

ICS 35.240.01

Leitfaden für die Entwicklung von Deep-Learning-Bilderkennungssystemen

Guideline for the development of deep learning image recognition systems

Guide pour le développement de l'apprentissage profond de systèmes de reconnaissance d'image

Zur Erstellung einer DIN SPEC können verschiedene Verfahrensweisen herangezogen werden:
Das vorliegende Dokument wurde nach den Verfahrensregeln einer PAS erstellt.

Gesamtumfang 17 Seiten

Inhalt

	Seite
Vorwort	3
Einleitung	4
1 Anwendungsbereich	5
2 Normative Verweisungen	5
3 Begriffe	5
4 Abkürzungen	7
5 Definition des Anwendungsszenarios	8
6 Kategorisierung des Bilderkennungsproblems	8
7 Eignungsmerkmale für Deep Learning	8
8 Technische Umsetzung	9
8.1 Allgemein	9
8.2 Problemdefinition	9
8.3 Erstellung des Datensatzes	9
8.4 Strukturierung der Daten zum Lernen des Deep-Learning-Bilderkennungssystems	10
8.5 Identifikation von potenziellen Deep-Learning-Architekturen	10
8.6 Ablaufstruktur von Lern-Experimenten	11
8.7 Modellselektion	11
8.8 Modell-Deployment	11
8.9 Qualitätssicherung des Deep-Learning-Bilderkennungssystems	11
8.9.1 Allgemein	11
8.9.2 Konsistenz	12
8.9.3 Repräsentativität	13
8.9.4 Genauigkeit	13
8.10 Kontinuierliches Lernen	14
Anhang A (normativ) Darstellung des Ablaufs der technischen Umsetzung eines DL-Systems	15
Literaturhinweise	17
Bilder	
Bild A.1 — Schematische Entwicklung eines DL-Systems	16

Vorwort

Diese DIN SPEC wurde nach dem PAS-Verfahren erarbeitet. Die Erarbeitung von DIN SPEC nach dem PAS-Verfahren erfolgt in DIN SPEC (PAS)-Konsortien und nicht zwingend unter Einbeziehung aller interessierten Kreise.

Die Erarbeitung und Verabschiedung des Dokuments erfolgten durch die nachfolgend genannten Initiator(en) und Verfasser:

- MindPeak GmbH: Felix Faber, Marc Päpper und Dr. Tobias Lang
- PSIORI GmbH: Dr. Sascha Lange und Lars Eickmeier
- FUSE-AI GmbH: Matthias Steffen und Dr. Sabrina Reimers-Kipping
- Hochschule Düsseldorf: Prof. Dr.-Ing. Thomas Zielke
- Deutsche Röntgengesellschaft e. V.: Dr. Bettina Baeßler
- Quality Match GmbH: Dr. Daniel Kondermann

Für dieses Thema bestehen derzeit keine Normen im Deutschen Normenwerk.

DIN SPEC (PAS) sind nicht Teil des Deutschen Normenwerks.

Für diese DIN SPEC wurde kein Entwurf veröffentlicht.

Trotz großer Anstrengungen zur Sicherstellung der Korrektheit, Verlässlichkeit und Präzision technischer und nicht-technischer Beschreibungen kann das DIN SPEC (PAS)-Konsortium weder eine explizite noch eine implizite Gewährleistung für die Korrektheit des Dokuments übernehmen. Die Anwendung dieses Dokuments geschieht in dem Bewusstsein, dass das DIN SPEC (PAS)-Konsortium für Schäden oder Verluste jeglicher Art nicht haftbar gemacht werden kann. Die Anwendung der vorliegenden DIN SPEC (PAS) entbindet den Nutzer nicht von der Verantwortung für eigenes Handeln und geschieht damit auf eigene Gefahr.

Es wird auf die Möglichkeit hingewiesen, dass einige Elemente dieses Dokuments Patentrechte berühren können. DIN ist nicht dafür verantwortlich, einige oder alle diesbezüglichen Patentrechte zu identifizieren.

Die kostenfreie Bereitstellung dieses Dokuments als PDF-Version über den Beuth WebShop wurde im Vorfeld finanziert.

Einleitung

Deep Learning (DL) bildet den neusten State-of-the-Art der Bilderkennung und wird mit Erfolg in vielen Gebieten angewendet, z. B. für autonomes Fahren, Gesichtserkennung und Spracherkennung [1]. Deep Learning ist eine Methode der künstlichen Intelligenz (KI) und basiert auf tiefen, künstlichen neuronalen Netzen. Anstatt mit viel Aufwand fest codierte Regeln durch einen menschlichen Experten festzulegen, werden mit einem Deep-Learning-Bilderkennungssystem statistische Muster aus Beispieldaten gelernt. Durch den Einsatz von Deep Learning werden viele Aufgaben in der Bilderkennung am Computer lösbar, die zuvor mit klassischen Methoden der Bildverarbeitung nicht zufriedenstellend bearbeitet werden konnten.

Eine der häufigsten Anwendungen von Deep Learning ist die Erkennung und Analyse von Bildern. Hierzu zählt z. B. die Erkennung von Objekten (Auto, Fußgänger), die Segmentierung von Bildbereichen (normales Gewebe, Tumorgewebe) [2] oder das Generieren von Bildunterschriften („ein großer weißer Vogel steht im Wald“) [3]. Die hohe Erkennungsgenauigkeit von Deep-Learning-Bilderkennungssystemen bietet vielversprechende Möglichkeiten für die Automatisierung langwieriger, monotoner Aufgaben und Erschließung neuer Anwendungsbereiche in der Bilderkennung.

