

Features to consider in land cover classification based on a machine-learning random forest method

Amarsaikhan Damdinsuren^{1,*}, Enkhmanlai Amarsaikhan²,
Enkhjargal Damdinsuren¹, Tsogzol Gurjav¹

¹*Division of GIS and Remote Sensing, Institute of Geography and Geoecology, Mongolian Academy of Sciences, Ulaanbaatar, Mongolia*

²*National Technical University, Ulaanbaatar, Mongolia*

*Corresponding author email: amarsaikhan@mas.ac.mn

Received: 29 April 2024 / Accepted: 13 November 2024 / Published online: 29 November 2024

ABSTRACT

In some modern digital image processing systems, different machine learning and high-level processing algorithms are being developed. A researcher can obtain high-quality thematic results by correctly and optimally determining the parameter values in the algorithm based on his or her own knowledge. This research aims to perform a land cover classification using a machine learning random forest (RF) classification algorithm on the SNAP Toolbox system and demonstrate how to adjust the classification parameters for efficient decision-making. In order to increase the classification accuracy, the study uses 3 different combinations of the integrated optical and synthetic aperture radar (SAR) data sets. The area in the southern part of Ulaanbaatar, the capital city, was selected as a test area for the study, and the land cover was classified into 6 different classes, such as the built-up area, forest, willow, grass, open land, and water. In the RF technique, the number of signatures were set as 5000, 10000, and 15000 and the related number of decision trees as 10, 50, and 100, respectively. As could be seen from the results of the study, the classification results of the combined multisource data sets had higher overall accuracies, and the result of 11 bands with the number of signatures set to 15000 and the decision trees set to 100, had the highest overall accuracy (92.68%).

Keywords: Land cover, SNAP Toolbox, Machine learning, RF

Машин сургалтын анамсаргүй ойн аргаар газрын бүрхэвчийг ангилахад харгалзан үзэх онцлогууд

Амарсайхан Дамдинсүрэн^{1,*}, Энхманлай Амарсайхан²,
Энхжаргал Дамдинсүрэн¹, Цогзол Гүржав¹

¹Газарзүй мэдээллийн систем, зайнаас тандан судлалын салбар, Газарзүй, геоэкологийн хүүрээлэн, Шинжлэх Ухааны Академи, Улаанбаатар, Монгол

²Үндэсний Техникийн Их Сургууль, Улаанбаатар, Монгол

*Холбоо барих зохиогчийн цахим хаяг: amarsaikhan@mas.ac.mn

Хүлээн авсан: 2024 оны 04 сарын 29 өдөр / Зөвшөөрөгдсөн: 2024 оны 11 сарын 13 өдөр /
Нийтлэгдсэн: 2024 оны 11 сарын 29 өдөр

ХУРААНГУЙ

Орчин үеийн дүрсийн тоон боловсруулалтын зарим системд, машин сургалтын болон өндөр түвшингийн янз бүрийн аргууд бий болж байгаа бөгөөд судлаач өөрийн мэдлэг дээр тулгуурлан тухайн алгоритм дахь параметруудын утгуудыг зөв, оновчтой тодорхойлж өгснөөр өндөр чанарын үзүүлэлт бүхий сэдэвчилсэн үр дүнг гарган авах боломжтой. Энэхүү судалгаа нь SNAP Toolbox систем дээр машин сургалтын санамсаргүй ойн ангиллын алгоритмыг ашиглан газрын бүрхэвчийн ангилал хийхэд, ангиллын параметруудыг хэрхэн тохируулахыг харуулах үндсэн зорилготой бөгөөд нарийвчлалыг сайжруулах үүднээс олон эх сурвалжийн мэдээний сувгуудын 3 өөр янзын хоршлолыг ашиглав. Судалгааны загвар талбай болгон нийслэл Улаанбаатар хотын өмнөд хэсгийг сонгон авч, газрын бүрхэвчийг барилгажсан талбай, ой, бургас, өвс ургамал, сул газар, ус гэсэн үндсэн 6 ангид хуваан ангилсан бөгөөд сигнатурын тоог 5000, 10000, 15000, тэдгээрт харгалзах шийдвэрийн модыг 10, 50, 100 гэж тус тус тодорхойлов. Судалгааны үр дүнгээс харахад, оптикийн ба радарын нийлмэл мэдээг ангилсан үр дүнгүүд илүү өндөр ерөнхий нарийвчлалтай байсан бөгөөд сигнатурын тоог 15000, шийдвэрийн модыг 100 гэж тодорхойлсон 11 сувгийн үр дүн хамгийн өндөр нарийвчлалтай (92.68%) байсан болно.

Түлхүүр үгс: Газрын бүрхэвч, SNAP, Машин сургалт, RF

1. ОРШИЛ

Орчин үед, цахилгаан, соронзон долгионы янз бүрийн мужид хүлээн авсан зайнаас тандсан мэдээг хотын газрын бүрхэвч, түүний орон зай, цаг хугацаа, хэтийн төлөвийн өөрчлөлтийн судалгаанд өргөнөөр ашиглах боллоо. Тандан судлалын анхдагч дүрс мэдээнээс сэдэвчилсэн мэдээллийг сугалан авч, газрын гадаргын өөрчлөлттэй холбоотой дүн шинжилгээ хийх, улмаар үр дүнг шийдвэр гаргах түвшинд ашиглахад төрөл бүрийн аргуудыг хэрэглэдэг [1].

