

Амны Хаалттай болон Амны Хаалтгүй Хүний Царайг Бодит Хугацаанд Таних Программын Хөгжүүлэлт

Нямдаваагийн Даваасүрэн^{1*}, Бадарчийн Луубаатар², Өлзийсайханы Баттулга¹, Дуламрагчаагийн Ууганбаатар¹, Батцогтын Мөнхбаяр¹

¹Шинжлэх Ухааны Академи, Математик, тоон технологийн хүрээлэн, Улаанбаатар 13330, Монгол улс

²Шинжлэх Ухаан Технологийн Их Сургууль, Мэдээлэл, Холбооны Технологийн Сургууль, Улаанбаатар 13341, Монгол улс

*Холбоо барих зохиогч: davaasuren_n@mas.ac.mn; ORCID:0000-0002-2843-3700

Өгүүллийн мэдээлэл: Хүлээн авсан: 2022.09.11; Зөвшөөрөгдсөн: 2022.10.01;
Нийтлэгдсэн: 2022.12.26

Хураангуй: Ковид 19 цар тахлын эрсдэлээс сэргийлэхийн тулд хүмүүс амны хаалт хэрэглэх болсон. Амны хаалт нь хүний нүүрний ихэнх хэсгийг далдалдагас үүдэн царай таних алгоритмууд амны хаалттай хүнийг муу таних хандлагатай байдаг. Үүнийг дагаад амны хаалттай хүнийг хэн бэ? гэдгийг таних царай танилтын алгоритмыг боловсруулах хэрэгцээ үүссэн. Бид аль болох үр дүнтэй, өндөр нарийвчлалтай, бодит хугацаанд ажиллаж болохуйц амны хаалттай болон амны хаалтгүй хүний таних алгоритмыг гарган авах, түүн дээр тулгуурлан программ хөгжүүлэх зорилготойгоор ажилласан бөгөөд энэхүү өгүүллээр бид бодит хугацаанд амны хаалттай болон амны хаалтгүй хүний царай таних алгоритм ба программын хөгжүүлэлтийн өнөөдрийг хүртэлх зарим үр дүнг танилцуулна.

Түлхүүр үгс: Casia-Webface, Facenet, Гүн сургалт

1. Оршил

Царай танилт нь амьдрал дээр маш өргөн хэрэглээтэй, олон жил судлагдаж буй сэдвүүдийн нэг юм. Гүн сургалтын аргууд хүч түрэн орж ирэхээс өмнө судлаачид уламжлалт дүрс боловсруулалтын аргуудыг ашиглан царай таних алгоритмыг боловсруулдаг байсан (Жишээлбэл: Gabor [1], Eigenface [2]). Эдгээр аргууд нь гараар үүсгэсэн шүүлтүүрүүдийг онцлог задлагч (feature extractor) болгон ашигладаг бөгөөд шүүлтүүрүүдийн задалсан онцлогуудыг царай танилтад ашигладаг. Гэвч эдгээр аргууд нь гэрэлд маш өртөмтгий, байрлалд инвариант биш зэрэг олон сул талуудтай. Уламжлалт дүрс боловсруулалтын аргуудыг ашигласан царай таних аргуудыг судлаачид хөгжүүлсээр LFW [3] өгөгдлийн сан дээрх хамгийн сайн нарийвчлалыг 95%-д хүргэсэн [4].

2012 онд Alexnet [5] загвар ImageNet тэмцээнд гүн сургалтын арга ашиглан тэргүүн байр эзэлснээс хойш гүн сургалтын аргууд компьютер хараанд маш өргөнөөр ашиглагдах болсон. Гүн сургалтыг ашиглан царай танилт дээр өндөр үр дүн үзүүлсэн анхны арга нь DeepFace [6] юм (LFW өгөгдлийн сан дээр DeepFace: 97.35%, Хүн: 97.53%). DeepFace нь дөрвөн сая өгөгдөл дээр суралцсан есөн давхарга бүхий загвар. DeepFace-н дараагаар гүн сургалтыг ашигласан DeepID2 [7], FaceNet [8], Baidu [9], VGGFace [10], SphereFace [11], Arcface [12], Cosface [13] зэрэг загварууд танилцуулагдсан нь бүгд LFW өгөгдлийн сан дээр 99%-с өндөр нарийвчлал үзүүлсэн.

