deepRain

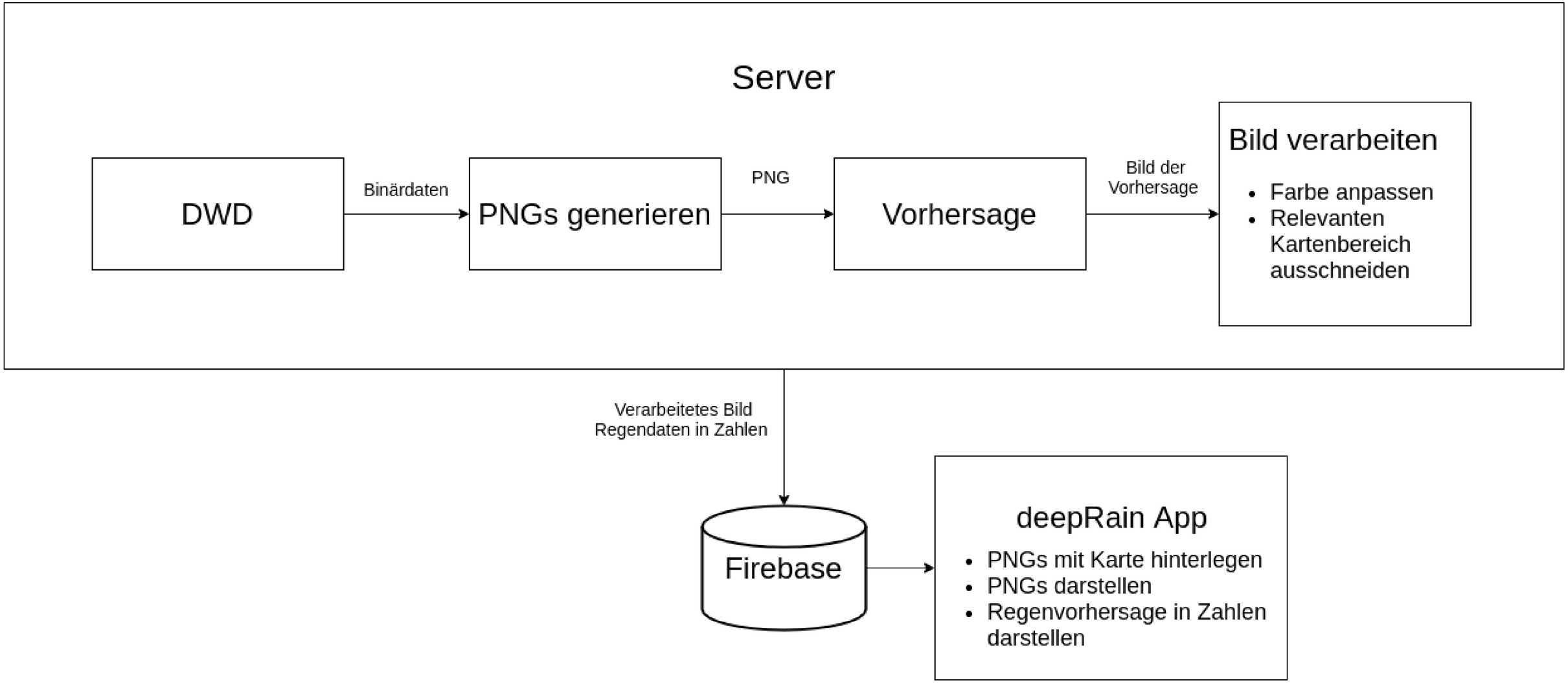
Simon Christofzik. Paul Sutter. Till Reitlinger.

Agenda

1. Gesamtsystem
2. Datenbeschaffung
3. Datenvalidierung
4. Die Netze
5. Validierung und Bewertung der Netze
6. Die deepRain App
7. Die Pipeline
8. Ausblick

d e e p R a i n

Gesamtsystem



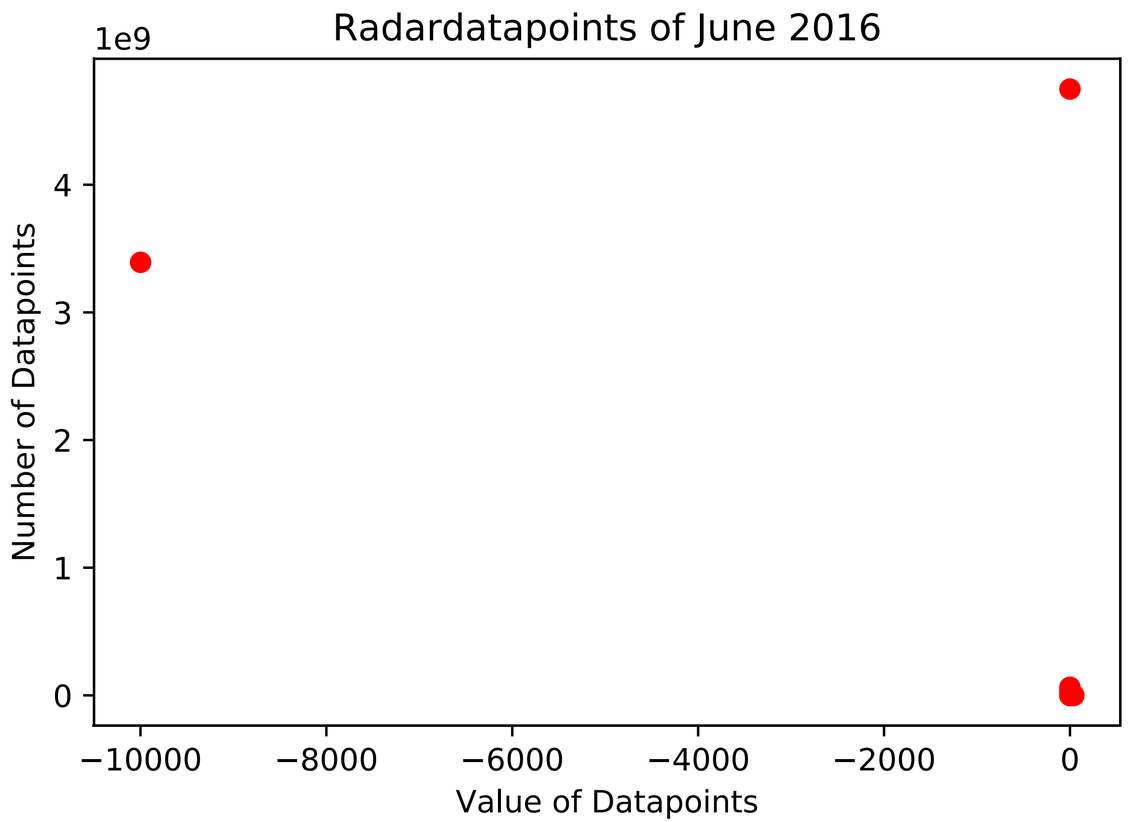
Datenbeschaffung

Beschaffung der Radardaten und erste Vorverarbeitung

Datenbeschaffung

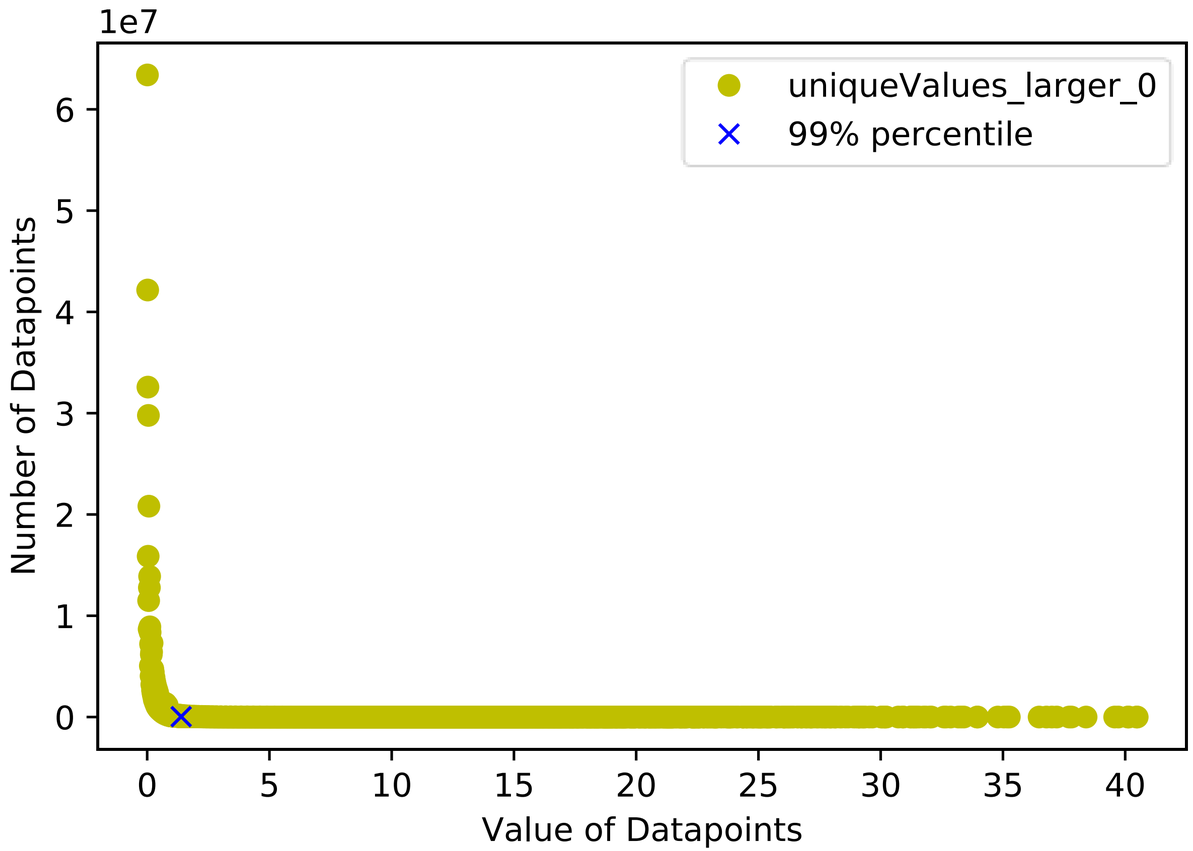
- Deutscher Wetterdienst stellt Radardaten in fünf minütiger Auflösung bereit
- Für das Training der Netzte müssen diese in Bilder umgewandelt werden
- Transformation von Radardatenpunkten zu Pixelfarbwerten
- Inspizierung der Daten

Explorative Datenanalyse

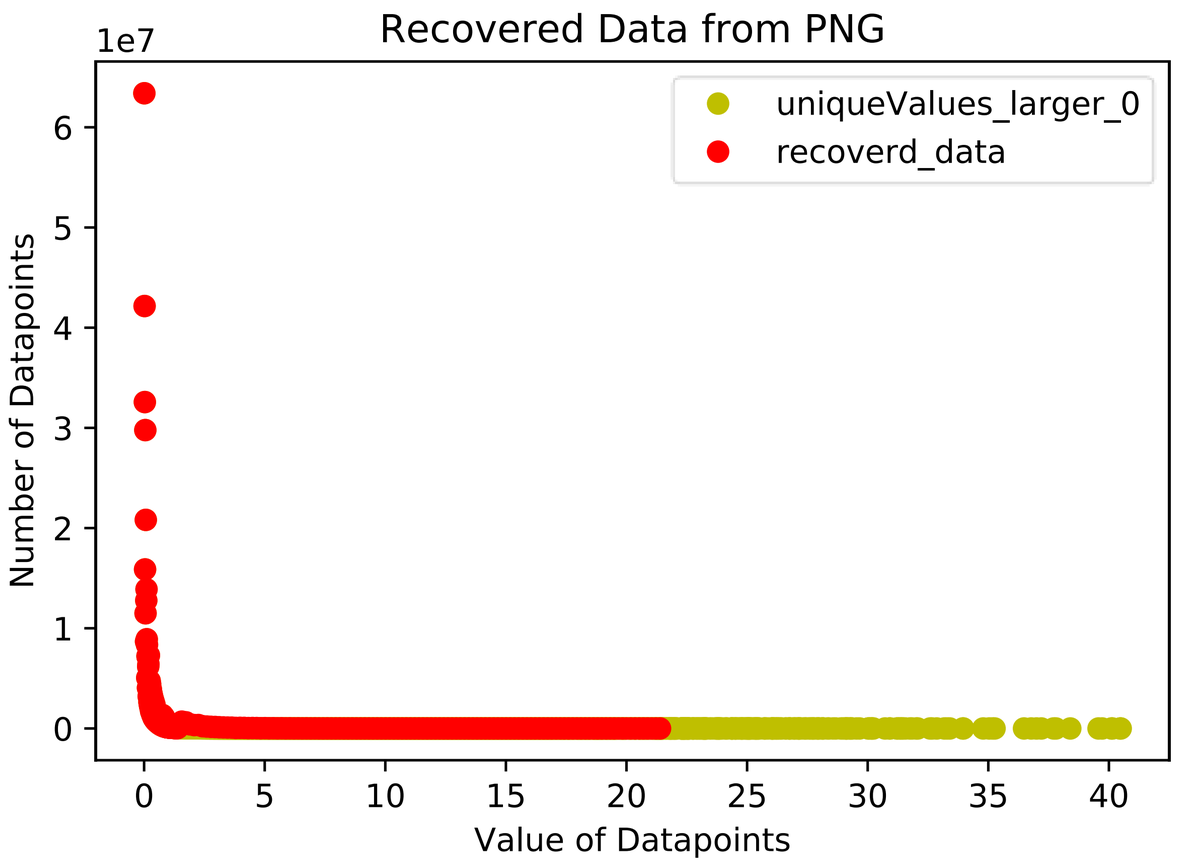


- Exemplarisch werden die Daten von Juni 2016 genauer untersucht
- -9999 = kein Daten verfügbar
- 0 = kein Regen

Datentransformation



- **Entwicklung einer Methode**
- **Ziel ist, möglichst wenig Datenpunkte zu verlieren**
- **99% Perzentil beinhaltet 138 verschiedene Werte**
- **Verbleibender Wertebereich: 117**



