

CKi ein CNN-Model in C++

by Simeon Stix

23. November 2023

Version: 1

Betreuer: Sven Nüesch

Inhaltsverzeichnis

1	Analyse	3
1.1	Aufgabenstellung	3
1.2	Zielgruppe	3
1.3	Anforderungen	3
1.3.1	Must-haves	3
1.3.2	Nice-to-haves	5
1.3.3	Use Cases	6
1.3.4	Use Case Diagramme	10
1.3.5	Ablaufdiagramme	11
1.4	Umsetzung	14
1.4.1	Entwicklungsumgebung	14
1.4.2	CNN	15
1.4.3	Testdaten	15
1.5	Abwägung des Nutzens und der Alternativen	16
1.5.1	Alternativen	16
1.5.2	Kosten	16
1.5.3	Nutzen	17
1.5.4	Effektivität	17
1.5.5	Vergleich	17
1.5.6	Nutzwertanalyse	17
2	Planung	19
2.1	Rollenverteilung	19
2.2	Aufgabenliste	19
2.3	Meilensteine	21
2.4	Gantt	21
3	Design	26
3.1	Konsole	26
3.1.1	Training	26
3.1.2	Verify	26
3.1.3	Prediction	27
3.2	Datenbank	27

3.2.1	UByte	27
3.3	Code	28
3.3.1	Klassendiagramm	28
3.3.2	Trainingsdaten	29
3.3.3	Tests	29
Literaturverzeichnis		31
Abbildungsverzeichnis		32

Kapitel 1

Analyse

1.1 Aufgabenstellung

Die Aufgabe, die sich das Projekt CKi stellt, ist die Wissenserweiterung des Entwicklers. Dabei wird ein Programm geschrieben, welches einzelnen handgeschriebenen Ziffern erkennen kann. Dies geschieht mittels KI. Dabei wird die KI komplett vom Entwickler geschrieben. Da keine modernen externen Grundlagen verwendet werden, wird das Produkt, das Programm, nicht die Geschwindigkeit einer modernen KI erreichen. Dies spielt jedoch keine Rolle, da dieses Projekt nicht wegen des Endprodukts durchgeführt wird.

1.2 Zielgruppe

Das Projekt CKi ist als solches nicht ausgelegt, einem realen Anwendungszweck zu entsprechen oder eine Lösung oder einen Lösungsansatz für einen solchen zu bieten. Diesbezüglich liegt der einzige Nutzen von CKi nicht in dessen Produkt, sondern nur im Wissensgewinn und Verständnis gewinn für den Entwickler in den Bereichen der künstlichen Intelligent oder genauer im Bereich des maschinellen Lernens mit einem *Convolutional Neural Network*, der Realisation von Anwendungen mit C++ und dessen Möglichkeiten Hardware direkt in die programminternen Abläufe einzubinden. Somit richtet sich CKi nicht nach dem Grundsatz ein bestmögliches nutzbares Produkt zu sein, sondern lediglich nach dem grössten Wissensgewinn für den Entwickler. Nach diesem Grundsatz ist die resultierende Zielgruppe der Entwickler und vereint so multiple Rollen des Projektes CKi in einer Person.

1.3 Anforderungen

1.3.1 Must-haves

Die Must-haves wurden aus dem Themenblatt, welches am 15.09.2023 bei Walter Schnyder eingereicht wurde, übernommen und mit weiterführenden Elaborationen versehen.

- Rückgabe in Prozentwerten, die die Wahrscheinlichkeit der Übereinstimmung mit dem digitalen Gegenstück der handgeschriebenen Zahl abbildet. "Welche Zahl wurde (vermutlich) aufgeschrieben?"

Erläuterung: Da ein simples neuronales Netzwerk für maschinelles Lernen durch ein "Netz aus Knotenpunkten" gebaut wird und jeder dieser Knotenpunkte, auch die Knotenpunkte, welche bei einem solchen neuronalen Netzwerk als Endschnittstellen fungieren, einzeln berechnet werden, erhält man, bei Anwendungsfall von CKi, eine multiple Anzahl von Prozentzahlen, welche zur Interpretation des gelieferten Ergebnisses verwendet werden können. Diese Rückgabe der einzelnen Prozentwerte erfolgt zum Beginn über eine Konsolenausgabe. Diese wird später, wie in *"Nice-to-haves"* unter GUI erläutert, in ein grafisches Nutzerinterface integriert und zu diesem Zeitpunkt evtl. auch interpretiert (wobei die einzelnen Prozentwerte weiterhin einsichtig bleiben sollten).

- CNN-Algorithmus (trainiert auf Zahlenwert)

Erläuterung: Ein CNN-Algorithmus oder auch Convolutional Neural Network wird beim maschinellen Lernen oft bei der Interpretation von Bildern genutzt. Dabei wird das Bild in kleinere Abschnitte unterteilt und „einzeln“ an den gehirnähnlich aufgebauten Algorithmus weitergegeben.

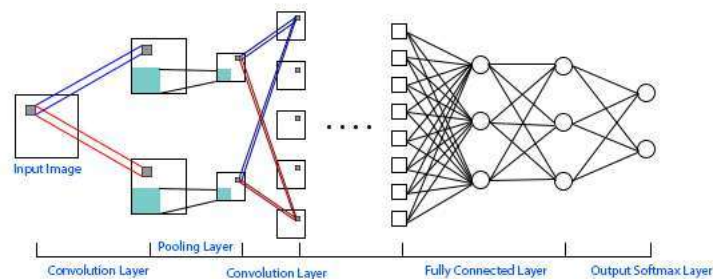


Abbildung 1.1: CNN Model Aufbau Grafik von researchgate.net

Im Projekt CKi wird dieser wie im Themenblatt beschrieben mit Bild von einzelnen handgeschriebenen Ziffern trainiert.

- Nutzer-Input → Nutzer darf eine Zahl zeichnen

Erläuterung: Ein solches CNN-Modell zu erbauen und zu trainieren ist zwar die Grundlage für dieses Must-have, jedoch sollte das Produkt auch erprobbar sein.

Diesbezüglich muss der Nutzer in der Lage sein, eine handschriftliche Ziffer an das neurale Netzwerk zu liefern. Hierbei ist die minimale Anforderung, dass der Nutzer in einer anderweitigen Applikation ein solches Bild erstellt hat und es nun interpretieren lassen kann. Wie in Nice-to-haves unter GUI beschreiben, wird diese Eingabemöglichkeit (eine Ziffer zu zeichnen oder eine schon vorhandene grafische Abbildung zu verwenden) in einem weiteren Entwicklungsschritt direkt in die grafische Nutzeroberfläche der Applikation integriert.

1.3.2 Nice-to-haves

Die Nice-to-haves wurden aus dem Themenblatt, welches am 15.09.2023 bei Walter Schnyder eingereicht wurde, übernommen und mit weiterführenden Elaborationen versehen. Zudem behalte ich mir als Verfasser dieses Dokumentes, als Entwickler des Projektes CKi und als Zielgruppe des Projektes CKi vor, diese Liste in gegebene Falle zu erweitern.

- GUI

Erläuterung: Ein GUI (oder auch Graphical User-Interface) ist die grafische Nutzeroberfläche der Applikation. Da die Applikation primär dem Wissensgewinn gewidmet ist und dementsprechend nicht für den Nutzer optimiert wird, hat eine solche Erweiterung nur eine geringe Priorität. Diese Nutzeroberfläche wird selbst bei der Umsetzung in einer möglichst simplen Form gehalten. Dabei sollte es folgende Bestandteile beinhalten:

- Eine Möglichkeit für den Nutzer eine anderweitig gezeichnete Ziffer interpretieren zu lassen
- Eine Möglichkeit für den Nutzer eine Ziffer zu zeichnen.
- Eine weiter-interpretierte Ausgabe der Interpretation des CNN.
- Eine Ausgabe des nicht interpretierten Prozentwertes der Interpretation des CNN.

