

BirdRecorder

Entwicklung und Erprobung eines Systems
zur Vermeidung von potenziellen Auswirkungen
auf Vögel durch die Windenergienutzung

Anton Kaifel, Ursula Amann, Marcel Zoll, Frank Sehnke,
Nico Klar, Kay Ohnmeiß und Martin Felder

BfN-Schriften

651

2023





Bundesamt für
Naturschutz

BirdRecorder

**Entwicklung und Erprobung eines Systems
zur Vermeidung von potenziellen Auswirkungen
auf Vögel durch die Windenergienutzung**

Anton Kaifel
Ursula Amann
Marcel Zoll
Frank Sehnke
Nico Klar
Kay Ohnmeiß
Martin Felder

Inhaltsverzeichnis

Inhaltsverzeichnis	3
Abbildungsverzeichnis	4
Tabellenverzeichnis	6
1 Zusammenfassung	7
2 Aufgabenstellung und Projektziel	8
3 Hardwareausstattung	9
3.1 Beschreibung des BirdRecorder-Systems IPC-1.....	9
3.2 Erkenntnisse zu Bildaufnahme und Anforderungen.....	11
3.3 Weiterentwicklung des Hardwaresystems	11
3.4 Schwenkneigekopf für Stereokameras.....	14
4 Software	15
4.1 Datenerfassung und -aufbereitung	15
4.2 Webbasierte Annotationsplattform	18
4.3 Erkennung der Vogelart	20
4.4 Stereometrie und Kalibrierung.....	27
4.4.1 Entfernungsberechnung.....	28
4.4.2 Validierung der Stereometrie	29
4.4.3 Ecken Validierung	31
4.4.4 Schachbrett Kalibrierung.....	31
4.4.5 Zusammenfassung Stereokalibrierung.....	33
4.5 Softwarearchitektur des Gesamtsystems	34
4.5.1 Statische Beobachtung	35
4.5.2 Stereobeobachtung.....	35
4.5.3 Schwenkneigekopf (SNK)	36
4.5.4 Koordinator	36
4.5.5 Monitoring	36
5 Evaluation	38
5.1 Verschneidung mit externen Vogelbeobachtungen	38
5.1.1 Datenvorverarbeitung	38
5.1.2 Statistik der detektierten Objekte	39
5.1.3 Verschneidung und Auswertung	42
5.1.4 Fazit Objekt- und Milanererkennung	46
6 Ausblick	48
Literaturverzeichnis	49

Abbildungsverzeichnis

Abb. 1:	Windmessmast Nord-West auf dem Windtestfeld in Stötten mit Kameras (rote Kreise) und Schaltschrank u.a. für Speicherung der Bilddaten (rotes Quadrat)..	10
Abb. 2:	BirdRecorder System IPC-2: Schematischer Aufbau mit IPC-Workstation, Framegrabbern und Schwenkneigekopf mit nachführbaren Stereokameras.....	12
Abb. 3:	Testinstallation des BirdRecorder Systems IPC-2 auf dem Dach des ZSW in Stuttgart mit klimatisiertem Schaltschrank, statisch montierten Kameras sowie Schwenkneigekopf mit Stereokamerapaar.....	13
Abb. 4:	Anordnung der Gesichtsfelder (FOV) der acht statischen Kameras des BirdRecorder Systems für einen 360 Grad Rundumblick.	13
Abb. 5:	Beispielbilder für die Objektarten Vögel, Fluggeräte und Sonstiges.	16
Abb. 6:	Prozesskette des BirdRecorders zur Vogel- und Vogelarterkennung (Beschreibung siehe Text).	16
Abb. 7:	Automatische Sortierung der Bilder für das Training der neuronalen Netze für die Objekt- und Arterkennung (Beschreibung siehe Text).....	17
Abb. 8:	Wie Abb. 7, jedoch für Bilder die überwiegend Vogelobjekte enthalten und anschließend noch manuell annotiert werden.	17
Abb. 9:	Webinterface der Annotationsplattform mit der Ansicht für Nicht-Ornithologen. .	18
Abb. 10:	Webinterface der Annotationsplattform mit der Ansicht für Ornithologen zur Annotation der Vogelart bzw. Artklasse.....	19
Abb. 11:	Verteilung manuell annotierter Objektarten für das Training des neuronalen Netzes NN-1.	19
Abb. 12:	128 zufällig ausgewählte Bilder aus dem unabhängigen Testdatensatz, welche die Fehlerverteilung des gesamten Testdatensatzes repräsentiert.....	23
Abb. 13:	Die 14 verbleibende, falsch kategorisierte Bilder nach Prüfung der auffälligen Labels.	24
Abb. 14:	Visualisierung der für das neuronale Netz entscheidungsrelevanten Bildbereiche mit der GradCam Methode (Beschreibung siehe Text).	25
Abb. 15:	Beispiel desselben Milans wie Abb. 14 nur drei Bilder bzw. 300 ms früher.....	25
Abb. 16:	Beispiel Bildsequenz für besonders interessante kritische Fehl kategorisierung nach Änderung der Fluglage des Milans (Beschreibung siehe Text).	26
Abb. 17:	Milan-Erkennungswahrscheinlichkeit einer Beispielsequenz ohne Tracking (blau) und mit Tracking (orange), für einen Tag mit schlechten Sichtverhältnissen und einem Milan im Abstand von 200 bis 700 Metern.	26
Abb. 18:	Wie Abb. 17, jedoch für einen Tag mit guten Sichtverhältnissen für einem Milan im Abstand von 200 bis 700 Metern.	27
Abb. 19:	Stereokamerapaar auf dem Schwenkneigekopf.....	28
Abb. 20:	Mögliche Fehlstellungen der Stereokameras (Gierwinkel), die durch Justierung ausgeglichen werden.	28
Abb. 21:	Entfernung von Referenzobjekten gegenübergestellt mit der geschätzten Entfernung, berechnet über die Pixelparallaxe (Pixelversatz).	30

Abb. 22: Fehler der Entfernungsmessung mittels Stereometrie in Abhängigkeit der Objektentfernung.	30
Abb. 23: Mögliche Verzerrungen der Bilder durch Neigungsdrift oder Rolldrift der Stereokameras.	31
Abb. 24: Automatisierte Ecken- und Kantendetektion für die Kamerakalibrierung Die dektektierten Schnittpunkte farblich gekennzeichnet.	32
Abb. 25: Entzerrte Aufnahme, nach Stereo Kalibrierung. Das Gitternetz visualisiert die Korrektur nach der Entzerrung und die Zahlen (blau) geben die Pixelkoordinaten im Bild wieder.	33
Abb. 26: Grundlegende Softwarearchitektur des BirdRecorder-Systems IPC-2.	35
Abb. 27: Dashboard-Visualisierung des Monitorings für BirdRecorder-System.	37
Abb. 28: Beispiel für die Korrektur der Altimetermessung für besenderte Vögel und Vergleich mit GPS Höhenmessungen.	38
Abb. 29: Visualisierung der Referenzdaten für die Vogeldetektion. Links: GPS, rechts: LRF. Koordinaten in Metern relativ zur Kameraposition.	39
Abb. 30: Verteilung der durch das neuronale Netz NN-2 bestimmten Klassen über die Bildfläche der Südkamera.	40
Abb. 31: Wie Abb. 30, aber für die Nordkamera.	40
Abb. 32: Anteil der Wochentage an der Gesamtanzahl der in eine Klasse fallenden ROIs. Unter "Vogel" sind hier alle Vögel incl. Milane zusammengefasst.	41
Abb. 33: Verteilung der Konfidenz über die Klassen. Da eine Klasse nur gewählt wird, wenn sie von allen vier Klassen die höchste Konfidenz hat, liegt der Minimalwert bei 0,25. Die Histogrammhöhe für den Konfidenzwert 1,0 sind der besseren Detaildarstellung wegen abgeschnitten.	41
Abb. 34: Darstellung der Verschneidung des Sichtfelds der Nordkamera mit GPS Besenderungsdaten von Rotmilanen.	43
Abb. 35: Flugroute eines kontinuierlich verfolgten Milans auf der Bildfläche der Südkamera am 17.9.2019 mit ROIs über einen Zeitraum von 85 s (Beschreibung siehe Text).	44
Abb. 36: Abschätzung einer Erkennungsrate für einzelne ROIs aller Kameras für die Kollokation mit den LRF-Daten (Erklärung siehe Text).	44
Abb. 37: Entwicklung der kumulierten Konfidenz für eine Milanklassifikation für die Sequenz aus Abb. 35.	45
Abb. 38: Wie Abb. 36, aber Erkennung bezieht sich nur auf die Klassen Vogel und Nicht-Vogel und nicht auf die Vogelart.	45
Abb. 39: Flugroute zweier kontinuierlich verfolgter Milane der Nordkamera am 18.05.2020 und die erkannten ROIs. Symbole wie in Abb. 35.	46

Tabellenverzeichnis

Tab. 1:	Anzahl der Bilder und deren Datenmenge, die mit dem BirdRecorder-System IPC-1 in den Jahren 2019, 2020 und 2021 auf dem Windenergie-testfeld in Stötten aufgenommen wurden.	11
Tab. 2:	Verteilung der Vogelarten für die 412.146 von Ornithologen annotierten Bilder. Die gelb markierten Vogelarten bzw. Artgattungen zählen in Baden-Württemberg zu den windkraftempfindlichen Vogelarten (Landesanstalt für Umwelt Baden-Württemberg, 2022).....	20
Tab. 3:	Auswertung der Entfernungsmessung mittels Stereometrie und resultierende Fehler nach Abgleich der Messungen mit den Koordinaten bzw. Entfernungen von bekannten Objekten.	29
Tab. 4:	Verteilung der klassifizierten ROIs.	39

1 Zusammenfassung

Das dreijährige Forschungsprojekt BirdRecorder, welches durch das BfN mit Mitteln des BMUV gefördert wurde, hatte die Entwicklung und Erprobung eines Systems zur Vermeidung von potentiellen Auswirkungen auf Vögel durch die Windenergienutzung zum Ziel. Der BirdRecorder stellt damit ein Antikollisionssystem für den vogelfreundlichen Betrieb von Windenergieanlagen dar. Die Zielartgruppe für die Erkennung von Vögeln mit Methoden der künstlichen Intelligenz (KI) für das BirdRecorder-System ist die Artgruppe der Milane die Rot- und Schwarzmilan umfasst. Es wurde ein erster Prototyp entwickelt und erprobt, der jedoch noch keine Schnittstelle zur Abschaltung von Windenergieanlagen aufweist.

Der BirdRecorder besteht aus fest installierten Kameras für die Erkennung von Objekten. Wird ein sich bewegendes Objekt im Luftraum um eine Windenergieanlage erkannt, wird dieses mit KI-Methoden klassifiziert, um zu entscheiden, ob es sich um einen Milan, einen anderen Vogel, ein Flugzeug oder ein sonstiges Objekt handelt. Die Reichweite der Detektion liegt bei bis zu 700 Metern. Wenn ein Milan erkannt wurde, wird das Objekt mit einem nachführbaren Stereokamerapaar mit Teleoptik erfasst und erneut klassifiziert. Die Stereokameras sind auf einem Schwenkneigekopf mit zweiachsiger Nachführung montiert. Ergibt die erneute Klassifikation mit den hochaufgelösten Stereobildern, dass es sich bei dem Objekt um einen Milan handelt, erfolgt die dreidimensionale Verortung und Verfolgung des Milans mit den Stereokameras. Damit ist die Aufzeichnung der Flugroute des Milans gewährleistet. Sollte dieser in den Gefahrenbereich einer Windenergieanlage einfliegen, könnte aufgrund der räumlichen Verortung des Milans diese abgeschaltet werden.

Für die Belernung der KI-Modelle wurden auf dem Windenergiefeld des ZSW in Stötten auf der Schwäbischen Alb über einen Zeitraum von zwei Jahren mehr als 16 Millionen Bilder aufgezeichnet. Die Daten wurden für das Training von KI-Modellen aufbereitet, sodass nach der Qualitätskontrolle insgesamt ca. 680.000 Bilder für das Training der KI-Modelle für die Objektdetektion zur Verfügung standen. Darin enthalten waren ca. 420.000 Bilder von Vögeln, die wiederum von Ornithologen manuell nach Arten und Artgruppen annotiert wurden. Dies ergab ca. 72.000 Bilder für die Artgruppe Milan. Der Rest teil sich in ca. 20 weitere Arten und Artgruppen auf. Mit den von Ornithologen annotierten Bildern wurden weitere KI-Modelle für die Erkennung der Artgruppe Milan trainiert.

Die Erkennungsrate der KI für die Objektdetektion liegt für die Klasse Vögel bei ca. 99 %. Die zweite KI für die Erkennung der Artgruppe Milan, kann auf Basis von Einzelbildern Milane von anderen Vögeln mit einer Genauigkeit von 93,1 % unterscheiden. Diese Erkennungsraten basieren auf Tests mit unabhängigen Bilddaten, die nicht für das Training der KI-Modelle eingesetzt wurden. Erfolgt die Erkennung von Milanen nicht nur auf Einzelbildern, sondern auf Bildserien von verfolgten Objekten beträgt die Erkennungswahrscheinlichkeit für die Artgruppe Milan schon nach wenigen Sekunden meist mehr als 98 %.

Mit der Entwicklung eines BirdRecorder-Prototypsystems konnte gezeigt werden, dass zum einen das Konzept von statisch montierten Kameras für die Objekterkennung und einem nachgeführten Stereokamerapaar für die Vogelarterkennung sehr gut geeignet ist und zum anderen mit den entwickelten KI-Methoden eine sehr hohe Erkennungswahrscheinlichkeit für die Artgruppe Milan erreicht wurde. Mit einer Reichweite von bis zu 700 m ist damit die rechtzeitige Abschaltung von Windenergieanlagen möglich, wenn ein Milan in den Gefahrenbereich der Anlage einfliegen sollte.

2 Aufgabenstellung und Projektziel

Vor dem Hintergrund der Vereinbarkeit von Naturschutz und der Nutzung der Windenergie und der schnellen technischen Entwicklungen bietet die Erkennung und Verortung von Vögeln an Windenergieanlagen (WEA) perspektivisch Lösungsansätze im Hinblick auf etwa ereignisbezogene Abschaltungen.

Auf dem Windenergie-Testfeld des Projekts „Wind Science and Engineering in Complex Terrain (WINSENT, FKZ 0324129)“ und der Naturschutzbegleitforschung am Windtestfeld (FKZ 3518 86 0100) soll daher als eigenständiges Forschungsmodul ein Erkennungssystem für windkraftempfindliche Vogelarten entwickelt und erprobt werden (BirdRecorder). Bei der Erkennung der Vogelart steht die Erkennung von Milanen (Rot- und Schwarzmilan) im Vordergrund.

Mit einem ersten Kamerasystem soll eine umfangreiche Bilddatenbasis von windkraftempfindlichen Vögeln erstellt werden, die am Standort des Windtestfeldes vorkommen. Diese Bilddaten bilden dann die Grundlage für die Entwicklung von Erkennungsalgorithmen mit Methoden des maschinellen Lernens, um Vogelarten möglichst sicher identifizieren zu können. Die dafür benötigten Labels für die Bilddaten sollen über ein Webinterface von Ornithologen vergeben werden. Die Erkennung konzentriert sich auf die Erfassung von Milanen (Rot- und Schwarzmilan), die zum einen am Windtestfeld häufig vorkommen und zum anderen zu den besonders kollisionsgefährdeten und planungsrelevanten Greifvögeln zählen (LAG VSW, 2015; Umweltministerkonferenz, 2020).

Mit weiteren hochauflösenden Kameras wird ein Prototypsystem für den BirdRecorder aufgebaut, der auf fest installierten Kameras und nachgeführten Stereokameras besteht. Der Prototyp des BirdRecorders soll programmiert und erprobt werden.

3 Hardwareausstattung

Für die Bildaufnahme und den Test der Vogelerkennung mit Methoden der Künstlichen Intelligenz (KI) wurden zwei verschiedene Hardwareversionen aufgebaut, programmiert und erprobt. Die erste Hardwareversion (nachfolgend IPC-1 genannt) umfasst zwei Kameras mit einer Auflösung von 13 Megapixeln (MP). Für die zweite Hardwareversion war ursprünglich eine Workstation vorgesehen, die im Turm einer der Forschungswindenergieanlagen (FWEA) betrieben werden sollte. Die Kameras sollten über den Umfang der FWEA verteilt installiert werden, ebenso das nachführbare Stereokamerapaar.

Da sich die Installation der FWEA im Vorhaben WINSENT verzögert hat, war es nicht möglich die Workstation im Turm einer der FWEA zu installieren. Stattdessen wurde für den Aufbau des BirdRecorder-Systems ein wetterfester und klimatisierter Schaltschrank entwickelt, der alle Komponenten für einen autonomen Betrieb enthält und an den bis zu acht statische Kameras und ein nachgeführtes Stereokamerapaar angeschlossen werden können. Für die Bildaufnahme und Steuerung der Stereokameras ist ein leistungsfähiger Industrie-PC integriert. Das System wird nachfolgend als IPC-2-System bezeichnet.

3.1 Beschreibung des BirdRecorder-Systems IPC-1

Für die Bildaufnahmen am Windmessmast Nord-West auf dem Windenergiefeld in Stötten wurde zunächst, in Anknüpfung an die vorbereitenden Aktivitäten aus dem Projekt NatForWinsent, ein Industrie-PC (IPC-1) mit einer Bildaufnahmeeinheit (Framegrabber) und zwei Kameras installiert und betrieben. Da keine schnelle Internetanbindung verfügbar war, wurden die Bilddaten auf einem lokalen Datenspeichersystem mit 21 TB Speicherkapazität (Network Attached Storage: NAS-System) gespeichert und regelmäßig am ZSW Standort Stuttgart auf entsprechende Massendatenspeichersysteme zur Weiterverarbeitung übertragen.

Der IPC-1 und das NAS-System wurden zusammen in einem wettergeschützten und klimatisierten Schaltschrank installiert, der auch alle notwendigen Installationen und Sicherungselemente für die angeschlossenen Geräte enthält. Im Mai 2019 erfolgte die Montage des Kamerasystems am Windmessmast Nord-West auf dem Windenergiefeld. In Abb. 1 ist der Windmessmast und die Installation der Kameras und der Schaltschrank zu sehen. Auf den untersten beidseitigen Auslegern in einer Höhe von 19 m sind die Kameras des IPC-1 BirdRecorder-Systems montiert (roter Kreise). Der Schaltschrank für den Betrieb der Kameras und Speicherung der Bilddaten ist im roten Quadrat zu erkennen.

Die optische Kalibrierung der Kameras wurde initial am Standort in Stuttgart durchgeführt und am Windmessmast nachjustiert. Dabei wurden verschiedene Aspekte berücksichtigt. Die Belichtung soll automatisch durchgeführt werden, da je nach Wetterlage, Tages- und Jahreszeit unterschiedliche Lichtverhältnisse herrschen. Dabei soll die Belichtungszeit jedoch nicht länger als 30 ms sein, um Bewegungsunschärfe von Objekten (Vögel) auf den Bildern zu vermeiden. Je weiter offen die Blende ist, desto besser können Vögel auch bei schwachen Lichtverhältnissen erkannt werden. Andererseits ist bei weit geöffneter Blende die Tiefenschärfe der Bilder nicht optimal. Nach der Analyse von Bildern mit verschiedenen Blenden- und Fokuseinstellungen, wurde für eine optimale Bildschärfe eine Blendenzahl von 6 eingestellt. Zudem ist ein automatischer Weißabgleich aktiviert. Die vorgenommenen Kameraeinstellungen werden in der Kamera selbst gespeichert und beim Einschalten der Kamera automatisch geladen.

Um die Bildaufnahmequalität zu verbessern, wurden im Mai 2020 neue Objektive an den bestehenden Kameras am Windmessmast installiert. Die bisher verwendeten Objektive

wiesen eine Vignettierung auf, das heißt eine leichte Abschattung der Bilder zum Rand hin. Mit den neu installierten Objektiven tritt dieser Effekt nicht mehr auf. Um vermehrt Bildmaterial mit groß abgebildeten Vögeln zu sammeln, wurde für eine der beiden Kameras ein Objektiv mit 25 mm Brennweite installiert, anstelle der bisher verwendeten Brennweite von 16 mm.



Abb. 1: Windmessmast Nord-West auf dem Windtestfeld in Stötten mit Kameras (rote Kreise) und Schaltschrank u.a. für Speicherung der Bilddaten (rotes Quadrat).