Агаар-сансрын анхдагч өгөгдлөөс газрын бүрхэвчийн талаарх сэдэвчилсэн мэдээллийг гарган авахад өргөн хэрэглэгддэг тоон боловсруулалтын аргуудын тоонд дүрс мэдээний ангилал зүй ёсоор багтана. Дүрс мэдээг ангилахад эрдэмтэд, судлаачид уламжлалт сургалттай, сургалтгүй аргуудаас гадна, янз бүрийн аргыг хэрэглэж, сайжруулан хөгжүүлсээр ирсэн бөгөөд ихэнх судлаачдын хувьд өндөр чанарын үр дүнг гарган авахад, объектод суурилсан, машин сургалттай, хиймэл оюуны зэрэг аргуудыг дангаар нь болон хосолмол хэлбэрээр ашиглаж байна [2].

Газрын бүрхэвчийн автомат ангилалтай холбоотой судалгааны хөгжлийг 1972 онд Landsat-1 дагуулыг хөөргөсөн цаг үеэс эхлэлтэй гэж үздэг. Тус дагуулын мэдээг ангилан, газар ашиглалтын өөрчлөлтийн талаарх анхны шахуу судалгааг Alexander [3] хийсэн ба Ази тивийн нутагт хамаарах ийм төрлийн судалгааг Welch нар [4] 1973 онд хийжээ. Манай улсын хэмжээнд, газрын бүрхэвчийн ангилал хийсэн судалгаа олон байдаг бөгөөд анхны судалгаануудын нэгэнд зүй ёсоор тооцогдох ажлыг М.Ганзориг нар [5] 1990 онд хийжээ. Уг судалгаанд, Сэлэнгэ аймгийн Цагаантолгойн нутгийн газар ашиглалтыг уламжлалт автомат ангиллын аргаар ангилжээ.

Олон төрлийн ангиллын аргуудыг ашиглан газрын бүрхэвчийг ангилж, улмаар харьцуулсан дүн шинжилгээ хийсэн судалгааг Д.Амарсайхан [6] 1996 онд анхлан хийсэн ба тус судалгаанд, Улаанбаатар хот орчмын Landsat-5 дагуулын мэдээг ашиглажээ. Шинэ зууны эхэн үеэс, радарын мэдээний хэрэглээ манай судлаачдын дунд нэмэгдэж, оптикийн мэдээтэй нийлэгжүүлэн ангилах оролдлогууд хийгдэж эхэлсэн юм. Олон эх сурвалжийн мэдээг ангилах анхны судалгааг Д.Амарсайхан нарын судлаачид [7] 2004 онд анхлан хийсэн бөгөөд уг судалгаанд SPOT дагуулын олон сувгийн ба ERS-2 дагуулын радарын нийлмэл мэдээг сайжруулсан ангиллын аргаар ангилсан байна. Сүүлийн үед, газрын бүрхэвчийг ангилах төрөл бүрийн өндөр түвшингийн аргууд бий болсон бөгөөд уг аргуудыг ашиглан Э.Нямжаргал нар [8], Э.Жаргалдалай нар [9], Б.Бямбадолгор нар [10] нэлээн дорвитой судалгаануудыг хийсэн байна.

Зайнаас тандсан мэдээ нь дийлэнх тохиолдолд олон хэмжээсээр илэрхийлэгдсэн их хэмжээний тоон өгөгдөл байх ба ийм мэдээг боловсруулахад мэргэжилтний мэдлэг, ур чадвар, ашиглаж буй программ хангамж, техникийн хүчин чадал зэрэг зүйлс чухал үүрэг гүйцэтгэнэ. Сүүлийн үеийн дүрс боловсруулалтын зарим системд, тухайн шинжээчид дэм болох зорилгоор машин сургалтын болон хиймэл оюуны аргын янз бүрийн алгоритмууд бий болсон бөгөөд эдгээр нь хүн-машины харилцааны зарчим дээр тулгуурлана. Өөрөөр хэлбэл, шинжээч өөрийн онол, практикийн мэдлэг дээр тулгуурлан тухайн алгоритм дахь параметруудын утгуудыг зөв, оновчтой тодорхойлж өгснөөр найдвартай үр дүнг гарган авна.

Энэхүү судалгаа нь Европын сансрын агентлаг (ESA)-ийн хөгжүүлсэн SNAP

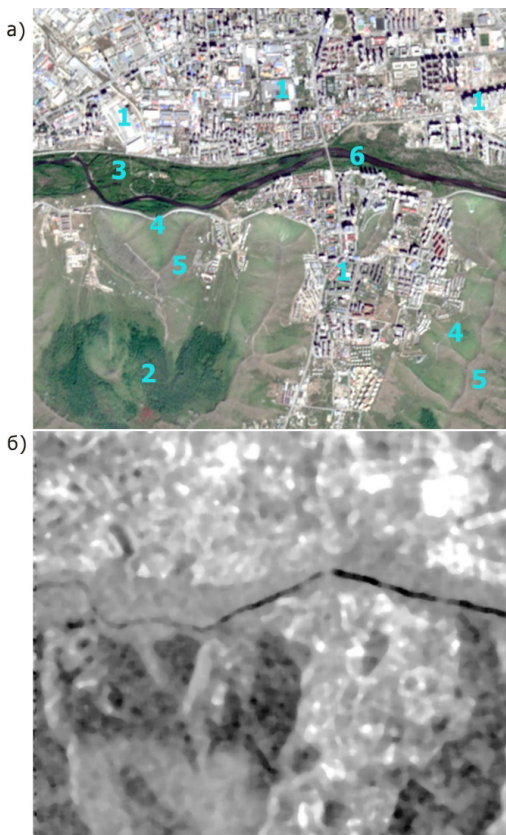
Toolbox систем [11] дээр машин сургалтын санамсаргүй ойн ангиллын алгоритмыг ашиглан ангилал хийхэд, параметруудийг нь хэрхэн тохируулахыг харуулах үндсэн зорилготой бөгөөд дүн шинжилгээнд, Улаанбаатар хотын өмнөд хэсгийн Sentinel-1,2 дагуулын нийлмэл мэдээг ашиглав. Судалгааны хүрээнд, ангиллын нарийвчлалыг сайжруулах үүднээс 3 өөр янзын сувгийн хоршлолыг ашигласан болно.