Bis zu dem Punkt, wo ein GUI realisiert wurde und in den Einsatz gestellt wird, sind nicht alle dieser Funktionen über die Konsole (das Interface zum Programm, welches vor dem GUI zum Einsatz kommt) verfügbar.

- GPU als Berechnungsplattform nutzen

Erläuterung: Die CPU in jedem Computer ist eine sehr „fokussierte“ Hardware, immer nur einen einzelnen Prozess kann diese berechnen. Dies ist für ein neuronales Netzwerk, welche Hunderte oder Tausende von Knotenpunkten in seinem Netzwerk hat und alle einzeln berechnet werden müssen, äusserst hinderlich. Die dedizierte Grafikkarte, wenn vorhanden, kann diesem Geschwindigkeitsverlust nachhelfen, da

eine solche GPU in der Lage ist, tausende Berechnungen gleichzeitig zu tätigen. Da im Projekt mit C++, einer Hardware-nahen Programmiersprache, gearbeitet wird, kann eine Anbindung an das Rechenpotential der GPU erreicht werden. Da dies jedoch ein äusserst schwieriger Prozess ist, sich diese Anbindung bei den unterschiedlichen Herstellern von GPUs unterscheiden kann und ich notwendig um ein funktionierendes CNN zu erstellen, wurde diese Optimierung als nicht notwendig eingestuft.

1.3.3 Use Cases

Die hier aufgelisteten Use Cases entsprechen den Use Cases nach den Must-haves und sind diesbezüglich ohne die grafische Benutzeroberfläche. Dies kann dazu führen, dass ein endgültiges Produkt nicht mehr kohärent zu den Use Cases steht. Im Allgemeinen sollte aber selbst eine solche Inkohärenz nicht in extremer Weise auftreten, da die Bedienung der Applikation als Konsole als auch als grafische Oberfläche in ähnlicher Weise auftreten sollte.

Nummer	Name	Akteur	Ablauf	Nachbedingungen	Ausnahmen	Anmerkungen
1	Interpretation	Nutzer, CNN	Siehe Ablaufdiagramm "Interpretation"	Der Nutzer sollte in der Konsole eine Auflistung aller möglichen Ziffern (0-9) und deren entsprechenden Wahrscheinlichkeit sehen. Zudem kann der Nutzer auch die kongruierende Zahl zur höchsten Möglichkeit auf einer speziellen Zeile in der Konsole ablesen.	Bei der Interpretation von Nutzer-eigenen Abbildungen kann es zu multi-plexen Fehlern kommen. Diese reichen von falschen Dateikodierung zu falscher Auflösung. Aufgrund dieser mannigfaltigen Möglichkeiten zu Fehlern können diese nicht alle hier erläutert werden. Eine Auflistung aller Fehler ist in der Dokumentation unter " <i>Fehler</i> " zu finden.	Offiziell ist zwar Training keine Vorbedingung zu Testet, jedoch macht die Anwendung der Applikation keinen Sinn, wenn man nicht erwarten kann, dass man ein realistisches oder sinnvolles Ergebnis erhält.

2	Training	Nutzer, CNN	Siehe Ablaufdiagramm "Training"	Der Nutzer erhält nach der Beendigung der Schulung des CNN-Models eine kurze Benachrichtigung, dass das Training abgeschlossen ist. Zudem erhält der Nutzer nach jedem einzelnen Datensatz den Output, dass nun x von y Datensätzen bearbeitet wurden.	Hierbei kann es zu zwei Fehlern kommen. Dabei handelt es sich beides um Fehler bei der Datenbank. Der erste ist ein Fehler, wenn die benötigten "Tabellen" der Datenbank nicht stimmig mit den Erwartungen sind und der andere, wenn keine Daten in der Datenbank zu finden sind. Für die genauen Fehlercodes ist die Dokumentation unter " <i>Fehler</i> " zu kontaktieren.	
---	----------	-------------	---------------------------------	--	--	--

3	Testen	Nutzer, CNN	Siehe Ablaufdiagramm "Testen"	Der Nutzer findet nach jedem Datensatz in der Test-Datenbank eine simple Gegenüberstellung der richtigen und der falschen interpretierten Ziffern. Diese weiterzuverarbeiten, ist dem Nutzer überlassen. Neben der Gegenüberstellung erhält der Nutzer auch die Benachrichtigung, dass nun x von y Datensätzen bearbeitet wurden.	Hier kann es zu zwei Fehlern kommen. Dabei handelt es sich beides um Fehler bei der Datenbank. Der erste ist ein Fehler, wenn die benötigten "Tabellen" der Datenbank nicht stimmgig mit den Erwartungen sind und der andere, wenn keine Daten in der Datenbank zu finden sind. Für die genauen Fehlercodes ist die Dokumentation unter " <i>Fehler</i> " zu kontaktieren.	Bei den Testdaten würde es evtl. Sinn machen, die einzelnen Konsolen Outputs auch in einer CSV-Datei festzuhalten, wobei dies mit Pipes in der Konsole dem Nutzer bereits offensteht. Offiziell ist zwar Training keine Vorbedingung zu Testen, jedoch macht das Testen nur begrenzten Sinn, wenn es noch nichts gibt, was sich zu testen lohnt.
---	--------	-------------	-------------------------------	---	--	--

1.3.4 Use Case Diagramme

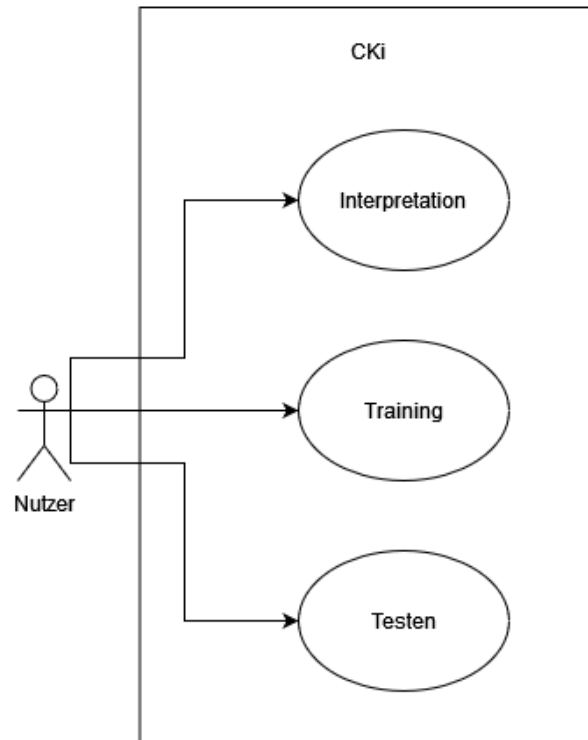


Abbildung 1.2: Use-Case-Diagramm

1.3.5 Ablaufdiagramme

Interpretation

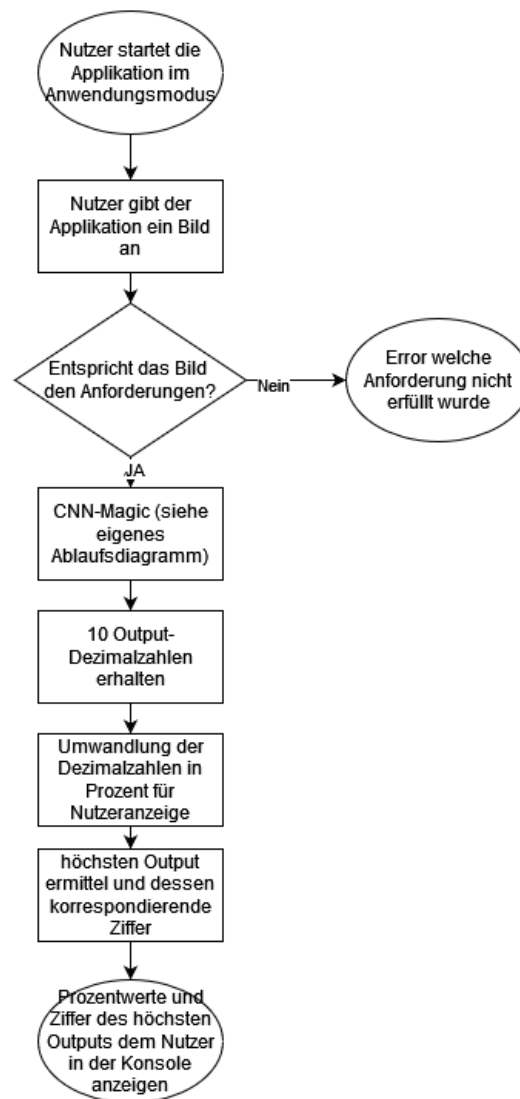


Abbildung 1.3: Ablaufdiagramm für die Interpretation von Nutzer-gelieferten Einzelbildern