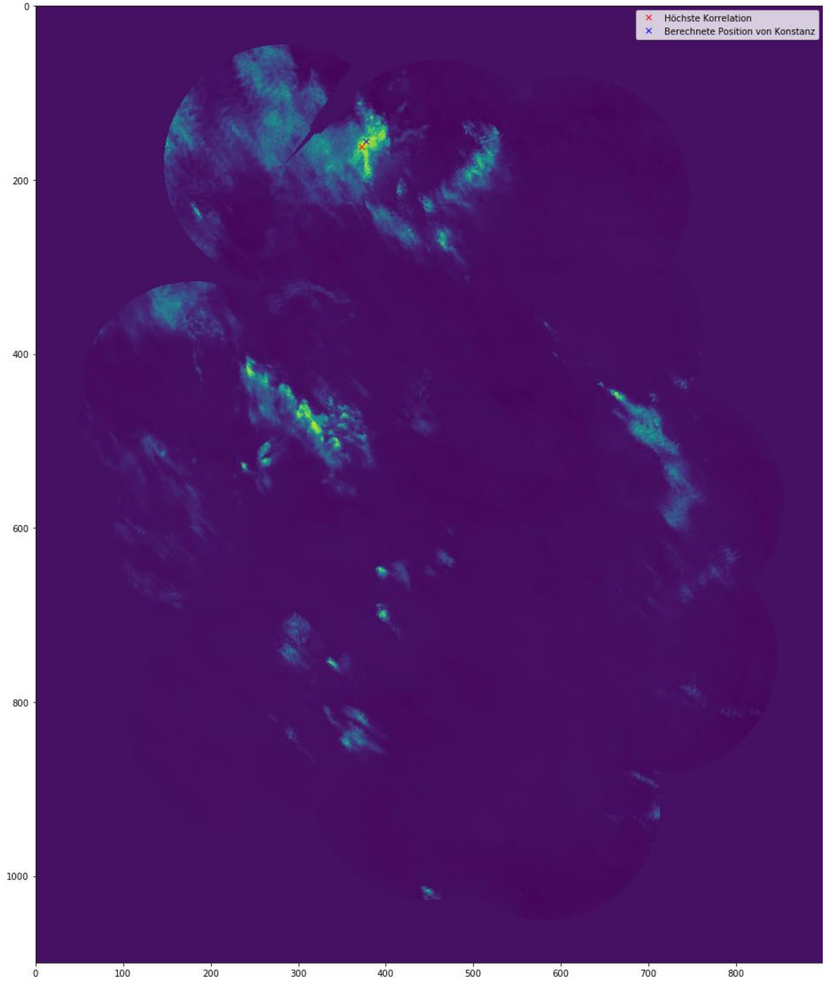
Datenvorverarbeitung

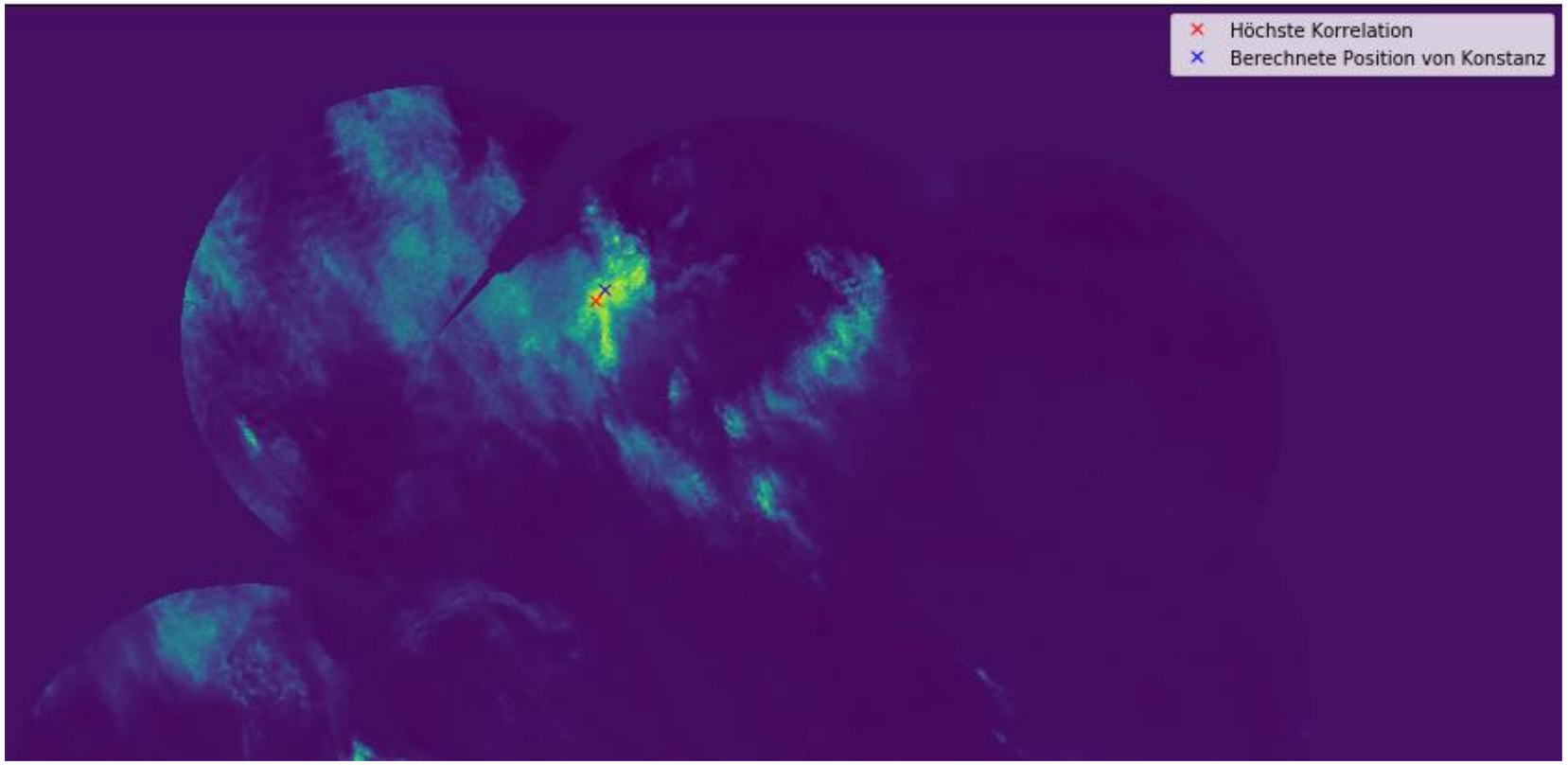
- **Es gehen keine Regendaten verloren**
- **Fast 100% der Datenpunkte können wieder hergestellt werden**

Datenqualität

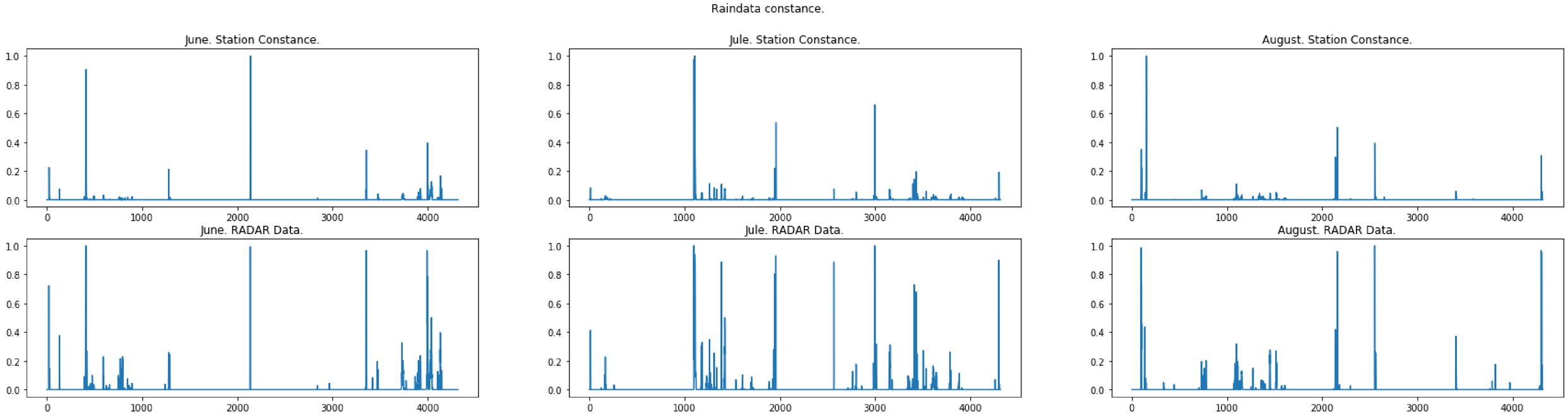
Validierung und überprüfung der Datenqualität

Datenvalidierung





Datenvalidierung



Datenqualität

- Es gibt Ausreißer
- Qualität ausreichend um Netzte zu trainieren
- Regenstärken weichen systematisch leicht ab
- Regenzeitpunkte stimmen gut überein

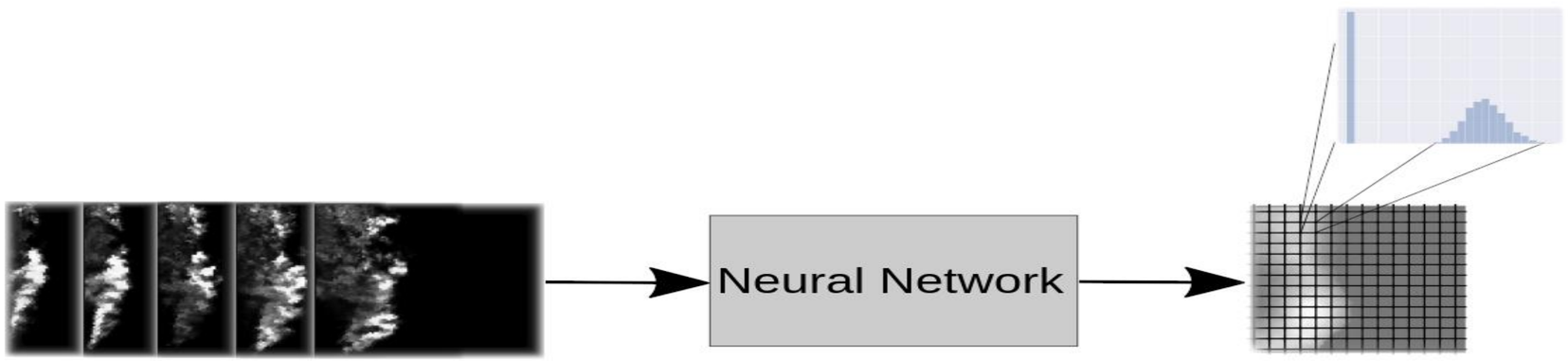
Die Netze

Vorgehensweise

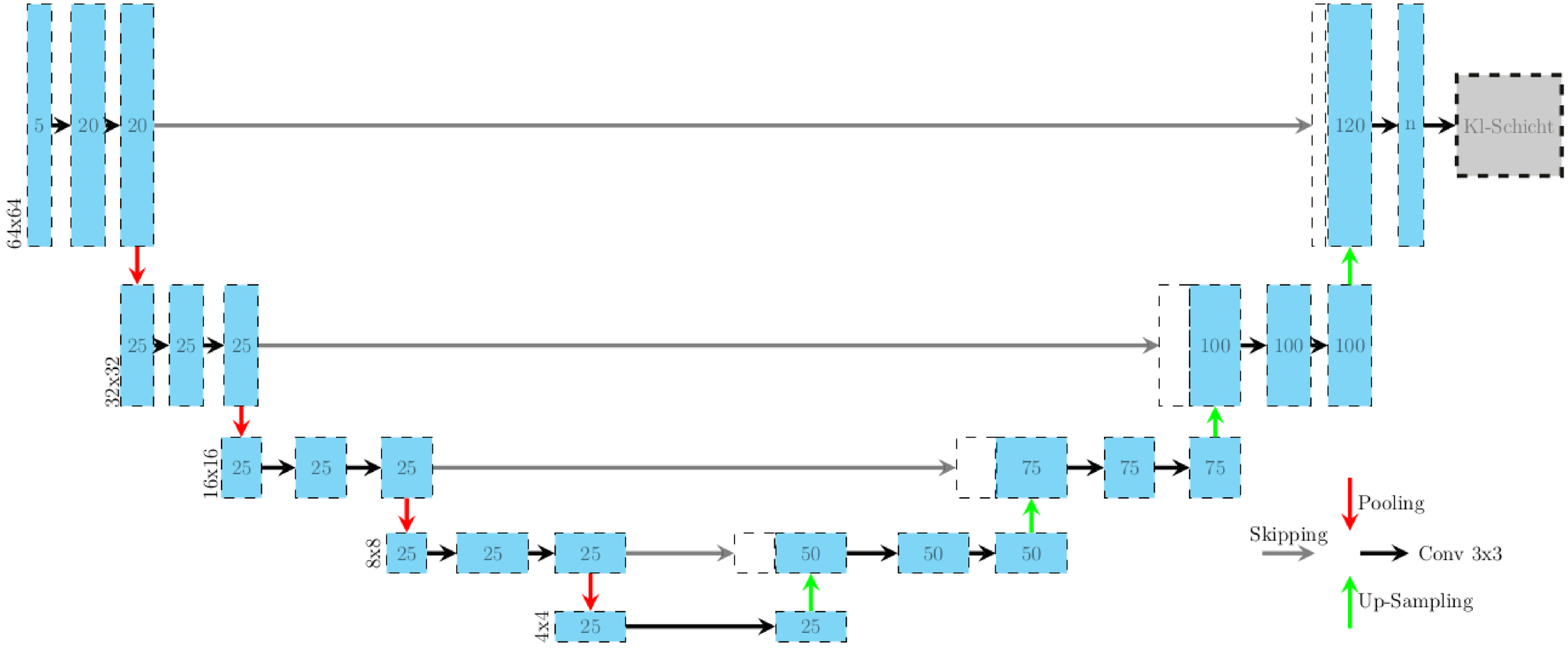
"It is scientific only to say what is more likely and what less likely" - Richard P. Feyman