2. СУДАЛГААНЫ АРГАЗҮЙ

Судалгаанд, Улаанбаатар хотын өмнөд хэсгийг загвар талбай болгон сонгон авсан бөгөөд баруунаас зүүн тийш 4160м, хойноос урагшаа 3700м тус газар нь барилгажсан талбай, ой, бургас, өвс ургамал, сул газар, ус зэрэг газрын бүрхэвчийн ангиудтай. Энэхүү талбайд ой, бургас, өвс ургамал зэрэг ангиуд, түүнчлэн барилгажсан талбай, задгай газар гэсэн ангиуд спектрийн шинж чанарын хувьд өөр хоорондоо өндөр давхцалтай бөгөөд эдгээрийг ялган ангилахад ихээхэн хүндрэлтэй.

Дүн шинжилгээнд, 2021 оны 8-р сарын 25-ны өдрийн Sentinel-2A дагуулын олон бүсчлэлийн мэдээний 10 суваг (B1, B2, B3, B4, B5, B6, B7, B8a, B11, B12) болон 2021 оны 8-р сарын 27-ны өдрийн Sentinel-1B дагуулын C сувгийн (5.55 см) VV туйлшралын 10 м орон зайн шийдтэй, Level 1 GRD бүтээгдэхүүний радарын өгөгдлийг тус тус ашигласан. Sentinel-2A дагуулын эх өгөгдлүүд нь өөр өөр орон зайн шийдтэй ба тэдгээрийг радарын мэдээтэй ижил нарийвчлалтай болгох үүднээс геометрийн засал хийж 10м-ийн шийдтэй болгов.

Зураг 1-д судалгааны талбайн Sentinel-2A, Sentinel-1B дагуулын мэдээнүүд болон газрын бүрхэвчийн ангиудын тархалтыг харуулав.



Зураг 7. Судалгааны талбайн зураг: (а) Sentinel-2A дагуулын мэдээ (Y=B3, H=B2, X=B1) (1-барилгажсан талбай, 2-ой, 3-бургас, 4-өвс ургамал, 5-сул газар, 6-ус) (б) Sentinel-2A дагуулын VV туйлшралын мэдээ.

Сүүлийн үед, машин сургалтын аргууд нэн эрчимтэй хөгжиж хүрээлэн буй орчин судлалын олон салбарт ихээхэн ашиглагдах боллоо [12]. Машин сургалтын загвар нь оролт болон гаралтын өгөгдөл нь мэдэгдэж буй жишээнүүдээс *суралцан*, шинээр оруулсан өгөгдлийг тодорхой үр дүнд хувиргах үйл явц бөгөөд оролтын өгөгдлийн хүлээгдэж буй үр дүнд хамгийн ойр байх дүрслэл буюу төлөвөөс суралцдаг. Машин сургалтын маш олон аргууд байдаг ба тэдгээрийн нэгийг нь

онцлон сайн гэж хэлэх боломжгүй ч, практикт хамгийн өргөн хэрэглэгддэг аргуудын тоонд санамсаргүй ойн арга багтана [13, 14].

Санамсаргүй ойн арга нь орчин үеийн өндөр түвшингийн удирдлагатай машин сургалтын арга бөгөөд ангиллын үндсэн нэгж нь шийдвэрийн мод байдаг. Уг мод нь анхдагч өгөгдлийн шинж чанарууд буюу ижил төстэй байдал дээр тулгуурлан сургалтын өгөгдлийг олон багцуудад санамсаргүй байдлаар хувааж, салаалсан мод хэлбэрийн дүрсийг үүсгэх бөгөөд тэдгээрээс гарсан шийдвэр бүрийг дундажлан нэгтгэх замаар эцсийн шийдийг тодорхойлдог [15, 16]. Тус алгоритм нь дараах алхмуудтай. Үүнд:

Санамсаргүй ойн аргад **P** тооны өгөгдлийн багцаас **m** тооны санамсаргүй өгөгдлийн дээжийг авна.

Шийдвэрийн модыг дээж тус бүр дээр байгуулна.

Шийдвэрийн мод бүрд үр дүнгийн гаралтын утга гарна.

Эцсийн үр дүнг ангиллын болон регрессийн хувьд, давуу шийд эсвэл дундаж утгад үндэслэн тооцно.

Санамсаргүй ойн арга нь нарийвчлал өндөртэй, олон хэмжээсээр илэрхийлэгдсэн их хэмжээний өгөгдөлтэй ажиллахад илүү тохиромжтой байдаг [17]. Энэхүү судалгааны хүрээнд, 4, 10, 11 хэмжээстэй дүрс мэдээг ашигласан бөгөөд сигнатурын тоог 5000, 10000, 15000, шийдвэрийн модыг 10, 50, 100 гэж сонгосон болно.

