HOCHSCHULE
MITTWEIDA
University of Applied Sciences

BACHELORARBEIT

Frau
Gizem Yalcin

Verlinkung von Videodateien auf der Metadatenebene im Anwendungsgebiet der Videoforensik

Mittweida, August 2024

Fakultät **Angewandte Computer- und Biowissenschaften**

BACHELORARBEIT

Verlinkung von Videodateien auf der Metadatenebene im Anwendungsgebiet der Videoforensik

Autor:
Gizem Yalcin

Studiengang:
IT-Forensik/Cybercrime

Seminargruppe:
CC20w1-B

Erstprüfer:
Prof. Dr. rer. pol. Ronny Bodach

Zweitprüferin:
Sven Becker, M.Sc.

Einreichung:
Mittweida, 12.08.2024

Verteidigung/Bewertung:
Mittweida, 2024

Faculty of **Applied Computer Sciences and Biosciences**

BACHELOR THESIS

Linking video files at the metadata level in the application area of video forensics

Author:
Gizem Yalcin

Course of Study:
IT-Forensic/Cybercrime

Seminar Group:
CC20w1-B

First Examiner:
Prof. Dr. rer. pol. Ronny Bodach

Second Examiner:
Sven Becker, M.Sc.

Submission:
Mittweida, 12.08.2024

Defense/Evaluation:
Mittweida, 2024

Bibliografische Beschreibung

Yalcin, Gizem:
Verlinkung von Videodateien auf der Metadatenebene im Anwendungsgebiet der Videoforensik. – 2024. – 50 S.
Mittweida, Hochschule Mittweida – University of Applied Sciences, Fakultät Angewandte Computer- und Biowissenschaften, Bachelorarbeit, 2024.

Referat

Metadaten speichern Informationen über Dateien. Das könnten im Fall von Videodateien Informationen sein, wie Zeitstempel und inhaltsbasierte Daten, wie Objekte oder Personen. Durch die Verknüpfung dieser Details können unter anderem Bewegungsprofile von Personen entstehen, die beispielsweise in der Strafverfolgung bei der Identifizierung und Verfolgung von Verdächtigen unterstützen können und somit zur Aufklärung von Straftaten beitragen. Ziel dieser Arbeit ist es, durch die Betrachtung verschiedener Methoden für die Verknüpfung von Metadaten aus Videodateien eine Erleichterung für videoforensische Analysen zu erreichen. Diese Betrachtung soll als theoretische Grundlage für die Entwicklung einer Software oder einem Werkzeug dienen.

Im Rahmen der Arbeit werden verschiedene methodische Ansätze zur Extraktion und Analyse von Metadaten detailliert beschrieben. Dabei wird zusätzlich ein besonderer Fokus auf die Kombination unterschiedlicher Methoden gelegt. Die Arbeit diskutiert auch die Herausforderungen und Grenzen dieser Technologien, insbesondere hinsichtlich Datenschutz und ethischer Aspekte. Ein wichtiger Teil dieser Bachelorarbeit ist die Untersuchung und Darstellung, wie verknüpfte Daten in der Videoforensik visuell aufbereitet werden können. Die Visualisierung spielt eine entscheidende Rolle, um komplexe Informationen verständlich zu vermitteln und Ermittlern zu ermöglichen, Muster und Zusammenhänge schnell zu erkennen. Veranschaulicht wird die Thematik der Arbeit anhand von exemplarisch ausgewählten realen Ereignissen, bei denen die videoforensische Analyse eine Rolle spielte.

Inhaltsverzeichnis

Abbildungsverzeichnis

Abkürzungsverzeichnis

CMS . . . . . Content Management System

CNN . . . . . Convolutional Neural Network

CSV . . . . . Comma-separated values

DAM . . . . . Digital Asset Management

DCNN . . . . . Deep Convolutional Neural Network

DSGVO . . . . . Datenschutz-Grundverordnung

EXIF . . . . . Exchangeable Image File Format

FBI . . . . . Federal Bureau of Investigation

FPN . . . . . Feature Pyramid Network

GPS . . . . . Global Positioning System

IPTC . . . . . International Press Telecommunications Council

IS . . . . . Islamische Staat

JEIDA . . . . . Japan Electronic Industries Development Association

JPEG . . . . . Joint Photographic Experts Group

NSU . . . . . Nationalsozialistischer Untergrund

R-CNN . . . . . Region-based Convolutional Neural Network

RPN . . . . . Region Proposal Network

SSIS . . . . . Shadow-Object Segmentation and Instance Shadow Tracking

TIFF . . . . . Tag Image File Format

UTF . . . . . Unicode Transformation Format

ViT . . . . . Vision Transformer

VR . . . . . Virtual Reality

XML . . . . . Extensible Markup Language

XMP . . . . . Extensible Metadata Platform

1 Einleitung

In der Welt der digitalen Forensik hat die Bedeutung von Videomaterial als Beweismittel in den letzten Jahren zugenommen, vor allem durch die weit verbreitete Nutzung von Überwachungskameras, Smartphones und anderen Geräten, die Videoaufnahmen ermöglichen [1]. Durch diese Herausforderung ist es notwendig Videodateien möglichst effizient zu nutzen und zu analysieren. Eine Möglichkeit hierzu ist die Verknüpfung von Videodateien auf der Metadatenebene. Diese Technik ermöglicht es, umfangreiche Informationen aus Videodateien zu extrahieren und unterschiedliche Videos aus verschiedenen Quellen zu verknüpfen.

Metadaten, die strukturierte Informationen zu Videodateien liefern, umfassen nicht nur technische Details wie Datum, Uhrzeit und Kameraeinstellungen, sondern auch kontextbezogene Daten, die Aufschluss über den Ort der Aufnahme und möglicherweise beteiligte Personen oder Objekte geben [2]. Die Fähigkeit, diese Daten effektiv zu verknüpfen und zu analysieren, kann eine große Rolle in der modernen Videoforensik spielen. Sie kann es Ermittlern ermöglichen, aus verschiedenen, oft unzusammenhängenden Videoquellen, einen zusammenhängenden Ablauf eines Ereignisses zu formen. In der forensischen Praxis kann die Verknüpfung von Videodateien auf der Metadatenebene möglicherweise dazu beitragen, komplexe Straftaten aufzudecken, indem sie verborgene Muster und Zusammenhänge offenlegt, die durch die Betrachtung eines einzelnen isolierten Videos sonst unerkannt geblieben wären. Darüber hinaus hat sich die Analyse von Metadaten als ein wichtiges Werkzeug erwiesen, um die Integrität und Authentizität von Beweismitteln in Strafverfahren zu gewährleisten [3].

1.1 Hintergrund und Motivation

Im heutigen digitalen Zeitalter, in der Videomaterial eine immer größere Rolle in der Sicherheit und Überwachung spielt, sind präzise und effiziente Methoden zur Analyse dieser Daten notwendig. Ein zentraler Bestandteil in diesem Zusammenhang ist die Nutzung von Metadaten, die in den Exchangeable Image File Format (EXIF)-Daten von Videoaufnahmen gespeichert werden. Diese technischen Metadaten, bieten wesentliche Informationen, die für die forensische Analyse und auch beispielsweise für die Erkennung von Bildmanipulationen von großem Wert sind [4]. Sie können Ermittlern helfen, die Authentizität [3] und Herkunft des Materials zu verifizieren und bieten wichtige Kontextinformationen [2] zu den Aufnahmen.

Neben den technischen Metadaten gewinnt die inhaltsbasierte Verknüpfung von Videodateien zunehmend an Bedeutung. Diese fortschrittliche Methode nutzt Algorithmen des maschinellen Lernens und der künstlichen Intelligenz, um Inhalte innerhalb verschiedener Videoquellen zu identifizieren und miteinander zu verbinden [5]. Durch die Analyse von visuellen Merkmalen können ähnliche Objekte, Personen oder Handlungen in unterschiedlichen Videos erkannt und zugeordnet werden. Diese inhaltsbasierte Vernetzung eröffnet neue Möglichkeiten für die forensische Analyse, da sie dabei hilft, komplexe Ereignisse über mehrere unterschiedliche Videoquellen hinweg zu rekonstruieren und somit tiefere Einblicke in die Umstände eines Vorfalls zu gewinnen.

In dieser Arbeit werden, sowohl die Möglichkeiten der EXIF-basierten Metadatenanalyse als auch die Ansätze der inhaltsbasierten Verknüpfung betrachtet. Diese duale Perspektive soll zusammengeführt werden, um die Idee von einem umfassenden System zu entwickeln, das in der Lage ist, aus den massiven Datenmengen, die moderne Überwachungssysteme und mobile Endgeräte produzieren, sinnvolle und verwertbare Informationen zu extrahieren. Die anschließende Verknüpfung soll dabei helfen, kontextbezogene Analysen über mehrere Aufnahmen hinweg durchzuführen.

1.2 Zielsetzung der Arbeit

Das primäre Ziel dieser Bachelorarbeit ist es, eine Analyse darüber durchzuführen, mit welchen Methoden eine Software oder ein Werkzeug entwickelt werden kann, das Videodateien effektiv und ggf. auch effizient miteinander verknüpft, um komplexe Analysen zu erleichtern. Der Ausgangspunkt dieser Untersuchung ist eine Betrachtung von bedeutenden Fallbeispielen, wie dem Boston Massaker, dem Anschlag auf dem Berliner Weihnachtsmarkt und dem Nationalsozialistischer Untergrund (NSU)-Fall. Diese Fälle dienen nicht nur als empirische Grundlage zur Veranschaulichung der realen Anwendbarkeit der Verknüpfungstechniken, sondern auch dazu, die spezifischen Herausforderungen und Anforderungen zu identifizieren, die bei der Verknüpfung von Videodateien entstehen.

Im Anschluss an die Fallstudien wird die Arbeit verschiedene methodische Ansätze zur Verarbeitung und Verknüpfung von Videodateien aufzeigen. Dies umfasst unter anderem fortschrittliche Techniken wie Personenerkennung, Objekterkennung und Gesichtserkennung. Auch die Verknüpfung auf Grundlage der EXIF-Daten wird betrachtet. Schließlich widmet sich die Arbeit der Frage, wie die durch die Verknüpfung gewonnenen Informationen möglichst optimal visualisiert und präsentiert werden können. Hierbei werden Ansätze wie der Einsatz von Virtual Reality (VR) oder 3D-Visualisierungen in Betracht gezogen, um eine Umgebung zu schaffen, die es den Benutzern ermöglicht, die verknüpften Daten intuitiv zu erkunden. Zusätzlich werden unter anderem Möglichkeiten wie interaktive Karten und strukturierte Listen oder Tabellen betrachtet, um die geografische und zeitliche Verteilung von Ereignissen übersichtlich darzustellen.

Um das oben genannte Ziel der Arbeit zu erreichen, beschäftigen sich die nachfolgenden Abschnitte zunächst mit den theoretischen Grundlagen. Dort werden die notwendigen Begriffe, Technologien und Herausforderungen dargestellt, die für das Verständnis der Video- und Metadatenanalyse notwendig sind. Es bietet das Fundament für die weiteren Untersuchungen, indem es die technischen Voraussetzungen und Anwendungsbereiche beschreibt. Das darauffolgende Kapitel betrachtet reale Vorfälle, wie das Bostoner Marathon-Attentat und den NSU-Fall. Dieses Kapitel veranschaulicht die praktische Relevanz und die Herausforderungen bei der Analyse und Verknüpfung von Videodateien, indem sie konkrete Anwendungsbeispiele und die in diesen Fällen verwendeten forensischen Methoden und die daraus resultierenden Probleme beleuchtet. Das vierte Kapitel, stellt die verschiedenen technischen Ansätze dar, die zur Extraktion, Analyse und Verknüpfung von Videodateien verwendet werden können. Dieses Kapitel betrachtet etablierte und innovative Technologien, die zur Verbesserung der Effizienz und Genauigkeit der forensischen Videoanalyse beitragen können. Anschließend folgt ein Kapitel, dass die Methoden zur Visualisierung der gewonnenen Daten darstellt, um sie für die Benutzer intuitiv und zugänglich zu machen. Diese Visualisierungen sind notwendig, um komplexe Datenmengen verständlich darzustellen. Das abschließende Kapitel befasst sich mit den Ergebnissen und Diskussion und bietet eine Reflexion über die in der Arbeit durchgeführten Analysen und betrachteten Methoden.

2 Theoretische Grundlagen

Metadaten, als strukturierte Informationen über eine Videodatei, bieten wichtige Hinweise auf deren Inhalt, Erstellungsprozess und Kontext [2]. Durch die Betrachtung und Auswertung dieser Metadaten besteht die Möglichkeit für Ermittler Videoinhalte effizienter zu analysieren, was zu einer verbesserten Aufklärung von Straftaten führen kann. Die Verknüpfung von Videodateien auf der Metadatenebene könnte somit in der modernen Videoforensik bei der Analyse von Videodateien eine Arbeitserleichterung darstellen.

Neben der Nutzung von Metadaten auf technischer Ebene ist auch die inhaltsbasierte Verknüpfung von Videodateien ein zentraler Aspekt in der Analyse von Videodateien. Diese Art der Verknüpfung bezieht sich auf die Anwendung fortgeschrittener Technologien wie beispielsweise Objekterkennung, Gesichtserkennung und Personenerkennung, um spezifische Inhalte innerhalb eines Videos zu identifizieren und zu analysieren. Im Folgenden werden die verschiedenen Aspekte der Verknüpfung von Videodateien auf Metadatenebene und deren Bedeutung für die Videoforensik detaillierter betrachtet.

2.1 Begriffsdefinitionen und Terminologie

Die Videoforensik ist ein Fachgebiet der digitalen Forensik, das sich auf die Untersuchung und Analyse von Videomaterial aus unterschiedlichen Quellen wie Sicherheitskameras, Mobiltelefonen, aber auch anderen Videoquellen konzentriert. Sie umfasst technische Verfahren zur bildtechnischen Verbesserung von Aufnahmen, die unter weniger idealen Umständen entstanden sind, sowie photogrammetrische Methoden zur räumlichen Analyse und Rekonstruktion von Szenen. Bei der Analyse von Videodateien können auch deren Metadaten betrachtet werden. [6]

Bei Metadaten handelt es sich um sogenannte „Daten über andere Daten“. Das bedeutet, das Videodateien zusätzliche Informationen über Videomaterial beinhalten können, welche über den bloßen Bildinhalt hinausgehen. Beispiele für solche Informationen sind die Beschreibung des Videos, der Codec, die Zeitstempel und auch Informationen zum Aufnahmegeräts. [7]

Metadaten lassen sich in drei Kategorien einteilen: technische, beschreibende und administrative Metadaten. Technische Metadaten beschreiben die technischen Aspekte einer Videodatei. Dazu gehören Informationen über die technischen Eigenschaften eines Bildes, wie ISO-Geschwindigkeit, Größe und Farbprofil. Sie können aber auch zusätzliche Informationen wie Besitz-, Kamera- und Beschreibungsangaben enthalten. Beschreibende Metadaten liefern Kontextinformationen wie den geografischen Standort der Aufnahme durch Global Positioning System (GPS)-Koordinaten, das Datum und die Uhrzeit der Aufnahme sowie Ereignisbeschreibungen. Sie umfassen Angaben wie Bildunterschriften, Titel und Standort, die das Auffinden von Bildern in Sammlungen erleichtern können. Administrative Metadaten bieten Informationen zur Herkunft des Bildes und beinhalten standardisierte Felder wie Nutzungseinschränkungen und Kontaktdaten des Erstellers. [4]

2.2 Bedeutung und Anwendungsbereiche

Grundsätzlich ist die Analyse von Metadaten in der Videoforensik von großer Bedeutung, da sie eine umfassende und tiefgehende Analyse von Videodateien ermöglicht. Die Analyse ist unter anderem wichtig für den Nachweis von Manipulationen, Anomalien oder Inkonsistenzen in den Metadaten. Beispielsweise können Zeitstempel auf eine mögliche Bearbeitung oder Änderung der Datei hinweisen, was in forensischen Untersuchungen ein entscheidender Beweis sein kann. Darüber hinaus spielen Metadaten eine zentrale Rolle bei der Organisation und Verwaltung großer Mengen von Videomaterial, was sie zu einem wichtigen Werkzeug in der digitalen Forensik machen kann. [4]

Die Rekonstruktion von Ereignissen durch Videodateien und deren Metadaten spielt eine wesentliche Rolle in der Videoforensik, besonders bei der Untersuchung komplexer Vorfälle wie terroristischer Anschläge. Metadaten, die technische, beschreibende und administrative Informationen enthalten, ermöglichen es, unterschiedliche Videoquellen zu verknüpfen und zu analysieren [4]. Mit der Verknüpfung der einzelnen Videodateien kann man möglicherweise Muster und Zusammenhänge erkennen, die bei isolierter Betrachtung einzelner Videodateien möglicherweise nicht sichtbar wären. Durch die Kombination dieser Daten können Ermittler die räumliche und zeitliche Abfolge von Ereignissen nachvollziehen, was bei der Beweisführung und die Ermittlungsarbeit helfen kann. [8]

2.3 Technologische Grundlagen

Um Videodateien auf Metadatenebene zu verknüpfen, sind Kenntnisse in mehreren Schlüsselbereichen notwendig. Gonzalez und Woods bieten in *Digital Image Processing* [9] eine umfassende Grundlage für die digitale Bildverarbeitung, die für die Verarbeitung und Analyse von Videodateien von Bedeutung ist. Techniken der digitalen Bildverarbeitung sind wichtig, um beispielsweise die Qualität von Videomaterial zu verbessern und dadurch die Analyse zu erleichtern. Durch die Verbesserung der Schärfe, Reduktion von Bildrauschen und Korrektur von Beleuchtungsproblemen wird das Bildmaterial klarer und Details werden sichtbarer, was für eine forensische Untersuchung von Vorteil sein kann.

Szeliski erweitert dieses Wissen durch *Computer Vision: Algorithms and Applications* [10], einem Werk, das sich den Algorithmen und Anwendungen der Computer Vision widmet, die für die Interpretation von Videos wichtig sind. Algorithmen können die effiziente Verarbeitung und Analyse großer Datenmengen erleichtern, da sie Aufgaben wie die Erkennung von Mustern oder spezifischen Ereignissen automatisieren. Diese Fähigkeit ist wichtig, um mit dem Anstieg der durch Videoüberwachung und Videos aus mobilen Endgeräten erzeugten Datenmengen gerecht zu werden und eine skalierbare Lösung für die Datenanalyse zu bieten. Darüber hinaus ermöglichen moderne Algorithmen, die auf Techniken des maschinellen Lernens und der künstlichen Intelligenz basieren, tiefere Einblicke in die Daten. Sie können Muster und Verbindungen erkennen, die nicht sofort offensichtlich sind, und tragen so dazu bei, komplexe Zusammenhänge zu entschlüsseln. Ein Grundlagenwerk hierfür ist die Ausarbeitung von Goodfellow, Bengio und Courville mit *Deep Learning* [11], das in die Thematik einführt.

Schließlich bietet *Database Systems: The Complete Book* von Garcia-Molina, Ullman und Widom [12] Einblicke in die Prinzipien des Datenmanagements, die für die Handhabung der durch Videos generierten Datenmengen von großer Bedeutung sind. Sie spielen eine entscheidende Rolle in

der Organisation und Zugänglichkeit von Videodateien, indem sie systematische Strukturen zur effizienten Speicherung und Kategorisierung bieten. Diese Strukturen ermöglichen es, Videodateien schnell und einfach zu identifizieren und auf sie zuzugreifen, was besonders in großen Datensätzen nützlich ist. Zusätzlich gewährleistet ein robustes Datenmanagement die Einhaltung von rechtlichen Anforderungen, wie Datenschutzbestimmungen und Aufbewahrungspflichten. Alle Datenmanagementprozesse sollten den geltenden Gesetzen und Richtlinien entsprechen, um rechtliche Probleme zu vermeiden und die Integrität der Daten zu wahren.

Im Bereich der Metadatenverarbeitung und -analyse innerhalb von Videoüberwachungssystemen liegt der Fokus aktueller Forschungen auf Schlüsselaspekten wie der Vorverarbeitung von Daten, der Extraktion relevanter Merkmale, der Verfolgung von Objekten und dem Verständnis von Verhaltensmustern. Fortschrittliche Methoden zur Hintergrundmodellierung, wie beispielsweise der Einsatz von Gaussian Mixture Models, sowie Techniken, die einen optischen Fluss für die Erstellung von Feature-Deskriptoren nutzen, stehen im Mittelpunkt dieser Untersuchungen. Ein besonderes Augenmerk liegt auf der Erkennung von abnormalem Verhalten und Ereignissen, die sowohl auf individueller als auch auf kollektiver Ebene auftreten können. Hierfür werden Deep Learning-Methoden integriert, um die Verarbeitung in Echtzeit und die Anomalieerkennung zu verbessern. [13]

Metadaten erleichtern außerdem die Identifizierung und Handhabung von Schlüsseldetails in Videoszenen, wie Standort, Zeit, Bewegungsmerkmale und vieles mehr, und unterstützen somit bei der schnellen Aktion und Entscheidungsfindung basierend auf Videoinhalten. Die Integration von Künstlicher Intelligenz und maschinellem Lernen verbessert die Präzision dieser Metadaten, ermöglicht eine tiefgreifende Analyse und trägt zur Entwicklung intelligenterer Videoüberwachungssysteme bei. [2]

2.4 Herausforderungen und Probleme

In der Videoforensik führt die steigende Datenmenge zu signifikanten Herausforderungen. Die enorme Zunahme an Videoaufnahmen aus beispielsweise Überwachungskameras und mobilen Geräten generieren Datenmengen, die gesichtet, analysiert und gespeichert werden müssen. Dieser wachsende Datenstrom erschwert nicht nur die schnelle Identifizierung relevanter Inhalte, sondern beansprucht auch erhebliche Speicherressourcen und Rechenleistung zur Verarbeitung der Daten. Darüber hinaus erhöhen die unterschiedlichen Datenformate die Komplexität der Datenverarbeitung, da unterschiedliche Formate und Kodierungsstandards angepasst und integriert werden müssen. Die Entwicklung effizienter Werkzeuge und Methoden zur Bewältigung dieser Probleme ist daher notwendig, um die Effektivität forensischer Analysen zu gewährleisten. [14]

Im Kontext der digitalen Forensik ist die Verarbeitungsgeschwindigkeit eine zentrale Herausforderung, besonders angesichts der wachsenden Datenmengen, die analysiert werden müssen. Moderne forensische Werkzeuge müssen in der Lage sein, große Mengen von Daten effizient und zeitnah zu verarbeiten. Die Notwendigkeit, diese Datenmengen in Echtzeit zu bearbeiten, erfordert fortschrittliche Algorithmen und leistungsfähige Hardware, um die Genauigkeit der Untersuchungen nicht zu beeinträchtigen. Die Entwicklung solcher Systeme, die hohe Durchsatzraten ohne Kompromisse bei der Datenintegrität ermöglichen, ist daher von entscheidender Bedeutung. [15]

Die Qualität der Daten beeinflusst ebenfalls die forensische Analyse. Videos, die unter schlechten Lichtverhältnissen, mit niedriger Auflösung oder mit ungeeigneten Kameraeinstellungen aufgenommen wurden, können die Identifizierung relevanter Details erschweren. Die Studie *Face Recognition in Poor-Quality Video: Evidence From Security Surveillance* zur Gesichtserkennung in Überwachungsvideos hebt hervor, dass diese Bedingungen die Erkennungsleistung erheblich beeinträchtigen, insbesondere wenn die Betrachter, also die Analysten, die Zielpersonen nicht persönlich kennen. Techniken zur Verbesserung der Bildqualität sind daher von entscheidender Bedeutung, um die Effektivität der Videoforensik zu gewährleisten und die Zuverlässigkeit von Beweismitteln zu erhöhen. [16]

Verändert man die Qualität von Aufnahmen, kann dies zur Verbesserung eines Videos dienen. Es gibt jedoch auch Manipulationen, die böswillig vorgenommen werden. Diese zu erkennen ist eine wichtige Aufgabe für Forensiker, die in der Lage sein müssen, selbst subtile Manipulationen zu entdecken und nachzuweisen. In der digitalen Forensik ist die Aufdeckung solcher Manipulationen von Bedeutung, da sie das Ergebnis einer Untersuchung verfälschen können. Fortgeschrittene Methoden, basierend auf maschinellem Lernen, sind zunehmend notwendig, um Anomalien und Inkonsistenzen in den Videos, die auf Manipulationen hindeuten könnten, zu identifizieren. [17]

Die Verarbeitung von Videodateien in der digitalen Forensik bringt auch ethische und rechtliche Fragen mit sich. Es wird besonders problematisch, wenn Ermittler ohne die erforderlichen rechtlichen Genehmigungen auf sensible Daten zugreifen. Dies kann tief in die Privatsphäre der Menschen eingreifen und deren Rechte verletzen. Ein weiteres Problem ist die mögliche Verzerrung durch Technologien wie die Gesichtserkennung, die bei bestimmten demografischen Merkmalen wie Hautfarbe, Geschlecht oder kulturellen Unterschieden, wie Kleidung oder Gesichtsausdruck, weniger zuverlässig sein können, was zu falschen Identifizierungen und somit zu unrechtmäßigen Festnahmen führen kann. Um diesen Herausforderungen gerecht zu werden ist es wichtig, dass forensische Untersuchungen strengen ethischen Richtlinien folgen. Es muss sichergestellt werden, dass die Privatsphäre gewahrt bleibt, keine Verzerrungen auftreten und die Ergebnisse der Analysen genau sind. [18]

2.5 Schematischer Ablauf

Vor diesem Hintergrund hat sich diese Arbeit das Ziel gesetzt, sich insbesondere mit der Herausforderung der stetig wachsenden Datenmenge in der Videoforensik zu beschäftigen. Eine manuelle Auswertung durch Analysten erweist sich in diesem Zusammenhang als ineffizient, da die große Anzahl an Videodateien und die erforderliche Detailtiefe der Analyse den Prozess erheblich verzögern würden. Um diesen Herausforderungen gerecht zu werden, ist der Einsatz automatisierter Verfahren unerlässlich. Solche Verfahren müssen in der Lage sein, große Datenmengen effizient und präzise zu verarbeiten, relevante Informationen aus den Videodateien zu extrahieren, diese zu analysieren und schließlich miteinander zu verknüpfen, um ein umfassendes Bild der Ereignisse zu erstellen. Die in dieser Arbeit betrachteten Grundlagen und Methoden sollen die Basis für die Entwicklung eines Werkzeugs oder einer Software schaffen, das Videodateien sowohl inhaltsbasiert als auch auf Grundlage ihrer EXIF-Daten verknüpft. Der schematische Ablauf eines solchen Programms kann in fünf Schritten beschrieben werden (siehe Abbildung 2.1) und beginnt mit der Datenerfassung und -eingabe (1). In diesem Schritt werden Videodateien aus verschiedenen Quellen wie Überwachungskameras, Smartphones und anderen Aufnahmegeräten gesammelt und in ein System importiert.

