

AI+ Video™ (1 Day)

Program Detailed Curriculum

Executive Summary

The AI + Video certification provides professionals with the expertise to leverage artificial intelligence in video production, enhancing both the creation and analysis of content. It focuses on AI-driven tools for editing, personalization, and real-time video insights. Participants will gain proficiency in automating processes, improving video quality, and utilizing data to boost audience engagement. Additionally, the certification addresses ethical considerations, privacy concerns, and the evolving role of AI in the media landscape. Targeted at content creators, marketers, and technology professionals, this certification ensures participants are well-equipped to navigate the growing intersection of AI and video technology.

Course Prerequisites

- **Basic Video Editing Skills:** Familiarity with video editing software is essential.
- **Understanding of AI Concepts:** Basic knowledge of artificial intelligence principles.
- **Familiarity with Data Analytics:** Comfort with data-driven decision-making and analysis.
- **Experience with Content Creation:** Hands-on experience in producing digital media.

Module 1

Foundation of AI in Video Integration

1.1 Basics of Video Processing

- **Introduction to Video Processing:** Introduction to video processing focuses on converting raw video into meaningful information. Techniques include frame manipulation, signal processing, and object detection to enable advanced video analysis using AI algorithms.
 - **Core Video Processing Techniques:** Core video processing techniques involve tasks like motion detection, video enhancement, and feature extraction. These methods help convert raw footage into a refined output that AI systems can analyze and interpret.
 - **Color Spaces, HDR, and Chroma Subsampling for AI Applications:** This section covers how color spaces, high dynamic range (HDR), and chroma subsampling impact video data interpretation. Choosing the right approach is critical for AI models to achieve accurate visual analysis and classification.
-

1.2 Introduction to AI in Video

- **Key AI Techniques in Visual Data Analysis:** AI techniques such as Convolutional Neural Networks (CNNs), image segmentation, and object detection enable systems to understand visual content. These methods are crucial for applications in security, healthcare, and autonomous systems.
 - **Comparing Machine Learning and Deep Learning in Video Processing:** Machine Learning uses handcrafted features for simpler tasks with lower resource requirements. Deep Learning leverages neural networks to learn from raw data, excelling in complex tasks but requiring more data and computational power.
 - **Common Applications in Media, Surveillance, and Enterprise:** AI enhances media through content recommendations, automated editing, and synthetic media. In surveillance, it powers threat detection and facial recognition. In enterprises, AI optimizes safety, retail analytics, manufacturing, and healthcare diagnostics.
-

1.3 Toolkits and Frameworks

- **OpenCV for Video Processing:** OpenCV is an open-source computer vision library with functions for real-time video processing. It supports tasks like object tracking and motion detection, making it a go-to tool for video AI applications.
 - **TensorFlow and PyTorch for AI-Driven Video Analytics:** TensorFlow and PyTorch are two major frameworks used for deep learning in video analytics. Both provide tools for classification, object detection, and video summarization, helping to build advanced video processing models.
 - **Setting Up a Video-AI Development Environment:** Setting up a video-AI development environment involves defining project goals, selecting appropriate hardware (e.g., GPUs), and using platforms like Google Cloud or NVIDIA DeepStream for scalable AI-based video processing.
-

1.4 AI-Enhanced Video Compression for Streaming

AI-based compression techniques reduce video file sizes without losing quality, optimizing bandwidth usage and improving user experience on streaming platforms. It involves neural networks to predict and compress video data effectively.

1.5 Case Study: YouTube's AI-Driven Transcoding System

YouTube's AI-powered transcoding system optimizes video delivery across devices and network conditions. Using adaptive bitrate streaming, it ensures smooth playback with reduced buffering and consistent quality, even for high-demand videos.

Preparing Video Data for AI

2.1 Data Preparation for AI Models

- **Extracting and Annotating Video Frames:** This subsection discusses methods for extracting video frames and annotating them, including tools and techniques to make video data usable for AI models.
 - **Organizing Datasets for Supervised Learning:** Here, the focus is on how to properly structure and label video data to prepare it for supervised learning tasks, ensuring accuracy and efficiency.
 - **Efficient Storage and Structuring Practices:** Best practices for storing and structuring large video datasets are highlighted, ensuring efficient data processing and scalable AI model training.
-

2.2 Preprocessing and Augmenting Frames

- **Normalizing Data and Scaling Input Sizes:** Explains normalization and scaling techniques for video data, which standardize input sizes and ranges, improving model performance and training efficiency.
 - **Cropping, Sampling, and Sequencing Frames:** Describes methods like cropping and sampling frames from videos, followed by sequencing to maintain chronological order, optimizing video data for model training.
 - **Augmentation for Motion Blur, Lighting, and Angle Variance:** Explores video data augmentation techniques, like adding motion blur and simulating lighting or angle changes, to make models more robust to real-world variations.
-