Training

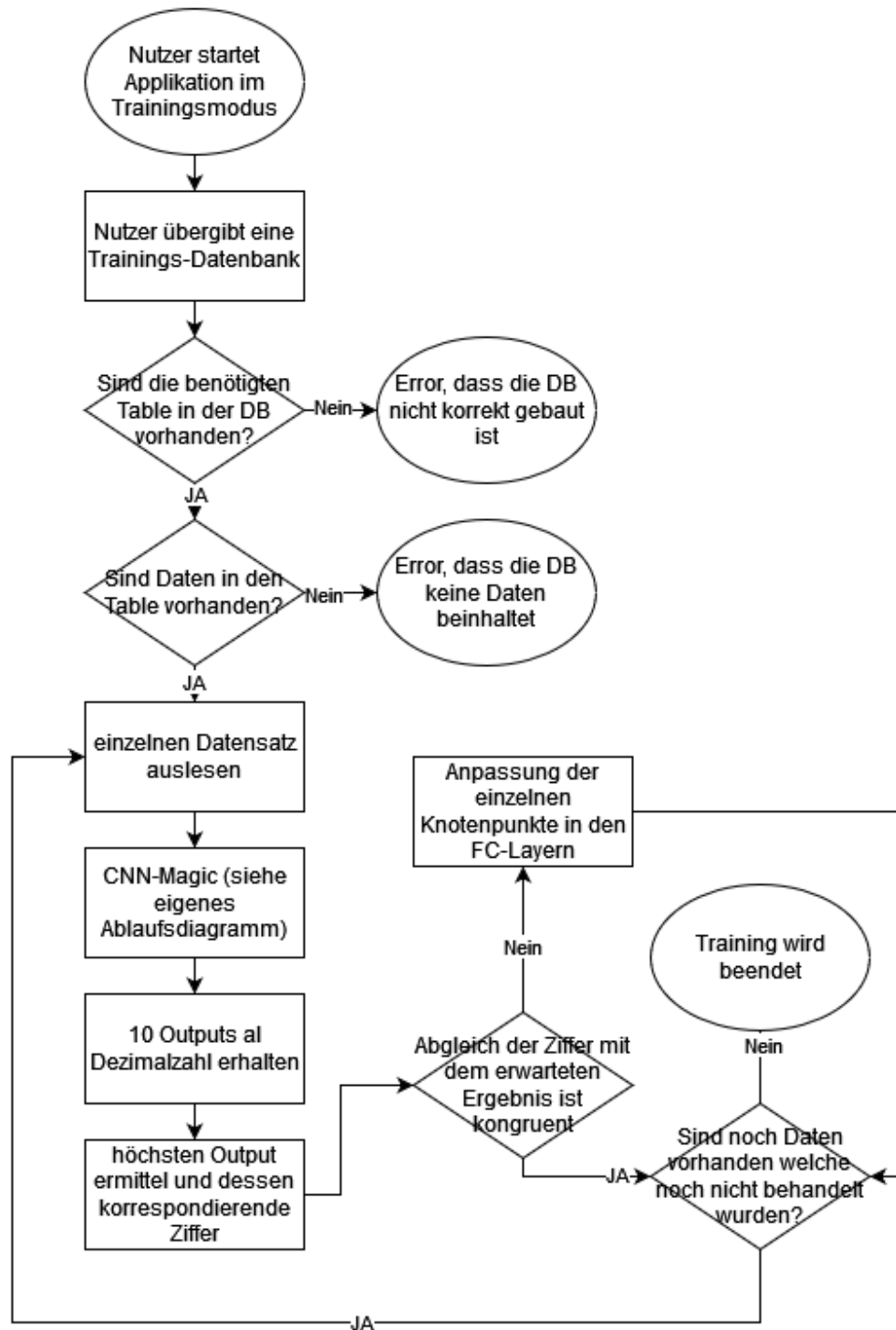


Abbildung 1.4: Ablaufdiagramm vom Training des CNN

Testen

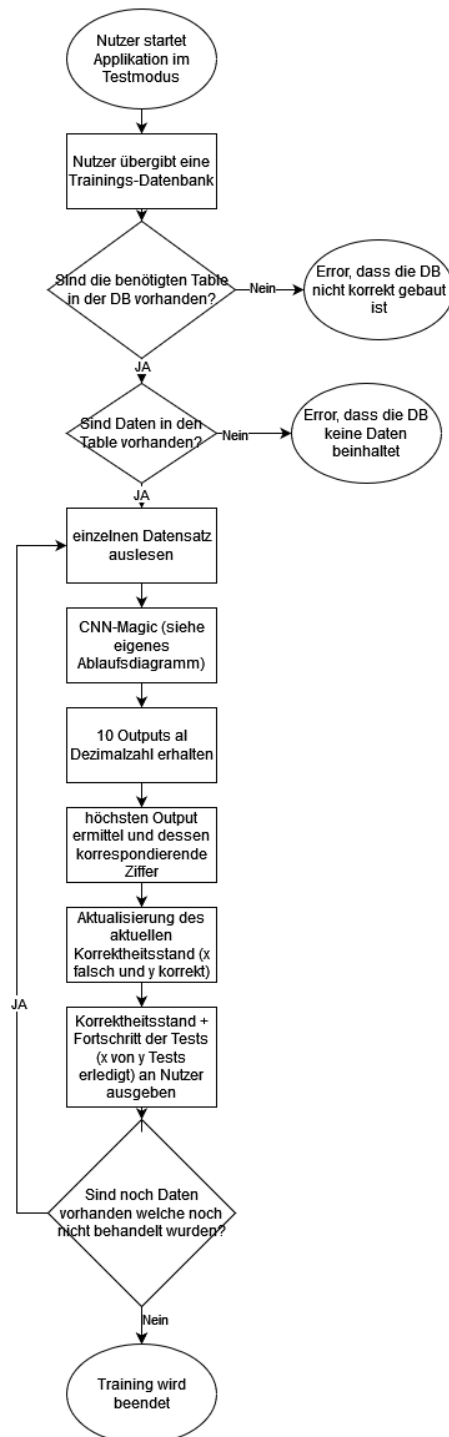


Abbildung 1.5: Ablaufdiagramm von einem Test des CNN

CNN-Magic

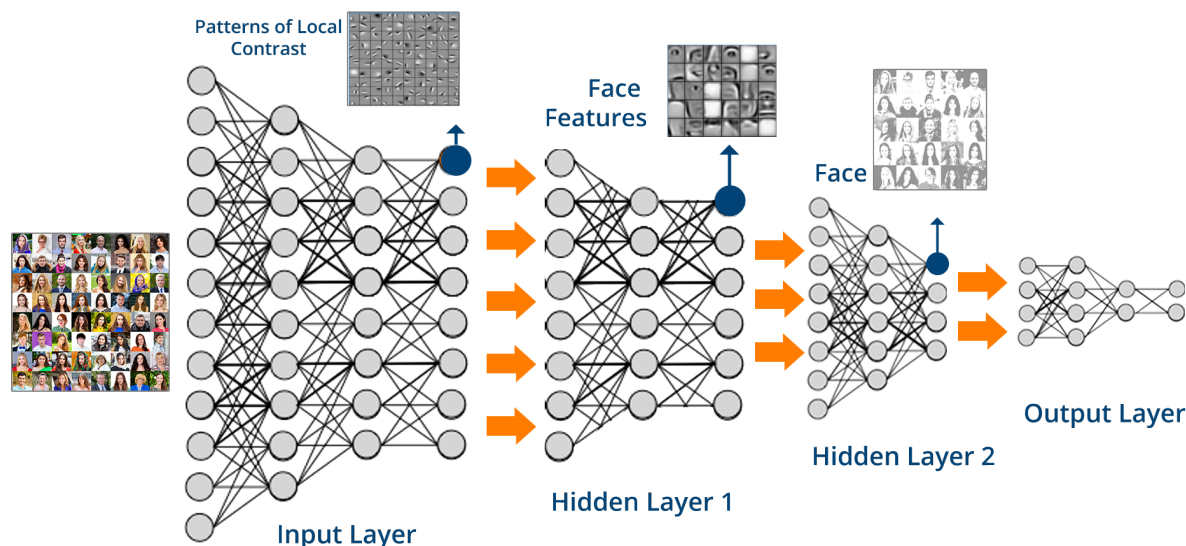


Abbildung 1.6: Visuelle Darstellung der Funktionsweise eines CNN von blog.goodaudience.com

1.4 Umsetzung

Wie im Themenblatt vom 15.09.2023 wird für die Umsetzung reines C++ verwendet.

1.4.1 Entwicklungsumgebung

Bei der Entwicklungsumgebung wird das JetBrains Produkt **CLion** zum Einsatz kommen. Dieses beinhaltet alle benötigten Funktionalitäten, die bei der Entwicklung des Produktes nötig sind. Gegebenen Falles könnte noch **SQLite Browser** zur Verwendung kommen, da dieses ein besseres (und für den Entwickler ein gewohntes) Interface zur Handhabung von lokalen Datenbanken bietet. Für die Version Control des Projektes wird lokal **Git** verwendet und zur Sicherung in der Cloud wird sowohl **GitHub** als auch **GitLab** verwendet. Die Entwicklungsumgebung mit IDE und Version Control wird primär auf dem Betriebssystem Windows verwendet. Es kann jedoch nicht garantiert sein, dass die Entwicklung nicht teilweise unter einer Distribution von Linux ablaufen wird. Hierbei sollten (evtl. abgesehen von der GUI) keine Kompatibilitätsprobleme mit sich führen.

Auflistung Versionen

- JetBrains Toolbox 2.1.1.18388, Windows 11, x64

- CLion (von JetBrains Toolbox) 2023.2.2
- Git 2.39.2.windows.1

1.4.2 CNN

Grundlagen

Der Kern des CNN lässt sich in drei unterschiedliche Arten von Schichten aufteilen. Zusätzlich gibt es noch zwei zusätzliche "Hilfs-Schichten, welche für die Ein- und Ausgaben zuständig sind. Die *Eingabeschicht* ist eine dieser zwei erwähnten Schichten. Diese akzeptiert im Falle von CKi jeweils ein Bild von speziell definierten Grössen und nimmt, gibt jeder Pixel in Graustufen (Dies ist entscheidend, weil so der Farbwert des Pixels von drei, zwei Byte grossen Zahlen zu nur einer zwei Byte grossen Zahl reduziert wird.) als ein sogenanntes Eingabeneuron. Auf die Eingabeschicht folgt die *Convolutional Layers* oder auch *Faltungsschichten*. Diese Faltungsschichten sind einer der Schlüsselbestandteile eines CNN. Die Convolutional Layer dienen dazu, bestimmte Merkmale und Schlüsselemente aus dem Bild zu extrahieren. Hierbei wird bei jedem der Convolutional Layer eine mathematische Faltungsoperationen durchführt. Nach den Faltungsschichten kommen die simplen *Pooling-Schichten*. Bei einer Pooling-Schicht wird die