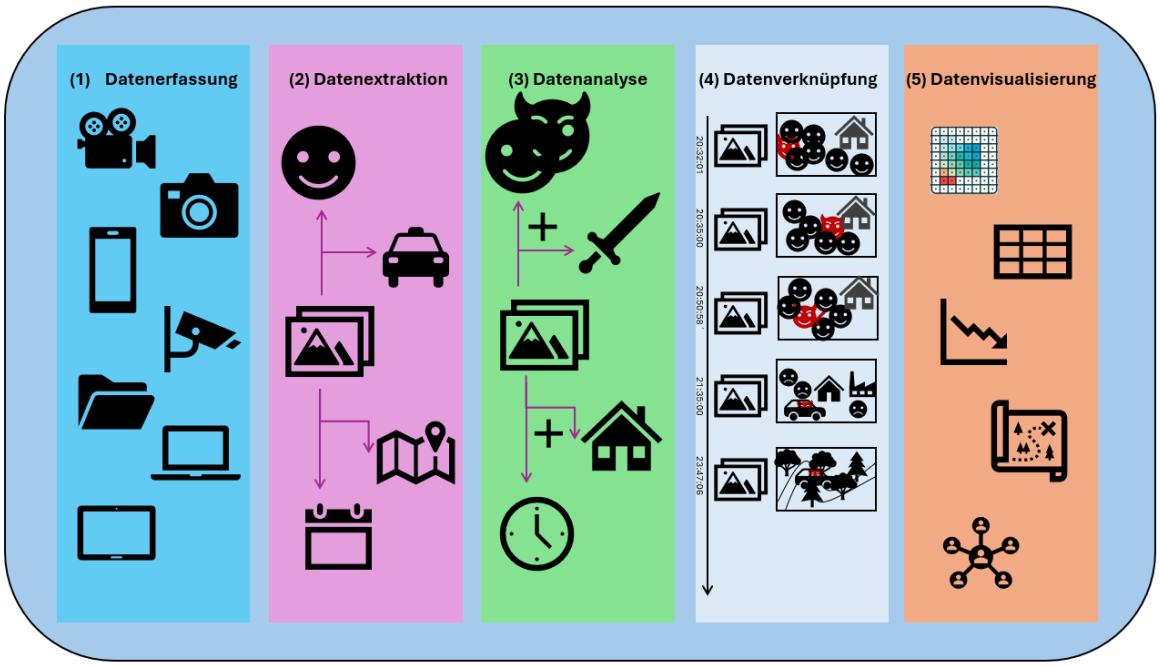

Abbildung 2.1: Schematischer Ablauf der Verknüpfung von Videodateien.

Anschließend erfolgt die Extraktion von Metadaten (2), bei der technische und inhaltsbasierte Metadaten automatisiert aus den Videodateien gewonnen werden. Diese Metadaten umfassen wichtige Informationen wie Datum, Uhrzeit, GPS-Daten, Kameramodell sowie erkannte Personen, Objekte und Szenen. Ein weiterer wesentlicher Schritt in diesem Prozess ist die Normalisierung der Daten, um sicherzustellen, dass die Metadaten aus verschiedenen Quellen in einem einheitlichen Format vorliegen und konsistent analysiert werden können. Die Datenextraktion wird in Abschnitt 4.1 konkretisiert.

Nach der Datenerfassung und Metadatenextraktion folgt die Metadatenanalyse (3). Hierbei werden zunächst die EXIF-Metadaten analysiert, um technische Details der Aufnahmen zu untersuchen. Gleichzeitig werden Algorithmen des maschinellen Lernens und der künstlichen Intelligenz eingesetzt, um inhaltsbasierte Metadaten zu analysieren und zu ergänzen. Diese Analyse ermöglicht die Erkennung und Identifizierung von Personen, Objekten und Szenen innerhalb der Videodateien, wodurch kontextuelle Informationen gewonnen werden können. Im Abschnitt 4.2 werden einzelne Methoden vorgestellt, die man für die Analyse von Metadaten verwenden kann.

Der nächste Schritt im Ablauf ist die Datenverknüpfung (4). Durch beispielsweise den Abgleich der Zeitstempel können Videodateien zeitlich miteinander verknüpft werden, um eine chronologische Sequenz der Ereignisse zu erstellen. Die Verknüpfung von Videos basierend auf geografischen Informationen mittels Geotagging und räumlicher Analyse ermöglicht es, räumliche Beziehungen zwischen den Aufnahmen zu erkennen. Zusätzlich erfolgt eine inhaltsbasierte Verknüpfung, bei der Personen, Objekte und Szenen, die in verschiedenen Videos erkannt wurden, miteinander in Beziehung gesetzt werden. Dies erlaubt eine umfassendere Analyse und das Auffinden von Zusammenhängen, die bei isolierter Betrachtung einzelner Videos möglicherweise verborgen geblieben wären. Abschnitt 4.3 beschäftigt sich mit insbesondere mit der Kombination der verschiedenen Me-

thoden. Es reicht nicht aus, sich nur auf eine einzelne Methode zu konzentrieren, wie beispielsweise die zeitliche Verknüpfung von Daten. Vielmehr müssen verschiedene Dimensionen wie Zeit, Ort und Inhalt kombiniert werden, um ein umfassendes und genaues Bild von einem Ereignis zu erhalten.

Der abschließende Schritt im schematischen Ablauf ist die Datenvisualisierung (5), das in Kapitel 5 genauer beschrieben wird. Um die verknüpften Daten übersichtlich darzustellen, werden verschiedene Visualisierungsmethoden eingesetzt. Zeitliche Visualisierungen, wie die Darstellung der verknüpften Videos auf einer Zeitleiste, ermöglichen die Nachverfolgung der Ereignisse in chronologischer Reihenfolge. Interaktive Karten werden verwendet, um die geografischen Verknüpfungen der Videodateien zu visualisieren. Netzwerke und Graphen helfen dabei, die Beziehungen zwischen verschiedenen Ereignissen und Personen darzustellen, während Listen und Tabellen eine strukturierte und detaillierte Analyse der verknüpften Daten ermöglichen. Diese vielfältigen Visualisierungstechniken tragen dazu bei, die Analyseergebnisse klar und verständlich zu präsentieren und erleichtern es Ermittlern, relevante Informationen schnell zu erfassen.

Der in dieser Arbeit vorgestellte schematische Ablauf bietet einen strukturierten Ansatz, um die Herausforderung der stetig wachsenden Datenmengen in der Videoforensik zu lösen. Durch die systematische Erfassung, Extraktion, Analyse, Verknüpfung und Visualisierung von Videodateien wird eine umfassende und präzise Auswertung ermöglicht. Diese automatisierten Verfahren erlauben es, große Datenmengen effizient zu verarbeiten und die relevanten Informationen gezielt zu extrahieren, um sie in einem sinnvollen Kontext zu verknüpfen. Damit wird eine Grundlage geschaffen, die den Anforderungen moderner forensischer Untersuchungen gerecht wird und eine detaillierte, zeit- und ortsübergreifende Analyse komplexer Ereignisse ermöglicht.

3 Bedeutende Ereignisse

Die Analyse und Auswertung von Video- bzw. Bilddateien hat sich im Laufe der Zeit als wichtig für die Aufklärung von Straftaten erwiesen. Beispielsweise können verschiedene Kameraperspektiven eine Situation aus unterschiedlichen Blickwinkeln dokumentieren und so die Identifizierung mehrerer Personen ermöglichen. Eine manuelle Auswertung dieser Daten, ohne Unterstützung durch spezialisierte Software, ist jedoch zeitaufwändig und ineffizient. [19]

Ohne eine gezielte Verknüpfung der Videodateien können wichtige Informationen leicht übersehen werden, da die Zusammenhänge zwischen verschiedenen Aufnahmen nicht erfasst werden. Dies führt dazu, dass relevante Beweismittel isoliert bleiben und die volle Rekonstruktion eines Ereignisses erschwert wird. Fehlt die systematische Verknüpfung, besteht die Gefahr, dass die Ermittlung fragmentiert verläuft und potenziell entscheidende Beweise ungenutzt bleiben. In komplexen Fällen, wie terroristischen Anschlägen oder groß angelegten Verbrechen, könnte dies die Strafverfolgung behindern und die Aufklärung von Straftaten verzögern oder sogar verhindern.

Um die Bedeutung und Anwendbarkeit der in dieser Arbeit vorgestellten Methoden zur Verknüpfung von Videodateien zu verdeutlichen, werden im Folgenden reale Ereignisse betrachtet, bei denen die forensische Analyse von Videodateien eine Rolle gespielt hat. Diese historischen Fallbeispiele zeigen die konkreten Herausforderungen und Anforderungen, die bei der Verarbeitung und Verknüpfung großer Datenmengen in der Videoforensik auftreten. Im Kapitel 5 wird nochmal Bezug auf diese Ereignisse genommen, um zu beschreiben, wie die gesammelten Videodateien sinnvoll visualisiert werden könnten, um einen Mehrwert für die Analyse zu schaffen.

3.1 Das Bostoner Marathon-Attentat

Abbildung 3.1: Bilder von Videoaufzeichnungen der Tsarnaev-Brüder. Quelle: [20]

Das Bostoner Marathon-Attentat, im Jahr 2013, war ein terroristischer Anschlag, der während des jährlichen Boston-Marathons stattfand. Am 15. April 2013 explodierten zwei Sprengsätze nahe der Ziellinie, wodurch drei Menschen getötet und über 200 weitere verletzt wurden. Die Ermittlungen ergaben, dass die Sprengsätze in Rucksäcken versteckt waren und von zwei Brüdern, Tamerlan und Dzhokhar Tsarnaev (siehe Abbildung 3.1), platziert wurden. Nach dem Angriff kam es zu einer groß angelegten Fahndung, die zu einer Schießerei mit den Verdächtigen und der Festnahme von Dzhokhar führte, während Tamerlan beim Schusswechsel getötet wurde. [21]

3.1.1 Ermittlungs- und kriminaltechnische Analyse

Die Ermittlungen nach dem Attentat wurden vom Federal Bureau of Investigation (FBI) geleitet. Am Ort des Anschlags konnten unzählige Beweismittel asserviert werden, welche in einer Lagerhalle gesammelt wurden. Darunter befanden sich unter anderem Akkus von Modelautos und Stücke von zwei Schnellkochtöpfen, was Bauteile einer Bombe sein können. Die Kochtöpfe konzentrierten hierbei die Explosion, um die maximale Auswirkung an Schaden zu erreichen. Um die Beweismittel zu analysieren und das Attentat weitestgehend rekonstruieren zu können, wurden zweimal täglich mit einem Flugzeug Beweismittel zum FBI geflogen. Unmittelbar nach dem Anschlag begannen die Ermittlungsbehörden außerdem mit dem Sammeln und Analysieren von Videoaufnahmen von Überwachungskameras und privaten Aufnahmen, die entscheidend für die schnelle Identifizierung der Tsarnaev-Brüder als Hauptverdächtige waren. Es konnten fast 13.000 verschiedene Videos und mehr als 120.000 Fotoaufnahmen gesammelt werden. 120 Analysten sollten die Aufnahmen nach Auffälligkeiten analysieren, wie beispielsweise Personen, die zum Zeitpunkt der Explosion keine Reaktion wie Erstaunen gezeigt haben. [22]

Die Videoforensik während der Untersuchung des Bostoner Marathon-Attentats umfasste eine intensive Analyse von Überwachungsaufnahmen und privaten Handyvideos, die eine zentrale Rolle bei der Identifizierung der Verdächtigen spielte. Ermittler mussten Terabytes an Videodateien in unterschiedlichen Formaten durchsehen, darunter kurze Handyvideos und stundenlange Aufnahmen von Überwachungskameras. Diese riesige Datenmenge erforderte den Einsatz spezialisierter Softwarelösungen, um die Daten effizient zu bewältigen. Videoanalyse-Tools halfen dabei, Videos zu komprimieren, Anomalien zu erkennen und relevante Details herauszufiltern. Ein wichtiger Schritt war die gezielte Analyse spezifischer Bereiche innerhalb der Aufnahmen, insbesondere nachdem der Ort der Bombendetonation identifiziert worden war. Eine Analysesoftware ermöglichte es, irrelevante Bewegungen auszublenden und Veränderungen, wie das Ablegen eines Rucksacks, schnell zu identifizieren. [23]

Das Ablegen des Rucksacks war eine Schlüsselaufnahme von einer Kamera des Geschäfts Lord & Taylor, die einen der Verdächtigen dabei zeigte, wo kurz darauf eine der Bomben explodierte [24]. Diese Information konnte durch eine Triage ermittelt werden, die es den Ermittlern erlaubte, relevante Videosequenzen schnell zu identifizieren und zu analysieren. In der Videoforensik bezeichnet der Begriff Triage den Prozess der schnellen Bewertung und Priorisierung von Videodateien, um die wichtigsten und relevantesten Informationen für eine weitere detaillierte Analyse herauszufiltern. Grundsätzlich mangelte es trotzdem an automatisierten Lösungen, um die Verdächtigen zu identifizieren und aufzuspüren. Der massive Zustrom von Mediendateien zeigte die Herausforderungen, die mit der Verarbeitung und Analyse einer solchen Datenflut verbunden waren. Die Herausforderung bestand nicht nur in der Menge des zu analysierenden Materials, sondern auch in der Notwendigkeit, dieses effizient zu betrachten, um relevante Informationen schnell zu extrahieren. [25]

Die erfolgreiche Aufklärung des Bostoner Marathon-Attentats zeigt, wie entscheidend die systematische Verknüpfung und Analyse von Videodateien in der modernen forensischen Arbeit ist. Ohne diese Technologien wäre es kaum möglich gewesen, die riesigen Datenmengen effizient zu verarbeiten und die Verdächtigen in kurzer Zeit zu identifizieren. Deshalb sollte man dieses Themenfeld weiter intensiv erforschen und entwickeln, da die Verknüpfung von Videodateien ein enormes Potenzial besitzt, um die Aufklärung von Straftaten noch effektiver und präziser zu gestalten.

3.2 Der Vorfall auf dem Berliner Weihnachtsmarkt

Abbildung 3.2: Bild vom Fahndungsaufruf zu Anis Amri. Quelle: [26]

Am 19. Dezember 2016 verübte Anis Amri (siehe Abbildung 3.2), ein 24-jähriger Tunesier, einen Terroranschlag in Berlin. Er steuerte einen entführten Lastwagen durch die Menschenmenge auf dem Weihnachtsmarkt am Breitscheidplatz, direkt neben der Kaiser-Wilhelm-Gedächtniskirche. Dieser Vorfall resultierte in 12 Todesopfern und 56 Verletzte. Der Islamische Staat (IS) hat den Anschlag für sich beansprucht, was auf ein terroristisch motiviertes Vorhaben hindeutet. Nach dem Anschlag floh Amri aus Deutschland und reiste durch mehrere europäische Länder. Am 23. Dezember 2016 wurde er von der Polizei in Italien gestellt und nach einem Schusswechsel getötet. Seine Flucht und der Tod in Italien markierten das Ende einer intensiven, grenzüberschreitenden Fahndung. [27]

3.2.1 Ermittlungs- und kriminaltechnische Analyse

Der Fall von Anis Amri hatte ermittlungs- und kriminaltechnisch weite Ausmaße. Bereits vor dem Attentat wurde der Einsatz von verdeckten Ermittlern und Informanten intensiv genutzt, um tiefergehende Einblicke in Amris Netzwerk und seine Planungen zu gewinnen. Diese Techniken waren entscheidend, um die Bedrohung durch Amri besser zu verstehen und präventive Maßnahmen zu ergreifen. Die Überwachung von Amris Telekommunikation spielte ebenfalls eine zentrale Rolle in den Ermittlungen. Die Auswertung seiner Telefonate, Chats und Internetaktivitäten, einschließlich seiner Versuche, Informationen über den Bau einer Bombe zu erlangen, lieferte wertvolle Erkenntnisse über seine Absichten und Fähigkeiten. Nach dem Attentat wurden umfangreiche forensische Untersuchungen am Tatort und am verwendeten Lastwagen durchgeführt. Diese Analysen dienten dazu, Beweise zu sichern und den genauen Ablauf des Anschlags zu rekonstruieren. Die Ergebnisse dieser Untersuchungen waren entscheidend für das Verständnis der durchgeführten Tat und die Identifikation von Sicherheitslücken. Die Videoüberwachung spielte eine wesentliche Rolle bei der Rekonstruktion der Ereignisse. Die Auswertung der Überwachungskameraaufnahmen vom Breitscheidplatz und anderer relevanter Orte ermöglichte es den Ermittlern, Amris Bewegungen vor und nach dem Anschlag genau zu verfolgen und zu dokumentieren. [28]

Nach dem Terroranschlag auf dem Berliner Breitscheidplatz erhielt das Bundeskriminalamt über ein spezielles Internetportal, genannt „Boston Cloud“, mehr als 600 Hinweise in Form von Fotos und Videos von der Bevölkerung. Obwohl die Plattform kurzzeitig von einem Hackerangriff betroffen war, umfasste das gesammelte Material neben relevanten Hinweisen auch irrelevante Inhalte wie beispielsweise Katzenvideos. Die meisten Aufnahmen, die ausgewertet wurden, stammten jedoch aus Überwachungskameras. [29]

Die systematische Auswertung und Verknüpfung der Videodateien spielte eine zentrale Rolle bei der Rekonstruktion der Ereignisse und der Nachverfolgung von Amris Bewegungen. Durch die gezielte Verknüpfung der Aufnahmen aus Überwachungskameras an verschiedenen Orten konnten die Ermittler Amris Aufenthaltsorte und seine Handlungen vor und nach dem Anschlag präzise nachvollziehen. Dies ermöglichte nicht nur eine lückenlose Dokumentation seiner Aktivitäten, sondern lieferte auch wertvolle Hinweise für die Identifikation von Sicherheitslücken und die Verbesserung zukünftiger Überwachungs- und Ermittlungsmethoden. Besonders in Fällen wie diesem, in denen eine große Menge an Videomaterial von der Öffentlichkeit eingereicht wird, ist es entscheidend, irrelevante Inhalte wie beispielsweise Katzenvideos schnell herauszufiltern. Diese schnelle und effiziente Trennung von nützlichen und unnötigen Informationen spart wertvolle Zeit und Ressourcen, die für die Analyse der relevanten Daten benötigt werden. Der Fall zeigte, wie wichtig die Verknüpfung von Videodateien in der modernen Terrorismusbekämpfung ist. Die Fähigkeit, relevante Informationen aus einer Vielzahl von Quellen schnell und effizient zusammenzuführen, war entscheidend für den Erfolg der Ermittlungen. Ohne diese Technologien wäre es deutlich schwieriger gewesen, den Täter zu identifizieren und seine Bewegungen nachzuvollziehen und ihn schlussendlich fassen zu können.

3.3 Der NSU-Fall

Abbildung 3.3: Beate Zschäpe, Uwe Mundlos und Uwe Böhnhardt. Quelle: [30]

Der NSU war eine rechtsextreme Terrorzelle in Deutschland, die zwischen 2000 und 2007 zehn Morde, drei Bombenanschläge und fünfzehn Banküberfälle verübte. Die Mitglieder Uwe Böhnhardt, Uwe Mundlos und Beate Zschäpe (siehe Abbildung 3.3) lebten von 1998 bis 2011 im Untergrund und erhielten dabei Unterstützung von einem Netzwerk aus Neo-Nazis und Informanten aus dem Sicherheitsdienst. Die Taten des NSU blieben lange unentdeckt, da die Ermittlungen zunächst in die falsche Richtung gingen und die Opfer fälschlicherweise mit kriminellen Milieus in Verbindung gebracht wurden. Erst nachdem sich Zschäpe 2011 den Behörden stellte, konnte der Zusammenhang zwischen den Morden und der Terrorgruppe aufgedeckt werden. Dies führte zu einer tiefgreifenden Untersuchung der Versäumnisse der deutschen Sicherheitsbehörden, die zuvor Hinweise auf rechtsextreme Hintergründe ignoriert hatten. [31]

3.3.1 Ermittlungs- und kriminaltechnische Analyse

Zu den angewandten Techniken zählten DNA-Analysen, die an den Tatorten gesicherte Spuren auswerteten, um die Anwesenheit der Täter zu beweisen und eine Verbindung zwischen den verschiedenen Tatorten herzustellen. Ballistische Untersuchungen unterstützten diese Befunde, indem sie zeigten, dass dieselben Waffen bei mehreren Morden verwendet wurden, was half, die Fälle miteinander zu verknüpfen. [32]

Die Relevanz von Videos im NSU-Fall war besonders hoch, da sie wesentliche Beweise lieferten, die zur Identifikation und späteren Verurteilung der Mitglieder des NSUs beitrugen. Eines der wichtigsten Videos war das sogenannte Bekennervideo, in dem die NSU ihre Verantwortung für eine Reihe von Morden, Bombenanschlägen und Banküberfällen übernahm. Dieses Video enthielt auch spöttische Darstellungen der Opfer und zeigte Ausschnitte von den Tatorten, was nicht nur zur Aufklärung der Taten beitrug, sondern auch die Gesinnung der Gruppe offenbarte. Darüber hinaus spielten Überwachungsvideos eine entscheidende Rolle bei der Rekonstruktion der Bewegungen und Aktivitäten der NSU-Mitglieder vor und nach den Verbrechen. Diese Aufnahmen wurden in verschiedenen Phasen der Ermittlung verwendet, um Verbindungen zwischen den Tätern und den Tatorten herzustellen. [31]

Der NSU-Fall zeigt, wie Videoaufnahmen nicht nur zur Täteridentifikation beitragen, sondern auch tiefere Einblicke in die ideologischen Hintergründe und die Brutalität der Taten bieten können. Das Bekennervideo der NSU war ein zentrales Beweismittel, das die rechtsextreme Motivation der Taten verdeutlichte und damit half, die öffentliche Wahrnehmung und die politischen Reaktionen auf die Verbrechen zu formen. Mit evtl. vorhandenen Videoaufzeichnungen von den Anschlägen könnte man nun gezielt die darauf abgebildeten Personen identifizieren und durch die Verknüpfung der Aufnahmen ihre Bewegungen verfolgen. Dies würde nicht nur die Nachverfolgung der Aktivitäten der bekannten Täter erleichtern, sondern könnte auch zur Identifikation weiterer Tatverdächtiger führen, die bisher unentdeckt geblieben sind.

3.4 Zusammenfassung

Die Betrachtung der drei Ereignisse – das Bostoner Marathon-Attentat, der Anschlag auf den Berliner Weihnachtsmarkt und der NSU-Fall – verdeutlicht die Bedeutung der Verknüpfung und Analyse von Videodateien in der modernen forensischen Arbeit. Diese Fälle zeigen, wie essenziell die systematische Auswertung von Videomaterial für die Aufklärung komplexer Straftaten ist. Durch die gezielte Verknüpfung von Videodateien können Ermittler wichtige Zusammenhänge erkennen, die ansonsten möglicherweise unentdeckt bleiben würden.