3.2 Erkenntnisse zu Bildaufnahme und Anforderungen

Die ersten am Windmessmast aufgenommenen Bilder waren farblich nicht ganz korrekt, da die Kameras noch nicht mit NIR-Sperrfiltern (nahes Infrarot) ausgestattet waren. Die benötigten NIR-Filter wurden im Juli 2019 für die Kameras am Windmessmast nachgerüstet.

Die beteiligten Ornithologen bewerteten hinsichtlich der Erkennung einer Vogelart in einer Bildfolge, dass eine Bildaufnahmefrequenz von vier Bildern pro Sekunde (4 fps) hierfür ausreichend ist.

Die lokale Rechenleistung der verwendeten Framegrabber war nicht ausreichend, um mehr als zwei Kameras pro Framegrabber zu betreiben. Dieses schlechte Verhältnis aus Kameras zu Framegrabbern ist für die geplante Rundumsicht des Zielsystems nicht zuträglich. Gleichzeitig ist auch kein Erweiterungspotential für das FPGA-Programm (Field Programmable Gate Array) auf dem Framegrabber vorhanden, was die Weiterentwicklung des Systems zusätzlich erschwert. Ein Wechsel auf die schnelle Datenübertragung über abgeschirmte Rundkabel (CoaXPress)Schnittstelle mit neuen hochauflösenden Kameras adressiert diese Anforderung und wird im nächsten Abschnitt beschrieben.

Tab. 1: Anzahl der Bilder und deren Datenmenge, die mit dem BirdRecorder-System IPC-1 in den Jahren 2019, 2020 und 2021 auf dem Windenergiefeld in Stötten aufgenommen wurden.

Jahr	Anzahl der Rohbilder	Datenmenge der Rohbilder
2019	3.979.890	30,36 TB
2020	12.024.807	91,74 TB
2021	2.787.329	21,27 TB

Das System BirdRecorder IPC-1 wurde über die nahezu gesamte Vorhabendauer auf dem Windtestfeld in Stötten auf dem Windmessmast Nord-West betrieben, um möglichst viele Bilddaten zu sammeln, zu bewerten und für die Belernung der KI zur Erkennung von Milanen einzusetzen. Die Anzahl der Bilder für die einzelnen Jahre und die damit einhergehende Datenmenge ist in Tab. 1 wiedergegeben.

Die Erkennung von Bewegungen (Blob-Detektion), die auf einem FPGA direkt auf dem Framegrabber aufgeführt wird hat sich besten bewährt. Erst wenn eine Bewegung im Bild einer Kamera in Echtzeit detektiert wird, werden die Vollbilder der Kamera auf dem NAS-Speichersystem abgespeichert. Damit stehen diese für die weitere Bearbeitung sowie für die Entwicklung und Optimierung der Methoden und Algorithmen zur Verfügung.

3.3 Weiterentwicklung des Hardwaresystems

Wie oben beschrieben, wurde das Zielsystem für den BirdRecorder (IPC-2) zum einen auf Basis der bisherigen Erfahrungen mit dem ersten Bildaufnahmesystem (IPC-1) und zum anderen als Standalone-System mit einem klimatisierten Schaltschrank entwickelt, aufgebaut und erprobt. Das IPC-2 System ist für den Anschluss von bis zu acht statisch montierten, hochauflösenden Kameras mit CoXPress Schnittstelle und bis zu zwei schwenkbaren Stereokamerapaaren konzipiert. In der ersten Ausbaustufe wurde ein Stereokamerapaar implementiert (Abb. 2). Das System enthält insgesamt drei Framegrabber für die Bildaufnahme. Je einen für bis zu vier statische Kameras und einen für den Anschluss und die Bildaufnahme von bis zu zwei Stereokamerapaaren. Die Schwenkneigekopfeinheiten werden über eine Ethernet-Schnittstelle an die Workstation angeschlossen und gesteuert.

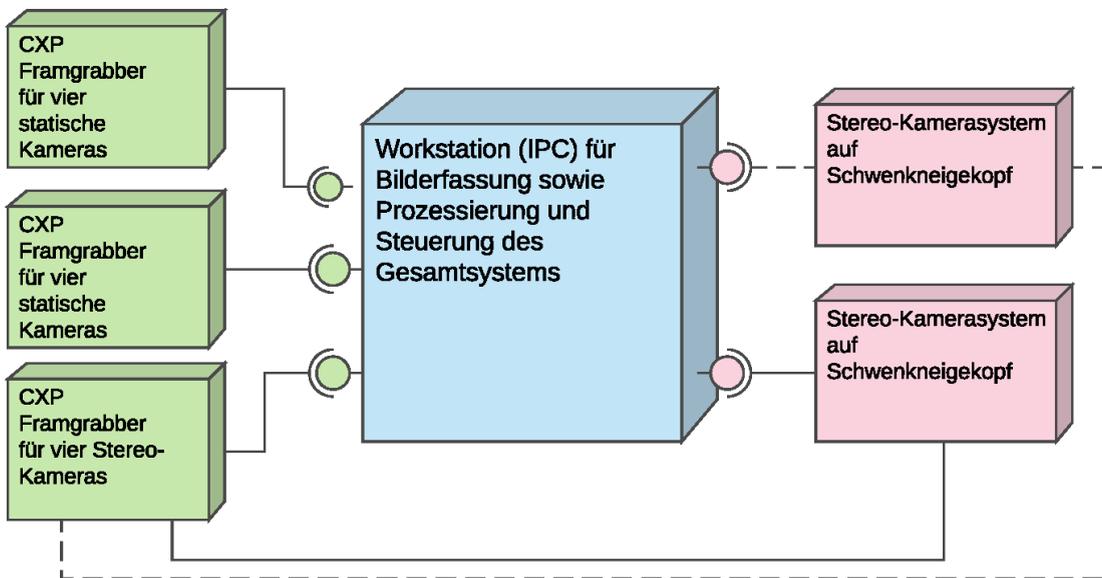


Abb. 2: BirdRecorder System IPC-2: Schematischer Aufbau mit IPC-Workstation, Frame-grabbern und Schwenkneigekopf mit nachführbaren Stereokameras.

Eine funktionale Übersicht des BirdRecorders IPC-2 Systems ist in Abb. 2 dargestellt. Die beiden Framegrabber der statischen Kameras nehmen die Bilder von jeweils vier statischen Kameras auf und senden die Objektpositionen und -größen sowie Bildausschnitte an die Workstation. Dort nehmen die neuronalen Netze eine Klassifizierung vor und es wird ein Tracking der einzelnen Objekte durchgeführt, d.h. eine Ermittlung und Verfolgung der Flugbahnen. Die Entscheidung, welches Objekt vom Stereokamerapaar auf dem Schwenkneigekopf weiterverfolgt (getrackt) werden soll, wird aus dem Erkennungsergebnis und der Erkennungswahrscheinlichkeit getroffen. Entsprechend wird dann die gewünschte Soll-Position des Objektes an den Schwenkneigekopf gesendet. Sobald das getrackte Flugobjekt in den Bildern der schwenkbar montierten Kameras mit Teleobjektiven erscheint, wird zum einen die Entfernung zum Flugobjekt ermittelt. Zum anderen wird mithilfe der Ist-Position des getrackten Objekts auf den Teleaufnahmen der Sollwert für den Schwenkneigekopf nachgeregelt, so dass das Flugobjekt möglichst zentral in der Bildmitte abgebildet wird. Aufgrund von Erkennungsergebnis, Erkennungswahrscheinlichkeit und Entfernung wird eine Entscheidung für weitere Maßnahme getroffen werden.

Die Framegrabber ermöglichen die Verwendung von hochauflösenden 26 Megapixel (MP) Kameras, von denen jeweils zwei auf bis zu zwei Schwenkneigeköpfen montiert werden und bis zu acht als statisch montierte Kameras eingesetzt werden können. Für die Weitwinkelaufnahmen werden Objektive mit 12 mm Brennweite verwendet, die einen Bildwinkel von 56° horizontal und vertikal erlauben. Bei den Teleobjektiven der Stereokameras mit einem Bildwinkel von 9,8° wird ein 1 Meter großes Objekt in 700 m Entfernung noch 42 Pixel groß abgebildet. Dies ist für die sichere Erkennung der Vogelart ausreichend, ebenso für die Entfernungsbestimmung mittels Stereometrie.



Abb. 3: Testinstallation des BirdRecorder Systems IPC-2 auf dem Dach des ZSW in Stuttgart mit klimatisiertem Schaltschrank, statisch montierten Kameras sowie Schwenkneigekopf mit Stereokamera paar.

Für einen 360 Grad Rundumblick mit den acht statischen Kameras werden diese so angeordnet und installiert, dass sich der Bildbereich der einzelnen Kameras leicht überlappt. Mit einem quadratischen Gesichtsfeld (FOV) der Kameras von 56 Grad und insgesamt acht Kameras ergibt sich ein Überlappungsbereich in der Horizontalen von 9,3 Grad. Die Anordnung der acht statischen Kameras und die Überlappung der FOV der einzelnen Kameras ist in Abb. 3 dargestellt. Die Überlappung der FOV der statischen Kameras ist notwendig, damit Objekte, die sich vom FOV einer Kamera in das FOV der benachbarten Kamera bewegen als das selbe Objekt identifiziert werden und über die verschiedenen statischen Kameras hinweg sicher getrackt werden können.



Abb. 4: Anordnung der Gesichtsfelder (FOV) der acht statischen Kameras des BirdRecorder Systems für einen 360 Grad Rundumblick.

Für die IPC-Workstation wurde ein Schaltschrank entwickelt und aufgebaut. Im Vergleich zur IPC-1 Version ist die neue IPC-2 Workstation wesentlich größer und leistungsfähiger. Der Schaltschrank für das BirdRecorder System IPC-2 ist in Abb. 3 rechts zu sehen,

3.4 Schwenkneigekopf für Stereokameras

Für die Auswahl eines für alle Belange passenden Schwenkneigekopfes wurde eine umfangreiche Marktrecherche unternommen. Insgesamt wurden mehr als 20 verschiedene Typen und Hersteller angefragt, Datenblätter ausgewertet und ein Bewertungsschema für die Auswahl erstellt. Auf Basis der Bewertungen kamen insgesamt sechs verschiedene Modelle in die engere Auswahl. Die entsprechenden Firmen wurden angefragt, ob ein Testgerät für eine erste Erprobung leihweise zur Verfügung gestellt werden kann. Dies war leider nur bei sehr wenigen Firmen der Fall.

Das ausgewählte und leihweise zur Verfügung gestellte Modell wurde sowohl hardwareseitig als auch die Softwareansteuerung intensiv getestet. Der Schwenkneigebereich des beschafften Schwenkneigekopfes erfüllt mit einem Azimut Schwenkbereich von +/- 188° die Anforderung für eine Rundumsicht und ermöglicht mit einem Neigebereich von 0° (horizontal) bis +120° das Tracking im gesamten Halbraum um seinen Standort herum. Mit einer Winkelauflösung von 0,003° für den Azimut- und Elevationswinkel ist das Tracking von Objekten sehr genau möglich. Der Schwenkneigekopf ist für den Außenbereich geeignet (Schutzart IP 67), ist für einen Temperaturbereich von - 30° C bis + 70° C spezifiziert und die Nutzlast beträgt bis zu 7,5 kg.

4 Software

Die im Projekt entwickelte Software untergliedert sich in einen Teil, der auf dem frei programmierbaren FPGA (Field Programmable Gate Array) der Framegrabber in Echtzeit ausgeführt wird. Der zweite Teil wird auf dem Industrie PC ebenfalls in Echtzeit ausgeführt und ist in der Programmiersprache C++ implementiert.

Die auf den FPGAs der Framegrabber laufende Software detektiert Bewegungen in den von den Kameras aufgezeichneten Bildern und liefert die Bilder und Bildausschnitte der sich bewegenden Objekte an den Industrie-PC. Damit wird bewerkstelligt, dass nur Bilder aufgezeichnet werden, wenn sich bewegende Objekte im Gesichtsfeld der statischen Kameras befinden. Somit wird die zu erfassende Bilddatenmenge frühzeitig auf die wesentlichen Bildinhalte reduziert. Die Blob-Analyse für die Bewegungsdetektion auf dem FPGA ist dabei so parametrisiert, dass die Falsch-Negativ-Rate nahe null ist, d.h. es werden in diesem ersten Schritt der Prozesskette möglichst alle Bewegungen erfasst und die entsprechenden Bildfolgen gespeichert, um diese später weiter zu verarbeiten.

Die C++-Software auf dem Industrie-PC übernimmt dann die nachgelagerte Steuerung des Schwenkneigekopfes und schreibt die Ergebnisse der Objekterkennung in eine Datenbank. Im Detail ist die Softwarearchitektur in Abschnitt 4.5 Softwarearchitektur des Gesamtsystems noch weiter erläutert.

4.1 Datenerfassung und -aufbereitung

Bei der initialen Datenerfassung durch das BirdRecorder System IPC-1 fand auf dem FPGA nur die Bewegungsdetektion statt. Die Detektion der relevanten Bildausschnitte (Blob-Analyse) erfolgte zeitversetzt am ZSW auf geeigneten Servern. Dafür wurden die mit einer Rate von 10 Bildern pro Sekunde (fps) aufgenommenen Bilder zuerst von dem Kamera-Rohbildformat in ein JPEG-Farbformat konvertiert. Aus diesen Bildern konnten dann relevante Bildausschnitte mittels Blob-Analyse bestimmt und extrahiert werden. Diese Bildausschnitte, auch ROIs genannt (Regions of Interest), werden für das Training und die Anwendung der neuronalen Netze für die Erkennung der Vogelart benötigt.

In Abb. 5 sind verschiedene Objektarten gezeigt, eingeteilt in die Klassen Vogel, Fluggerät und Sonstiges. Im Nachfolgenden werden die ROIs auch Bilder genannt. Auf Basis dieser Bilder wird die Vogelart bestimmt. Einen Überblick über die dafür notwendigen Schritte gibt die in Abb. 6 dargestellte Prozesskette zur Vogel- und Vogelarterkennung. Dabei kommen verschiedene neuronale Netze zum Einsatz. Auf Basis der vorhandenen Bilddaten kann eine automatische Sortierung (gelb) durchgeführt werden. Anschließend erfolgt das Annotieren der vorsortierten Bilder durch Nicht-Ornithologen (hellblau) und das Training des ersten Neuronales Netz NN-1 zur Vogelerkennung (grün), das zwischen Vögeln, Fluggeräten und sonstigen Objekten unterscheidet. Nach dem Annotieren von Vogelbildern durch Ornithologen (hellblau) wird ein zweites neuronales Netz NN-2 (dunkelblau) trainiert, das die Vogelarterkennung durchführt.

Das neuronale Netz zur automatischen Sortierung wird auf Grundlage von nicht annotierten Bildern trainiert (unüberwachtes Lernen). Die automatische Sortierung, wie nachfolgend beschrieben, vereinfachte das anschließende annotieren der Bilder durch Nicht-Ornithologen wesentlich. Das Training der neuronalen Netze zur Objekterkennung (NN-1) und zur Vogelarterkennung (NN-2) erfolgt mithilfe der annotierten Bilddaten (überwachtes Lernen).

Für das Training des NN-1 wurden aus dem ca. 1.000.000 Kamerabilder umfassenden Bildmaterial des Zeitraums 31.07.2019 bis 03.09.2019 350.000 ROIs ausgewählt und ausgeschnitten. Um den Aufwand beim händischen Annotieren der 350.000 ROIs zu

reduzieren, wurde auf eine Methode aus dem Bereich des unüberwachten Lernens zurückgegriffen, einem sogenannten tiefen convolutional Autoencoder. Dieser versucht, die gegebenen 256 x 256 Pixel großen ROIs mit Farbinformation auf eine Repräsentation deutlich geringerer Dimensionalität abzubilden und von dieser Codierung wieder das Originalbild zu reproduzieren. Dabei extrahiert der Autoencoder die den Datensatz charakterisierenden Merkmale, wodurch eine Sortierung der Bilddaten entsprechend dieser Merkmale erreicht wird.

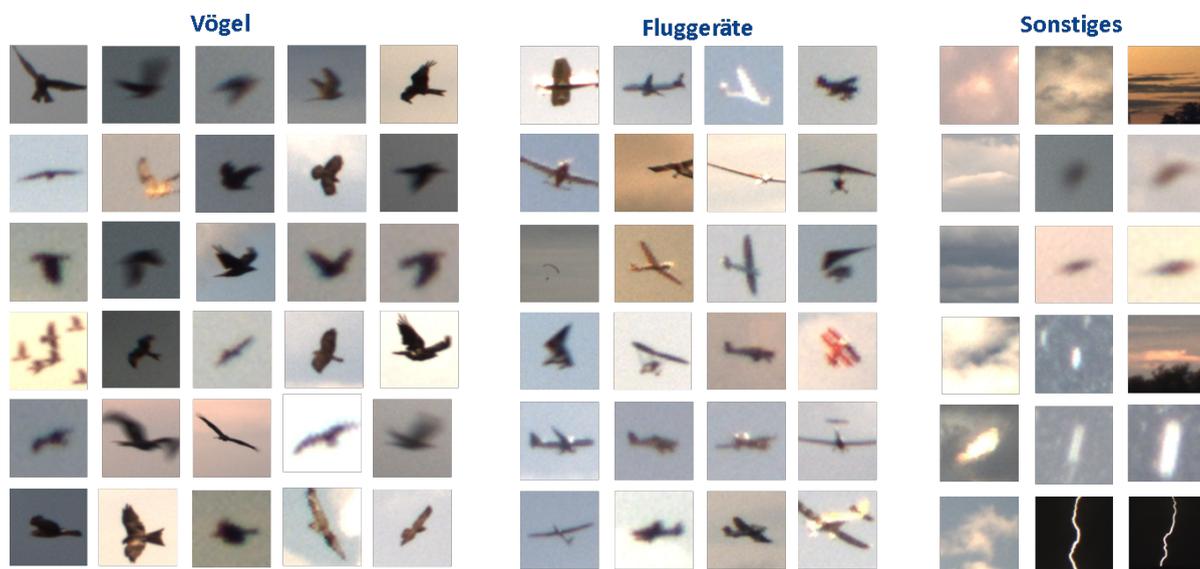


Abb. 5: Beispielbilder für die Objektarten Vögel, Fluggeräte und Sonstiges.

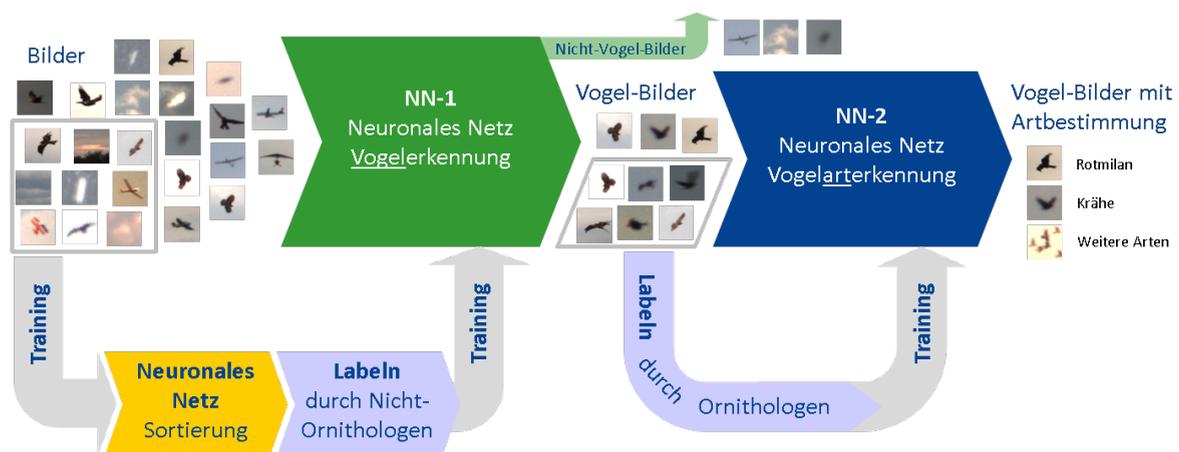


Abb. 6: Prozesskette des BirdRecorders zur Vogel- und Vogelarterkennung (Beschreibung siehe Text).

Der Ergebnisvektor der Merkmale kann zur Visualisierung auf zwei Dimensionen projiziert werden. Hierzu wird das sog. t-SNE (t-distributed stochastic neighbor embedding) (Maaten and Hinton, 2008; Van Der Maaten, 2014) verwendet. Die Projektion mit der t-SNE Methode für die vorliegenden Bilddaten sind in Abb. 7 zu sehen. In der Mitte ist jeweils die 2D-Projektion des Ergebnisses des Autoencoders zu sehen. Bei einer manuellen Betrachtung sind zwei Cluster erkennbar, die eindeutig keine Vögel enthalten (orange umrandet). Die

Bilder in diesen beiden Clustern werden für das Training der Arterkennung nicht weiter verwendet. Die Bilder des grün umrandeten Clusters in Abb. 8 werden manuell in die Klassen Vogel, Fluggerät und Sonstiges annotiert und für das Training der neuronalen Netze verwendet.