3. СУДАЛГААНЫ ҮР ДҮН, ХЭЛЭЛЦҮҮЛЭГ

Зайнаас тандсан мэдээний спектрийн ангиллыг хийхийн тулд, газрын бүрхэвчийн ангиудыг төлөөлсөн сигнатурыг дүрсийн янз бүрийн хэсгүүдээс сонгох шаардлагатай ба үүнд

ангилах цэгийн байршил, хоорондын орон зайн хамаарал бас нэгэн чухал үзүүлэлт болдог. Манай судалгаанд, барилгажсан талбай, ой, бургас, өвс ургамал, сул газар, ус гэсэн 6 ангийг төлөөлсөн пикселүүдийг дүрсийн өөр өөр хэсгүүдээс сонгон авав. Уг ангиуд спектрийн зайн хувьд хэр зэрэг саланги байгааг ковариацийн матриц дээр суурилсан сарнилын диаграмм дээр шалгасан бөгөөд эцсийн байдлаар сонгон авсан сигнатурууд нь 192-1685 пикселээс бүрдэж байлаа. Дүн шинжилгээнд дараах сувгуудын хоршлолыг ашигласан болно. Үүнд:

1. 4 суваг (Sentinel-2 дагуулын B2, B3, B8a, B11-р сувгууд).
2. 10 суваг (Sentinel-2 дагуулын B1, B2, B3, B4, B5, B6, B7, B8a, B11, B12-р сувгууд).
3. 11 суваг (Sentinel-2 дагуулын 10 суваг ба Sentinel-1B дагуулын мэдээ).

Ангиллын сэдэвчилсэн үр дүнг газрын үнэн зөв мэдээтэй харьцуулан, ерөнхий нарийвчлалыг шалгахын тулд жигд бус тархалттай, батлагдсан байрлал бүхий пикселүүдийг сонгон авч, тохирлын матрицыг тодорхойлсон бөгөөд нийт 1853 (барилгажсан талбай-861, ой-207, бургас-162, өвс ургамал-179, сул газар-306, ус-138) цэвэр пикселийг сонгов.

3.1. 4 сувгийн ангиллын үр дүн

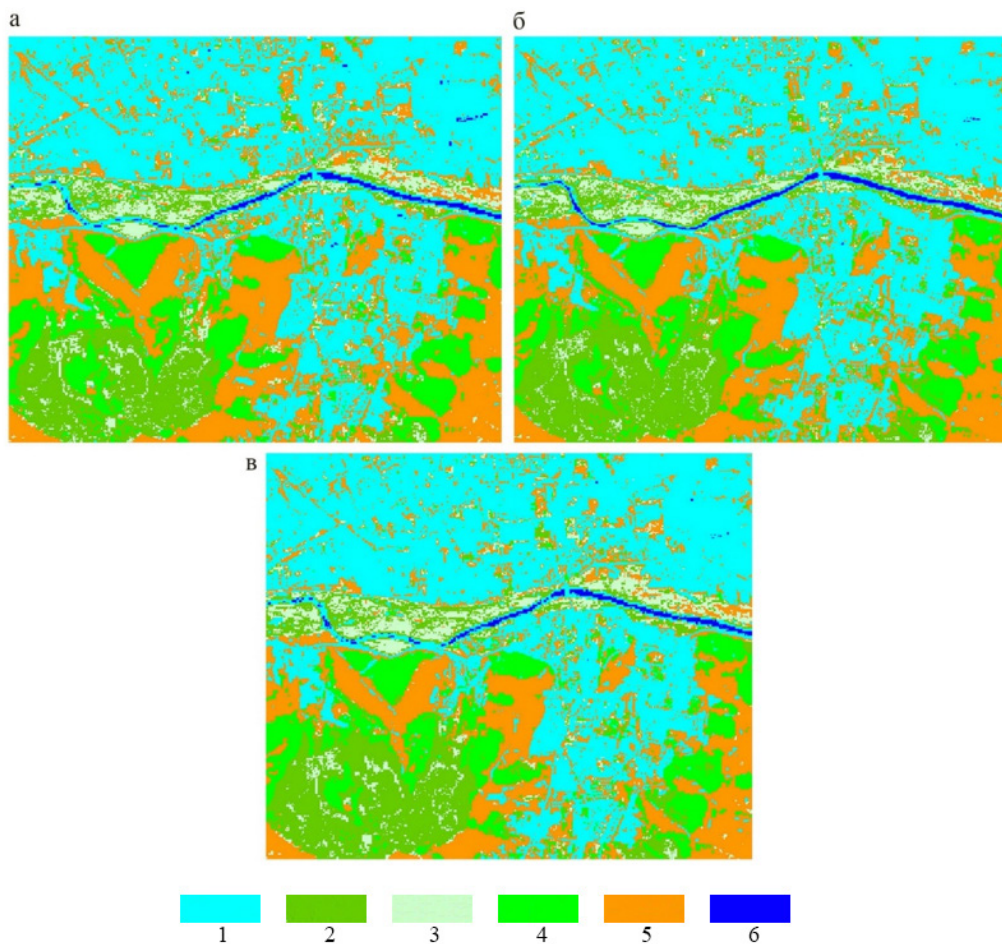
Sentinel-2A дагуулын 2, 3, 8a, 11-р сувгуудын мэдээ нь цахилгаан соронзон долгионы үзэгдэх ногоон гэрэл, улаан гэрэл, ойрын нэл улаан туяа, богино долгионы нэл улаан туяаны мужуудад хамаарах бөгөөд эдгээр долгионы мужуудад манай судалгааны талбай дахь газрын бүрхэвчийн ангиуд нарны энергийг янз бүрээр ойлгоно. Хэрэв тухайн ангиудын спектрийн утгууд ойр байвал ангиудыг өөр хооронд нь ялгахад бэрхшээлтэй байх ба утгуудын ялгаа их

байх тусам ангиллын нарийвчлал өндөр байдаг.