- Ausgang des Netzwerkes soll eine Wahrscheinlichkeitsverteilung sein
  - Multinomiale Verteilung
  - Bernoulliverteilung kombiniert mit der Negativ Binomialverteilung

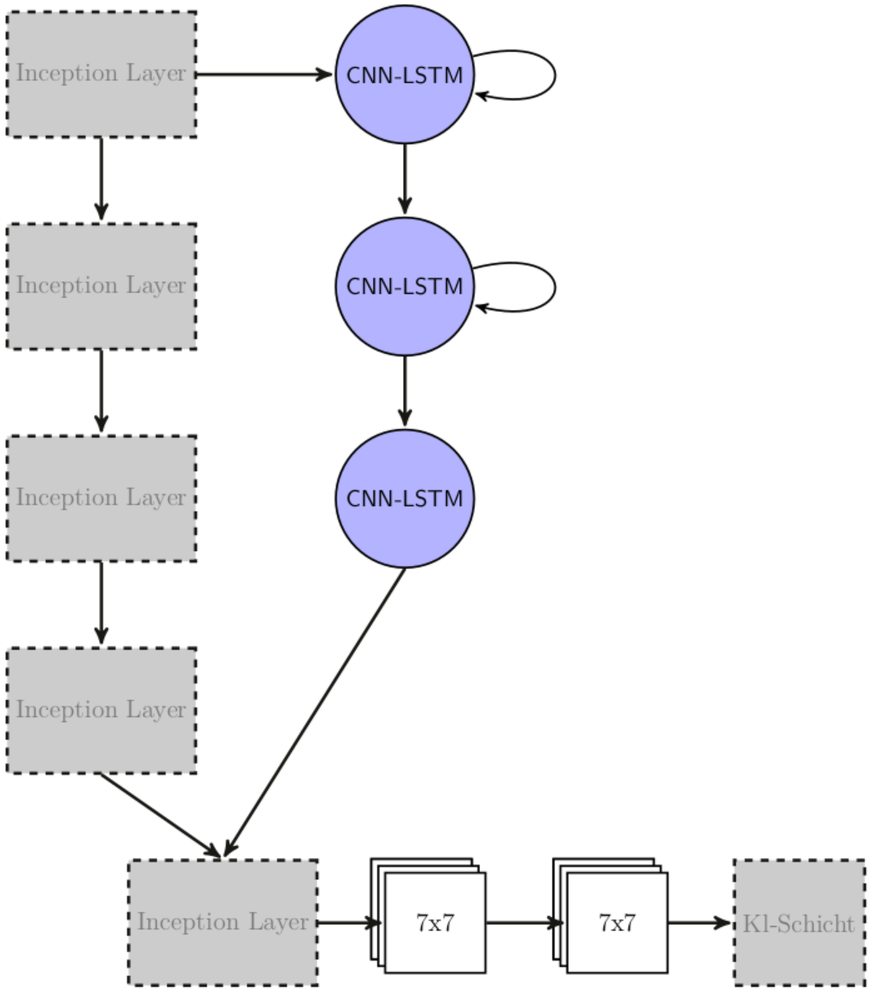
- Input: Bilderserie der Größe 96 x 96
- Output: Matrix von Verteilungen der Größe 64 x 64



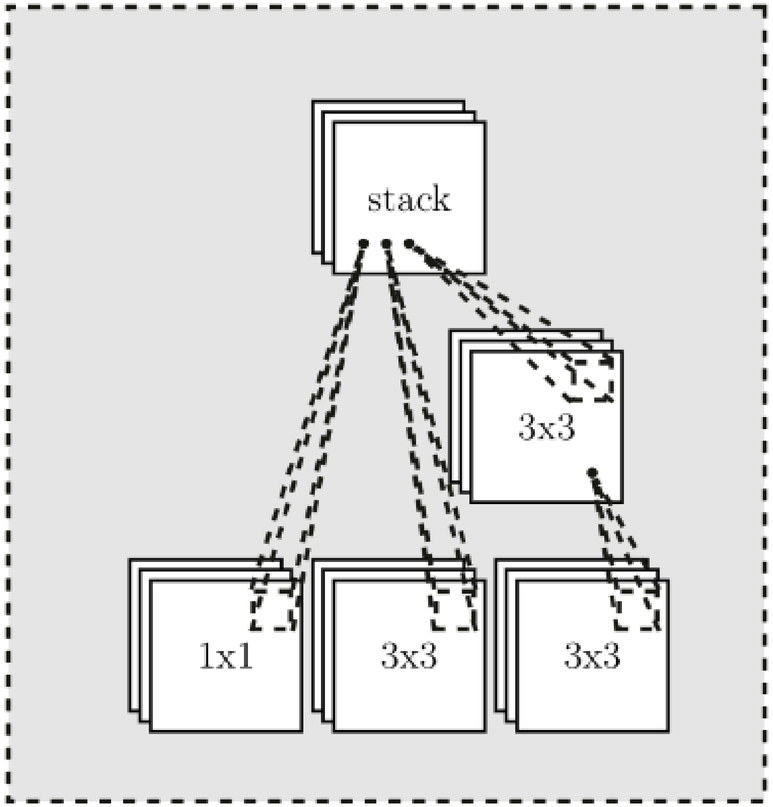
Netzwerktopologie



Netzwerktopologie

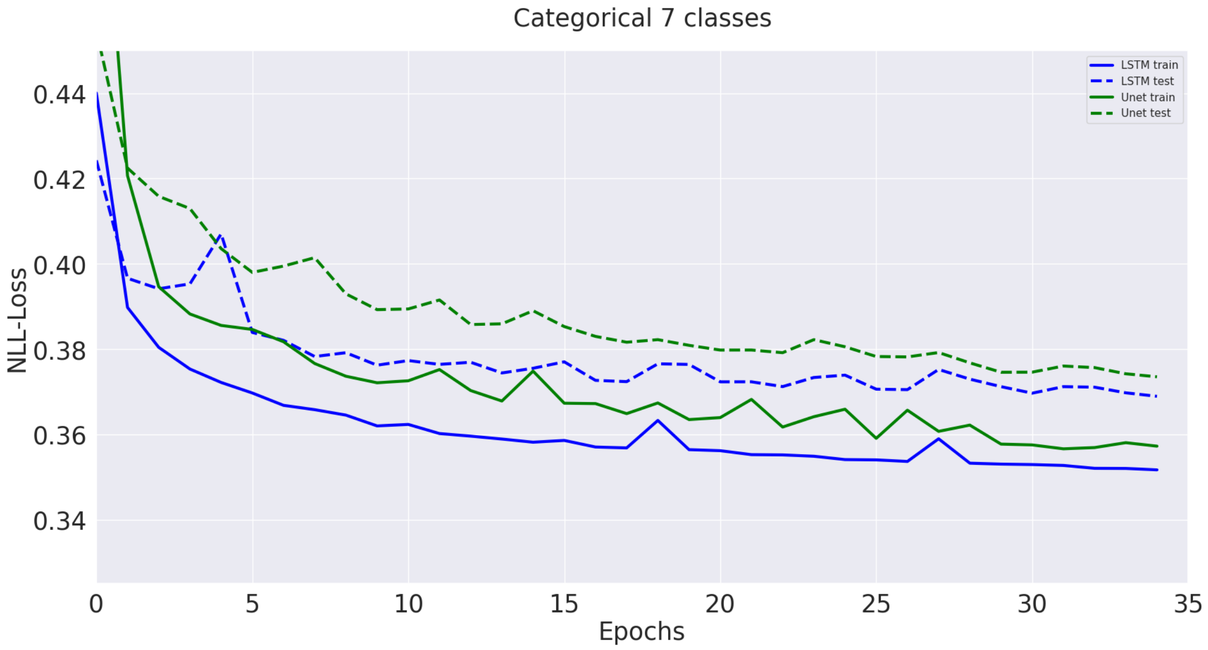


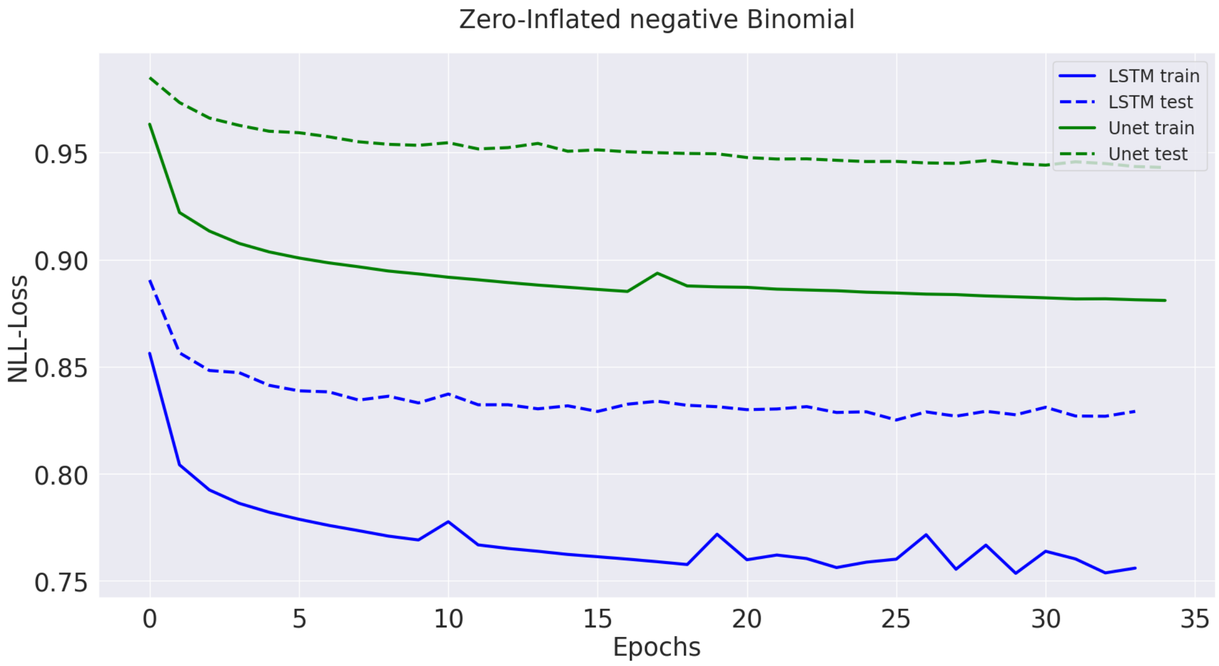
- Inception Layer



Training

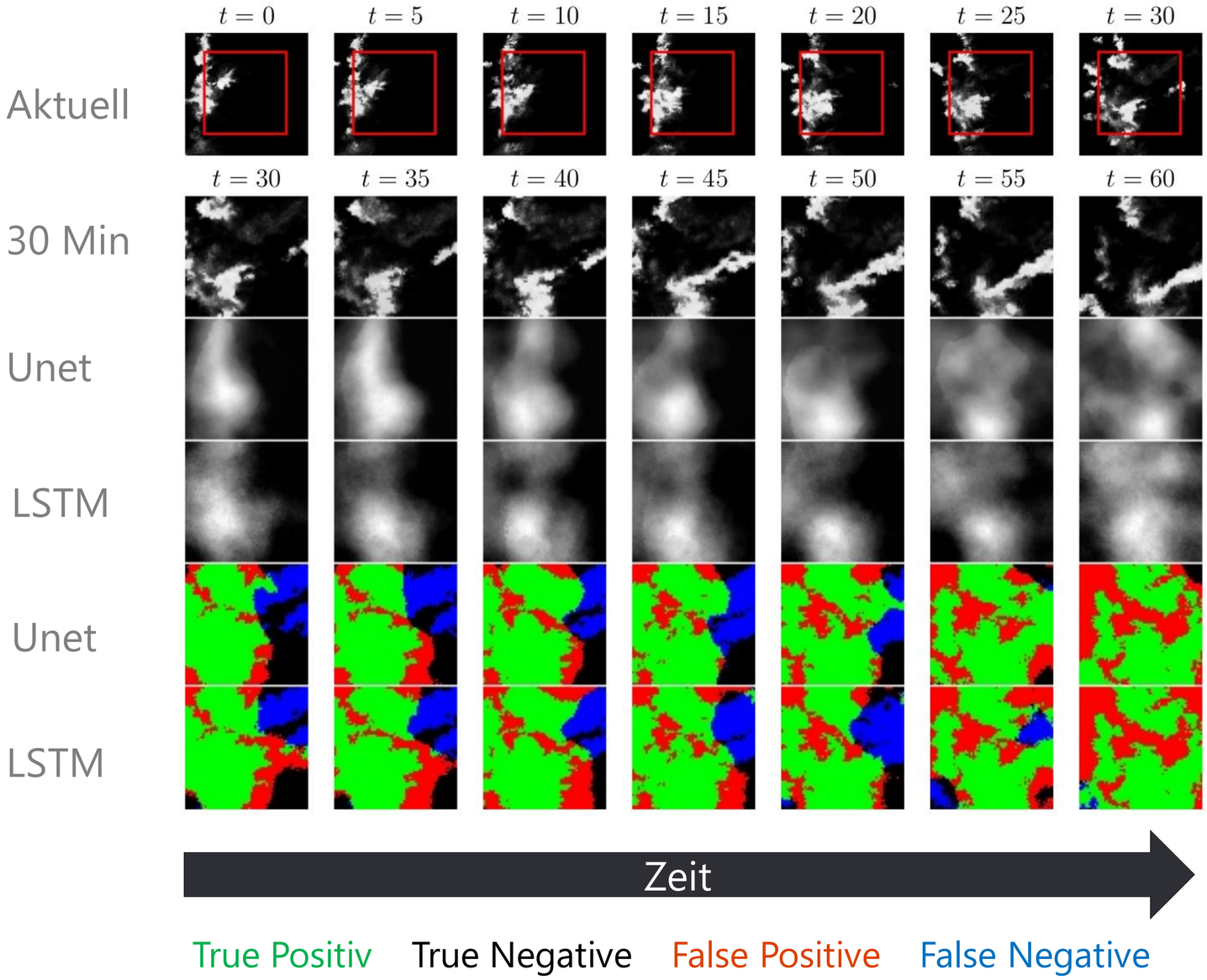
- Daten über 10 Jahre
- Ca. 1 051 200 Daten
- 80 % der Daten entfallen durch aussortieren
- 25% der Daten als Testdaten ~ca 2 Jahre





