

Kurzfassung

Zur automatischen Klassifikation von Luftbildern haben sich in den letzten Jahren überwachte Methoden bewährt. Für diese Methoden werden für jedes zu klassifizierende Bild ausreichend Trainingsdaten benötigt, welche manuell generiert werden müssen und dadurch teuer und zeitaufwändig sind. Durch immer mehr aufgenommene Bilddaten, deren Verteilungen der zu klassifizierenden Daten voneinander abweichen, wird die Beschaffung der Trainingsdaten zum Problem. Abhilfe können hier Verfahren der Domain Adaptation schaffen. Auf verfügbaren Trainingsdaten einer sogenannten Source Domain wird ein Klassifikator trainiert, der an die Target Domain, den aktuell betrachteten Datensatz für den keine Trainingsdaten zur Verfügung stehen, angepasst wird.

In dieser Arbeit wird ein Verfahren zur Domain Adaptation mit Hilfe des Expectation-Maximization (EM)-Algorithmus entwickelt. Der EM-Algorithmus bestimmt dabei die Parameter eines Gaußschen Mischmodells. Es wurden drei unterschiedliche Varianten implementiert und anhand zweier realer Datensätze getestet. In der ersten Variante der ersten Methode wird ein Gaußsches Mischmodell unüberwacht an die Verteilung der Daten der Target Domain trainiert. Als Initialparameter des EM-Algorithmus dienen dabei die Parameter des auf den Source Daten trainierten Klassifikators, was zu einer Verbesserung der Ergebnisse führen soll. Eine zweite Variante der ersten Methode verwendet die Parameter des Source Klassifikators nicht nur als Initialparameter des EM-Algorithmus, sondern auch zur initialen Klassifikation der Target Daten. Anhand der nun mit Klassenlabels versehenen Trainingsdaten der Target Domain, wird ein neuer Klassifikator trainiert. Die letzte implementierte Variante, Methode 2, bedient sich einem iterativen Ansatz. Ausgehend von einem auf den Source Datensatz trainierten Klassifikator werden iterativ Source Trainingsdaten mit Target Daten ausgetauscht. In jedem Iterationsschritt wird der Klassifikator neu trainiert, sodass sich der Klassifikator schrittweise an die Verteilung der Target Domain anpassen kann.

Insgesamt liefern die Methoden unterschiedliche Ergebnisse. Variante 1 der ersten Methode liefert keinen Informationsgewinn, während die zweite Variante in mehr als der Hälfte der Versuche eine Verbesserung der Klassifikationsergebnisse erzielt. Die iterative Methode liefert nur in wenigen Fällen einen positiven Transfer.