2.3 Storage and Workflow Management

- **Cloud Storage Strategies (AWS, GCP, etc.):** Focuses on leveraging cloud storage platforms like AWS and GCP for scalable, cost-effective management of large video datasets.
 - **Version Control for Video Datasets:** Describes the importance of version control for tracking changes to video data, ensuring dataset integrity, collaboration, and reproducibility.
 - **Workflow Automation Using ML Tools:** Highlights the automation of repetitive tasks in video data workflows using machine learning tools, improving efficiency and accuracy.
-

2.4 Use Case: Building AI-ready Video Datasets for Autonomous Driving Applications

This use case outlines steps to create video datasets for autonomous driving, focusing on data collection, preprocessing, and handling edge cases to train AI models.

2.5 Case Study: Tesla's In-house Pipeline for Labeling Driving Scenarios Across Multiple Geographies Using Video Footage

A detailed case study on how Tesla processes and labels video footage for its autonomous driving system using an in-house pipeline.

2.6 Hands-On: Video Annotation Using CVAT Tool and Organizing Them for Model Training

Describes a hands-on tutorial using the CVAT tool for annotating video data and organizing it for training object detection models.

Module 3

Machine Learning for Video Analysis

3.1 Video Classification and Tagging

- **Scene Recognition and Classification Models:** Scene recognition models use machine learning to classify video scenes based on spatial, temporal, and contextual features, helping to identify specific settings like urban landscapes or sports events within video data.
 - **Syncing Audio and Visual Components:** Syncing audio and visual components ensures seamless integration between sound and visuals in video content, enhancing viewer engagement and clarity through methods like manual syncing and automatic software tools.
 - **Tagging with Multiple Labels in Diverse Video Content:** Tagging multiple labels across video content enhances searchability, discoverability, and audience targeting, using techniques like semantic tagging, long-tail keywords, and AI-powered suggestions to improve content categorization.
-

3.2 Object Detection and Movement Tracking

- **Working with YOLO, SSD, and Faster R-CNN:** YOLO, SSD, and Faster R-CNN are advanced models for object detection in videos. They differ in speed and accuracy, with YOLO excelling in real-time applications, while Faster R-CNN focuses on precision.
 - **Tracking Motion Using Deep SORT and SORT Algorithms:** SORT and Deep SORT track multiple objects across frames in video sequences, with Deep SORT enhancing SORT by using deep learning features for improved performance, particularly in occlusions and crowded environments.
 - **Managing Occlusions and Rapid Movements:** Occlusions and rapid movements present challenges in tracking, causing issues like loss of track and misidentification. Solutions like appearance-based tracking, high-speed detection, and motion prediction help overcome these obstacles.
-

3.3 Action and Behavior Recognition

- **Using RNNs and LSTMs for Temporal Activity Detection:** RNNs and LSTMs are neural networks designed for sequential data, capturing temporal dependencies and enabling accurate detection and classification of activities, essential for understanding dynamic actions over time in video.
 - **Identifying Human and Group Behaviors:** Human and group behavior recognition detects individual actions and interactions, employing deep learning and temporal models like LSTMs to analyze behaviors, interactions, and emotions in video for security and social analysis.
 - **Crowd Monitoring and Anomaly Recognition:** Crowd monitoring uses AI to track and analyze crowd dynamics in real-time, with anomaly recognition detecting unusual behaviors, such as panic or aggression, using techniques like motion tracking, density estimation, and event-based analysis.
-

3.4 Use Case: Smart Surveillance Systems Detecting Abandoned Objects in Real Time

AI-driven surveillance systems automatically detect unattended objects in crowded areas, triggering real-time alerts to security teams, ensuring timely intervention to prevent safety threats in locations like airports or shopping malls.

3.5 Case Study: Dubai Smart City's AI Implementation for Object Recognition

Dubai Smart City uses AI to enhance surveillance at public transport hubs, integrating object recognition, behavior analysis, and real-time alerts to improve security and operational efficiency, minimizing risks at transit stations.

3.6 Hands-On: Train YOLOv8 on Sample Security Footage

In this hands-on session, users train YOLOv8 to detect and track objects, such as people in security footage, using a simple visual platform, reducing the need for manual video review and enhancing security monitoring.