Vorhersage

- Negativ Binomial
- Pixelwerte 0 - 255
- Multinomialverteilung als 7-Klassenproblem
- Wahl des Mittelwertes für die Vorhersage

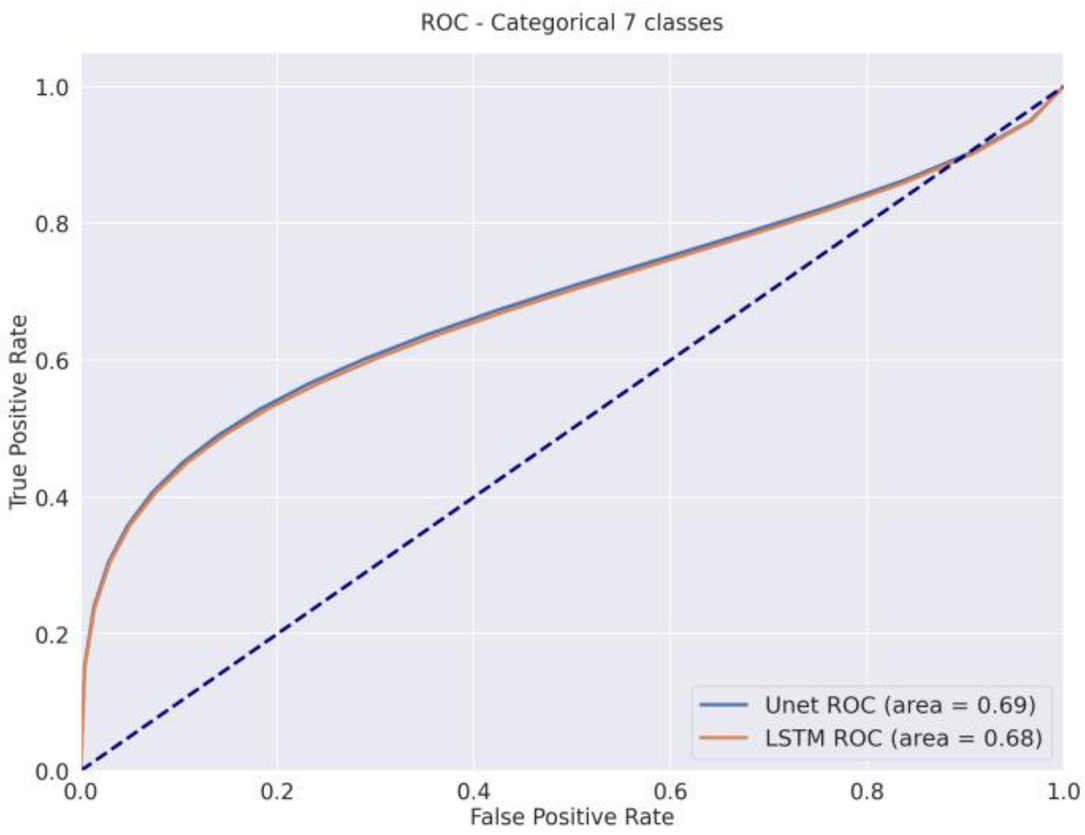


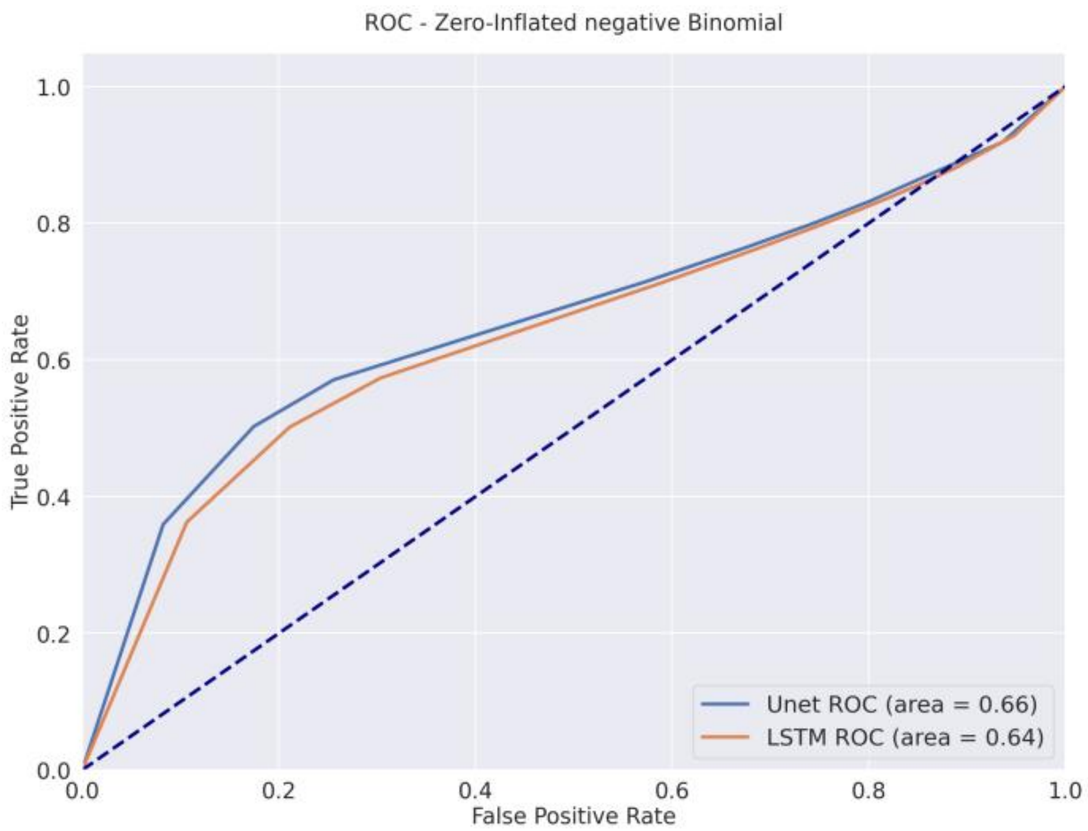
Validierung und Bewertung der Netze

Vergleich der Netze und deren Qualität

Binäre Regenvorhersage

- ROC erstellt mit 20 verschiedenen Schwellwerten





Binäre Regenvorhersage

- Einfache Baseline:
  Wetter verändert sich nicht
- Confusionmatrix für Schwellwert 0.5
- Besser als die Baseline!

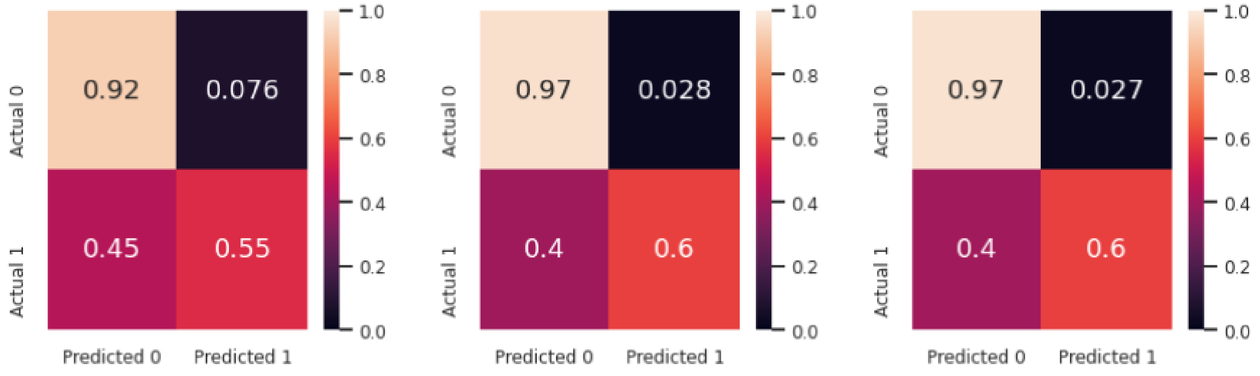


Abbildung 26: Confusionmatrix für die Multinomialverteilung und dem Schwellwert 0.5. Links die einfache Baseline, mittig für die LSTM-Architektur und rechts für die Unet-Architektur.

Verteilung	Person wird Nass	Person wird nicht Nass
Multinomialeverteilung	**39,9** %	**60,1** %
ZINB	43,0 %	57,0 %
Einfache Baseline	44,8 %	55,2 %

Tabelle 1: Wahrscheinlichkeit trocken zu bleiben

Regenklassifikation

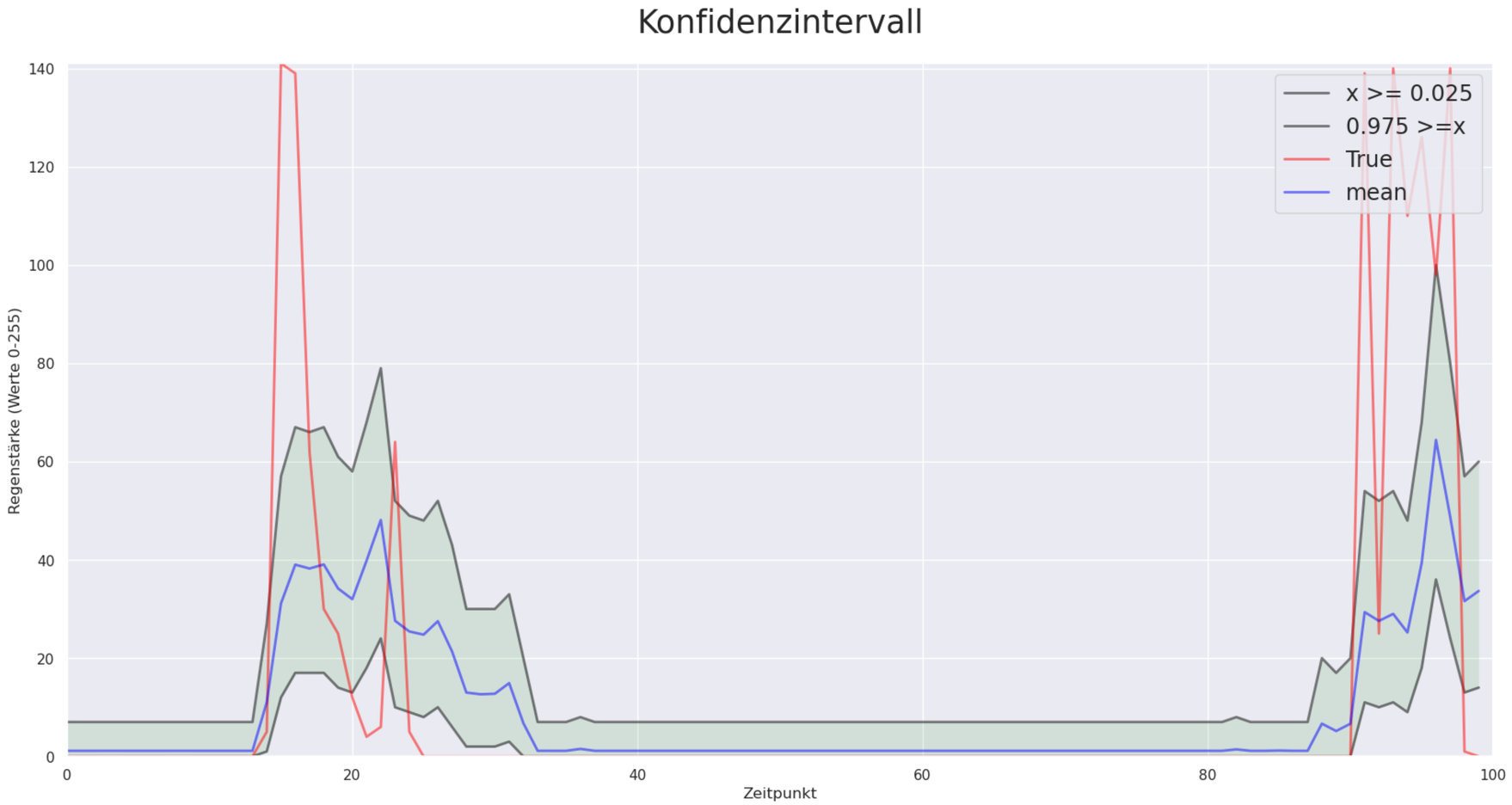
- ZINB Umwandlung von diskret in 7 Klassen
- Wie oft wird das korrekte Label vorhergesagt?
- ZINB ist falls es Regnet (etwas) besser!
- Multinomial ist für binäre Regenvorhersage besser!

	LSTM		Unet	
Klasse	ZNIB	Multi	ZNIB	Multi
0	97 %	97 %	**98** %	97 %
1	**10** %	0 %	**10** %	0 %
2	**42** %	21 %	41 %	23 %
3	30 %	33 %	**34** %	**34** %
4	5 %	**21** %	7 %	20 %
5	0 %	**8** %	0 %	6 %
6	0 %	0 %	0 %	0 %

Regenintensität

Motivation

- Wie gut ist der Mittelwert als Schätzung?

