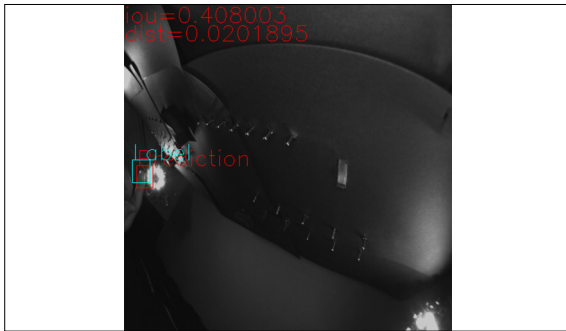


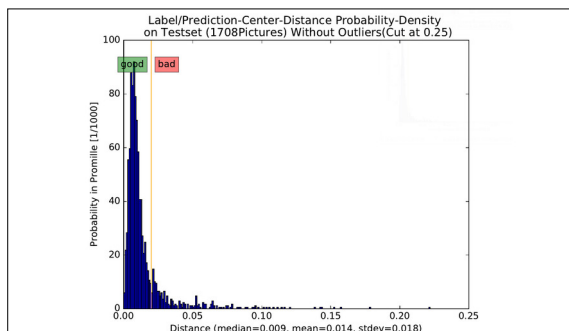
Student	Heinz Hofmann
Examinator	Prof. Dr. Guido Schuster
Themengebiet	Sensor, Actuator and Communication Systems

## Yolo auf Finger

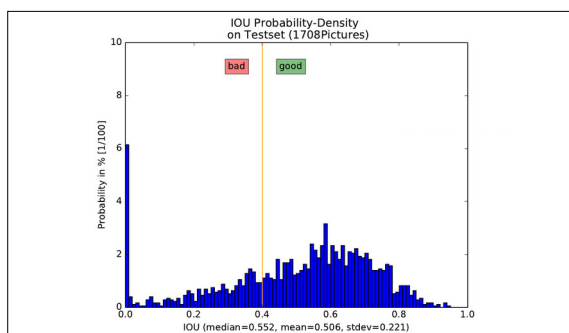
### Fingererkennung in 2D-Bildern mit State of the Art Deeplearning Konzept



Beispiel eines Bildes mit Label(türkis) und Prediction(rot)



Histogramm der Distanz zwischen Label und Prediction



Histogramm der IOU zwischen Label und Prediction

**Aufgabenstellung:** Das Ziel dieser Projektarbeit bestand darin, herauszufinden, ob Yolo geeignet wäre, die Fingerspitzen einer Hand in einem Bild zu klassifizieren und genau zu detektieren. Die Genauigkeit sollte bei maximal 0.1mm liegen. Yolo ist eine Möglichkeit, mittels Deep-Learning Objekte in einem Bild zu klassifizieren und gleichzeitig deren genaue Position zu detektieren. Daher auch der Ausdruck Yolo (You only look once). Yolo wurde als Konzept gewählt, weil es in diesem Bereich dem aktuellen Stand der Technik entspricht. Gerade die Geschwindigkeit dieses Netzwerks wurde als extrem hoch angepriesen (bis zu 45fps). Diese Geschwindigkeit ist für die letztendliche Anwendung von hoher Wichtigkeit, weil es sich um eine Echtzeitanwendung handeln soll.

**Vorgehen:** Mithilfe der Apparatur und Software von Tabea Méndez wurden Daten generiert. Um diesen Aufwand klein zu halten, wurden nur Daten vom rechten Zeigefinger generiert. Gleichzeitig wurde in Tensorflow die Architektur von Yolo nachgebaut. Dies wäre theoretisch nicht nötig gewesen, da fertige Architekturen in Keras oder Darknet online zur Verfügung stehen. Um aber einen Lerneffekt im Erstellen von Neuronalen Netzwerken zu erzielen, wurde trotzdem alles von Grund auf selber aufgebaut. Rund um die Kernarchitektur von Yolo wurden das Datenhandling, die Kostenfunktionen aber auch sämtliche Validierungen und Tests zweimal erstellt. Einmal für das Pretraining der Kerngewichte auf dem ImageNet Klassifizierungsdatenset und einmal für das "echte" Training auf den selber generierten Daten. Sobald dies alles aufgebaut und lauffähig war, wurde noch so viel wie möglich experimentiert und gleichzeitig letzte Fehler behoben. Es sollte herausgefunden werden, mit welchen Änderungen und Einstellungen das Lernresultat noch optimiert werden könnte.

**Fazit:** Das Pretraining und auch das Training hatten seine Tücken. Das originale Yolo-Netzwerk war extrem gross und brauchte entsprechend nahezu den ganzen RAM-Speicher einer GPU. Deswegen blieb nur noch begrenzt Platz für Daten übrig. Diese Probleme konnten umgangen werden, hatten jedoch zur Folge, dass die Bilder am Netzwerk-Input von 1280x960 auf 448x448 verkleinert werden mussten, um das Netzwerk zum Laufen zu bringen. Dies hatte zur Folge, dass ein Pixel bereits bis zu 1,5mm entsprechen konnte. Trotzdem wurde mit rund 84% der Predictions nur eine Genauigkeit von 15mm erreicht, was in etwa 10 Pixeln entsprach. Mit diesem Ergebnis wurde zwar das Ziel der Aufgabenstellung (0.1mm) um Faktor 150 verpasst, allerdings in 84% der Fälle Vorhersagen gemacht, welche aus subjektiver menschlicher Sicht "gut" aussehen (Das heisst ein menschlicher Betrachter würde die Vorhersage mindestens als grob richtig beurteilen.). Dies ist ein erstaunliches Resultat, wenn man bedenkt, dass zum Trainieren nur rund 18'000 Bilder verwendet wurden. Es ist anzunehmen, dass in naher Zukunft eine Genauigkeit von bis zu 1mm erreicht werden könnte. Dies sollte mit einer Verbesserung der Datengewinnung, entsprechend viel mehr Daten und mit einem verbesserten Konzept von Yolo möglich sein.