Dimension eines Bildes reduziert (Downsampling). Dies kann durch zwei Arten geschehen, entweder durch Max-Pooling, dabei wird aus einer Liste von Werten nur der höchste übermittelt oder durch Durchschnitts-Pooling, wobei der Durchschnitt dieser Liste berechnet und übermittelt wird. Durch diese Datenreduktion wird Rechenleistung gespart und so das Ergebnis schneller und stabiler geliefert. Vor der Ausgabeschicht gibt es noch die *FC Layers* oder auch *Vollständig verknüpfte Schichten*. Bei diesen ist jeder Knotenpunkt mit jedem Knotenpunkt in der nächsten Schicht verbunden. Dies ist das Herz des gesamten Modells. Dabei wird in jedem Knotenpunkt ein neuer Wert berechnet, um in der letzten Schicht, der *Ausgabeschicht* diese in Wahrscheinlichkeit zu "konvertieren".

Implementation

Bei der Implementation eines CNN-Modells wurde für CKi der Entschluss getroffen, eine C++-Klasse für jede benötigte Art von Layer zu schreiben. Danach muss ein Grundgerüst, eine Main-Klasse realisiert werden, die alle Klassen miteinander kombiniert. Dies erleichtert es in einem späteren Schritt die grosse Umstellung von einer Konsolen-Anwendung hin zu einer grafisch passierten Anwendung (auch wenn immer noch grosse Teile der Main-Klasse ausgetauscht werden müssen.). Für die Speicherung der Berechnungswerte für die FC-Layer ist die Abwägung zu treffen, ob es sinnvoll ist, diese in einer lokalen Datenbank zu speichern oder in eine konkret hierfür entworfene Datei zu schreiben.

1.4.3 Testdaten

Die Testdaten für CKi können in unzähliger Ausführung auf Kaggle gefunden werden.

Speicherung

Im gegebenen Fall würde es für die Schulung, des CNN, Sinn ergeben, die einzelnen Bilder nicht über Ordner- oder Archivstrukturen zu speichern, sondern in eine oder mehrere lokale Datenbanken abzuspeichern.

1.5 Abwägung des Nutzens und der Alternativen

1.5.1 Alternativen

Die möglichen Alternativen zu einer kompletten Realisation eine CNN oder irgendeinem neuronalen Netzwerk in C++ ohne Bibliotheken oder Frameworks für neuronale Netze sind endlos. Selbst wenn man bereits bestehende Produkte begutachtet, wird man auf GitHub schnell fündig. Bei einer Suche auf GitHub mit dem Query "mnist digit recogniser" findet man nach Stand 20.10.2023 02:13 178 Repositorien (je eines in C, C++, C#, Rust, drei in Java + Kotlin und 53 in Python). Selbst wenn man unter dem Beschluss steht, dass man das CNN selber bauen wolle, würde man für jede beliebige Sprache eine Bibliothek oder Framework finden, die einem diese Aufgabe erleichtern würde.