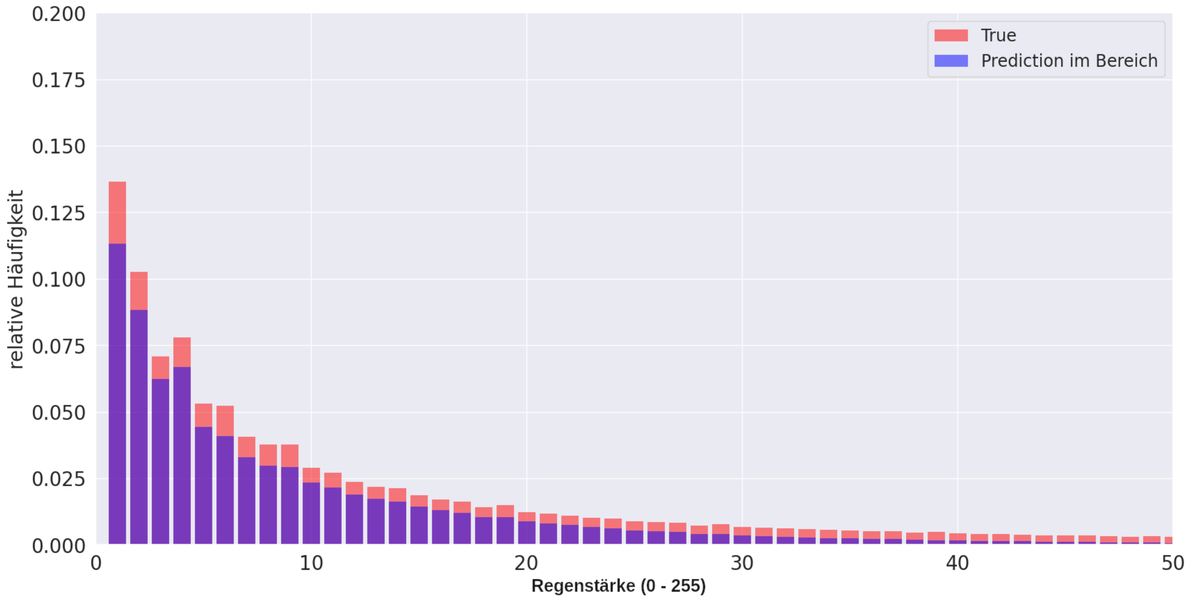


Regenintensität

Negativ Binomialverteilung

- Wie oft liegt der "wahre" Wert im Konfidenzintervall?
- Schlechte Vorhersage für starker Regen

Histogramm für das Unet



Architektur	inklusive kein Regen	exklusive kein Regen	Konfidenzintervall
UNET	**97,0** %	**73,0** %	97,5 %
	94,0 %	43,0 %	95 %
	90,0 %	35,0 %	90 %
LSTM	96,0 %	71,0 %	97,5 %
	95,0 %	**45,0** %	95 %
	93,0 %	**40,0** %	90 %

Tabelle 2: Relative Anzahl der wahren Werte im Konfidenzintervall für verschiedene Intervalle und Architekturen

- Je kleiner das Intervall, desto seltener ist der "wahre" Wert enthalten!

Fazit

Positiv

- Binäre Vorhersage ist besser als die Baseline!
- Wahrscheinlichkeitsverteilung für Regen
- Vorhersage von mehreren Regenstärken

Negativ

- Schwache Struktur in der Vorhersage der Regenfront
- Kein deutlich besseres Netzwerk
- Starker Regen wird schlecht Vorhergesagt

Die deepRain App

Die Visualisierung der Vorhersagen

Die App

DeepRain Funktionen

Regenwarnung

Der User soll vor Regen gewarnt werden. Es soll eingestellt werden können, wie früh die Regenwarnung kommt.

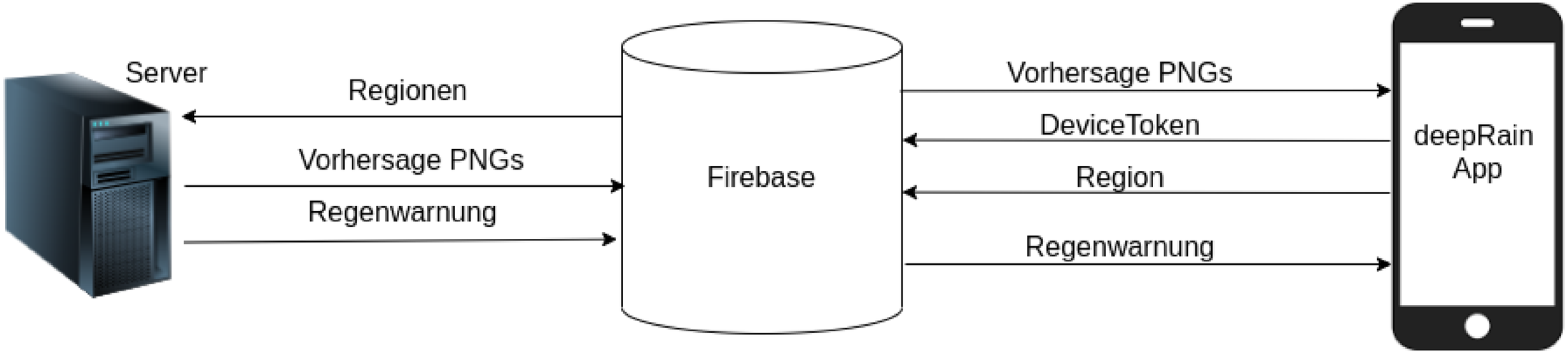
Regenvorhersage in Zahlen

Die aktuellste berechnete Regenvorhersage soll immer in Zahlen zur Verfügung stehen.

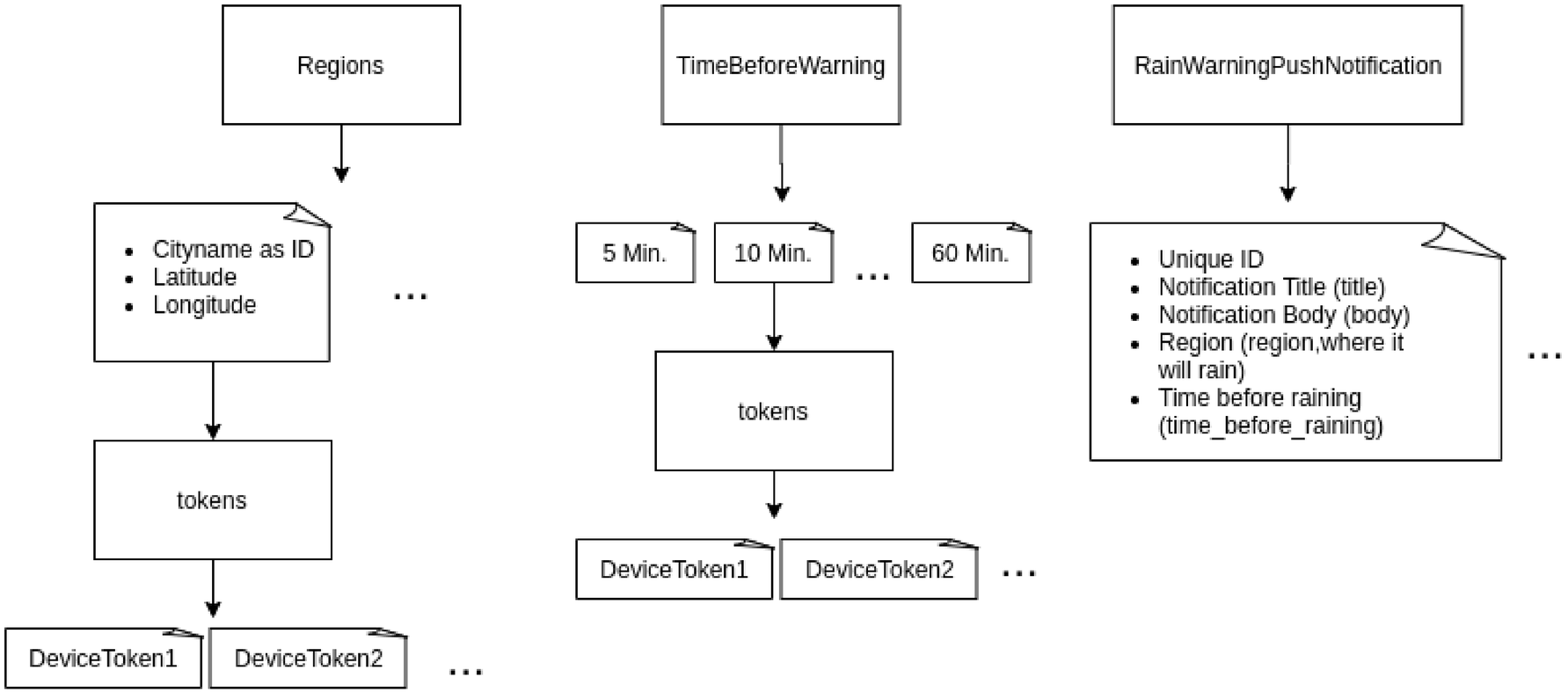
Regenvorhersage Karte

Die aktuellste berechnete Regenvorhersage soll immer als Karte zur Verfügung stehen.