Сигнатурын тоог 5000, 10000, 15000, тэдгээрт харгалзах шийдвэрийн модыг 10, 50, 100 гэж тодорхойлон, санамсаргүй ойн аргаар Sentinel-2A дагуулын 4 сувгийн мэдээг ангилсан үр дүнг Зураг 2-д харуулав. Зургаас харахад, судалгааны талбай дахь хамгийн том ангиуд болох барилгажсан талбай нь сул газар, ус зэрэг ангиудтай, ой нь бургас, өвс ургамал гэсэн ангиудтай давхцал хамгийн их

байгаа бол сул газар, ногоон ургамал гэсэн ангиудын хувьд харьцангуй бага хэмжээний давхцалтай байгаа нь харагдаж байна.

Дээр дурдсан давхцалууд, сигнатурын тоог 5000, 15000 гэж авсан дүнгүүд дээр их байгаа бол 10000 дээр харьцангуй багасжээ. Энэ нь ялангуяа Туул голын ангилагдсан зураг дээр тодорхой харагдаж байх бөгөөд уг мэдээн дээр голын ус тасралтгүй шахам үргэлжилж байна.



Зураг 2. 4 сувгийн мэдээг ангилсан дүн (1-барилгажсан талбай, 2-ой, 3-бургас, 4-өвс ургамал, 5-сул газар, 6-ус)

Нарийвчлалын үнэлгээнд, сонгон авсан 1853 жигд бус байрлал бүхий, цэвэр пикселийг ашиглан үнэлгээ хийхэд, сигнатурын тоог 5000, 15000 гэж авсан дүнгүүд дээр ерөнхий нарийвчлал нь 88.05%, 88.86%, харин сигнатурын тоог 10000 гэж авсан дүн дээр нарийвчлал нь бага зэрэг сайжирч 89.78% болсон байлаа.

3.2. 10 сувгийн ангиллын үр дүн

Sentinel-2A дагуулын 10 сувгийн мэдээ нь цахилгаан соронзон долгионы хөх гэрэл, ногоон гэрэл, улаан гэрэл, ойрын нэл улаан туяа, богино долгионы нэл улаан туяаны мужуудыг хамаарна. Гэхдээ Sentinel-2A дагуулын 4 сувгийн мэдээтэй харьцуулахад улаан гэрэл болон ойрын нэл улаан туяаны завсрын сувгуудыг агуулахаас гадна, богино долгионы нэл улаан туяаны муж нь илүү урт долгионы уртыг хамарна.

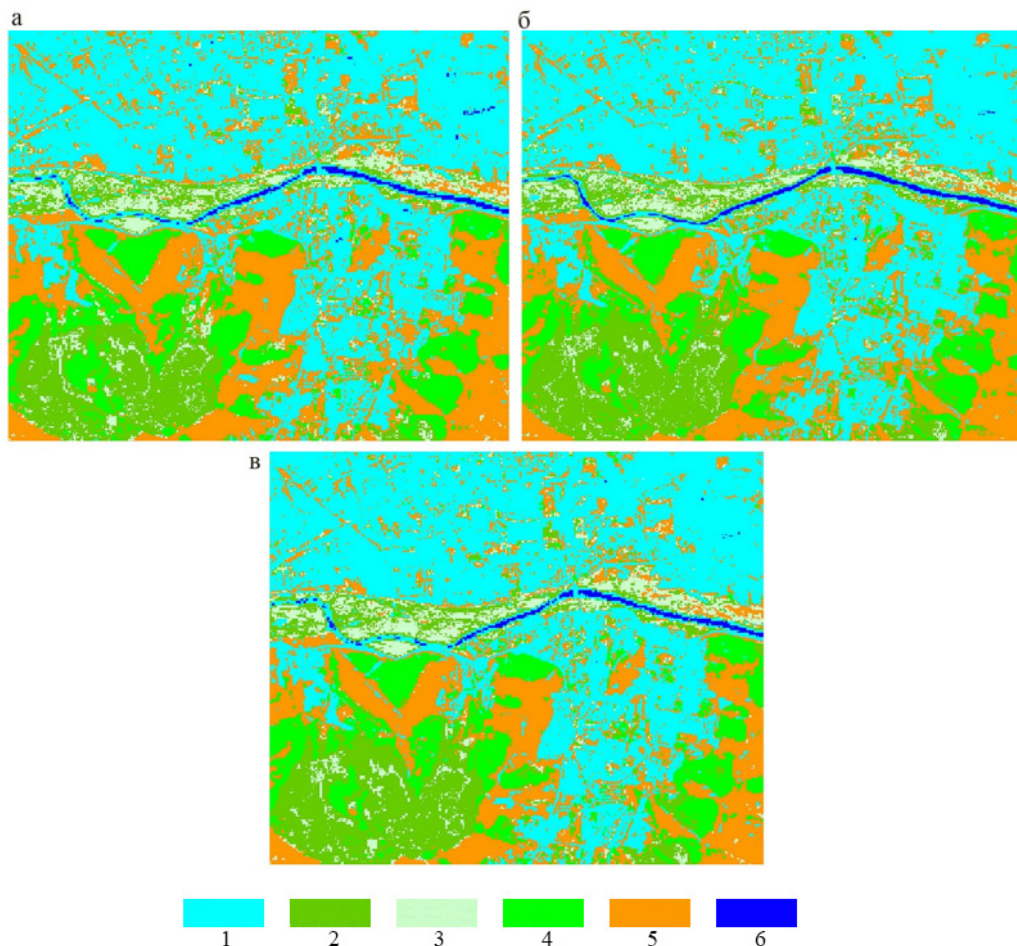
Сигнатурын тоог 5000, 10000, 15000, тэдгээрт харгалзах шийдвэрийн модыг 10, 50, 100 гэж тодорхойлон, санамсаргүй ойн аргаар Sentinel-2A дагуулын 10 сувгийн мэдээг ангилсан үр дүнг Зураг 3-д харуулав. Зургаас харахад, Зураг 2 дээр өндөр давхцалтай байсан барилгажсан талбай ба сул газар, түүнчлэн ой ба бургас гэсэн ангиудын өөр хоорондын ялгаа илүү сайжирсан байх бөгөөд энэ нь сигнатурын тоог 10000, 15000 гэж авсан дүнгүүд дээр илүү тод ажиглагдаж байна. Харин Туул голын хувьд, сигнатурын тоог 15000 гэж авсан дүн дээр голын ус тасралтгүй шахам үргэлжилжээ. Нарийвчлалын үнэлгээнд, сонгон авсан 1853 жигд бус байрлал бүхий пикселийг ашиглан үнэлгээ хийхэд Зураг 2-д үзүүлсэн үр дүнгүүдээс сайжирч, сигнатурын тоог 5000, шийдвэрийн модыг 10 гэж авсан дүн дээр ерөнхий нарийвчлал нь 90,01%, 10000 гэж авсан дүн дээр нарийвчлал нь 90,89%, харин 15000 гэж авсан дүн дээр ерөнхий нарийвчлал нь 91.78% болсон байлаа.