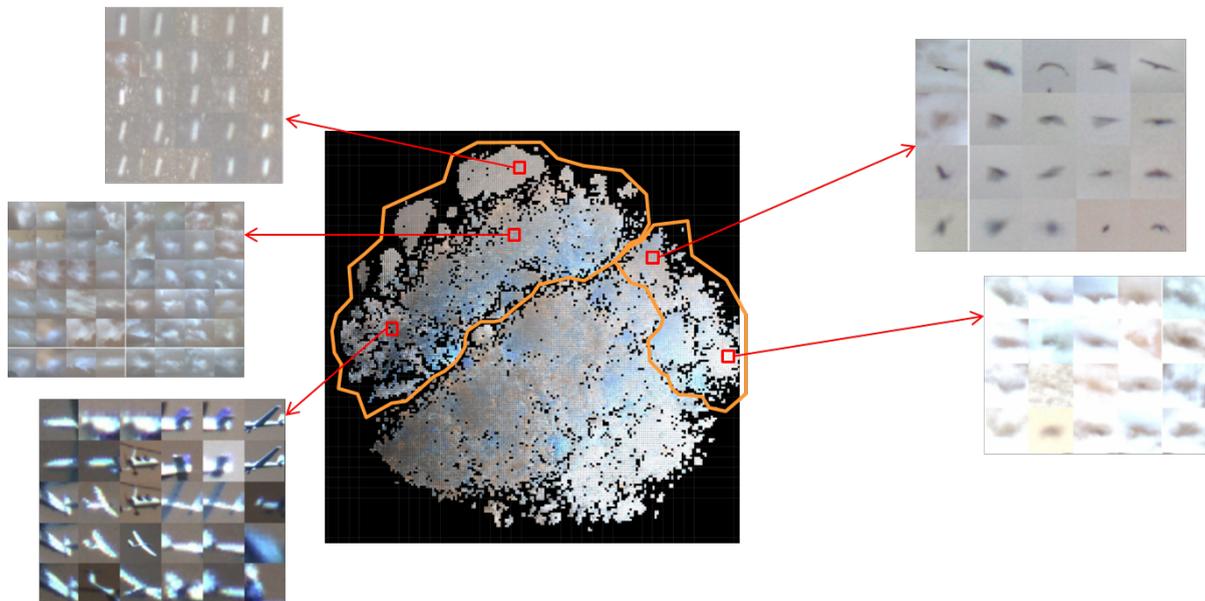


Abb. 7: Automatische Sortierung der Bilder für das Training der neuronalen Netze für die Objekt- und Arterkennung (Beschreibung siehe Text).

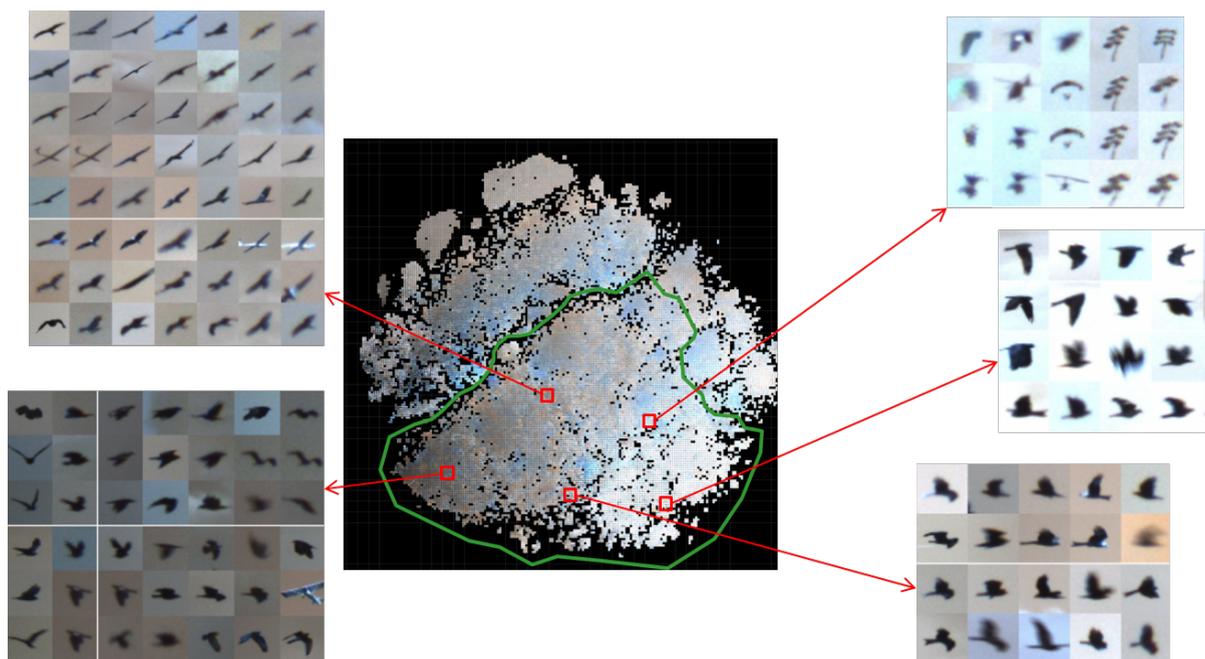


Abb. 8: Wie Abb. 7, jedoch für Bilder die überwiegend Vogelobjekte enthalten und anschließend noch manuell annotiert werden.

Durch eine geeignete Maskierung und Clusterung des Projektionsraums konnten 53.289 Bilder des zu annotierenden Datensatzes bereits automatisch der Klasse Nicht-Vögel zugeordnet werden. Die übrigen Bilder aus dem Cluster Vögel und Fluggeräte wurden Nicht-

Ornithologen zum Annotieren vorlegt. Für eine Qualitätsprüfung der automatischen Sortierung wurden die 53.289 als Nicht-Vögel gekennzeichnete Bilder manuell überprüft. Dies ergab, dass nur 340 falsch automatisch annotiert wurden, was 0,64 % entspricht. Durch die Vorsortierung und Clusterung konnte die Überprüfung mit überschaubarem Zeitaufwand durchgeführt werden.

4.2 Webbasierte Annotationsplattform

Für das Training von neuronalen Netzen zur Objekterkennung sind entsprechende Trainingsdaten in Form von annotierten Bildern notwendig. Zur Bestimmung der Vogelarten auf den Bildern wurde die Expertise von Ornithologen benötigt. Die Vogelbilder liegen in teilweise hunderte Bildern langen Sequenzen desselben Vogels vor. Eine Sequenz entspricht einer Bildfolge vom Beginn bis zum Ende einer detektierten Flugbewegung und beinhaltet ein oder mehrere verschiedene Objekte (z.B. Vogel, Fluggerät). Verbreitete Hilfsmittel zum annotieren von Bildern berücksichtigen diesen Anwendungsfall nicht. Durch die Bereitstellung eines entwickelten Webinterfaces konnten die Ornithologen somit alle Bilder einer Vogelflugtrajektorie mit derselben Annotation versehen. Dies steigert die Effizienz des Annotationsvorgangs durch Reduzierung des Zeitaufwands unter gleichzeitiger Erhöhung der Anzahl an annotierten Bilder. Ist ein Vogel auf einem Einzelbild nicht oder nur unsicher erkennbar, so ist er im Kontext der Sequenz dennoch sicher erkennbar und kann mit dem korrekten Label versehen werden.

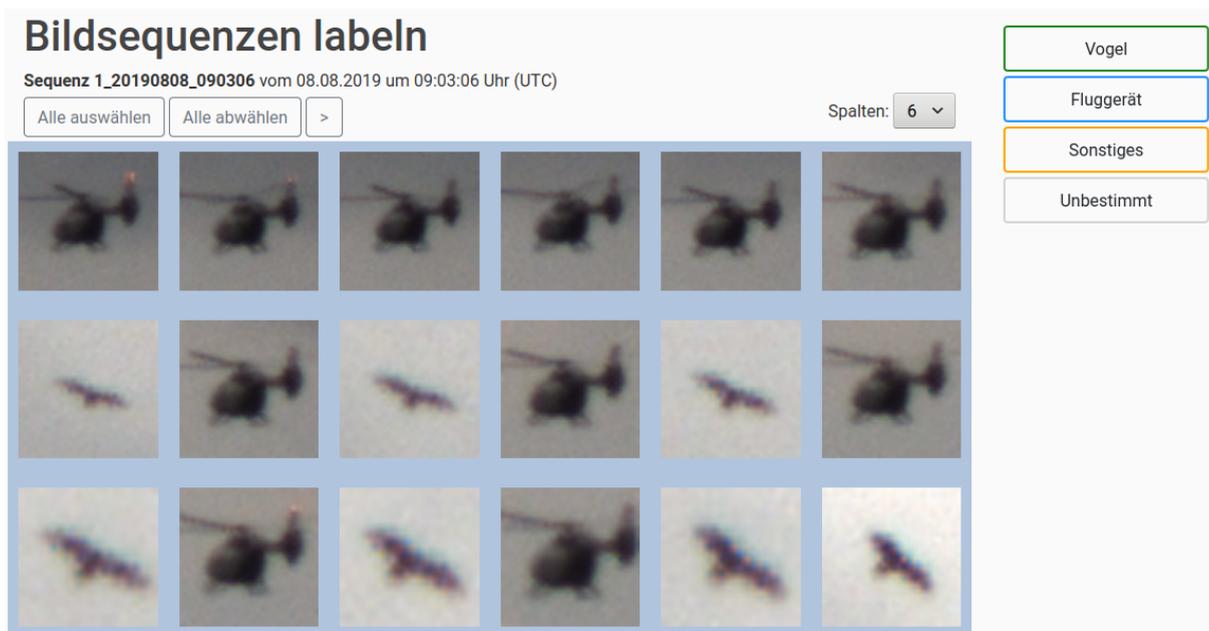


Abb. 9: Webinterface der Annotationsplattform mit der Ansicht für Nicht-Ornithologen.

Das Webinterface zum Annotieren der Bilder besteht aus einer Nicht-Ornithologen-Ansicht und einer Ornithologen-Ansicht und sind in Abb. 9 bzw. Abb. 10 gezeigt. Die Handhabung des Interfaces wurde in enger Abstimmung mit den beteiligten Ornithologen optimiert. Des Weiteren wurden in Kooperation mit den Ornithologen, die bei der Ornithologen-Ansicht zur Verfügung stehenden Vogelarten und Kategorien eingehend diskutiert und festgelegt.

Von den Nicht-Ornithologen wurden mit Hilfe dieses Webinterfaces ca. 680.000 Bilder annotiert. Die rund 412.000 als Vogel annotierten Bilder wurden anschließend durch die beteiligten Ornithologen nach Vogelart annotiert. Die Verteilung der annotierten Objektarten und Vogelarten zeigt Tab. 2. Die gelb markierten Vogelarten bzw. Artgattungen in Tab. 2

zählen in Baden-Württemberg zu den windkraftempfindlichen Vogelarten (Landesanstalt für Umwelt Baden-Württemberg, 2022).

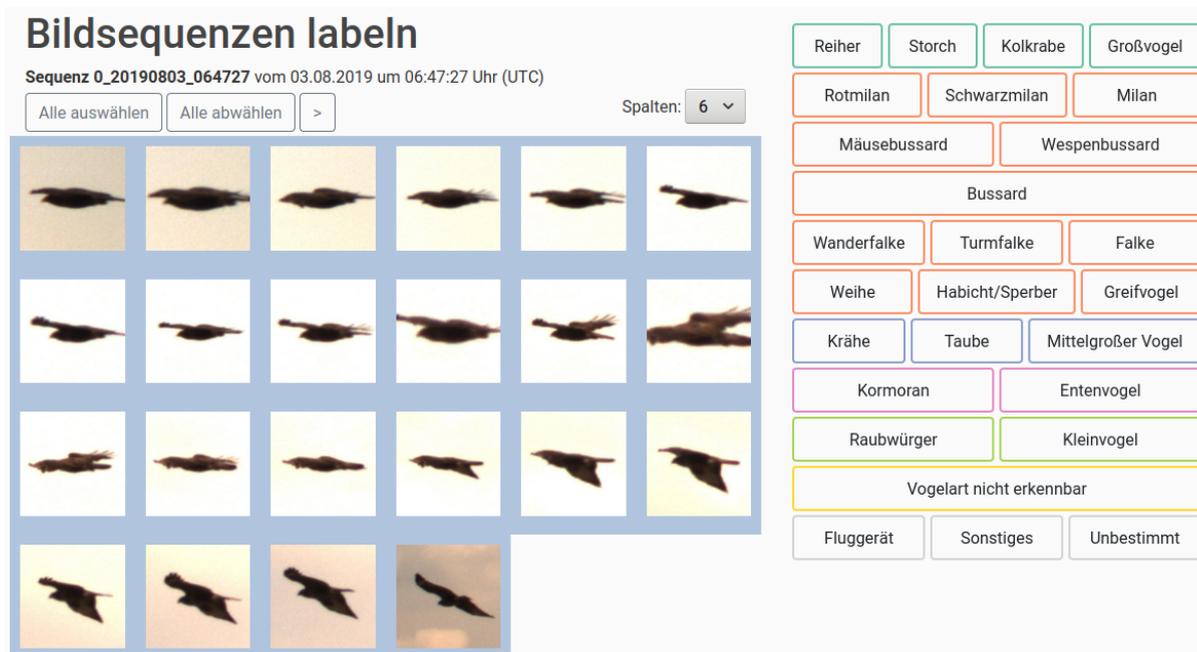


Abb. 10: Webinterface der Annotationsplattform mit der Ansicht für Ornithologen zur Annotation der Vogelart bzw. Artklasse.

Label	Bilder	Anteil [%]
Vogel	412146	61.0
Fluggerät	132949	19.7
Sonstiges	119965	17.7
Unbestimmt	11043	1.6
Summe	676103	100

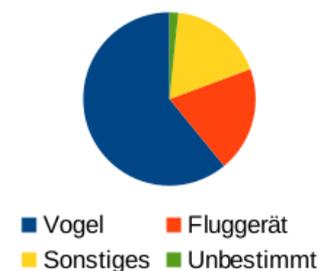


Abb. 11: Verteilung manuell annotierter Objektarten für das Training des neuronalen Netzes NN-1.

Tab. 2: Verteilung der Vogelarten für die von Ornithologen annotierten Bilder (Beschreibung siehe Text).

Label	Bilder	Bildsequenzen	Anteil [%]
Taube	89.013	524	21,597
Rotmilan	71.833	678	17,429
Krähe	70.690	497	17,152
Vogelart nicht erkennbar	65.119	1.528	15,800
Greifvogel	50.321	1.407	12,210
Mäusebussard	14.978	220	3,634
Milan	14.393	455	3,492
Kleinvogel	14.203	165	3,446
Falke	10.298	163	2,499
Turmfalke	5.411	162	1,313
Bussard	1.908	80	0,463
Mittelgroßer Vogel	1.460	29	0,354
Kolkrabe	934	7	0,227
Schwarzmilan	617	6	0,150
Weihe	429	6	0,104
Reiher	348	6	0,084
Großvogel	125	4	0,030
Wanderfalke	41	1	0,010
Habicht/Sperber	5	1	0,001
Entenvogel	20	1	0,005
Gesamt	412.146	5.940	100

4.3 Erkennung der Vogelart

Die Vogelarterkennung durch neuronale Netze ist in zwei aufeinander folgende Schritte aufgeteilt (Abb. 6). Im ersten Schritt erfolgt durch das NN-1 (Neuronales Netz 1) die Objekterkennung. Dabei werden Vögel von Nicht-Vögeln getrennt. Im zweiten Schritt erfolgt die Vogelarterkennung mittels NN-2 (Neuronales Netz 2) bei der zwischen den Vogelarten bzw. Artgattungen unterschieden wird. Im Rahmen dieses Vorhabens konzentriert sich die Erkennung der Vogelart auf die die Gattung Milan, die Rot- und Schwarzmilan umfasst.

Nachfolgend werden das Training und die Ergebnisse der beiden Erkennungsalgorithmen mit neuronalen Netzen erläutert. Das Training erfolgte auf Basis der annotierten Bilder, deren Verteilungen in Abb. 11 und Tab. 2 dargestellt sind.

Für beide neuronalen Netze wurden verschiedene Netzwerkarchitekturen und Belernungs-

verfahren getestet. Primär wurden Algorithmen herangezogen, die auf dem ImageNet Benchmark Datensatz zur Zeit die Top 10 belegen (Chen et al., 2021; Li et al., 2022, p. 2; Moon et al., 2021).

Für den hier vorliegenden Anwendungsfall – NN-1 für Erkennung von Vogelobjekten – hat ein sogenanntes Residual Network (Rajagopal et al., 2020; Shrestha and Mahmood, 2019) mit drei Convolutional Layern (Convolutional Neural Networks: CNN) pro Residualblock die beste Erkennungsrate generiert (Abdel-Hamid et al., 2013; Tustison et al., 2019).

Für die Unterscheidung der detektierten Objekte in die Klassen Vogel, Fluggerät und Sonstiges wurde auf Basis der ROIs das NN-1 mit den entsprechenden ca. 680.000 Farbbildern trainiert, in denen die Objekte skaliert dargestellt sind. Die Skalierung erfolgte so, dass das zu klassifizierende Objekt den Bereich 64 x 64 Pixel möglichst voll ausfüllt.

Die Erkennungsraten von NN-1 für die o.g. Klassen auf Basis unabhängiger Testdaten weist eine Genauigkeit von 98,2 % auf.

Die falsch klassifizierten ROIs teilen sich hierbei folgendermaßen auf:

- 1,1 % falsch **positiv**: Nicht-Vogel-Objekt als Vogel erkannt.
- 0,7 % falsch **negativ**: Vogel-Objekt nicht als Vogel erkannt.

Die Falsch-Positiv-Rate ist zu einem Großteil dadurch erhöht, dass aufgrund der Nähe zum nördlich des Windtestfelds gelegenem Flugplatz Messelberg viele Aufnahmen von Gleitschirmfliegern im Datensatz vorhanden sind. Ist die Person – Pilot des Gleitschirms – die oftmals nur als kleiner Punkt unterhalb des Gleitschirms zu sehen ist, nicht erkennbar oder im ROI abgeschnitten, ist je nach Blickwinkel der Gleitschirm der Silhouette eines Vogels ähnlich und deshalb für das neuronale Netz schwer von einem Vogel zu unterscheiden.

Für einen Zeitraum in der Morgendämmerung bis in den späteren Vormittag hinein ist ein Video der Einzelbildklassifikation unter <https://youtu.be/zZ7NZteZOng> verfügbar.

Nach der vorangegangenen Filterung durch NN-1 handelt es sich bei den verbliebenen Bildern nur noch um Bilder der Objektklasse Vögel. Die Aufgabe für NN-2 besteht darin, die Milane (Rot- und Schwarzmilan) möglichst sicher zu erkennen. Die sukzessive Prozessierung der Bilder mit NN-1 und NN-2 hat den Vorteil, dass NN-2 sich rein auf die Unterscheidung von Vogelarten bzw. in der wichtigsten Form zwischen Milanen und anderen Vögeln konzentrieren kann.

Für NN-2 wurden Experimente mit Farb- und Grauwertbildern sowie verschiedenen neuronalen Netzarchitekturen unternommen. Die besten Ergebnisse lieferten Trainings auf Basis von Grauwertbildern, auf denen die Vögel ohne Skalierung abgebildet sind. Die Architekturen für NN-2 sind im Vergleich zum NN-1 deutlich kleiner (weniger Feature Maps und Schichten). Dies ist u.a. auf die geringere Menge an Trainingsbeispielen zurückzuführen.

Wie aus Tab. 2 zu erkennen ist, umfasst der Datensatz 71.833 Rotmilane, 617 Schwarzmilane und 14.393 nicht näher spezifizierte Milane. Es liegen also 86.843 ROIs für die Klasse Milan vor. Da diese aus Sequenzen stammen, die mit bis zu 10 Bildern pro Sekunde aufgenommen wurden, sind sie mit einem hohen Grad an Redundanz belegt.

