

Projektbericht: Skin Cancer Detection

Florence Lopez, Jonas Einig, Julian Späth
Department of Computer Science, University of Tübingen

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Einleitung

Hautkrebs gilt als eine der häufigsten Krebserkrankungen der Welt. Jährlich erkranken etwa 18.000 Menschen in Deutschland an dieser Krankheit, wobei Hautkrebs allgemein etwa für ein Prozent der Krebstodesfälle verantwortlich ist Krebsgesellschaft (2012). Findet eine Erkennung der malignen Hautläsionen frühzeitig statt, so ist es in den meisten Fällen möglich einen tödlichen Verlauf der Krankheit zu verhindern. Daher sind frühzeitige Erkennungssysteme sehr wichtig für die Bekämpfung von Hautkrebs.

Eine Ergänzung zum regelmäßigen Arztbesuch und dem damit verbundenem Hautscreening, können daher neuronale Netze bieten, die aufgrund von medizinischen Datenbanken lernen können, eine maligne von einer benignen Hautläsion zu unterscheiden. Die genaue Implementierung und das Training solcher neuronalen Netzen werden wir im Folgenden genauer erklären.

Problemstellung und Zielsetzung

Im Rahmen des Machine Learning Praktikums beschäftigen wir uns mit der folgenden Problemstellung: unser Ziel ist es, einen Klassifikator zu entwickeln, der Bilder von Hautläsionen in maligne und benigne Läsionen unterteilen kann. Maligne Hautläsionen sind die

Läsionen, die für den Menschen gefährlich bis sogar tödlich werden könne, während benigne Hautläsionen die gutartigen Läsionen darstellen. Unser Projekt basiert dabei auf der Arbeit von Esteva et al. (2017), wobei wir die originale Problemstellung jedoch etwas abgewandelt haben. Während Esteva et al. (2017) viele verschiedene Arten von Hautläsionen unterschieden haben, wollen wir lediglich zwischen zwei Klassen, nämlich den gutartigen und den bösartigen Läsionen, unterscheiden, was die Problemstellung etwas vereinfacht.

Unser Ziel war es eine möglichst hohe Genauigkeit zu erreichen und vor allem die Anzahl der falsch Negativen möglichst gering zu halten. Denn im Zweifel soll der Klassifikator eine Läsion als maligne klassifizieren, auch wenn sie eigentlich benigne ist, anstatt eine maligne Läsion, die tödlich verlaufen könnte, zu verharmlosen und als benigne zu klassifizieren. Im Folgenden werden wir genauer auf die Methodik eingehen, die hinter unserem Klassifikator steckt und welche Ergebnisse dieser auf unbekannte Bilder von Hautläsionen liefert.

Methoden und Tools

- Python, Tensorflow, Scikit Learn, NumPy
- Skalierung der Bilder
- Augmentierungsmethoden
- Trainingsparameter (Anzahl Durchläufe, Lernrate, Loss-Funktionen)
- Evaluationsmethoden: Berechnung des Scores, etc.
- Aufteilung des Datensatzes in Training, Test, Validierung
- genaue Erklärung des Shufflings, da Datensatz nicht ausbalanciert

Ergebnisse

- besten Trainingsdurchlauf genau beschreiben → Lernrate, Loss, ...
- Evaluierungswerte aus diesem Durchlauf nennen und beschreiben bzw. erklären
- viele Bildchen :-)
- Bilder aus Tensorboard
- Matlab-Plots und Tabellen

Analysemethoden

- MCC, F1, F2 → genaue Beschreibung, was Scores machen
- threshold-Bestimmung

MCC**Genauigkeit****Aussicht****Einbindung in mobile App**

- Einbindung in App → unsere Fortschritte dort beschreiben
- Verknüpfung des Klassifikators mit speziellen Kameras fürs Handy
- Vorteil von ML in Medizin: Bewusstsein für Skin Cancer kann erhöht werden, einfachere Methode grad in ländlichen Gegenden (wo es nicht viele Ärzte gibt)
- genauere Unterscheidung wie auch im Paper noch möglich (dass einzelne Subtypen auch erkannt und unterschieden werden)

Literatur

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., and Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115.

Krebsgesellschaft, D. (2012). Patientenratgeber Hautkrebs.