

# Comparing CNNs and Random Forests for Landsat Image Segmentation Trained on a Large Proxy Land Cover Dataset

## Supplementary Material

**Table S1. Class correspondence of ALUMv8 land use classes and simplified land cover classes \*.**  
\* Secondary ALUM class names (corresponding to ALUM codes ending in 0) are listed in **bold**.

ALUMv8 tertiary code	ALUMv8 tertiary name	Simplified land cover class	ALUMv8 tertiary code	ALUMv8 tertiary name	Simplified land cover class
110-113	<b>Nature conservation</b> , Strict nature reserves, Wilderness area, National park	Forest	420-424	<b>Grazing irrigated modified pastures</b> , Irrigated woody fodder plants, Irrigated pasture legumes, Irrigated legume/grass mixtures, Irrigated sown grasses	Grassland
114	Natural feature protection	Unknown	430-439	<b>Irrigated cropping</b> , Irrigated cereals, Irrigated beverage and spice crops, Irrigated hay and silage, Irrigated oilseeds, Irrigated sugar, Irrigated cotton, Irrigated alkaloid poppies, Irrigated pulses, Irrigated rice	Crop
115-116	Habitat/species management area, Protected landscape	Forest	440-449, 450-454	<b>Irrigated perennial horticulture</b> , Irrigated tree fruits, Irrigated olives, Irrigated tree nuts, Irrigated vine fruits, Irrigated shrub berries and fruits, Irrigated perennial flowers and bulbs, Irrigated perennial vegetables and herbs, Irrigated citrus, Irrigated grapes, <b>Irrigated seasonal horticulture</b> , Irrigated seasonal fruits, Irrigated seasonal flowers and bulbs, Irrigated seasonal vegetables and herbs, Irrigated turf farming	Horticulture
117, 120-125, 130-132	Other conserved area, <b>Managed resource protection</b> , Biodiversity, Surface water supply, Groundwater, Landscape, Traditional Indigenous uses, <b>Other minimal use</b> , Defence land - natural areas, Stock route	Unknown	460-465, 510-515, 520-528	<b>Irrigated land in transition</b> , Degraded irrigated land, Abandoned irrigated land, Irrigated land under rehabilitation, No defined use - irrigation, Abandoned irrigated perennial horticulture, <b>Intensive horticulture</b> , Production nurseries, Shadehouses, Glasshouses, Glasshouses - hydroponic, Abandoned intensive horticulture, <b>Intensive animal production</b> , Dairy sheds and yards, Feedlots, Poultry farms, Piggeries, Aquaculture, Horse studs, Saleyards/stockyards, Abandoned intensive animal production	Unknown
133-134	Residual native cover, Rehabilitation	Forest	530-538, 540-541	<b>Manufacturing and industrial</b> , General purpose factory, Food processing factory, Major industrial complex, Bulk grain storage, Abattoirs, Oil refinery, Sawmill, Abandoned manufacturing and industrial, <b>Residential and farm infrastructure</b> , Urban residential	Built-up
210	<b>Grazing native vegetation</b>	Grassland	542-543	Rural residential with agriculture, Rural residential without agriculture	Unknown
220-222	<b>Production native forests</b> , Wood production forestry, Other forest production	Forest	544-545, 550-555	Remote communities, Farm buildings/infrastructure, <b>Services</b> , Commercial services, Public services, Recreation and culture, Defence facilities - urban, Research facilities	Built-up
310-314	<b>Plantation forests</b> , Hardwood plantation forestry, Softwood plantation forestry, Other forest plantation, Environmental forest plantation	Plantation	560-567	<b>Utilities</b> , Fuel powered electricity generation, Hydro electricity generation, Wind electricity generation, Solar electricity generation, Electricity substations and transmission, Gas treatment storage and transmission, Water extraction and transmission	Unknown
320-325	<b>Grazing modified pastures</b> , Native/exotic pasture mosaic, Woody fodder plants, Pasture legumes, Pasture legume/grass mixtures, Sown grasses	Grassland	570-575	<b>Transport and communication</b> , Airports/aerodromes, Roads, Railways, Ports and water transport, Navigation and communication	Built-up
330-338	<b>Cropping</b> , Cereals, Beverage and spice crops, Hay and silage, Oilseeds, Sugar, Cotton, Alkaloid poppies, Pulses	Crop	580-584	<b>Mining</b> , Mines, Quarries, Tailings, Extractive Industry not in use	Bare
340-349, 350-353	<b>Perennial horticulture</b> , Tree fruits, Olives, Tree nuts, Vine fruits, Shrub berries and fruits, Perennial flowers and bulbs, Perennial vegetables and herbs, Citrus, Grapes, <b>Seasonal horticulture</b> , Seasonal fruits, Seasonal flowers and bulbs, Seasonal vegetables and herbs	Horticulture	590-595	<b>Waste treatment and disposal</b> , Effluent pond, Landfill, Solid garbage, Incinerators, Sewage/sewerage	Unknown
360-365	<b>Land in transition</b> , Degraded land, Abandoned land, Land under rehabilitation, No defined use, Abandoned perennial horticulture	Unknown	610-614, 620-623, 630-633, 640-643, 650-654	<b>Lake</b> , Lake - conservation, Lake - production, Lake - intensive use, Lake - saline, <b>Reservoir/dam</b> , Reservoir, Water storage - intensive use/farm dams, Evaporation basin, <b>River</b> , River - conservation, River - production, River - intensive use, <b>Channel/aqueduct</b> , Supply channel/aqueduct, Drainage channel/aqueduct, Stormwater, <b>Marsh/wetland</b> , Marsh/wetland - conservation, Marsh/wetland - production, Marsh/wetland - intensive use, Marsh/wetland - saline	Water
410-414	<b>Irrigated plantation forests</b> , Irrigated hardwood plantation forestry, Irrigated softwood plantation forestry, Irrigated other forest plantation, Irrigated environmental forest plantation	Plantation	660-663	<b>Estuary/coastal waters</b> , Estuary/coastal waters - conservation, Estuary/coastal waters - production, Estuary/coastal waters - intensive use	Unknown