Гэвч эдгээр царай таних гүн сургалтын аргуудаас амны хаалттай царай танилт дээр хамгийн өндөр нарийвчлал үзүүлсэн нь RMFRD өгөгдлийн сан дээр VGGFace: 68.17% [14] байна. Бид аль болох үр дүнтэй, өндөр нарийвчлалтай, бодит хугацаанд ажиллаж болохуйц амны хаалттай болон амны хаалтгүй хүнийг таних алгоритмыг гарган авах, түүн дээр тулгуурлан программ хөгжүүлэх зорилготойгоор ажилласан бөгөөд энэхүү өгүүллээр бид бодит хугацаанд амны хаалттай болон амны хаалтгүй хүний царай таних алгоритм ба программын хөгжүүлэлтийн өнөөдрийг хүртэлх зарим үр дүнг танилцуулна.

2. Судалгаанд ашигласан өгөгдөл ба материал

Бид загварыг сургах өгөгдлийн сангаар олон нийтэд нээлттэй Casia-Webface [15] өгөгдлийн санг сонгосон. Уг өгөгдлийн сан нь 10575 ялгаатай хүний 494414 зураг бүхий өгөгдлийн сан юм. Бид уг өгөгдлийн санг дараах алхмуудаар боловсруулсан.

1. SSD ашиглан зургуудаас хүмүүсийн царайг хайж олох, олдсон царайнуудыг хадгалах.
2. Хадгалж авсан царайнуудыг Facenet ашиглан шүүлтүүрдэх. Өмнөх ажилд бид Naar-eye ашиглан зураг бүрээс нүд илрүүлж, хоёроос бага тооны нүд илэрсэн зургуудыг устгах замаар өгөгдлийн санг шүүлтүүрдсэн бол энэ удаад нээлттэй эхийн Facenet царай таних загварыг ашигласан. Casia-Webface өгөгдлийн сан нь 10575 ялгаатай хүний 494414 ширхэг зургаас бүрдэх бөгөөд хүн тус бүрд харгалзах хавтас дотор тухайн хүнд харгалзах зургууд байрладаг байхаар зохион байгуулагдсан өгөгдлийн сан юм. Өөрөөр хэлбэл 10575 ширхэг хавтас бүхий өгөгдлийн сан гэсэн үг юм. Facenet загварыг ашиглан Casia-Webface өгөгдлийн санг хавтас хавтсаар шүүнэ. Хавтас доторх сонгогдсон зургийн шигтгээ векторыг Facenet загварыг ашиглан гаргаж авах ба уг шигтгээ вектор болон хавтас дахь бусад зургуудаас гарган авсан шигтгээ векторуудын хоорондох Евклидийн зайн дундаж нь 1.25-аас их бол сонгогдсон зургийг устгах замаар цэвэрлэсэн. Хавтас доторх зургууд бүгд нэг хүний зураг тул тэдгээрийн шигтгээ векторуудын хоорондох Евклидийн зайн дундаж нь бага байх ёстой. Энэхүү боловсруулалтын дараа өгөгдлийн сан 10567 ялгаатай хүний 462506 зураг болж бага зэрэг багассан.
3. Шүүгдэн үлдсэн цэвэр өгөгдлүүдийг MaskTheFace [18] ашиглан амны хаалт зохиомлоор зүүлгэж амны хаалттай өгөгдлийн сан гарган авсан. MaskTheFace dlib-н царайг 64 цэгээр загварчлах аргыг ашиглан царайг загварчилж загварчилсан царайнаас ам, хамрын хэсгүүдийг олж, амны хаалтыг харгалзуулан зүүлгэдэг. Хэрвээ dlib-н ашигладаг царай олох арга өгөгдсөн зургаас царай олж чадахгүй бол тус зураг бас устана. Тиймээс энэхүү боловсруулалтын дараа өгөгдлийн сангийн хэмжээ 10567 ялгаатай хүний 393037 болж багассан.



Зураг 1: MaskTheFace ашиглан амны хаалт зүүлгэхээс өмнөх ба зүүлгэсний дараах зургуудын жишээ.