Um die Varianz in den Bildern zu erhöhen und gleichzeitig den Datensatz zu vergrößern, wurden die Bilder leicht verschoben, rotiert, skaliert, geschert und in der Helligkeit oder dem Farbton verändert. Die transformierten Bilder zeigen immer noch das selbe Objekt bzw. hier denselben Vogel. Diese Technik wird als Datenaugmentierung bezeichnet und kann im Bereich des Maschinellen Lernens helfen, robustere Modelle zu erhalten (Cao et al., 2020; Gayer et al., 2019; Tustison et al., 2019). Für ein gutes Erkennungsergebnis wurde auf die

Augmentierungen und Transformationen der Bilder in einem hohen Maße zurückgegriffen.

Mit der gewählten Architektur, Anpassungen in der Struktur und Regularisierung, sowie der Augmentierung, wurden verschiedene Experimente für das Training von neuronalen Netzen durchgeführt. Das im Rahmen der Experimente und Auswertungen gewählte neuronale Netz, kann auf einem Testdatensatz Milane (Rot- und Schwarzmilan) mit einer Genauigkeit von 93,1 % von anderen Vögeln unterscheiden. Bei den 6,9 % nicht korrekt zugeordneten ROIs handelt es sich größtenteils um Bilder von Vögeln, deren Art auch für Ornithologen nicht auf dem Einzelbild erkennbar ist, sondern nur mit Hilfe der gesamten Bildfolge innerhalb einer erfassten Flugbahn. Abb. 12 zeigt 128 zufällig ausgewählte Bilder aus einem unabhängigen Testdatensatz, die so ausgewählt wurden, dass die Fehlerverteilung des Testdatensatzes erhalten bleibt. Die grün umrandeten Bilder wurden hierbei durch das Netz richtig klassifiziert, gelb, orange und rot umrandete Bilder wurden falsch klassifiziert. Die Farbe kennzeichnet hierbei die Wahrscheinlichkeit, wie sicher sich das Netz bei der Fehlklassifizierung ist. Die Farbskala reicht dabei von gelb für bis 75 % über orange für 75 % bis 85 % und rot für mehr als 95 % Wahrscheinlichkeit für die Fehlklassifikation.

Insgesamt umfasst der Testdatensatz 13.440 ROIs. Hiervon wurden 925 falsch klassifiziert und davon 106 mit einer Wahrscheinlichkeit $> 95\%$. Diese mit hoher Wahrscheinlichkeit falsch klassifizierten ROIs werden im Folgenden als kritisch bezeichnet, da dies sehr negative Auswirkungen auf die Erkennungswahrscheinlichkeiten haben kann, sobald das Tracking berücksichtigt wird. Um eine robuste Annotierung zu erhalten, werden bei einem Tracking die Wahrscheinlichkeiten für die Einzelbilder nach dem Satz von Bayes (Peyrolón, 2020) über die Sequenz statistisch korrekt aufsummiert und somit die sog. Verbundwahrscheinlichkeit bestimmt. Dies setzt allerdings voraus, dass bei der Einzelbildklassifizierung möglichst keine falschen Klassifizierungen mit hoher Wahrscheinlichkeit auftreten.

Daher wurden alle kritischen Falschklassifizierungen (ca. 600) des gesamten Datensatzes ermittelt und den Ornithologen zur erneuten Prüfung vorgelegt. Bei der Prüfung wurden einige falsch gelabelte ROIs gefunden. Werden diese ROIs aus dem Testdatensatz entfernt, reduziert sich die Zahl der kritischen Falschklassifizierungen im Testdatensatz von 106 auf 14 (Abb. 13), die als kritisch einzustufen sind.

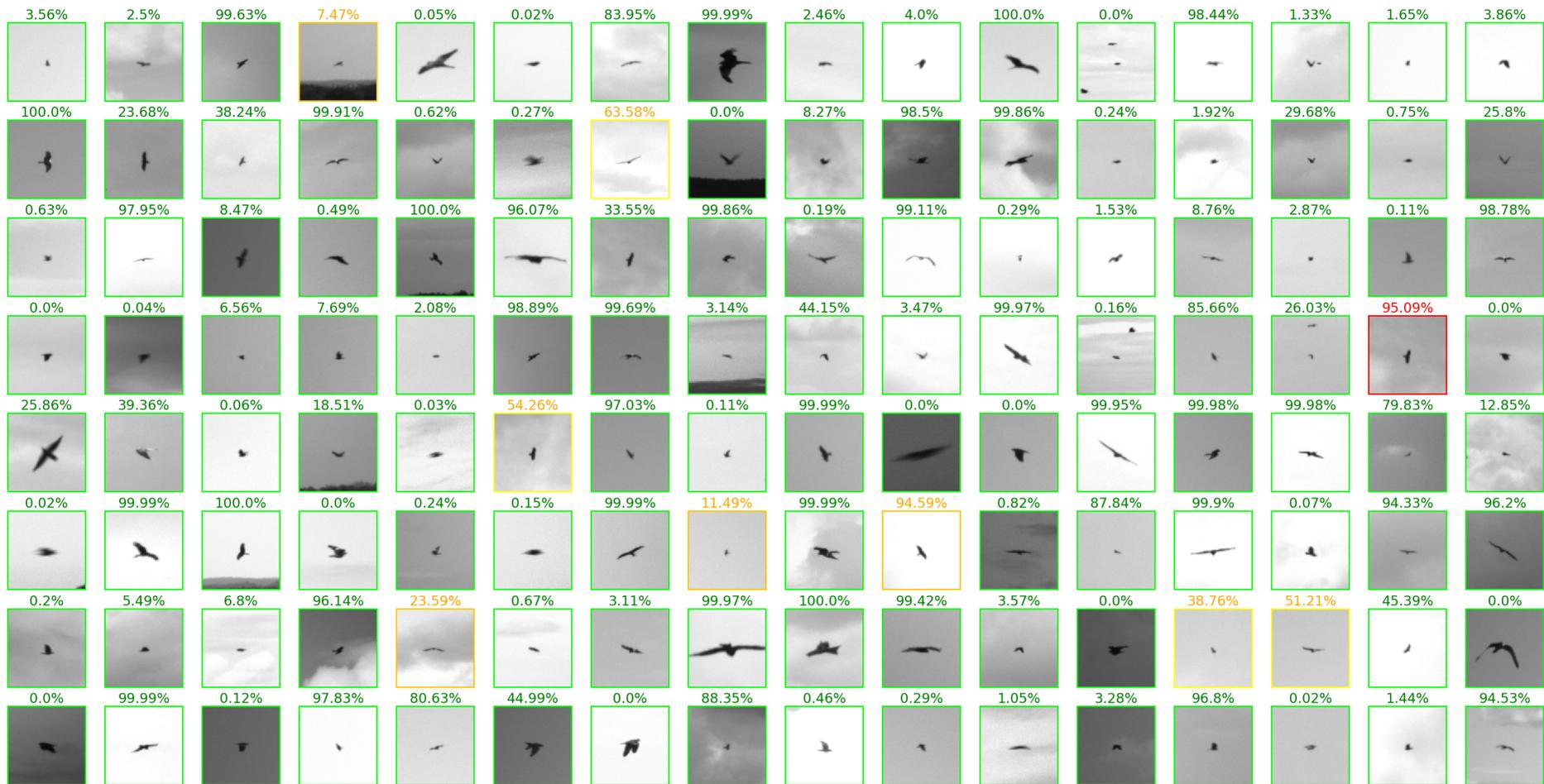


Abb. 12: 128 zufällig ausgewählte Bilder aus dem unabhängigen Testdatensatz, welche die Fehlerverteilung des gesamten Testdatensatzes repräsentiert.

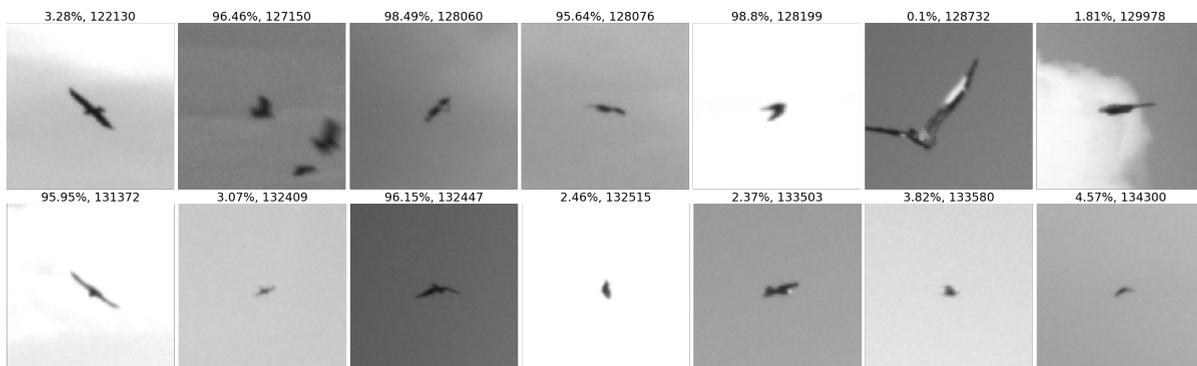


Abb. 13: Die 14 verbleibende, falsch kategorisierte Bilder nach Prüfung der auffälligen Labels.

Allgemein ist zu beobachten, dass das neuronale Netz NN-2 Probleme mit ROIs hat, auf denen viele Vögel gleichzeitig zu sehen sind. Es fällt dem Netz hierbei schwer, die Klasse des mittleren Vogels korrekt zu bestimmen. Auch ROIs in denen der Vogel größer als die gewählte Grundauflösung ist und somit Teile abgeschnitten sind, werden häufiger falsch klassifiziert. Sehr interessant ist der Fall in Abb. 14. Hier ist ein Milan zu sehen dessen spitze Schwanzfedern entweder fehlen oder nur sehr gering ausgeprägt sind (keine Schwalbenschwanzform erkennbar). Dieser Vogel wird dennoch vom Netz innerhalb der gesamten Sequenz mit hoher Wahrscheinlichkeit als Milan erkannt. Dann verändert der Milan seine Fluglage und der Schwanz erscheint (durch die Pixelinterpolation) nach außen gewölbt, für das Netz ein sehr wichtiges Merkmal, um Bussarde zu erkennen, wie auch in Abb. 15 und Abb. 16 zu erkennen. Die Visualisierung der farblich gekennzeichneten, entscheidungsrelevanten Bereiche des Bildes wurde mit der GradCam Methode erreicht (Selvaraju et al., 2017).

In Abb. 16 ist diese Bildfolge im Detail zu sehen. Die linke Zahl über jedem ROI gibt die ROI-ID der aufeinander folgenden Bildausschnitte aus der betreffenden Sequenz wieder (Aufnahme mit 10 Bildern pro Sekunde). Die rechte Zahl gibt die Wahrscheinlichkeit des Netzes an, dass es sich um einen Milan handelt (1,0 hohe, 0,5 mittlere, 0,0 sehr geringe Wahrscheinlichkeit). Der Milan wird vom Netz in der Sequenz mit hoher Wahrscheinlichkeit als Milan erkannt. Dann aber dreht er sich und der Schwanz scheint (durch die Pixelinterpolation) nach außen gewölbt. Die Schwanzform ist für das Netz ein sehr wichtiges Merkmal, um Bussarde von Milanen zu unterscheiden.

Eine prototypische Implementierung zeigt, dass die Erkennungswahrscheinlichkeit für Milane unter Einbeziehung des Trackings stark verbessert werden kann. Dazu zeigt Abb. 17 die Erkennungswahrscheinlichkeiten einer Beispielsequenz, die im Juni 2020 aufgenommen wurde. Das zugehörige Video ist unter <https://youtu.be/jOx8DOQKnr4> verfügbar (links ohne Tracking, rechts mit Tracking). Diese Bildsequenz ist ein Beispiel für schlechte Sichtverhältnisse, da es an dem Tag sehr neblig war. Dennoch würde der Milan als Objekt mit einem Durchmesser von ca. 8 Pixeln erkannt. Ab einem Durchmesser von 14 Pixeln wurde der Milan korrekt als solcher erkannt. Dieses Label verlor der Milan erst bei einem Durchmesser von 6 – 7 Pixeln. Unter der Annahme, dass der Milan eine Flügelspannweite von ca. einem Meter hat, würde dies bedeuten, dass er als Objekt bei einer Distanz 580 m vom System wahrgenommen wurde, er bei ca. 330 m als Milan erkannt wurde und erst bei ca. 700 m nicht mehr gesehen wurde. Anzumerken ist hier, dass die Aufnahmen mit den statisch montierten Kameras mit Weitwinkelobjektiv erfolgten. Für die Stereokameras mit Teleobjektiv ist die Erkennungswahrscheinlichkeit für größere Entfernungen wesentlich höher.



Abb. 14: Visualisierung der für das neuronale Netz entscheidungsrelevanten Bildbereiche mit der GradCam Methode (Beschreibung siehe Text).



Abb. 15: Beispiel desselben Milans wie Abb. 14 nur drei Bilder bzw. 300 ms früher.



Abb. 16: Beispiel Bildsequenz für besonders interessante kritische Fehl kategorisierung nach Änderung der Fluglage des Milans (Beschreibung siehe Text).

Abb. 18 zeigt eine weitere Beispielsequenz, die Ende März 2020 aufgenommen wurde. Sie ist ebenfalls als [Video verfügbar](#) und zeigt exemplarisch eine Situation mit guten Sichtverhältnissen. Die Entfernungsschätzungen sind sehr ähnlich zur vorigen Beispielsequenz. Allerdings ist in Abb. 18 zu erkennen, dass das neuronale Netz NN-2 durch die besseren Sichtverhältnisse eine höhere Milan-Erkennungswahrscheinlichkeit hat und auch über den gesamten Verlauf der Bildsequenz beibehält.

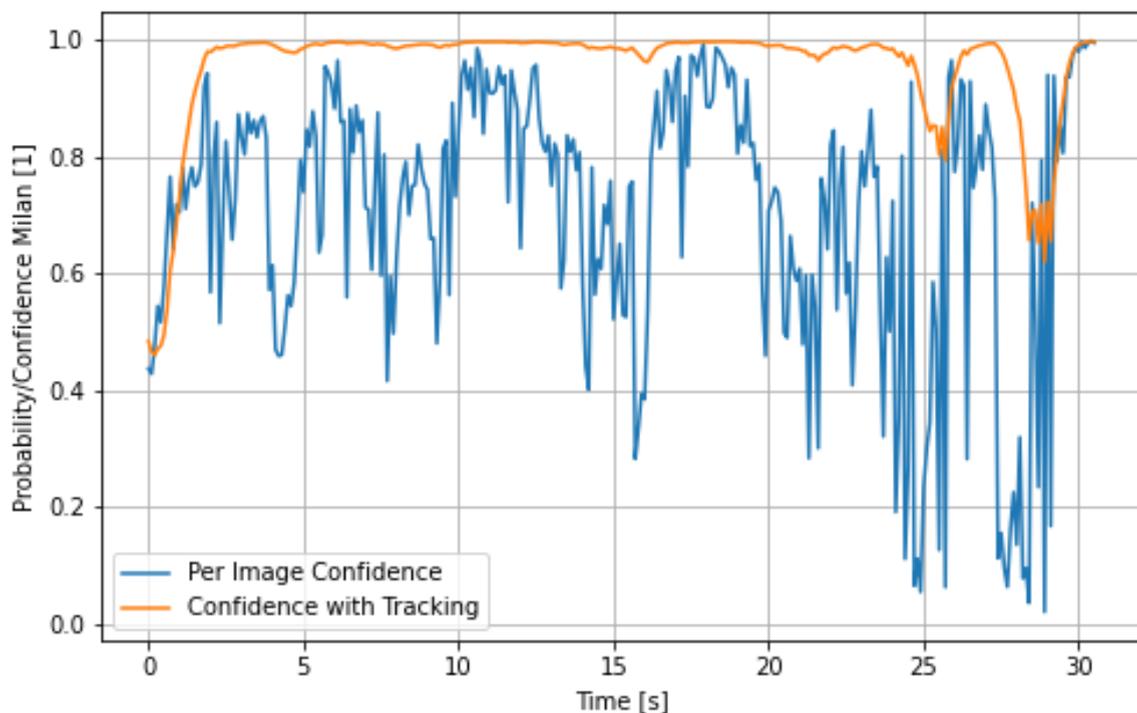


Abb. 17: Milan-Erkennungswahrscheinlichkeit einer Beispielsequenz ohne Tracking (blau) und mit Tracking (orange), für einen Tag mit schlechten Sichtverhältnissen und einem Milan im Abstand von 200 bis 700 Metern.

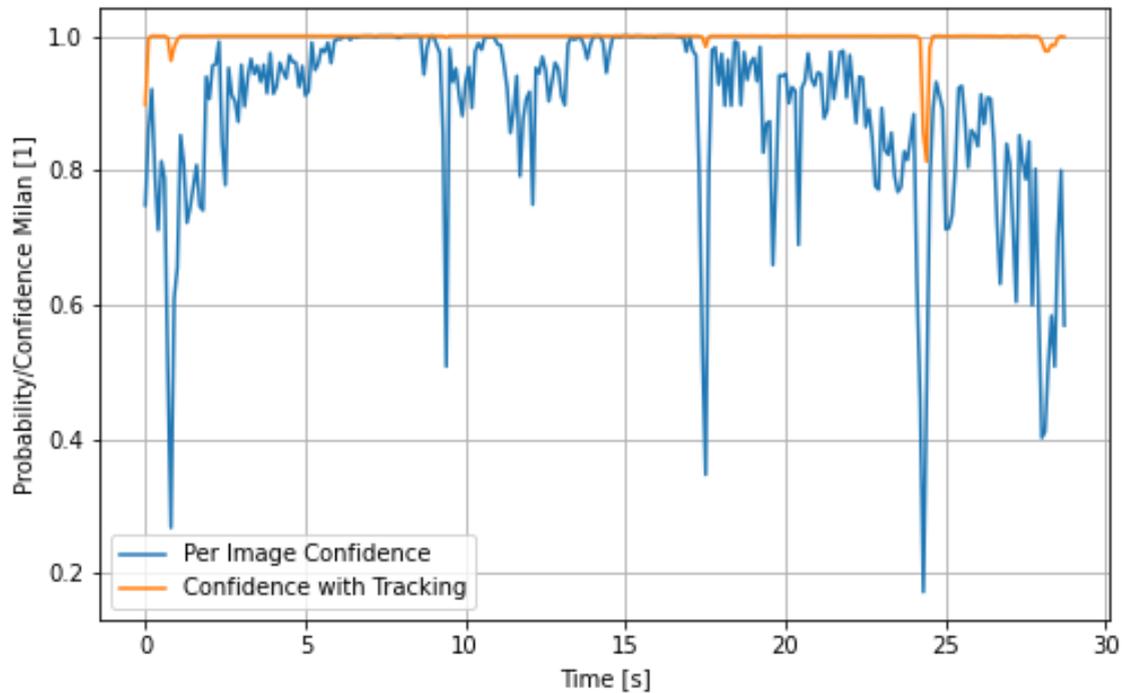


Abb. 18: Wie Abb. 17, jedoch für einen Tag mit guten Sichtverhältnissen für einem Milan im Abstand von 200 bis 700 Metern.

4.4 Stereometrie und Kalibrierung

Um eine möglichst präzise Entfernungsschätzung der von den Stereokameras erfassten Objekte zu gewährleisten, muss vorab eine Kalibrierung der Stereokameras vorgenommen werden. Abb. 19 zeigt den Aufbau der Stereokameras auf dem Schwenkneigekopf. Die Kameras werden unten mittig auf dem Schwenkneigekopf auf den entsprechenden Montagebügeln montiert. Für die Justierung der Kameras sind drei Verbindungselemente mit Gewindestangen und Wantenspannern für die Feineinstellung vorhanden. Als Referenz für die Justierung dienten verschiedene Objekte mit bekannter Verortung, welche über den Schwenkneigekopf zentriert angefahren wurden.



Abb. 19: Stereokamera paar auf dem Schwenkneigekopf.

4.4.1 Entfernungsberechnung

Die Entfernung in Metern wird über 3D-Triangulation bestimmt. Hierbei werden aus der ermittelten Parallaxe (die Verschiebung der korrespondierenden Punkte zwischen den beiden Bildern) die Tiefeninformation zwischen Kamera und Objekt nach Gleichung 1 ermittelt.

$$Z = f * T / d$$

Gleichung 1: Bestimmt die Entfernung Z ein Objektes aufgrund der Brennweite f , des Kamera Abstands T zueinander und der Disparität d .

Da dieser Zusammenhang nicht linear ist, ist die akkurate Justierung und deren Validierung des Gesamtsystems insbesondere für spätere Objekte in der Ferne von hoher Relevanz, da kleine Abweichung im Gierwinkel zu großen Fehleinschätzungen der Entfernung führen (Abb. 20).