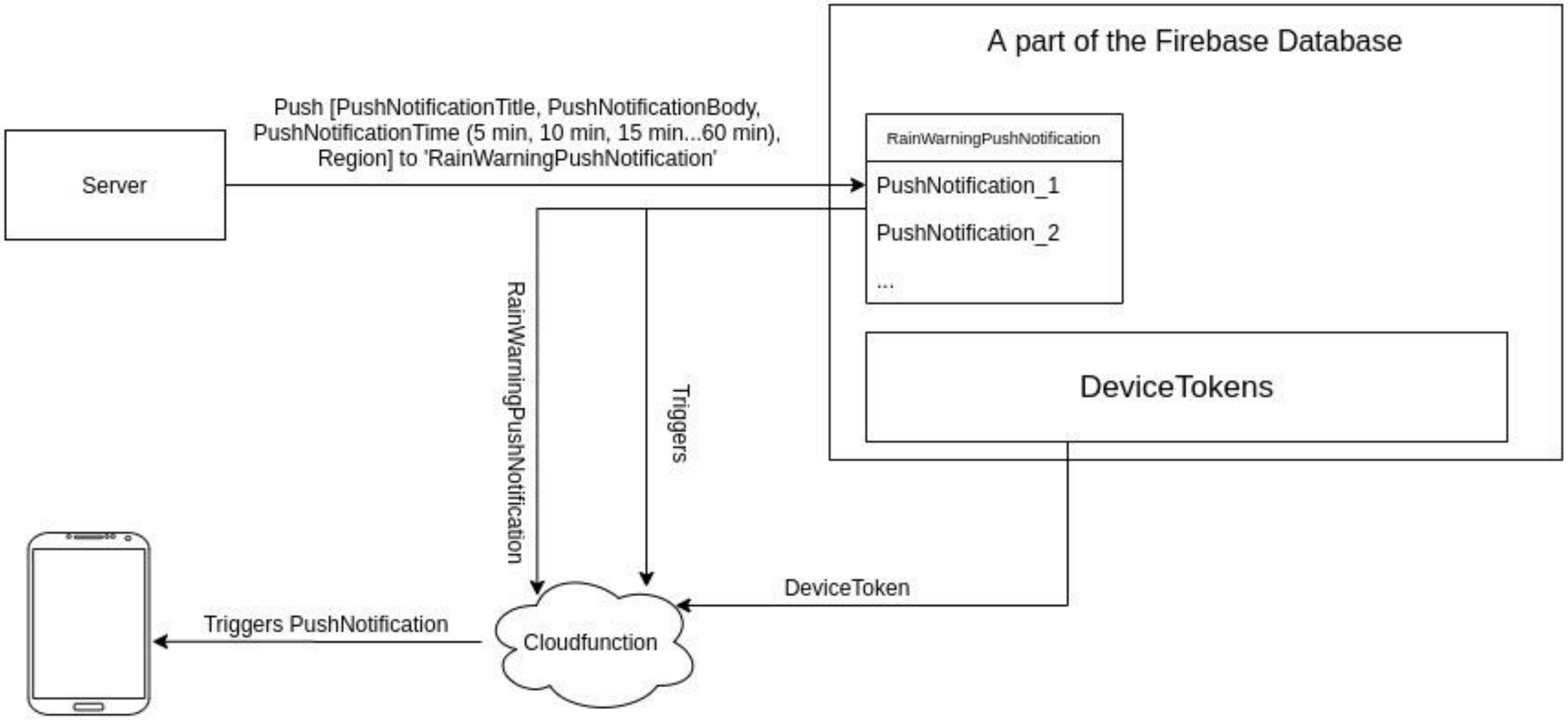
Komponenten Überblick



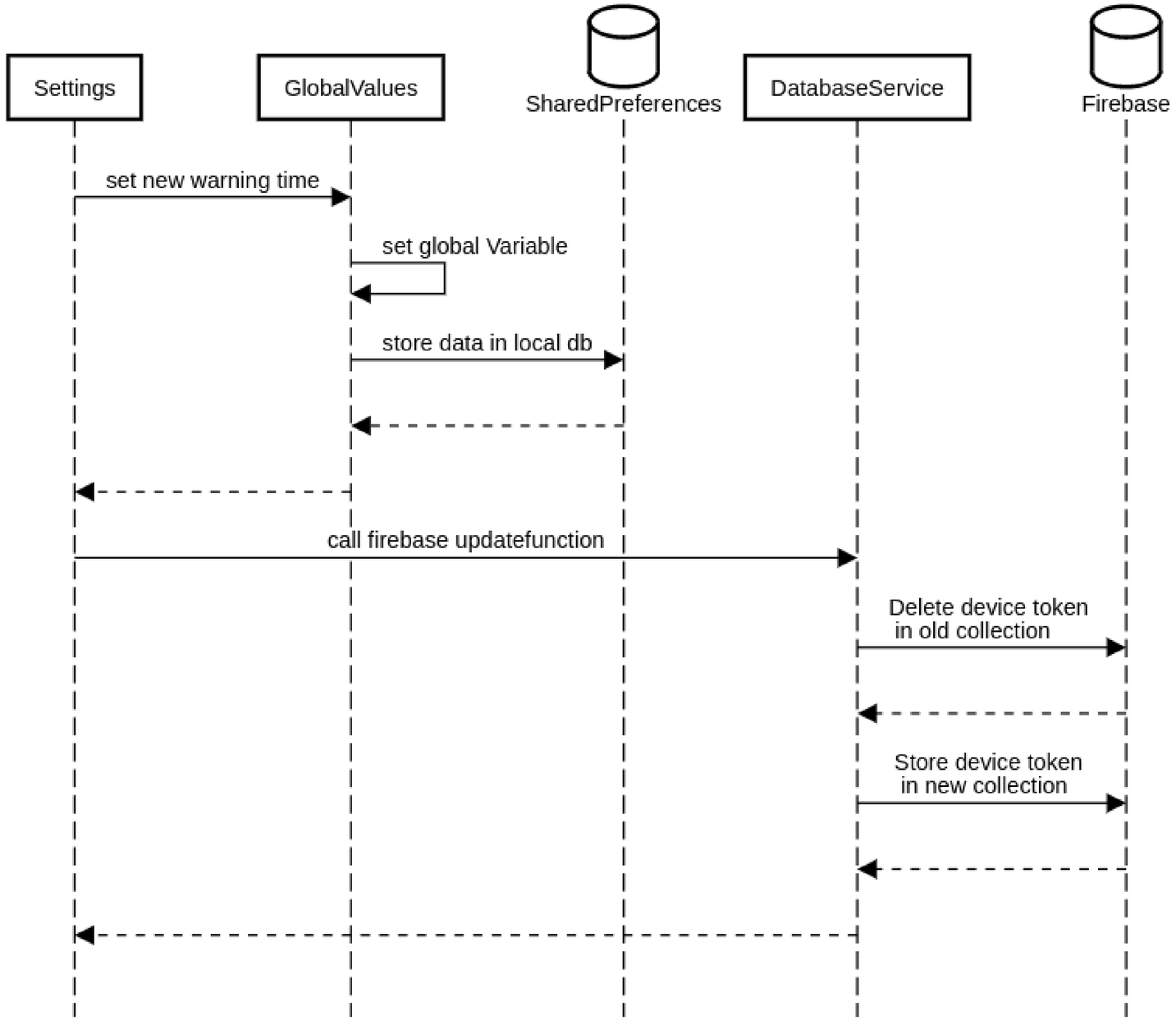
Datenbankaufbau

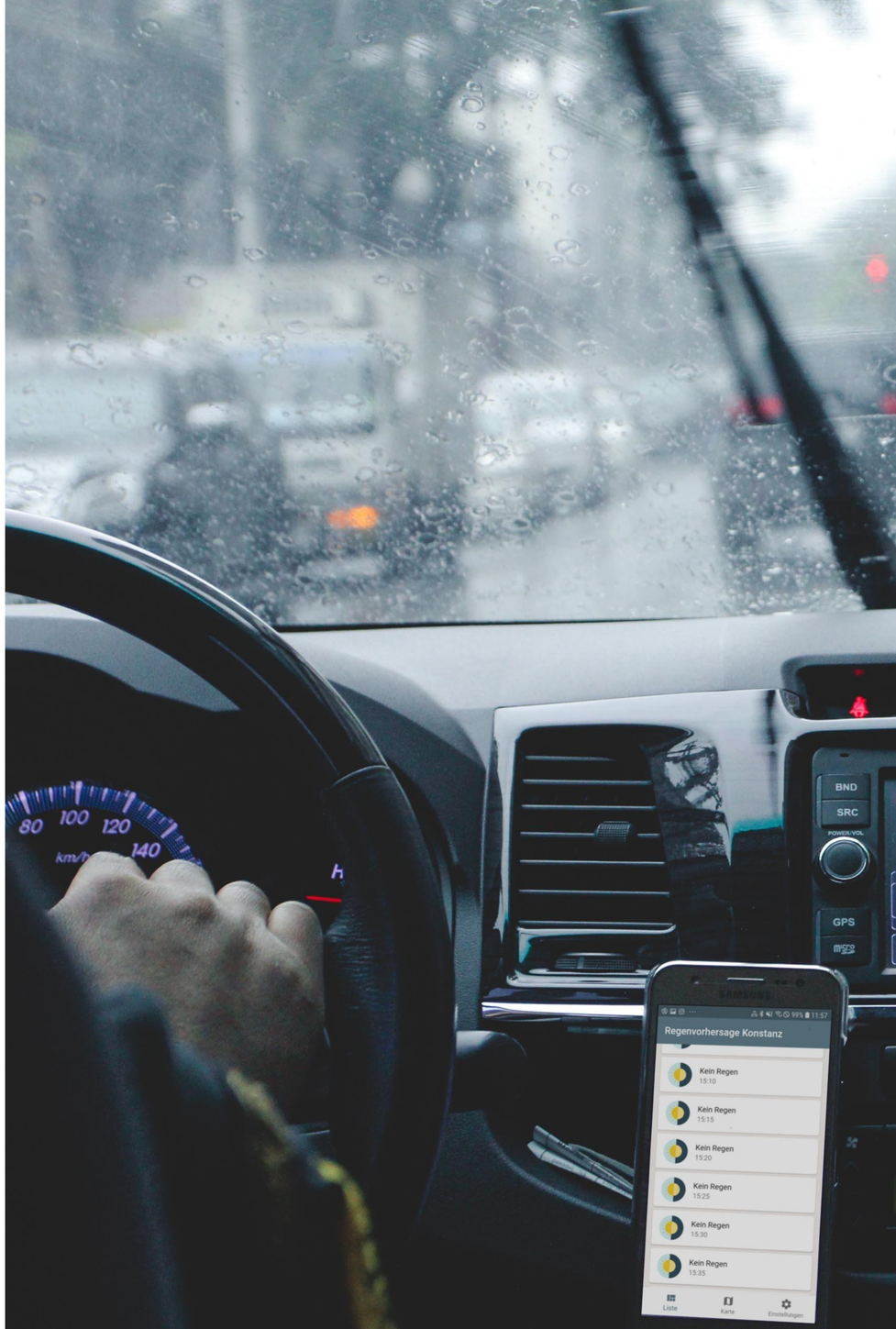


Funktionsweise der Regenwarnung



Beispiel Datenfluss





Berechnung der Regenintenistät

- Listen mit allen Pixel Koordinaten als Input
- Berechnung des Pixels in dem Bild mit eigenem Algorithmus
- Bei jedem Appstart berechnung des Farbwertes in diesem Pixel für alle Vorhersagen
- Kommunikation dieser Daten innerhalb der App mit Stream

