



**AFCEA Bonn e.V. Studienpreis 2017**  
Kernthesen der Arbeit

<b>Titel der Arbeit:</b>	Investigating Multiscale Class Activation Mapping for ConvNet-based Car Detection
<b>Tag der Einreichung:</b>	26.09.2016
<b>Hochschule:</b>	Rheinische Friedrich-Wilhelms-Universität Bonn
<b>Name des Verfassers:</b>	Vincent Casser
<b>Betreuender Professor:</b>	PD Dr. Volker Steinhage
<i>Kurze Beschreibung (1 Seite !) der Kernthesen.</i>	
<i>Was ist die Quintessenz der Arbeit?</i>	
<p>Das <b>Ziel der Bachelor-Arbeit</b> von Herrn Casser war die experimentelle Untersuchung von <i>Deep Learning</i> zur Erkennung und Lokalisation von Automobilen in Bildaufnahmen von Straßenszenen (s. Abb. 6 - 10 der Arbeit). Dies ist ein hochaktuelles Forschungsgebiet mit praktischen Bezügen u.a. zu Fahrerassistenzsystemen, Verkehrsüberwachung etc.</p> <p>Im Rahmen der Arbeit wurde ein tiefes Konvolutionsnetz auf umfangreichen Bilddatensätzen trainiert (s. S. 9 der Arbeit). Ca. 150.000 Bilder stammen aus dem Internet: <i>Stanford Cars Data Set</i>, <i>Chin. Univ. of Hong Kong CompCars Data Set</i> und <i>Google Images</i>. Viele dieser Bilddaten zeigen Autos in Werbebildern und Einzelaufnahmen. So scheute Herr Casser keine Mühen und erstellte ca. 660 eigene Aufnahmen aus dem realen Straßenverkehr im Raum Köln/Bonn über ein im Fahrzeugfond installiertes Smartphone! Anschließend wurden Lokalisierungsansätze evaluiert, die auf dem Ansatz des <i>Class Activation Mapping (CAM)</i> von Zhou et al. aufbauen und diesen in zwei Schritten erweitern.</p> <p>Die Arbeit von Herrn Casser bearbeitet folgende <b>Kernthesen</b>:</p> <ol style="list-style-type: none"> <li>1. Die <b>Erweiterung</b> von CAM auf <b>Multiskalen</b> reduziert das Problem der fehlenden Skalierungsinvarianz und verbessert dadurch die Lokalisation.</li> <li>2. Die <b>Sliding-Window-Erweiterung</b> von CAM erhöht die Auflösung und dadurch die Qualität der Lokalisation.</li> </ol> <p>Die experimentelle Evaluation zeigt, dass die vorgeschlagenen CAM-Erweiterungen selbst bei den realen Straßenbildern zu einer Verbesserung von durchschnittlich 6% in der beschriebenen Metrik führen (Abs. 4.3, Tab. 5). Der Lokalisationsansatz ist robust und in der Lage, gute Ergebnisse auch in anspruchsvollen Szenarien zu erzielen.</p> <p>Somit bestehen die <b>Beiträge der Arbeit</b></p> <ol style="list-style-type: none"> <li>1. technisch in der Erweiterung von CAM um Multiskalen und Sliding-Window Techniken, was die Lokalisation verbessert (Kap. 3),</li> <li>2. in der umfassenden Evaluation und des Vergleichs des konventionellen und des CAM-basierten Lokalisationsansatzes auf Lab-Bildern und unter realen Bedingungen (Kap. 4),</li> <li>3. anwendungsbezogen in der Entwicklung eines robusten Lokalisationsverfahrens für die Erkennung von Automobilen in realen Szenarien.</li> </ol> <p>Zwar ist die Berechnung eines Konvolutionsnetzes selbst auf GPUs auch heutzutage noch rechenintensiv, aber mit der Nutzung auf neuer, optimierter Hardware und insbesondere der Umsetzung der in der Arbeit vorgeschlagenen Optimierungsmöglichkeiten lässt sich die Laufzeit des Verfahrens erheblich senken (Abs. 4.4). Im Vergleich zum konventionellen Ansatz ohne CAM kann zudem bei vergleichbarer Qualität eine geringere Laufzeit erzielt werden.</p> <p>Ein wichtiger Bestandteil des vorgestellten Ansatzes ist, dass keine expliziten Lokalisationsdaten, sondern nur das <i>Labeling</i> von Bildausschnitten (Auto/nicht-Auto) für das Training notwendig sind (<i>weakly-supervised localization</i>). Das verwendete Konvolutionsnetz wird damit nur als Klassifizierer trainiert, was deutlich mehr Bilddatensätze für das Training verfügbar macht und Zeit und Kosten senken kann.</p> <p>Insgesamt wurden so nicht nur in Bezug auf die anwendungsbezogene Problemstellung bemerkenswerte Ergebnisse erzielt, sondern auch ein allgemeines Verfahren für <i>weakly-supervised localization</i> vorgestellt, das sich voraussichtlich auch für andere Domänen und mehr Objektklassen eignet (s. Abs. 4.5 und Abb. 13).</p> <p>Die Bachelorarbeit von Herrn Casser wurde von beiden Gutachtern mit der Bestnote <i>sehr gut</i> bewertet. Ausgehend von der Bachelorarbeit von Herrn Casser befindet sich eine Einreichung mit dem Titel „<i>Investigating Multiscale Class Activation Mapping for ConvNet-based Car Detection</i>“ mit Herrn Casser als Erstautor im Begutachtungsprozess der German Conference on Artificial Intelligence.</p> <p>Herr Casser ist Stipendiat der Studienstiftung und hat zwischenzeitlich Zusagen für ein Masterstudium aus Oxford, Cambridge, Yale und Harvard.</p>	

