

**ШИНЖЛЭХ УХААН ТЕХНОЛОГИЙН ИХ СУРГУУЛЬ**  
**Мэдээлэл холбооны технологийн сургууль**



**БИЕ ДААЛТЫН АЖЛЫН**  
**ТАЙЛАН**

**Машин сургалт (F.CSC336) 2024-2025**  
**оны хичээлийн жилийн намар**

**Бие даалтын ажлын нэр: Корпус үүсгэн хэлний загвар байгуулах**

**Хичээл заасан багш:**  
**Бие даалтын ажил гүйцэтгэсэн:**

**Ж.Оргил**  
**С.Тэмүүжин B221960002**  
**М.Жамсран B221960028**

**Улаанбаатар хот. 2024 он**

# U-Net: Convolutional Networks for Biomedical Image Segmentation

Olaf Ronneberger, Philipp Fischer, болон Thomas Brox

Компьютерийн Шинжлэх Ухааны Тэнхим, BIOS Биологийн Сигнал Судлалын Төв, Фрайбургийн Их Сургууль, Герман

**[ronneber@informatik.uni-freiburg.de](mailto:ronneber@informatik.uni-freiburg.de)**

**<http://lmb.informatik.uni-freiburg.de/>**

Товч хураангуй: Гүнзгий сүлжээнүүдийн амжилттай сургалтанд олон мянган аннотлогдсон сургалтын жишээ шаардлагатай гэсэн ойлголт түгээмэл байдаг. Энэ судалгааны бид өгөгдлийн арвин нэмэгдүүлэлтэнд тулгуурлан байгаа аннотлогдсон өгөгдлийг илүү үр дүнтэй ашиглах боломжийг олгодог сүлжээ ба сургалтын аргыг танилцуулж байна. Архитектур нь орчны мэдээллийг олж авах агшилтын шугам, нарийвчлалтай локализацийг хангах тэгш хэмтэй өргөтгөлийн шугамаас бүрдэнэ. Бид ийм сүлжээг цөөн хэдэн зургаас end-to-end сургалтаар сургаж, электрон микроскопийн стекийн мэдрэлийн бүтэцийн сегментац хийх ISBI тэмцээнд өмнөх хамгийн сайн аргыг (Sliding-Window Convolutional Network) давсан болохыг харуулсан. Ижил сүлжээг гэрэлтүүлэлтийн микроскопийн зургууд дээр (фазын эсэргүүцэл болон DIC) сургаснаар 2015 оны ISBI эсийн мөрдөлтийн уралдаанд тухайн төрлүүдэд хамаагүй өндөр оноо авсан. Түүнчлэн, уг сүлжээ хурдтай бөгөөд 512x512 хэмжээтэй зургийг сүүлийн үеийн GPU дээр секундээс бага хугацаанд сегментчилдэг. Бүрэн хэрэгжүүлэлт (Caffe дээр суурилсан) болон сургагдсан сүлжээнүүдийг <http://lmb.informatik.uni-freiburg.de/people/ronneber/u-net> хаягаар авах боломжтой.

Thanks for

<https://www.digitalocean.com/community/tutorials/unet-architecture-image-segmentation>

[www.geeksforgeeks.org/u-net-architecture-explained/](http://www.geeksforgeeks.org/u-net-architecture-explained/)

<https://viso.ai/deep-learning/u-net-a-comprehensive-guide-to-its-architecture-and-applications/>