Pipeline

Vom DWD bis zur App

Pipeline

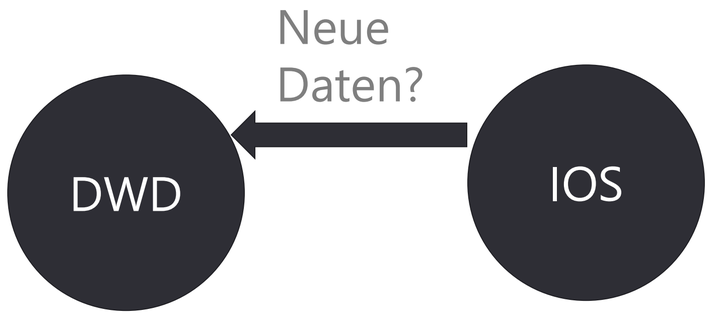
DWD

Pipeline

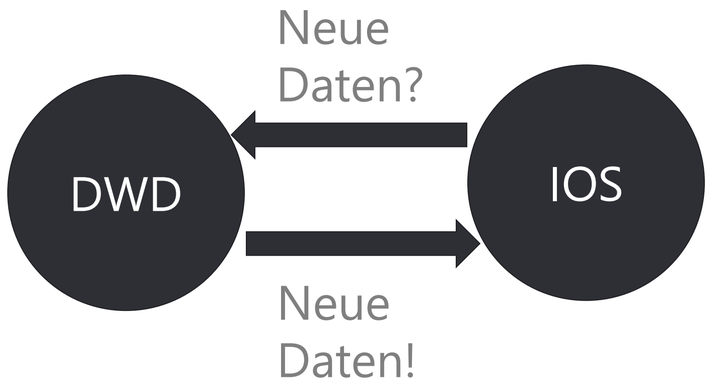
DWD

IOS

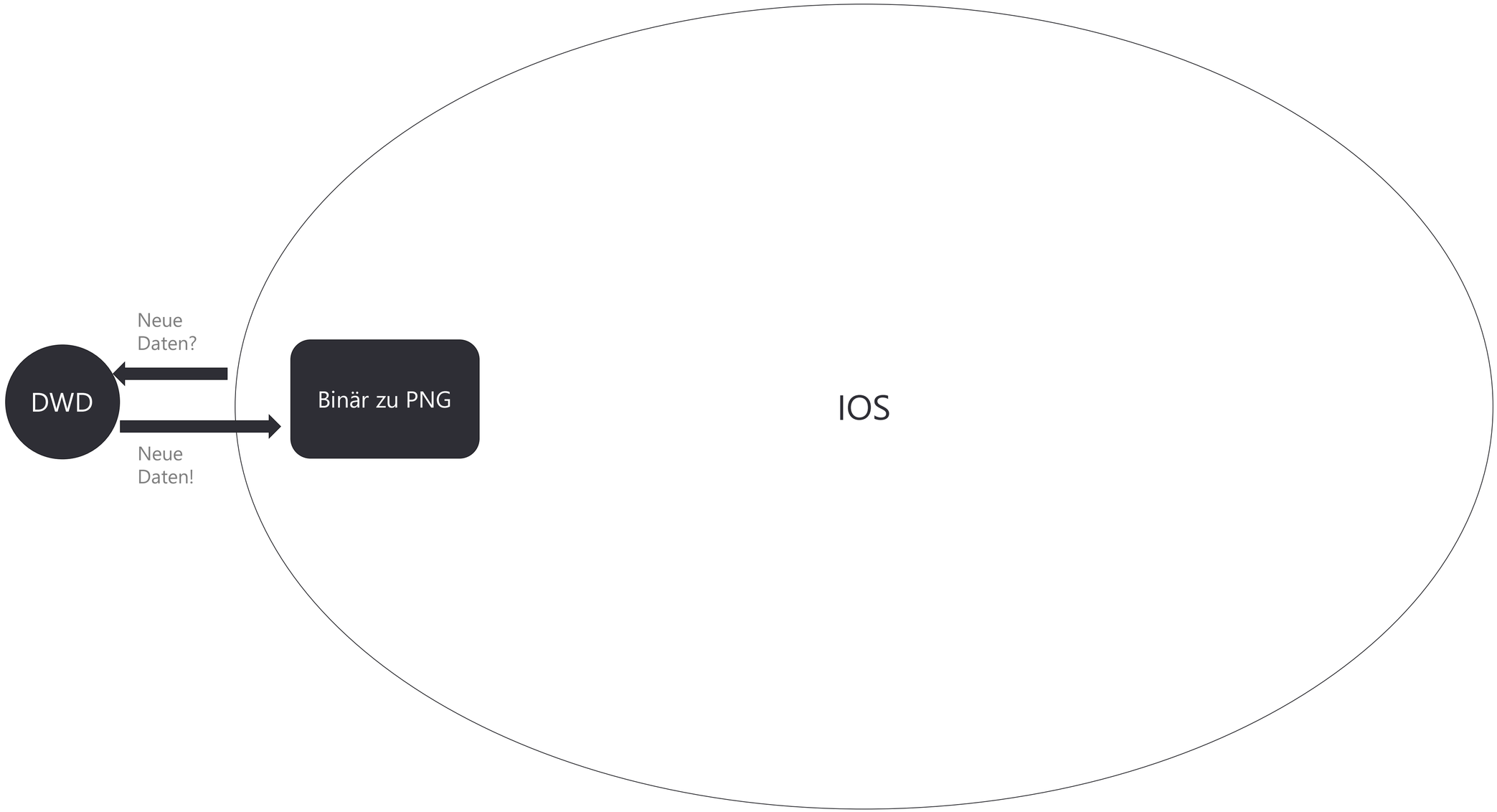
Pipeline



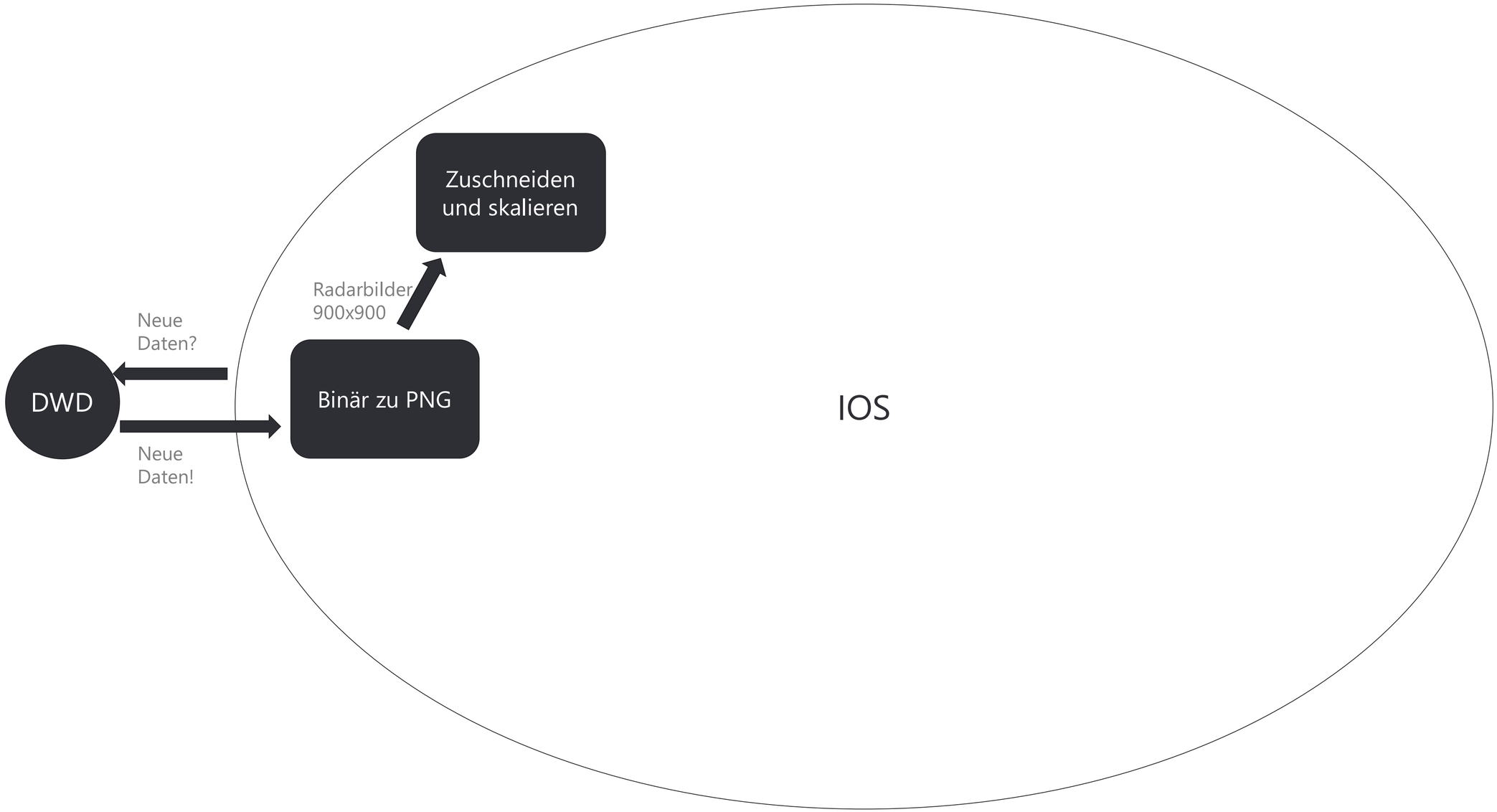
Pipeline



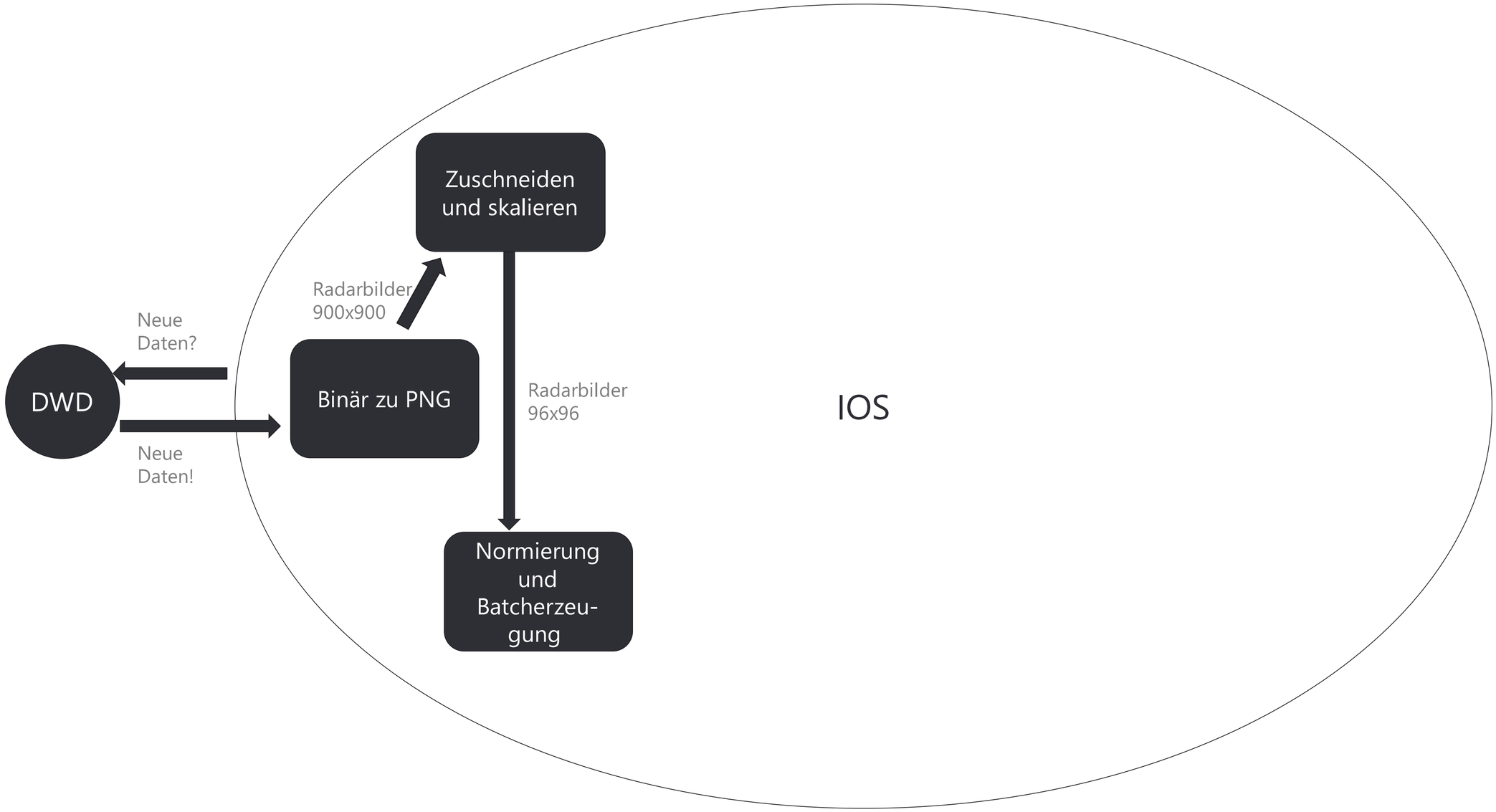
Pipeline



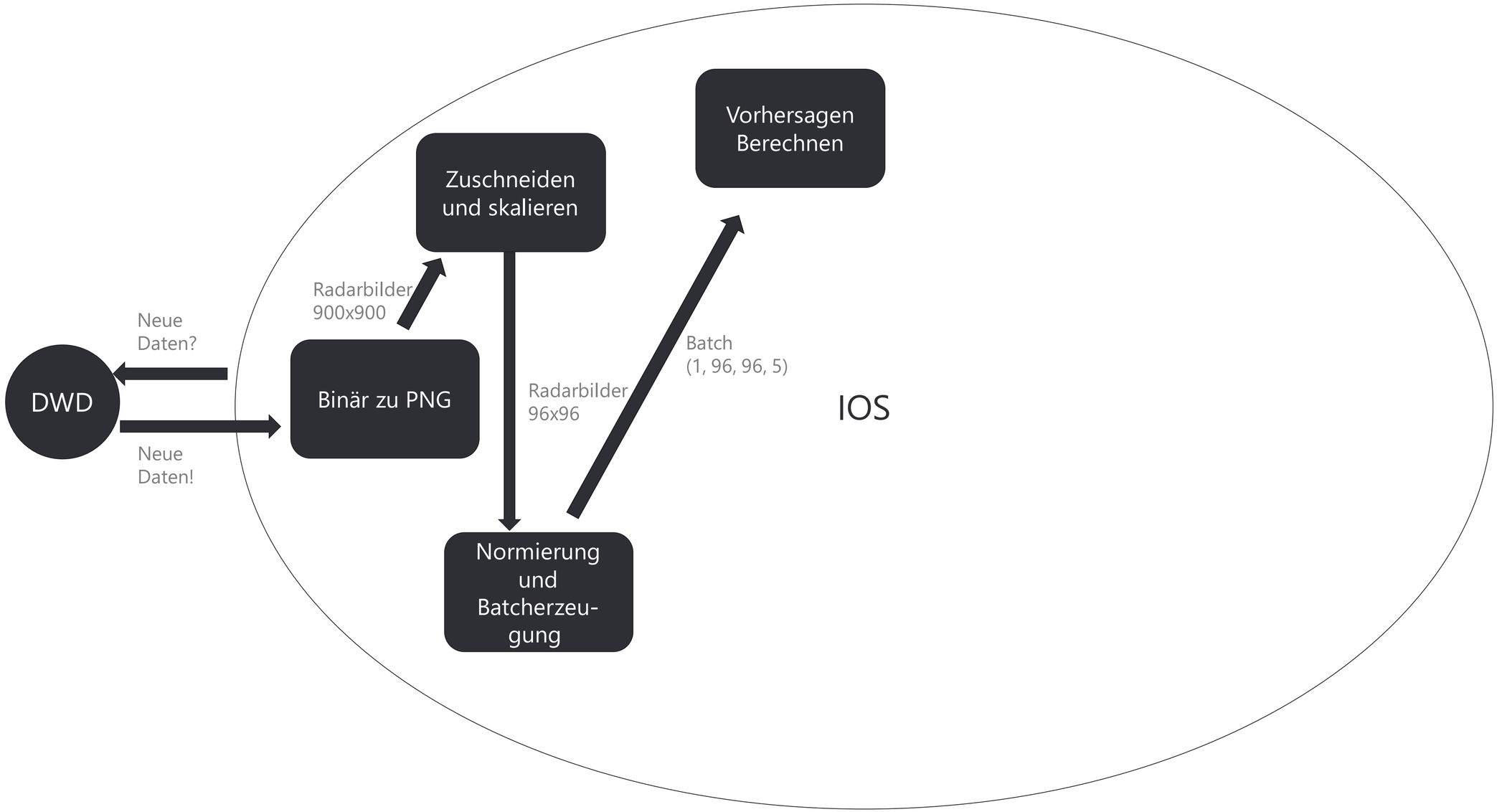
Pipeline


Zuschneiden
und skalieren
Radarbilder
900x900
Neue
Daten?
DWD
Binär zu PNG
Neue
Daten!
IOS

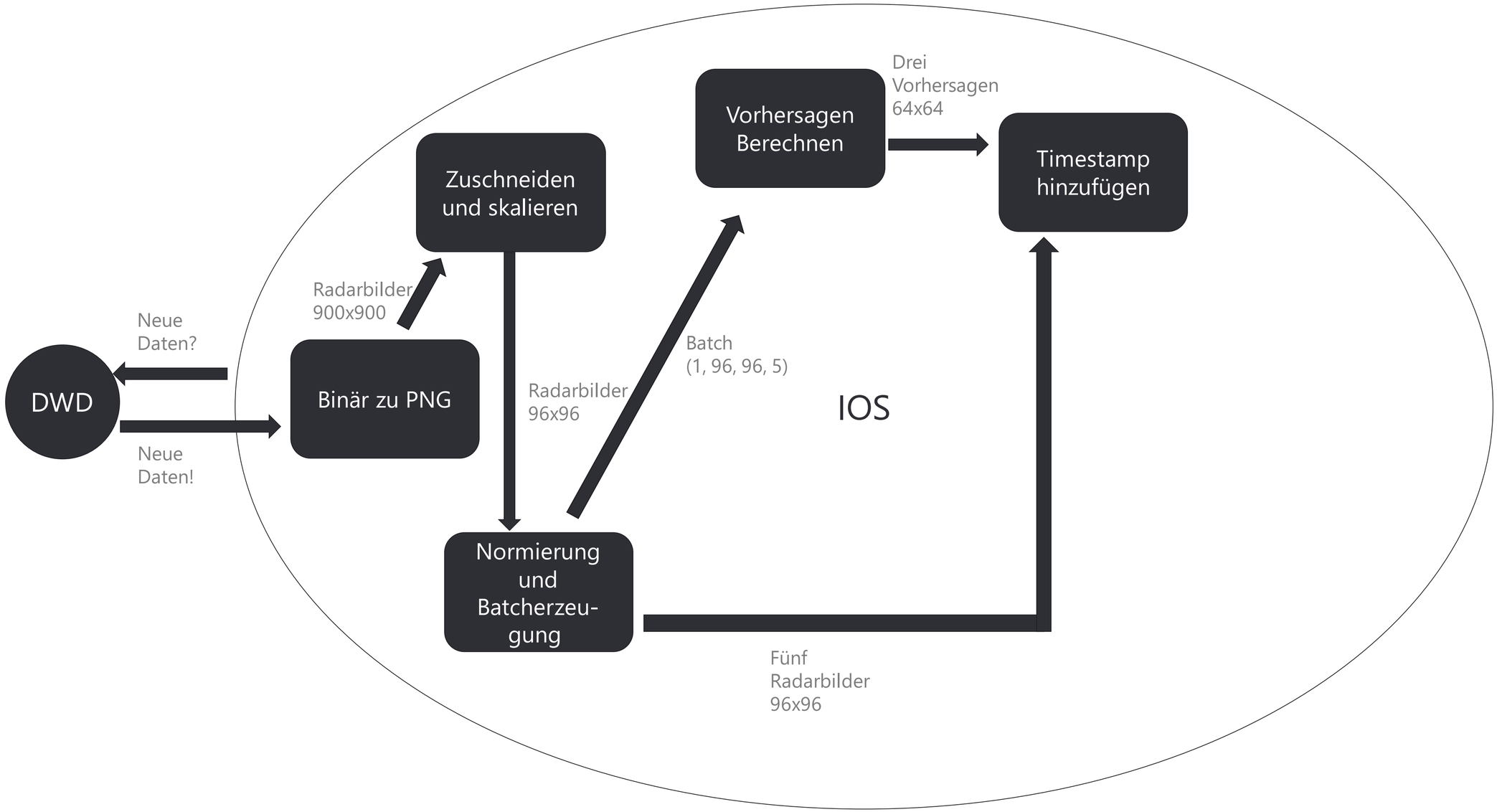
Pipeline


Zuschneiden und skalieren
Radarbilder 900x900
Neue Daten?
DWD
Binär zu PNG
Neue Daten!
Radarbilder 96x96
IOS
Normierung und Batcherzeu-gung

