

BLACK BOX KI



KÜNSTLICHE INTELLIGENZ – WAS STECKT DAHINTER?

Künstliche Intelligenz (KI) ist ein Bereich der Informatik, der Maschinen dazu befähigt, Aufgaben zu übernehmen, für die normalerweise menschliche Intelligenz erforderlich ist. Dazu gehören Mustererkennung, Entscheidungsfindung und das Lernen aus Daten. Ein gutes Beispiel dafür ist eben unsere Alters- und Geschlechtsbestimmung anhand von Handbildern. Anstatt jede Hand einzeln manuell zu analysieren, kann eine KI tausende von Handbildern auswerten, Gemeinsamkeiten entdecken und daraus Vorhersagen ableiten. Doch bevor die KI die Vorhersage treffen kann, müssen die Daten entsprechend vorbereitet werden – hier kommt die Normalisierung ins Spiel.

NORMALISIERUNG – WARUM BILDER VERGLEICHBAR SEIN MÜSSEN

Stell dir vor, du möchtest das Alter von Händen erkennen, aber die Bilder haben unterschiedliche Größen, Helligkeiten oder Farbtöne. Das kann die KI verwirren! Deshalb gibt es die Normalisierung.

WARUM IST DAS WICHTIG?

Wenn wir Vorhersagen anhand von Händen machen, könnten manche Bilder heller wirken, weil sie bei Tageslicht aufgenommen wurden, während andere dunkler sind. Auch Größenunterschiede könnten das Modell verwirren.

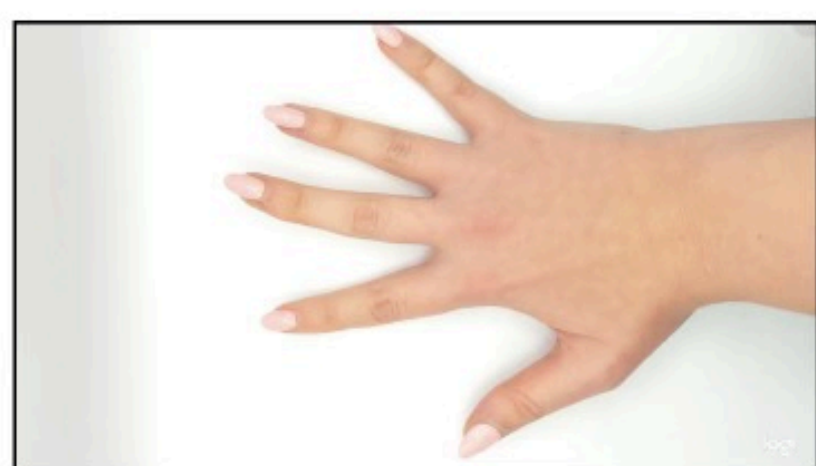
LÖSUNG:

Die KI normalisiert die Bilder, indem sie:

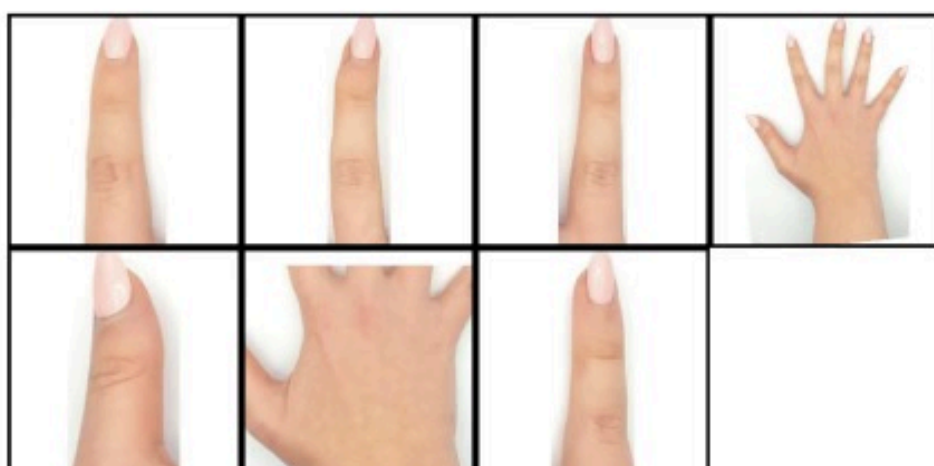
- ✓ Helligkeit und Kontraste anpasst
- ✓ Die Bilder auf eine einheitliche Größe bringt
- ✓ Farbwerte standardisiert
- ✓ Rotationen vereinheitlichen