Im Fall des Bostoner Marathon-Attentats erwies sich die Verwendung von spezialisierten Videoanalyse-Tools als entscheidend, um die riesigen Mengen an gesammeltem Videomaterial effizient zu bewältigen und die Verdächtigen schnell zu identifizieren. Die gezielte Analyse spezifischer Bereiche in den Aufnahmen ermöglichte es, relevante Informationen herauszufiltern und die Bewegungen der Täter nachzuverfolgen. Ähnlich zeigte der Anschlag auf den Berliner Weihnachtsmarkt, wie die systematische Verknüpfung von Videoaufnahmen aus Überwachungskameras eine zentrale Rolle bei der Nachverfolgung von Amris Bewegungen vor und nach dem Anschlag spielte. Die schnelle Filterung irrelevanter Inhalte, wie etwa von der Öffentlichkeit eingereichte Katzenvideos, war entscheidend, um den Fokus auf die wichtigen Beweismittel zu legen und die Ermittlungen zu beschleunigen. Auch im NSU-Fall zeigte sich, dass Videoaufnahmen, insbesondere das Bekennervideo, entscheidend waren, um die ideologischen Hintergründe der Taten aufzudecken und die öffentliche Wahrnehmung zu beeinflussen. Die Verknüpfung weiterer potenzieller Videoaufzeichnungen hätte möglicherweise zur Identifikation zusätzlicher Tatverdächtiger geführt.

In allen drei Fällen hätte eine noch bessere Verknüpfung der Videodateien die Effizienz der Ermittlungen weiter steigern können und möglicherweise zu einer schnelleren Aufklärung beigetragen. Insbesondere die Visualisierung der verknüpften Daten bietet ein enormes Potenzial. Durch die Verwendung interaktiver Karten, Zeitleisten oder Netzwerkgrafiken könnten Ermittler komplexe Zusammenhänge schneller erkennen und nachvollziehen. Diese Ereignisse verdeutlichen, dass die Weiterentwicklung von Technologien zur Verknüpfung und Visualisierung von Videodateien wichtig ist. Eine verbesserte automatisierte Analyse könnte nicht nur die Nachverfolgung von Täterbewegungen erleichtern, sondern auch die Identifikation weiterer Tatverdächtiger ermöglichen. Diese Optimierung wird zunehmend realisierbar, da sich Algorithmen stetig weiterentwickeln und die Rechenleistung moderner Computer kontinuierlich verbessert wird. Dadurch eröffnet sich die Möglichkeit, Straftaten noch effektiver und präziser aufzuklären und die Sicherheit der Bevölkerung nachhaltig zu erhöhen.

4 Methoden zur Verknüpfung von Videodateien

Die effiziente Verknüpfung von Videodateien auf Metadatenebene ist ein Aspekt der modernen Videoforensik, der positive Auswirkungen auf die Aufklärung von Straftaten haben kann. Dies wurde verdeutlicht durch die Betrachtung der vorangehenden Ereignisse. Dieses Kapitel orientiert sich am schematischen Ablauf, der im Abschnitt 4.1 beschrieben wurde. Es werden zunächst Methoden und zur Extraktion von Metadaten, anschließend Modelle zur Analyse von Videodateien und dann einige Möglichkeiten zur Verknüpfung von Videodateien vorgestellt. Die Datenerfassung wird nicht ausführlich beschrieben, sondern nur kurz erläutert. Unterschiedliche Methoden, die zur Verknüpfung von Videodateien eingesetzt werden können, werden vorgestellt und beleuchtet. Hierbei handelt es sich sowohl etablierte als auch innovative Technologien, die in diesem Bereich Anwendung finden. Von grundlegenden technischen Ansätzen bis hin zu fortschrittlichen Verfahren unter Einsatz von Algorithmen des maschinellen Lernens und künstlicher Intelligenz werden die verschiedenen Techniken dargestellt, die es ermöglichen, Metadaten aus Videodateien zu extrahieren, zu analysieren und anschließend zu verbinden. Diese Verknüpfung verbessert nicht nur die Möglichkeit, relevante Informationen aus großen Mengen an Videomaterial effizient zu filtern und zu interpretieren, sondern trägt auch dazu bei, komplexe Sachverhalte und Zusammenhänge innerhalb der gesammelten Daten zu erkennen.

4.1 Extraktion von Metadaten

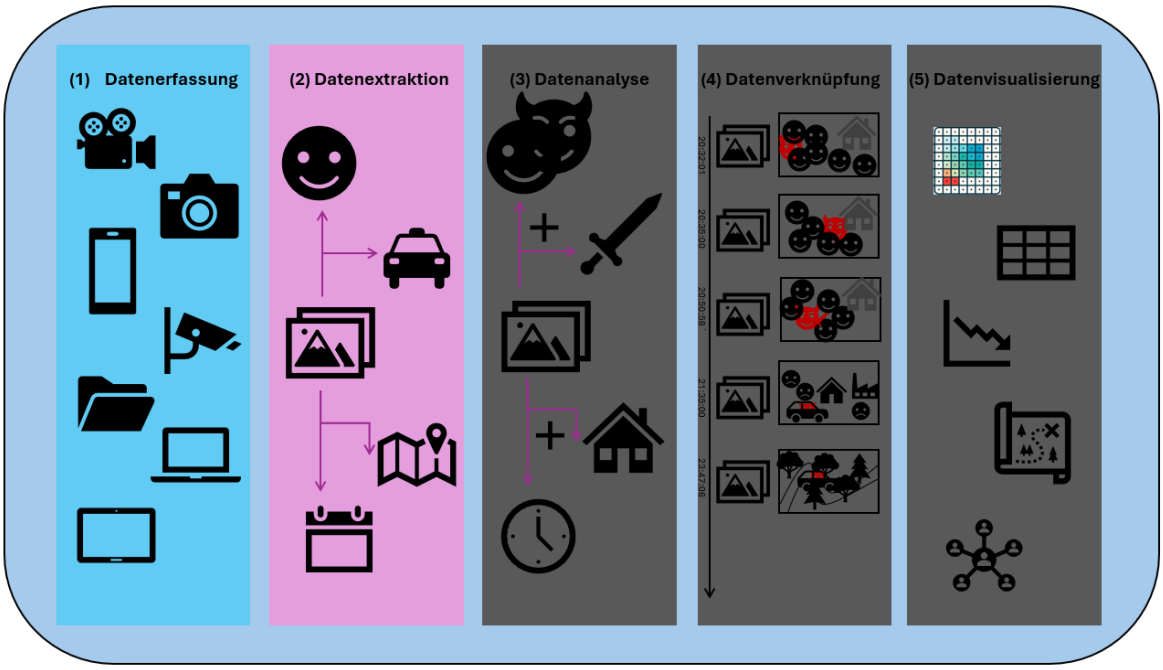

Abbildung 4.1: Schematischer Ablauf - Datenerfassung und -extraktion.

In diesem Abschnitt wird kurz erläutert, wie Videodateien erfasst werden können und anschließend ausführlich beschrieben, wie Metadaten methodisch aus Videodateien extrahiert werden, welche Herausforderungen dabei auftreten und wie diese Daten zur Unterstützung einer forensischen Untersuchung genutzt werden können. Im bereits vorgestellten schematischen Ablauf befinden wir uns bei in ersten beiden Teilen (siehe Abbildung 4.1).

Zunächst muss eine Datenerfassung erfolgen, um die Videodateien zu sammeln. Diese gesammelten Daten bilden die Grundlage für die Extraktion von Metadaten, die technische Details wie Datum, Uhrzeit, Kameramodell und GPS-Koordinaten sowie inhaltsbasierte Informationen wie erkannte Personen oder Objekte umfassen können. Möglichkeiten, Videodateien mit Metadaten zu erlangen, sind:

- **Überwachungskameras**: Aufzeichnungen von Überwachungskameras in öffentlichen oder privaten Bereichen.
- **Mobile Endgeräte**: Videos, die mit Smartphones oder Tablets aufgenommen wurden.
- **Drohnenaufnahmen**: Videos, die von Drohnen während Überwachungsflügen oder Einsätzen aufgenommen wurden.
- **Bodycams**: Aufzeichnungen von Körperkameras, die von Polizei oder Sicherheitspersonal getragen werden.
- **Dashcams**: Videomaterial von Kameras, die in Fahrzeugen installiert sind.

EXIF-Metadaten sind standardisierte Informationen, die direkt in eine Datei eingebettet sind und technische Details über die Aufnahme eines Videos oder Fotos enthalten. Dazu gehören Daten wie Kameratyp, Datum und Uhrzeit der Aufnahme, Blendenöffnung, Belichtungsdauer, ISO-Wert, und, wenn verfügbar, GPS-Daten. Diese Informationen sind besonders nützlich, um die Authentizität und Herkunft eines Videos zu überprüfen, was in forensischen Untersuchungen von großer Bedeutung sein kann. [7]

Inhaltsbasierte Metadaten dagegen werden durch die Analyse des Inhalts des Videos selbst gewonnen und umfassen Daten über die im Video erkannten Objekte, Personen, Texte oder andere visuelle und akustische Merkmale. Diese Art von Metadaten wird häufig durch fortschrittliche Technologien wie maschinelles Lernen und künstliche Intelligenz erzeugt, die in der Lage sind, komplexe Muster und Merkmale in den Videodateien zu erkennen und zu interpretieren. Inhaltsbasierte Metadaten sind entscheidend für Aufgaben wie die Ereignisrekonstruktion und das Verständnis von Verhaltensmustern, da sie es ermöglichen, Verbindungen zwischen unterschiedlichen Videos zu erkennen, die über bloße technische Daten hinausgehen.

Die Extraktion und Analyse dieser Metadaten sind von großem Wert für forensische Zwecke. Sie ermöglichen es Ermittlern, das Videomaterial präzise zu datieren, dessen Ursprung zu bestimmen und zu authentifizieren, sowie den Kontext zu verstehen, in dem die Aufnahmen entstanden sind. In der Praxis können Metadaten verwendet werden, um verschiedene Videodateien effektiv miteinander zu verknüpfen und so ein umfassenderes Bild eines Ereignisses oder einer Straftat zu rekonstruieren.

4.1.1 EXIF-Metadaten

Die Speicherung und Analyse von Metadaten in Dateien folgt einem Standard namens EXIF [33], der wichtig für die Verknüpfung von Videodateien auf Metadatenebene ist. Dieser von der Japan Electronic Industries Development Association (JEIDA) entwickelte Standard wurde mehrfach überarbeitet, um technologische Entwicklungen zu integrieren und die Anforderungen der digitalen Bildverarbeitung zu erfüllen. Die vier Versionen, die insbesondere für die Bild- und Videoforensik von Relevanz waren, sind EXIF 2.1, EXIF 2.2, EXIF 2.3 und EXIF 3.0.

Die Version EXIF 2.1, das im Jahr 1998 eingeführt wurde, definierte grundlegende Metadaten wie Kamerainformationen, Aufnahmeeinstellungen und Datumsstempel. Diese Informationen sind essenziell für die Identifikation und Analyse von Bildern in der digitalen Forensik. Mit EXIF 2.2, das 2002 eingeführt wurde, erweiterte sich der Standard um Felder, die auf die Optimierung von Druckanwendungen ausgerichtet sind, einschließlich spezifischerer Farbmanagementinformationen. EXIF 2.3, das im Jahr 2009 veröffentlicht wurde, verbesserte die Integration von GPS-Daten erheblich, was die Genauigkeit der Geolokalisierung von Aufnahmen verbessert. Die aktuelle Version, EXIF 3.0, das 2023 herausgebracht wurde, stellt einen bedeutenden Fortschritt dar, indem es die Unterstützung für die Unicode Transformation Format (UTF)-8-Kodierung einführt. Dies ermöglicht die Verwendung von Unicode-Zeichen in den Metadaten und erhöht somit die globale Zugänglichkeit und Nützlichkeit des Standards. [33]

Metadaten lassen sich in fünf Hauptkategorien unterteilen, die jeweils spezifische Aspekte der aufgezeichneten Metadaten darstellen und organisieren. Die fünf Hauptkategorien sind *allgemeine Informationen*, *bildspezifische Daten*, *Tonwiedergabe*, *weitere technische Daten* und *Thumbnail-spezifische Informationen*. Es werden nachfolgend exemplarisch einige EXIF-Daten vorgestellt, die von Relevanz sein können, da nicht alle den gleichen Wert für die forensische Analyse haben. [33]

Allgemeine Informationen - Softwareversion:
Dieses Feld in den Metadaten gibt die Version der Software an, die zur Bearbeitung oder Erzeugung des Bildes verwendet wurde. Anhand der angegebenen Softwareversion kann man auch ein frühestes Datum der Erstellung der Datei ableiten, da die diese erst nach dem entsprechenden Release verwendet worden sein kann. Wenn man die Timestamps der Bildaufnahme mit der angegebenen Softwareversion verknüpft und feststellt, dass das Release-Datum der Softwareversion nach den Timestamps liegt, entsteht ein Widerspruch. Ein solcher Widerspruch kann darauf hindeuten, dass die Metadaten manipuliert wurden oder dass das Bild zu einem anderen Zeitpunkt oder mit einer anderen Softwareversion bearbeitet wurde als angegeben. Dies ist ein wertvoller Aspekt in der forensischen Analyse, um die Authentizität und Integrität von Bilddaten zu überprüfen.

Bildspezifische Daten - Kamera-Hersteller und Modell:
Im EXIF-Standard beziehen sich diese Felder auf die spezifischen Metadaten, die den Hersteller und das genaue Modell der Kamera angeben, mit der das Bild aufgenommen wurde. Diese Informationen werden automatisch von der Kamera in die Bilddatei eingebettet. Wenn mehrere Bilder denselben Hersteller und dasselbe Kameramodell aufweisen, kann dies darauf hindeuten, dass sie alle von demselben Gerät stammen, was in forensischen Untersuchungen dazu beitragen kann, eine gemeinsame Herkunft zu identifizieren. Darüber hinaus können diese Informationen mit anderen Metadaten wie Zeitstempeln, Standortdaten oder Softwareversionen kombiniert werden, um ein umfassenderes Bild der Entstehungsgeschichte von Bildern zu zeichnen.

Bildspezifische Daten - Datum und Uhrzeit:
Eine der wichtigsten Metadaten sind die Zeitstempel und das Datum einer Aufnahme. Im EXIF-Standard beziehen sich diese Felder auf die Angaben zum Datum und zur Uhrzeit, zu denen ein Bild aufgenommen wurde. Diese Informationen werden automatisch von der Kamera generiert und in den Metadaten des Bildes gespeichert, was eine genaue zeitliche Einordnung der Aufnahme ermöglicht. In der forensischen Analyse sind diese Zeitstempel besonders entscheidend, um die Reihenfolge der Bildaufnahmen zu rekonstruieren und Unstimmigkeiten zu erkennen, die auf Manipulationen hinweisen könnten, wie etwa falsche Zeitstempel im Vergleich zu anderen Metadaten oder bekannten

Ereignissen. Auch für die Verknüpfung von Metadaten spielen diese Werte eine große Rolle. Wenn mehrere Bilder denselben Zeitstempel oder einen engen Zeitrahmen aufweisen, können diese Bilder miteinander verknüpft werden, was auf eine gemeinsame Aufnahmeumgebung oder ein Ereignis hindeuten kann. Zudem ermöglicht die Verknüpfung von Zeitstempeln mit anderen Metadaten, wie dem Standort, eine Rekonstruktion der Abfolge von Ereignissen.

Bildspezifische Daten - Standortdaten:
Auch die Standortdaten sind von sehr großer Bedeutung. Sie beziehen sich auf die geografischen Koordinaten, die in den Metadaten eines Bildes gespeichert werden und den genauen Ort der Aufnahme festhalten. Diese Informationen werden oft von Kameras oder Smartphones mit GPS-Funktionalität automatisch erfasst und umfassen in der Regel Längen- und Breitengrade. In der forensischen Analyse sind Standortdaten besonders nützlich, um die Herkunft eines Bildes zu bestätigen oder zu widerlegen. Wenn die Standortdaten nicht mit dem erwarteten Aufnahmeort übereinstimmen oder mit anderen Bildern derselben Serie in Konflikt stehen, könnte dies auf eine Manipulation oder Fälschung der Metadaten hinweisen. Der Wert dieser Daten für die Verknüpfung von Metadaten ist sehr hoch. Standortdaten ermöglichen es, Bilder geografisch zu verorten und somit Bilder miteinander zu verknüpfen, die am selben Ort oder in geografischer Nähe zueinander aufgenommen wurden. Besonders interessant wird es, wenn man die GPS-Daten mit den Zeitstempeln verknüpft. Dadurch lassen sich Bewegungsprofile erstellen, die zeigen, wann und wo sich eine Person oder ein Objekt bewegt hat. Gleichzeitig ermöglicht diese Verknüpfung auch die Erkennung von Manipulationen: Wenn die dargestellte Tageszeit im Zeitstempel nicht mit den Lichtverhältnissen oder dem erwarteten Umfeld in den GPS-Daten übereinstimmt, könnte dies auf eine Fälschung der Metadaten hinweisen. Solche Inkonsistenzen sind wichtige Indikatoren für die Überprüfung der Authentizität.

Thumbnail-spezifische Informationen - Thumbnail-Daten:
Thumbnail-Daten sind kleine, komprimierte Version eines aufgenommenen Bildes, die in den EXIF-Metadaten eingebettet ist. Diese Thumbnails dienen in erster Linie der schnellen Vorschau des Bildes, ohne dass das vollständige, oft viel größere Originalbild geladen werden muss. Sie werden von der Kamera oder der verwendeten Software automatisch erzeugt und im EXIF-Header der Bilddatei gespeichert. Videodateien können auch Thumbnail-Daten enthalten, hier werden Thumbnails jedoch oft als separate Metadaten oder Vorschauen eingebettet, die in Form von Standbildern aus dem Video generiert werden. Der Wert dieser Thumbnail-Daten liegt in ihrer Fähigkeit für die Verknüpfung von Metadaten, einen zusätzlichen Anhaltspunkt für die Authentizität und Integrität des Bildes zu bieten. Da das Thumbnail direkt aus der Originaldatei generiert und eine Vorschau darstellt, kann es genutzt werden, um Veränderungen oder Manipulationen am Hauptbild zu erkennen. Zudem können Thumbnails bei der Verknüpfung von Videodateien helfen, indem sie es ermöglichen, schnell zu überprüfen, ob zwei Aufnahmen tatsächlich identisch oder zumindest sehr ähnlich sind, ohne aufwendige Analysen der großen Originaldateien durchführen zu müssen. Dies ist besonders nützlich, wenn man große Mengen an Daten durchsuchen muss.

4.1.1.1 Speichern der EXIF-Metadaten

In Bildformaten wie Joint Photographic Experts Group (JPEG) und Tag Image File Format (TIFF) sind EXIF-Daten weit verbreitet. JPEG ist dabei das am meisten genutzte Format für digitale Fotos und speichert umfassende Informationen, die in der Fotografie elementar sind, von Kameraeinstellungen bis hin zu Geo-Tagging-Informationen. TIFF wird oft in der professionellen Fotografie verwendet und

unterstützt ebenfalls detaillierte EXIF-Metadaten, die eine hohe Bildqualität ohne Datenverlust bieten. Verschiedene herstellerspezifische Dateiformate wie beispielsweise CR2 (Canon) und NEF (Nikon) speichern Bilder direkt vom Sensor und behalten umfangreiche EXIF-Informationen für Nachbearbeitung und Bildanalyse.

Bei Videoformaten ist die Einbettung von EXIF-ähnlichen Metadaten weniger verbreitet. Formate wie MOV und MP4 können Metadaten aufnehmen, die über die Aufnahmebedingungen Auskunft geben, was bei der Bearbeitung und Analyse von Videoinhalten hilfreich ist. Auch das AVI-Format, entwickelt von Microsoft, kann in begrenztem Umfang technische Metadaten speichern, die jedoch nicht so detailliert oder standardisiert sind wie bei den zuvor genannten Formaten.

Jedoch beinhalten nicht alle Mediendateien EXIF-Daten, was verschiedene Ursachen haben kann, die sowohl technischer als auch benutzerdefinierter Natur sein können. Benutzer oder Software können EXIF-Daten bewusst entfernen, um sensible Informationen wie den Standort oder das Aufnahmedatum zu schützen, bevor Bilder oder Videos online gestellt werden. Einige Kameras oder Mobiltelefone bieten auch die Möglichkeit, das Speichern von EXIF-Daten zu deaktivieren, eine Funktion, die von datenschutzbewussten Nutzern bevorzugt wird. Darüber hinaus können Bildbearbeitungsprogramme beim Bearbeiten oder Konvertieren von Bildern in ein anderes Format EXIF-Daten entfernen oder verändern. Ältere oder weniger fortschrittliche Kameratechnologien haben außerdem nicht die Fähigkeit, umfangreiche Metadaten zu erfassen und zu speichern. Zusätzlich entfernen viele Online-Plattformen und soziale Medien aus Sicherheits- und Datenschutzgründen EXIF-Daten standardmäßig beim Hochladen von Bildern [34].

4.1.1.2 Extraktion der EXIF-Metadaten

Um Videodateien miteinander verknüpfen zu können, müssen zunächst die Metadaten extrahiert werden. Dieses Vorgehen kann abhängig von der Skalierung der Anforderungen unterschiedlich angegangen werden. Für die Analyse einer einzelnen Datei bieten spezialisierte Softwarelösungen wie ExifTool[1] einen direkten und effektiven Weg, um detaillierte Metadaten zu gewinnen. Benutzer können dieses Tool über einfache Befehle in einer Kommandozeilenumgebung nutzen, um Zugang zu einer Vielzahl von Informationen zu erhalten, die von technischen Details wie Kameraeinstellungen und Dateiformaten bis hin zu GPS-Daten reichen, sofern diese verfügbar sind.

In Szenarien, in denen große Mengen von Videodateien verarbeitet werden müssen, wie es oft bei umfangreichen forensischen Untersuchungen der Fall ist, erfordert die Situation einen automatisierten und skalierbaren Ansatz. Hier kommt die Batch-Verarbeitungsfähigkeit von ExifTool ins Spiel, die es ermöglicht, Metadaten aus einer Vielzahl von Dateien effizient zu extrahieren. Durch die Erstellung von Batch-Skripten können Nutzer ganze Verzeichnisse durchsuchen und Metadaten in organisierten Formaten wie Comma-separated values (CSV) für weitere Analysen speichern. Dies reduziert den manuellen Aufwand erheblich und kann eine schnelle Durchsicht großer Datenmengen ermöglichen.

[1] https://exiftool.org/

Für Entwickler, die spezifische Anforderungen haben oder in einem kontrollierten Umfeld arbeiten, bieten Programmierbibliotheken wie Pillow[2] in Python die Möglichkeit, maßgeschneiderte Skripte zu erstellen. Diese Skripte können EXIF-Daten aus Bildern extrahieren und analysieren, was besonders nützlich ist, wenn bestimmte Metadatenfelder für die forensische Untersuchung oder die Datenaufbereitung erforderlich sind. Die Flexibilität der Programmierung ermöglicht es, tiefergehende Analysen durchzuführen und Daten in einer Weise zu verarbeiten, die mit generischen Tools nicht möglich wäre.

Cloud-basierte Dienste wie AWS Rekognition[3] oder Google Cloud Video Intelligence[4] bieten eine weitere Dimension der Datenverarbeitung, indem sie die Möglichkeiten der Cloud-Nutzung ausschöpfen. Diese Dienste können nicht nur Standardmetadaten extrahieren, sondern auch fortschrittliche Inhaltserkennungen durchführen, wie das Erkennen von Gesichtern, Objekten und Szenen innerhalb von Videos. Solche Technologien sind besonders sinnvoll in Umgebungen, in denen die Menge der zu analysierenden Daten die Kapazitäten lokaler Systeme übersteigt und wo Echtzeitanalysen gefordert sind. Es ist jedoch kritisch zu hinterfragen, ob man Daten eines Strafprozesses auf einer Cloud hochladen darf und möchte.

4.1.2 Inhaltsbasierte Metadaten

Inhaltsbasierte Metadaten sind spezielle Daten, die Informationen über den Inhalt von Mediendateien wie Videos und Bilder enthalten. Diese können Objekte, Personen, Texte, Handlungen oder andere erkennbare Elemente innerhalb der Medien umfassen. Im Gegensatz zu technischen Metadaten, die Informationen über die Datei selbst liefern, beziehen sich inhaltsbasierte Metadaten direkt auf die visuellen oder akustischen Inhalte der Medien.

