Teaching Machines to Measure Economic Activities from Satellite Images: Challenges and Solutions

발표: 양현주 (서강대)

토론: 이종관 (이화여대)

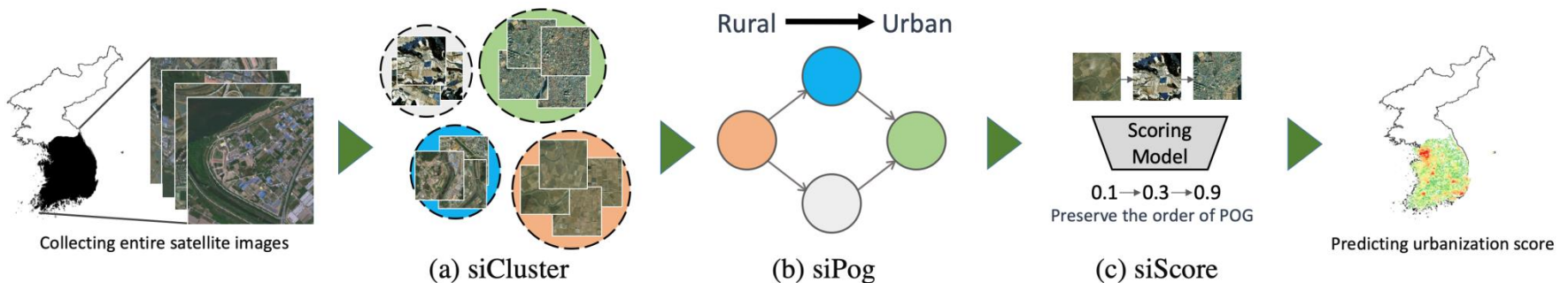한국개발정책학회

2021. 5. 14(금)

Review

- Measuring development without ground-truth data

  – Rank *relative scores* of economic development

- Highly relevant for studying areas without ground-truth

  – North Korea, developing countries,...

- Significant improvement from existing methods

  – Such as nightlight intensity data

- Potentially very useful source of data in urban economics

  – Can answer various research questions that were impossible to study before

Review

- **1st stage** (siCluster)

  – Collect satellite images → cluster them using unsupervised learning

- **2nd stage** (siPog)

  – Define order of the clusters, capturing the level of development

- **3rd stage** (siScore)

  – Assign a differentiable score

Collecting entire satellite images

(a) siCluster

Rural → Urban

(b) siPog

Scoring Model

0.1→0.3→0.9
Preserve the order of POG

(c) siScore

Predicting urbanization score

Comments and Questions

1. Order of the approach
   - 현 방법은 먼저 cluster 를 한 후에 rank와 score를 결정
   - 초기에 cluster를 어떻게 설정하는지가 매우 중요할 것으로 보임.
   - How to choose the number of clusters? (same score within cluster?)
   - Are the results stable with different clusters? (e.g., K-means clustering)
   - More fundamentally, why cluster first?

Comments and Questions

2. What aspects are NOT captured?

   – Can capture who there are? (e.g., share of college-educated?)

3. Lagged effect?

   – Model evaluation에서 "시점"도 고려해볼 필요

   – Satellite image 가 일정 시간 이후의 economic activity 를 더 잘 capture 할 가능성

Comments and Questions

4. Ordinal/cardinal

    –   Can the assigned scores have cardinal meaning?

    –   Cluster1 is 5% more developed compared to cluster2 ?

5. Mainly used as a dependent variable?

    –   Sometime people may want to use it as explanatory variable in regressions?

    –   Important to think about what economic aspects these proxies capture

Comments and Questions

6. North Korea Example

    – Event Study 를 적용함에 있어서 앞서 말한 시점 문제를 고민해 볼 필요

    – Satellite image itself could be controlled by the dictatorship