3.3. 11 сувгийн ангиллын үр дүн

Sentinel-2 дагуулын 10 сувгийн мэдээ нь цахилгаан соронзон долгионы ямар мужуудад хамаардгийг дээр дурдсан билээ. Харин Sentinel-1B дагуулын мэдээ нь радарын 5.55см долгионы мужид хүлээн авсан тоон өгөгдөл байдаг бөгөөд уг долгионы уртад манай загвар талбай дахь газрын бүрхэвчийн ангиудад янз бүрийн сарнилын механизм үйлчилнэ. Тухайлбал, байшин барилга гэсэн ангийн хувьд давхар ойсон сарнил давамгайлах бол ойн ангийн хувьд эзлэхүүнт сарнил давамгайлна.

Сигнатурын тоог 5000, 10000, 15000, тэдгээрт харгалзах шийдвэрийн модыг 10, 50, 100 гэж тодорхойлон, санамсаргүй ойн аргаар 11 сувгийн мэдээг ангилсан үр дүнг Зураг 4-д харуулав. Зургаас үзэхэд, Зураг 2 ба 3 дээр давхцалтай байсан ангиудын өөр хоорондын ялгаа илүү сайжирсан байна. Жишээ нь, барилгажсан талбай, бургас, сул газар, ус гэсэн ангиудын хувьд бусад ангиудтай давхцал нь харьцангуй багасаж, илүү үнэн зөв ангилагдсан байдалтай харагдаж байна. Туул голын хувьд, бүх дүнгүүд дээр голын ус тасралтгүй шахам үргэлжилж байгаа бол сигнатурын тоог 15000 гэж сонгосон зураг дээр барилгажсан талбай илүү нарийвчлал сайтай ангилагдсан нь харагдаж байна.

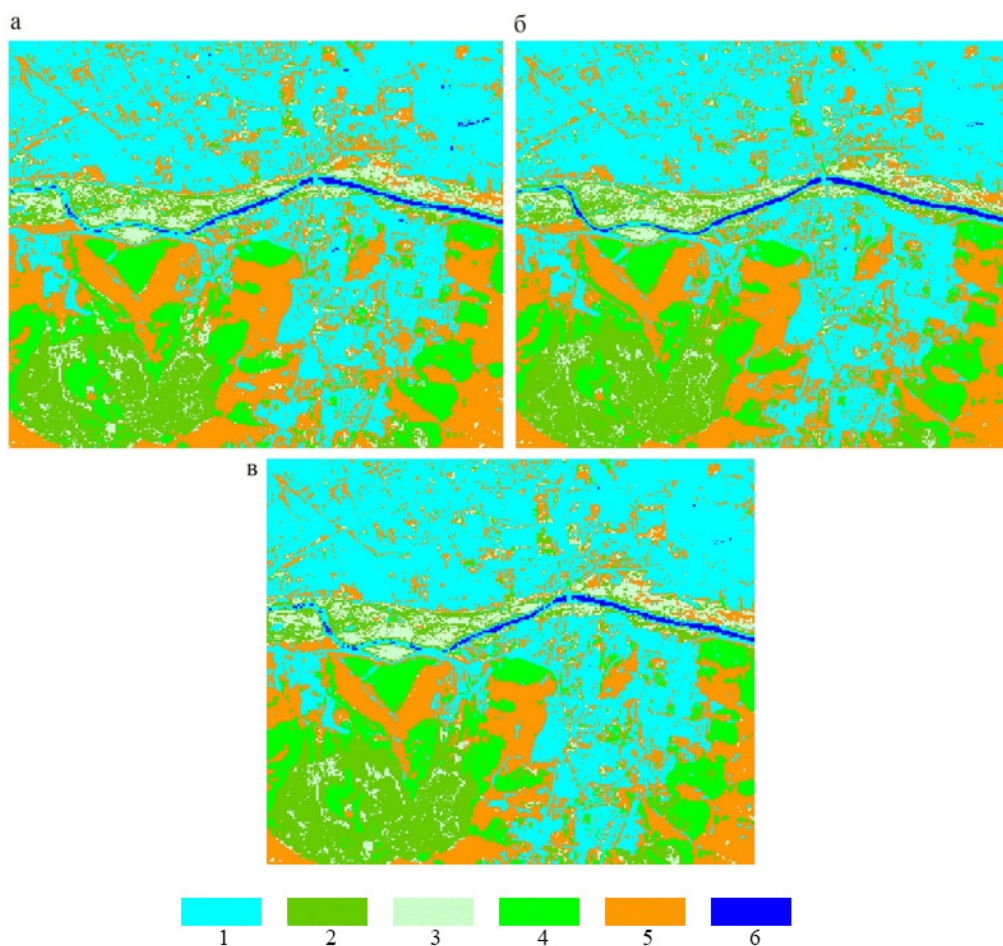
Нарийвчлалын үнэлгээнд, сонгон авсан 1853 жигд бус байрлал бүхий, цэвэр пикселийг ашиглан үнэлгээ хийхэд Зураг 3-д үзүүлсэн үр дүнгүүдээс сайжирч, сигнатурын тоог 5000, шийдвэрийн модыг 10 гэж сонгосон дүн дээр ерөнхий нарийвчлал нь 91,18%, сигнатурын тоог 10000, шийдвэрийн модыг 50 гэж авсан дүн дээр нарийвчлал нь 91,85%, харин сигнатурын тоог 15000, шийдвэрийн модыг 100 гэж сонгосон дүн дээр ерөнхий нарийвчлал нь 92.68% болсон байв.



