

# Klassifizierung von Früchten

Henry Apostle, Sibel Gül, Robert Johanson und Darwin Willers

20. Februar 2020

## Zusammenfassung

## 1 Einleitung

## 2 Daten

Der verwendete Datensatz stammt von zenodo.org. Ursprünglich bestand er aus 44.406 Bildern verteilt auf 15 Klassen. Da wir eine zusätzliche Unterscheidung zwischen grünen und roten Äpfeln machten, kamen wir auf 16 Klassen. Pro Klasse haben wir anschließend 1.000 Bilder zufällig ausgewählt und somit den Datensatz auf eine Größe von 16.000 Bildern reduziert. Die 16 Klassen sind: grüner Apfel, roter Apfel, Banane, Sternfrucht, Guave, Kiwi, Mango, Zuckermelone, Orange, Pfirsich, (Nashi-)Birne, Kaki, Drachenfrucht, Pflaume, Granatapfel und Tomate. Abbildung 1 zeigt je ein exemplarisches Bild aus jeder Klasse. Alle Bilder sind 320x258 Pixel groß. Weiterhin ist allen Bildern gemeinsam, dass das Obst auf einem grau/silbernen Blech abfotografiert wurde. Unterschiede gibt es in der Belichtung, der Anzahl und der Anordnung der Früchte auf dem Blech. Des Weiteren wurden einige Bilder im Nachhinein gedreht und verzerrt [genauer], während auf anderen noch eine Hand zu sehen ist (siehe Abbildung 2). Auffallend am Datensatz ist, dass es sich bei mehreren Klassen um rundes grünes Obst handelt, nämlich bei grüner Apfel, Guave, Zuckermelone, Birne und oftmals auch Granatapfel. Bei der Zuckermelone gibt es manche Bilder, die wir als grünen Apfel klassifiziert hätten (z.B siehe Abbildung 3). Bei der Birne handelt es sich um die Nashi Birne, die eben nicht der gewöhnlichen Form einer Birne entspricht und die man optisch auch leicht mit grünen Äpfeln verwechseln kann. Des Weiteren sind die Granatäpfel auf den Bildern oft sehr unreif und somit eher grün.

## 3 Methodik

Als wesentliches Merkmal zur Unterscheidung der Früchte dient uns die Farbe. Form haben wir an dieser Stelle vernachlässigt, da auf den Bildern das Obst

in unterschiedlicher Anzahl und Anordnung gegeben ist. Manchmal liegen die Früchte nebeneinander, manchmal sind sie sogar gestapelt. Somit müsste man erst ein Programm schreiben, dass eine einzelne Frucht lokalisiert, um dann anschließend ihre Form auszuwerten. Dies fanden wir zu kompliziert, besonders da das alleinige Betrachten der Farbe schon für sehr gute Ergebnisse ausreicht. Unser Programm durchläuft mehrere Schritte: Zuerst wird das Obst vom Hintergrund mittels Binarisierung getrennt, anschließend wird das 'ausgeschnittene' Bild mittels 3D-Histogramm in farbliche Komponenten zerlegt und zuletzt wird mittels Nearest-Neighbour-Klassifikator bestimmt zu welcher Obstsorte das Bild gehört. Im Folgenden gehen wir auf die einzelnen Schritte genauer ein:

### 3.1 Das Ausschneiden

Um das Obst von dem Blech zu unterscheiden, haben wir beobachtet, dass die Farben der Früchte deutlich saturierter sind, als das Grau/Silber des Blechs. Deshalb transformiert unser Programm das Bild zuallererst in den HSV-Raum, wo wir dann nur den S-Kanal (die Saturation) betrachten. Darauf wenden wir anschließend das Otsu-Verfahren an. [Otsu macht diesdas]. [siehe Abbildung 4]

### 3.2 3D-Histogramm

[So funktioniert ein 3D-Histogramm]

[Das Normieren erwähnen]

### 3.3 Nearest-Neighbor Klassifikator

[So funktioniert der NN Klassifikator]



Abbildung 1: Die 16 verschiedenen Früchte, die klassifiziert werden. Von oben links nach unten rechts sind das: grüner Apfel, roter Apfel, Banane, Sternfrucht, Guave, Kiwi, Mango, Zuckermelone, Orange, Pfirsich, Birne, Kaki, Drachenfrucht, Pflaume, Granatapfel und Tomate.