1.5.2 Kosten

Bezüglich der Kosten kann das Projekt CKi nicht mit den Alternativen mithalten. C++ ist keine Sprache, in welcher man schnell ein Produkt hervorbringt. Dies belegen etliche Studien und Artikel (Programming Languages Table, An Empirical Comparison of Seven Programming Languages, Programming Languages by Energy Efficiency, Development time in various languages) Zudem, dass keine dezidierten Bibliotheken für maschinelles Lernen zum Einsatz kommen, erhöht den Entwicklungsaufwand, Zeitaufwand und so die Kosten. Wenn man die Stundenzahl aus der Planung mit einem Stundensatz von 50 CHF quantifizieren, käme man auf einen gesamten Kostenaufwand von 8'600.- + Lizenzkosten. Zuzüglich kämen noch Hardware Anschaffungen für die Entwicklung, diese können jedoch im Falle vom Projekt CKi vernachlässigt werden.

Senior Developer	2h	100.00 Fr.
Junior Developer	89h	4450.00 Fr.
Projektleitung	39h	1950.00 Fr.
Hardware	0	0.00 Fr.
Software	Lizenz	827.00 Fr.
Möglicher Zusatz-Entwicklungskosten	42h	2100.00 Fr.
Total		9427.00 Fr.

1.5.3 Nutzen

Wie bereits im Abschnitt ?? erwähnt, ist dieses Produkt auf den Wissensgewinn ausgerichtet und nicht auf das Produkt selbst. Dementsprechend ist in der geplanten Umsetzung der Projektes CKi der maximale Nutzen gewährleistet. Leider lässt sich dieser Nutzen nicht/schlecht quantifizieren.

1.5.4 Effektivität

Der Vergleich in Effektivität der Anwendung, die aus dem Projekt CKi resultiert und den professionell entwickelten Bibliotheken in Python, etc. zieht, wird CKi selbst mit GPU-Berechnung und dem Geschwindigkeitsvorteil von C++ gegen diese Bibliotheken in Bezug auf Run-Time-Length verlieren.

1.5.5 Vergleich

Bestehendes Produkt		Eigene Kreation	
Sektion	Anzahl	Sektion	Anzahl
Anschaffungskosten	0.- Fr.	Kreation	8'600.- Fr.
Wiederholende Kosten	0.- Fr.	wiederholende Kosten	0.- Fr.
Absolut	0.- Fr.	absolut	8'600.- Fr.

Eine kurze Zusammenfassung der Kosten-Nutzen-Analyse zeigt den desolaten Zustand von CKi. Es werden Kosten für die Arbeitszeit von 8'600 CHF anfallen. Die Effektivität ist niedriger, als wenn dasselbe Produkt mit einer vorhandenen Bibliothek geschrieben werden würde (und die Entwicklungsdauer wäre ebenfalls geringer). Der Nutzen des Endproduktes ist nichtig, da dieses nie zur wirklichen Anwendung kommen wird. Weiters gibt es unzählige kostenlose Alternativen, die es nur zu verwenden gilt. Das Projekt CKi kann und wird sich nur auf den Wissensgewinn für den Entwickler stützen, in der Hoffnung, dass dieser die Kosten-Politik aufwiegen wird.

1.5.6 Nutzwertanalyse

Bei der Nutzwertanalyse werden unterschiedliche Alternativen systematisch und statistisch nach gewissen Kategorien verglichen. Konkret werden hier unterschiedliche Möglichkeiten der Implementierung miteinander verglichen, dabei wird eine Punkteskala von eins bis zehn verwendet. Wichtig bei der Implementation ist es, dass eine höhere Punktzahl nicht bedeutet, wie hoch diese Kategorie ist. Eine Entwicklungsdauer von 10 bedeutet nicht, dass diese besonders lang ist, sondern dass diese Option im Bereich Entwicklungsdauer zehn Punkte erhalten hat (also besonders schnell zum Implementieren ist).

Kriterien		Python mit Tensflow		C++	
Entwicklungsdauer	10 %	10	1	2	0.2
Komplexität	15 %	10	1.5	4	0.6
Sprachkenntnisse	10 %	3	0.3	4	0.4
Frustration	10 %	5	0.5	7	0.7
Typisierung	5 %	1	0.05	9	0.45
Run-Time-Length	10 %	10	1	9	0.9
Testing	5 %	6	0.3	7	0.35
Objektorientiert	5 %	5	0.25	9	0.45
Wissensgewinn	30 %	2	0.6	10	3
Total	100 %	52	5.5	61	7.05

Kapitel 2

Planung

2.1 Rollenverteilung

Rolle	Person
Leiter	Simeon Stix
Entwickler	Simeon Stix
Betreuer	Sven Nüesch

2.2 Aufgabenliste

1. *Planung:*

- Aufgabenliste erstellen
- Zeiteinteilung
- Meilensteine definieren
- Gantt erstellen

2. *Analyse:*

- Recherche CNN
- Recherche Testdaten
- Recherche C++ Details
- Anforderungen definieren
- Use Cases
- weitere Analysen
- Schreiben Analyse

3. *Design:*

- GUI konzipieren
- Wireframes zeichnen
- Konsolenbefehle definieren
- Konsolen Output definieren

4. *Realisation:*

- Input Layer realisieren
- Convolutional Layer realisieren
- Pooling Layer realisieren
- Dense Layer Klasse schreiben
- Output Layer erstellen
- CNN-Model erstellen
- CNN-Model trainieren
- GUI auf CNN-Model setzen
- GPU-Anbindung implementieren
- Kommentare schreiben
- weitere Verbesserungen vornehmen
- Buffer Zeit programmieren