Module 4

Generative AI in Video

4.1 Generating Synthetic Video with GANs

- **Using GANs to Synthesize New Video Content:** Learn how GANs generate new video content by synthesizing frames, performing video-to-video translation, and creating action sequences, revolutionizing video creation and enhancing creative possibilities.
 - **Applying Style Transfer in Moving Frames:** Discover how style transfer techniques applied to video frames can create visually appealing, artistic transformations while maintaining smooth motion and consistency across frames.
 - **Frame Interpolation Using Deep Learning:** Deep dive into how deep learning models generate intermediate frames to enhance video fluidity, improve frame rates, and enable slow-motion effects, offering superior results compared to traditional methods.
-

4.2 AI-Driven Animation and Avatars

- **Lip-Sync and Voice-to-Motion Generation:** Explore the process of synchronizing avatars' lip movements and body motions with speech using AI, enabling realistic character animations and voice-based interactions in virtual environments.
 - **Creating Realistic Avatars and Character Movement:** Understand the AI-driven techniques used to craft realistic avatars, incorporating detailed movements, expressions, and behaviors to create highly interactive and lifelike digital characters.
 - **Generating Video from Textual Input:** AI-powered text-to-video technology transforms written content into dynamic video sequences, using NLP, GANs, and deep learning to create videos for marketing, education, and entertainment applications.
-

4.3 Ethical Use of Generative Content

- **Understanding Deepfakes and Manipulation Risks:** Learn about deepfake technology, its potential dangers, and how it can be misused for misinformation, fraud, and manipulation, posing serious ethical challenges.
 - **Detection Methods and Digital Watermarking:** Discover methods for detecting deepfakes, including machine learning algorithms, facial forensics, and blockchain, and how digital watermarking ensures content authenticity and combats manipulation.
 - **Responsible AI Usage Guidelines:** Explore best practices and guidelines to ensure the responsible and ethical use of generative AI, promoting transparency, fairness, privacy protection, and accountability in AI-generated media.
-

4.4 Use Case: Auto-Generation of Product Explainer Videos Using Avatars and Synthesized Narration

Explore how AI-driven tools automate the creation of product explainer videos, streamlining content production, reducing costs, and enabling businesses to scale marketing and training efforts.

4.5 Case Study: Synthesia's Solution Enabling Businesses to Create AI-Driven Training and Marketing Videos

Discover how TechBridge Solutions utilized Synthesia's AI platform for scalable, multilingual video production, enhancing efficiency and expanding their global reach with localized content.

4.6 Hands-On: Generate a Deepfake or AI Avatar Using AKOOL

Learn how to generate deepfakes or AI avatars using AKOOL, with a focus on face alignment, identity swapping, and the ethical considerations of AI content creation.

Enhancing Video with AI

5.1 Super-Resolution and Restoration

- **Improving Live-Stream Quality with AI:** AI improves live-streaming quality by adjusting resolution, enhancing bitrate, and reducing lag, ensuring smooth and high-definition streaming, even in low-bandwidth or fluctuating network environments.
 - **Deblurring and Stabilizing Shaky Video:** AI removes motion blur and stabilizes shaky video, ensuring clear and smooth playback. Machine learning techniques are applied to restore clarity, particularly in dynamic and rapidly moving scenes.
 - **Creating Smooth Motion via Frame Prediction:** AI predicts intermediate frames to create smooth motion between frames, enhancing video fluidity, especially for slow-motion, low-frame-rate, or choppy footage, ensuring seamless transitions and reduced stuttering.
-

5.2 Real-Time Video Enhancement

- **Improving Live-Stream Quality with AI:** AI improves live-streaming quality by adjusting resolution, enhancing bitrate, and reducing lag, ensuring smooth and high-definition streaming, even in low-bandwidth or fluctuating network environments.
 - **Deblurring and Stabilizing Shaky Video:** AI removes motion blur and stabilizes shaky video, ensuring clear and smooth playback. Machine learning techniques are applied to restore clarity, particularly in dynamic and rapidly moving scenes.
 - **Creating Smooth Motion via Frame Prediction:** AI predicts intermediate frames to create smooth motion between frames, enhancing video fluidity, especially for slow-motion, low-frame-rate, or choppy footage, ensuring seamless transitions and reduced stuttering.
-

5.3 Making Video More Inclusive

- **Auto-Generated Captions and Subtitles:** AI generates captions and subtitles by transcribing speech into text, enhancing accessibility for deaf or hard-of-hearing viewers and enabling content comprehension in noisy environments or for non-native speakers.
 - **Real-Time Sign Language Interpretation:** AI provides real-time sign language interpretation, using computer vision and NLP to translate gestures into text or speech, ensuring accessible communication for the deaf and hard-of-hearing community during live events.
 - **Voiceover and Audio Description for Accessibility:** AI-driven voiceovers and audio descriptions narrate visual elements for visually impaired users, ensuring content accessibility by describing key visual details, actions, and settings in videos.
-

5.4 Use Case: Streaming Platforms Using AI to Enhance Resolution and Reduce Latency for Mobile Users

AI improves mobile streaming by dynamically adjusting resolution and reducing latency based on network conditions, ensuring smooth, high-quality playback despite fluctuating mobile network performance.