3. Арга зүй

Манай багийнхан 2021 онд амны хаалттай хүний царайг нүд, дух орчмын хэсгүүдийг ашиглан таних арга боловсруулсан [16]. Yolov5 объект илрүүлэх аргыг ашиглан зургаас хүний царайг амны хаалттай амны хаалтгүйгээр нь ангилан олох ба царайнуудыг олсны дараа амны хаалттай царайнуудыг өөрсдийн сургасан загвараар, амны хаалтгүй царайнуудыг Facenet, DeepFace зэрэг царай таних аргуудыг ашиглан таних төсөөлөлтэй байсан. Хүнийг нүд дух орчмын хэсгүүдээр таних бидний загвар нь Resnet 50-д суурилсан бөгөөд 24 сая параметртэй. Таних нарийвчлал 86%.



Зураг 2: Амны хаалттай, амны хаалтгүй хүнийг таних програмын кодын ажиллах зарчмын анхны төсөөлөл.

Гэвч бидний боловсруулсан арга нь гурван ширхэг гүн сургалтын том загваруудыг агуулсан тул бодит хугацаанд ажиллахын тулд маш их тооцоолон бодох хүчин чадал шаардана. Цаашлаад программ хангамж болж хөгжүүлэгдэх загвар тул дундаж үзүүлэлттэй компьютер дээр ч бодит хугацаанд ажилладаг байх ёстой. Тиймээс бид эхлээд Зураг 2-т үзүүлсэн алгоритмын гүн сургалтын загваруудыг илүү хөнгөн буюу параметрийн тоо цөөн загваруудаар солих боломжийг эрэлхийлсэн. Хүмүүсийг амны хаалттай хаалтгүйгээр нь ангилах хэсгийн Yolov5 загварыг арай жижиг загвар болох SSD [17]-ээр сольсон. Зургаас амны хаалттай болон амны хаалтгүй царайнуудыг олох асуудал нь гүн сургалтын объект илрүүлэх аргуудын хувьд тийм ч төвөгтэй даалгавар биш юм. Гүн сургалтын объект илрүүлэх ихэнх арга уг асуудлыг маш өндөр нарийвчлалтайгаар шийдэж чадна. Тийм болохоор ажиллах хурдны хувьд бараг ялгаагүй ч гэсэн санах ойд эзлэх зай бага тул SSD загварыг сонгосон. Гэсэн ч алгоритм гурван гүн сургалтын загвартай хэвээрээ л байсан тул бид гүн сургалтын загваруудын тоог багасгах боломжийг эрэлхийлсэн. Зураг 3-т гүн сургалтын царай таних загваруудыг амны хаалттай зургууд дээр туршсан [14] үр дүнг харууллаа.



Зураг 3: Гүн сургалтын царай таних аргуудыг амны хаалттай RMFRD өгөгдлийн сан дээр үнэлсэн нарийвчлал.

Эдгээр загваруудыг RMFRD (Real World Masked Face Recognition Dataset) өгөгдлийн сан дээр үнэлсэн ба уг өгөгдлийн сан нь 525 ялгаатай хүний амны хаалттай 5000 зургийг агуулсан өгөгдлийн сан юм. Уг загваруудын амны хаалттай хүнийг таних нарийвчлалыг үнэлэхдээ зөвхөн амны хаалттай зургуудыг ашигласан, RMFRD өгөгдлийн сан нь нэлээд “бохир” өгөгдлийн сан юм. Зураг 4-т RMFRD өгөгдлийн сангийн 108 дугаартай хавтаснаас түүвэрлэн авсан зургуудыг харууллаа. Зураг 4-т үзүүлсэн зургууд бүгд нэг хүний зургууд боловч хүн нүдээр хараад ч нэг хүний зургууд гэж хэлэхэд хүндрэлтэй байгааг харж болно.

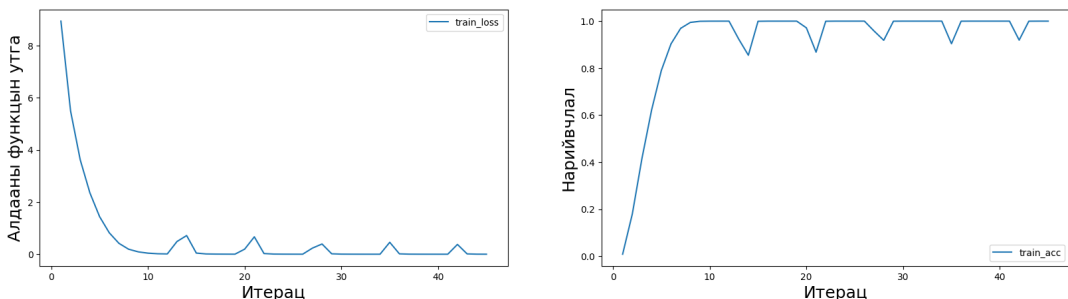


Зураг 4: RMFRD 108-р хавтас дахь ижил хүний зургууд.