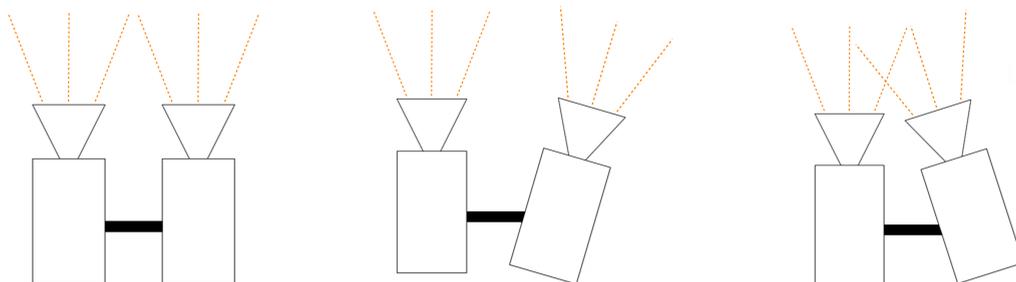


Abb. 20: Mögliche Fehlstellungen der Stereokameras (Gierwinkel), die durch Justierung ausgeglichen werden.

4.4.2 Validierung der Stereometrie

Zur Validierung der Triangulation und auch für den späteren Betrieb wurde ein Algorithmus entwickelt, welcher aus Bildausschnitten nach einer Vorverarbeitungsstufe (Graustufen, Schwellwert, Dilatation, Erosion) die Parallaxe der Stereobilder über eine Autokorrelation berechnet. Anschließend werden verortete Objekte über den Algorithmus ausgewertet und die ermittelten Entfernungen nach Gleichung 1 errechnet.

Tab. 3: Auswertung der Entfernungsmessung mittels Stereometrie und resultierende Fehler nach Abgleich der Messungen mit den Koordinaten bzw. Entfernungen von bekannten Objekten.

Nr.	Pixelversatz [px]	Entfernung (Maps) [m]	Entfernung (berechnet) [m]	Fehler [m]
1	126	81	79,8	-1,2
2	70	142,5	143,6	1,1
3	64	157,5	157	-0,5
4	58	173	173,3	0,3
5	52	204,5	193,3	-11,2
6	44	225	228,4	3,4
7	36	256	279,2	23,2
8	41	256	245,1	-10,9
9	19	559	528,9	-30,1
10	16	709	628,1	-80,9
11	12	803	837,5	34,5
12	14	819,5	717,9	-101,6
13	13	912,5	773,1	-139,4
14	10	1155	1005	-150,0
15	9	1321	1116,7	-204,3
16	8	1375	1256,3	-118,8
17	9	1377	1116,7	-260,3

Tab. 3 zeigt diese Auswertung. Im relevanten Entfernungsbereich bis 800 Meter wurde ein mittlerer Fehler von 17,9 Metern ermittelt, welcher oft nur auf wenige Pixel an Ungenauigkeit im Algorithmus oder auch auf GPS Abweichungen, bei den verorteten Objekten zurück zu führen ist. Bereits ab einer Entfernung von ca. 550 Meter beträgt der Fehler der Entfernungsmessung nur noch 30 Meter. Somit ist die Entfernungsschätzung geeignet, um mit Blick auf die rechtzeitige Abschaltung einer Windenergieanlage eine hinreichende Genauigkeit der Entfernungsmessung mittels Stereometrie zu gewährleisten. Auch hier wird noch einmal darauf hingewiesen, dass sich der Fehler bei Objekten in größeren Entfernungen verstärkt und fortpflanzt und sich sogleich minimiert, wenn sich diese Objekte dem Erkennungssystem nähern. Die Entfernungsmessungen in Tab. 3 sind in Abb. 21 auch nochmals als Grafik in Abhängigkeit vom Pixelversatz dargestellt. Es ist zu erkennen, dass die Fehler sich ab 600 Metern Entfernung erhöhen. Die Indexierung der Punkte für

Entfernungen über 600 Meter dient nur der besseren Identifikation der einzelnen vermessenen Objekte nach Tab. 3. Abb. 22 zeigt den Fehler der Entfernungsmessung mittels Stereometrie in Abhängigkeit von der Entfernung eines Objektes nach Tab. 3. Mit einem mittleren Fehler von 17,9 Metern für den relevanten Entfernungsbereich bis 800 Metern, ist dies für Anwendung des BirdRecorders als Antikollisionssystem – aktuell nur für Milane – zur bedarfsgerechten Abschaltung von Windenergieanlagen bei weitem ausreichend.

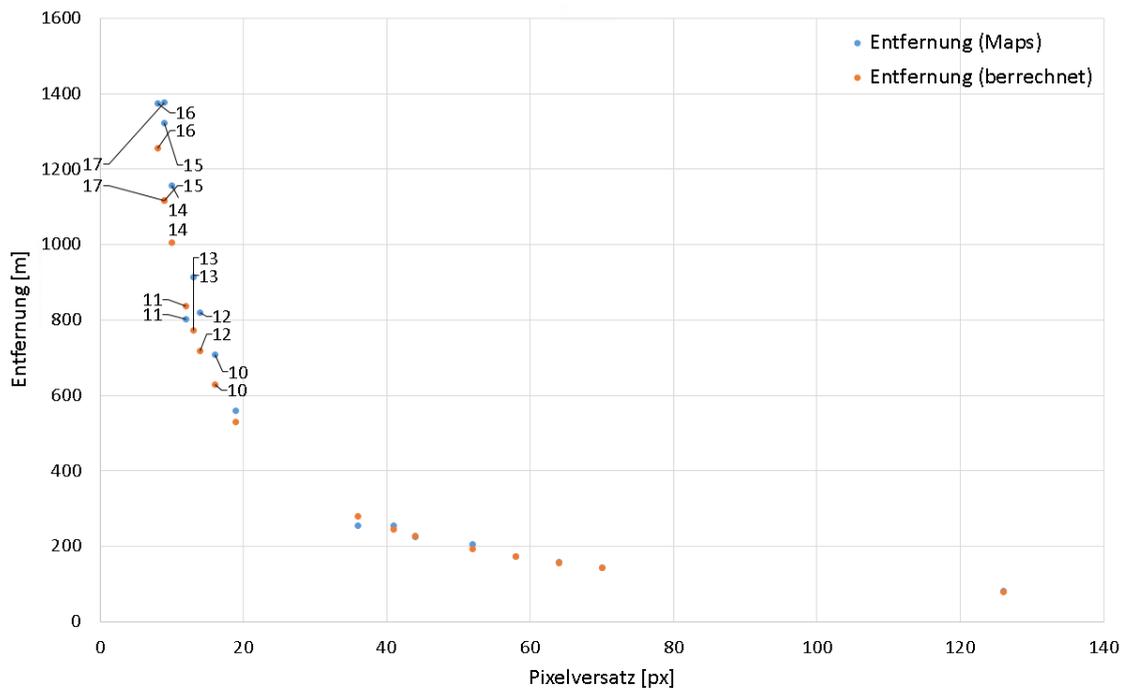


Abb. 21: Entfernung von Referenzobjekten gegenübergestellt mit der geschätzten Entfernung, berechnet über die Pixelparallaxe (Pixelversatz).

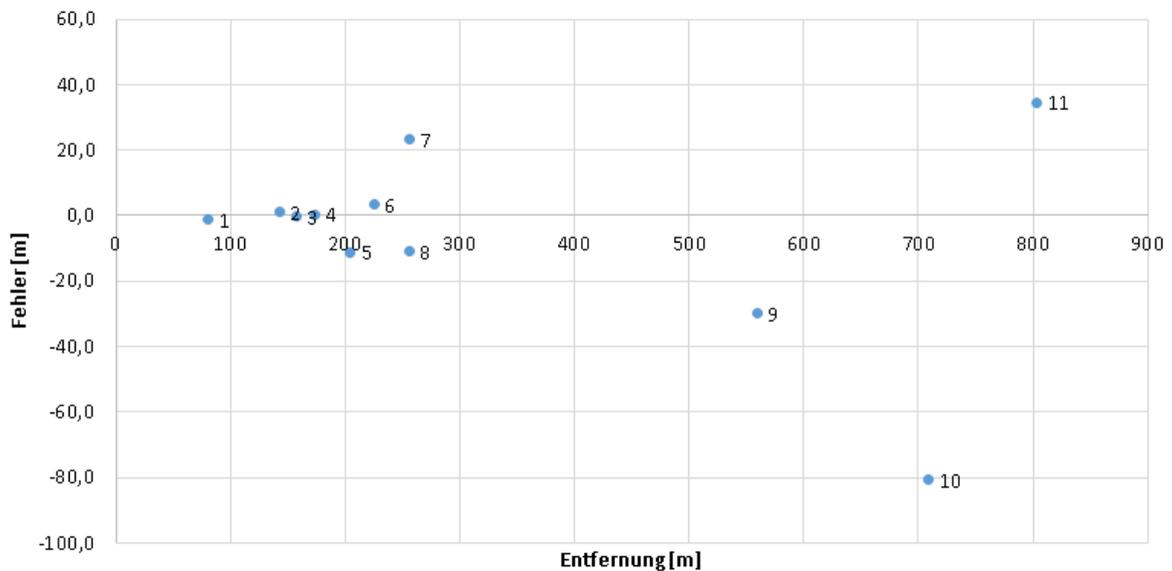


Abb. 22: Fehler der Entfernungsmessung mittels Stereometrie in Abhängigkeit der Objektentfernung.

Mit der damit erreichten Kalibrierung der Stereokameras, können Objekte somit problemlos verarbeitet werden. Da der Schwenkneigekopf aber einen Zeitversatz für das genaue Anfahren eines Objektes bis zur Bildaufnahme aufweist, ist das Zielobjekt nicht immer zwangsläufig im Bild zentriert abgebildet. Daher bedurfte es einer weiteren digitalen Verzerrungskorrektur für die Kalibrierung der Stereokameras, welche über eine Schachbrettmuster-Kalibrierung vorgenommen wurde und nachfolgend näher beschrieben wird.

4.4.3 Ecken Validierung

Eine Problematik des Stereosystems ist die Verzerrung der Bildaufnahmen durch Unzulänglichkeiten der eingesetzten Objektive und Kameras. Die vertikale Verzerrung konnte über die Gewindestangen korrigiert werden. Des Weiteren wurde aber zusätzlich ein Rotationsunterschied (Rollwinkel) der Kameras zueinander festgestellt. Auch wenn dieser sehr viele kleiner als 1° ist, kommt es im Zusammenspiel mit der horizontalen Drift (Neigungswinkel) zu Verzerrungen der korrespondierenden Stereobilder (Abb. 23).

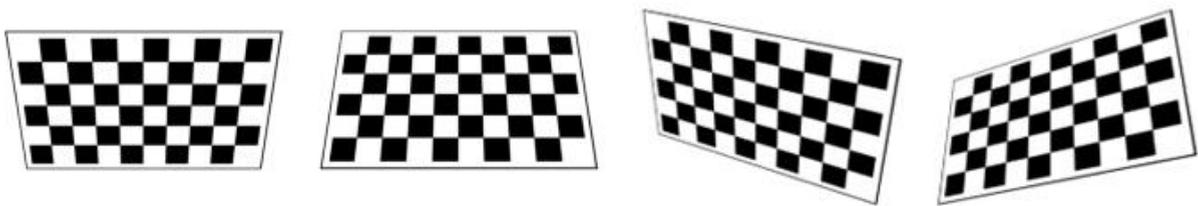


Abb. 23: Mögliche Verzerrungen der Bilder durch Neigungsdrift oder Rolldrift der Stereokameras.

4.4.4 Schachbrett Kalibrierung

Für die Kalibrierung wurde ein Schachbrettmuster auf einer etwa ein Quadratmeter großen Platte eingesetzt. Der Kalibrierungsprozess besteht aus mehreren Bildaufnahmen, bei der das Referenzmuster in allen Bildbereichen abgebildet werden muss und welcher automatisiert über den Schwenkneigekopf vorgenommen wird. Dieser Prozess kann somit bei Standortwechsel oder Neuausrichtung des Systems leicht wiederholt werden. Anschließend werden die Aufnahmen mittels Ecken und Kantendetektion analysiert, und die intrinsischen sowie extrinsischen Kameraparameter ermittelt (Abb. 24).



Abb. 24: Automatisierte Ecken- und Kantendetektion für die Kamerakalibrierung Die detektierten Schnittpunkte farblich gekennzeichnet.

Bei den intrinsischen Parametern wurde festgestellt, dass die Einzelkameras vernachlässigbare Verzerrungen aufweisen. Eine Fehlstellung der Stereoausrichtung und die daraus resultierende Verzerrung wurden über die Kalibrierung wie erwartet nachgewiesen. Der letzte Schritt der automatisierten Kalibrierungsprozesskette ist die digitale Entzerrung der Stereoaufnahmen (Abb. 25).

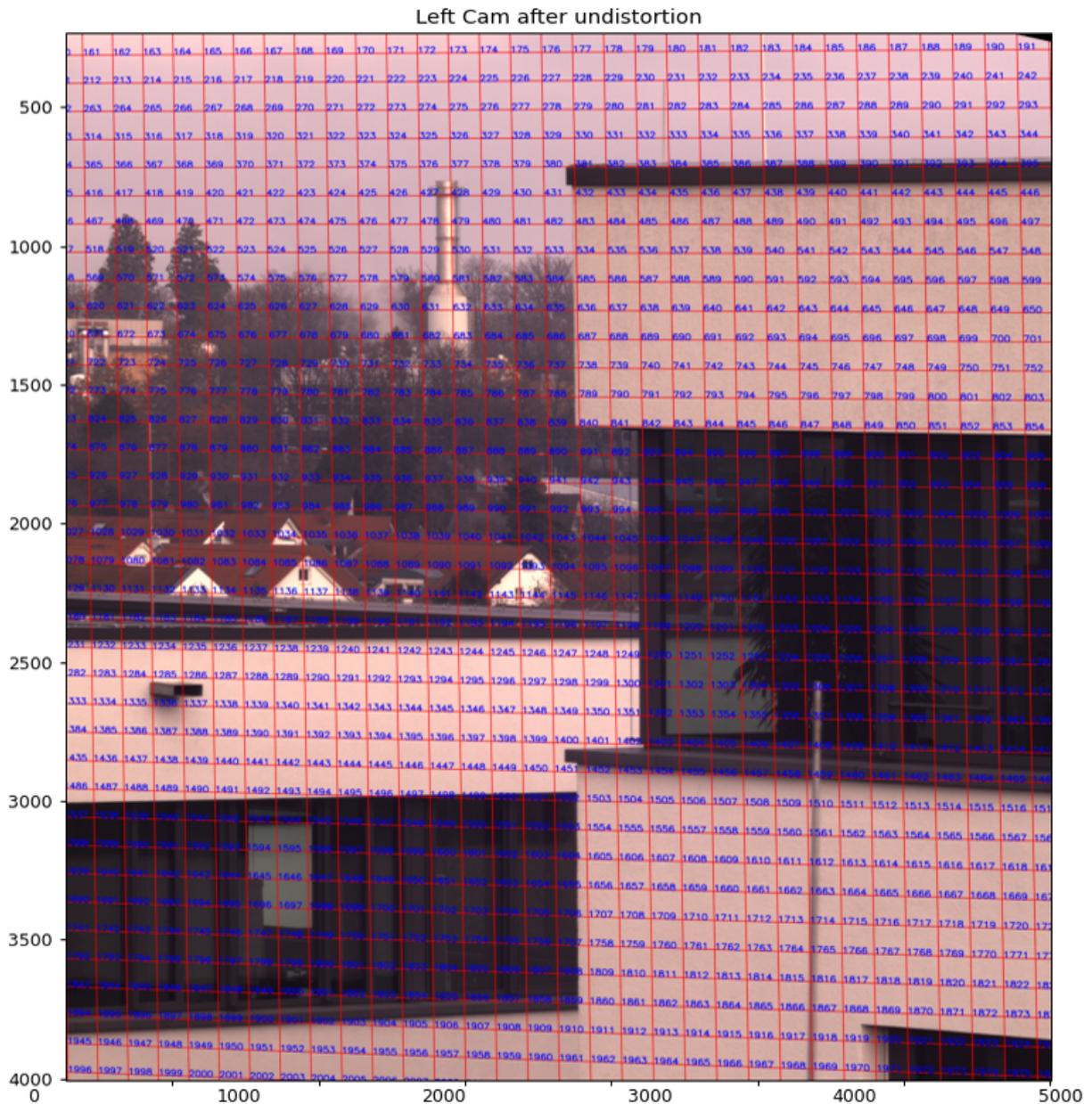


Abb. 25: Entzerrte Aufnahme, nach Stereo Kalibrierung. Das Gitternetz visualisiert die Korrektur nach der Entzerrung und die Zahlen (blau) geben die Pixelkoordinaten im Bild wieder.

4.4.5 Zusammenfassung Stereokalibrierung

Auch nach der Kalibrierung und der digitalen Bildentzerrung sind kleine aber meist vernachlässigbare Fehler für die Entfernungsmessung mittels Stereometrie aufgetreten. Diese sind auf Fehler in der Kalibrierung, der automatisierten Ermittlung der Kameraparallaxe oder minimale Veränderungen der Kamera Justierung zurückzuführen. Gerade letzteres ist eine Problematik, welche im Langzeitbetrieb immer wieder Auftreten kann und zu einer Nachjustierung, inklusive neuer Kalibrierung des Systems führt. Dies verursacht im operationellen Betrieb des BirdRecorder-Systems einem entsprechenden Wartungsaufwand. Es ist für die Qualitätssicherung auf jeden Fall notwendig, Methoden für eine möglichst vollautomatische und regelmäßige Kalibrierung der Stereokameras zu entwickeln und zu

implementieren. Werden durch die regelmäßigen Kalibrierungen jedoch nicht zu vernachlässigende Abweichungen festgestellt, so ist eine manuelle Nachjustierung der Stereokameraeinheit notwendig.

4.5 Softwarearchitektur des Gesamtsystems

Die Softwarearchitektur für den finalen Prototyp des BirdRecorder-System IPC-2 mit bis zu acht statischen Kameras und bis zu zwei Stereokamerapaaren musste gegenüber dem ersten BirdRecorder-System IPC-1 für die Bildaufnahme komplett neu entwickelt und implementiert werden. Zum einen waren für die neuen hochauflösenden Kameras neue leistungsfähigere Framegrabber notwendig und zum andern wurde das System unter dem Betriebssystem Ubuntu Linux entwickelt. Das IPC-1 System dagegen wurde unter einem Microsoft Windows- Betriebssystem betrieben.

Die zu Grunde liegenden Softwarearchitektur ist in Abb. 26 dargestellt. Diese besteht aus den folgenden vier Softwaremodulen als Hauptkomponenten für den Betrieb des BirdRecorder-Systems:

1. Statische Beobachtung: Bildaufnahme, Objekterkennung und Tracking der einzelnen Objekte in den Bildern der statisch montierten Kameras mit 360 Grad Rundumsicht. Die Klassifizierung der erkannten Objekte erfolgt mit dem neuronalen Netz NN-1 (Abb. 6)
2. Stereometrie-Beobachtung: Stereobildaufnahme und Bestimmung der Entfernung des beobachteten Objektes sowie die Erkennung und Klassifizierung des Objektes mit dem neuronalen Netz NN-2 für die Erkennung von Milanen (Abb. 6).
3. Schwenkneigekopf (SNK): Objektverfolgung und Steuerung des SNK aufgrund der erkannten Objekte in den statischen Kameras und den Stereokameras.
4. Koordinator: Dieses Softwaremodul ist für das reibungslose Zusammenspiel und die zeitliche Koordination der drei anderen Softwaremodule verantwortlich.

Die vier Softwaremodule und deren Zusammenspiel für den operationellen Betrieb des BirdRecorder-Systems werden nachfolgend noch näher beschrieben. Alle Softwaremodule sind über Konfigurationsdateien parametrisierbar, so dass die Parameter sehr einfach angepasst werden können.