5. *Dokumentieren:*

- Schreiben

6. *Testen:*

- Testliste kreieren
- Tests implementieren
- Tests durchführen

7. *Deployment:*

- Projekt als EXE bauen
- Build Anleitung kreieren

8. *Präsentation:*

- Präsentation erstellen

9. *Buffer allgemein*

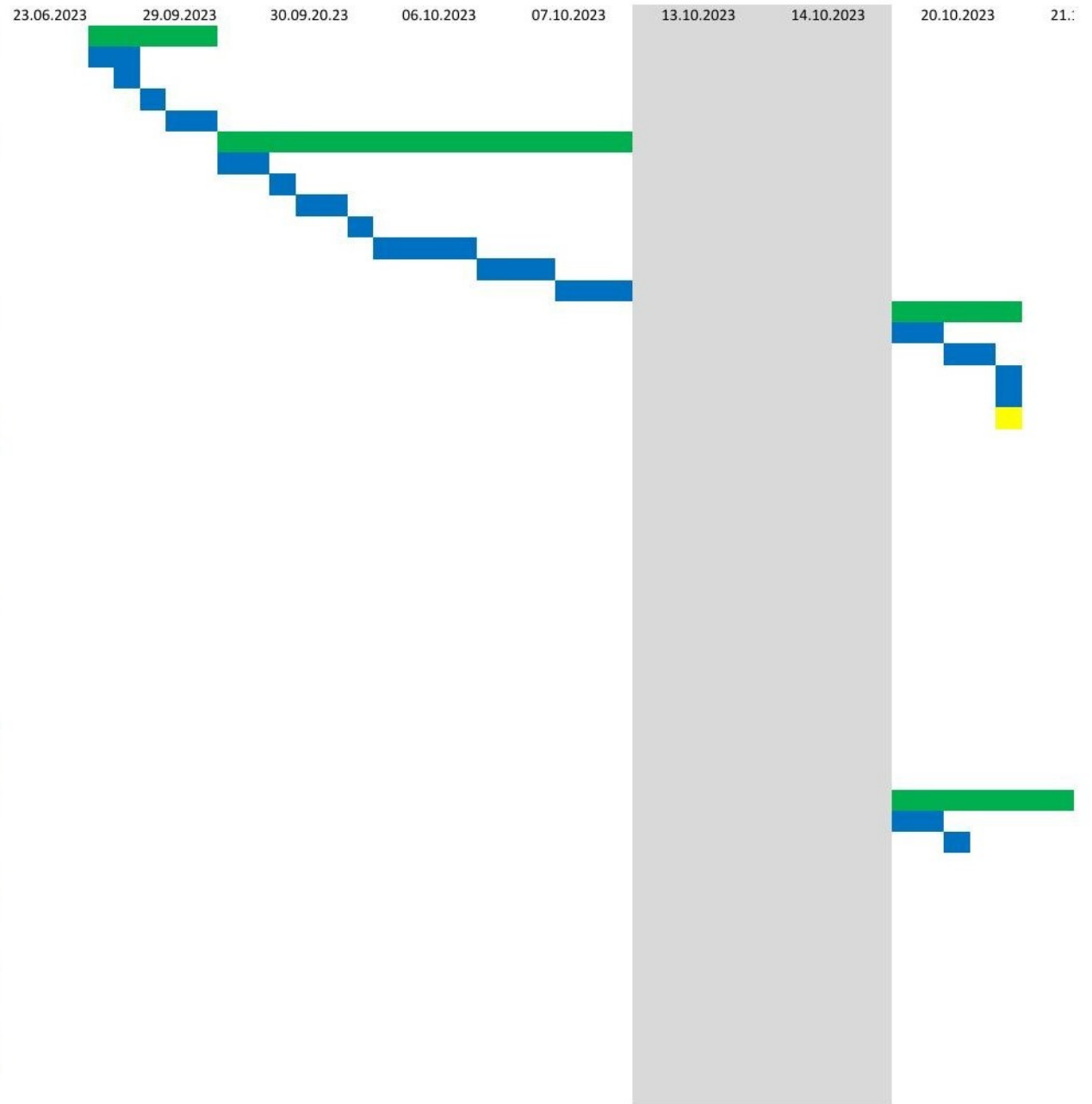
2.3 Meilensteine

Meilenstein-Name	Datum
Vorbereitungen abgeschlossen	20.10.2023
CNN-Model Kreation abgeschlossen	10.11.2023
Dokumentation abgeschlossen	01.12.2023
Produkt abgeschlossen	05.01.2024
Projekt abgeschlossen	10.02.2024
Projekt abgegeben	23.02.2024

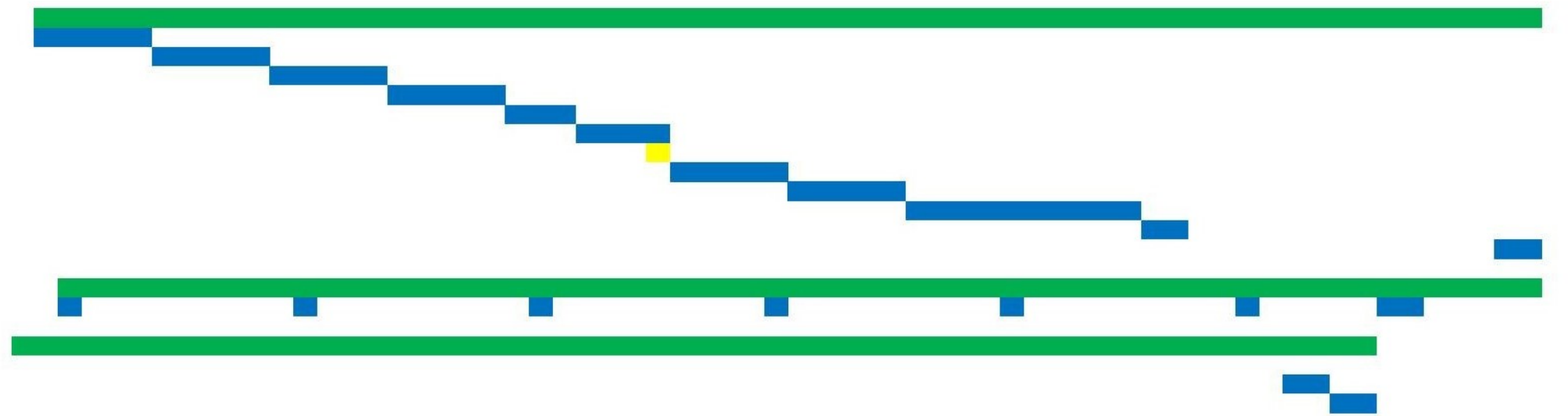
2.4 Gantt

Gross-Tasks	Tasks	Zeit [h]	Datum
Planung		5	
	Taskliste erstellen	1.5	
	Zeiteinteilung	1	
	Meilensteine definieren	0.5	
	Gantt erstellen	2	
Analyse		16	
	Reserche CNN	2	
	Recherche Testdaten	1	
	Recherche C++ Details	2	
	Anforderungen definieren	1	
	Use-Cases	4	
	weitere Analysen	3	
	Schreiben Analyse	3	
Design		5	
	Gui konsipieren	2	
	Wireframes zeichnen	2	
	Konsolenbefehle definieren	0.5	
	Konsolen Output definieren	0.5	
Meilenstein	Vorbereitungen abgeschlossen		20.10.23
Realisation		75	
	Input Layer realisieren	5	
	Convolutional Layer realisieren	5	
	Pooling Layer realisieren	5	
	Dense Layer Klasse schreiben	5	
	Output Layer erstellen	3	
	CNN-Model erstellen	4	
Meilenstein	CNN-Model kreation abgeschlossen		10.11.2023
	CNN-Model trainieren	5	
	GUI auf CNN-Model setzen	5	
	GPU-Anbindung implementieren	10	
	Kommentare schreiben	2	
	weitere Verbesserungen vornehmen	6	
	Bufferzeit Programmieren	20	
Dokumentieren		15	
	Schreiben	15	
Meilenstein	Dokumentation abgeschlossen		16.12.2023
Testen		7	
	Testliste kreieren	2	
	Tests implementieren	3	
	Tests durchführen	2	
Meilenstein	Produkt abgeschlossen		05.01.2024
Deployment		2	
	Projekt als EXE bauen	0.5	
	Build Anleitung kreieren	1.5	
Meilenstein	Projekt abgeschlossen		10.02.2024
Präsentation		3	
	Präsentation erstellen	3	
Buffer Allgemein		44	
Meilenstein	Projekt aggegeben		23.02.2024