## 4 Experimente

Pro Klasse hatten wir 800 Trainings- und 200 Testbilder. Insgesamt also 12.800 Trainings- und 3.200 Testbilder. Unser Programm hat die 3.200 Testbilder zu 96,34% korrekt zugeordnet. Im Folgenden gehen wir die einzelnen Komponenten unseres Systems durch und schauen, wie sich die Trefferquote ändert, wenn wir sie austauschen:

Wenn wir die Bilder nicht ausschneiden, sondern direkt in 3D-Histogramme verwandeln und mit dem Nearest-Neighbour Klassifikator klassifizieren erhalten wir immer noch eine Trefferquote von 93,09%. Tatsächlich selbst wenn wir den 3D-Histogramm durch den kanalweisen Mittelwert ersetzen, erhalten wir immer noch eine Trefferquote von 51,28% auf den nicht ausgeschnittenen Bildern. Dies erklären wir uns dadurch, dass wir einerseits eine sehr hohe Menge an Trainingsbildern haben und andererseits die Bilder im Großen und Ganzen doch ziemlich ähnlich sind. Wenn wir aber den Mittelwert aber auf den ausgeschnittenen Bildern anwenden liegt die Trefferquote sogar bei 77,5%. Dies zeigt, dass es durchaus etwas bringt, die Bilder auszuschneiden.

[Andere Deskriptoren betrachten]

[Andere Klassifikatoren (k-Nearest Neighbour) betrachten]

[Man kann dann auch Bilder zeigen, die unser Programm falsch erkennt. Die sind im Ergebnisse.pdf aufgelistet.]

## 5 Schluss

## Literatur

<https://zenodo.org/record/1310165>

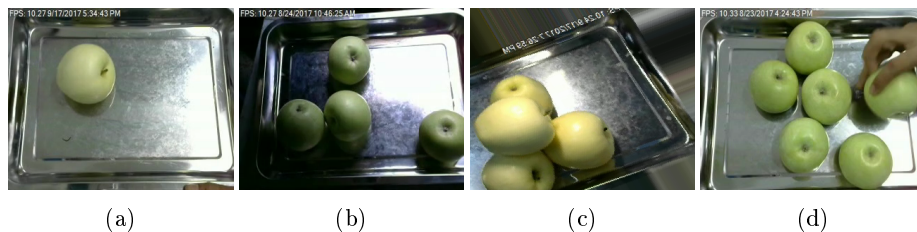


Abbildung 2: Auf diesen vier Bildern sind Unterschiede zu erkennen, die man innerhalb einer Klasse begegnen kann. Alle Bilder zeigen grüne Äpfel. Bei (a) und (b) erkennt man Unterschiede in der Anzahl der Äpfel und in der Belichtung, (c) wurde gedreht und verzerrt und auf (d) ist eine Hand zu sehen.

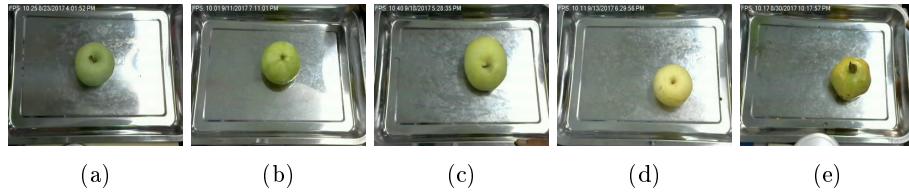


Abbildung 3: Hier sind fünf grüne runde Früchte zu sehen, welche eine Gefahr für die Klassifikation sein könnten: (a) grüner Apfel, (b) Guave, (c) Zuckermelone, (d) Birne und (e) Granatapfel.

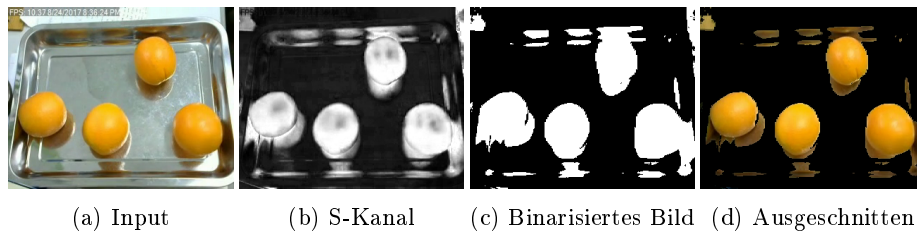


Abbildung 4: Auf diesen vier Bildern ist der Prozess des Ausschneidens verdeutlicht. Bild (a) zeigt ein Beispielbild. Dieses wird in den HSV-Raum transformiert und Bild (b) zeigt den entsprechenden Saturation Kanal. Darauf wird das Otsu-Verfahren angewandt und es entsteht das binarisierte Bild (c). Aus (a) und (c) lässt sich das ausgeschnittene Bild (d) ableiten.