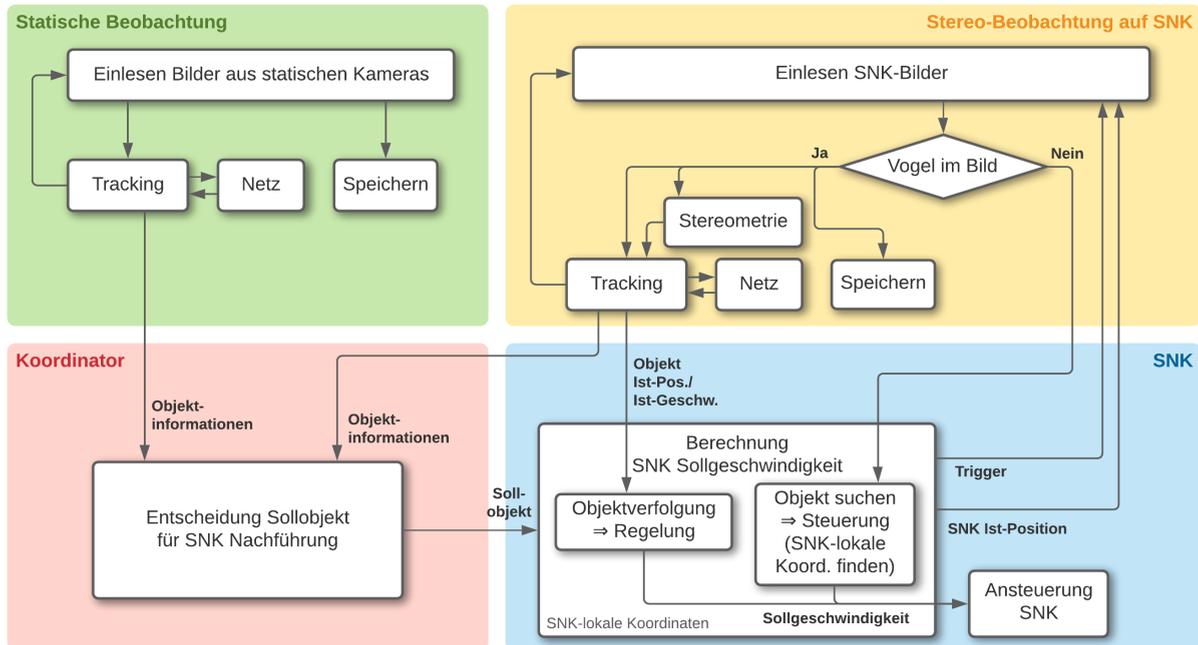


Abb. 26: Grundlegende Softwarearchitektur des BirdRecorder-Systems IPC-2.

4.5.1 Statische Beobachtung

Dieses Softwaremodul ist für die statische Beobachtung eines 360 Grad Rundumblicks mit bis zu acht hochauflösenden Kameras verantwortlich. Jeweils vier Kameras werden mit einem Framegrabber betrieben. Die Objekterkennung selbst erfolgt in Echtzeit auf den Framegrabbern mit einer eigens dafür entwickelten Software, die auf den jeweiligen FPGAs des Framegrabbers in Echtzeit die Objekterkennung realisiert. Der Framegrabber stellt die erfassten Daten und Bilder dem Softwaremodul über DMA (direkt access memory) zur Verfügung. Daraus wird eine Liste von Objekten generiert, die mit dem NN-1 klassifiziert werden. Interessante Objekte (hier Milane) werden dann über einen speziellen Algorithmus getrackt, so dass diese auch über die verschiedenen Kameras kontinuierlich verfolgt werden können. Dazu sind die statischen Kameras so angeordnet, dass diese einen Überlappungsbereich aufweisen (Abb. 4). Die Koordinaten der Objekte und alle weiteren Metainformationen werden über gemeinsame Speicherobjekte an die weiteren Softwaremodule übergeben.

4.5.2 Stereobeobachtung

Dieses Softwaremodul ist für die Bildaufnahme der Stereokamera paare zuständig. Die Stereokameras sind dieselben hochauflösenden Kameramodelle wie die statischen Kameras. Die Stereokameras sind an einen weiteren Framegrabber mit FPGA-Programm für die Objektdetektion angeschlossen, der vier Kameraeingänge für den Anschluss von bis zu zwei Stereokamera paaren ausweist. Das FPGA Programm auf dem Framegrabber macht wie schon bei den statischen Kameras die Objektdetektion und stellt Bilddaten und Objekt-ROIs über DMA zur Verfügung. Auf die ROIs wird das neuronale Netz NN-2 angewendet, um die Objektklassifikation in die Klassen Milan, Vogel, Fluggerät und Sonstiges vorzunehmen. Im Zusammenspiel mit dem Softwaremodul für die statischen Kameras und den Schwenkneigekopf wird ein erkanntes Objekt getrackt und mit jeder zehnten Bildaufnahme die Entfernungsbestimmung mittels Stereometrie vorgenommen. Bei einer Bildaufnahme rate von 10 fps wird die Entfernungsbestimmung somit einmal pro Sekunde

vorgenommen. Dies kann bei Bedarf, etwa bei sehr schnell fliegenden Vögeln, auch auf zwei- bis fünfmal pro Sekunde erhöht werden. Die erfassten Daten werden wiederum über die Objekt Registry dem Softwaremodul Koordinator übergeben, der damit den Gesamttablauf steuert.

4.5.3 Schwenkneigekopf (SNK)

Das Softwaremodul SNK erhält die notwendige Objektinformationen von den Modulen statische und Stereobeobachtung. Aufgrund der Daten wird die Geschwindigkeit des zu trackenden Objekts bestimmt und die Sollkoordinaten des Azimut- und Elevationswinkels berechnet. Diese werden dem SNK übermittelt, der dann die entsprechende Position anfährt. Ist die Sollposition erreicht, werden weitere Bildaufnahmen über das Softwaremodul Stereometrie getriggert. Dieser Regelkreis geht so lange bis ein Objekt den Beobachtungsraum verlassen hat (z.B. Radius bzw. Entfernung > 700 m) oder ggf. durch ein Hindernis, das Objekt sich nicht mehr im einsehbaren Bereich der Stereokameras befindet.

4.5.4 Koordinator

Das Softwaremodul Koordinator ist für das reibungslose Zusammenspiel der beschriebenen Softwaremodule zuständig. Es koordiniert den gesamten Ablauf des operationellen Betriebs des BirdRecorders. Dies beinhaltet vor allem den zeitlichen Ablauf aufgrund der Objektinformationen der statischen Kameras und der Stereometrie. Er ermittelt auf Basis der verfügbaren Objektinformationen das „interessanteste“ Objekt, das dann für das Tracking mit den Stereokameras ausgewählt wird. Wichtig ist auch, dass während ein Objekt mit den Stereokameras getrackt wird, die anderen Objekte von den statischen Kameras parallel dazu ebenfalls weiter erfasst und getrackt werden. Somit steht unmittelbar nach dem Ende des Trackings eines einzelnen Objekts mit den Stereokameras das nächste interessante Objekt für die Stereometrieerfassung zur Verfügung.

Alle erfassten Bilder (ROIs) werden auch abgespeichert und ebenso alle relevanten Metadaten der Bilder in eine lokale Datenbank für spätere Auswertungen gespeichert. Darauf können dann weitere Qualitätssicherungsalgorithmen angewendet werden und die Objekterkennung und Klassifizierung entsprechend nachvollzogen werden.

4.5.5 Monitoring

Für das BirdRecorder System (Soft- und Hardware) wurde ein Monitoringsystem implementiert. Dieses ermöglicht das Aufzeichnen von Metriken vieler Systemkomponenten. Gleichzeitig werden durch die BirdRecorder-Software erfasste Werte, wie beispielsweise die Temperaturen der Kameras und Framegrabber, sowie interne Parameter und aktuelle Informationen der Objekterkennung und des Trackings überwacht.

Alle Metriken werden mittels eines Dashboards visualisiert (Abb. 27). In Abb. 27 oben sind relevante Betriebssystem-Metriken wie CPU und Speicherauslastung, sowie die Füllstände der eingebundenen Festplatten zu sehen. Unten die zeitlichen Verläufe der Kamera- und Framegrabber-Temperaturen, sowie Anzahl der Objekte in der Objekt-Registry. Lücken in den Zeitreihen beruhen auf Nachtabschaltung und Anpassungen der Software. Bei der Annäherung an eigens definierte Grenzwerte, werden entsprechende Warnungen über zwei Nachrichtensysteme verschickt, so dass rechtzeitig auf die Situation reagiert werden kann. Gleichzeitig werden post hoc Analysen von Ausfallsituationen ermöglicht und eine Ursachenforschung vereinfacht.

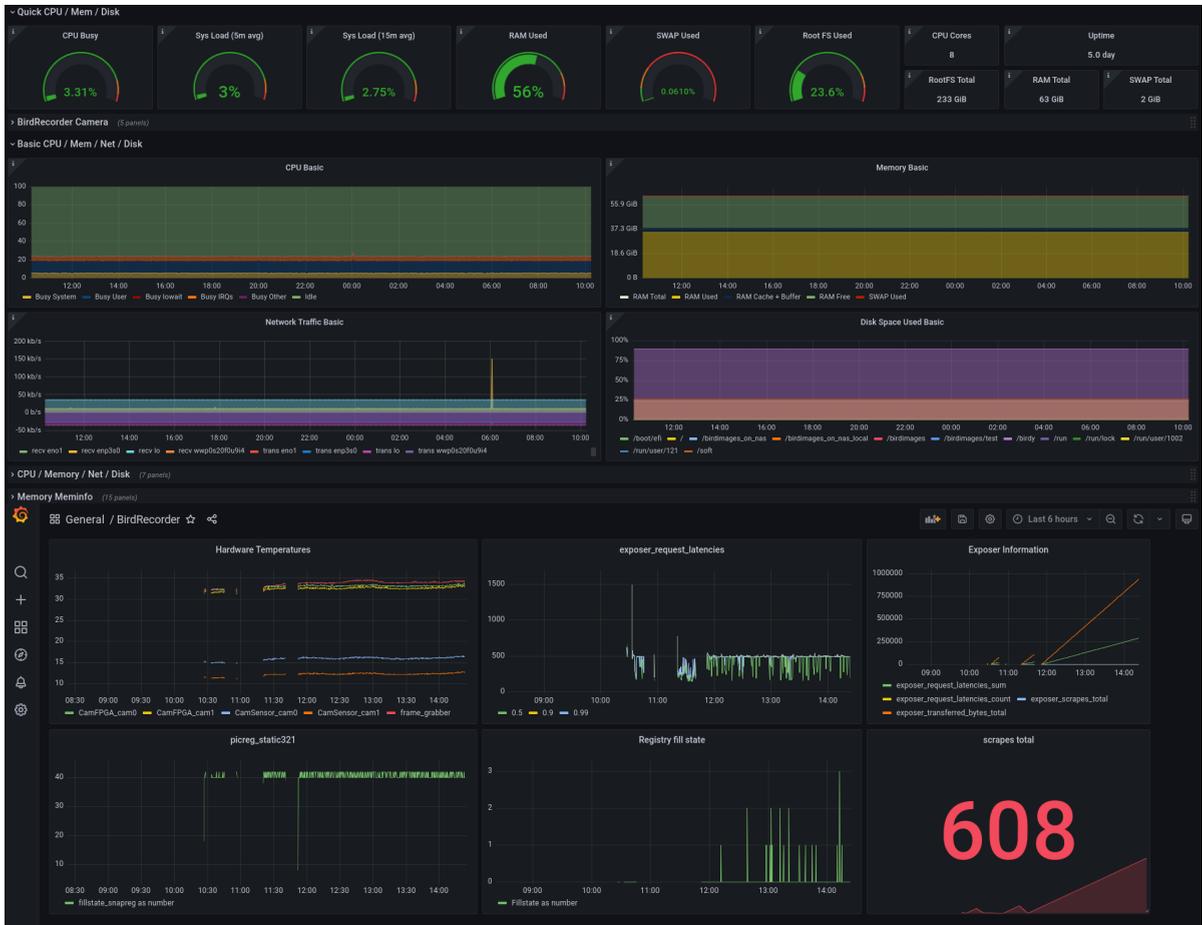


Abb. 27: Dashboard-Visualisierung des Monitorings für BirdRecorder-System.

5 Evaluation

Die in Tab. 1 beschriebenen Bildaufnahmen von mehr als 18 Millionen Rohbildern wurden von den lokalen NAS Speichersystemen in das Datenarchiv am ZSW übertragen und für die Weiterverarbeitung aufbereitet. Die nachfolgenden Auswertungen erstrecken sich auf die Daten der Jahre 2019 und 2020.

5.1 Verschneidung mit externen Vogelbeobachtungen

Im Juni 2019 wurde von der Schweizerische Vogelwarte Sempach im Rahmen des Vorhabens NatForWINSSENT ein Rotmilan namens Donzi besendert. Im nächsten Jahr kam ein weiterer besendertes Rotmilan („Franzi“) hinzu, der sich zeitweilig in der Nähe des Testfelds aufhielt. Die GPS-Messdaten wurden dem ZSW zur Verfügung gestellt. Weiterhin wurden im Rahmen einiger Messkampagnen auf dem Windtestfeld ornithologische Beobachtungen mit Hilfe eines Laser-Rangefinders (LRF) durchgeführt. Der LRF wird dabei mittels GPS eingemessen und berechnet anhand von Winkel und Abstand ebenfalls 3D-Koordinaten der beobachteten Vögel, die in eine Datenbank am ZSW aufgenommen wurden.

5.1.1 Datenvorverarbeitung

Die Besendungsdaten enthalten sowohl eine GPS-Höhenmessung als auch ein druckbasiertes Altimeter. Während das Altimeter sehr schnell auf Druckänderungen reagiert, wird die GPS-Messung weniger oft durchgeführt und weist einige Ausreißer auf. Daher empfehlen die Ornithologen die Verwendung der Altimeterdaten. Diese müssen jedoch mit dem wetterbedingten Bodendruck korrigiert werden. In der Nähe des Testfelds bietet sich hierfür das auf dem Messmast Nord-West in 96 m Höhe (entspricht 761 m ü. NN) installierte Barometer an. Der Druck p in der für das Altimeter angenommenen Standardatmosphäre beträgt auf dieser Höhe 925,1 hPa. Unter vereinfachten Annahmen gilt dann für die Altitude H die Korrekturformel

$$H_{\text{corr}} = H + (p - 925,1 \text{ hPa}) * 8,62 \text{ m}$$

Die entsprechenden Werte sind beispielhaft für einen Tag in Abb. 28 dargestellt. Wie zu erkennen, passen die korrigierten Werte im Mittel besser zu den zeitlich niedriger aufgelösten GPS-Höhenmessungen, während sie gleichzeitig nicht deren Ausreißer nachvollziehen.

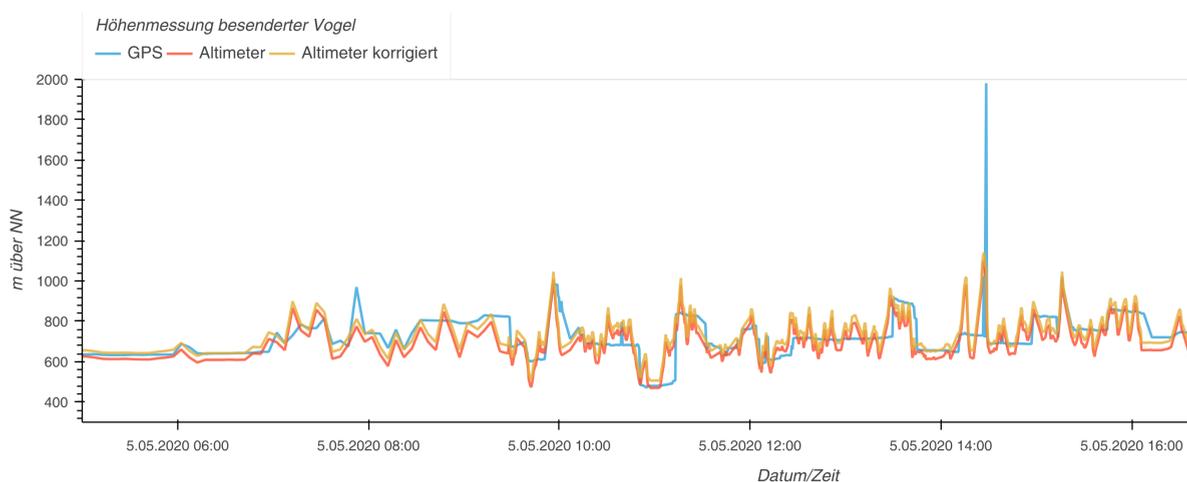


Abb. 28: Beispiel für die Korrektur der Altimetermessung für besenderte Vögel und Vergleich mit GPS Höhenmessungen.

Die LRF-Daten müssen nicht korrigiert werden. Laut Vogelwarte Sempach gibt es hier allerdings teilweise Probleme mit der Azimutausrichtung des Gerätes, da der Kompass manchmal einen systematischen Offset um einige Grad aufweist. Dies lässt sich leider ohne systematische Kalibrierung nicht korrigieren, und verringert möglicherweise die Genauigkeit der Verschneidung etwas.

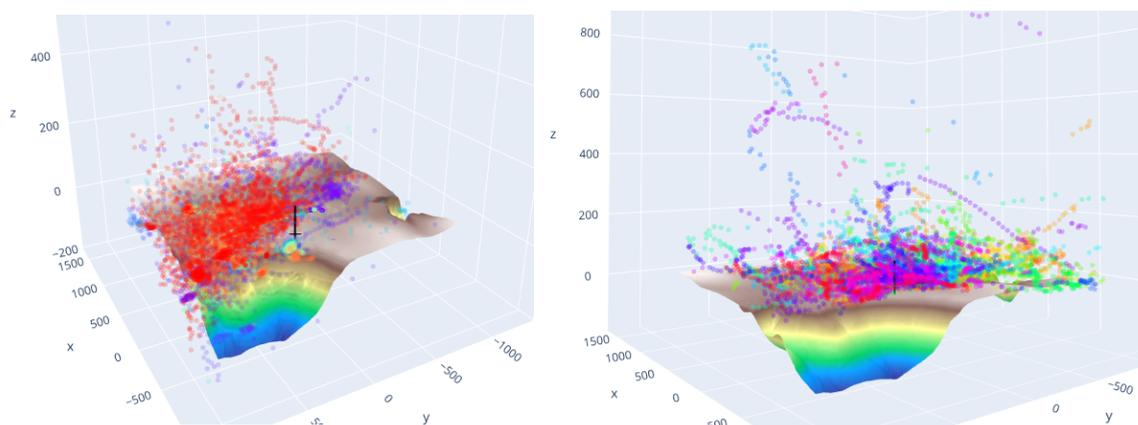


Abb. 29: Visualisierung der Referenzdaten für die Vogeldetektion. Links: GPS, rechts: LRF. Koordinaten in Metern relativ zur Kameraposition.

Um einen besseren Eindruck von der Datenqualität zu bekommen, wurden diese drei-dimensional visualisiert (Abb. 29). Dabei zeigt die x-Achse nach Norden. Die Farben deuten den zeitlichen Zusammenhang – soweit als möglich – an. Der Messmast mit den Kameras ist als schwarze Struktur angedeutet. In Abb. 29 ist zu erkennen, dass die GPS-Daten stark streuen, gelegentlich unter der Geländeoberfläche liegen und selten Strukturen aufweisen, die als Flugbahnen zu identifizieren sind. Im Gegensatz dazu sind die Flugbahnen in den LRF-Daten deutlich zu erkennen und wurden auch von den Beobachtern als zusammenhängende Sequenzen markiert.

5.1.2 Statistik der detektierten Objekte

Um die Verarbeitung der Bilder und Detektion zu erleichtern, wurden alle Metadaten in eine SQL-Datenbank eingepflegt. Insgesamt waren zu Projektende ca. 15 Mio. Fotos der beiden Kameras sowie ca. 23,8 Mio. detektierte Objekte (ROIs) in der Datenbank registriert.

Tab. 4: Verteilung der klassifizierten ROIs.

Klasse	Anteil	Bemerkung
Milan	9,45 %	Rot- und Schwarzmilan wurden zu einer Klasse zusammengefasst
Anderer Vogel	15,87 %	Alle als Vogel erkannte Objekte außer Milan
Fluggerät	14,09 %	Meist Kleinflugzeuge vom nahen Flugplatz Messelberg
Sonstiges	60,59 %	Hauptsächlich Wolken, Insekten, Lichtartefakte

Die folgenden Auswertungen beziehen sich auf das neuronale Netz NN-2, das in Abschnitt 4.3 näher beschrieben wurde. Dieses wurde auf alle ROIs angewandt und die ROIs jeweils der Klasse mit der größten Konfidenz zugewiesen. Die resultierende Verteilung ist aus Tab. 4 ersichtlich.