**Table S2: Label data statistics – no. pixels of each land cover class for tile +14, –40 and study area <sup>+</sup>**  
<sup>+</sup> Study area was clipped to  $19968 \times 19968$  pixels to allow division evenly into  $128 \times 128$  pixel tiles. Note: weighted-mean F1 scores and class accuracies were calculated using class statistics (excluding the no data class) for the whole study area.

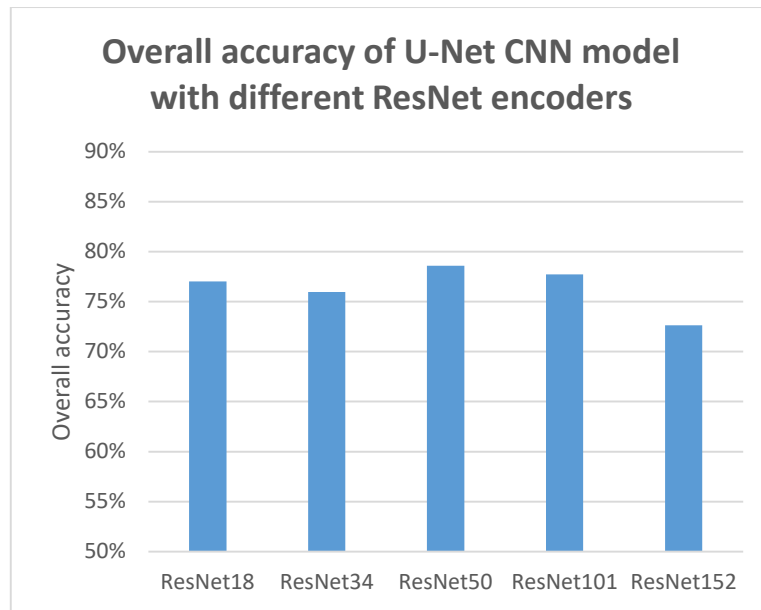
Code	Land Cover class	Tile +14, -40		Study area	
		Count	Percentage	Count	Percentage
0	No data	92,020	0.6%	14,196,184	3.6%
1	Forest	4,821,084	30.1%	112,447,908	28.2%
2	Grassland	6,080,880	38.0%	130,618,500	32.8%
3	Horticulture	39,224	0.2%	1,658,220	0.4%
4	Crop	2,531,988	15.8%	118,481,376	29.7%
5	Plantation	1,817,784	11.4%	5,676,428	1.4%
6	Bare	6,160	0.04%	379,236	0.1%
7	Water	390,076	2.4%	7,554,004	1.9%
8	Built-up	220,784	1.4%	7,709,168	1.9%
<b>TOTAL</b>		<b>16,000,000</b>	<b>100.0%</b>	<b>398,721,024</b>	<b>100.0%</b>