1 Anwendungsbereich

Dieses Dokument stellt die Voraussetzungen, unter denen Bilderkennungsprobleme mit Hilfe eines Deep-Learning-Bilderkennungssystems bearbeitet werden können, auf. Es erlaubt Entscheidungsträgern, Kenntnisse über die Anwendungsmöglichkeiten eines Deep-Learning-Bilderkennungssystems und seine Struktur zu gewinnen.

Mit Hilfe dieses Dokumentes kann die Schätzung des Aufwandes und des Nutzens eines Deep-Learning-Bilderkennungssystems unterstützt werden sowie eine genauere Erfolgsprognose erstellt werden.

Dieses Dokument gibt Leitlinien zur praktischen Umsetzung eines Deep-Learning-Bilderkennungssystems vom Vorgehen bei der Datensammlung über die Strukturierung der Daten zum Lernen der KI-Bilderkennung bis zur Ablaufstruktur von Lern-Experimenten und zur Qualitätssicherung.

Dieses Dokument ist für Entscheidungsträger von KI-Projekten zur Bilderkennung als auch für deren Umsetzer gedacht.

Dieses Dokument adressiert keine spezifischen Angaben zu den Themengebieten aktives Lernen, mentales Lernen und kontinuierliches Lernen.

ANMERKUNG Dieses Dokument beschreibt das allgemeine Vorgehen für die Entwicklung von Deep-Learning-Bilderkennungssystemen. Für spezifische Anwendungsfälle wie beispielsweise die biometrische Erkennung von Personen verweisen wir auf weiterführende Normen und Standards wie in diesem Falle die des ISO/IEC JTC 1/SC 37 „Biometrie“ (ISO/IEC TR 24714-1, ISO/IEC 2382-37 usw.) und des CEN/TC 224 „Persönliche Identifikation, elektronische Signatur, maschinenlesbare Karten sowie zugehörige Geräteschnittstellen und Verfahren“ (DIN EN 17054 usw.).

2 Normative Verweisungen

Es gibt keine normativen Verweisungen in diesem Dokument.

3 Begriffe

Für die Anwendung dieses Dokuments gelten die folgenden Begriffe.

ISO und IEC stellen terminologische Datenbanken für die Verwendung in der Normung unter den folgenden Adressen bereit:

- ISO Online Browsing Platform: verfügbar unter <http://www.iso.org/obp>
- IEC Electropedia: verfügbar unter <http://www.electropedia.org/>

DIN und DKE stellen terminologische Datenbanken für die Verwendung in der Normung unter den folgenden Adressen bereit:

- DIN-TERMinologieportal: verfügbar unter <https://www.din.de/go/din-term>
- DKE-IEV: verfügbar unter <http://www.dke.de/DKE-IEV>

3.1

Bildererkennung

Erfassung und Analyse eines Bildes, der Objekte, aus denen es besteht, sowie deren Eigenschaften und räumlichen Beziehungen durch eine Funktionseinheit

Anmerkung 1 zum Begriff: Bildererkennung schließt Szenenanalyse ein.

[QUELLE: DIN ISO/IEC 2382-28:1998, 28.01.14]

3.2

Bilderkennungsproblem

Problem ein bestimmtes Ziel der Bildererkennung zu lösen, z. B. ein Bild zu klassifizieren

3.3

Crowdsourcing

Methode, um Daten über das Internet von Personen generieren zu lassen

3.4

Deep Learning

Lernen statistischer Muster durch tief verschachtelte neuronale Netze

3.5

Deep-Learning-Architektur

DL-Architektur

Beschreibung der verschiedenen genutzten Recheneinheiten des neuronalen Netzes sowie deren Verbindung

Anmerkung 1 zum Begriff: Aus der Architektur ergibt sich die Anzahl der lernbaren Gewichte im Netzwerk. Je höher diese Anzahl, desto schwierigere Probleme sind lösbar; aber desto mehr Daten sind dazu notwendig.

3.6

Deep-Learning-Bildererkennungssystem

DL-System

das zu entwickelnde ComputermodeLL, welches das Bilderkennungsproblem löst

3.7

Domäne

spezifischer Wissensbereich oder bereichsspezifisches Expertenwissen; genauer der Bereich, in dem das DL-Modell eingesetzt wird

[QUELLE: DIN ISO/IEC 2382-28:1998, 28.01.04 – modifiziert – „; genauer der Bereich, in dem das DL-Modell eingesetzt wird“ wurde hinzugefügt.]

3.8

Grundwahrheiten

zu den Daten zugehörige ideale Ausgabewerte

BEISPIEL Grundwahrheiten können Klassifikationen (Bild 1 - Hund, Bild 2 - Katze) sein.

3.9

Initialdaten

Gesamtmenge der Daten, die zum Start der Entwicklung des DL-Systems zur Verfügung stehen oder gesammelt werden, und die in Trainings-, Test- und Validierungsdaten aufgeteilt werden

Anmerkung 1 zum Begriff: In der Regel bestehen Initialdaten aus Bildern und den zugehörigen Grundwahrheiten/Annotationen.