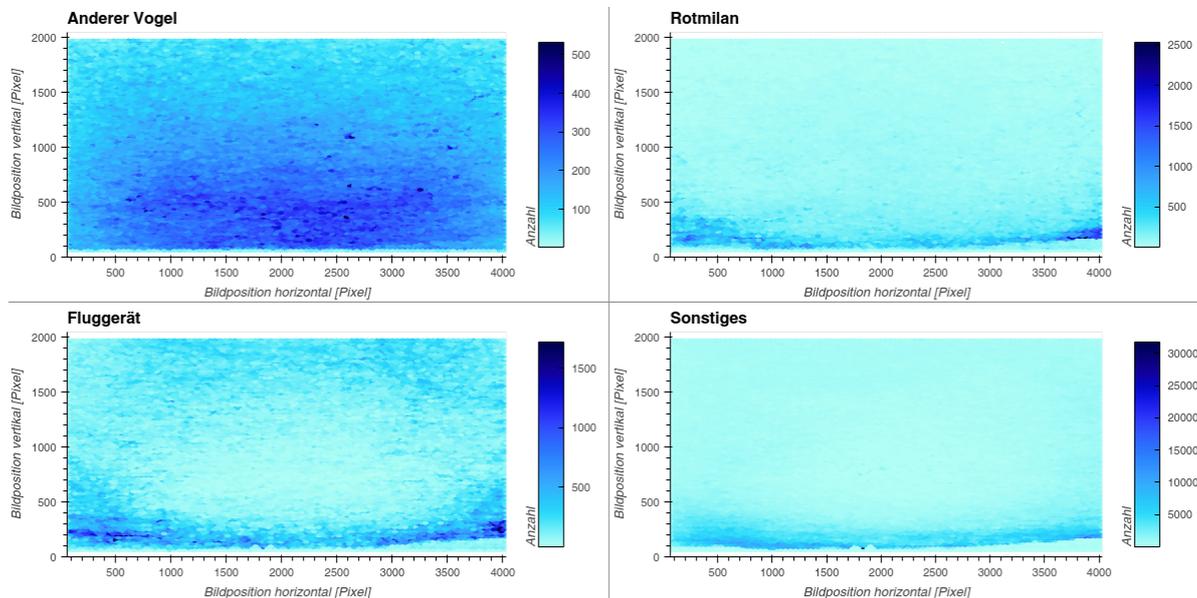


Abb. 30: Verteilung der durch das neuronale Netz NN-2 bestimmten Klassen über die Bildfläche der Südkamera.

Bei Betrachtung der räumlichen Verteilung einzelner Klassen über die Bildfläche der Kameras (Abb. 30) fällt auf, dass die meiste Aktivität sich in der Nähe des Horizonts abspielt. Lediglich die Objekte der Klasse „anderer Vogel“ treten relativ gut verteilt auf. Die Artefakte, die bei den anderen Klassen in Horizontnähe zu beobachten sind, entstehen durch die Maskierung des nicht rein aus Himmel bestehenden Bildanteils. Diese Maske war nur vor dem 20.05.2020 nötig, danach wurde die Kamera neu ausgerichtet. Der höchste Wert auf dem Diagramm für die Klasse „Sonstiges“ liegt bei horizontaler Position bei ca. 1800. Hier hat sich offenbar ein Baum bei starkem Wind wiederholt über die Maske hinausgeschoben. Dass diese ROIs praktisch ausnahmslos als „Sonstiges“ klassifiziert wurde, zeigt die hohe Reproduzierbarkeit der Detektion. Bei der Nordkamera ergibt sich ein ähnliches Bild (Abb. 31); Auch wurde die Maske an wenigen Stellen von einem Objekt überragt, mit demselben Effekt auf die Klassifikation.

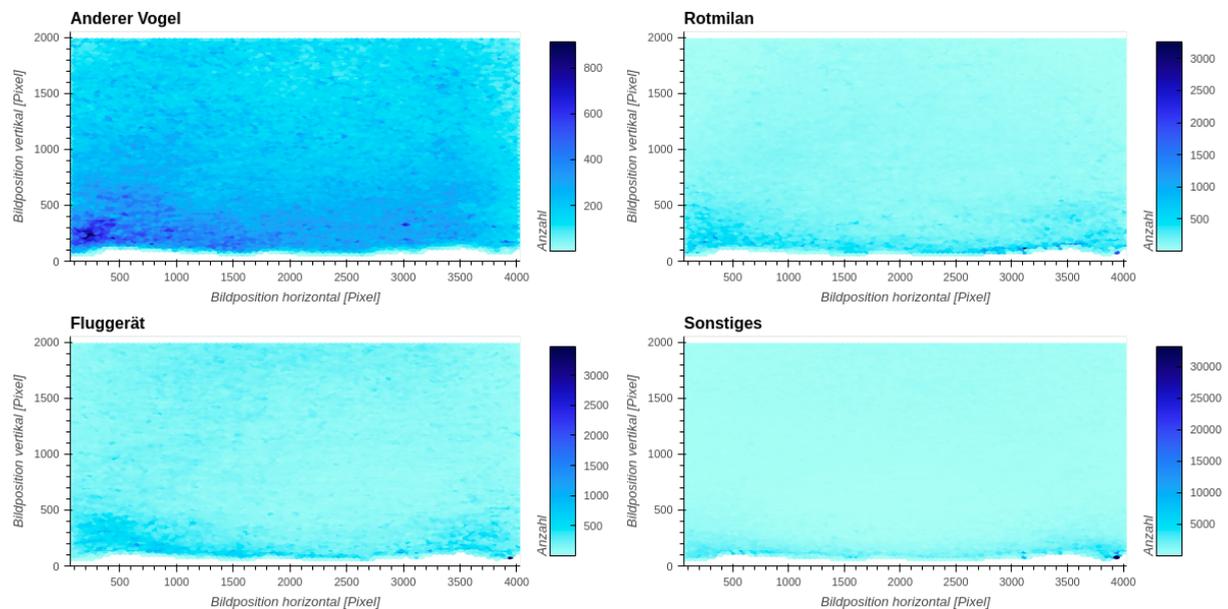


Abb. 31: Wie Abb. 30, aber für die Nordkamera.

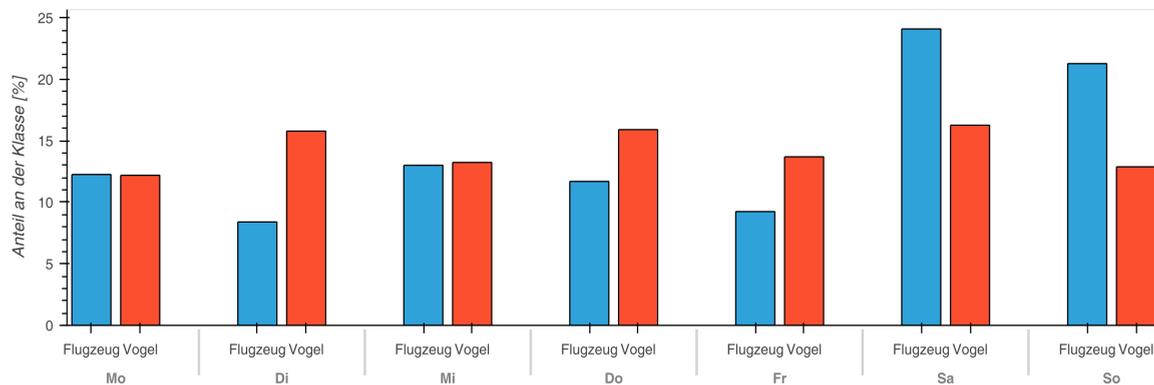


Abb. 32: Anteil der Wochentage an der Gesamtanzahl der in eine Klasse fallenden ROIs. Unter "Vogel" sind hier alle Vögel incl. Milane zusammengefasst.

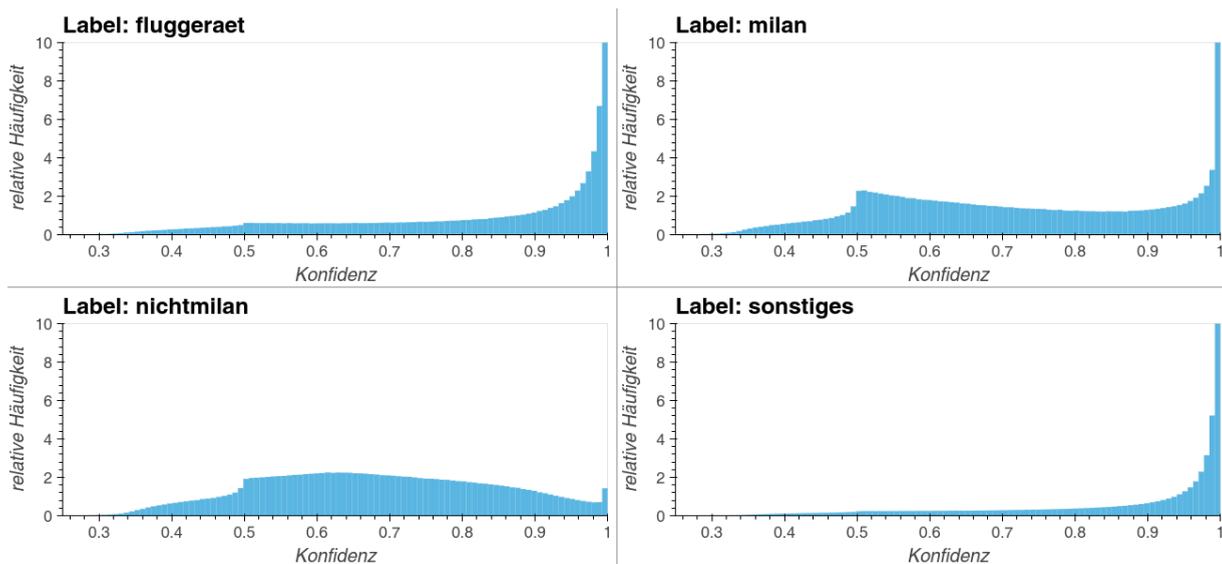


Abb. 33: Verteilung der Konfidenz über die Klassen. Da eine Klasse nur gewählt wird, wenn sie von allen vier Klassen die höchste Konfidenz hat, liegt der Minimalwert bei 0,25. Die Histogrammhöhe für den Konfidenzwert 1,0 sind der besseren Detaildarstellung wegen abgeschnitten

Die wochentägliche Verteilung der Labels ist ebenfalls plausibel (Abb. 32). An Wochenenden herrscht reger Flugbetrieb auf dem nahe gelegenen Flugplatz Messelberg, der sich in der vermehrten Detektion von Fluggeräten widerspiegelt. Die als Vögel erkannten ROIs verteilen sich hingegen relativ gleichmäßig über die Woche.

In Abb. 33 ist die beobachtete Konfidenz für die einzelnen Klassen dargestellt. Wie zu erkennen, werden die „Sonstigen“ Labels gut erkannt, und auch die „Fluggeräte“ weisen nur eine relativ geringe Wahrscheinlichkeitsmasse bei niedrigen Werten auf. Die Unterscheidung zwischen „Milan“ und „Nichtmilan“ (also „sonstiger Vogel“) fällt dem neuronalen Netz NN-2 jedoch deutlich schwerer, weswegen bei 0,5 eine Kante entsteht. Diese deutet darauf hin, dass hier eine Entscheidung zwischen zwei der vier Klassen gefällt wurde. Offenbar gibt es Fälle, in denen der Milan für das NN-2 eindeutig zu erkennen ist, jedoch aufgrund der Heterogenität der „Nichtmilan“-Klasse kaum Fälle, in denen diese mit sehr hoher Konfidenz gewählt wurde. Das ist die Auswirkung des Kompromisses zwischen Anzahl der Beobachtungen und Klassenschärfe. Wären genügend gelabelte ROIs für alle Vogelarten vorhanden, könnten diese als Einzelklassen sicherer erkannt werden, so wie im Fall der „Milan“-Klasse.

5.1.3 Verschneidung und Auswertung

Um zu einer Aussage über die Zuverlässigkeit des Vogelerkennungssystems zu gelangen, müssen die GPS und/oder LRF-Vergleichsdaten mit den klassifizierten ROIs zeitlich und räumlich verschnitten (kollokiert) werden. Da es mit den Weitwinkel-Kamerabildern nicht möglich ist, eine 3D-Verortung der aufgenommenen Vögel zu berechnen, müssen umgekehrt die 3D-Vergleichsdaten auf die 2D-Kameraebene projiziert werden. Die als „Sonstiges“ klassifizierten ROIs werden hierbei nicht betrachtet, da sie sich wie oben gezeigt, zuverlässig erkennen lassen und für die Auswertung hier nicht weiter betrachtet werden.

Als wichtigstes Verschneidungskriterium wird der zeitliche Abstand zwischen Beobachtung und ROI angesehen. Die Uhren aller beteiligten Geräte erhalten entweder Internetzeit oder GPS-Zeit und sollten daher gut übereinstimmen. Die Pixeldistanz auf der Bildebene ist ebenfalls wichtig, weist aber aufgrund von Ungenauigkeiten in Kameraausrichtung, Kameraprojektion, GPS-Koordinaten, Höhenmessung und LRF-Kompass teilweise Diskrepanzen auf. Sie wird daher sekundär berücksichtigt.

Die GPS-Daten der besenderten Vögel stellten sich trotz der durchgeführten barometrischen Korrekturen der Höhenangaben als wenig geeignet für die automatisierte Auswertung heraus. Wie bereits in Abb. 29 zu erkennen, liegen fast keine Daten für die Südkamera vor, während im Norden die meisten Daten wahrscheinlich von sitzenden Vögeln im nahegelegenen Baumbestand stammen, die jedoch durch die Ungenauigkeiten der Höhenmessung häufig nicht im Kamerasichtfeld verortet werden. Dieser Effekt ist deutlich in Abb. 34 zu erkennen. Die Farben entsprechen Sequenzen der ROI-Daten, da die Sender keine Sequenzinformation liefern. Die Scheiben sind proportional der Blobgröße der ROIs skaliert, „Rm“ steht jeweils für eine GPS-Sendermessung. Übereinstimmende Farben bezeichnen Kollokationen. D.h. nur wenige kollokierte GPS-Daten lassen sich eindeutig Flugrouten im freien Raum zuordnen, die meisten befinden sich am Horizont oder darunter. Aufgrund der zu erwartenden Ungenauigkeiten wurde von einer statistischen Auswertung dieser Kollokationen der GPS-Daten Abstand genommen.

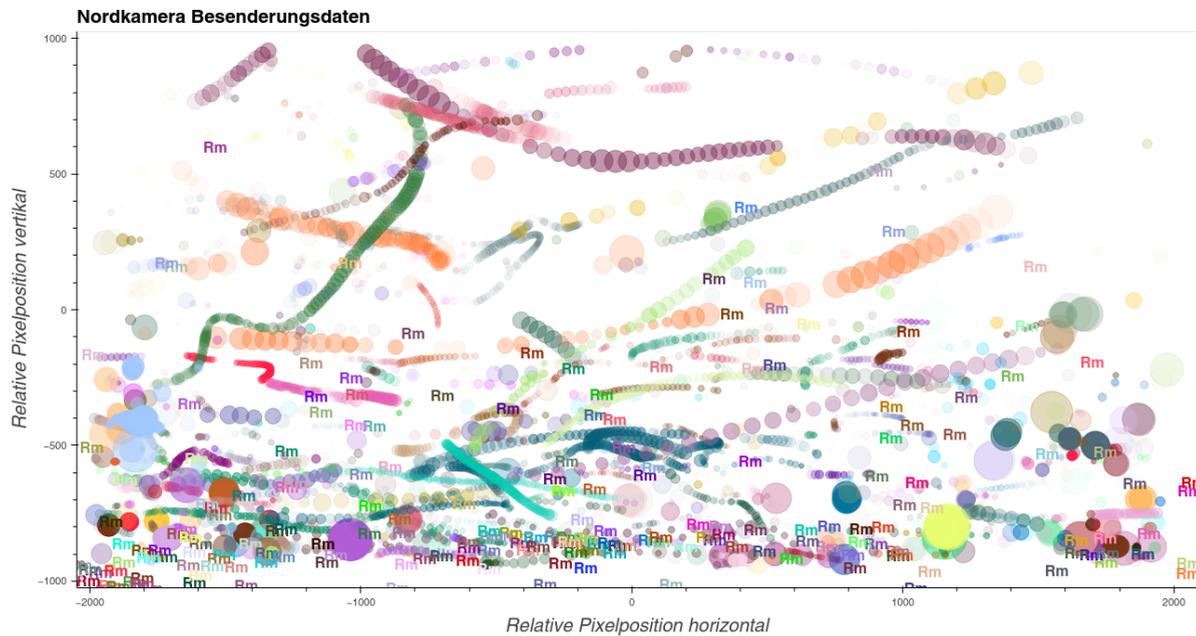


Abb. 34: Darstellung der Verschneidung des Sichtfelds der Nordkamera mit GPS Besenderungdaten von Rotmilanen.

Ein Beispiel einer Verschneidung mit LRF-Daten ist in Abb. 35 zu sehen. Die roten Zahlen in Abb. 35 geben die vom LRF an dieser Stelle gemessene Entfernung zum Vogel wieder, gefüllte Punkte sind weitere, nicht dargestellte ROIs, skaliert mit ihrer Blobgröße. Der Milan startet unten rechts und wird zunächst nicht richtig erkannt, bewegt sich aber dann auf die Kamera zu, wobei sich die Erkennungswahrscheinlichkeit verbessert. D.h. in Abb. 35 ließ sich die vom LRF gemessene Flugbahn des Rotmilans sehr gut mit der Kamera verfolgen. Alle ROIs der betreffenden Sequenzen sind hier dargestellt. Für die statistische Auswertung wurden nur die nahe den LRF-Punkten gelegenen Kollokationen verwendet.

Um eine zuverlässige Aussage über die Erkennungsrate des Systems zu erhalten, müssten sehr viele Kollokationen in verschiedenen Entfernungen verglichen werden. Leider wurden größtenteils Rotmilane vermessen, mit den Besenderungdaten sogar ausschließlich. Durch die im ersten Prototypen des BirdRecorder-Systems (IPC-1) noch fehlende Stereometrie ist es auch nicht möglich, die ROIs ohne externe Vergleichsdaten einer Entfernung zuzuordnen. Abb. 36 zeigt die Approximation einer Erkennungsrate auf Einzelbildern. Die Kurve in Abb. 36 wurde durch das Verhältnis von geglätteten Entfernungshistogrammen von korrekten erkannten ROIs (OK) zu allen betrachteten ROIs (OK, FN = Milan als anderer Vogel erkannt, FP = Anderer Vogel als Milan erkannt) berechnet. Mit XX markierte ROIs wurden aufgrund der Filterung der Klasse Sonstiges nicht berücksichtigt. Die vertikalen Punktwolken entstehen dadurch, dass die jeweils zeitlich und räumlich einer LRF-Messung benachbarten ROIs deren Entfernung zugewiesen bekamen.

Wie aus Abb. 36 folglich ersichtlich nimmt die Erkennungsrate bis etwa 400 m mehr oder weniger linear ab. Danach sind die Ergebnisse aufgrund der schlechten Statistik nicht belastbar. Bei den fälschlicherweise als Milan klassifizierten Vögeln ab 450 m (gelb) handelt es sich zumeist um Mäusebussarde.

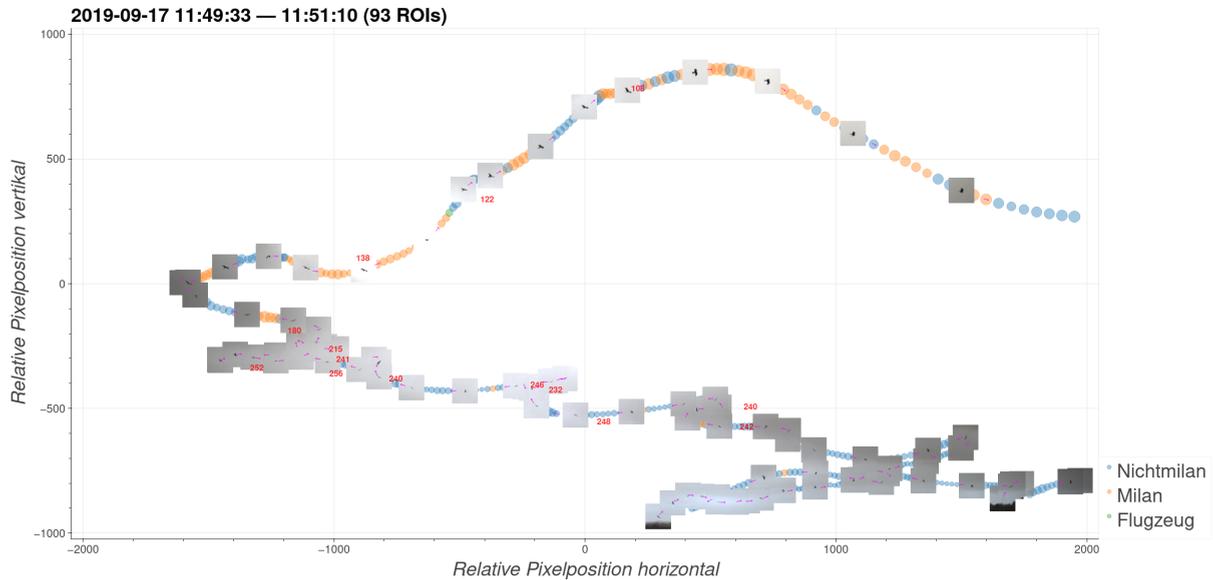


Abb. 35: Flugroute eines kontinuierlich verfolgten Milans auf der Bildfläche der Südkamera am 17.9.2019 mit ROIs über einen Zeitraum von 85 s (Beschreibung siehe Text).