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, None, None, 128)	7040
conv2d_2 (Conv2D)	(None, None, None, 64)	73792
conv2d_3 (Conv2D)	(None, None, None, 32)	18464
conv2d_transpose_1 (Conv2DTr	(None, None, None, 32)	9248
conv2d_transpose_2 (Conv2DTr	(None, None, None, 64)	18496
conv2d_transpose_3 (Conv2DTr	(None, None, None, 128)	73856
conv2d_4 (Conv2D)	(None, None, None, 9)	10377
Total params: 211,273		
Trainable params: 211,273		
Non-trainable params: 0		

**Figure S1: Autoencoder model summary. Input images are  $128 \times 128$  pixels. Output is a pixel-based land cover classification of nine classes (eight land cover classes plus no data).**

**Table S3: Model results – most accurate U-Net CNN and Random Forests model predictions in bold.**

Model name	Epoch	Model type	Input	Per class F1 scores								Weighted-mean F1	Overall Acc.
				Forest	Grassland	Hort.	Crop	Plantation	Bare	Water	Built-up		
ms_r1_ae_6b_e70.h5	70	Autoencoder	6bands	78%	80%	10%	18%	65%	10%	54%	43%	59%	72%
<b>ex4.1_rf_rsp_e200_d30.sav</b>	<b>n/a</b>	<b>Random Forests</b>	<b>6bands</b>	<b>83%</b>	<b>77%</b>	<b>4%</b>	<b>50%</b>	<b>59%</b>	<b>8%</b>	<b>35%</b>	<b>23%</b>	<b>68%</b>	<b>73%</b>
bm_r1_unet_6b_e79.h5	79	Simple U-Net	6bands	86%	80%	29%	58%	83%	1%	47%	40%	74%	78%
bm_r1_urn18_6b_e45.h5	45	U-Net ResNet18	6bands	85%	80%	31%	57%	81%	24%	43%	42%	73%	77%
ms_r1_urn34_6b_e60.h5	60	U-Net ResNet34	6bands	86%	79%	19%	40%	85%	23%	53%	39%	68%	76%
bm_r2_urn50_6b_e36.h5	36	U-Net ResNet50	6bands	85%	81%	21%	65%	82%	20%	51%	43%	76%	79%
bm_r1_urn101_6b_e39.h5	39	U-Net ResNet101	6bands	85%	80%	27%	63%	81%	21%	50%	40%	75%	78%
bm_r1_urn152_6b_e39.h5	39	U-Net ResNet152	6bands	83%	75%	9%	57%	79%	17%	44%	32%	70%	73%
bm_r1_urn50_b1_e19.h5	19	U-Net ResNet50	Blue	84%	78%	12%	57%	77%	10%	49%	39%	72%	75%
bm_r1_urn50_b2_e26.h5	26	U-Net ResNet50	Green	82%	79%	24%	63%	74%	14%	49%	37%	73%	75%
bm_r1_urn50_b3_e23.h5	23	U-Net ResNet50	Red	84%	78%	17%	49%	77%	11%	46%	31%	69%	74%
bm_r1_urn50_b4_e12.h5	12	U-Net ResNet50	NIR	74%	73%	0%	61%	28%	0%	49%	24%	67%	67%
bm_r1_urn50_b5_e22.h5	22	U-Net ResNet50	SWIR1	85%	78%	9%	48%	73%	0%	52%	22%	68%	73%
bm_r1_urn50_b6_e33.h5	33	U-Net ResNet50	SWIR2	83%	78%	15%	50%	74%	1%	51%	27%	69%	74%
bm_r1_urn50_bcdev_e17.h5	17	U-Net ResNet50	bcdev	71%	65%	0%	60%	47%	0%	36%	17%	63%	63%
bm_r1_urn50_sdev_e27.h5	27	U-Net ResNet50	sdev	67%	59%	1%	62%	66%	0%	38%	20%	61%	62%
bm_r1_urn50_edev_e16.h5	16	U-Net ResNet50	edev	73%	75%	6%	61%	71%	0%	38%	24%	68%	70%
bm_r1_urn50_RGB_e34.h5	34	U-Net ResNet50	RGB	86%	79%	18%	55%	83%	13%	48%	43%	72%	77%
bm_r1_urn50_NS1S2_e38.h5	38	U-Net ResNet50	NS1S2	86%	81%	26%	62%	78%	19%	52%	44%	75%	78%
bm_r1_urn50_RGBN_e34.h5	34	U-Net ResNet50	RGBN	86%	79%	20%	52%	80%	24%	53%	34%	71%	76%
bm_r1_urn50_TMADs_e29.h5	29	U-Net ResNet50	TMADs	85%	80%	23%	66%	74%	10%	49%	20%	75%	77%
<b>bm_r1_urn50_9b_e34.h5</b>	<b>34</b>	<b>U-Net ResNet50</b>	<b>9bands</b>	<b>87%</b>	<b>80%</b>	<b>28%</b>	<b>67%</b>	<b>82%</b>	<b>18%</b>	<b>54%</b>	<b>47%</b>	<b>77%</b>	<b>79%</b>