Ein Beispiel für eine funktionierende inhaltliche Verknüpfung von Objekten und Mediendateien ist die automatische Speicherung von inhaltsbasierten Metadaten durch moderne Geräte, wie zum Beispiel iPhones. Gibt man in der Fotogalerie in der Suche einen Begriff wie Katze ein, so erscheinen hier fast ausschließlich Treffer mit Katzen. Diese Geräte erfassen und speichern Informationen zu erkannten Gesichtern und Objekten durch die integrierte Software. Diese Automatisierung erleichtert den Benutzern das Verwalten und Wiederfinden von Medien basierend auf ihrem Inhalt und Kontext, was die Relevanz und Zugänglichkeit der aufgenommenen Medien erheblich steigert. [35]

Eine ähnliche Automatisierung und inhaltsbasierte Verknüpfung von Metadaten sollte auch in der forensischen Analyse Einzug halten, da sie einen erheblichen Mehrwert bieten würde. In der Forensik geht es oft darum, große Mengen an digitalen Beweisen effizient zu sichten und relevante Verbindungen zwischen verschiedenen Dateien herzustellen. Inhaltsbasierte Metadaten, die automatisch durch fortschrittliche Algorithmen generiert werden, könnten es Ermittlern ermöglichen, schneller und präziser Zusammenhänge zwischen verschiedenen Bildern und Videos zu erkennen. Dies würde nicht nur die Durchsuchung und Analyse beschleunigen, sondern auch die Wahrscheinlichkeit erhöhen, kritische Hinweise zu entdecken, die sonst möglicherweise übersehen worden wären.

[2] https://python-pillow.org/
[3] https://aws.amazon.com/de/rekognition/
[4] https://cloud.google.com/video-intelligence

4.1.2.1 Speichern inhaltsbasierte Metadaten

Die Speicherung von inhaltsbasierten Metadaten kann auf verschiedene Arten erfolgen, um die Anforderungen von Medienverwaltungssystemen, Archivierungsprozessen und analytischen Anwendungen zu erfüllen. Diese Metadaten werden häufig direkt in den Mediendateien eingebettet oder in begleitenden Datenbanken gespeichert.

Eine gängige Methode zur Speicherung dieser Metadaten ist die Einbettung direkt in die Mediendateien selbst, wobei Standards wie EXIF, International Press Telecommunications Council (IPTC) und Extensible Metadata Platform (XMP) genutzt werden. EXIF ist vor allem bekannt für die Speicherung von Daten in Bildern, die von digitalen Kameras aufgenommen wurden, und kann Informationen wie Datum, Uhrzeit und Kamerainformationen beinhalten. IPTC und XMP sind weiter gefasst und ermöglichen die Speicherung einer Vielzahl von deskriptiven Daten, die über die technischen Spezifikationen hinausgehen, wie Keywords, Beschreibungen und Urheberrechtshinweise. Diese Metadatenstandards sind in der Medienbranche besonders nützlich, da sie die Portabilität der Dateien gewährleisten, was die Organisation und Suche erleichtert, unabhängig davon, wo und wie die Dateien gespeichert sind.

Der IPTC-Standard [36] umfasst eine Reihe von Spezifikationen für die Einbettung von Metadaten in digitale Medien. Diese Metadaten können eine breite Palette von Informationen enthalten, darunter Titel, Beschreibung, Urheber, Erstellungsdatum und -ort sowie weitere spezifische Informationen über das Urheberrecht. Dieser Standard wird häufig von Fotografen und Medienorganisationen genutzt, um sicherzustellen, dass wichtige Informationen über das Bildmaterial wie die Quelle, das Aufnahmedatum und andere relevante Kontextinformationen zusammen mit den Bildern gespeichert und verbreitet werden. Der große Vorteil von IPTC ist seine Fähigkeit, eine Vielzahl von Daten in einer strukturierten Form zu speichern, die von vielen Bildbearbeitungs- und Verwaltungssystemen direkt unterstützt wird.

XMP ist ein von Adobe Systems entwickelter Standard, der eine erweiterbare Plattform für die Erstellung, Verarbeitung und den Austausch von standardisierten und benutzerdefinierten Metadaten bietet. XMP ist in Extensible Markup Language (XML) formatiert und kann in die Dateien selbst eingebettet oder als externe Datei gespeichert werden. Der XMP-Standard ist flexibler als IPTC, da er nicht nur vorgefertigte Felder verwendet, sondern auch die Erstellung benutzerdefinierter Felder erlaubt, was ihn besonders vielseitig macht. Er kann in einer Vielzahl von Dateitypen verwendet werden, einschließlich JPEG, TIFF und vielen Formaten von Adobe-Anwendungen wie Photoshop, Illustrator und InDesign. Eine der Stärken von XMP ist seine Fähigkeit, über unterschiedliche Dateitypen und Medienformate hinweg konsistent zu funktionieren, was die Interoperabilität zwischen verschiedenen Anwendungen und Diensten verbessert.

Für umfangreichere Mediensammlungen oder in Umgebungen, in denen schneller Zugriff und komplexe Abfragen erforderlich sind, werden inhaltsbasierte Metadaten oft in Datenbanken gespeichert. Relationale Datenbanken können verwendet werden, um strukturierte Informationen effizient zu speichern und zu verwalten, wobei jede Art von inhaltsbasiertem Metadatum als eine Spalte in einer Tabelle repräsentiert wird. Dies ermöglicht schnelle Suchvorgänge und komplexe Abfragen, die über einfache Dateisuchfunktionen hinausgehen. Zusätzlich zur Speicherung in Dateien und Datenbanken können inhaltsbasierte Metadaten auch in dedizierten Metadaten-Repositories wie

beispielsweise Adobe Experience Manager gespeichert werden, die als Teil eines Content Management System (CMS) fungieren. Solche Systeme können nicht nur die Metadaten speichern, sondern bieten auch Tools zur Bearbeitung, Verwaltung und Analyse dieser Daten, was die Effizienz der Datenverarbeitung und -nutzung erheblich steigert.

4.1.2.2 Automatisierte Erstellung inhaltsbasierter Metadaten

Die automatisierte Erstellung von inhaltsbasierten Metadaten durch den Einsatz von Computer Vision und maschinellem Lernen erleichtert die Art und Weise, wie große Mengen an Bild- und Videodateien verwaltet und durchsucht werden. Modernen Methoden sind in der Lage, komplexe Muster und Objekte innerhalb von Medien zu erkennen. Modelle wie das Convolutional Neural Network (CNN) und Deep Learning werden speziell trainiert, um Gesichter, Objekte, Szenen und sogar Emotionen in Medieninhalten zu identifizieren und zu klassifizieren. Diese Technologien bieten eine grundlegende Basis für automatisierte Tagging-Systeme, die in vielen Digital Asset Management (DAM) Systemen und CMS integriert sind. Diese Tools analysieren eingehende Medien und ordnen automatisch Tags und Kategorien zu, die später für die Organisation und Suche verwendet werden können.

In der praktischen Implementierung dieser Systeme ist die Vorbereitung der Daten ein wichtiger erster Schritt. Dabei müssen Bilder und Videos aus verschiedenen Quellen gesammelt und in einer strukturierten Form organisiert werden, die für die Analyse zugänglich ist. Anschließend erfolgt die Auswahl und das Training der Modelle. Abhängig von den spezifischen Anforderungen können Entwickler vorhandene vortrainierte Modelle nutzen oder eigene Modelle auf Frameworks wie TensorFlow[5] oder PyTorch[6] trainieren. Diese Frameworks bieten umfangreiche Unterstützung für das Training von Modellen, die speziell auf die Erkennung bestimmter Inhalte in den Medien ausgelegt sind. Die Effizienzsteigerung steht dabei im Vordergrund, da manuelle Tagging-Verfahren oft zeitaufwändig und inkonsistent sind. Das Trainieren der Modelle und die daraus resultierende Automatisierung bietet eine schnelle und konsistente Erstellung von Metadaten, verbessert die Suchfähigkeit durch präzise und umfangreiche Metadaten und erleichtert das Auffinden spezifischer Medien in großen Sammlungen.

Zusätzlich zu TensorFlow und PyTorch gibt es öffentlich verfügbare Frameworks bzw. Projekte, die für die automatische Erstellung von inhaltsbasierten Metadaten genutzt werden können:

- **OpenCV**[7] bietet eine Vielzahl von Funktionen für die Bild- und Videoanalyse, insbesondere die Gesichts- und Objekterkennung.
- **MediaPipe**[8] ist ein von Google entwickeltes Framework, das vorgefertigte Lösungen für Gesichtserkennung, Hand-Tracking, Pose-Schätzung liefert.
- **YOLO**[9] wird unter anderem für Echtzeit-Objekterkennungsaufgaben verwendet.

Diese Projekte bieten eine solide Grundlage für die Entwicklung von Systemen, die in der Lage sind, umfangreiche und akkurate inhaltsbasierte Metadaten automatisch zu generieren, wodurch die Verwaltung und Nutzung digitaler Medienressourcen signifikant verbessert wird.

[5] https://www.tensorflow.org/
[6] https://pytorch.org/
[7] https://opencv.org/
[8] https://github.com/google-ai-edge/mediapipe
[9] https://github.com/ultralytics/ultralytics

4.2 Analyse von Metadaten

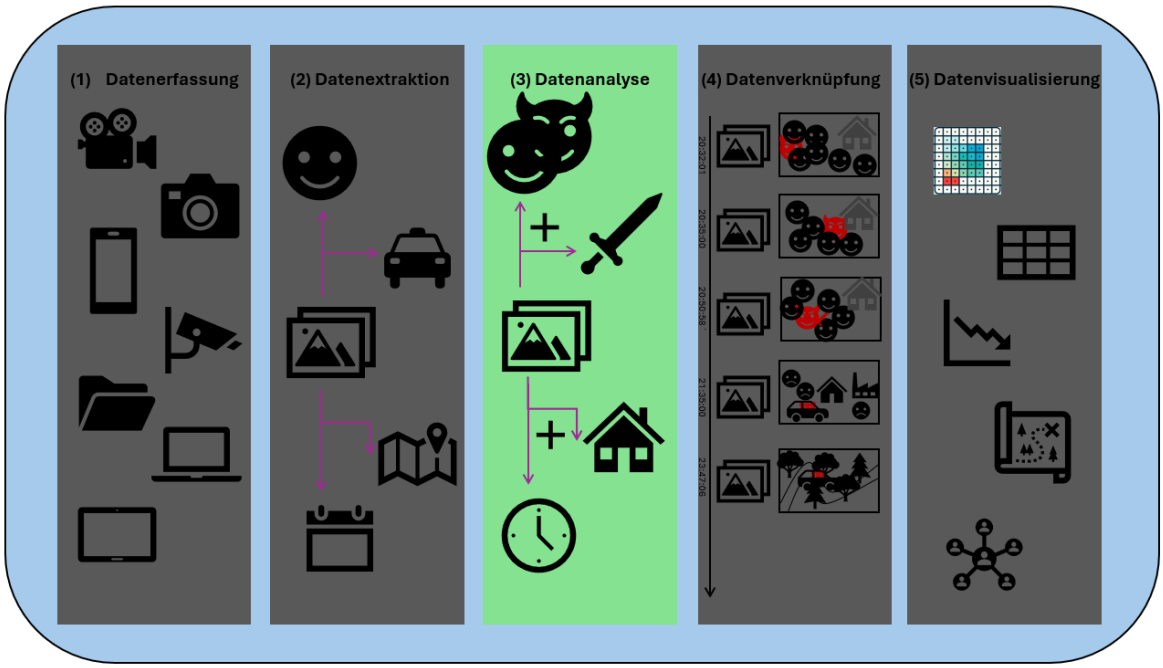

Abbildung 4.2: Schematischer Ablauf - Datenanalyse.

Im Rahmen dieses Abschnitts wird die metadatenbasierte Analyse von Videodateien im Detail behandelt (siehe Abbildung 4.2). Es geht insbesondere darum, wie man mit den bereits vorhandenen Informationen weitere Details ermitteln kann. Der effiziente Einsatz von Metadaten kann nicht nur die Navigation in umfangreichen Videobeständen erleichtern, sondern ermöglicht auch Einsichten, die weit über die Möglichkeiten traditioneller Verwaltungssysteme hinausgehen. Im Fokus stehen zwei zentrale Methoden der metadatenbasierten Analyse: die Analyse der EXIF-Metadaten und die inhaltsbasierte Analyse. Jede dieser Methoden wird in den folgenden Unterabschnitten detailliert betrachtet, wobei spezielle Techniken und Prozesse vorgestellt werden.

4.2.1 Datenanalyse mittels EXIF-Metadaten

Die Datenanalyse auf Basis der EXIF-Metadaten nutzt die im EXIF eingebetteten Informationen, um technische Details aus digitalen Aufnahmedateien zu analysieren und zu ggf. Metadaten zu ergänzen. Diese Art der Daten ist besonders wertvoll, weil sie es ermöglicht, Zusammenhänge zwischen verschiedenen Aufnahmen zu identifizieren, die unter ähnlichen Bedingungen oder an ähnlichen Orten gemacht wurden. Durch die Betrachtung von Zeitstempeln, die Analyse von Geotags und räumlichen Informationen können den bereits vorhandenen Informationen weitere wertvolle Daten hinzugefügt werden, die für die Rekonstruktion und dem Verständnis von Ereignissen, die Überprüfung von Zeitlinien oder die Durchführung von Standortanalysen nützlich sein können.

4.2.1.1 Betrachtung von Zeitstempeln

Die Betrachtung von Zeitstempeln ist eine Möglichkeit zur Analyse von Videodateien, die auf den enthaltenen Aufnahmezeitpunkten in den EXIF-Metadaten basiert. Zeitstempel geben den genauen Zeitpunkt der Aufnahme eines Videos oder Fotos an und sind daher eine entscheidende Informationsquelle, um Zusammenhänge zwischen verschiedenen Aufnahmen zu erkennen.

Aktuelle Technik:
In der digitalen Forensik ist die Darstellung von Bildern und Videos entlang einer Zeitleiste inzwischen eine Standardausstattung vieler forensischer Tools. Diese Funktion ermöglicht es Ermittlern, zeitlich geordnete Abfolgen von Ereignissen visuell darzustellen und zu analysieren. Bekannte Tools wie beispielsweise Autopsy [37] bieten diese Funktionalität an, wodurch die Untersuchung von zeitlichen Abläufen und Verknüpfung von Zeitstempeln erleichtert wird. Da die visuelle Darstellung von Ereignissen entlang einer Zeitleiste in den meisten forensischen Tools bereits etabliert ist, gibt es derzeit keine signifikant neuen oder innovativen Ansätze in diesem Bereich.

Anwendungsgebiete:
Die Betrachtung von Zeitstempeln kann dafür verwendet werden, Ereignisse zu rekonstruieren, die zeitlich synchronisiert abgelaufen sind, oder um Muster in der zeitlichen Abfolge von Ereignissen zu identifizieren. Zum Beispiel kann bei einer Betrachtung von mehreren Zeitstempeln festgestellt werden, ob Aufnahmen, die auf unterschiedlichen Geräten oder von verschiedenen Nutzern gemacht wurden, Teile desselben Ereignisses darstellen. Dies ist besonders nützlich in Bereichen wie der Sicherheitsüberwachung, der Forensik und bei der Erstellung von zeitbasierten medialen Dokumentationen. Ein präziser zeitlicher Abgleich kann auch aufzeigen, ob bestimmte Ereignisse gleichzeitig an verschiedenen Orten stattgefunden haben. Ein weiterer Nutzen ergibt sich aus der Verfolgung von Bewegungen oder Aktivitäten über große geografische Gebiete hinweg. Beispielsweise kann in der Sicherheitsüberwachung festgestellt werden, ob verdächtige Aktivitäten, die an verschiedenen Standorten aufgezeichnet wurden, zeitlich zusammenhängen und möglicherweise Teil eines koordinierten Ereignisses sind. Durch die Synchronisation von Videoaufnahmen verschiedener Überwachungskameras kann ein vollständiges Bild der Bewegungen und Aktionen innerhalb eines bestimmten Zeitraums erstellt werden.

Herausforderungen:
Die Genauigkeit der Zeitstempel ist wichtig für die Analyse und anschließende Verknüpfung von Mediendateien. Zeitstempel müssen präzise und synchronisiert sein, damit Ereignisse korrekt rekonstruiert und analysiert werden können. Abweichungen in den Zeitstempeln, die durch unterschiedliche Systemeinstellungen oder Gerätefehler verursacht werden, können zu fehlerhaften Analysen führen. Deshalb ist es wichtig, Methoden zur Zeitstempel-Normalisierung und -Korrektur zu verwenden, um diese Ungenauigkeiten zu minimieren.

Die Verarbeitungsgeschwindigkeit stellt eine weitere Herausforderung dar, insbesondere wenn große Mengen an Mediendateien analysiert werden müssen. Die Effizienz der Algorithmen zur Extraktion und Verknüpfung von Zeitstempeln beeinflusst direkt die Geschwindigkeit der Datenverarbeitung. Fortschritte in der Hardware, sowie Optimierungstechniken wie Parallelverarbeitung tragen dazu bei, die Verarbeitungsgeschwindigkeit zu erhöhen. Trotzdem bleibt die Balance zwischen Geschwindigkeit und Genauigkeit eine ständige Herausforderung.

Eine weitere Herausforderung ergibt sich aus der Praxis vieler Social Media Plattformen, die Metadaten, einschließlich Zeitstempeln, beim Hochladen von Mediendateien zu entfernen [34]. Dies kann die Verknüpfung und Analyse von Dateien erheblich erschweren, da wichtige zeitliche Informationen verloren gehen.

Datenschutzprobleme entstehen insbesondere bei der Erfassung und Analyse personenbezogener Daten. Zeitstempel können Rückschlüsse auf die Aktivitäten und Bewegungen von Personen zulassen, was sensible Informationen offenlegen kann. In der Forensik und Sicherheitsüberwachung ist es daher wichtig, Datenschutzrichtlinien zu beachten und Mechanismen zur Anonymisierung oder Verschlüsselung der Daten zu implementieren, um die Privatsphäre zu schützen.

4.2.1.2 Geotagging und räumliche Analyse

Geotagging und räumliche Analysen sind fortschrittliche Methoden zur Erweiterung der Metadaten von Videos und Fotos, indem geografische Informationen hinzugefügt werden. Geotagging bezieht sich auf den Prozess, in dem jeder Aufnahme geografische Koordinaten zugewiesen werden, die den genauen Aufnahmeort festlegen. Diese Informationen werden in der Regel während der Aufnahme durch GPS-fähige Geräte erfasst und direkt in den EXIF-Metadaten der Dateien gespeichert. Die räumliche Analyse nutzt diese Daten, um Muster und Beziehungen innerhalb der geografischen Informationen zu untersuchen. Durch die Analyse der Standorte und ihrer Beziehungen zueinander kann man wichtige Einblicke gewinnen, etwa die Bewegungsmuster von Menschen in Gebieten.

Aktuelle Technik:
Moderne Geotagging-Techniken nutzen GPS-Sensoren, die in den meisten Smartphones und Digitalkameras integriert sind, um automatisch geografische Koordinaten in die Metadaten von Fotos und Videos einzubinden. Darüber hinaus spielen visuelle Analysealgorithmen eine wichtige Rolle, insbesondere wenn GPS-Daten nicht verfügbar sind. Diese Algorithmen analysieren visuelle Merkmale wie markante Gebäude oder Landschaften, um den Aufnahmeort zu bestimmen. Machine Learning und Bildverarbeitungstechniken werden häufig verwendet, um diese Merkmale zu identifizieren und mit einer Datenbank bekannter Orte abzugleichen. [38]

Ein weiterer Fortschritt ist das Reverse Geocoding, das geografische Koordinaten in verständliche Ortsnamen umwandelt. Dies erleichtert das Verständnis des geografischen Kontexts und die Integration in geografische Informationssysteme. Reverse Geocoding kann sowohl auf konventionelle als auch auf Online-Weise durchgeführt werden. Online-Reverse-Geocoding-Dienste werden beispielsweise von Google Maps oder Bing Maps angeboten und sind über Web-Schnittstellen zugänglich. [39]

Moderne Smartphones verfügen über GPS-Empfänger, die es während der Aufnahme eines Bildes ermöglichen, genaue Standortdaten zu erfassen. Diese Standortdaten werden oft automatisch in die Metadaten von Fotos und Videos eingebettet, die mit der Smartphone-Kamera aufgenommen werden. Dies macht es einfach, die geografische Herkunft von Medieninhalten zu verfolgen und zu organisieren. Die Studie *Smartphone GPS accuracy study in an urban environment* zeigt, dass die Genauigkeit von einem iPhone 6 in urbanen Umgebungen im Bereich von 6 bis 13 Metern liegt. Neben GPS nutzen Smartphones auch Wi-Fi und Mobilfunkdaten zur Positionsbestimmung, was besonders in städtischen Gebieten hilfreich ist. Solche Technologien ermöglichen es, genaue und zuverlässige Geotagging-Daten für Fotos und Videos zu erzeugen und verbessern somit die Möglichkeiten zur präzisen Standortverfolgung und -organisation erheblich. [40]

Anwendungsgebiete:
Im Bereich der Sicherheitsüberwachung und Strafverfolgung kann Geotagging die präzise Verknüpfung von Videos und Bildern mit Standortinformationen ermöglichen, was die Überwachung und

Verfolgung von Bewegungen in Echtzeit erleichtern kann. Sicherheitsbehörden könnten so ein umfassendes Bild von Ereignissen an verschiedenen Orten erstellen und die Aufklärung von Straftaten sowie die Koordinierung von Einsatzkräften unterstützen.

Im Kontext von Terroranschlägen kann die Analyse von getaggten Mediendateien dazu beitragen, Bewegungsprofile von Verdächtigen zu erstellen. Durch die Analyse von Bewegungsprofilen über längere Zeiträume hinweg könnten Muster und Treffpunkte identifiziert werden, die möglicherweise auf kriminelle Aktivitäten hinweisen. Sicherheitsbehörden können Überwachungsvideos aus verschiedenen Kamerasystemen an verschiedenen Standorten verknüpfen, um die Bewegungen von Verdächtigen vor, während und nach einem Anschlag zu verfolgen oder auch oft besuchte Plätze festzustellen. Dies hilft nicht nur bei der Identifizierung der Täter, sondern auch bei der Aufdeckung möglicher Mittäter, Unterschlüpfe und Unterstützer.

Herausforderungen:
Eine große Herausforderungen ist die Genauigkeit der Geodaten. GPS-Daten können durch verschiedene Faktoren wie Gebäude, dichte Vegetation oder schlechtes Wetter beeinträchtigt werden, was zu Ungenauigkeiten führt. Diese Ungenauigkeiten können die Verlässlichkeit der erstellten Bewegungsprofile und räumlichen Analysen erheblich beeinträchtigen.

Wie auch bei der Betrachtung von Zeitstempeln stellt ein weiteres Problem der Verlust von Metadaten dar. Viele Social-Media-Plattformen und andere Online-Dienste entfernen Metadaten, einschließlich Geotags, beim Hochladen von Bildern und Videos [34]. Dadurch gehen geografische Informationen verloren, was die Verknüpfung und Analyse von Mediendateien erschwert.

Datenschutz und Sicherheit sind ebenfalls zentrale Herausforderungen. Die Verwendung von getaggten Daten kann sensible Informationen über die Aktivitäten und Aufenthaltsorte von Personen offenlegen, was erhebliche Datenschutz- und Sicherheitsbedenken aufwirft. Es ist daher wichtig, strenge Datenschutzrichtlinien und Anonymisierungstechniken zu implementieren, um die Privatsphäre der betroffenen Personen zu schützen.

4.2.2 Inhaltsbasierte Analyse

Im Gegensatz zu Ansätzen, die primär auf EXIF-Metadaten basieren, nutzt die inhaltsbasierte Verknüpfung fortschrittliche Technologien der künstlichen Intelligenz, insbesondere aus den Bereichen des maschinellen Lernens und der Computer Vision, um Objekte, Gesichter, Personen und Anomalien innerhalb von Videodateien automatisch zu erkennen und zu kategorisieren. Diese Techniken ermöglichen es, Videodateien tiefgehend nach ihrem tatsächlichen Inhalt zu organisieren und durchzusuchen. Die inhaltsbasierte Verknüpfung erweitert zudem die Möglichkeiten der Datenanalyse durch die Einbindung von Techniken wie der Schattenerkennung, die nicht nur zur Erkennung von Objekten beitragen, sondern auch zur Schätzung von Uhrzeiten und geografischen Details basierend auf Schattenwürfen. Abschließend wird die Rolle von dynamischen Metadaten beleuchtet, die die Interaktion mit Videodateien in Echtzeit fördern und somit neue Wege für Benutzerinteraktion und -engagement eröffnen. Insgesamt bietet die inhaltsbasierte Verknüpfung weitreichende Möglichkeiten, Videodateien auf innovative Weise zu nutzen, indem sie tiefere und kontextualisierte Einblicke in die Inhalte liefert.