Das sorgt dafür, dass alle Bilder unter ähnlichen Bedingungen betrachtet werden. HandScan AI muss jedoch nur die Bilder auf eine einheitliche Größe bringen und richtig rotieren, da vieles bereits durch das Kamera SetUp abgefangen wird. Zusätzlich wurde die Hand in sieben Regionen unterteilt, damit die HandScan AI einzelne Bereiche – wie Daumen oder Zeigefinger – separat analysieren kann. Dadurch können Unterschiede innerhalb einzelner Handbereiche gezielt erfasst und die Vorhersage auf einer detaillierteren Ebene durchgeführt werden.

1. AUFGENOMMES BILD



2. NORMALISIERTE UND SEGMENTIERTE REGIONEN



Sobald die Daten vorbereitet sind, kann die KI daraus Embeddings erstellen – eine mathematische Repräsentation der Hand.

EMBEDDINGS – WENN HÄNDE IN ZAHLEN VERWandelt WERDEN

Für uns ist eine Hand eine Kombination aus Fingern, Handfläche und Gelenken. Doch für eine KI ist sie zunächst nur ein Bild mit vielen Pixeln. Damit die KI dieses Bild verarbeiten kann, muss sie es in eine mathematische Form bringen – hier kommen Embeddings ins Spiel.

WAS SIND EMBEDDINGS?

Ein Embedding ist eine kompakte Zahlenrepräsentation eines Objekts in Form eines Vektors. Die Embeddings werden nicht manuell bestimmt, sondern von einem weiteren KI-Modell berechnet: einem Convolutional Neural Network (CNN). Dieses neuronale Netz ist darauf spezialisiert, Muster in Bildern zu erkennen. Es analysiert die Hand, extrahiert relevante Merkmale und wandelt sie in eine kompakte Zahlenform um – das Embedding.

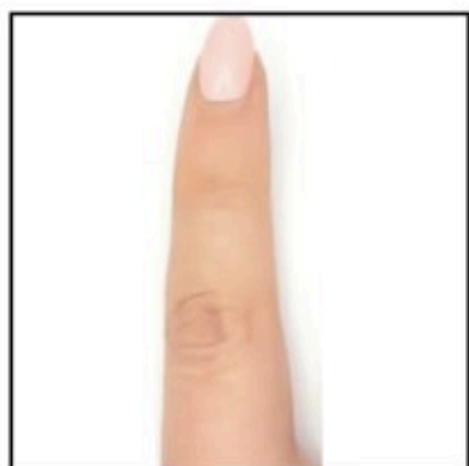
WIE VERGLEICHT DIE KI ZWEI HÄNDE? – KOSINUS DISTANZ

Sobald die Hände bzw. die Regionen in Zahlenform (Embeddings) vorliegen, kann die KI deren Ähnlichkeit berechnen.

Dafür nutzen wir die Kosinus-Distanz. Statt einfach nur die Abstände zwischen zwei Punkten zu messen, wie es bei der euklidischen Distanz der Fall wäre, schaut sich die Kosinus-Distanz an, wie ähnlich zwei Vektoren in ihrer Richtung sind.

Stell dir vor, du hast zwei Pfeile auf einem Blatt Papier gezeichnet. Je mehr sie in die gleiche Richtung zeigen, desto ähnlicher sind sie – selbst wenn einer länger oder kürzer ist als der andere. Die Kosinus-Distanz misst genau diesen Winkel zwischen den beiden Pfeilen. Wenn zwei Hand-Embeddings fast in die gleiche Richtung zeigen, bedeutet das, dass sie ähnliche Merkmale haben – also vermutlich von ähnlich alten Personen gleichen Geschlechtes stammen. Ist der Winkel zwischen den Vektoren groß, unterscheiden sich die Hände stärker voneinander.

