

School mapping in ESRI imagery

Andrija Gorup, Marin Kačan, Siniša Šegvić

Data

OSM data for schools and non-schools

- 51 countries (including Vietnam)

Anditi schools

- Hand-curated set of ~1000 schools in Vietnam
- 946 schools in the chosen 26 districts

Tiles covering 26 Vietnamese districts (for dense inference)

- non-overlapping: ~27,500 tiles
- overlapping: ~109,000 tiles

OpenStreetMap data

Country	School tiles	Non-school tiles
Brazil	12,113	56,690
Ukraine	1,228	44,797
Thailand	5,213	38,317
Indonesia	9,940	34,872
Vietnam	1,574	15,384
Other	27,662	179,941
Total	57,730	370,001

Initial (UNICEF) training setup

Models

- frozen foundation model (DinoV2_ViT-L) + linear classifier (LogReg, SVM, ...)
- convolutional networks
 - ResNet-18/50
 - Xception

Initial data augmentations

- randomly applied 90 degree rotation
- randomly applied horizontal flip
- randomly applied vertical flip
- resize 500px -> 224px

Initial results

Initial (UNICEF) set of 42 countries

- train set size: ~27,000 schools; ~180,000 non-schools

Model	F1 (pp)
frozen DinoV2_ViT-L + LR	46.16
frozen DinoV2_ViT-L + SVC	14.34
frozen DinoV2_ViT-L + xgboost	29.98
Resnet18	66.31
Resnet50	66.34
Xception	66.30
<u>Convnext-small</u>	<u>72.50</u>
Convnext-large	72.64

Validation of data augmentation procedures

Setup: ResNet-18, initial set of 42 countries

Model	F1 (pp)	Δ F1 (pp)
No resize (500px)	68.21	+1.9
RandomCrop (352px)	66.35	+0.04
CenterCrop (352px)	66.83	+0.52
ColorJitter (0.1,0.1,0.1,0.02)	65.56	-0.75
ColorJitter (0.2,0.2,0.2,0.05)	65.49	-0.82
GaussianBlur	64.46	-1.85
Continuous rotation + CenterCrop 352px	68.79	+2.48
Weighted loss	54.83	-8.48

Evaluation on OSM Vietnam val

Setup: ResNet-18, initial data augmentations

Training dataset	# schools in train	# non-schools in train	Total train	F1 (pp)
Initial 42 countries (no Vietnam)	~27,000	~180,000	~207,000	46.68
Countries neighbouring Vietnam	~22,000	~134,500	~156,500	48.77
Initial + neighbouring countries	~45,000	~283,500	~328,500	54.12
Vietnam train only	~1,300	~12,200	~13,500	56.07

Two-stage training, eval on OSM Vietnam val

Train on all countries (w/out Vietnam) + fine-tune

Setup:

- ConvNext-small, continuous rotation

Two-stage training, eval on OSM Vietnam val

Train on all countries (w/out Vietnam) + fine-tune

Setup:

- ConvNext-small, continuous rotation

Fine-tuning dataset	F1 (pp)
No fine-tuning	56.43
OSM Vietnam train	71.30
Anditi	60.68
OSM Vietnam train + Anditi	67.19

2-stage training, fine-tune and eval on Anditi

2-fold cross-validation (50:50 train/val split)

- **F1: 93.91 pp**
- Recall: 93.11 pp
- Precision: 94.73 pp

Outdated imagery

ESRI imagery outdated compared to Google Maps

- New schools (visible on Google Maps), might not be visible in ESRI images

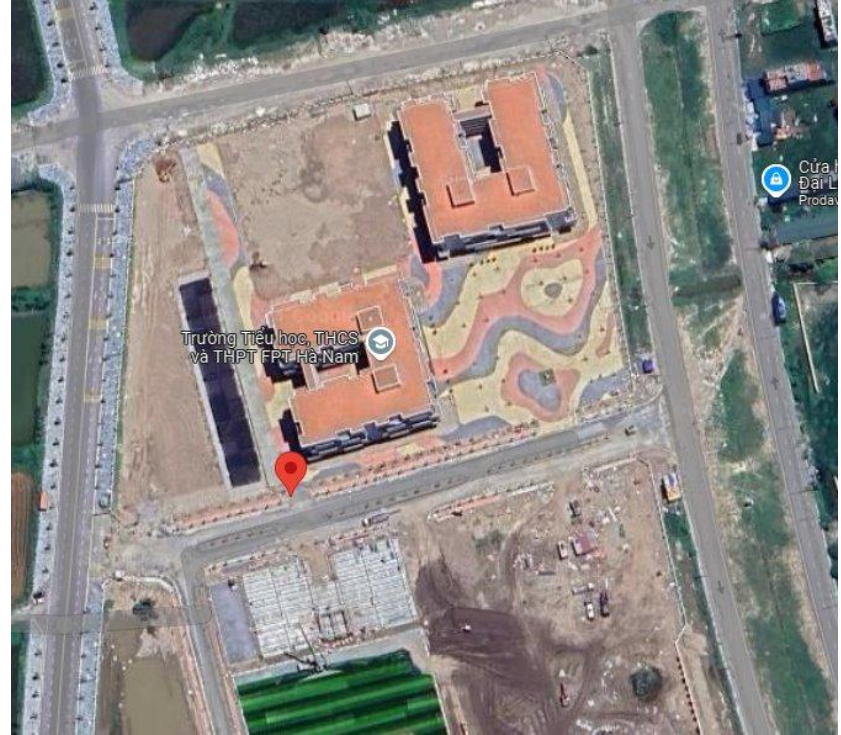
Outdated imagery

ESRI imagery outdated compared to Google Maps

- New schools (visible on Google Maps), might not be visible in ESRI images



ESRI



GOOGLE

Outdated imagery

ESRI imagery outdated compared to Google Maps

- New schools (visible on Google Maps), might not be visible in ESRI images



ESRI



GOOGLE

Deliverable: Dense inference on 26 districts

Run our best model on tiles covering the chosen 26 districts

- non-overlapping tiles: ~27,500 inhabited tiles
- **overlapping tiles: ~109,000 inhabited tiles**
 - non-max suppression for overlapping predictions

Deliverable: Dense inference on 26 districts

Run our best model on tiles covering the chosen 26 districts

- non-overlapping tiles: ~27,500 inhabited tiles
- **overlapping tiles: ~109,000 inhabited tiles**
 - non-max suppression for overlapping predictions

Performance estimate

- % of Anditi schools contained in the 26 districts retrieved by this approach
- Anditi train:
 - “easy” schools (659), recognized by the model w/out Anditi fine-tuning
 - **96.1% schools retrieved**
- Anditi val
 - “hard” schools (289)
 - **65.1% schools retrieved**

Deliverable: Dense inference on 26 districts

Run our best model on tiles covering the chosen 26 districts

- non-overlapping tiles: ~27,500 inhabited tiles
- **overlapping tiles: ~109,000 inhabited tiles**
 - non-max suppression for overlapping predictions

Retrieved schools: ~2,900