4.2.2.1 Objekterkennung

Objekterkennung ermöglicht es, Objekte innerhalb von Bild- und Videodateien automatisch zu identifizieren und zu klassifizieren. Mit Hilfe von modernen Technologien lernen Modelle aus Millionen von Bildern und können eine Vielzahl von Objekten identifizieren. Die Verlinkung von Mediendateien erfolgt durch das Tagging der erkannten Objekte in den Dateien, wobei jedes Objekt mit einem Label versehen wird, das in den Metadaten der Datei oder einer Datenbank gespeichert ist. Diese Labels erleichtern die schnelle und effiziente Suche nach spezifischen Objekten oder Szenarien innerhalb eines umfangreichen Bestands an Mediendateien. In Sicherheitssystemen zum Beispiel kann Objekterkennung genutzt werden, beispielsweise Fahrzeuge zu identifizieren und Videos zu verlinken, die ähnliche Merkmale aufweisen.

Aktuelle Technik:
Ein bekanntes Modell der Objekterkennung ist YOLOv8, das mehrere architektonische Verbesserungen gegenüber seinen Vorgängern aufweist. In früheren Modellen der Objekterkennung musste oft manuell festgelegt werden, wie die Grenzen zwischen verschiedenen Objekten in einem Bild definiert werden sollen. Diese Grenzen, auch als *Bounding Boxes* bekannt, sind rechteckige Umrahmungen, die um die erkannten Objekte gezogen werden, um ihre Position und Größe im Bild zu markieren. YOLOv8 implementiert Mechanismen, die die Notwendigkeit für manuelle Einstellungen von *Bounding Boxes* eliminiert und somit den Rechenaufwand und die Komplexität des Modells reduziert. Dies führt zu einer höheren Effizienz und Genauigkeit bei der Objekterkennung. Ein weiteres Merkmal von YOLOv8 ist seine hohe Verarbeitungsgeschwindigkeit. Durch die Integration von Technologien, die die Berechnungen optimieren und den Rechenaufwand minimieren, kann YOLOv8 mehrere Bilder pro Sekunde analysieren. Dies macht es besonders geeignet für Echtzeitanwendungen, bei denen schnelle und präzise Objekterkennung erforderlich ist, wie z.B. in der Videoüberwachung oder in autonomen Fahrsystemen. [41]

RetinaNet[10] ist ein fortschrittliches Modell für die Objekterkennung, das speziell entwickelt wurde, um die Herausforderungen von Klassenungleichgewichten in Datensätzen zu bewältigen. Es kombiniert eine Residual Networks-Architektur mit einem Feature Pyramid Network (FPN) und nutzt eine innovative Verlustfunktion namens *Focal Loss*, um die Leistung bei der Erkennung schwer zu findender Objekte zu verbessern. Residual Networks ist bekannt für seine tiefen Netzwerkarchitekturen, die es ermöglichen, sehr tiefe Modelle zu trainieren, ohne dass das Problem des Verschwindens von Gradienten auftritt. Das FPN ist eine Schlüsselkomponente von RetinaNet, die es ermöglicht, Merkmale auf verschiedenen Auflösungsstufen zu extrahieren und zu kombinieren. Diese pyramidenartige Struktur hilft dem Modell, sowohl große als auch kleine Objekte effektiv zu erkennen. Focal Loss ist eine von RetinaNet eingeführte Verlustfunktion, die das Problem der Klassenungleichgewichte in Datensätzen adressiert. In vielen realen Datensätzen sind einige Objekte häufiger als andere aufzufinden, was zu einem Ungleichgewicht führt. Focal Loss reduziert die Gewichtung von leicht zu klassifizierenden Beispielen, sodass das Modell mehr Fokus auf schwer zu klassifizierende, seltene Objekte legen kann. Diese gezielte Gewichtung hilft dabei, die Gesamtgenauigkeit und Leistung des Modells zu verbessern. [42]

[10] https://github.com/fizyr/keras-retinanet

Faster Region-based Convolutional Neural Network (R-CNN) ist ein Zwei-Stufen-Detektor, der für seine hohe Genauigkeit bekannt ist. Die wesentlichen Innovationen von Faster R-CNN umfassen die Einführung eines Region Proposal Network (RPN). Es erzeugt Vorschläge für mögliche Objektregionen, indem es Ankerboxen auf verschiedenen Skalen und Seitenverhältnissen verwendet. Diese Vorschläge werden dann nach ihrer Wahrscheinlichkeit, ein Objekt zu enthalten, bewertet und sortiert. Nachdem die Regionsvorschläge vom RPN generiert wurden, verwendet Faster R-CNN eine Schicht, um feste Merkmalskarten für jede vorgeschlagene Region zu extrahieren. Diese Karten werden dann durch voll verbundene Schichten geführt, um die Klassifizierung der Objekte und die genauen Grenzen der Vorschläge zu bestimmen. Durch die Einführung des RPN und die gemeinsame Nutzung von Features kann Faster R-CNN schnell und präzise Regionenvorschläge generieren und klassifizieren. Es wird in vielen Überwachungssystemen verwendet, um Anomalien zu erkennen und detaillierte Analysen durchzuführen. [43]

Anwendungsgebiete:
Die Objekterkennung spielt eine zentrale Rolle bei der Verlinkung von Mediendateien, da sie die automatische Identifizierung und Kategorisierung von Inhalten in Bildern und Videos ermöglicht. In der Sicherheitsüberwachung wird sie verwendet, um Personen, Fahrzeuge und andere relevante Objekte zu identifizieren, was es ermöglicht, Bewegungsmuster zu verfolgen und potenzielle Bedrohungen zu erkennen. In der forensischen Analyse hilft die Objekterkennung, Beweismittel zu verknüpfen und Ereignisse zu rekonstruieren, indem sie Objekte in verschiedenen Mediendateien identifiziert.

Auch die Verkehrsüberwachung profitiert von der Objekterkennung, indem sie Fahrzeuge und deren Bewegungen in verschiedenen Videos identifiziert und verknüpft. Dies ermöglicht eine bessere Überwachung des Verkehrsflusses und die Identifizierung von Verkehrsverstößen. Ein Beispiel wäre die automatische Erfassung von Geschwindigkeitsübertretungen, bei der Systeme die Geschwindigkeit von Fahrzeugen messen und feststellen, ob sie die erlaubte Höchstgeschwindigkeit überschreiten. Hierbei müsste man ein fahrendes Fahrzeug aufnehmen, das zwei Markierungen in einem festen Abstand überfährt. Mit der Formel für die Geschwindigkeit (Geschwindigkeit = Strecke / Zeit), kann dann anhand der Aufnahme, diese berechnet werden. Ein weiteres Beispiel ist die Erkennung von Fahrzeugen, die unerlaubt auf Busspuren fahren. Ein mögliches Szenario wäre auch die Objekterkennung zu verwenden, um in Echtzeit Waffen in öffentlichen Räumen zu identifizieren. Wenn eine Waffe erkannt wird, könnte sofort ein Alarm ausgelöst werden, um die Sicherheitskräfte zu informieren und schnelle Maßnahmen zu ermöglichen.

Herausforderungen:
Eine Herausforderung ist die Genauigkeit der Erkennung, die stark von der Qualität und Vielfalt der Trainingsdaten abhängt. Schlechte Bildqualität, ungünstige Beleuchtung und ungewöhnliche Perspektiven können die Erkennungsleistung erheblich beeinträchtigen. Eine weitere Herausforderung ist die Verarbeitungsgeschwindigkeit, insbesondere wenn große Datenmengen in Echtzeit analysiert werden müssen. Effiziente Algorithmen und leistungsfähige Hardware sind notwendig, um die erforderliche Rechenleistung bereitzustellen. Ein weiteres Problem ist die Generalisierbarkeit der Modelle. Viele Objekterkennungsmodelle sind auf spezifische Szenarien und Datensätze trainiert und können Schwierigkeiten haben, in anderen Kontexten oder auf neuen Datensätzen gleichermaßen gut zu funktionieren. Das ist dann der Fall, wenn ein Modell zu stark an seine Trainingsdaten angepasst wurde (Overfitting). Datenschutz und ethische Bedenken spielen ebenfalls eine zentrale Rolle, da die

Erfassung und Analyse von Bild- und Videodateien sensible Informationen über Personen enthalten können. Es ist daher wichtig, strenge Datenschutzrichtlinien einzuhalten und sicherzustellen, dass die Technologie verantwortungsvoll eingesetzt wird.

4.2.2.2 Personenerkennung

Personenerkennung ist eine Technologie und ein Prozess, der es ermöglicht, Individuen in digitalen Bildern oder Videos zu identifizieren und zu verfolgen. Diese Identifizierung kann durch verschiedene Methoden und Algorithmen erreicht werden, die sich auf Merkmale wie Gesichtszüge, Körperform, Bewegungsmuster und andere biometrische Daten stützen.

Die wohl bekannteste Form der Personenerkennung ist die Gesichtserkennung, bei der spezifische Gesichtszüge wie Augen, Nase, Mund und deren relative Positionen zueinander analysiert werden. Neben der Gesichtserkennung gibt es die Körpererkennung, die die gesamte Körperform und Silhouette einer Person analysiert. Diese Methode kann besonders nützlich sein, wenn das Gesicht einer Person verdeckt ist oder sie sich bewegt. Eine weitere Methode ist die Gang-Erkennung, bei der das einzigartige Bewegungsmuster einer Person beim Gehen analysiert wird.

Aktuelle Technik:
Die Gesichtserkennungstechnologie hat in den letzten Jahren bedeutende Fortschritte gemacht, insbesondere durch den Einsatz fortschrittlicher Deep-Learning-Algorithmen, wie CNNs. Sie sind bekannt für ihre Fähigkeit, hierarchische Merkmale aus Rohbilddaten zu extrahieren und komplexe Muster zu erkennen, die für eine präzise Identifikation von Gesichtern erforderlich sind. Eine Einschränkung traditioneller CNN ist jedoch ihre begrenzte Fähigkeit, räumliche Hierarchien und Beziehungen innerhalb von Gesichtsmerkmalen effektiv zu erfassen. Dies kann zu einer reduzierten Erkennungsgenauigkeit führen, insbesondere bei Bildern mit variierenden Gesichtsausdrücken, Beleuchtungen oder Perspektiven. Eine Weiterentwicklung dieser Technologie sind die Deep Convolutional Neural Network (DCNN), die durch tiefere Architekturen in der Lage sind, noch kompliziertere Merkmale und Muster zu erlernen, was zu einer höheren Erkennungsgenauigkeit führt. CNNs und DCNNs sind keine unterschiedlichen oder komplementären Konzepte, sondern vielmehr miteinander verbunden, denn DCNNs bauen auf der Grundstruktur von CNNs auf und erweitern diese um zusätzliche Tiefe. [44]

Capsule Networks sind ein neuer Ansatz, der die Einschränkungen traditioneller CNN überwindet, indem er hierarchische räumliche Beziehungen innerhalb von Gesichtsmerkmalen erfasst. Diese Netzwerke bewahren räumliche Hierarchien effektiver und bieten somit eine vielversprechende Alternative nuancierter und zuverlässigere Gesichtserkennungsmethoden. Ein weiterer Ansatz sind Siamese Networks, die darauf spezialisiert sind, Ähnlichkeiten und Unterschiede zwischen Gesichtsabbildungen zu lernen. Sie bieten einen robusten Rahmen für Verifizierungs- und Identifikationsaufgaben, indem sie Paare von Bildern analysieren und feststellen, ob sie dieselbe Person darstellen. Während CNNs eine solide Grundlage für allgemeine Objekterkennung bieten, ermöglichen Capsule Networks eine robustere Erfassung von räumlichen Beziehungen, und Siamese Networks bieten eine spezielle Architektur für Vergleichsaufgaben. Sie adressieren somit unterschiedliche Herausforderungen in der Bildverarbeitung. [44]

Die moderne Körpererkennungstechnologie nutzt fortschrittliche Methoden der Computer Vision, um Menschen in Bildern und Videos präzise zu erkennen und zu analysieren. Ein zentrales Verfahren ist die Bildsegmentierung, die ein Bild in verschiedene Segmente unterteilt und so den menschlichen Körper von anderen Objekten im Bild trennt. Die Bildklassifikation weist jedem Bild eine spezifische Klasse wie Mensch oder Tier zu und wird oft in Überwachungssystemen eingesetzt. Die Pose-Schätzungstechnologie, wie sie von Modellen wie MoveNet[11] verwendet wird, bestimmt die Position von Körperteilen in 2D oder 3D und findet Anwendung in Animation, Sportwissenschaften und Rehabilitation. Diese Technik bietet detaillierte Informationen über die Körperhaltung und Bewegungen, was die Analyse und Rekonstruktion von Bewegungsabläufen erleichtert. [45]

Die Körpererkennung mittels thermischer Bildgebung hat sich in den letzten Jahren weiterentwickelt und bietet zahlreiche Anwendungen in verschiedenen Bereichen. Diese Technologie nutzt Infrarotkameras, um die Wärmesignaturen des menschlichen Körpers zu erfassen, was besonders in schlecht beleuchteten oder rauen Umgebungen von Vorteil ist. Fortschritte in der Technologie haben zur Entwicklung von Algorithmen geführt, die speziell für die Erkennung von Menschen in thermischen Bildern optimiert sind. Beispielsweise wurde das ThermalYOLO[12]-Modell entwickelt, um Personen in thermischen Bildern präzise zu erkennen und zu verfolgen. Dieses Modell nutzt maschinelles Lernen, in Form von einem Training, das mithilfe großer Mengen von thermischen Bilddaten trainiert wird, um die Genauigkeit der Erkennung zu verbessern und können in smarten Umgebungen, die durch eine hohe Dichte an Sensoren, Vernetzung und Automatisierung gekennzeichnet sind, und Überwachungssystemen eingesetzt werden. [46]

Die Technologie der Gang-Erkennung ermöglicht es, Personen anhand ihrer individuellen Gehweise zu identifizieren. Wie auch bei der Körpererkennung und der Gesichtserkennung ist die Anwendung von CNNs eine der aktuellen Technologien. CNNs werden häufig in der Bildverarbeitung eingesetzt, um Gang-Erkennungsaufgaben zu bewältigen, da sie effektiv visuelle Merkmale aus Gangbildern extrahieren und analysieren kann. Sie sind besonders leistungsfähig, weil vielschichtige neuronale Netze nutzen, um Merkmale zu extrahieren und Muster in den Daten zu erkennen. Diese Schichten ermöglichen es, sowohl einfache als auch komplexe Merkmale aus den Eingabedaten zu extrahieren, was besonders nützlich für die Gang-Erkennung ist, da es viele Variationen in der Gehweise einer Person gibt, die erfasst und analysiert werden müssen. [47]

Eine neue Technologien ist der Vision Transformer (ViT)[13]. Diese Technologie nutzt Aufmerksamkeitsmechanismen, um bedeutende Bildregionen zu identifizieren und detaillierte Ganganalysen durchzuführen. Das Gait-ViT[14] Modell ist ein spezifisches ViT-Modell, das für die Gang-Erkennung entwickelt wurde. Es beginnt damit, Gait Energy Images zu erzeugen, indem die Bilder eines Gangzyklus gemittelt werden. Diese Images werden dann in flache 2D-Patches umgewandelt und in eine Sequenz eingebettet, die in den Vision Transformer eingespeist wird. Der Transformer Encoder verarbeitet diese Sequenz, und schließlich führt ein mehrschichtiges Perzeptron die Klassifikation durch. Das Gait-ViT Modell erreicht eine Erkennungsgenauigkeit von über 99 Prozent auf verschiedenen Datensätzen. [48]

[11] https://github.com/tensorflow/tfjs-models/tree/master/pose-detection/src/movenet
[12] https://github.com/MAli-Farooq/Thermal-YOLO-And-Model-Optimization-Using-TensorFlowLite
[13] https://github.com/google-research/vision_transformer
[14] https://github.com/cosmaadrian/gait-vit

Das COMBI-Projekt nutzt das OpenPose-Framework[15], um die Gelenkpunkte des menschlichen Körpers in den aufgenommenen Bildern und Videos zu identifizieren. OpenPose arbeitet mit umfangreichen Datensätzen und maschinellem Lernen, um ein Modell zu trainieren, das genaue Vorhersagen über die Positionen der Gelenke treffen kann. Anhand dieser Vorhersagen werden sogenannte Rigs (digitale Skelettmodelle) erstellt, die die Bewegungsmuster und Gelenkpositionen der Personen im Video abbilden. Um diese Rigs mit den Bewegungsmustern der Verdächtigen in den Videos abzugleichen, wird zunächst ein 3D-Modell des aufgenommenen Bereichs erstellt. Virtuelle Kameras werden im 3D-Modell an den Positionen der tatsächlichen Überwachungskameras ausgerichtet. Dies ermöglicht genaue Messungen und Vergleiche, indem die Perspektive der Überwachungskameras nachgeahmt wird. Die erstellten Rigs werden dann in das 3D-Modell importiert und entlang der Achsen gemäß der Position des Täters im Videobild ausgerichtet. Ein möglicher Treffer wird angenommen, wenn die Extremitäten und Gelenke des Rigs perfekt mit denen des Täters übereinstimmen. [49]

Anwendungsgebiete:
Die Anwendungsgebiete der Personenerkennung in Bezug auf die Verknüpfung von Mediendateien auf Metadatenebene ähneln stark denen der Objekterkennung. In der Sicherheitsüberwachung ermöglicht die Personenerkennung die Identifizierung und Verfolgung von Personen in verschiedenen Überwachungsvideos. Durch die Verknüpfung von Aufnahmen, die dieselbe Person an verschiedenen Orten und zu verschiedenen Zeiten zeigen, können Bewegungsmuster und verdächtige Aktivitäten effizient nachverfolgt werden. Personenerkennung wird auch in Zugangskontrollsystemen eingesetzt, um sicherzustellen, dass nur autorisierte Personen Zugang zu bestimmten Bereichen haben. Durch die Verknüpfung von Videoaufnahmen mit Metadaten wie Zeitstempeln und Zugangsberechtigungen können Sicherheitsverantwortliche nachvollziehen, wer wann und wo Zugang hatte, und potenzielle Sicherheitslücken erkennen. In der Verkehrsüberwachung kann Personenerkennung verwendet werden, um Fußgängerbewegungen zu analysieren und ihre Sicherheit zu gewährleisten. Durch die Verknüpfung von Aufnahmen, die dieselben Personen an verschiedenen Verkehrsüberwachungsstellen zeigen, können Behörden Bewegungsmuster erkennen und Maßnahmen zur Verbesserung der Verkehrssicherheit ergreifen.

Herausforderungen:
Die Herausforderungen ähneln denen der Objekterkennung. Dies umfasst die Genauigkeit der Erkennung, die stark von der Qualität der Videodateien und den Lichtverhältnissen abhängt. Insbesondere im Bereich der Gesichtserkennung kommt noch der Aspekt hinzu, dass Kriminelle häufig ihr Gesicht verdecken, Masken oder ähnliche Kleidung tragen, was die Identifikation erschweren kann. Es ist wichtig, strenge Datenschutzrichtlinien zu befolgen und sicherzustellen, dass die Technologie verantwortungsbewusst eingesetzt wird, um den Datenschutz- und ethische Bedenken entgegenzuwirken. Auch die enormen Datenmengen und die Generalisierbarkeit der Erkennungsmodelle sind von Bedeutung. Man braucht erhebliche Rechenressourcen, um die Menge an Daten mit einem geeigneten Modell verarbeiten zu können.

[15] https://github.com/CMU-Perceptual-Computing-Lab/openpose

4.2.2.3 Anomalieerkennung

Die Anomalieerkennung in Videos hat das Ziel, ungewöhnliche oder abweichende Muster innerhalb von Videodateien zu identifizieren, die auf unvorhergesehene Ereignisse, Fehler oder möglicherweise verdächtige Aktivitäten hinweisen könnten. Diese Technologie verwendet fortschrittliche Algorithmen und Techniken, um die zeitliche Abfolge von Videobildern zu untersuchen und Aktivitäten oder Ereignisse zu erkennen, die vom üblichen Verhalten abweichen.

Aktuelle Technik:
Die aktuelle Technik zur Anomalieerkennung basiert zunehmend auf komplexen neuronalen Netzen und kombiniert verschiedene fortschrittliche Methoden, um die Erkennungsgenauigkeit zu verbessern. DifferNet[16] ist ein Beispiel für eine solche moderne Methode, die für unbeaufsichtigte, Bild-basierte Anomalieerkennung entwickelt wurde. Es kombiniert CNNs mit Normalizing Flows, um Anomalien zu identifizieren. DifferNet nutzt ein AlexNet-CNN[17] im Hintergrund, um Merkmale aus den Trainingsbildern zu extrahieren. Diese Merkmale werden dann in einem speziellen Raum mithilfe eines *Normalizing Flow*-Modells abgebildet. In diesem Raum kann man berechnen, wie wahrscheinlich es ist, dass eine neue Bildprobe den fehlerfreien Trainingsbildern ähnelt. Wenn ein Bild eine Anomalie enthält, hat es in der Regel eine geringere Wahrscheinlichkeit, in diesem Raum gut zu passen, als die fehlerfreien Bilder. Das Ziel des Trainings ist es, die Parameter des Modells so zu optimieren, dass die Wahrscheinlichkeit der typischen, fehlerfreien Merkmale maximiert wird, was die Erkennung von Anomalien erleichtert. [50]

Anomalib[18] ist eine umfassende Bibliothek zur Anomalieerkennung, die eine Sammlung von State-of-the-Art-Algorithmen für die visuelle Anomalieerkennung bereitstellt. Entwickelt von der OpenVINO-Toolkit-Community, zielt Anomalib darauf ab, leistungsstarke Anomalieerkennungsmodelle zu entwickeln, zu bewerten und bereitzustellen, sowohl auf öffentlichen als auch auf privaten Datensätzen. Die Bibliothek unterstützt eine Vielzahl von Anomalieerkennungsalgorithmen und bietet zudem Werkzeuge, die die Entwicklung und Implementierung benutzerdefinierter Modelle erleichtern. Zu den Hauptmerkmalen von Anomalib gehört eine modulare Schnittstelle und eine Kommandozeilenschnittstelle, die Training, Inferenz, Benchmarking und Hyperparameter-Optimierung unterstützen. Diese Schnittstellen machen die Bibliothek sowohl für Entwickler als auch für Benutzer ohne tiefere Programmierkenntnisse zugänglich. [51]

SplatPose[19] ist eine neuere Methode, die auf die Erkennung von Anomalien in 3D-Objekten abzielt, die aus unterschiedlichen Perspektiven aufgenommen wurden. Diese Methode verwendet einen Ansatz namens *3D Gaussian Splatting*, um die Pose von Objekten aus Multi-View-Bildern genau zu schätzen und Anomalien in diesen Ansichten zu erkennen. SplatPose zeichnet sich durch seine Effizienz sowohl beim Training als auch bei der Inferenz aus und erreicht dabei hohe Genauigkeit selbst mit weniger Trainingsdaten im Vergleich zu anderen Methoden.

Anwendungsgebiete:
In der Sicherheitsüberwachung ermöglicht die Anomalierkennung die automatische Erkennung ungewöhnlicher Aktivitäten oder Verhaltensweisen in Videomaterial, wodurch Sicherheitskräfte schnell

[16] https://github.com/marco-rudolph/differnet
[17] https://github.com/dansuh17/alexnet-pytorch
[18] https://github.com/openvinotoolkit/anomalib
[19] https://github.com/m-kruse98/SplatPose

auf potenzielle Bedrohungen reagieren können. Durch die Verknüpfung von Videos, die ähnliche Anomalien zeigen, können umfassendere Muster und verdächtige Bewegungen über verschiedene Standorte hinweg identifiziert werden. Anomalieerkennung kann auch im Verkehrsbereich eingesetzt werden, um Verstöße wie Geschwindigkeitsüberschreitungen, illegale Wendevorgänge oder das Überfahren von roten Ampeln zu identifizieren. Durch die Verknüpfung von Überwachungsvideos von verschiedenen Verkehrskameras können Ermittler die Bewegungen von verdächtigen Fahrzeugen nachverfolgen. Sie kann helfen, potenziell terroristische Aktivitäten zu erkennen, indem sie ungewöhnliche Verhaltensweisen in der Nähe sensibler Orte wie Flughäfen, Regierungsgebäuden oder öffentlichen Verkehrsmitteln, aber auch Konzerten, Sportereignissen oder Demonstrationen identifiziert.

Herausforderungen:
Eine der größten Herausforderungen in diesem Bereich ist die Definition dessen, was als anormal betrachtet wird. Bei klaren Regelverstößen, bei Verkehrsdelikte, wie beispielsweise illegales Wenden, ist die Definition einfach: Anomalien sind Aktivitäten, die gegen festgelegte Regeln oder Gesetze verstoßen. Bei menschlichem Verhalten ist die Definition jedoch komplexer und kontextabhängig. Was in einem Kontext als ungewöhnlich oder verdächtig gelten kann, könnte in einem anderen völlig normal sein. Zum Beispiel könnte das Verweilen an einer Bushaltestelle zu ungewöhnlichen Zeiten als Anomalie betrachtet werden, aber diese Bewertung hängt stark vom Kontext ab.