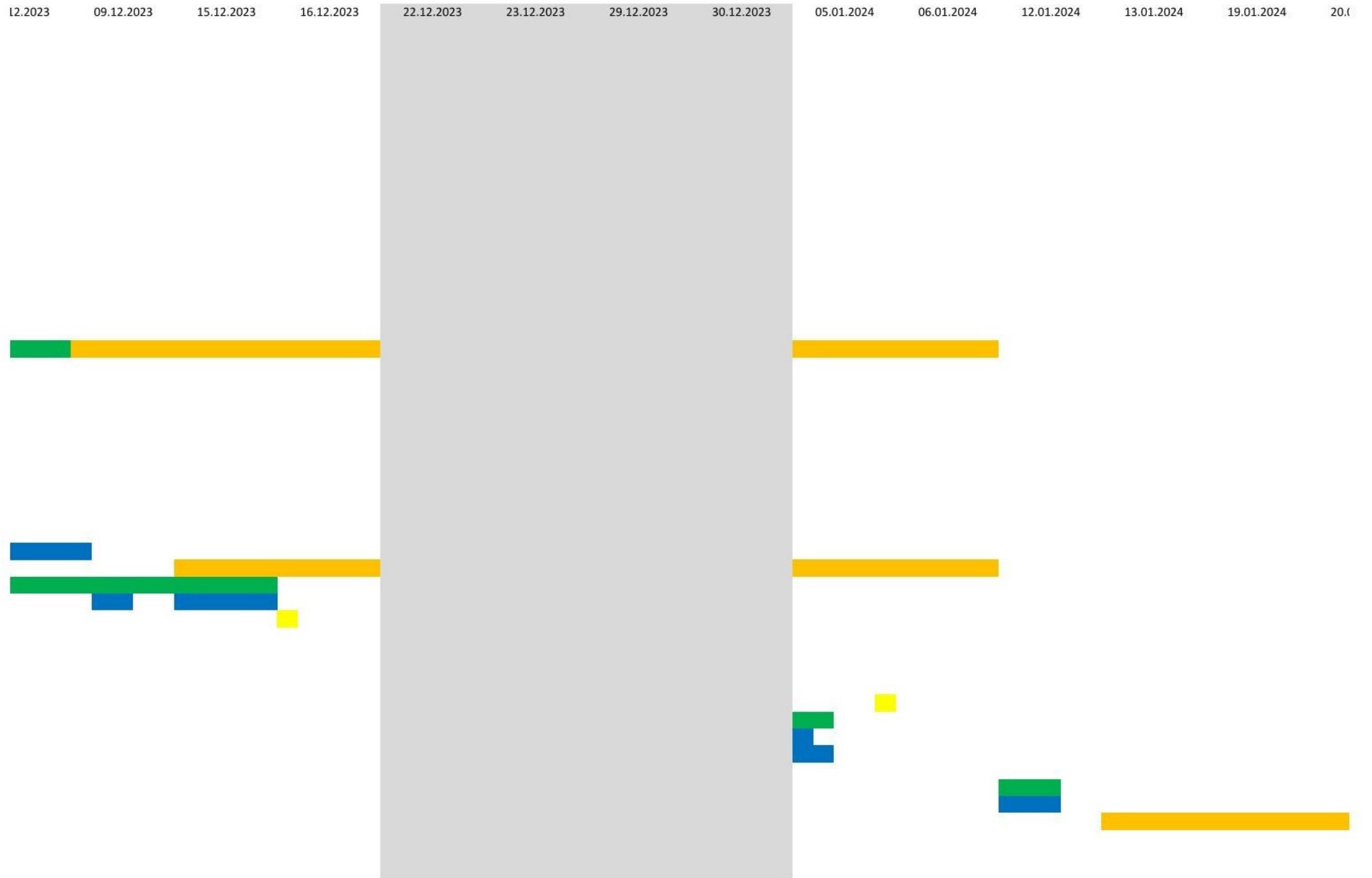
Total 172



10.2023 27.10.2023 28.10.2023 03.11.2023 04.11.2023 10.11.2023 11.11.2023 17.11.2023 18.11.2023 24.11.2023 25.11.2023 01.12.2023 02.12.2023 08.12.2023



12.2023 09.12.2023 15.12.2023 16.12.2023 22.12.2023 23.12.2023 29.12.2023 30.12.2023 05.01.2024 06.01.2024 12.01.2024 13.01.2024 19.01.2024 20.01.2024



01.2024 26.01.2024 27.01.2024 02.02.2024 03.02.2024 09.02.2024 10.02.2024 16.02.2024 17.02.2024 23.02.2024



Kapitel 3

Design

3.1 Konsole

Das Projekt CKi ist in seinem Grundaufbau für die Konsole konzipiert. Der grundlegende Befehl soll wie folgt aussehen:

```
$ ki [options] [file]
```

Dabei gibt es drei Ausführungen von diesem Befehl.

3.1.1 Training

Input

```
$ ki —training [file]  
$ #Dieser Befehl nimmt ubyte-Dateien.
```

Output

Nach jedem Datensatz wird angezeigt, wie viele der zum Training gegebenen Daten bereits abgearbeitet wurden.

3.1.2 Verify

Input

```
$ ki —verify [file]  
$ #Dieser Befehl nimmt ubyte-Dateien.
```

Output

Nach jedem Datensatz wird angezeigt, wie viele der zum Testen gegebenen Daten bereits abgearbeitet wurden. Zudem wird angezeigt, wie viele der Datensatz richtig bzw. falsch erkannt wurden. Eine interessante Erweiterung wäre es, bei diesem Befehl (oder in einer leicht abgewandelten Form) ein kleines Fenster zu öffnen, um dort direkt die Zahl zu zeichnen. Dies hätte den Vorteil, dass Parameter wie Grösse direkt bekannt und kontrolliert werden können.

3.1.3 Prediction

Input

```
$ ki [ file ]  
$ #Dieser Befehl nimmt jpg-Dateien.
```

Output

Als Output werden alle Ziffern von 0 bis 9 mit den entsprechenden Prozentwerten zurückgegeben (Dies kommt daher, dass die KI die Wahrscheinlichkeit der Übereinstimmung für jede Ziffer berechnet.). Zudem wird auch angezeigt, welche Ziffer die höchste Übereinstimmung hat. Da diese die erkannte Ziffer darstellt.

3.2 Datenbank

Die Applikation benötigt keine Datenbank. Zur Speicherung der Test- und Trainings-Datensätze werden sogenannte ubyte-Dateien eingesetzt. Diese enden auf die Dateierweiterung ".ubyte".

3.2.1 UByte

Die Bilddateien im MNIST-Datensatz werden im sogenannten übyte"(unsigned byte) Format gespeichert. Dies bedeutet, dass die Pixelwerte der Bilder als Bytes gespeichert werden. Jedes Bild im Datensatz wird als Reihe von Bytes dargestellt, wobei jedes Byte den Graustufenwert eines Pixels repräsentiert. Hinzu kommt, dass auch die dazugehörigen Labels in einer übyteDatei gespeichert werden.

Aufgrund dieser Speicherung müssen die Datensets vor deren Verwendung eingelesen und interpretiert werden.

Für diese Interpretation muss die standardisierte Höhe und Weite für jedes auszulesende Bild multipliziert werden. Danach müssen so viele Bytes aus der entsprechenden übyteDatei ausgelesen werden und als vorzeichenlose ganze Zahlen gespeichert. Denn jede dieser Zahlen stellt einen Pixel dar.

Für eine genaue Anleitung für das Einlesen und Interpretieren in Python ist die Webseite AndroidKT zu kontaktieren.

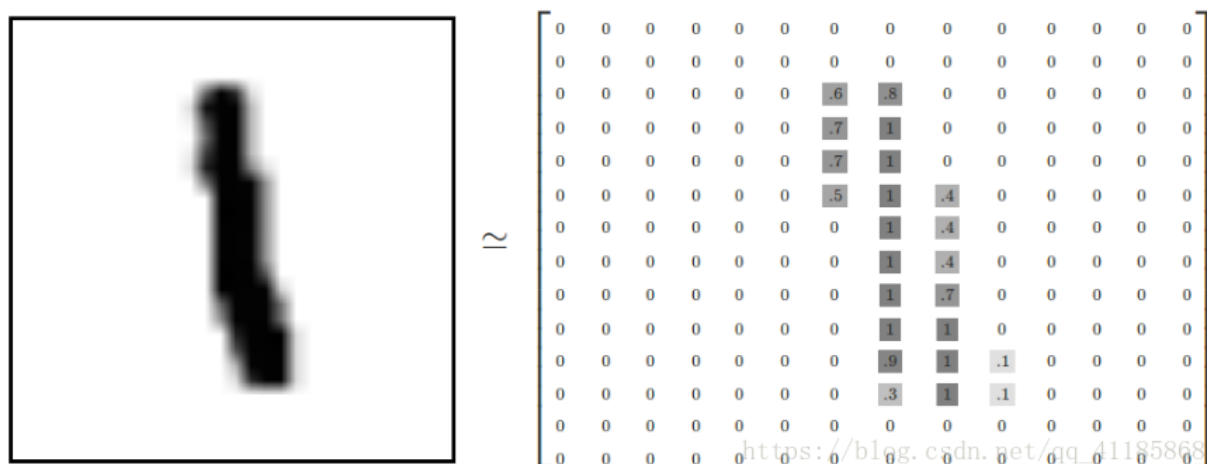


Abbildung 3.1: Darstellung einer Ziffer und deren Code von blog.csdn.net

3.3 Code

3.3.1 Klassendiagramm

Es gibt drei Klassen in CKi (Convolutional Layer, Pooling Layer, Dense Layers). Ein Convolutional Layer ist eine Art von Filter, welcher auf das Bild gelegt wird, um eine vereinfachte Erkennung zu ermöglichen. Die Pooling Layers sind verantwortlich das Bild zu verkleinern und so eine schnellere Verarbeitung zu ermöglichen. Die Dense Layers oder auch Fully Connected Layers sind das eigentlich das Gehirn der KI. Zusätzlich gibt es die Main-Klasse, diese ist zuständig die oben genannten Klassen miteinander zu verbinden und so das Netzwerk aufzubauen. Zusätzlich handhabt es die Nutzereingaben und die Auslese der Trainings-/Testdaten.

