Machine Learning, AI & satellite imagery: what impact on humanitarian mapping?

Benjamin Herfort

GIScience Research Group

Heidelberg University
A Sketch of a “smarter” crowdsourcing workflow

- automation can play a role during information extraction and quality assessment (validation!)
- mappers are essential in all steps of the workflow
- let’s not name it “artificial intelligence”, but “intelligent assistance”
Example: Towards MapSwipe 2.0

Targets:

- Improve data quality (building sensitivity and building precision)
- Increase mapping speed or map larger areas
- Rely on human skills when needed, rely on automation when possible

Potential for Machine Learning:

- Pre processing, e.g. filtering cloudy areas or waterbodies
- Image Analysis and Building Detection
- Intrinsic Data Quality Analysis
### Example: Towards MapSwipe 2.0

<table>
<thead>
<tr>
<th>(1) automated building detection</th>
<th>(2) building detection post processing</th>
<th>(3) building footprint validation</th>
<th>(4) OSM upload</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very high resolution satellite imagery (e.g. from Bing, Digital Globe, OpenAerialMap)</td>
<td>Automatically generated building footprints (raw)</td>
<td>building footprint post processing (filter, simplify, rectangular, circular, remove duplicates)</td>
<td>validated building footprints (e.g. by 3 independent users)</td>
</tr>
<tr>
<td>Area of Interest (e.g. bounding box, polygon geometry)</td>
<td></td>
<td>enhanced building footprints (e.g. simplified, rectangular)</td>
<td>Upload to OpenStreetMap Database</td>
</tr>
<tr>
<td>existing building footprints in OSM</td>
<td>external building footprint dataset with appropriate licence</td>
<td>MapSwipe 2.0 crowdsourced building footprint validation</td>
<td>invalidated building footprints (e.g. by 3 independent users)</td>
</tr>
</tbody>
</table>

https://docs.google.com/document/d/1lVXquSJA_AzgHjIVMI3ny2QQqDpsnOddnt_zYaZBLOQ/edit?usp=sharing
Example: AI & HOT Tasking Manager

- AI to prioritize mapping tasks
  - E.g. estimate number of features to be mapped in a task

- AI to validate OSM data
  - Compare between estimates and actuals
  - Tell me where I missed something
  - Detect areas with most quality issues

Example: AI & HOT Tasking Manager

Challenges

Technical:
- ensure that tools and data are open and accessible
- understand and communicate quality of AI datasets
- Integrate AI data into OSM database and processes
- build the infrastructure (an API that can spin up and deploy a model) to consistently and quickly run ML models

Non-Technical:
- increased need for experienced mappers for validation
- machines are only effective at a small number of mapping tasks → don’t forget local mapping
- Increased need for technical know-how for project managers → understand AI data and results
- Potential source of new bias
Thank you for the attention.
herfort@uni-heidelberg.de

more infos:
https://heigit.org/