Wie bei der Erkennung von Personen und Objekten sind auch hier die Genauigkeit der Erkennung, Datenschutz- und ethische Bedenken, die enormen Datenmengen und die Generalisierbarkeit der Erkennungsmodelle von großer Bedeutung. Die Herausforderung besteht darin, zuverlässige und präzise Erkennungsergebnisse zu erzielen, während gleichzeitig die Privatsphäre geschützt und ethische Richtlinien eingehalten werden.

4.2.2.4 Schattenerkennung

Schattenerkennung in Videodateien ist eine Technik, die darauf abzielt, Schatten von Objekten in Videosequenzen zu identifizieren und zu analysieren. Diese Technik ist wichtig in der Bildverarbeitung und Computer Vision, um die Genauigkeit von Überwachungs- und Analysesystemen zu verbessern. Schatten können oft als Störfaktoren auftreten, die die Erkennung und Verfolgung von Objekten beeinträchtigen. Durch die Erkennung und genaue Analyse von Schatten kann jedoch nicht nur die Objektverfolgung verbessert werden, sondern es können auch zusätzliche Informationen wie die Lichtquelle und der Zeitpunkt der Aufnahme abgeleitet werden.

Aktuelle Technik:
Die aktuelle Technik zur Schattenerkennung hat sich durch den Einsatz fortschrittlicher Methoden der Bild- und Videobearbeitung erheblich weiterentwickelt. Ansätze wie der Triple-cooperative-Ansatz [52] und der Shadow-Object Segmentation and Instance Shadow Tracking (SSIS)-Track-Rahmen [53] nutzen die räumlich-zeitlichen Informationen aus mehreren Videobildern, um die Genauigkeit und Konsistenz der Schattenerkennung zu verbessern. Diese Techniken sind besonders nützlich für Anwendungen, bei denen eine präzise und robuste Schattenerkennung in dynamischen Szenen erforderlich ist, wie z.B. in der Videoüberwachung und beim autonomen Fahren.

Der Triple-cooperative-Ansatz ist besonders effektiv bei der Bewältigung von Herausforderungen, die durch Bewegungen und wechselnde Lichtverhältnisse in Videos entstehen. Der Triple-cooperative-Ansatz besteht aus drei kooperierenden Netzwerken, die zusammenarbeiten, um die Schattenerkennung zu verbessern. Das erste Netzwerk, das Schattenerkennungsnetzwerk, identifiziert Schatten in den einzelnen Frames eines Videos. Es verwendet unter anderem CNNs, um die charakteristischen Merkmale von Schatten zu erkennen und sie von anderen Objekten und Hintergründen zu unterscheiden. Dieses Netzwerk stellt sicher, dass Schatten präzise und zuverlässig erkannt werden, was die Grundlage für die weiteren Schritte bildet. Nachdem die Schatten in den Frames erkannt wurden, übernimmt das Schattenverfolgungsnetzwerk die Aufgabe, die Bewegungen der Schatten über die Zeit zu verfolgen. Es stellt sicher, dass die erkannten Schatten konsistent über mehrere Frames hinweg verfolgt werden, was besonders wichtig ist, um die räumlich-zeitliche Konsistenz zu wahren. Das dritte Netzwerk im Triple-cooperative-Ansatz ist das Schattenentfernungsnetzwerk. Dieses Netzwerk entfernt die erkannten Schatten aus den Frames, um die darunter liegenden Objekte und Szenen klarer sichtbar zu machen. Dies ist besonders nützlich in Anwendungen wie der Videoüberwachung, wo Schatten die Sicht auf wichtige Details verdecken können. Durch die Entfernung der Schatten können die analysierten Videos genauer ausgewertet und interpretiert werden. Er wurde auf einem neuen Datensatz evaluiert, der speziell für die Schattenerkennung in Videos entwickelt wurde. Dieser Datensatz umfasst 120 Videos mit insgesamt 11.685 Frames und hochqualitativen Annotationen. Die Ergebnisse der Experimente zeigen laut den Autoren, dass der Triple-cooperative-Ansatz die Leistung der Schattenerkennung in Videos signifikant verbessert. Durch die Zusammenarbeit der drei Netzwerke können präzisere und konsistentere Ergebnisse erzielt werden, selbst in komplexen und dynamischen Szenen. [52]

Der SSIS-Track-Rahmen ist eine fortschrittliche Methode zur Schattenerkennung, der gleichzeitig die Erkennung, Segmentierung und Verfolgung von Schatten-Objekt-Paaren in Videos durchführt. Diese Methode nutzt transformerbasierte Modelle, um eine präzise und konsistente Schattenerkennung zu gewährleisten. Zunächst werden Schatten und ihre zugehörigen Objekte in den Video-Frames identifiziert und segmentiert, wodurch die räumlichen Beziehungen zwischen ihnen erfasst werden. Anschließend werden diese Paare über die Zeit hinweg verfolgt, um die Konsistenz der Erkennung aufrechtzuerhalten. Die transformerbasierten Modelle lenken die räumlich-zeitliche Aufmerksamkeit auf die relevanten Merkmale der Schatten und Objekte, was die Genauigkeit der Erkennung und Verfolgung erheblich verbessert. Diese Methode bietet mehrere Vorteile, darunter eine hohe räumlich-zeitliche Konsistenz und Robustheit gegenüber Veränderungen in der Beleuchtung und den Bewegungen der Objekte. Der SSIS-Track-Rahmen stellt somit eine Methode der Schattenerkennung in dynamischen Szenen dar, was ihn besonders nützlich für Anwendungen wie Videoüberwachung und autonomes Fahren macht. [53]

Um die Aufnahmeuhrzeit eines Fotos oder Videos anhand von Schatten bestimmen zu können, kann die Analyse der Schattenlängen und -winkel in Verbindung mit Werkzeugen wie SunCalc[20] genutzt werden. Diese Methode, auch als Chronolocation bezeichnet, kann helfen, den ungefähren Zeitpunkt der Aufnahme zu ermitteln, wenn andere Methoden, wie die Betrachtung von Uhren oder Datumseinblendungen, nicht anwendbar sind. Zunächst wird die Länge des Schattens und die Höhe des Objekts, das den Schatten wirft, gemessen. Wichtig ist das Verhältnis dieser beiden Werte, nicht die absoluten Längen. Dann wird SunCalc, ein Online-Tool, verwendet, das die Position der Sonne zu einem bestimmten Zeitpunkt und an einem bestimmten Ort simuliert. Durch Eingabe des

[20] https://www.suncalc.org/

Standorts und der ungefähren Zeit kann man die Schattenlänge so lange anpassen, bis sie mit der gemessenen Schattenlänge im Bild übereinstimmt. Damit diese Methode genau ist, müssen zusätzlich zur natürlichen Lichtquelle einige Bedingungen erfüllt sein: Der Ort und das Datum der Aufnahme müssen bekannt sein, das Objekt und sein Schatten müssen in einem Winkel von ungefähr 90 Grad zur Kamera stehen, der Schatten muss auf einer ebenen Fläche fallen und das Bild sollte nicht durch spezielle Kameraobjektive wie Fischaugenobjektive verzerrt sein. Wichtig ist auch, dass die Schattenlänge zu zwei verschiedenen Tageszeiten gleich sein kann: einmal am Vormittag und einmal am Nachmittag. Daher muss die Richtung des Schattens berücksichtigt werden, um die genaue Uhrzeit zu bestimmen. [54]

Anwendungsgebiete:
In der Sicherheitsüberwachung kann die Schattenerkennung eine wichtige Rolle spielen, um die präzise Erfassung von Position und Bewegung von Objekten und Personen zu verbessern. Durch die Analyse zusätzlicher Details wie Szenengeometrie und Lichtverhältnisse kann die Genauigkeit der Überwachungssysteme gesteigert werden. In der forensischen Analyse hilft die Schattenerkennung dabei, den Zeitpunkt und die Bedingungen von Aufnahmen genauer zu bestimmen. Dies ist besonders bei der Rekonstruktion von Tatorten oder Ereignisabläufen von großer Bedeutung, da es Ermittlern ermöglicht, eine genauere zeitliche und räumliche Zuordnung der Ereignisse vorzunehmen. Bei der Verkehrsüberwachung unterstützt die Schattenerkennung die präzise Bestimmung der Position von Fahrzeugen und Fußgängern sowie die Erkennung von Verkehrsverstößen.

Herausforderungen:
Eine große Herausforderungen ist die Variabilität der Lichtverhältnisse. Unterschiedliche Beleuchtungssituationen, sei es durch Tageslicht, künstliche Beleuchtung oder Wetterbedingungen, können die Form, Intensität und Richtung von Schatten stark beeinflussen. Diese Möglichkeit erschwert die gleichmäßige Erkennung und Analyse von Schatten in verschiedenen Aufnahmen. Ein weiteres Problem stellt die Komplexität der Szenengeometrie dar. In komplexen Szenen mit vielen Objekten und Personen können Schatten überlagert oder verzerrt werden, was die genaue Zuordnung und Analyse erschwert. Zudem können bewegliche Lichtquellen oder sich ändernde Umgebungsbedingungen die Schattenerkennung weiter komplizieren. Weitere Herausforderungen decken sich mit den bereits vorgestellten Methoden. Diese sind die Genauigkeit der Erkennung, Datenschutz- und ethische Bedenken, die enormen Datenmengen und die Generalisierbarkeit der Erkennungsmodelle von Bedeutung.

4.3 Datenverknüpfung

Nachdem die Daten extrahiert und analysiert wurden, können sie nun verknüpft werden (siehe Abbildung 4.3), um tiefere Einblicke zu gewinnen und komplexe Zusammenhänge aufzudecken. Es gibt mehrere Methoden, wie diese Verknüpfungen effizient durchgeführt werden können, um beispielsweise Bewegungsmuster zu rekonstruieren oder Anomalien in einem größeren Kontext zu verstehen.

Zeitstempel spielen eine zentrale Rolle bei der Verknüpfung von Daten aus unterschiedlichen Quellen. Durch die Zuordnung von Ereignissen zu spezifischen Zeitpunkten können chronologische Abfolgen erstellt werden, die es ermöglichen, den Ablauf von Ereignissen präzise zu rekonstruieren.

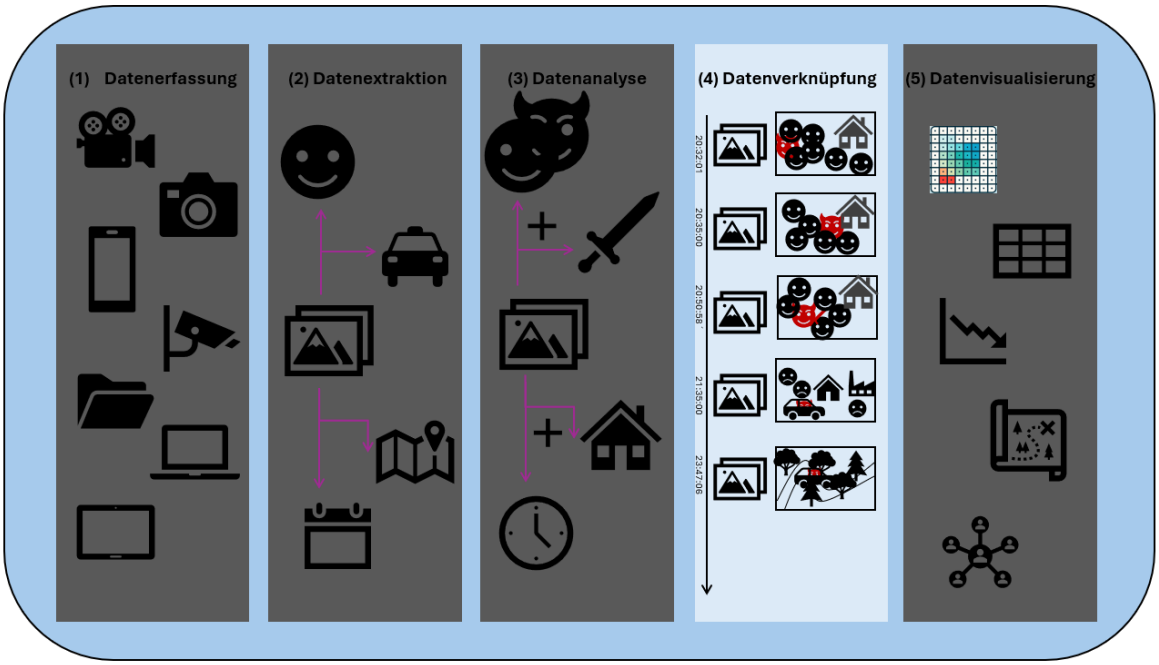

Abbildung 4.3: Schematischer Ablauf - Datenverknüpfung.

Diese Methode ist besonders nützlich, um Ereignisse in Videos oder Bildern mit anderen zeitgebundenen Daten wie Sensorinformationen oder Log-Dateien zu verbinden. Eine detaillierte Analyse von Bewegungsmustern wird möglich, wenn Zeitstempel mit GPS-Daten und erkannten Objekten oder Personen verknüpft werden. Diese Methode erlaubt es, die Bewegungen von Personen oder Objekten über Zeit und Raum hinweg zu verfolgen. Beispielsweise kann in der Strafverfolgung nachvollzogen werden, wie sich eine verdächtige Person an einem Tatort bewegt hat. Diese Verknüpfung kann auch dazu dienen, verdächtige Muster zu identifizieren, wie etwa das wiederholte Auftreten einer Person an verschiedenen Tatorten zu unterschiedlichen Zeiten.

Die Analyse von Schatten in Bildern oder Videos, kombiniert mit den entsprechenden GPS-Daten, ermöglicht es, die genaue Tageszeit einer Aufnahme zu bestimmen. Indem die Position und Länge von Schatten in Bezug auf die Sonnenposition analysiert wird, kann überprüft werden, ob die Zeitangaben in den Metadaten korrekt sind. Diese Methode ist besonders wertvoll, um die Authentizität von Videoaufnahmen zu überprüfen oder Zeitstempel zu rekonstruieren, wenn diese fehlen oder manipuliert wurden. Durch die Kombination von Anomaliedetektion mit Informationen über erkannte Personen kann die Effizienz in der Strafverfolgung erheblich gesteigert werden. Wenn eine Anomalie, wie etwa ein ungewöhnliches Verhaltensmuster, identifiziert wird, kann diese sofort mit den erkannten Personen in Verbindung gebracht werden. Dies ermöglicht eine schnelle Identifizierung potenzieller Verdächtiger, was besonders in zeitkritischen Situationen wie bei der Aufklärung von Straftaten von großem Vorteil ist.

Die Kombination erkannter Objekte oder Personen in verschiedenen Videodateien ist eine effektive Methode, um Zusammenhänge zwischen Aufnahmen zu identifizieren und Bewegungsmuster zu analysieren. Wenn eine Person oder ein Objekt in einem Video erkannt und identifiziert wird, können diese Informationen als Metadaten gespeichert werden. Moderne Algorithmen ermöglichen es, diese Metadaten in anderen Videos zu suchen und dieselbe Person oder dasselbe Objekt erneut zu erkennen. Dies eröffnet die Möglichkeit, Videodateien miteinander zu verknüpfen und ein

umfassendes Bewegungsprofil zu erstellen. Beispielsweise könnte eine Überwachungskamera in einer Stadt eine verdächtige Person aufzeichnen. Diese Person wird durch Gesichtserkennung oder andere biometrische Merkmale identifiziert, und die Identität wird in den Metadaten gespeichert. Anschließend können mit Hilfe von verschiedenen Algorithmen andere Videos durchsucht werden, um festzustellen, ob diese Person in weiteren Aufnahmen erscheint. Wenn sie erkannt wird, werden die Videos miteinander verknüpft, und es entsteht eine zeitliche und räumliche Sequenz, die zeigt, wie sich die Person durch verschiedene Bereiche der Stadt bewegt hat.

Durch diese Beispiele wird deutlich, dass eine Kombination mehrerer Methoden oft die beste Herangehensweise darstellt, um die Verknüpfung und Analyse von Videodateien zu optimieren. Durch die Nutzung verschiedener Techniken können Schwächen einzelner Methoden kompensiert und die Gesamteffizienz und Genauigkeit der forensischen Analyse erheblich verbessert werden. Wenn EXIF-Metadaten wie Zeitstempel fehlen oder unvollständig sind, kann die Schattenerkennung eine wertvolle Alternative bieten. In Situationen, in denen GPS-Daten nicht verfügbar sind, kann die Identifikation signifikanter Objekte im Video zur Bestimmung des Aufnahmeorts herangezogen werden. Beispielsweise kann ein markantes Gebäude, eine bekannte Landschaft oder ein einzigartiges Straßenschild im Hintergrund eines Videos Hinweise auf den Standort geben.

4.4 Zusammenfassung

Trotz der Unterschiede in den spezifischen Ansätzen bleiben die Grundprobleme bei der Analyse von Videodateien in der Regel gleich. Insbesondere stellen die erforderlichen Rechenressourcen eine bedeutende Hürde dar. Die Verarbeitung und Analyse großer Mengen an Mediendateien erfordern in der Regel erhebliche Rechenkapazitäten, insbesondere bei der Anwendung komplexer Algorithmen der inhaltsbasierten Metadatenextraktion. Die Qualität der Videodateien spielt dabei eine entscheidende Rolle, denn minderwertige Aufnahmen können die Genauigkeit der Analysen erheblich beeinträchtigen. Für EXIF-basierte Methoden ist das Vorhandensein von Metadaten eine Voraussetzung. Fehlen diese Daten, sind alternative Ansätze erforderlich, um dennoch brauchbare Informationen zu extrahieren.

Ein weiterer kritischer Aspekt ist der Datenschutz. Die Erfassung und Analyse von Mediendateien und den darin enthaltenen Metadaten wirft wichtige Fragen zum Schutz der Privatsphäre und zur ethischen Nutzung dieser Technologien auf. Die Verarbeitung von Mediendateien umfasst oft die Erfassung persönlicher Informationen wie das Verhalten, die Bewegungen und die Identität von Personen. Werden diese Daten ohne Zustimmung oder Wissen der betroffenen Personen gesammelt und verwendet stellt das ein datenschutzrechtliches Problem dar. Besondere Bedenken ergeben sich, wenn diese Daten mit anderen Datenquellen kombiniert werden, wodurch umfassende Profile erstellt werden können, die die Privatsphäre der betroffenen Personen erheblich beeinträchtigen. Um diese Herausforderungen anzugehen, müssen strenge rechtliche Rahmenbedingungen und Datenschutzrichtlinien eingehalten werden. In vielen Regionen gibt es Gesetze wie die Datenschutz-Grundverordnung (DSGVO) der Europäischen Union, die den Schutz personenbezogener Daten regeln und sicherstellen, dass solche Daten nur mit ausdrücklicher Zustimmung der Betroffenen und unter Einhaltung strenger Sicherheitsmaßnahmen verarbeitet werden dürfen.

Neben den rechtlichen Anforderungen spielt auch die ethische Dimension eine wichtige Rolle. Es ist entscheidend, dass die Technologie nicht diskriminierend eingesetzt wird und keine bestimmten Gruppen benachteiligt. Beispielsweise muss die Gesichtserkennungstechnologie und auch die Personenerkennung so entwickelt und implementiert werden, dass sie unabhängig von Hautfarbe, Geschlecht oder anderen persönlichen Merkmalen gleichbleibend genau funktioniert. Diskriminierung und Verzerrungen in den Datenanalysen können zu ungerechten Ergebnissen führen, wie etwa falschen Identifizierungen oder ungerechtfertigten Verdächtigungen. Ein ethisch verantwortungsbewusster Ansatz sollte auch die Minimierung der Datenerfassung beinhalten. Das bedeutet, dass nur die unbedingt notwendigen Daten erfasst und verarbeitet werden sollten. Bei der Implementierung von Schattenerkennungssystemen könnte dies bedeuten, dass nur Metadaten und keine vollständigen Videoaufzeichnungen gespeichert werden, um die Privatsphäre der betroffenen Personen zu schützen, was jedoch beispielsweise die Rekonstruktion von Straftaten einschränken kann.

Trotz der genannten Herausforderungen bieten die vorgestellten Methoden ein großes Potenzial für die Verknüpfung und Analyse von Videodateien. Die systematische Kombination verschiedener Ansätze kann die Effizienz und Genauigkeit forensischer Untersuchungen erheblich verbessern und somit einen wertvollen Beitrag unter anderem zur modernen Sicherheitsüberwachung und Tatortrekonstruktion leisten. Durch die Kombination von EXIF-Metadaten und inhaltsbasierten Analysetechniken kann der Prozess der Datenanalyse erheblich beschleunigt werden. Statt sich auf manuelle Überprüfungen zu verlassen, können automatisierte Systeme große Mengen an Videodateien in kürzerer Zeit analysieren. Dies ist besonders wichtig in dringenden Fällen, in denen schnelle Ergebnisse entscheidend sind, um weitere Maßnahmen zu ergreifen oder Ermittlungen voranzutreiben.

5 Visualisierung der verknüpften Informationen

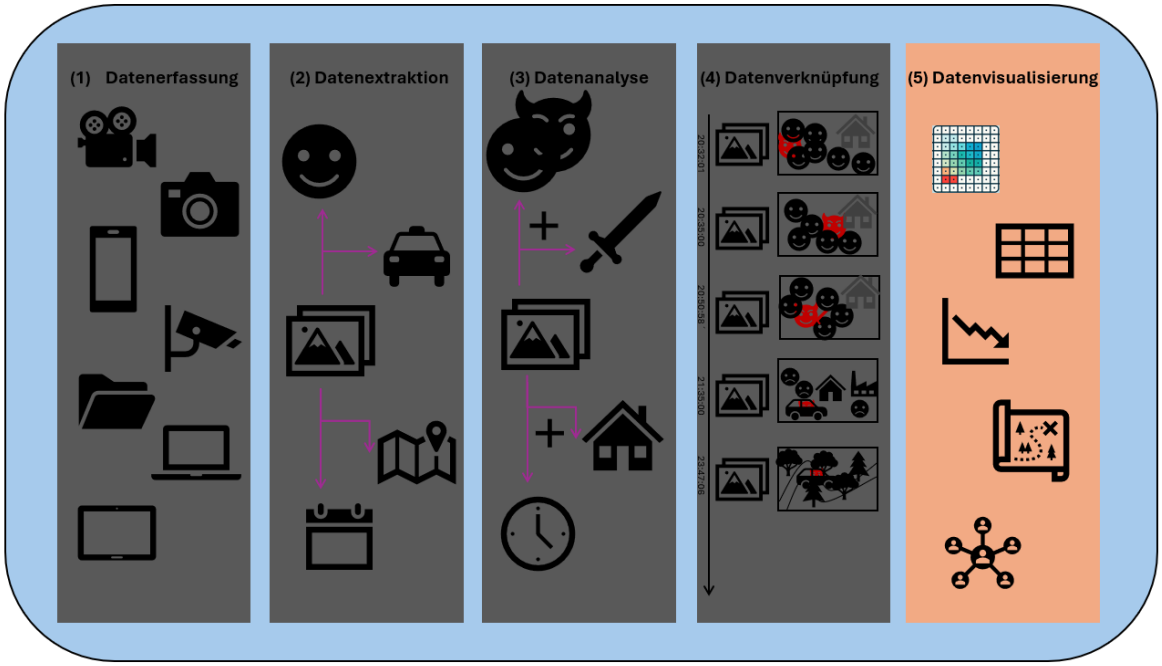

Abbildung 5.1: Schematischer Ablauf - Datenvisualisierung.

Nachdem relevante Daten extrahiert und verknüpft wurden, ist es entscheidend, diese Informationen so aufzubereiten, dass sie für Ermittler und andere Anwender verständlich und nutzbar sind. Hierfür sollten die Daten mit verschiedenen Methoden visualisiert werden (siehe Abbildung 5.1). In diesem Kapitel werden verschiedene Methoden und Ansätze vorgestellt, die es ermöglichen, die gewonnenen Erkenntnisse effektiv darzustellen. Ziel ist es, durch geeignete Visualisierungen komplexe Zusammenhänge zu verdeutlichen und die Interpretation der Daten zu erleichtern. Dabei werden sowohl bewährte als auch neue Visualisierungstechniken betrachtet, um eine umfassende Übersicht der Möglichkeiten zu bieten.

5.1 Zeitliche Visualisierungen

Um die zeitliche Visualisierung realisieren zu können, ist es notwendig, dass Zeitstempel entweder direkt extrahiert worden sind oder aber aus der Analyse gewonnen werden. Danach kann man einen chronologischen Ablauf bilden, indem man die Dateien nach den Zeitstempeln sortiert. Die Videodateien können dann in diesem Kontext verknüpft betrachtet werden. Eine Filterung auf einen Zeitraum mit Start und Ende ist ebenfalls sinnvoll. Die Betrachtung der Daten in einem zeitlichen Kontext Visualisierung bietet eine zentrale Methode, um die Abfolge und Dynamik von Ereignissen in der Videoforensik darzustellen. Durch die chronologische Anordnung der verknüpften Videoinformationen können komplexe Zusammenhänge und Abläufe besser verständlich gemacht werden. Diese Methode ermöglicht es, den zeitlichen Verlauf von Ereignissen detailliert nachzuvollziehen und unterstützt die Identifikation von Mustern und Anomalien. Diese Art der Darstellung ist bereits ein etablierter Standard in den Forensik.