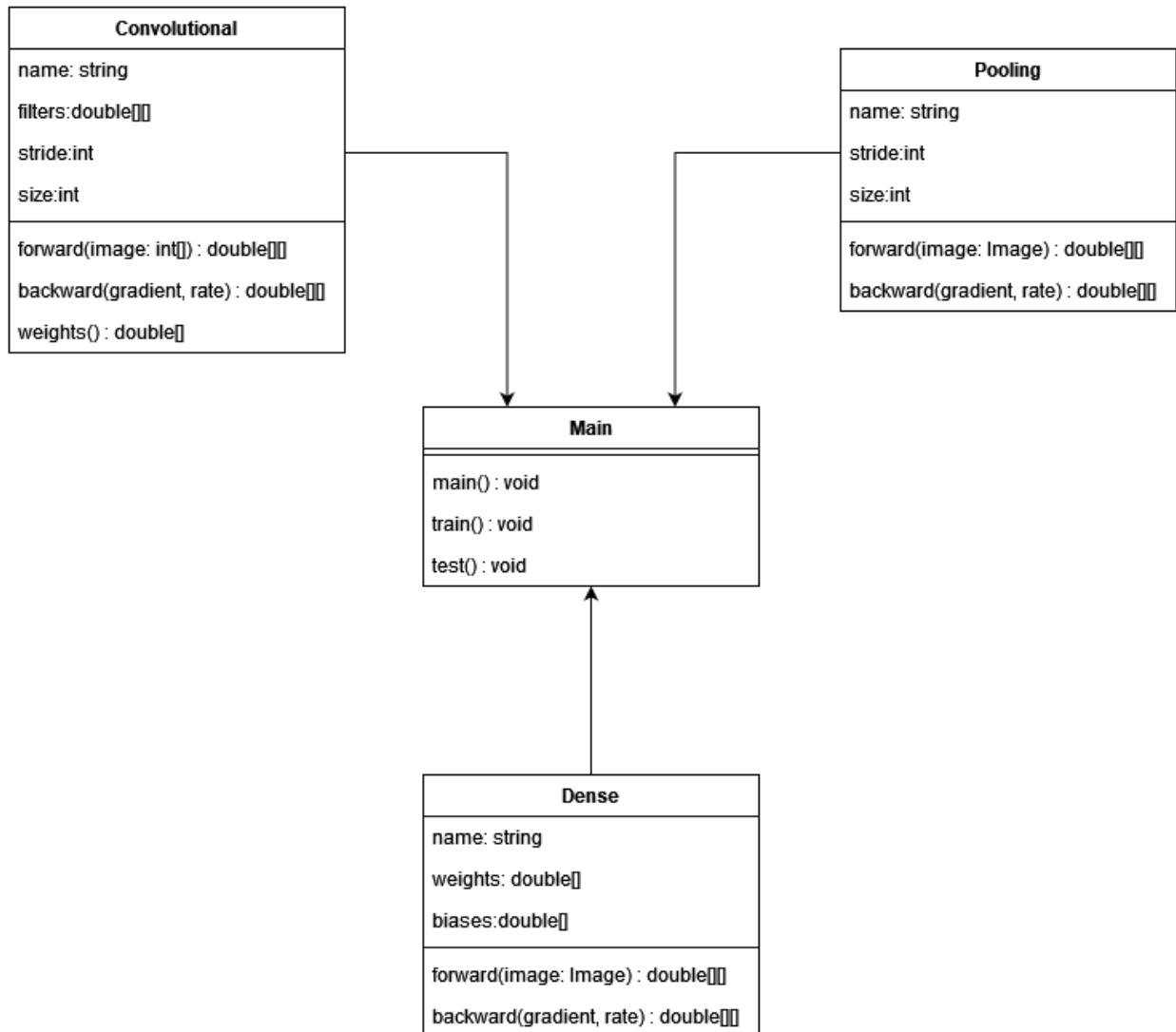


Abbildung 3.2: Das Klassendiagramm, Grundaufbau der Applikation

3.3.2 Trainingsdaten

Die Trainingsdaten sind das MNIST-Datenset mit den handschriftlichen Zahlen. (<http://yann.lecun.com/exdb/mnist/>)

3.3.3 Tests

Im Projekt CKI gibt es drei Arten von Tests. Es gibt das simple Ausprobieren. Da es nicht allzu viele Nutzerschnittstellen gibt, kann man diese ausführen und begutachten, ob diese mit der Beschreibung übereinstimmen. Bei einer dieser Schnittstellen wird die KI getestet. Dies geschieht, indem der KI ihr unbekannte Datensätze zu sehen bekommt und

das Endergebnis mit einem vordefinierten Ergebnis abgeglichen wird. Egal ob das Ergebnis korrekt oder inkorrekt erkannt worden ist, wird es statistisch aufgenommen. Am Ende wird dem Nutzer eine Prozentzahl der korrekten Erkennungen präsentiert. Diesbezüglich ist dies kein Test, in dem die Applikation versagen könnte, es ist eine reine Leistungsüberprüfung, ob mehr Training vonnöten ist. Zusätzlich zu diesen zwei Testmöglichkeiten gibt es noch die Unit-Tests. Diese werden dazu genutzt, um einzelne Funktionen und Klassen noch vor deren Verwendung zu überprüfen.

Zusätzlich gegenüber diesen drei Tests wäre es überaus interessant zu begutachten und zu vergleichen, wie hoch die Leistungsunterschiede zwischen CKi und einem in Python mit TensorFlow gebauten Programm mit vergleichbarer Grundstruktur sind. Diesbezüglich wird neben CKi auch ein kleines vergleichbares Modell in Python realisiert, um diese Daten zu sammeln.

Literaturverzeichnis

- [1] AndroidKT. Extract images from mnist idx3 ubyte file format in python, 11.06.2023. Accessed on 23 November 2023.
- [2] ibm. What are convolutional neural networks? Accessed on 23 November 2023.
- [3] Amrita Pathak. Convolutional neural networks (cnns): Eine einföhrung, 24.09.2023. Accessed on 23 November 2023.
- [4] Sumit Saha. A comprehensive guide to convolutional neural networks — the eli5 way, 15.12.2018. Accessed on 23 November 2023.
- [5] Christopher J.C. Burges Yann LeCun, Corinna Cortes. The mnist database of hand-written digits, 22.06.2002. Accessed on 23 November 2023.

Abbildungsverzeichnis

1.1	CNN Model Aufbau Grafik von researchgate.net	4
1.2	Use-Case-Diagramm	10
1.3	Ablaufdiagramm für die Interpretation von Nutzer-gelieferten Einzelbildern	11
1.4	Ablaufdiagramm vom Training des CNN	12
1.5	Ablaufdiagramm von einem Test des CNN	13
1.6	Visuelle Darstellung der Funktionsweise eines CNN von blog.goodaudience.com	14
3.1	Darstellung einer Ziffer und deren Code von blog.csdn.net	28
3.2	Das Klassendiagramm, Grundaufbau der Applikation	29