Зураг 3. 10 сувгийн мэдээг ангилсан дүн (1-барилгажсан талбай, 2-ой, 3-бургас, 4-өвс ургамал, 5-сул газар, 6-ус)

Zheng нар [18] Канадын Ниагра хотын газрын бүрхэвчийг зайнаас тандсан мэдээ ашиглан санамсаргүй ойн аргаар ангилахад, шийдвэрийн мод нь их байх тусам ангиллын нарийвчлал илүү сайжирч байсан бол Amini нар [19] Ираны Исфахан хотын газрын бүрхэвчийг Landsat дагуулын цуврал мэдээ ашиглан санамсаргүй ойн аргаар ангилж үзэхэд, сувгийн тоог нэмэгдүүлэхэд ангиллын нарийвчлал мөн сайжирч байсан байна. Sun нар [20] санамсаргүй ойн оновчтой параметрыг судлах зорилгоор БНХАУ-

ын Нанжин хотын ойролцоох газрын бүрхэвчийг Sentinel-2 мэдээ ашиглан ангилахад, шийдвэрийн модны тоо 800 үед хангалттай өндөр нарийвчлал бүхий үр дүн гарчээ. Манай судалгааны үр дүнг нэгтгэн Хүснэгт 1-д үзүүлэв. Хүснэгтээс харахад, сигнатурын тоог 5000, шийдвэрийн модыг 10 гэж авсан 4 сувгийн хоршлолын ангиллын ерөнхий нарийвчлал хамгийн тааруу байгаа бол сигнатурын тоог 15000, шийдвэрийн модыг 100 гэж сонгосон 11 сувгийн хоршлолыг ангилсан дүнгийн нарийвчлал нь хамгийн өндөр байна.



Зураг 4. 11 сувгийн мэдээг ангилсан дүн (1-барилгажсан талбай, 2-ой, 3-бургас, 4-өвс ургамал, 5-сул газар, 6-ус

Хүснэгт 1. Ангиллын үр дүнгүүдийн ерөнхий нарийвчлал

	Сигнатур-5000 шийдвэрийн мод-10	Сигнатур-10000 шийдвэрийн мод-50	Сигнатур-15000 шийдвэрийн мод-100
4 суваг	88.05%	89.78%	88.86%
10 суваг	90.01%	90.89%	91.78%
11 суваг	91.18%	91.85%	92.68%

4. ДҮГНЭЛТ

Судалгааны хүрээнд, SNAP Toolbox систем дээр санамсаргүй ойн алгоритмыг ашиглан газрын бүрхэвчийн ангилал хийхэд, параметруудыг хэрхэн зөв сонгон тохируулахыг харуулахыг зорьсон болно. Дүн шинжилгээнд, Sentinel-2A дагуулын олон бүсчлэлийн мэдээний 10 суваг, Sentinel-1B дагуулын C сувгийн VV туйлшралаар хүлээн авсан радарын өгөгдлийг тус тус ашиглан Улаанбаатар хотын өмнөд хэсгийн нутгийг барилгажсан талбай, ой, бургас, өвс ургамал, сул газар, ус гэсэн үндсэн 6 ангид хуваан ангилсан. Ангиллын нарийвчлалыг сайжруулах үүднээс газрын бүрхэвчийг 4 сувгийн, 10 сувгийн, 11 сувгийн хоршлолуудыг ашиглан ангилсан ба сигнатурын тоог 5000, 10000, 15000, тэдгээрт харгалзах шийдвэрийн модыг 10, 50, 100 гэж тус тус тодорхойлсон болно.

Нарийвчлалын үнэлгээнд, жигд бус тархалттай, үнэн зөв байрлал бүхий 1853 цэвэр пикселийг сонгон авч, тохирлын матрицыг ашиглав. Судалгааны үр дүнгээс харахад, сигнатурын тоог 5000, шийдвэрийн модыг 10 гэж авсан 4 сувгийн хоршлолын ангиллын ерөнхий нарийвчлал хамгийн бага (88,05%) байсан бол сигнатурын тоог 15000, шийдвэрийн модыг 100 гэж сонгосон 11 сувгийн хоршлолын дүн дээр ерөнхий нарийвчлал нь хамгийн өндөр (92,68%) байлаа. Энэ нь SNAP Toolbox систем дээр санамсаргүй ойн аргаар ангилал хийхэд тухайн өгөгдлийн бүтцээс шалтгаалан параметруудыг хэрхэн зөв сонгон хэрэгтэйг харуулж байгаа юм.