Eine der grundlegendsten Formen der zeitlichen Visualisierung ist die Verwendung von Zeitachsen. Hierbei werden Ereignisse entlang einer horizontalen Linie platziert, die den zeitlichen Fortschritt repräsentiert. Zeitachsen können einfach oder interaktiv gestaltet sein, wobei letztere Benutzern die Möglichkeit bieten, detaillierte Informationen zu einzelnen Ereignissen durch Anklicken abzurufen. Diese Zeitachsen sind besonders effektiv, wenn sie in Kombination mit der Anzahl der Mediendateien verwendet werden, die mit jedem Ereignis verbunden sind. So kann eine Zeitachse nicht nur die zeitliche Abfolge, sondern auch die Menge an verfügbaren Video- und Bildmaterialien anzeigen. Eine weitere Variation ist das Gantt-Diagramm, das ursprünglich für das Projektmanagement entwickelt wurde. Es stellt die Dauer und Überlappung von Ereignissen dar und bietet eine visuelle Darstellung von Zeiträumen. Dies ist besonders hilfreich, um parallele Abläufe und deren zeitliche Beziehungen zueinander zu analysieren.

Die zeitliche Visualisierung bietet viele Vorteile. Sie sorgt für Übersichtlichkeit, da Ereignisse klar und strukturiert dargestellt werden, was es erleichtert, Muster und Abweichungen zu erkennen. Zudem ermöglicht die chronologische Anordnung, den Ablauf von Ereignissen leicht nachvollziehen und verstehen zu können. Interaktive Visualisierungen bieten die Möglichkeit zur tiefen Erkundung der Daten, da Benutzer in der Lage sind, spezifische Details zu untersuchen und Verbindungen zwischen Ereignissen herzustellen. Jedoch gibt es auch Nachteile. Bei großen und komplexen Datensätzen kann die Darstellung unübersichtlich werden, was die Interpretation erschwert. Die Handhabung und Visualisierung großer Datenmengen erfordert leistungsfähige Tools und Technologien, um eine flüssige und effiziente Nutzererfahrung zu gewährleisten. Einfache zeitliche Visualisierungen bieten möglicherweise nicht genügend Tiefe oder Details für umfassende Analysen, was die Notwendigkeit zusätzlicher Visualisierungstechniken erfordert.

5.2 Interaktive Karten

Interaktive Karten bieten eine zusätzliche Dimension, indem sie die geografischen Daten der Videodateien visualisieren. Die dafür notwendigen GPS-Daten können entweder direkt aus den Metadaten extrahiert werden oder anhand der bereits vorgestellten Methode durch signifikante Objekte im Bild selbst bestimmt werden. GPS-Daten können dann genutzt werden, um die Bewegungen von Personen oder Objekten auf einer Karte darzustellen. Es können dann durch die Verknüpfung von Objekten oder Personen und den GPS-Daten Bewegungen verfolgt werden. Diese räumliche Visualisierung erleichtert das Verständnis komplexer räumlicher Zusammenhänge und unterstützt die Identifikation von Mustern und Anomalien in den Bewegungsdaten. Sie ist besonders nützlich bei der Verfolgung von Fluchtwegen oder der Lokalisierung von Tatorten. Die Darstellungsformen von GPS-Daten gehören bereits in gängigen Anwendungen der Forensik zum Standard.

Es gibt verschiedene Variationen der Darstellung von interaktiven Karten. Eine häufig verwendete Variante sind Standardkarten, die geografische Informationen darstellen und es den Nutzern ermöglichen, durch Zoom- und Schwenkfunktionen die Karte zu erkunden. Diese Karten können zusätzlich mit Markierungen und Annotationen versehen werden, um spezifische Ereignisse oder Datenpunkte hervorzuheben. Eine weitere Variante sind Heatmaps, die Bereiche mit hoher Aktivität oder Dichte durch Farbkodierungen anzeigen. Diese Methode ist besonders nützlich, um Hotspots von Aktivitäten oder Vorfällen schnell zu identifizieren. 3D-Karten sind eine weitere innovative Variation, die es

ermöglicht, geografische Daten in einer dreidimensionalen Umgebung darzustellen. Diese Karten bieten eine realistischere Perspektive und können besonders hilfreich sein, um Bewegungen und Interaktionen in urbanen oder komplexen Umgebungen zu analysieren.

Die Verwendung interaktiver Karten ermöglicht eine intuitive und benutzerfreundliche Erkundung großer und komplexer Datensätze. Durch die visuelle Darstellung geografischer Daten können Muster und Anomalien leichter erkannt werden. Interaktive Funktionen wie Zoom, Schwenken und das Abrufen zusätzlicher Informationen durch Klicken auf Markierungen erhöhen die Nutzbarkeit und erleichtern die detaillierte Analyse. Zudem können interaktive Karten in Echtzeit aktualisiert werden, was sie besonders nützlich für die kontinuierliche Überwachung und Analyse von Ereignissen macht. Es gibt jedoch auch einige Nachteile bei der Verwendung interaktiver Karten. Die Erstellung und Pflege solcher Karten erfordert erhebliche technische Ressourcen und Expertise. Die Integration und Verarbeitung großer Datenmengen kann anspruchsvoll sein und leistungsfähige Hardware sowie fortschrittliche Softwarelösungen erfordern. Darüber hinaus kann die visuelle Komplexität interaktiver Karten bei umfangreichen Datensätzen überwältigend wirken, was die Interpretation erschweren kann. Schließlich besteht die Gefahr, dass durch unzureichende Datensicherheit sensible Informationen offengelegt werden könnten.

5.3 Netzwerke und Graphen

Ein weiteres Werkzeug zur Visualisierung sind Netzwerke und Graphen. Diese Methode ermöglicht es, Beziehungen und Interaktionen zwischen verschiedenen Akteuren oder Ereignissen darzustellen. Durch die Verwendung von Netzwerkanalysen können Verbindungen zwischen unterschiedlichen Datenpunkten visualisiert werden, was besonders hilfreich ist, um Zusammenhänge in großen Datenmengen zu erkennen. Netzwerke und Graphen bieten eine klare und strukturierte Übersicht über komplexe Beziehungsgeflechte und können dazu beitragen, versteckte Verbindungen aufzudecken. Hierfür muss die inhaltliche Analyse der Dateien durchgeführt worden sein. Die getaggten Objekte oder Personen können dann entsprechend verfolgt oder deren Beziehungen betrachtet werden.

Es gibt verschiedene Ansätze, Netzwerke und Graphen in der Videoforensik zu verwenden. Soziale Netzwerkanalyse untersucht die sozialen Beziehungen und Interaktionen zwischen verschiedenen Personen. In der Videoforensik kann diese Darstellung helfen, Netzwerke von Verdächtigen zu identifizieren und deren Beziehungen und Kommunikationsmuster zu analysieren. Bewegungsgraphen stellen die Bewegungen von Personen und Objekten innerhalb eines bestimmten Bereichs dar, wobei Knoten spezifische Positionen und Kanten die Pfade zwischen diesen Positionen repräsentieren. Diese Graphen können verwendet werden, um die Routen von Verdächtigen zu verfolgen und Muster in ihren Bewegungen zu erkennen. Ereignisgraphen verknüpfen verschiedene Ereignisse miteinander, basierend auf zeitlichen und räumlichen Zusammenhängen, und helfen, die Abfolge von Ereignissen zu rekonstruieren und kausale Beziehungen zu identifizieren. Interaktionsgraphen modellieren die Interaktionen zwischen verschiedenen Entitäten, wie z.B. Personen und Objekten, und helfen zu verstehen, wie verschiedene Elemente eines Verbrechens zusammenhängen.

Der Einsatz von Netzwerken und Graphen bietet eine klare und visuelle Darstellung komplexer Beziehungen, die es Ermittlern erleichtern kann, Zusammenhänge zu erkennen und zu analysieren. Sie ermöglichen die Identifizierung von Mustern und Anomalien in den Daten, die auf verdächtiges Verhalten hinweisen können. Zudem können sie Informationen aus verschiedenen Datenquellen

integrieren, was eine umfassendere Analyse ermöglicht. Netzwerke können dynamisch aktualisiert werden, um neue Informationen und Entwicklungen in Echtzeit zu berücksichtigen. Allerdings gibt es auch einige Nachteile. Die Erstellung und Analyse von Netzwerken und Graphen kann komplex und zeitaufwendig sein, insbesondere bei großen Datenmengen. Die Genauigkeit der Analyse hängt stark von der Qualität und Vollständigkeit der verfügbaren Daten ab, da unvollständige oder fehlerhafte Daten zu falschen Schlussfolgerungen führen können. Die Verarbeitung und Analyse großer Netzwerke erfordert erhebliche Rechenressourcen, was die Effizienz und Geschwindigkeit der Ermittlungen beeinträchtigen kann. Schließlich müssen die Ergebnisse der Netzwerkanalyse sorgfältig interpretiert werden, um Missverständnisse und Fehlinterpretationen zu vermeiden.

5.4 Listen und Tabellen

Auch traditionelle Methoden wie Listen und Tabellen haben weiterhin ihre Relevanz in der forensischen Analyse. Sie bieten eine strukturierte und übersichtliche Darstellung von Datenpunkten und ermöglichen den schnellen Vergleich und die Analyse spezifischer Informationen. Listen und Tabellen sind besonders nützlich, um große Datenmengen systematisch zu organisieren und zugänglich zu machen. Für diese Art der Visualisierung können alle Metadaten herangezogen und verknüpft werden. Es ist jedoch fraglich, ob beispielsweise eine Koordinate als Zahl von einem Anwender einfach interpretiert werden kann.

Eine Methode ist die Erstellung von Ereignislisten, die chronologisch Vorfälle aufzeichnen, die in den Videodateien identifiziert wurden. Diese Listen enthalten in der Regel Informationen wie Datum, Uhrzeit, Ort und eine kurze Beschreibung des Ereignisses, was dabei hilft, den zeitlichen Ablauf von Ereignissen nachzuvollziehen und Muster zu erkennen. Eine andere Anwendung sind Objekttabellen, die Details zu verschiedenen Objekten erfassen, die in den Videos identifiziert wurden. Solche Tabellen können Informationen wie Objektart, Farbe, Größe, Position im Bild und andere relevante Merkmale enthalten. Sie unterstützen die systematische Analyse und das Tracking von Objekten über verschiedene Videosequenzen hinweg. Ein weiteres Beispiel sind Personenlisten, die die identifizierten Personen in den Videodateien dokumentieren. Diese Listen können Daten wie Namen, sofern diese bekannt sind, körperliche Merkmale, Kleidung und Bewegungsmuster enthalten. Sie sind besonders nützlich für die Verfolgung von Personenbewegungen und zur Verknüpfung von Personen mit bestimmten Ereignissen.

Listen und Tabellen zeigen eine klare und strukturierte Darstellung von Daten, was die Übersichtlichkeit erhöht und die Analyse erleichtert. Durch die systematische Organisation von Informationen ermöglichen sie eine schnelle und effiziente Datenverarbeitung und -analyse. Sie erleichtern den direkten Vergleich von Datenpunkten, was die Identifikation von Mustern und Anomalien unterstützt. Zudem bieten sie eine gute Möglichkeit zur Dokumentation von Untersuchungsergebnissen, was für die Nachvollziehbarkeit und spätere Überprüfungen wichtig ist. Die Nachteile sind jedoch, dass Listen und Tabellen nur begrenzt tiefergehende Informationen vermitteln können. Komplexe Zusammenhänge und dynamische Interaktionen lassen sich nicht immer adäquat darstellen. Sie sind statische Darstellungsformen und eignen sich weniger für die Darstellung dynamischer Prozesse oder Veränderungen über die Zeit hinweg. Das Erstellen und Pflegen von Listen und Tabellen kann zeitaufwendig sein, insbesondere bei großen Datenmengen. Es erfordert sorgfältige Arbeit, um

sicherzustellen, dass die Daten korrekt und vollständig erfasst werden. Im Vergleich zu grafischen Darstellungen wie Netzwerken oder Diagrammen bieten Listen und Tabellen weniger visuelle Anreize und können bei umfangreichen Datenmengen schnell unübersichtlich werden.

5.5 VR und 3D-Darstellungen

Es gibt verschiedene Ansätze, wie VR und 3D-Darstellungen, die in der Videoforensik eingesetzt werden können. Eine Methode ist die Tatortrekonstruktion, bei der VR-Technologie verwendet wird, um Einsatzorte in einer dreidimensionalen, virtuellen Umgebung darzustellen. Ermittler können die Szene in 360 Grad betrachten und sich frei darin bewegen, um verschiedene Perspektiven einzunehmen. Dies hilft, den Tatort detaillierter zu analysieren und wichtige Hinweise zu identifizieren. Eine andere Methode ist die Erstellung von 3D-Modellen von Objekten und Personen, die in den Videos identifiziert wurden. Diese Modelle können aus Videodateien oder durch 3D-Scans erstellt werden und bieten eine detaillierte Ansicht, die in zwei Dimensionen nicht möglich wäre. VR ermöglicht es auch, Ereignisse zu simulieren und verschiedene Szenarien durchzuspielen. Ermittler können Bewegungen und Interaktionen von Personen nachstellen, um die Plausibilität von Zeugenaussagen zu überprüfen oder um mögliche Abläufe eines Verbrechens zu verstehen.

Durch die Verknüpfung von Metadaten, beispielsweise in Form von GPS-Daten und Zeitstempeln kann im virtuellen Raum eine Echtzeitsimmulation erzeugt werden. Durch beispielsweise Personenerkennung könnten Personen, die von besonderem Interesse sind, farblich oder durch einen Rahmen hervorgehoben werden, damit diese in komplexen Geschehen besser beobachtet werden können. Ein weiteres Ergebnis der Verknüpfung von Zeitstempeln und GPS-Daten ist Sammlung von Videos aus verschiedenen Kameras. Diese Aufnahmen können synchronisiert werden, um ein umfassenderes Bild eines Ereignisses zu erstellen. In einer virtuellen Umgebung könnte dies bedeuten, dass Aufnahmen von mehreren Kameras, die ein und dasselbe Ereignis aus unterschiedlichen Blickwinkeln aufgezeichnet haben, kombiniert und in Echtzeit dargestellt werden.

Die Vorteile dieser Technologie sind, dass VR eine immersive Erfahrung bietet, die es Ermittlern ermöglicht, tief in die Szenarien einzutauchen und Details zu erkennen, die in zweidimensionalen Darstellungen möglicherweise übersehen werden. Die interaktive Natur von VR und 3D-Darstellungen ermöglicht es, Szenarien aus verschiedenen Blickwinkeln zu betrachten und Hypothesen zu testen. 3D-Modelle bieten eine detaillierte Analyse von Objekten und Personen, was die Genauigkeit der forensischen Untersuchung verbessert. Die Möglichkeit, Ereignisse zu simulieren, hilft bei der Überprüfung von Zeugenaussagen und der Rekonstruktion von Tatabläufen. Allerdings gibt es auch einige Nachteile. Die Erstellung und Nutzung von VR und 3D-Darstellungen erfordert erhebliche finanzielle und technische Ressourcen. Hochwertige VR-Ausrüstung und spezialisierte Software sind teuer. Die Entwicklung und Implementierung von VR und 3D-Technologien erfordert technisches Know-how und kann komplex sein. Nicht alle Ermittlungsbehörden haben Zugang zu den notwendigen Ressourcen, um VR und 3D-Darstellungen effektiv zu nutzen. Die Genauigkeit und Nützlichkeit von VR und 3D-Darstellungen hängen stark von der Qualität der zugrunde liegenden Daten ab. Schlechte Daten können zu ungenauen Modellen und Simulationen führen.

5.6 Zusammenfassung und Einordnung

Die vorgestellten Methoden zur Visualisierung verknüpfter Videoinformationen bieten vielfältige Möglichkeiten, um komplexe Daten und große Datenmengen verständlich und nutzbar zu machen. Zeitliche Visualisierungen, interaktive Karten, Netzwerke und Graphen, Listen und Tabellen sowie VR und 3D-Darstellungen ergänzen sich gegenseitig und ermöglichen eine umfassende Analyse. Die Wahl der geeigneten Visualisierungstechniken hängt von den spezifischen Anforderungen und den verfügbaren Daten ab. Durch die Kombination dieser Methoden können Ermittler tiefere Einblicke gewinnen, Muster und Anomalien erkennen und fundierte Entscheidungen treffen, was die Effektivität und Effizienz forensischer Untersuchungen steigern kann.

Die bereits betrachteten bedeutenden Ereignisse haben gezeigt, wie wichtig die Analyse von Videodateien in der Videoforensik ist. Durch die Verknüpfung könnte man möglicherweise diese Analysen noch effizienter gestalten. Insbesondere bei komplexen Fällen wie dem Boston-Marathon-Attentat können verschiedene Visualisierungstechniken zur Aufklärung beitragen. Eine Objekttabelle wäre nützlich gewesen, um Details zu verschiedenen Gegenständen vor der Explosion zu erfassen, die in den Videos identifiziert wurden, wie z.B. der Rucksack, und die Kleidungsstücke der Verdächtigen. Netzwerke und Graphen, speziell Personengraphen, hätten die Beziehungen und Interaktionen zwischen den Verdächtigen und möglichen Komplizen sichtbar machen können. Interaktive Karten wären sinnvoll für die präzise Nachverfolgung der Bewegungen der Täter gewesen, sowohl während als auch nach der Tat. Diese Kombination von Visualisierungstechniken hätte die Identifizierung der Tsarnaev-Brüder durch die Analyse ihrer Bewegungen und ihres Verhaltens in der Menge erleichtern und somit möglicherweise die Ermittlungen beschleunigen und vereinfachen können.

Auch im Fall von Anis Amri, dem Täter des Berliner Weihnachtsmarkt-Anschlags, hätten verschiedene Visualisierungstechniken helfen können. Interaktive Karten wären nützlich gewesen, um die Bewegungen von Amri vor, während und nach der Tat nachzuverfolgen. Durch die Nutzung von GPS-Daten und interaktiven Kartensystemen hätte die Route des Täters präzise nachgezeichnet werden können, was bei der Identifizierung von Fluchtwegen und möglichen Komplizen von großem Nutzen ist. Heatmaps auf diesen Karten hätten angezeigt, in welchen Bereichen Amri sich länger aufgehalten hat, was wertvolle Hinweise auf seine Planung und mögliche Unterstützer hätte geben können. Eine Personenliste wäre hilfreich gewesen, um Details zu den beobachteten Aufenthaltsorten und Interaktionen von Amri systematisch zu dokumentieren, um Muster und Zusammenhänge zu erkennen. Zudem hätte eine Objekttabelle genutzt werden können, um Details zu Gegenständen zu erfassen, die Amri vor und während der Tat bei sich hatte, wie etwa Kleidung, der Lastwagen und mögliche weitere Ausrüstungsgegenstände.

Netzwerke und Graphen, insbesondere die soziale Netzwerkanalyse, hätten bei den Ermittlungen zum NSU eine Hilfe darstellen können, um das große Geflecht an Unterstützern und Mitgliedern besser darzustellen. Soziale Netzwerkanalysen helfen, die Kommunikationswege und Verbindungen zwischen den Tätern und ihren Unterstützern offenzulegen. Diese Visualisierungen bieten eine klare und strukturierte Übersicht über die komplexen Beziehungsgeflechte und können dazu beitragen, versteckte Verbindungen und unterstützende Netzwerke aufzudecken. Bewegungsgraphen hätten genutzt werden können, um die Aktivitäten und Bewegungen der NSU-Mitglieder über verschiedene Tatorte und Zeiträume hinweg darzustellen, was entscheidend für die umfassende Aufklärung ihrer

Taten ist. Ereignisgraphen wären sinnvoll gewesen, um die zeitlichen und räumlichen Zusammenhänge zwischen verschiedenen Aktionen und Tatorten der NSU darzustellen, was die Rekonstruktion der Abfolge von Ereignissen erleichtert.

Zusammengefasst tragen diese Visualisierungstechniken nicht nur zur Effizienz und Genauigkeit forensischer Untersuchungen bei, sondern auch zur Aufklärung und Dokumentation von komplexen Straftaten. Durch die gezielte Anwendung dieser Methoden können Ermittler fundierte Entscheidungen treffen, schneller auf kritische Ereignisse reagieren und letztendlich die öffentliche Sicherheit verbessern.

6 Ergebnisse und Diskussion

In diesem Kapitel werden die zentralen Ergebnisse der theoretischen Untersuchung und die praktischen Anwendungsmöglichkeiten der vorgestellten Methoden zur Verknüpfung von Videodateien auf der Metadatenebene im Bereich der Videoforensik zusammengefasst. Ein Schwerpunkt dieses Kapitels liegt auf der Diskussion zukünftiger Entwicklungen, die in der Videoforensik zu erwarten sind. Hierzu zählen technologische Fortschritte, die Weiterentwicklung von Algorithmen zur Metadatenverarbeitung sowie rechtliche und ethische Aspekte, die bei der Anwendung dieser Technologien eine Rolle spielen. Abschließend folgt ein Gesamtfazit zu dem Thema.

6.1 Zusammenfassung

In dieser Arbeit wurde zunächst ein schematischer Ablauf zur Verknüpfung von Videodateien auf Metadatenebene beschrieben, der aus fünf wesentlichen Schritten besteht und eine zentrale Rolle in der forensischen Analyse von Videodateien spielt.

- **Datensammlung:** Der erste Schritt im Prozess ist die systematische Sammlung der relevanten Videodateien. Diese Daten können aus verschiedenen Quellen stammen. Die Datensammlung bildet die Grundlage für die anschließende Analyse, indem sie sicherstellt, dass alle potenziell relevanten Videos in den Prozess einbezogen werden.
- **Extraktion von Metadaten:** Nach der Sammlung der Daten erfolgt die Extraktion von Metadaten aus den Videodateien. Für die Extraktion wurden Frameworks, wie ExifTool für EXIF-Daten bzw. TensorFlow und PyTorch für die inhaltsbasierten Metadaten vorgestellt, die eine automatisierte und effiziente Auslese der benötigten Informationen ermöglichen. Diese Metadaten bilden die Grundlage für die nachfolgenden Analyseschritte.
- **Datenanalyse:** In diesem Schritt werden die extrahierten Metadaten analysiert, um Muster und relevante Informationen zu identifizieren. Es kommen Methoden des maschinellen Lernens und der inhaltsbasierten Analyse zum Einsatz, um beispielsweise Personen oder Objekte zu erkennen und deren Bewegungen nachzuvollziehen. CNNs sind im Bereich der inhaltsbasierten Metadatengewinnung wie Objekterkennung und Personenerkennung ein Standard. Bei der Anomalierkennung wurden Anomalib und DifferNet vorgestellt. Die Schattenerkennung wurde durch den Triple-cooperative-Ansatz und dem SSIS-Track-Rahmen beschrieben.
- **Datenverknüpfung:** Die Verknüpfung der analysierten Daten ist der nächste zentrale Schritt. Hierbei werden die gewonnenen Informationen miteinander verbunden, um ein kohärentes Gesamtbild zu erstellen. Diese Verknüpfung ist entscheidend, da sie es ermöglicht, verschiedene Perspektiven und zeitliche Abfolgen zusammenzuführen und so Zusammenhänge zwischen verschiedenen Videoaufnahmen herzustellen, die in isolierter Betrachtung nicht erkennbar wären.
- **Datenvisualisierung:** Abschließend werden die Ergebnisse der Verknüpfung und Analyse in einer verständlichen und interaktiven Form visualisiert. Die Visualisierung kann in Form von interaktiven Karten, Zeitachsen oder 3D-Umgebungen erfolgen, um die geografische und zeitliche Verteilung von Ereignissen klar darzustellen. Diese Darstellungen erleichtern es den Ermittlern, komplexe Daten zu interpretieren und fundierte Entscheidungen zu treffen.

Die Analyse von Videodateien ist grundsätzlich von entscheidender Bedeutung, wie es in den behandelten bedeutenden Ereignissen bereits deutlich wurde. Durch die systematische Untersuchung von Videomaterial konnten in der Vergangenheit wichtige Hinweise gewonnen und komplexe Sachverhalte aufgeklärt werden. Doch die Verknüpfung dieser Videodateien auf Metadatenebene geht noch einen entscheidenden Schritt weiter. Sie bietet nicht nur die Möglichkeit, verschiedene Perspektiven und zeitliche Abfolgen zusammenzuführen, sondern erleichtert und beschleunigt die gesamte forensische Arbeit erheblich. Durch die gezielte Verknüpfung können Ermittler effizienter arbeiten, indem sie relevante Informationen schneller identifizieren und Zusammenhänge erkennen, die sonst möglicherweise übersehen worden wären. Dies führt zu einer erheblichen Steigerung der Genauigkeit und Geschwindigkeit der Analyse, was insbesondere in zeitkritischen Untersuchungen von unschätzbarem Wert ist.