Энэ мэт маш олон жишээ байгаа нь загваруудын нарийвчлалыг багасгах хамгийн том хүчин зүйл юм. Эдгээрийг харгалзан үзвэл 68% гэдэг бол тийм ч муу нарийвчлал биш гэж бид үзсэн. Мөн бүртгэлийн систем, баталгаажуулах систем зэрэг царай таних систем хэрэглэдэг бодит амьдралын асуудлууд дээр амны хаалттай хүнийг таних даалгавар нь RMFRD өгөгдлийн сан дээр танилт хийхээс амархан даалгавар юм. Учир нь дээрх тохиолдлуудад тухайн хүн камерт ойрхон байх ба ихэнх тохиолдолд яг камерын өдөөс харж зургаа авхуулна. Мөн тухайн хүний өгөгдлийн сан дахь зураг ч гэсэн ихэнх тохиолдолд дээрх шаардлагыг хангасан байна. Эдгээрийг харгалзан үзээд бид нээлтэй эхийн царай таних загваруудыг амны хаалттай зургууд ашиглан дахин сургаж гарч ирсэн загварыг хүн амны хаалттай байх эсэхээс үл хамааран танилт хийх загвар болгон ашиглах шийдэлд хүрсэн. Энэхүү шийдэл нь программыг илүү хурдан, хөнгөн болгоно. Тиймээс бид зохиомлоор амны хаалт зүүлгэж дахин боловсруулсан Casia-Webface өгөгдлийн санг ашиглан Facenet загварыг дахин сургасан.

4. Үр дүн

Бид Facenet загварыг дахин сургахдаа амны хаалттай болон амны хаалтгүй өгөгдлийг бүгдийг нь ашигласан буюу 10567 ялгаатай хүний 855543 ширхэг зургаар сургалтыг явуулсан. Сургалт явуулах цагийг хэмнэх зорилгоор туршилтын болон баталгаажуулах өгөгдөл ашиглаагүй. Сургалтыг Математик, тоон технологийн хүрээлэнгийн тооцон бодох төвийн техник хангамжуудыг ашиглан явуулсан. Оптимайзэрээр Адам-г сонгон авсан бөгөөд 0.0005 суралцах хурдтайгаар 1024 багцын хэмжээтэйгээр TESLA-V100 график тооцоолуурын тусламжтай 45 итерац сургасны дараах сургалтын үзүүлэлтүүдийг Зураг 5-т харуулав. Нэг итерац хийхэд 4083.88 секунд буюу 1.13 цаг зарцуулсан бөгөөд сургалт нийт 50.85 цаг үргэлжилсэн.



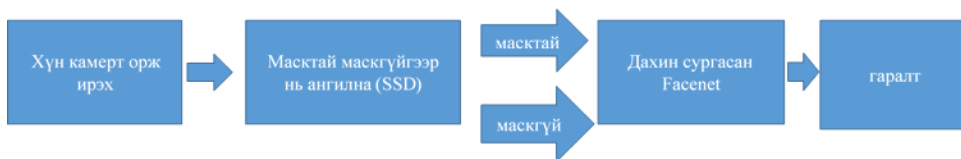
Зураг 5: Сургалтын үеийн алдааны функцийн утга болон нарийвчлал.

12 дахь итерацаас хойш загвар ямар нэгэн локал минимум руу орж, гараад байгаа юм шиг дүр зураг үзүүлсэн тул сургалтыг зогсоосон. Сургалтын өгөгдөл дээрх хамгийн өндөр нарийвчлал 0.9999076609825572 хүрсэн. Загварыг мөн сургалтад оролцоогүй 1000 хүний хүн тус бүрийн нэг зураг агуулсан өгөгдлийн санг ашиглан үнэлсэн. Үнэлэхдээ 1000 хүнд амны хаалт зохиомлоор зүүлгэн амны хаалттай өгөгдөл гарган авсан. Амны хаалттай өгөгдлүүдийг таних өгөгдөл болгосон ба амны хаалтгүй өгөгдлүүдийг харьцуулах (reference) өгөгдөл болгон авч загварыг үнэлэхэд загварын нарийвчлал 99.6% байсан. FaceNet загварын нээлттэй эхийн хоёр хувилбар байдаг. Тэдгээрийг өөрсдийн загварыг үнэлсэн өгөгдлийн сан дээр үнэлж, өөрсдийн загвартай харьцуулсан үр дүнг Хүснэгт 1-д харууллаа.