АШИГЛАСАН МАТЕРИАЛ

[1]. R. B. Salama, "Remote sensing of soils and plants imagery," in Encyclopedia of Agrophysics, J. Gliński, J. Horabik, and J. Lipiec, Eds. Springer, Dordrecht, 2011. https://doi.org/10.1007/978-90-481-3585-1_132.

- [2]. A. Enkhmanlai, E. Nyamjargal, D. Amarsaikhan, O. Munkhdulam, and B. Batbileg, "Estimation and mapping of pasture biomass in Mongolia using machine learning methods," *Geocarto International*, vol. 38, no. 1, pp. 1–17, 2023. <https://doi.org/10.1080/10106049.2023.2195824>.
- [3]. R. H. Alexander, "Land use classification and change analysis using ERTS-1 imagery in CARETS". Symposium on significant results of remote sensing obtained from ERTS-1. Sec. B. pp. 923-930, 1973.
- [4]. R. Welch, C.W.Pannell, C.P.Lo, "Land use in northeast China, 1973: A view from Landsat-1". *Annals of the Association of American Geographers*, 65(4), 595–596. 1973.
- [5]. M. Ganzorig, D. Amarsaikhan, and B. Enkhtuvshin, "An application of remote sensing and GIS technique in Mongolia," in Proceedings of European 'International Space Year' Symposium 1992, Munich, Germany, pp. 387–390, 1992.
- [6]. D. Amarsaikhan, "Update of a GIS by RS data using a knowledge-based approach". Abstract of PhD Dissertation, Mongolian Academy of Sciences, pp.25, 1997.
- [7]. D. Amarsaikhan, T. Douglas, "Data fusion and multisource data classification". *International Journal of Remote Sensing*, No.17, Vol.25, pp.3529-3539. <https://doi.org/10.1080/0143116031000115111>.
- [8]. Э. Нямжаргал, С. Наранбат, Д. Амарсайхан, "Object-oriented classification of satellite imagery and lidardata". Хүрэлтогоот-2018 эрдэм шинжилгээний хурлын эмхэтгэл, Улаанбаатар, pp.146-151, 2018.
- [9]. Э. Жаргалдалай, А. Мөнх-Эрдэнэ, Д. Амарсайхан, ба Г. Адъяа-Очир,

- "Олон бүсчлэлийн хиймэл дагуулын мэдээ ашиглан газрын бүрхэвч болон газар ашиглалтыг машин сургалттай аргаар ангилсан дүн," Монгол Орны Газарзүй-Геоэкологийн Асуудлууд, vol. 42, ШУА-ийн Газарзүй-Геоэкологийн хүрээлэн, 2021.
- [10]. Б. Бямбадолгор, Д. Амарсайхан, and Э. Нямжаргал, "MODIS дагуулын мэдээ ашиглан Монгол орны ойт хээр болон хээрийн бүсийн ургамлын биомассыг үнэлэн зураглах нь," Монгол Орны Газарзүй-Геоэкологийн Асуудлууд, vol. 44, ШУА-ийн Газарзүй-Геоэкологийн хүрээлэн, 2023.
- [11]. ESA Sentinel Toolbox, ESA, 2015. [Online]. Available: <http://step.esa.int>.
- [12]. G. C. Rosset and A. Klarsfeld, "Diversity, equity, and inclusion in artificial intelligence: An evaluation of guidelines," *International Journal of Applied Artificial Intelligence*, 2023. <https://doi.org/10.1080/08839514.2023.2176618>.
- [13]. G. James, D. Witten, T. Hastie, and R. Tibshirani, *Introduction to Statistical Learning with Applications in R*. Springer, 2013.
- [14]. H. Shihab, D. Stowa, K. Chang, D. Roberts, and H. Goulias, "From land cover to land use: Applying random forest classifier to Landsat imagery for urban land-use change mapping," *Geocarto International*, pp. 83–99, 2021. <https://doi.org/10.1080/10106049.2021.1897345>.
- [15]. C. Zheng, A. Abd-Elrahman, V. Whitaker, and C. Dalid, "Prediction of strawberry dry biomass from UAV multispectral imagery using multiple machine learning methods," *Remote Sensing*, vol. 14, no. 18, 4511, 2022. <https://doi.org/10.3390/rs14184511>
- [16]. Y. Tian, H. Huang, G. Zhou, and Q. Zhang, "Aboveground mangrove biomass estimation in Beibu Gulf using machine learning and UAV remote sensing," *Science of The Total Environment*, vol. 781, 146816, 2021. <https://doi.org/10.1016/j.scitotenv.2021.146816>.
- [17]. S. Chakraborty, "Bayesian additive regression tree for seemingly unrelated regression with automatic tree selection". *Handbook of Statistics, Vol. 35*, 229-251, 2016.
- [18]. G. Zheng, L. Zhong, Y. Li, H. Guan, "A random forest based method for urban object classification using lidar data and aerial imagery," *23rd International Conference on Geoinformatics, Wuhan, China*, 1-4, 2015.
- [19]. S. Amini, M. Saber, H. Rabiei-Dastjerdi, and S. Homayouni, "Urban land use and land cover change analysis using random forest classification of Landsat time series," *Remote Sensing*, vol. 14, no. 11, p. 2654, 2022. <https://doi.org/10.3390/rs14112654>.
- [20]. J. Sun and S. Ongsomwang, "Optimal parameters of random forest for land cover classification with suitable data type and dataset on Google Earth Engine," *Frontiers in Earth Science*, vol. 11, 1188093, 2023. <https://doi.org/10.3389/feart.2023.1188093>.