5.5 Case Study: DeOldify's Impact in Reviving Historical Video Archives by Upscaling and Colorizing Black-and-White Footage

DeOldify uses AI to upscale and colorize historical footage, improving its visual appeal. This technology helps preserve cultural heritage, making old videos more engaging and accessible for modern audiences.

5.6 Hands-On: Use AI4Video to Enhance a Sample Low-Resolution Black-and-White Video and Visualize Improvement

AI4Video enhances low-resolution black-and-white videos by upscaling resolution, reducing noise, and adding color, transforming old footage into clearer, more vibrant content suitable for modern high-definition displays.

Module 6

Interactive and Immersive AI Video

6.1 AI in AR and Mixed Reality

- **Using AI to Overlay AR Graphics Based on Camera Feed:** Discusses AI's role in overlaying graphics on live camera feeds, enhancing object tracking, depth perception, and scene understanding for interactive AR experiences with applications in gaming and education.
 - **Estimating Depth for Anchoring 3D Objects:** Focuses on AI techniques for depth estimation, ensuring accurate positioning and scale of virtual objects in real-world environments, which is essential for realistic AR/MR experiences in dynamic spaces.
 - **Real-time Special Effects with Visual Recognition:** This section explores how visual recognition technologies enable dynamic, interactive special effects in real-time. It covers applications across entertainment, gaming, and live events, using AI and computer vision to enhance immersion and personalization.
-

6.2 Intelligent Video Editing

- **Scene Segmentation and Cut Detection:** Explains how AI automates the detection of scene transitions and cut points in videos, facilitating faster editing by recognizing visual, motion, and audio cues, critical for organizing long footage.
 - **Auto-Generating Highlights and Trailers:** Describes how AI identifies key moments in video content to create compelling highlight reels or trailers, improving engagement by automatically assembling the most exciting or important segments.
 - **Storyboarding Powered by AI Analysis:** This section discusses how AI revolutionizes the traditional storyboarding process by automating scene and shot identification, visual representation generation, and integration of audio and dialogue, streamlining pre-production in video, film, and animation.
-

6.3 Viewer Engagement & Adaptation

- **Interactive Video Branching:** Introduces AI-driven interactive video that lets viewers make choices that influence the narrative, providing a dynamic, personalized storytelling experience across entertainment, training, and marketing.
 - **Personalized Edits:** Focuses on how AI customizes video content for individual viewers by adjusting elements like text, pacing, product placements, and visual themes, improving relevance based on viewer profiles and context.
 - **Emotion Recognition for Engagement:** This section covers how AI-based emotion recognition enhances video content by detecting viewer emotions through facial expressions, voice, and body language. It enables dynamic adjustments in real-time to optimize engagement and retention.
-

6.4 Use Case: Live Sports Broadcasters Using AR to Overlay Player Stats During Gameplay:

Shows how AR enhances live sports broadcasts by overlaying real-time player stats, enriching the viewer experience and allowing deeper analysis of the game's dynamics.

6.5 Case Study: NFL and AWS Collaboration to Deliver Real-Time Performance Insights via Augmented Visuals

Examines the partnership between the NFL and AWS to provide augmented reality-based real-time player stats, enhancing fan engagement and providing deeper insights into live game performance.

6.6 Hands-On: Creating a Highlight Video from a Video Clip Using Clipchamp

Guides users through creating a professional-looking highlight video using Clipchamp, demonstrating how AI simplifies the editing process by selecting key moments and enhancing them with audio, transitions, and text.

