

gbsv Mini-Challenge 2

2023-11-27

In dieser Mini-Challenge werden LE3 und LE4 von gbsv geprüft. In Data Science und Machine Learning werden oft Merkmale (Features) von Bildern und Signalen generiert, um basierend darauf zu analysieren oder zu lernen. Diesem Thema widmen wir uns mittels klassischer Signal- und Bildverarbeitung.

Jede:r Studierende:r hat eine individualisierte Aufgabenstellung. Die Abgabe soll ebenfalls einzigartig sein.

Die Programmiersprache und die Code-Dokumentation darf frei gewählt werden. Sofern nicht anders erwähnt, dürfen vorhandene Bibliotheken verwendet werden. Gebe die Quellen deiner Daten und ggf. deines Codes an.

Checkpoints:

- Aufgabenstellung Rebuttal: In dieser Phase kann die Aufgabenstellung gereviewt und überarbeitet werden. Sofern die Studierenden entsprechende Argumente haben. Bis ca. 1 Woche nach MC-Start am Tag der Sprechstunde. 5. Dez. 2023
- Review ausgewählte Daten: Reviewe deine ausgewählten Daten bis 3 Wochen nach MC-Start mit der Fachexpertin. Das Review kann schriftlich oder in der Sprechstunde erfolgen. 12. Dez. 2023
- Anonymisierte Abgabe im Peer-Grading-Tool: Code, Resultate und Report gemäss Vorlage. Es können aktuell max. 100MB aufs Peer-Grading Tool geladen werden. 12. Jan. 2024
- Peer-Grading von anderen Abgaben. In letzten Sprechstunde im Semester oder bis 19. Jan 2024.

Die Termine sind jeweils um 23:59.

1. Mustersuche in Bild und Signal (LE3)

1.1. Korrelation in Signalen

Suche 1 Signal, welches wiederkehrende Muster enthält. Analysiere dann mittels Auto-Korrelation die wiederkehrenden Muster innerhalb deines Signals. Kann die Periodizität deines Musters via Auto-Korrelogramm sichtbar gemacht werden? Diskutiere deine

Methoden- und Parameterwahl sowie die Resultate in ca. 150 Wörtern.

Schneide nun ein Stück deines Signals aus und versuche es via Kreuzkorrelation im Ursprungssignal zu finden. Woran erkennst du, dass die Stelle passt? Beschreibe in 1-2 Sätzen. Verändere nun dein ausgeschnittenes Stück etwas und schaue, ob es immer noch via Kreuzkorrelation gefunden werden kann. Welche Arten von Veränderungen werden toleriert? Welche nicht? Diskutiere die Resultate in ca. 150 Wörtern.

Achtung: dies ist eine offene Aufgabenstellung. Setzt euch selbst einen Rahmen für die Beschränkung oder kommt in die Sprechstunde, sofern euch die Grenzen nicht klar sind. Treffende Datenwahl, auf den Punkt gebrachte Kreativität, massvolle Vielfalt und kritisches Denken sind gesucht.

1.2. Segmentierung, morphologische Operationen, Objekteigenschaften in Bildern

Suche 1 Bild, welches mehrere ähnliche Objekte enthält. Diese Objekte sollen mittels geeigneter Methoden segmentiert werden. Die Resultate sollen als gelabelte Bilder (binär oder pro Klasse 1 Label) gespeichert werden. Zeige dabei, wie du die Labelmasken mittels morphologischer Operationen verbessert hast. Erkläre hier für jede angewendete Operation in 1-2 Sätzen, warum du diese Operation anwendest. Zeige auch in Einzelbildern die Zwischenresultate deiner angewendeten Operationen und dass du nur minimal die Objekte verändert hast (z.B. keine Verschiebungen, Verkleinerungen oder Vergrößerungen). Extrahiere am Ende deine einzelnen Objekte, zähle und vermesse 2-3 Eigenschaften deiner extrahierten Objekte mittels geeigneten Methoden. Erkläre pro Eigenschaft in 1-2 Sätzen, warum du diese gewählt hast und ob die Resultate brauchbar sind.

Erstelle dann ein möglichst minimales aber repräsentatives Skeleton eines deiner Objekte und gebe die Anzahl Pixel des Skeletons aus.

Diskutiere deine Erkenntnisse und Resultate in ca. 150 Wörtern.

Weiterführende Links:

skimage: Label image regions

skimage: Segment human cells (in mitosis)

skimage: Measure region properties

2. Feature Deskriptoren in Bildern (LE4)

2.1. Keypoint Matching

Suche ein paar Bilder mit dem gleichen Sujet/Objekt aus, welche das Objekt von unterschiedlichen Blickwinkeln, aus anderer Perspektive, aus unterschiedlicher Distanz

oder rotiert zeigen. Wende dann den dir zugeordneten Keypoint Deskriptor {'Histogram-of-Gradients (HOG)'} an, um mindestens zwei deiner Bilder via Keypoints zu "matchen". Wähle dafür geeignete Parameter und eine geeignete Anzahl Keypoints. Erläutere deine Entscheidungen in 1-2 Sätzen. Zeige deine Resultate visuell und stelle 2-3 Beobachtungen auf.

Diskutiere in ca. 150 Wörtern wie robust der dir zugeordnete Algorithmus {'Histogram-of-Gradients (HOG)'} ist in Bezug auf Transformationen oder unterschiedlicher Aufnahmeverhältnisse (Licht, ...) und dessen rechnerische Effizienz. Zeige mindestens eine dieser Eigenschaften anhand deiner Beispieldaten. Diskutiere die Resultate und Erkenntnisse in 2-3 Sätzen.

Achtung: dies ist eine offene Aufgabenstellung. Setzt euch selbst einen Rahmen für die Beschränkung oder kommt in die Sprechstunde, sofern euch die Grenzen nicht klar sind. Treffende Datenwahl, auf den Punkt gebrachte Kreativität, massvolle Vielfalt und kritisches Denken sind gesucht.

3. Peer-Grading

Nach Abgabe der Mini-Challenge hast du 1 Woche Zeit drei dir zugeordnete Abgaben von anderen zu bewerten. Die Zuordnung erfolgt via FHNW Peer-Grading-Tool (siehe Link unten). Orientiere dich für die Bewertung an den vorgegebenen Bewertungskriterien (siehe Excel-Datei oder Peer-Grading-Tool). Die Note 5 bedeutet, dass alles erfüllt ist, wie du es von einem guten Data Scientist in der Praxis erwarten würdest. Du startest als Baseline mit der Note 5. Entdeckst du Fehler, geht die Note nach unten. Der Note 5.5 nähert man sich, wenn die Erwartungen übertroffen wurden. Der Note 6 nähert man sich, wenn die Leistung ausserordentlich ist und kritisches Denken, Variabilität, eigene Ideen und Kreativität beinhaltet. Siehe auch Checkliste für Bewertung. Die Benotung soll auf Zehntel gerundet sein. Wer auf Zehntel gerundet mit 0.1 Abweichung die Endnote von der Fachexpertin trifft, kriegt einen Bonus von 0.2 Noten für die eigene Abgabe. In der Sprechstunde vom 16. Januar widmen wir uns dem Peer-Grading.

Weiterführende Links:

FHNW Peer-Grading-Tool