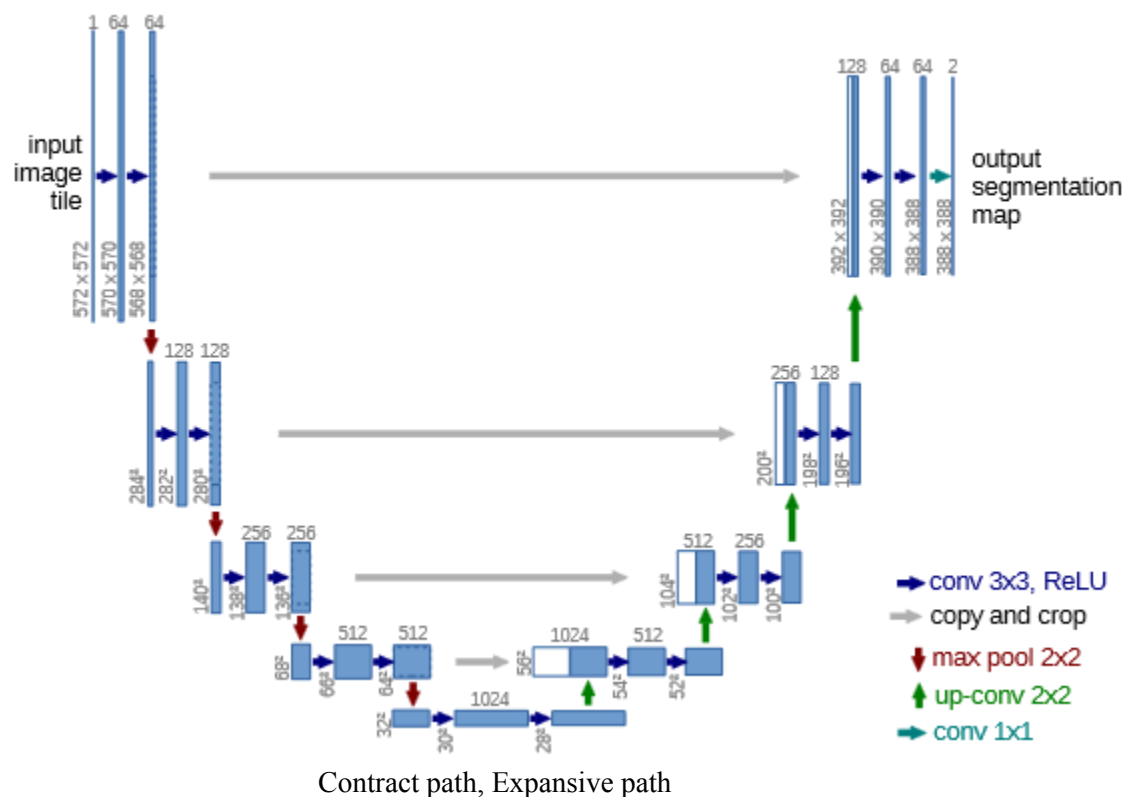
<https://latexeditor.lagrida.com/>

## 1. Introduction

Сүүлийн хоёр жилийн хугацаанд deep convolutional сүлжээнүүд олон төрлийн дүрс таних даалгавруудад бусад сүүлийн үеийн технологийн үр дүнг даван гаргаж ирсэн, жишээ нь [7,3]. Конволюцийн сүлжээнүүд дээр үеэс ашиглагдаж байсан ч [8], тэдгээрийн амжилт нь сургалтын багцын болон сүлжээний хэмжээнээс шалтгаалж хязгаарлагдаж байв. Krizhevsky et al [7] амжилт нь ImageNet өгөгдлийн сангийн 1 сая сургалтын зургийг ашиглан, 8 давхарга, хэдэн сая параметр

бүхий том сүлжээг хяналттай сургахад оршиж байв. Үүнээс хойш, бүр том, гүнзгий сүлжээнүүдийг сургаж эхэлсэн [12].

Конволюцийн сүлжээг ихэвчлэн ангиллын даалгаварт ашигладаг бөгөөдхариулт нь ганцхан ангиллын шошго байдаг. Гэвч олон дүрслэл боловсруулалтын даалгаварт, ялангуяа биомедиал зургийн боловсруулалтад хүссэн хариулт нь локализацийг агуулсан байх ёстой буюу ангиллын шошгыг зургийн бүр пикселд оноох шаардлагатай болдог. Мөн биомедиал даалгаварт олон мянган сургалтын зураг хангалтгүй байх нь түгээмэл. Иймд, Ciresan et al [1] тухайн пикселийн орчны хэсгийг (patch) оролт болгон өгч, тус бүрийн ангиллын шошгийг таамаглах зорилгоор оролт нь sliding window сүлжээг сургажээ. Энэ арга нь, нэгдүгээрт, сүлжээний локализаци хийнэ; хоёрдугаарт, patch илэрхийлэгдэх сургалтын өгөгдөл нь сургалтын зурагны тооноос их байж чадсан. Үүний үр дүнд бий болсон сүлжээ ISBI 2012-ын EM сегментацийн уралдаанд маш өндөр оноогоор түрүүлсэн байна.