1. SEGMENTIERTES BILD

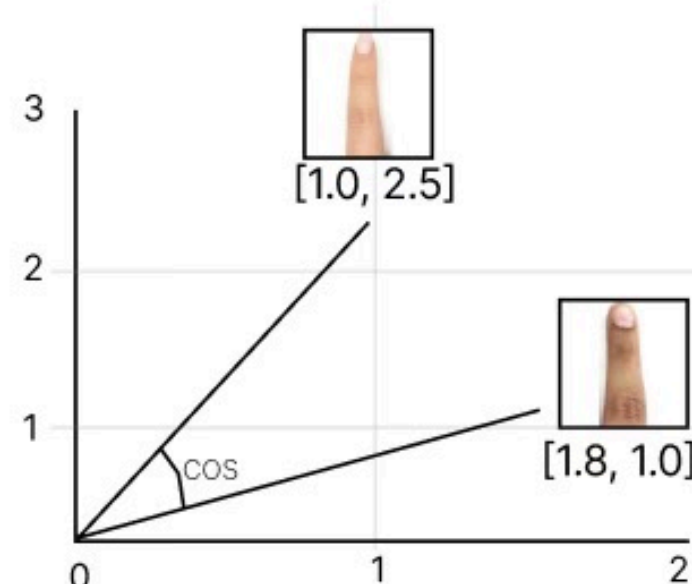


2. ZAHLENREPRÄSENTATION (EMBEDDING)

Zeigefinger → [1.0, 2.5]

vereinfachte Darstellung

3. DISTANZBERECHNUNG



K-NEAREST NEIGHBOR ALGORITHMUS – WIE DIE KI ÄHNLICHE HÄNDE FINDET

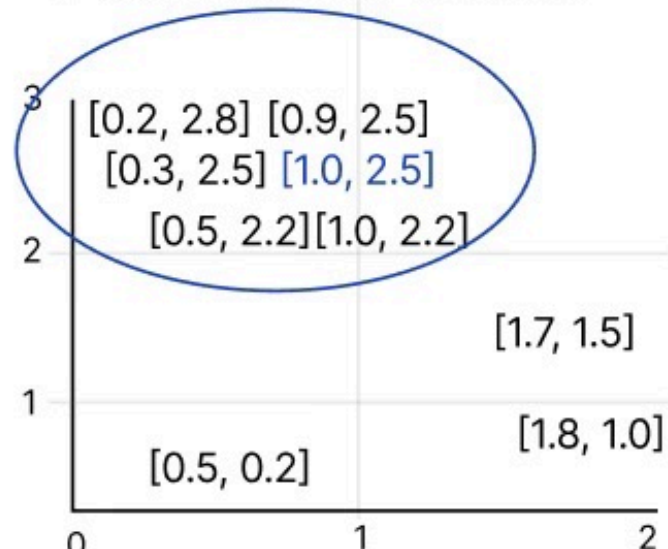
Der k-nearest neighbor Algorithmus (k-NN) ist einer der einfachsten und effektivsten Methoden, um eine Vorhersage für das Alter und das Geschlecht zu treffen. HandScan AI hat bereits viele Handbilder gesehen – für jede dieser Hände bzw. Regionen liegen Embeddings, das konkrete Alter und das Geschlecht vor. Diese Daten sind in einer großen Sammlung gespeichert.

Nachdem die Kosinus-Distanz zwischen den Hand-Embeddings berechnet wurde, sucht die KI die k nächsten Nachbarn – also die Hände, die den kleinsten Winkelunterschied zum neuen Handbild haben. In dem Beispiel unten sowie bei HandScan AI werden die fünf ähnlichsten Nachbarn berücksichtigt, indem k=5 gewählt wurde.

SO FUNKTIONIERT ES:

1. Die HandScan AI speichert alle Handregionen als Embeddings in einer Datenbank.
2. Wenn eine neue Hand analysiert wird, berechnet sie die Distanz zu allen gespeicherten Handabschnitten.
3. Sie sucht die k nächsten Nachbarn (die k ähnlichsten Hände).
4. Wenn die meisten dieser Nachbarn z. B. einer bestimmten Altersgruppe angehören, ordnet die KI die neue Hand ebenfalls dieser Gruppe zu.
 - a. Für das Alter: Hier wird ein gewichteter Mittelwert über das Alter der Nachbarn in Abhängigkeit der Distanz gebildet.
 - b. Für das Geschlecht: Hier wird der Modus in Abhängigkeit der Distanz gewählt.