6.2 Zukunftsausblick und Einschätzung

Die Verknüpfung von Videodateien auf der Metadatenebene steht vor zahlreichen spannenden Entwicklungen, die das Potenzial haben, die Effizienz und Genauigkeit der forensischen Analyse erheblich zu verbessern. Im Folgenden werden einige der vielversprechendsten Trends und technologischen Fortschritte erläutert.

Künstliche Intelligenz und maschinelles Lernen:
Die Integration fortschrittlicher Algorithmen des maschinellen Lernens und der künstlichen Intelligenz wird eine entscheidende Rolle bei der Weiterentwicklung der Metadatenverknüpfung spielen. Zukünftige Systeme könnten in der Lage sein, komplexe Muster und Zusammenhänge in Videodateien noch besser zu erkennen und automatisch zu analysieren. Insbesondere Deep Learning-Modelle könnten genutzt werden, um die Genauigkeit der Erkennung von Objekten, Personen und Ereignissen weiter zu steigern und so die Verknüpfung der Daten zu optimieren. Dennoch wird die manuelle Tätigkeit eines Ermittlers oder Sachbearbeiters nicht vollständig zu ersetzen sein. Insbesondere in Strafverfahren ist es dringend notwendig, dass ein Mensch die Daten überprüft und bewertet, um die Genauigkeit und Rechtmäßigkeit der Analysen sicherzustellen.

Verbesserte Datenintegration und -harmonisierung:
Eine weitere Möglichkeit der zukünftigen Entwicklung ist die Verbesserung der Integration und Harmonisierung von Daten aus verschiedenen Quellen. Dies umfasst nicht nur die Verknüpfung von Metadaten aus unterschiedlichen Videoformaten oder Bildern, sondern auch die Einbindung weiterer Datenquellen wie soziale Medien, GPS-Daten und Sensorinformationen. So könnten beispielsweise automatisierte Tools entwickelt werden, die anhand von Schlagworten gezielt nach relevanten Bildern und Videos auf Social Media Plattformen suchen und diese herunterladen. Diese Technologie würde es Ermittlern ermöglichen, umfangreiche Daten aus sozialen Medien in ihre Analysen einzubeziehen. Jedoch stellt sich hierbei die Herausforderung, dass viele Social Media Plattformen die EXIF-Daten der hochgeladenen Bilder und Videos entfernen, um die Privatsphäre der Nutzer zu schützen. Dies kann die Analyse und Verknüpfung der Daten erschweren, da wichtige Metadaten wie Datum, Uhrzeit und Standortinformationen fehlen. Ein weiteres Problem ist die Echtheit dieser Inhalte. In der digitalen Welt, insbesondere auf Social Media Plattformen, besteht ein Risiko, dass Inhalte manipuliert werden, sei es durch die Erstellung von *Fake News* oder die Verwendung von gefälschten Videos, die beispielsweise mithilfe von Technologien wie Deepfakes erstellt wurden. Diese gefälschten Inhalte können realistische, aber falsche Informationen verbreiten, die forensische Ermittlungen in die

verfälschen könnten. Daher wird es immer wichtiger, Mechanismen zu entwickeln, die nicht nur die Authentizität von Bildern und Videos überprüfen können, sondern auch gezielt gefälschte Inhalte erkennen.

Echtzeitanalyse und -verarbeitung:
Die Fähigkeit zur Echtzeitanalyse und -verarbeitung von Videodateien wird in der Zukunft immer wichtiger werden. Technologien, die es ermöglichen, Metadaten in Echtzeit zu extrahieren und zu verknüpfen, könnten dazu beitragen, sofortige Entscheidungen zu treffen und schneller auf kritische Ereignisse zu reagieren. Dies wäre insbesondere in Bereichen wie der öffentlichen Sicherheit und bei der Terrorismusbekämpfung von großem Nutzen. Allerdings muss dabei kritisch hinterfragt werden, ob eine solche Überwachung nicht einen zu großen Eingriff in die Privatsphäre darstellt. Die kontinuierliche Überwachung und Echtzeitanalyse könnten das Gefühl der ständigen Beobachtung bei den Bürgern verstärken und die persönlichen Freiheiten einschränken. Es ist daher unerlässlich, dass klare rechtliche Rahmenbedingungen und ethische Richtlinien geschaffen werden, um sicherzustellen, dass die Balance zwischen öffentlicher Sicherheit und individuellen Datenschutzrechten gewahrt bleibt.

Cloud-basierte Lösungen und verteilte Systeme:
Die Nutzung von Cloud-basierten Lösungen und verteilten Systemen könnte die Verarbeitung großer Mengen an Videodateien effizienter und kostengünstiger gestalten. Durch den Einsatz von Cloud-Technologien können Analyseprozesse skaliert und Ressourcen flexibel an die Anforderungen angepasst werden. Dies würde es ermöglichen, auch sehr große Datenmengen zeitnah zu verarbeiten und zu analysieren. Allerdings stellt die Nutzung von Cloud-Diensten in Strafprozessen eine erhebliche Herausforderung dar. Die Server dieser Dienste befinden sich oft in anderen Ländern, was rechtliche und datenschutztechnische Probleme aufwerfen kann. Selbst wenn die Daten in Deutschland gehostet werden, müssen sie so geschützt werden, dass sie keinem unbefugten Dritten in die Hände fallen können. Dies erfordert strenge Sicherheitsmaßnahmen und die Einhaltung gesetzlicher Vorschriften, um die Integrität und Vertraulichkeit der Daten zu gewährleisten.

Datenschutz und ethische Richtlinien:
Eine der größten Entwicklungen, aber gleichzeitig auch ein kritischer Punkt, ist die Etablierung robuster Datenschutz- und ethischer Richtlinien parallel zur Weiterentwicklung der Technologien zur Metadatenverknüpfung. Zukünftige Systeme müssen in der Lage sein, den Schutz der Privatsphäre zu gewährleisten und gleichzeitig die rechtlichen Rahmenbedingungen einzuhalten. Die Entwicklung von Technologien, die eine anonyme Analyse und Verknüpfung von Daten ermöglichen, könnte hierbei eine wichtige Rolle spielen. Der Schutz der Privatsphäre und der personenbezogenen Daten ist von zentraler Bedeutung, insbesondere bei der Verarbeitung und Analyse von Videodateien, die sensible Informationen enthalten können. Es ist unerlässlich, dass alle verwendeten Technologien und Verfahren den strengen Datenschutzgesetzen und -verordnungen entsprechen, wie beispielsweise der DSGVO in der Europäischen Union. Dies beinhaltet die Implementierung technischer und organisatorischer Maßnahmen, die sicherstellen, dass Daten nur von autorisierten Personen eingesehen und verarbeitet werden können.

Ein weiterer wesentlicher Aspekt ist die Schaffung von Systemen, die in der Lage sind, Daten anonym zu analysieren und zu verknüpfen. Durch den Einsatz von Techniken zur Anonymisierung und Pseudonymisierung können personenbezogene Daten geschützt und gleichzeitig wertvolle Analysen durchgeführt werden. Diese Technologien könnten beispielsweise dafür sorgen, dass identifizierende

Informationen entfernt oder verschlüsselt werden, bevor die Daten in Analyseprozesse einfließen. Dies würde es ermöglichen, die Privatsphäre der betroffenen Personen zu schützen, ohne die Effektivität der forensischen Untersuchungen zu beeinträchtigen. Es muss außerdem klar kommuniziert werden, wie und zu welchem Zweck Daten erfasst, verarbeitet und genutzt werden. Betroffene Personen sollten über ihre Rechte informiert werden und die Möglichkeit haben, der Verarbeitung ihrer Daten zu widersprechen oder diese zu korrigieren. Dies erfordert eine enge Zusammenarbeit zwischen Technologieentwicklern, Datenschutzbeauftragten und rechtlichen Experten, um sicherzustellen, dass die entwickelten Systeme sowohl effektiv als auch ethisch vertretbar sind. Allerdings gibt es Situationen, wie beispielsweise die Videoüberwachung auf öffentlichen Veranstaltungen oder Festen, in denen ein Widerspruch gegen die Datenerfassung praktisch nicht umsetzbar ist. In solchen Fällen bleibt betroffenen Personen oft nur die Wahl, der Veranstaltung fernzubleiben, wenn sie der Überwachung nicht zustimmen wollen.

6.3 Fazit

Durch den vorgestellten schematischen Ablauf, der die Schritte der Datensammlung, Metadatenauswertung, Analyse, Verknüpfung und Visualisierung umfasst, bietet die Arbeit eine strukturierte und fundierte Herangehensweise zur effizienten Verarbeitung und Interpretation von Videodateien. Die in der Arbeit beschriebenen Techniken und Modelle können Ermittlern helfen, schneller und genauer auf relevante Informationen in großen Datensätzen zuzugreifen. Darüber hinaus hat die Arbeit auch die Herausforderungen bei der Verknüpfung und Analyse von Videodateien nicht außer Acht gelassen. Die Betrachtung dieser Herausforderungen ist entscheidend, um mögliche Hindernisse frühzeitig zu erkennen und Strategien zu deren Überwindung zu entwickeln. Die Arbeit zeigt, dass die Auseinandersetzung mit technischen, rechtlichen und praktischen Herausforderungen nicht nur die Qualität der forensischen Analyse verbessern kann, sondern auch dazu beiträgt, dass die entwickelten Methoden in der Praxis erfolgreich angewendet werden.

Selbst wenn sich die Inhalte zukünftiger forensischer Untersuchungen ändern oder neue Technologien entwickelt werden, bleibt der in dieser Arbeit vorgestellte Ansatz flexibel und anpassungsfähig. Er kann kontinuierlich erweitert und angepasst werden, um den sich verändernden Anforderungen gerecht zu werden. Somit leistet diese Arbeit einen Beitrag zur Weiterentwicklung der forensischen Analyse dar und bietet ein Fundament, auf dem zukünftige Methoden und Techniken aufbauen können.

Quellenverzeichnis

[1] P. Bestagini, K. M. Fontani, S. Milani u. a., „An overview on video forensics“, in *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, 2012, S. 1229–1233.

[2] *Describing the details that matter: the importance of video metadata - Secure Insights*. Adresse: https://www.axis.com/blog/secure-insights/video-metadata/ (besucht am 26. 05. 2024).

[3] *IT-forensische Relevanz von Metadaten in Bilddateien*. Adresse: https://www.dr-datenschutz.de/it-forensische-relevanz-von-metadaten-in-bilddateien/ (besucht am 03. 08. 2024).

[4] R. G. Mani, R. Parthasarathy, S. Eswaran und P. B. Honnavalli, „A Survey on Digital Image Forensics: Metadata and Image forgeries“, 2022.

[5] M. Mühling, M. Meister, N. Korfhage u. a., „Content-based video retrieval in historical collections of the German Broadcasting Archive“, *International Journal on Digital Libraries*, Jg. 20, Nr. 2, S. 167–183, Juni 2019, ISSN: 1432-1300. DOI: https://doi.org/10.1007/s00799-018-0236-z.

[6] *Video-Forensik - Medien-SachverständigerMedien-Sachverständiger*. Adresse: https://www.medien-sachverstaendiger.de/video-forensik/ (besucht am 12. 05. 2024).

[7] *Details, auf die es ankommt: die Bedeutung von Video-Metadaten*. Adresse: https://www.axis.com/blog/secure-insights-de/details-auf-die-es-ankommt-die-bedeutung-von-video-metadaten/ (besucht am 12. 05. 2024).

[8] *FLORIDA1: Analyse von Videomassendaten im Kontext terroristischer Anschläge • CRISIS PREVENTION • Fachportal für Gefahrenabwehr, Innere Sicherheit und Katastrophenhilfe*. Adresse: https://crisis-prevention.de/sicherheit/florida1-analyse-von-videomassendaten-im-kontext-terroristischer-anschlaege.html (besucht am 12. 05. 2024).

[9] R. Gonzalez und R. Woods, *Digital Image Processing : Global Edition*. Pearson Deutschland, Okt. 2017, ISBN: ISBN 9781292223049.

[10] R. Szeliski, *Computer Vision: Algorithms and Applications* (Texts in Computer Science). Springer International Publishing, 2022, ISBN: 978-3-030-34372-9.

[11] I. Goodfellow, Y. Bengio und A. Courville, *Deep Learning* (Adaptive Computation and Machine Learning series). MIT Press, 2016, ISBN: 978-0-262-03561-3.

[12] H. Garcia-Molina, J. Ullman und J. Widom, *Database Systems: The Complete Book* (Pearson International edition). Pearson Prentice Hall, 2009, ISBN: 978-0-13-187325-4.

[13] G. Sreenu und M. A. Saleem Durai, „Intelligent video surveillance: a review through deep learning techniques for crowd analysis“, *Journal of Big Data*, Jg. 6, Nr. 1, S. 48, Juni 2019, ISSN: 2196-1115. DOI: https://doi.org/10.1186/s40537-019-0212-5.

[14] F. Focus, *Digital Forensics as a Big Data Challenge*, en-US, Aug. 2017. Adresse: https://www.forensicfocus.com/articles/digital-forensics-as-a-big-data-challenge/ (besucht am 26. 05. 2024).

[15] V. Roussev, C. Quates und R. Martell, „Real-time digital forensics and triage“, *Digital Investigation*, Jg. 10, Nr. 2, S. 158–167, 2013, ISSN: 1742-2876. DOI: https://doi.org/10.1016/j.diin.2013.02.001.

[16] A. M. Burton, S. Wilson, M. Cowan und V. Bruce, „Face recognition in poor-quality video: Evidence from security surveillance.“, *Psychological Science*, Jg. 10, Nr. 3, S. 243–248, 1999, Place: United Kingdom Publisher: Blackwell Publishing, ISSN: 1467-9280(Electronic),0956-7976(Print). DOI: https://doi.org/10.1111/1467-9280.00144.

[17] G. Mercier, F. Markatopoulou, R. Cozien u. a., „Detecting Manipulations in Video“, in *Video Verification in the Fake News Era*, V. Mezaris, L. Nixon, S. Papadopoulos und D. Teyssou, Hrsg., Cham: Springer International Publishing, 2019, S. 161–189, ISBN: 978-3-030-26752-0. DOI: https://doi.org/10.1007/978-3-030-26752-0_6.

[18] *United States - Privacy Protection - Ethical Digital Forensics – Balancing Investigation Procedures With Privacy Concerns*. Adresse: https://www.mondaq.com/unitedstates/privacy-protection/1306880/ethical-digital-forensics--balancing-investigation-procedures-with-privacy-concerns (besucht am 26. 05. 2024).

[19] *Überwachungstechnik: Video-Forensik, die neue Waffe der US-Polizei - WELT*. Adresse: https://www.welt.de/wissenschaft/article115536885/Video-Forensik-die-neue-Waffe-der-US-Polizei.html (besucht am 23. 04. 2024).

[20] *FBI releases photos of marathon suspects. Vindication for surveillance video? - CSMonitor.com*. Adresse: https://www.csmonitor.com/USA/2013/0418/FBI-releases-photos-of-marathon-suspects.-Vindication-for-surveillance-video (besucht am 04. 08. 2024).

[21] I. COMMUNITY, C. I. AGENCY, D. O. JUSTICE und D. O. H. SECURITY, „Unclassified Summary of Information Handlung and Sharing Prior to the April 15, 2013 BOSTON MARATHON BOMBINGS“, Techn. Ber., 2014. Adresse: https://oig.justice.gov/reports/2014/s1404.pdf.

[22] *Boston Marathon bombing: The attack, the arrest, the recovery | 60 Minutes Full Episodes - YouTube*. Adresse: https://www.youtube.com/watch?v=loeRbGg7Ef0&t=1498s (besucht am 27. 04. 2024).

[23] *How video analytics helps reconstruct Boston Marathon bombings - Route Fifty*. Adresse: https://www.route-fifty.com/cybersecurity/2013/04/how-video-analytics-helps-reconstruct-boston-marathon-bombings/281331/ (besucht am 27. 04. 2024).

[24] *Boston Marathon Bombing and the Video Forensic Process*. Adresse: https://www.videoforensicexpert.com/boston-marathon-bombing-and-the-video-forensic-process/ (besucht am 27. 04. 2024).

[25] *Advancements in Face Recognition: Reflecting on the Boston Marathon Bombing*. Adresse: https://roc.ai/2023/04/17/looking-back-at-boston-marathon-bombing-a-decade-of-face-recognition-advancement/ (besucht am 27. 04. 2024).

[26] *Nach dem Anschlag von Berlin - Öffentliche Fahndung nach Anis Amri*. Adresse: https://www.deutschlandfunk.de/nach-dem-anschlag-von-berlin-oeffentliche-fahndung-nach-100.html (besucht am 04. 08. 2024).

[27] *Berlin Christmas market attack: Tunisian man who dined with Anis Amri on eve of massacre arrested as probe continues | The Independent | The Independent*. Adresse: https://www.independent.co.uk/news/world/europe/berlin-christmas-market-lorry-attack-isis-anis-amri-tunisian-man-arrested-accomplice-investigation-network-germany-a7510306.html (besucht am 28. 04. 2024).

[28] D. J. Geerlings, *Schlussbericht des Parlamentarischen Untersuchungsausschusses I („Fall Amri")*, März 2022. Adresse: https://www.landtag.nrw.de/Dokumentenservice/portal/WWW/dokumentenarchiv/Dokument/MMD17-16890.pdf;jsessionid=9DBF0F1CB528E605D762286539241FBB.

[29] *Deutscher Bundestag - Zeuge berichtet über Videoauswertung*. Adresse: https://www.bundestag.de/webarchiv/presse/hib/2020_03/687534-687534 (besucht am 28. 04. 2024).

[30] *NSU-Morde - Aufklärung durch Geheimnisverrat?* Adresse: https://www.deutschlandfunk.de/nsu-morde-aufklaerung-durch-geheimnisverrat-100.html (besucht am 04. 08. 2024).

[31] L. McGowan, „Right-Wing Violence in Germany: Assessing the Objectives, Personalities and Terror Trail of the National Socialist Underground and the State's Response to It", en, *German Politics*, Jg. 23, Nr. 3, S. 196–212, Juli 2014, ISSN: 0964-4008, 1743-8993. Adresse: https://doi.org/10.1080/09644008.2014.967224.

[32] *Viele Spuren, keine Verdächtigen*. Adresse: https://www.fr.de/politik/viele-spuren-keine-verdaechtigen-11077293.html (besucht am 28. 04. 2024).

[33] *Exchangeable Image File Format (Exif) Family*, eng, web page, Nov. 2023. Adresse: https://www.loc.gov/preservation/digital/formats/fdd/fdd000618.shtml (besucht am 22. 06. 2024).

[34] R. Padilha, F. Andaló, B. Lavi, L. Pereira und A. Rocha, „Temporally Sorting Images from Real-World Events", *Pattern Recognition Letters*, Jg. 147, S. 212–219, Mai 2021. DOI: https://doi.org/10.1016/j.patrec.2021.04.027.

[35] *Identifizieren von Personen und Haustieren in der App „Fotos" auf dem iPhone*, de. Adresse: https://support.apple.com/de-de/guide/iphone/iph9c7ee918c/ios (besucht am 23. 06. 2024).

[36] *IPTC Standard - IPTC*. Adresse: https://iptc.org/standards/photo-metadata/iptc-standard/ (besucht am 06. 07. 2024).

[37] *Autopsy: Timeline Analysis*. Adresse: https://www.sleuthkit.org/autopsy/timeline.php (besucht am 14. 07. 2024).

[38] R. Bill, J. Blankenbach, M. Breunig u. a., „Geospatial Information Research: State of the Art, Case Studies and Future Perspectives", *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, Jg. 90, Nr. 4, S. 349–389, Aug. 2022, ISSN: 2512-2819. DOI: https://doi.org/10.1007/s41064-022-00217-9.

[39] *„Geocoding"-Dienst | Maps JavaScript API | Google for Developers*. Adresse: https://developers.google.com/maps/documentation/javascript/geocoding?hl=de#ReverseGeocoding (besucht am 14. 07. 2024).

[40] K. Merry und P. Bettinger, „Smartphone GPS accuracy study in an urban environment", *PLOS ONE*, Jg. 14, Nr. 7, e0219890, Juli 2019, Publisher: Public Library of Science. DOI: https://doi.org/10.1371/journal.pone.0219890.

[41] *Mastering Object Detection with YOLOv8*. Adresse: https://keylabs.ai/blog/mastering-object-detection-with-yolov8/ (besucht am 14. 07. 2024).

[42] T.-Y. Lin, P. Goyal, R. B. Girshick, K. He und P. Dollár, „Focal Loss for Dense Object Detection“, *CoRR*, Jg. abs/1708.02002, 2017. DOI: https://doi.org/10.48550/arXiv.1708.02002.

[43] S. Ren, K. He, R. B. Girshick und J. Sun, „Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks“, *CoRR*, Jg. abs/1506.01497, 2015. DOI: https://doi.org/10.48550/arXiv.1506.01497.

[44] J. K. Iype und S. Sebastian, „Comparative Analysis of State-of-the-Art Face Recognition Models: FaceNet, ArcFace, and OpenFace Using Image Classification Metrics“, in *Computational Sciences and Sustainable Technologies*, S. Aurelia, C. J., A. Immanuel, J. Mani und V. Padmanabha, Hrsg., Cham: Springer Nature Switzerland, 2024, S. 222–234, ISBN: 978-3-031-50993-3.

[45] *Body Detection using Computer Vision*. Adresse: https://blog.xmartlabs.com/blog/computer-vision-techniques-for-body-detection/ (besucht am 28. 07. 2024).

[46] M. Lupión, A. Polo-Rodríguez, P. M. Ortigosa und J. Medina-Quero, „ThermalYOLO: A Person Detection Neural Network in Thermal Images for Smart Environments“, in *Proceedings of the International Conference on Ubiquitous Computing & Ambient Intelligence (UCAmI 2022)*, J. Bravo, S. Ochoa und J. Favela, Hrsg., Cham: Springer International Publishing, 2023, S. 772–783, ISBN: 978-3-031-21333-5.

[47] L. Konz, A. Hill und F. Banaei-Kashani, „ST-DeepGait: A Spatiotemporal Deep Learning Model for Human Gait Recognition“, *Sensors*, Jg. 22, Nr. 20, 2022, ISSN: 1424-8220. DOI: https://doi.org/10.3390/s22208075.

[48] J. N. Mogan, C. P. Lee, K. M. Lim und K. S. Muthu, „Gait-ViT: Gait Recognition with Vision Transformer“, *Sensors*, Jg. 22, Nr. 19, 2022, ISSN: 1424-8220. DOI: https://doi.org/10.3390/s22197362.

[49] S. Becker, M. Heuschkel, S. Richter und D. Labudde, „COMBI: Artificial Intelligence for Computer-Based Forensic Analysis of Persons“, *KI - Künstliche Intelligenz*, Jg. 36, Nr. 2, S. 171–180, Sep. 2022, ISSN: 1610-1987. DOI: https://doi.org/10.1007/s13218-022-00761-x.

[50] A. L. B. V. e Silva, F. Simões, D. Kowerko, T. Schlosser, F. Battisti und V. Teichrieb, *Attention Modules Improve Modern Image-Level Anomaly Detection: A DifferNet Case Study*, 2024. DOI: https://doi.org/10.48550/arXiv.2401.08686. eprint: 2401.08686.

[51] S. Akcay, D. Ameln, A. Vaidya, B. Lakshmanan, N. Ahuja und U. Genc, *Anomalib: A Deep Learning Library for Anomaly Detection*, Feb. 2022. DOI: https://doi.org/10.48550/arXiv.2202.08341.

[52] Z. Chen, L. Wan, L. Zhu u. a., *Triple-cooperative Video Shadow Detection*, 2021. DOI: https://doi.org/10.48550/arXiv.2103.06533.

[53] Z. Xing, T. Wang, X. Hu, H. Wu, C.-W. Fu und P.-A. Heng, *Video Instance Shadow Detection*, _eprint: 2211.12827, 2024. Adresse: https://arxiv.org/abs/2211.12827.

[54] N. Waters, *Unsure When a Video or Photo was Taken? How to Tell by Measuring the Length of Shadows*, en-GB, Mai 2021. Adresse: https://www.bellingcat.com/resources/2021/05/18/unsure-when-a-video-or-photo-was-taken-how-to-tell-by-measuring-the-length-of-shadows/ (besucht am 04. 08. 2024).

Eidesstattliche Erklärung

Hiermit versichere ich – Gizem Yalcin – an Eides statt, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe.

Sämtliche Stellen der Arbeit, die im Wortlaut oder dem Sinn nach Publikationen oder Vorträgen anderer Autoren entnommen sind, habe ich als solche kenntlich gemacht.

Diese Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt oder anderweitig veröffentlicht.

Mittweida, 12. August 2024	
Ort, Datum	Gizem Yalcin