**Figure S2: Accuracy of U-Net CNN model using different ResNet encoders with ImageNet weights.**

**Table S4: U-Net CNN model (bm\_r1\_urn50\_9b\_e34.h5) v. label data confusion matrices for test tile.**

[illegible]

**Table S5: Random Forests modelling results for 6 band Landsat geomedian input. Best results highlighted~**

~ By sampling method: RSP (random sampling of pixels), SRS-Eq (stratified random sampling with equal samples for all classes), SRS-Prop (stratified random sampling proportional to area occupied by each class).

Sampling Method	Trees	Max. Depth	Per class F1 scores								Weighted-mean F1	Overall Accuracy
			Forest	Grassland	Horticulture	Crop	Plantation	Bare	Water	Built-up		
RSP	50	10	82.3%	55.2%	2.7%	42.9%	68.6%	3.1%	26.4%	7.9%	57.7%	54.7%
RSP	50	20	84.1%	75.7%	7.6%	48.1%	64.5%	9.4%	30.6%	16.9%	67.1%	71.0%
RSP	50	30	83.1%	77.1%	5.5%	49.4%	58.0%	6.7%	34.0%	21.9%	67.7%	72.7%
RSP	50	40	83.1%	77.0%	4.3%	48.9%	57.9%	6.6%	34.6%	21.4%	67.5%	72.7%
RSP	50	50	83.1%	77.1%	3.8%	49.2%	58.1%	7.4%	32.7%	21.1%	67.6%	72.7%
RSP	100	10	83.1%	55.4%	2.5%	42.9%	69.4%	3.9%	29.1%	8.1%	58.1%	55.0%
RSP	100	20	84.0%	75.6%	7.2%	48.0%	63.8%	8.5%	29.9%	16.7%	67.0%	71.0%
RSP	100	30	83.2%	77.3%	5.1%	49.6%	57.9%	7.3%	33.2%	22.5%	67.9%	72.9%
RSP	100	40	83.1%	77.3%	4.2%	49.4%	58.0%	6.8%	31.8%	22.0%	67.7%	72.8%
RSP	100	50	83.1%	77.3%	3.6%	49.5%	57.6%	9.0%	34.0%	21.4%	67.8%	72.8%
RSP	200	10	82.7%	55.4%	2.5%	44.2%	69.4%	4.9%	27.9%	8.2%	58.4%	55.2%
RSP	200	20	84.2%	76.0%	7.9%	47.9%	64.6%	9.7%	30.8%	18.0%	67.1%	71.3%
<b>RSP</b>	<b>200</b>	<b>30</b>	<b>83.3%</b>	<b>77.4%</b>	<b>4.2%</b>	<b>49.6%</b>	<b>58.5%</b>	<b>7.8%</b>	<b>35.2%</b>	<b>22.9%</b>	<b>68.0%</b>	<b>73.1%</b>
RSP	200	40	83.1%	77.3%	3.3%	49.6%	57.6%	7.7%	33.8%	22.6%	67.8%	72.9%
RSP	200	50	83.2%	77.3%	9.3%	49.7%	58.3%	5.9%	33.2%	22.0%	67.9%	73.0%
RSP	300	10	83.0%	56.4%	2.8%	43.1%	68.8%	4.2%	29.5%	8.0%	58.5%	55.6%
RSP	300	20	84.0%	76.1%	6.6%	48.6%	64.0%	11.1%	27.1%	17.9%	67.3%	71.4%
RSP	300	30	83.1%	77.5%	8.5%	49.5%	58.2%	9.2%	29.8%	22.7%	67.8%	73.0%
SRS-Eq	50	10	82.3%	58.1%	2.7%	42.2%	68.4%	3.9%	30.9%	8.3%	58.6%	56.1%
SRS-Eq	50	20	83.3%	62.4%	3.7%	41.7%	71.8%	2.4%	25.8%	9.0%	60.2%	58.8%
SRS-Eq	50	30	82.5%	59.1%	3.5%	40.7%	70.9%	1.8%	24.0%	8.0%	58.4%	56.6%