Module 7

AI in Video Surveillance and Compliance

7.1 Security and Monitoring Systems

- **Crowd Facial Recognition Systems:** AI-powered systems detect and identify faces in crowds, improving crowd control, identity verification, threat detection, predictive modeling, and ensuring privacy safeguards while optimizing security and resource allocation.
 - **Intrusion and Motion-Based Alerts:** Intrusion detection systems use AI algorithms and sensors to detect unauthorized access or suspicious movements, triggering alerts and enabling real-time analysis for immediate security responses.
 - **Recognizing and Logging License Plates:** AI-driven Automatic Number Plate Recognition systems capture and interpret vehicle plates, automating security, access control, and toll collection while providing real-time alerts and logs for surveillance and compliance.
-

7.2 Automated Content Moderation

- **Flagging and Filtering Sensitive Content:** AI systems automatically detect and filter harmful content, including violence, explicit materials, and threats, helping maintain security, compliance, and community guidelines across platforms.
 - **Managing Uploads on Large-Scale Platforms:** AI-powered moderation systems manage vast content volumes, flagging harmful uploads in real time, ensuring compliance, and minimizing human intervention across social media and video-sharing platforms.
 - **Visual Content Review Workflows:** Describes AI-driven and human-assisted workflows to moderate visual content, ensuring compliance with platform guidelines and legal regulations, utilizing automated detection and human review for effective moderation.
-

7.3 Addressing Privacy and Ethics

- **Compliance with GDPR and Surveillance Laws:** Explores how AI surveillance systems must adhere to GDPR and other privacy laws, ensuring lawful data processing, transparency, and protection of personal data in surveillance technologies.
 - **Privacy-Respecting AI Solutions:** Discusses strategies like data anonymization, edge computing, and real-time data masking to create AI systems that respect privacy while maintaining surveillance effectiveness and compliance with data protection regulations.
 - **Reducing Algorithmic Bias in Surveillance:** Outlines strategies to reduce bias in AI surveillance, including using diverse datasets, conducting audits, incorporating fairness constraints, and ensuring human oversight to promote fairness and accuracy in surveillance systems.
-

7.4 Use Case: Automated Real-Time Access Control in Corporate Offices Using Facial Authentication

Describes how facial recognition technology improves corporate office security, automating access control, providing real-time authentication, and eliminating the need for cards or PINs, enhancing both security and efficiency.

7.5 Case Study: Amazon Go's Cashier-less Stores Using Computer Vision for Security and Consumer Behavior Tracking

Explores Amazon Go's cashier-less stores using computer vision for security, theft prevention, and consumer behavior tracking, offering seamless shopping experiences while optimizing operational efficiency and store security.

7.6 Hands-On: Implement Facial Detection and Access Control Simulation Using OpenCV and A Basic Recognition Model

Guides users through creating a facial recognition access control system using Google Teachable Machine, providing practical experience in training models and simulating real-time access control without coding.

Future of AI + Video

8.1 Trends and Emerging Technologies

- **Running AI Video Analysis at the Edge:** AI video analysis at the edge involves processing data locally on devices, reducing latency, enhancing privacy, and improving efficiency by minimizing cloud dependency. Applications span surveillance, autonomous driving, and smart cities.
 - **Reducing Latency in Video Inference:** Reducing latency is vital for real-time AI video applications. Techniques like model optimization, hardware acceleration, edge computing, and 5G networks enable faster video processing, improving user experience in live streaming and surveillance.
 - **AI-powered Summarization and Captions:** AI-driven summarization condenses long videos into concise versions, while AI captions enhance accessibility by transcribing spoken words. These tools revolutionize video consumption, enabling faster digestion, improved engagement, and multi-language support.
-

8.2 AI Applications by Industry

- **Retail Analytics through Footfall and Heatmaps:** AI-based footfall analytics measures customer traffic, while heatmaps visualize movement within stores. These tools optimize store layouts, improve customer engagement, and enhance inventory management by tracking behaviors, peak hours, and product interactions.
 - **AI in Medical Imaging and Diagnostics:** AI improves medical imaging by enhancing image quality, automating detection, and enabling real-time monitoring. It aids in faster diagnoses, predictive analytics, and personalized medicine, supporting healthcare professionals in making accurate decisions.
 - **City-scale Public Safety Systems:** AI-powered public safety systems use real-time surveillance, predictive policing, and disaster response to enhance city safety. By analyzing video and sensor data, these systems proactively identify risks, improving response times and overall safety.
-

8.3 Careers and Professional Growth

- **Job Profiles, Responsibilities, and Required Skills:** Key job roles include AI video specialists, content creators, data analysts, and marketing experts. Required skills range from machine learning, video editing, and data analysis to creative content development and consumer behavior understanding.
 - **Career Growth Tracks:** Career growth in AI and video spans from entry-level roles to expert positions. Professionals can advance from technical roles like AI engineers to managerial and strategic positions, such as AI video architects and CTOs.
 - **Opportunities in Freelancing, Startups, and Enterprises:** Freelancers, startups, and enterprises offer various opportunities in AI and video, including video production, content creation, and analysis. Freelancers work on projects globally, while startups provide dynamic innovation roles, and enterprises ensure stability.
-