

Projektbericht: Skin Cancer Detection

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

In der Medizin unterscheidet man bei Hautverletzungen oder Hautveränderungen zwischen benignen und malignen Läsionen. Die malignen Läsionen gelten dabei als die bösartigen Hautveränderungen, die auf Hautkrebs hindeuten. Wir entwickelten daher auf der Grundlage einer anderen Arbeit einen Klassifikator, der aus einem Datensatz von Hautläsionen zwischen Benignen und Malignen unterscheidet. Durch diesen Klassifikator soll es zukünftig einfacher sein, bösartige Hautläsionen schneller zu erkennen, um diese effizient zu behandeln. Daher arbeiteten wir zusätzlich noch an einer mobilen Anwendung, welche es Nutzern ermöglicht, ihre Haut vorerst ohne eine ärztliche Analyse nach malignen und benignen Hautläsionen zu untersuchen. Natürlich kann eine solche Anwendung aber keinen Mediziner ersetzen, sondern lediglich als zusätzliche Unterstützung dienen.

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

Zur Erstellung des Klassifikators nutzten wir, wie Esteva et al. (2017) auch das öffentlich zur Verfügung stehende GoogleNet Inception v3, welches ein convolutional neural network ist und aus zahlreichen verschiedenen Schichten und Neuronen besteht (Szegedy et al., 2016). Dieses neuronale Netz wurde auf den Bildern des ImageNet vortrainiert (Russakovsky et al., 2015) und eignet sich daher gut für unser Klassifizierungsproblem. Das Training und die damit verbundene Vorverarbeitung programmierten wir mittels Python, in Kombination mit Tensorflow, Scikit Learn und NumPy.

Als Datensatz nutzten wir die ISIC-Datenbank, welche aus insgesamt 13.768 Bildern von sowohl malignen als auch benignen Hautläsionen besteht und außerdem der einzige frei erhältliche Hautläsions-Datensatz war. Leider stellte sich heraus, dass die Bilder dieser Datenbank sehr unausgeglich waren: der Anteil der malignen Läsionen war viel geringer als der Anteil der benignen Läsionen. Dieses Problem lösten wir durch eine spezielle Art der Randomisierung, welche später noch genauer erläutert wird. Ein weiteres Problem, das die Bilder des ISIC-Datensatzes brachten, war die Tatsache, dass sie verschieden groß waren. Für das richtige Trainieren des Netzes, mussten wir sie daher vorverarbeiten und wir entschieden uns dazu, alle auf die kleinste verfügbare Größe runterzuskalieren. Zusätzlich verkleinerten wir den Ausschnitt des Bildes, indem wir das maximale zentrale Quadrat auswählten und auf 299x299 px runterskalierten. Somit hatten wir eine homogene Menge an Bildern, die wir zu Trainingszwecken in das neuronale Netz füttern konnten.

*** HIER DANN AUGMENTIERUNGSMETHODEN ***

- 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.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., et al. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2818–2826.