SRS-Eq	50	40	82.7%	58.8%	3.4%	41.3%	71.2%	1.7%	23.8%	8.1%	58.6%	56.7%
SRS-Eq	50	50	82.6%	58.6%	3.3%	41.3%	71.1%	1.7%	23.9%	8.0%	58.5%	56.6%
SRS-Eq	100	10	82.3%	57.9%	2.7%	41.9%	68.5%	3.8%	30.9%	8.3%	58.4%	55.9%
<b>SRS-Eq</b>	<b>100</b>	<b>20</b>	<b>83.5%</b>	<b>63.3%</b>	<b>3.7%</b>	<b>42.2%</b>	<b>72.1%</b>	<b>2.4%</b>	<b>24.9%</b>	<b>9.3%</b>	<b>60.7%</b>	<b>59.4%</b>
SRS-Eq	100	30	82.8%	59.9%	3.6%	42.1%	71.3%	1.9%	24.2%	8.3%	59.2%	57.4%
SRS-Pr	50	10	82.4%	56.3%	2.7%	42.6%	68.6%	6.0%	29.8%	8.2%	58.1%	55.3%
SRS-Pr	50	20	84.0%	75.3%	6.0%	47.7%	64.7%	7.8%	29.4%	16.5%	66.7%	70.6%
SRS-Pr	50	30	83.2%	77.2%	6.2%	49.5%	58.2%	6.6%	35.1%	22.2%	67.8%	72.9%
SRS-Pr	50	40	83.0%	77.0%	8.3%	48.9%	57.6%	7.5%	33.3%	21.4%	67.5%	72.6%
SRS-Pr	50	50	82.9%	77.0%	5.6%	49.4%	57.6%	7.2%	32.3%	21.3%	67.6%	72.6%
SRS-Pr	100	10	83.0%	56.9%	2.4%	43.2%	68.8%	5.8%	31.4%	8.4%	58.7%	55.9%
SRS-Pr	100	20	84.0%	75.6%	7.5%	48.4%	63.7%	11.2%	30.9%	16.9%	67.1%	71.0%
<b>SRS-Pr</b>	<b>100</b>	<b>30</b>	<b>83.1%</b>	<b>77.4%</b>	<b>5.3%</b>	<b>49.8%</b>	<b>58.1%</b>	<b>7.6%</b>	<b>29.7%</b>	<b>22.9%</b>	<b>67.9%</b>	<b>72.9%</b>
<b>SRS-Pr</b>	<b>100</b>	<b>40</b>	<b>83.2%</b>	<b>77.3%</b>	<b>4.5%</b>	<b>49.8%</b>	<b>58.1%</b>	<b>6.0%</b>	<b>31.2%</b>	<b>22.4%</b>	<b>67.9%</b>	<b>72.9%</b>
SRS-Pr	100	50	83.2%	77.2%	4.1%	49.3%	58.3%	8.3%	31.1%	22.0%	67.7%	72.8%
SRS-Pr	200	10	82.9%	55.9%	2.6%	43.4%	68.6%	4.1%	32.3%	8.0%	58.4%	55.4%
SRS-Pr	200	20	84.0%	75.9%	6.5%	48.2%	64.1%	8.5%	28.7%	18.2%	67.1%	71.3%
SRS-Pr	200	30	82.8%	77.3%	2.6%	49.1%	56.0%	6.2%	29.9%	21.3%	67.4%	72.6%

**Table S6: U-Net CNN model statistics – no. pixels of land cover classes for tile +14, -40 and study area.**

Code	Land cover class	Tile +14, -40		Study area	
		Count	Percentage	Count	Percentage
0	No data	47	0.0%	12,512	0.0%
1	Forest	4,046,773	25.3%	104,296,078	26.2%
2	Grassland	6,990,386	43.7%	124,703,621	31.3%
3	Horticulture	82,332	0.5%	2,403,360	0.6%
4	Crop	2,281,698	14.3%	133,518,367	33.5%
5	Plantation	2,119,612	13.2%	6,246,536	1.6%
6	Bare	23,206	0.1%	1,401,930	0.4%
7	Water	261,034	1.6%	18,574,094	4.7%
8	Built-up	194,912	1.2%	7,564,526	1.9%
<b>TOTAL</b>		<b>16,000,000</b>	<b>100.0%</b>	<b>398,721,024</b>	<b>100.0%</b>