Хүснэгт 1: Загваруудын амны хаалттай өгөгдөл дээрх харьцуулалт.

	Сургасан өгөгдлийн сан	Нарийвчлал
Facenet (20180408-102900)	Casia-Webface	45.11%
Facenet (20180402-114759)	VGGFace2	60.49%
Бидний сургасан хувилбар	Casia-Webface + амны хаалттай Casia-Webface	99.6%

Бид гарган авсан загвар дээрээ суурилан Зураг 6-т үзүүлсэн зарчмаар ажилладаг царай таних бодит хугацааны программын анхан шатны хувилбарыг хийсэн. Уг алгоритм нь хоёр ширхэг гүн сургалтын загварыг агуулсан ба өмнөх алгоритмаас хамаагүй бага тооцоолон бодох хүчин чадал шаарддагаараа илүү давуу талтай.



Зураг 6: Амны хаалттай, амны хаалтгүй хүнийг таних сайжруулсан программын кодын ажиллах зарчмын блок схем.

Одоогийн байдлаар, туршилт зүгшрүүлэлтийн явцад гүн сургалтын загварууд огт алдаа гаргаагүй. Python хэлний PyQt5 санг ашиглан хэрэглэгчийн интерфэйс хийсэн ба одоогийн байдлаар хэрэглэгчийн интерфэйс нь хүн нэмж бүртгэх, бүртгэлийг харах гэсэн хоёр функцтэй байна.

5. Дүгнэлт

Бид энэхүү ажлаар тооцоолон бодох хүчин чадал аль болох бага шаардах гүн сургалтын амны хаалттай, хаалтгүй хүний царай таних программ хангамжийг хөгжүүлэхийг зорьсон. Одоогоор 2 гигабайт багтаамжтай график тооцоолуур (NVIDIA GeForce GTX 1050) дээр 25 FPS-тэйгээр ажиллах боломжтой алгоритм, түүнд харгалзах загваруудыг сургаж, гаргаж авлаа.

Хүний хувийн мэдээлэл хамгаалах тухай хуулийг даган мөрдөх зорилгоор төрийн албаны цаг бүртгэлийн системд хүний царай, хурууны хээ зэрэг өгөгдлүүдийг ашиглах боломжгүй болсон. Гэхдээ энэхүү хуулийн 10.2-т “Ажил олгогч хөдөлмөрийн дотоод журамд заасны дагуу ажилтныг таньж, баталгаажуулах үйлдлийг хялбарчлах зорилгоор ажилтны зөвшөөрлөөр биеийн давхцахгүй өгөгдөл /гарын хурууны хээ/-ээс бусад биометрик мэдээллийг ашиглаж болно” гэсэн байна. Гүн сургалтын загвараас гарах шигтгээ вектор нь энэхүү шаардлагыг хангаж байгаа гэж бид үзэж байна. Өгөгдлийн санд хүний зургийн оронд тухайн зургаас гүн сургалтын загварын үүсгэсэн 128 хэмжээстэй шигтгээ векторуудыг хадгалж, тэдгээрийг ашиглан царай таньж болно. Уг 128 хэмжээстэй шигтгээ вектор алдагдсан ч гэсэн энэхүү вектороос буцааж хүний царайны зургийг үүсгэх боломжгүй. Тиймээс ийм төрлийн царай таних аргыг цаг бүртгэлийн системд ашиглаж болно гэж бид дүгнэж байна.