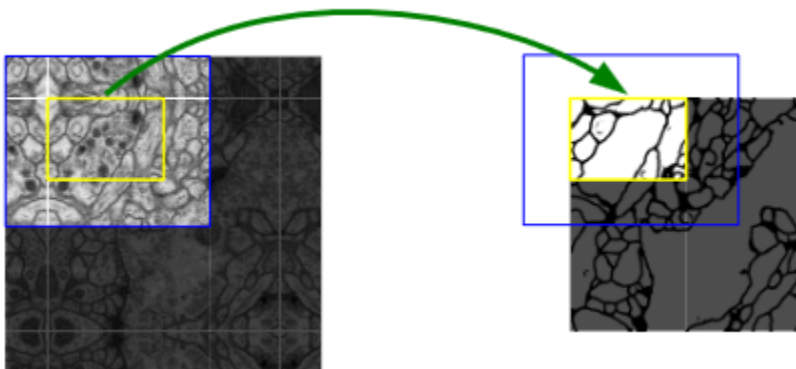


Зураг 1. U-net архитектур (хамгийн бага нарийвчлалд 32x32 пикселийн жишээ). Цэнхэр хайрцаг бүр олон сувгийн онцлог газрын зургийг илэрхийлнэ. Хайрцгийн дээд хэсэгт сувгийн тоо бичигдсэн байна. Хэмжээний х-у хэмжээ нь хайрцгийн зүүн доод буланд өгөгдсөн. Цагаан хайрцгууд нь хуулбарласан онцлог газрын зургуудыг илэрхийлнэ. Сумнууд нь янз бүрийн үйлдлүүдийг зааж байна.

Мэдээж Ciresan et al[1] стратеги хоёр сул талтай. Нэгдүгээрт, уг арга нэлээн удаан бөгөөд учир нь сүлжээг тус бүрийн patch дээр тусад нь ажиллуулах шаардлагатай ба давхардсан patch-аас ихээхэн давтамж үүсдэг. Хоёрдугаарт, локализацийн нарийвчлал ба орчны мэдээллийг ашиглах

хоёрын хооронд сонголт хийх шаардлага гардаг. Том patch нь локализацийн нарийвчлалыг бууруулдаг олон тооны max-pooling давхаргыг шаардах бол, жижиг patch нь сүлжээнд зөвхөн бага хэмжээний орчны мэдээлэл харагдахад хүргэдэг. Сүүлийн үеийн аргууд нь [11,4] олон давхаргын шинж чанарыг ашиглан гарцын ангилагчийг санал болгож байгаан сайн локализаци ба орчны мэдээллийг зэрэгцүүлэн ашиглах боломж бүрдсэн.

Энэ судалгаанд бид илүү гоёмсог архитектур болох "full convolutional сүлжээ" [9]-д тулгуурлан ажиллаж байна. Бид уг архитектурыг маш цөөн сургалтын зургууд дээр ажиллаж, илүү нарийвчлалтай сегментац гаргахуйц байдлаар өөрчилж, өргөжүүлсэн; Үүнийг 1-р зурагт харуулсан. [9]-ийн гол санаа нь contracting сүлжээг дараалсан давхаргуудаар өргөтгөх бөгөөд энд pooling операторуудыг upsampling операторуудаар сольдог. Иймээс эдгээр давхаргууд хариултын нарийвчлалыг нэмэгдүүлдэг. Локализаци хийхийн тулд contracting path өндөр нарийвчлалтай шинж чанаруудыг дээшлүүлсэн гарцтай хослуулдаг. Дараалсан конволюцийн давхарга нь энэ мэдээлэл дээр үндэслэн илүү нарийвчлалтай хариултыг угсрахыг сургах боломжтой.



Зураг 2. Бүхэл зургуудыг саадгүй сегментчлэхийн тулд overlap-tile стратеги ашигласан (энд EM стекийн мэдрэлийн бүтэцийн сегментац харуулсан). Шар хэсгийн сегментацийн таамаглал нь оролтод цэнхэр хэсгийн дүрслэлийн өгөгдлийг шаарддаг. Дутагдаж буй оролтын өгөгдлийг тусгал (mirror) аргаар нөхөж тооцдог.

Манай архитектурын нэг чухал өөрчлөлт нь дээшлүүлэлтийн хэсэгт олон тооны шинж чанарын суваг (feature channel)-уудыг нэмсэн явдал бөгөөд энэ нь сүлжээнд орчны мэдээллийг өндөр нарийвчлалтай давхаргуудад дамжуулах боломжийг олгодог. Үүний үр дүнд өргөтгөлийн шугам нь агшилтын шугамтай бараг тэгш хэмтэй болж, U хэлбэрийн архитектур үүсгэдэг. Сүлжээнд бүрэн холбогдсон давхарга байхгүй бөгөөд зөвхөн тус бүрийн конволюцийн хүчинтэй хэсгийг ашигладаг, өөрөөр хэлбэл, сегментацийн зурагт оролтын зурагт бүтэн орчны мэдээлэлтэй байгаа пикселүүд л багтдаг. Энэ стратеги нь overlap-tile аргыг ашиглан дурын хэмжээтэй зургийг саадгүй сегментчлэх боломжийг олгодог (Зураг 2-г үзнэ үү). Зургийн хилийн хэсгийн пикселүүдийг таамаглахын тулд дутуу байгаа орчныг оролтын зургийг толин тусгалд шилжүүлэх замаар нөхдөг. Энэ давхцах хуваалтын стратеги нь сүлжээг том хэмжээний зурагт хэрэглэхэд чухал ач холбогдолтой бөгөөд эс бөгөөс нарийвчлал нь GPU санахуйн хязгаарлалттай байх байв.