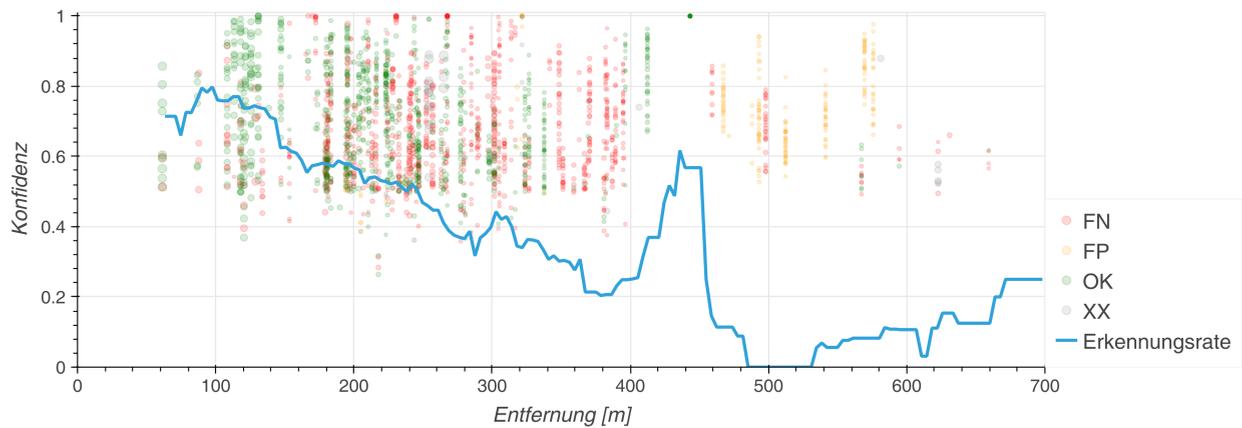


Abb. 36: Abschätzung einer Erkennungsrate für einzelne ROIs aller Kameras für die Kollokation mit den LRF-Daten (Erklärung siehe Text).

Mit Hilfe des in Abb. 17 beschriebenen Tracking-Algorithmus kann die Sicherheit der Klassifikation erhöht werden. Allerdings lässt der Algorithmus an sich keinen Rückschluss darauf zu, ab wann der Vogel „sicher“ als Milan eingestuft werden kann, da bei entsprechend niedrigen Einzelwahrscheinlichkeiten auch wieder eine „sichere“ Einstufung als anderer Vogel erfolgen kann (Abb. 37). Der Index in Abb. 37 läuft linear über das im Titel angegebene Zeitintervall, bei einer Bildaufnahmefrequenz von 10 Bildern/s. Die Posterior-Wahrscheinlichkeit wird nach dem Satz von Bayes mit dem in Abschnitt 4.3 beschriebenen Tracking-Algorithmus berechnet.

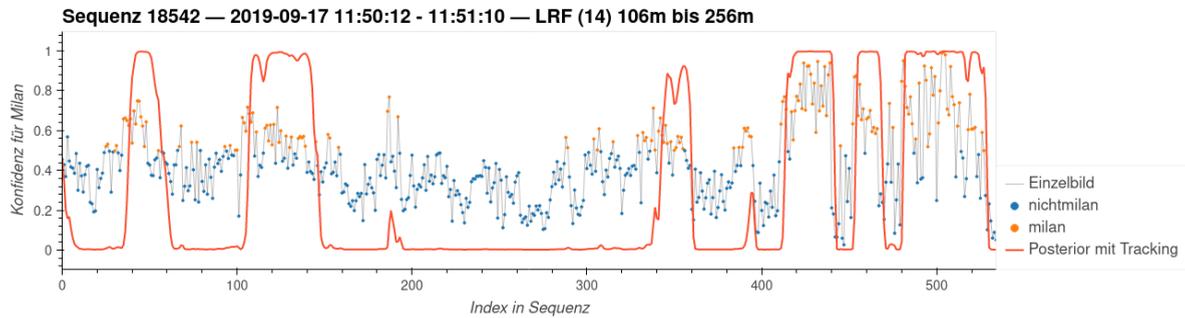


Abb. 37: Entwicklung der kumulierten Konfidenz für eine Milanklassifikation für die Sequenz aus Abb. 35.

Für den BirdRecorder besteht die Aufgabe dabei lediglich darin, zu entscheiden, wenn ein Objekt in den Bildaufnahmen der statischen Kameras als untersuchungswürdig erscheint und dann entsprechend weiter getrackt werden soll. Dies wäre im vorliegenden Fall (Abb. 37) bereits ganz am Anfang der Sequenz erfolgt, denn das System war sich bereits anfangs sicher, dass es sich bei dem Objekt zumindest um einen Vogel handelt.

Aus dieser Betrachtung resultiert eine andere Interpretation der Klassifikationsrate: Geht es nur darum zu entscheiden, ob es sich um einen Vogel (OK) handelt oder nicht (FN), ergibt sich Abb. 38. Ein Vogel wird also bis mindestens 300 m Entfernung noch mit etwa 80%-iger Sicherheit erkannt. Über diese Entfernung hinaus gibt es weitere Vogeldetektionen, aber keine Vergleichsdaten mehr, um diese zu überprüfen.

Abb. 39 zeigt ein weiteres Beispiel der Flugrouten von zwei Rotmilanen, die sicher erkannt wurden (vgl. auch Abb. 35).

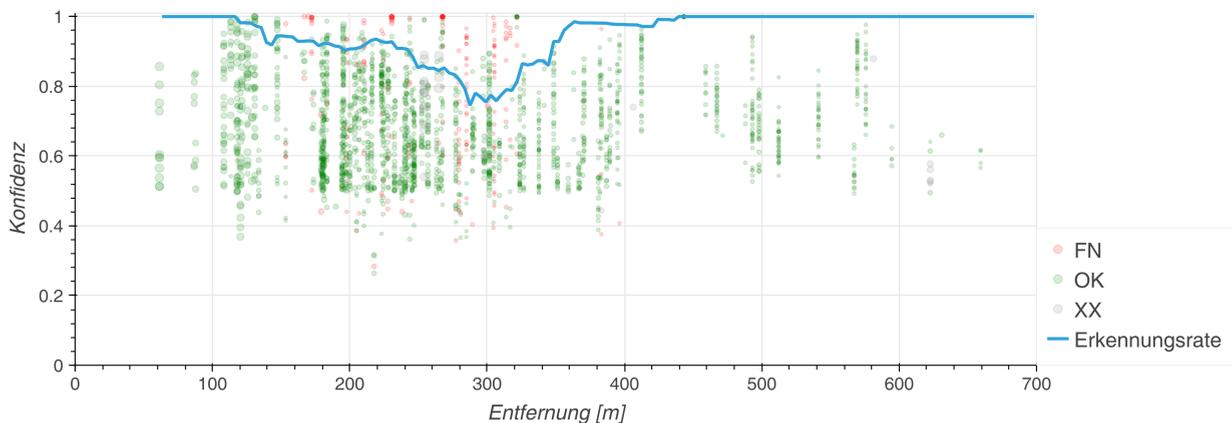


Abb. 38: Wie Abb. 36, aber Erkennung bezieht sich nur auf die Klassen Vogel und Nicht-Vogel und nicht auf die Vogelart.

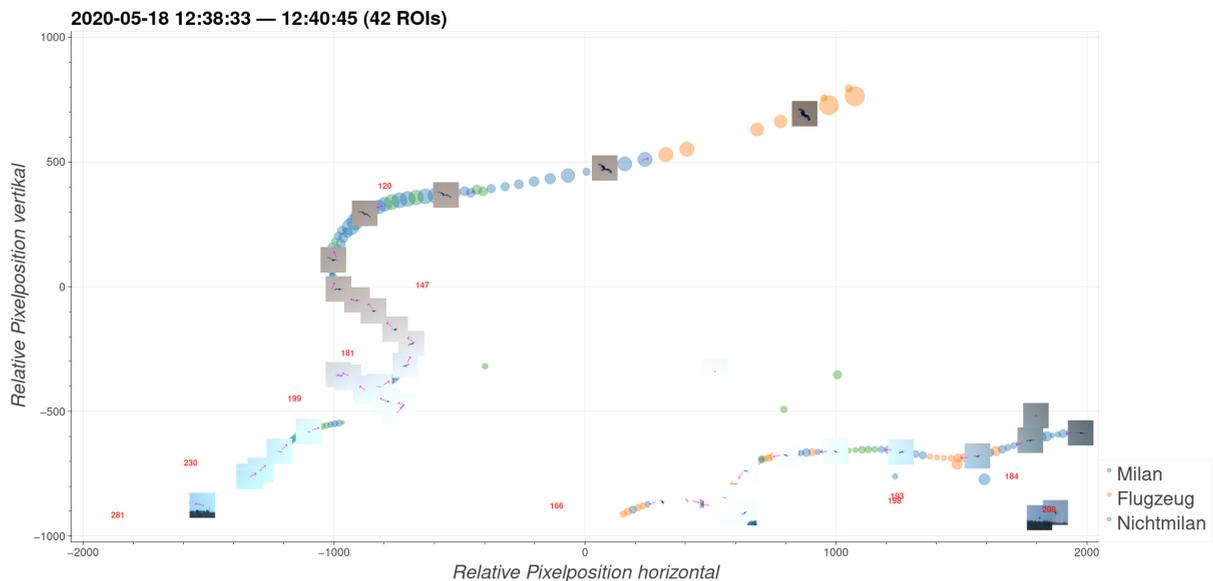


Abb. 39: Flugroute zweier kontinuierlich verfolgter Milane der Nordkamera am 18.05.2020 und die erkannten ROIs. Symbole wie in Abb. 35.

5.1.4 Fazit Objekt- und Milanerkennung

Die in Abschnitt 5.1.3 dargestellten Auswertungen beziehen sich auf die mit dem BirdRecorder-System IPC-1 aufgenommen Bilder und Daten im Zeitraum Juli 2019 bis August 2021. Das IPC-1 System ist noch nicht das finale Prototypsystem BirdRecorder IPC-2 zum Ende des Vorhabens. D.h. die Erkennungsraten erreichen noch nicht die finalen Erkennungsraten des IPC-2 Systems mit den hochauflösenden Kameras. Die Objekterkennung und Klassifikation der Vogelarten (hier Erkennung der Klasse Milan) mit den statischen Kameras dient dazu, interessante Objekte für das Tracking mit der Stereokameras zu erfassen und die initialen Koordinaten für das Tracking zu bestimmen.

Erst nach der erstmaligen Erfassung von einem Objekt mit den Stereokameras, erfolgt das finale Tracking und die finale Erkennung der Vogelart. Im Vergleich zu den gezeigten Auswertungen in Abschnitt 5.1.3 erfolgt die Bilddatenerfassung der statischen Kameras des IPC-2 Systems zum einen mit der doppelten Kameraauflösung und zum anderen mit einer Teleobjektivoptik der Stereokameras mit nur 9,6° Gesichtswinkel. Die bisherigen statischen Kameras haben einen Gesichtswinkel von 60° und weisen mit 12 MP eine nur halb so hohe Kameraauflösung aus.

Mit den 12 MP Kameras konnten wir für die Objektdetektion bei einer Falsch-Negativ-Rate von 0,7 % eine Reichweite von 300 m bis 400 m erreichen (Abschnitt 4.3). Dies bedeutet, dass nur 0,7 % der Vogel-Objekte nicht als Vogel erkannt wurden. Mit den 26 MP Kameras in dem aktuellen BirdRecorder-System IPC-2 wird die Reichweite für die Objektdetektion bei gleichbleibender Falsch-Negativ-Rate verdoppelt, so dass die angestrebten 700 m Reichweite erreicht werden. Mit den Teleobjektiven der Stereokameras wird die Auflösung der Objekte nochmals mindestens um den Faktor 4 erhöht. Mit den bisherigen 12 MP Kameras nach Abb. 36 wurde eine Erkennungsrate auf Basis der Einzelbilder von ca. 80 % im Entfernungsbereich von bis zu 150 m erreicht. Nach der Objekterkennung und Auswahl von interessanten Objekten auf Basis des neuronalen Netzes NN-1 (Objekterkennung und Klassifikation) erfolgt das Tracking von Einzelobjekten mit den Stereokameras. Auf die Stereobildaufnahmen wird wiederum das neuronale Netz NN-2 angewendet und die Verbundwahrscheinlichkeit nach dem Satz von Bayes (Abb. 17, Abb. 18 und Abb. 37)

berechnet. Mit der doppelten Kameraauflösung der statischen Kameras und der um ca. Faktor 8 erhöhten Auflösung der Stereokameras, ist mit dem BirdRecorder-System IPC-2 eine Erkennungswahrscheinlichkeit bis 800 m Reichweite von ca. 90 % für die Klasse Milan zu erwarten. Dies ist eine eher konservative Abschätzung, die das angestrebte Entwicklungsziel mit einer Reichweite von 700 m übersteigt. Diese Distanz ist nach Ammermann et al., 2020 ausreichend, um Windenergieanlagen bei Annäherung von Milanen rechtzeitig abzuschalten.

Die weiterführende Erprobung und Qualifizierung des BirdRecorder-System IPC-2 ist auf Basis der hier präsentierten und durch das BfN mit Mitteln des BMUV geförderten Arbeiten im Laufe des Jahres 2022 und 2023 geplant. Hierbei werden insbesondere der Aufbau und Betrieb des Systems mit Stereokamera-Tracking, Validierung der Erkennungsraten und Nachtraining der neuronalen Netze wesentliche Arbeiten sein. Dies beinhaltet auch die Installation des BirdRecorder aus dem Windenergie-testfeld in Stötten, um weitere Erfahrungen für den Langzeitbetrieb des Systems zu sammeln.

6 Ausblick

Das abgeschlossene Vorhaben BirdRecorder soll im Vorhaben Bird- und BatRecorder 2.0 (BBR 2.0): „Weiterentwicklung des BirdRecorders mit neuester Sensortechnologie und Erweiterung für nachtaktive Vögel und Fledermäuse zur Minderung von Kollisionen von Vögeln und Fledermäusen an Windenergieanlagen“ fortgesetzt und weiterentwickelt werden. Hierzu wurde im Rahmen des 7. Energieforschungsprogramms der Bundesregierung Innovationen für die Energiewende im August 2021 eine entsprechende Antragskizze und im Mai 2022 ein entsprechender Antrag auf Zuwendung eingereicht. Das Vorhaben konnte zum 1. Januar 2023 gestartet werden.

Literaturverzeichnis

- Abdel-Hamid, O., Deng, L., Yu, D., 2013. Exploring Convolutional Neural Network Structures and Optimization Techniques for Speech Recognition.
- Ammermann, K., Bruns, E., Ponitka, J., Schuster, E., Sudhaus, D., Tucci, F., 2020. Technische Systeme zur Minderung von Vogelkollisionen an Windenergieanlagen. BfN Skripten 571 30.
- Cao, P., Li, X., Mao, K., Lu, F., Ning, G., Fang, L., Pan, Q., 2020. A novel data augmentation method to enhance deep neural networks for detection of atrial fibrillation. *Biomed. Signal Process. Control* 56, 101675. <https://doi.org/10.1016/j.bspc.2019.101675>
- Chen, L., Li, S., Bai, Q., Yang, J., Jiang, S., Miao, Y., 2021. Review of Image Classification Algorithms Based on Convolutional Neural Networks. *Remote Sens.* 13, 4712. <https://doi.org/10.3390/rs13224712>
- Gayer, A., Chernyshova, Y., Sheshkus, A., 2019. Effective real-time augmentation of training dataset for the neural networks learning, in: Nikolaev, D.P., Radeva, P., Verikas, A., Zhou, J. (Eds.), Eleventh International Conference on Machine Vision (ICMV 2018). Presented at the Eleventh International Conference on Machine Vision, SPIE, Munich, Germany, p. 64. <https://doi.org/10.1117/12.2522969>
- LAG VSW, L. der V., 2015. [LAG VSW] Länderarbeitsgemeinschaft der Vogelschutzwarten | Windenergie [WWW Document]. URL <http://www.vogelschutzwarten.de/windenergie.htm> (accessed 6.30.22).
- Landesanstalt für Umwelt Baden-Württemberg, L., 2022. Erfassung 2020 und Bewertung 2015 Vögel - Artenschutz und Windkraft [WWW Document]. Artenschutz Windkraft. URL https://www.lubw.baden-wuerttemberg.de/natur-und-landschaft/artenschutz-und-windkraft/-/document_library_display/bFsX3wOA3G54/view/886319 (accessed 6.28.22).
- Li, Y., Wu, C.-Y., Fan, H., Mangalam, K., Xiong, B., Malik, J., Feichtenhofer, C., 2022. MViTv2: Improved Multiscale Vision Transformers for Classification and Detection. Presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4804–4814.
- Maaten, L. van der, Hinton, G., 2008. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2605.
- Moon, J., Hossain, M.B., Chon, K.H., 2021. AR and ARMA model order selection for time-series modeling with ImageNet classification. *Signal Process.* 183, 108026. <https://doi.org/10.1016/j.sigpro.2021.108026>
- Peyrolón, P., 2020. Der Satz von Bayes: Wahrscheinlichkeitstheorie für Finanzen und Betriebswirtschaft. Springer-Verlag.
- Rajagopal, A., Ramachandran, A., Shankar, K., Khari, M., Jha, S., Lee, Y., Joshi, G.P., 2020. Fine-Tuned Residual Network-Based Features With Latent Variable Support Vector Machine-Based Optimal Scene Classification Model for Unmanned Aerial Vehicles. *IEEE Access* 8, 118396–118404. <https://doi.org/10.1109/ACCESS.2020.3004233>
- Selvaraju, R.R., Das, A., Vedantam, R., Cogswell, M., Parikh, D., Batra, D., 2017. Grad-CAM: Why did you say that? (No. arXiv:1611.07450). arXiv. <https://doi.org/10.48550/arXiv.1611.07450>
- Shrestha, A., Mahmood, A., 2019. Review of Deep Learning Algorithms and Architectures. *IEEE Access* 7, 53040–53065. <https://doi.org/10.1109/ACCESS.2019.2912200>

- Tustison, N.J., Avants, B.B., Lin, Z., Feng, X., Cullen, N., Mata, J.F., Flors, L., Gee, J.C., Altes, T.A., Mugler, I., John P., Qing, K., 2019. Convolutional Neural Networks with Template-Based Data Augmentation for Functional Lung Image Quantification. *Acad. Radiol.*, Special Issue: Pulmonary Functional Imaging 26, 412–423. <https://doi.org/10.1016/j.acra.2018.08.003>
- Umweltministerkonferenz, U., 2020. Ländereinigung zum Artenschutz [WWW Document]. UMK Beschließt Einheitliche Stand. Zur Ermittl. Signifikant Erhöhten Tötungsrisikos. URL <https://www.fachagentur-windenergie.de/aktuelles/detail/laendereinigung-zum-artenschutz/> (accessed 6.30.22).
- Van Der Maaten, L., 2014. Accelerating t-SNE using tree-based algorithms. *J. Mach. Learn. Res.* 15, 3221–3245.

Die „BfN-Schriften“ sind eine seit 1998 unperiodisch erscheinende Schriftenreihe in der institutionellen Herausgeberschaft des Bundesamtes für Naturschutz (BfN) in Bonn. Sie sind kurzfristig erstellbar und enthalten u.a. Abschlussberichte von Forschungsvorhaben, Workshop- und Tagungsberichte, Arbeitspapiere oder Bibliographien. Viele der BfN-Schriften sind digital verfügbar. Printausgaben sind auch in kleiner Auflage möglich.

DOI 10.19217/skr651