detectors.
- Evaluation Metrics: Intersection over Union (IoU), Mean Average Precision (mAP).
- Challenges in Object Detection: Varying object scales, occlusions, cluttered scenes.
- **Case Study**: Analyzing common object detection errors in a surveillance system.

Module 3: Two-Stage Object Detectors: R-CNN Family

- R-CNN: Region Proposals, CNN Feature Extraction, SVM Classification, Bounding Box Regression.
- Fast R-CNN: RoI Pooling, shared computations.
- Faster R-CNN: Region Proposal Network (RPN) for end-to-end training.
- Anchor Boxes: Understanding their role in object proposal generation.
- Non-Maximum Suppression (NMS): Filtering overlapping bounding boxes.
- **Case Study**: Implementing Faster R-CNN for pedestrian detection in urban environments.

Module 4: Single-Stage Object Detectors: YOLO (You Only Look Once)

- YOLO Philosophy: Unified detection, real-time performance.

- YOLO Architecture (v1-v3): Grid-based detection, predicting bounding boxes and class probabilities directly.
- Loss Function in YOLO: Coordinating, confidence, and classification losses.
- Advantages and Disadvantages of YOLO: Speed vs. accuracy tradeoffs.
- Optimizing YOLO for Specific Use Cases: Input resolution, anchor box tuning.
- **Case Study**: Real-time traffic sign detection using a custom YOLOv3 model.

Module 5: Advanced YOLO Architectures (YOLOv5, YOLOv7, YOLOv8, YOLOv11)

- Evolution of YOLO: Architectural improvements, data augmentation, training strategies.
- YOLOv5 & YOLOv7: Focus on efficiency, performance, and ease of use.
- YOLOv8 & YOLOv11: State-of-the-art performance, including instance segmentation capabilities.
- Training and Inference with YOLO: Practical implementation details.
- Benchmarking YOLO Models: Performance comparison and selection criteria.
- **Case Study**: Developing a high-speed defect detection system on a manufacturing assembly line using YOLOv8.

Module 6: Mask R-CNN for Instance Segmentation

- Introduction to Instance Segmentation: Pixel-level masks for each object instance.
- Mask R-CNN Architecture: Extending Faster R-CNN with a mask branch.
- RoI Align: Preserving spatial information for accurate masks.
- Training Mask R-CNN: Multi-task loss (classification, bounding box, mask).
- Applications of Mask R-CNN: Fine-grained object understanding.
- **Case Study**: Precise segmentation of organs in medical images for diagnostic purposes.

Module 7: Semantic Segmentation Networks

- Semantic Segmentation Defined: Pixel-wise classification of categories.
- Fully Convolutional Networks (FCN): End-to-end pixel prediction.
- U-Net Architecture: Encoder-decoder structure with skip connections for medical imaging.
- DeepLab Family: Atrous convolution, Atrous Spatial Pyramid Pooling (ASPP).
- Loss Functions for Segmentation: Cross-entropy, Dice loss.
- **Case Study**: Automated lane detection and road segmentation for autonomous driving.

Module 8: Data Preparation and Annotation for Object Detection & Segmentation

- Dataset Collection Strategies: Sourcing and curating image/video data.
- Image Annotation Tools: LabelImg, VGG Image Annotator (VIA), COCO Annotator.
- Bounding Box Annotation: Rectangular and rotated boxes.
- Polygon Annotation for Segmentation Masks: Pixel-level precision.
- Data Augmentation Techniques: Geometric and photometric transformations.
- **Case Study**: Building a custom dataset for vehicle detection in challenging weather conditions.

Module 9: Training Methodologies and Best Practices

- Transfer Learning: Leveraging pre-trained models for faster convergence.
- Fine-tuning Strategies: Freezing layers, learning rates, differential learning rates.
- Hyperparameter Tuning: Optimizing learning rate, batch size, epochs.
- Regularization Techniques: Dropout, L1/L2 regularization.
- Monitoring Training Progress: TensorBoard, validation metrics.

- **Case Study**: Fine-tuning a pre-trained YOLO model for detecting specific product categories in a retail environment.

Module 10: Model Optimization and Deployment

- Model Compression: Quantization, pruning, knowledge distillation.
- Inference Optimization: ONNX, TensorRT, OpenVINO.
- Edge AI Deployment: Running models on resource-constrained devices.
- Model Export and Integration: Integrating models into applications.
- Scalability Considerations: Handling large-scale deployments.
- **Case Study**: Deploying a real-time object detection model on an NVIDIA Jetson board for smart camera applications.

Module 11: Advanced Topics in Object Detection

- Feature Pyramid Networks (FPN): Multi-scale feature representation.
- Anchor-Free Detectors: Centernet, CornerNet.
- Transformer-based Detectors: DETR and its variants.
- Small Object Detection Techniques: Specialized architectures, data augmentation.
- Weakly Supervised and Semi-Supervised Detection: Reducing annotation effort.
- **Case Study**: Improving detection of small, distant vehicles in autonomous driving scenarios.

Module 12: Advanced Topics in Image Segmentation

- Panoptic Segmentation: Unifying semantic and instance segmentation.
- PointRend: High-quality mask prediction at object boundaries.
- Multi-modal Segmentation: Combining visual data with other sensor inputs.
- 3D Segmentation: Processing volumetric data for medical or industrial applications.
- Generative Models for Segmentation: Using GANs for data augmentation or refinement.
- **Case Study**: Automating the precise segmentation of building footprints from satellite imagery.

Module 13: Addressing Real-world Challenges

- Handling Class Imbalance: Weighted loss, focal loss, oversampling.
- Robustness to Occlusion and Lighting Variations: Data augmentation, advanced architectures.
- Real-time Performance vs. Accuracy Trade-offs: Model selection and optimization.
- Upcoming Scheduled Sessions

Start Date	End Date	Location	Price
23 Nov 2026	04 Dec 2026	Online	$2,000
23 Nov 2026	04 Dec 2026	Physical	$3,000
23 Nov 2026	04 Dec 2026	Physical	$6,000
23 Nov 2026	04 Dec 2026	Physical	$6,000
23 Nov 2026	04 Dec 2026	Physical	$12,000

Start Date	End Date	Location	Price
23 Nov 2026	04 Dec 2026	Physical	$9,000
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$12,000

Ready to enroll or request a customized program? Book online or contact us:

Register Online Now

Email: info@fineskilltrainingcenter.com | Website: www.fineskilltrainingcenter.com