1. DISTANZVERGLEICH



2. NÄCHSTEN NACHBARN

1. Hand, weiblich, 28
2. Hand, weiblich, 30
3. Hand, männlich, 28
4. Hand, weiblich, 22
5. Hand, männlich, 23

Je nachdem wie groß k gewählt wurde, werden hier k nächste Nachbarn gesucht (hier beispielhaft k=5)

3. VORHERSAGE

Alter durch gewichteten Mittelwert:
28

Geschlecht durch Modi:
weiblich

gewichteter Mittelwert: 22, 23, 28, 28, 30 → 28
Modi: 3x weiblich, 2x männlich → weiblich

Damit die KI gute Vorhersagen treffen kann, braucht sie allerdings eine große Menge an Trainingsdaten.

TRAININGSDATEN – DAS FUNDAMENT EINER GUTEN KI

KI-Systeme sind nur so gut wie die Daten, mit denen sie trainiert wurden. Für die Alters- und Geschlechtsbestimmung von Händen haben wir unser Modell mit dem 11K Hands Dataset (Mahmoud Afifi, "11K Hands: Gender recognition and biometric identification using a large dataset of hand images." Multimedia Tools and Applications, 2019) trainiert – einer Sammlung von über 11.000 Handbildern, bei denen das Alter und Geschlecht bereits bekannt ist.

Da wir uns auf die dorsale Handansicht und nicht auf die Handinnenseite konzentrieren und eine konsistente Normalisierung sicherstellen wollten, haben wir den Datensatz entsprechend gefiltert.

DATENVERTEILUNG NACH FILTERUNG:

- 78 Personen, 40 % Männer, 60 % Frauen
- Alter: 20 bis 75 Jahre, 78 % zwischen 20 und 22 Jahren
- Männer: 70 % linke Hand, 30 % rechte Hand
- Frauen: 56 % linke Hand, 44 % rechte Hand

Einfluss auf das Modell – BIAS:

- Fokus auf junge Erwachsene (Bias in der Altersverteilung)
- Mehr Frauen als Männer (geschlechtsspezifischer Bias)
- Mehr linke als rechte Hände, besonders bei Männern

Da der Großteil der Trainingsdaten von jungen Erwachsenen stammt, erkennt HandScan AI das Alter in diesem Bereich besonders präzise. Für ältere Personen könnte die Vorhersage ungenauer sein, da hierfür weniger Vergleichsdaten vorliegen. Ebenso beeinflusst die ungleiche Verteilung der Geschlechter die Genauigkeit: Da mehr weibliche als männliche Hände im Datensatz enthalten sind, kann das Modell weibliche Hände möglicherweise besser klassifizieren. Zudem könnte die höhere Anzahl an linken Händen dazu führen, dass die Vorhersagen für rechte Hände weniger exakt ausfallen.

Mit jeder neuen Anwendung wird unser Datensatz erweitert, da zusätzliche Handbilder mit bekannten Alters- und Geschlechtsangaben hinzukommen. Der Bias lässt sich langfristig ausgleichen, da HandScan AI durch Personen mit unterschiedlichem Alter und Geschlecht genutzt wird. Dadurch wird das Modell kontinuierlich präziser und robuster und kann für eine größere Nutzergruppe verwendet werden.

Und das Beste: Mit jeder neuen Anwendung wächst unser System! Jedes neue Handbild hilft der KI, noch präzisere Vorhersagen zu treffen. Je mehr Daten sie sieht, desto besser wird sie. Daher ist es großartig, dass du die Anwendung genutzt hast!



FAZIT

1. Künstliche Intelligenz hilft uns, Muster in Daten zu erkennen – zum Beispiel das Alter aus Handbildern.
2. Normalisierung sorgt dafür, dass alle Bilder vergleichbar sind.
3. Embeddings wandeln Hände in Zahlenvektoren um, sodass die KI sie mathematisch analysieren kann.
4. Der k-nearest neighbor Algorithmus findet ähnliche Hände in der Datenbank, um Alters- und Geschlechtsprognosen zu machen.
5. Trainingsdaten sind die Basis für eine zuverlässige KI – je mehr und besser sie sind, desto präziser die Vorhersagen.