Ном зүй

- [1] C. Liu, and H. Wechsler, “Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition,” *Image processing, IEEE Transactions on*, vol. 11, no. 4, pp. 467–476, 2002, doi: <https://doi.org/10.1109/TIP.2002.999679>.
- [2] M. Turk, and A. Pentland, “Eigenfaces for recognition,” *Journal of cognitive neuroscience*, Vol 3, no. 1, pp. 71–86, 1991, doi: <https://doi.org/10.1515/ijsl.1991.87.71>.
- [3] G. B. Huang, M. Ramesh, and T. Berg, “Labeled faces in the wild: A database for studying face recognition in unconstrained environments,” *Technical Report, University of Massachusetts, Amherst*, pp. 07-49, 2007.
- [4] D. Chen, X. Cao, and F. Wen, “Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification,” *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3025–3032, 2013, doi: <https://doi.org/10.1109/CVPR.2013.389>.
- [5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Advances in neural information processing systems*, pp. 1097–1105, 2012.
- [6] Y. Taigman, M. Yang, and M. Ranzato, “Deepface: Closing the gap to human-level performance in face verification,” *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1701–1708, 2014, doi: <https://doi.org/10.1109/CVPR.2014.220>.
- [7] Y. Sun, Y. Chen, and X. Wang, “Deep learning face representation by joint identification-verification,” *Advances in neural information processing systems*, pp. 1988–1996, 2014.
- [8] F. Schroff, D. Kalenichenko, and J. Philbin, “Facenet: A unified embedding for face recognition and clustering,” *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 815–823, 2015, doi: <https://doi.org/10.1109/CVPR.2015.7298682>.
- [9] J. Liu, Y. Deng, T. Bai, and Z. Wei, “Targeting ultimate accuracy: Face recognition via deep embedding,” *arXiv preprint arXiv:1506.07310*, 2015.
- [10] O. M. Parkhi, A. Vedaldi, A. Zisserman, “Deep face recognition,” *BMVC*, Vol. 1, no. 3, pp. 6, 2015, doi: <https://doi.org/10.5244/C.29.41>.
- [11] W. Liu, Y. Wen, and Z. Yu, “Sphereface: Deep hypersphere embedding for face recognition,” *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 212–220, 2017, doi: <https://doi.org/10.1109/CVPR.2017.713>.
- [12] J. Deng, J. Guo, and N. Xue, “Arcface: Additive angular margin loss for deep face recognition,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4690–4699, 2019, doi: <https://doi.org/10.1109/CVPR.2019.00482>.
- [13] H. Wang, Y. Wang, and Z. Zhou, “Cosface: Large margin cosine loss for deep face recognition,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5265–5274, 2018, doi: <https://doi.org/10.1109/CVPR.2018.00552>.
- [14] Y. B. Chandra, and G. K. Reddy, “A Comparative Analysis Of Face Recognition Models On Masked Faces,” *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, Vol. 9, no. 10, 2020, doi: [https://doi.org/10.1016/S0958-2118\(20\)30175-0](https://doi.org/10.1016/S0958-2118(20)30175-0).
- [15] D. Yi, Z. Lei, Sh. Liao, and S. Z. Li, “Learning Face Representation from Scratch,” *CoRR*, 2014.
- [16] Н. Даваасүрэн, Ө. Баттулга, Б. Луубаатар, Д. Ууганбаатар, “Амны хаалттай хүнийг таних гүн сургалтын загвар”, *Хүрэлтогоот-2021 эрдэм шинжлэлийн хурлын эмтгэл*, pp. 48-55, 2021, doi: <https://doi.org/10.1055/a-1347-3272>.
- [17] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.Y. Fu, and A. C. Berg, “SSD: Single shot multibox detector”, *XECCV*, 2016, doi: https://doi.org/10.1007/978-3-319-46448-0_2.
- [18] A. Anwar, and A. Raychowdhury, “Masked Face Recognition for Secure Authentication”, *arXiv*, 2020.

Real Time Face Recognition Software Development for People With and Without Face Masks

Davaasuren Nyamdavaa^{1*}, Luubaatar Badarch², Battulga Ulziisaikhan¹,
Uuganbaatar Dulamragchaa¹, Munkhbayar Battsogt¹

¹*Institute of Mathematics and Digital Technology, Mongolian Academy of Sciences, Ulaanbaatar 13330, Mongolia*

²*School of Information and Communication Technology, Mongolian University of Science and Technology, Ulaanbaatar 13341, Mongolia*

**Corresponding author: davaasuren_n@mas.ac.mn; ORCID:0000-0002-2843-3700*

Article Info: Received: 2022.09.11; Accepted: 2022.10.01; Published: 2022.12.26

Abstract: In recent years, people began to wear masks to prevent the spread of Covid-19, which is widespread all over the world. Face recognition algorithms tend to be relatively inadequate for people wearing masks because masks cover most of a person's face. Our goal is to develop software with high accuracy and precision on the task which is recognizing people with and without face masks. In this paper, we are introducing some results of our work.

Key words: Casia-Webface, Facenet, Deep learning