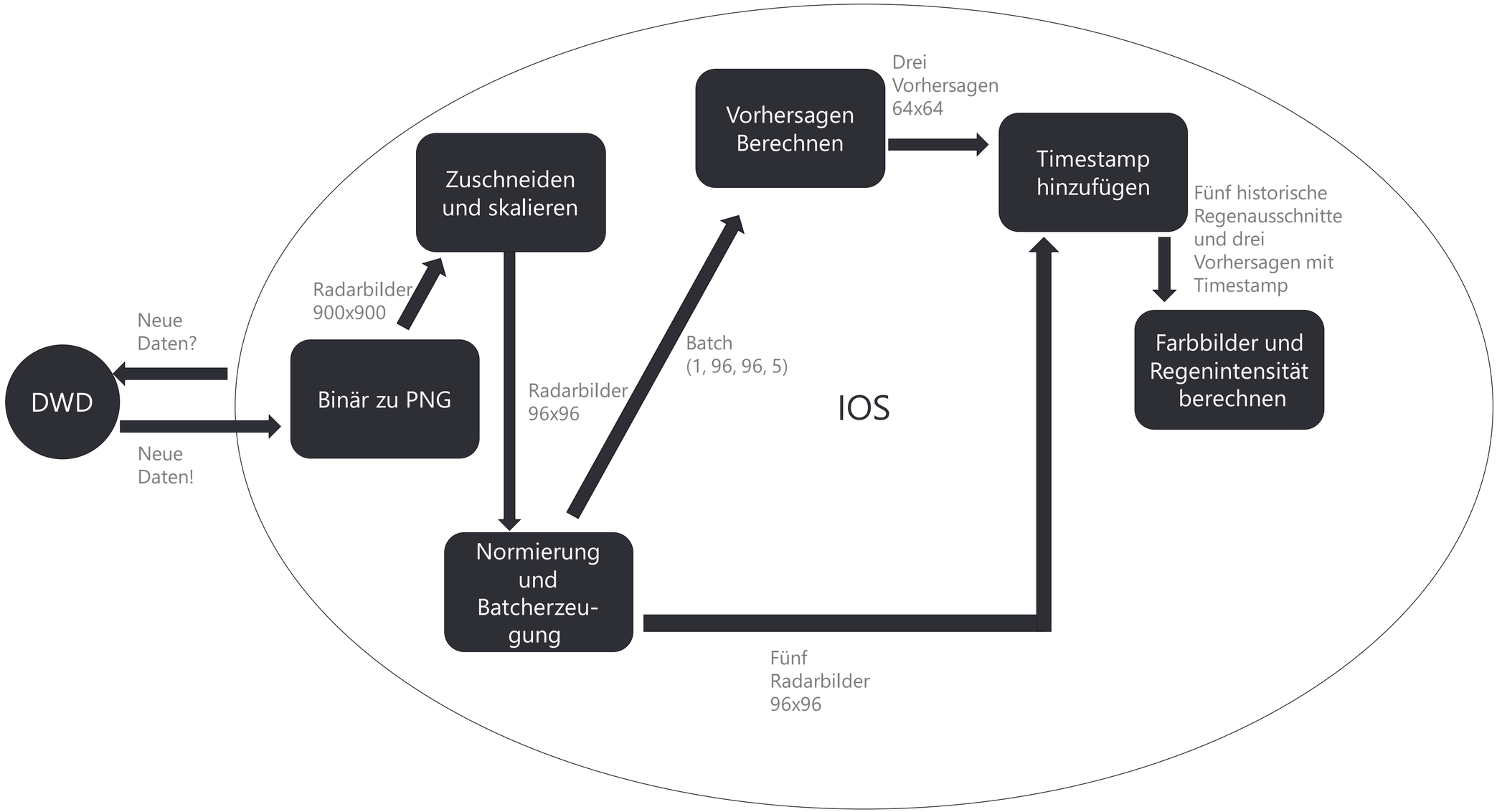
Pipeline


Vorhersagen Berechnen
Zuschneiden und skalieren
Radarbilder 900x900
Neue Daten?
DWD
Binär zu PNG
Neue Daten!
Radarbilder 96x96
Batch (1, 96, 96, 5)
IOS
Normierung und Batcherzeu-gung

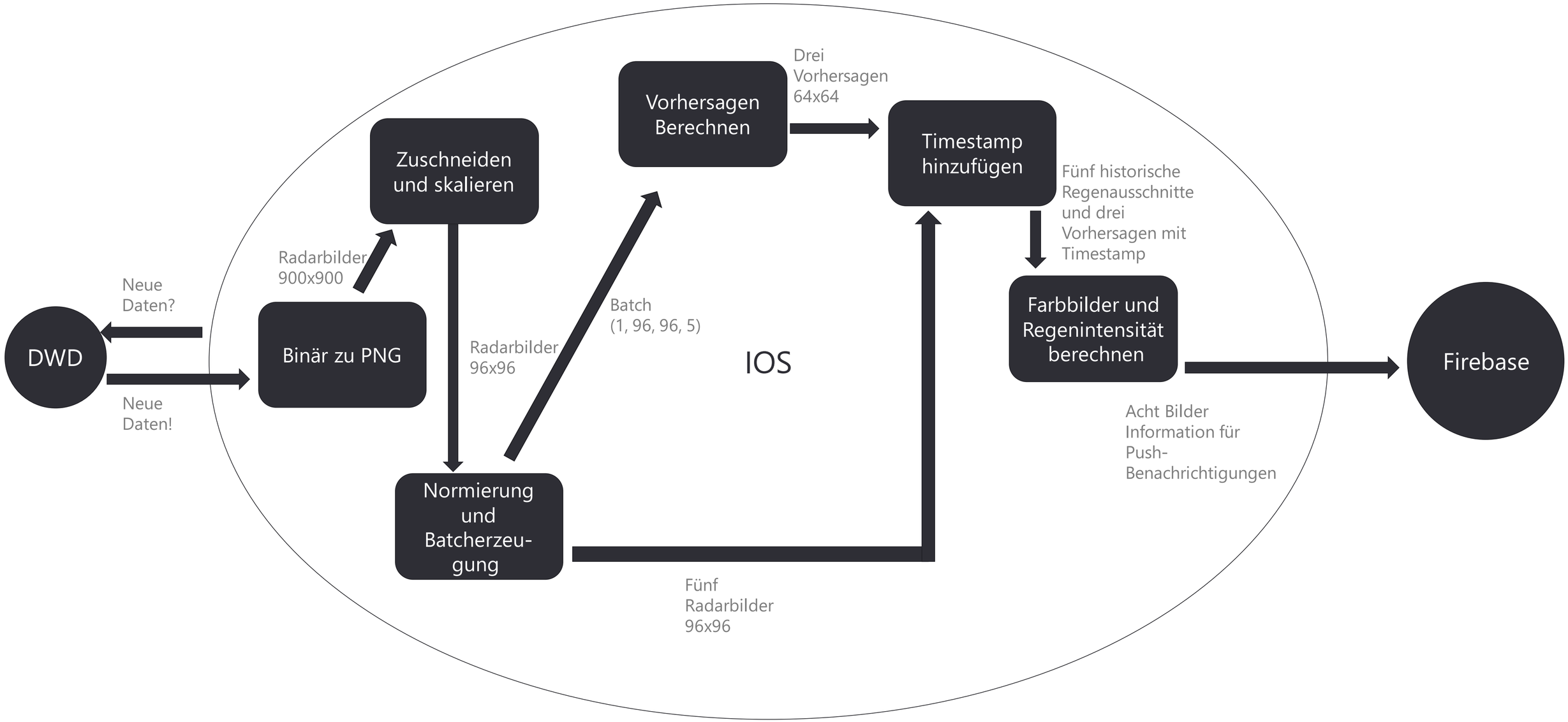
Pipeline


DWD
Neue Daten?
Neue Daten!
Binär zu PNG
Radarbilder 900x900
Zuschneiden und skalieren
Radarbilder 96x96
Normierung und Batcherzeu-gung
Batch (1, 96, 96, 5)
Vorhersagen Berechnen
Drei Vorhersagen 64x64
Timestamp hinzufügen
Fünf Radarbilder 96x96
IOS

Pipeline


DWD
Neue Daten?
Neue Daten!
Binär zu PNG
Radarbilder 900x900
Zuschneiden und skalieren
Radarbilder 96x96
Normierung und Batcherzeu-gung
Batch (1, 96, 96, 5)
Vorhersagen Berechnen
Drei Vorhersagen 64x64
Timestamp hinzufügen
Fünf historische Regenausschnitte und drei Vorhersagen mit Timestamp
Farbbilder und Regenintensität berechnen
Fünf Radarbilder 96x96
IOS

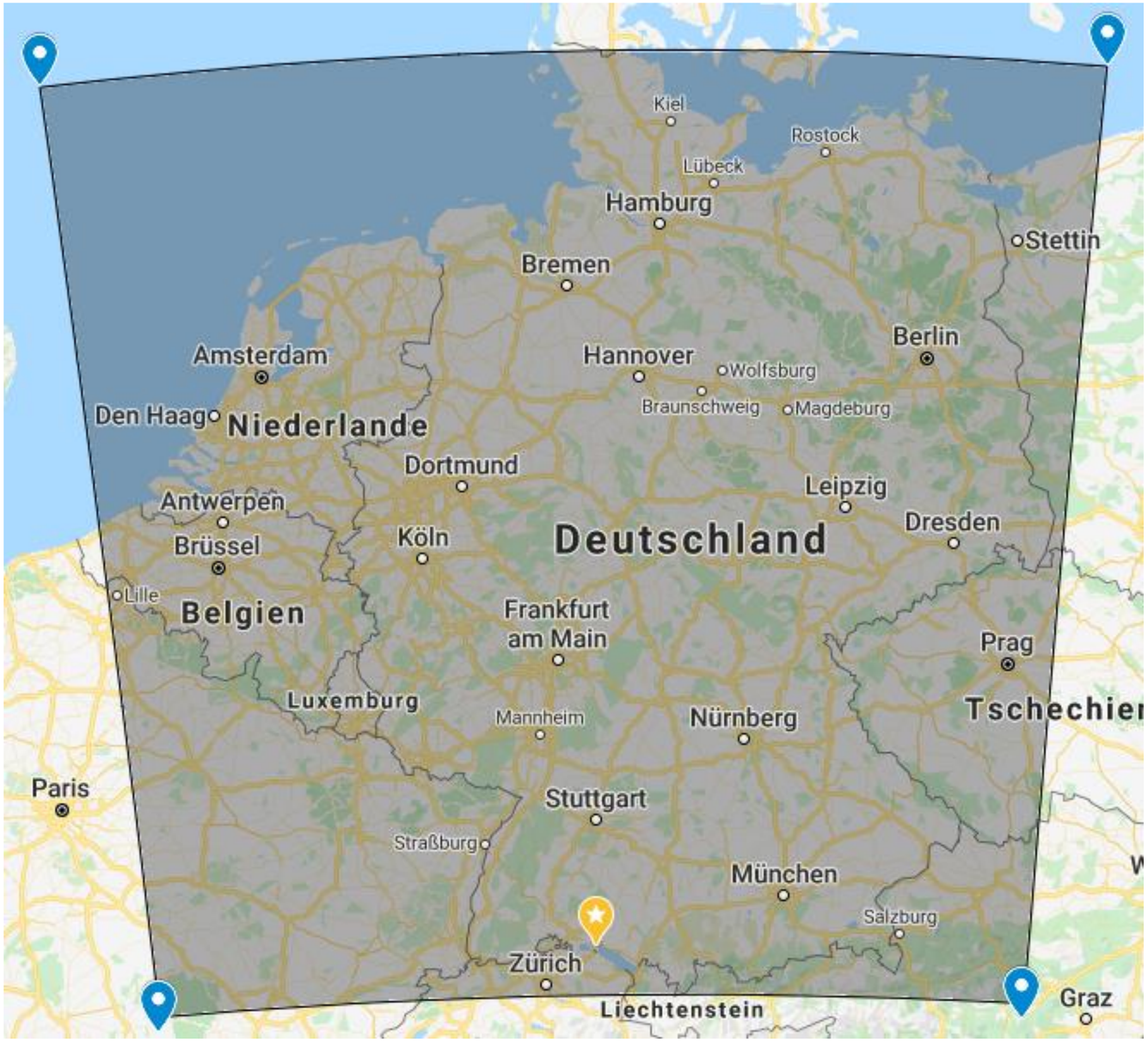
Pipeline

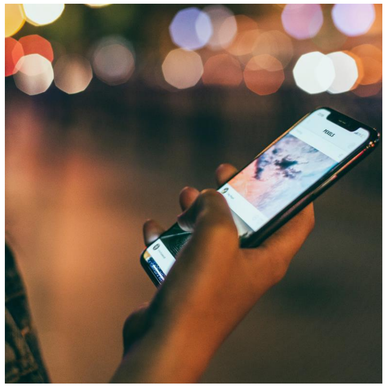
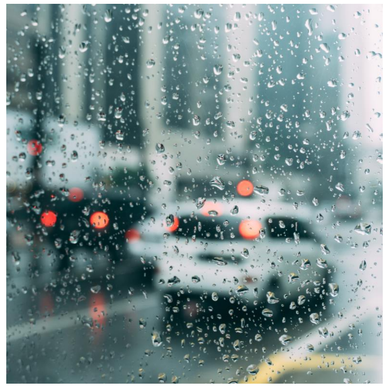
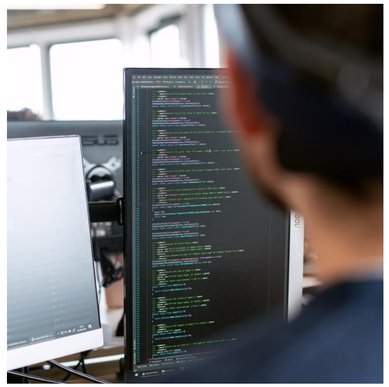
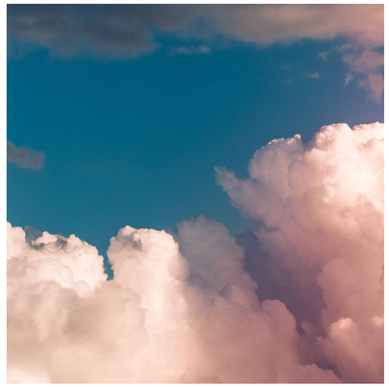

DWD
Neue Daten?
Neue Daten!
Binär zu PNG
Radarbilder 900x900
Zuschneiden und skalieren
Radarbilder 96x96
Normierung und Batcherzeu-gung
Batch (1, 96, 96, 5)
Vorhersagen Berechnen
Drei Vorhersagen 64x64
Timestamp hinzufügen
Fünf historische Regenausschnitte und drei Vorhersagen mit Timestamp
Farbbilder und Regenintensität berechnen
Acht Bilder Information für Push-Benachrichtigungen
Fünf Radarbilder 96x96
IOS
Firebase

Ausblick

Was noch zu tun ist.

Entzerrung der Vorhersage Bilder





Fragen?