Манай даалгавруудад ашиглах боломжтой сургалтын өгөгдөл маш бага тул, өгөгдлийн арвин нэмэгдүүлэлт хийхийн тулд байгаа сургалтын зургууд дээр уян деформацийг (elastic

deformations) ашиглаж байна. Энэ нь сүлжээнд эдгээр деформацид үл хувиралтай байдлыг сурахад тусалдаг ба аннотлогдсон зургийн цуглуулгад эдгээр хувирлуудыг харуулах шаардлагагүй болгодог. Энэ нь биомедикал сегментацид онцгой ач холбогдолтой, учир нь эдийн деформаци нь хамгийн түгээмэл хувирлуудын нэг байсан бөгөөд бодитой деформацийг үр дүнтэй симуляц хийх боломжтой байдаг. Үл хувиралтын суралцахад өгөгдлийн нэмэгдүүлэлтийн ач холбогдлыг Dosovitskiy нар [2] хяналтгүй шинж чанарын суралцах судалгааны хүрээнд харуулсан байдаг.

Олон эсийн сегментацийн даалгаварт өөр нэгэн сорилт нь ижил ангиллын хүрэлцсэн объектийг тусгаарлах явдал байдаг; Үүнийг Зураг 3-т үзүүлсэн болно. Үүнийг шийдэхийн тулд бид хүрэлцсэн эсүүдийн хоорондох тусгаарлах суурь шошгонд алдагдлын функцэд өндөр жин оноодог жинлэгдсэн алдагдлыг ашиглахыг санал болгож байна.

Үүний үр дүнд бий болсон сүлжээ нь биомедикал сегментацийн олон төрлийн асуудалд ашиглагдах боломжтой. Энэ судалгаанд бид EM стекийн мэдрэлийн бүтцийн сегментацийн үр дүнг (ISBI 2012-оос эхэлсэн уралдаан) харуулсан бөгөөд Ciresan нарын [2] сүлжээг давсан үр дүнд хүрсэн. Мөн бид 2015 оны ISBI эсийн мөрдөлтийн уралдаанаас авсан гэрэлтүүлэлтийн микроскопийн зургуудаар эсийн сегментацийн үр дүнг үзүүлж байна. Энд бид хамгийн хүнд 2D гэрэлтүүлэлтийн хоёр өгөгдлийн багц дээр томоохон давуу байдлаар тэргүүлсэн.

## 2. Network Architecture

Сүлжээний архитектурыг Зураг 1-д үзүүлсэн. Энэ нь агшилтын шугам (зүүн талд) болон өргөтгөлийн шугамаас (баруун талд) бүрдэнэ. Агшилтын шугам нь конволюцийн сүлжээний ердийн архитектурыг дагадаг. Энэ нь хоёр 3x3 конволюцийг (padding-гүй) давтан хэрэглэснээс бүрддэг бөгөөд тус бүрийн дараа засагдсан шугаман нэгж (ReLU) болон 2x2 макс-пуулын үйлдлийг 2 алхамтайгаар хийгдэж, доош нь дээж авдаг. Дээж авалтын алхам бүрт шинж чанарын сувагны тоог хоёр дахин нэмэгдүүлдэг. Өргөтгөлийн шугамын алхам бүр нь шинж чанарын зургийг дээшлүүлж, 2x2 конволюци ("up-convolution") хийснээр шинж чанарын сувгийг хагас бууруулж, агшилтын шугамаас таслагдсан шинж чанарын зурагтай холбож, хоёр 3x3 конволюци ба ReLU-г дагаж хийгддэг. Конволюци бүрийн үед ирмэгийн пикселүүд алдагддаг тул таслалт зайлшгүй шаардлагатай. Эцсийн давхаргад 1x1 конволюци ашиглаж, тус бүрийн 64 бүрэлдэхүүнтэй шинж чанарын векторыг шаардлагатай ангийн тоонд хувиргадаг. Нийтдээ сүлжээнд 23 конволюцийн давхарга бий.

Гарцын сегментацийн зургийг саадгүй хуваах боломжтой болгохын тулд (Зураг 2-г үз), оролтын хэсгийн хэмжээг сонгохдоо бүх 2x2 макс-пуулын үйлдлүүдийг  $x$  ба  $y$  хэмжээ нь тэгш тоотой давхаргад хийх нь чухал юм.

## 3. Training

Оролтын зургууд болон тэдгээрийн харгалзах сегментацийн зургийг Caffé [6]-ийн санамсаргүй градиент бууралтын аргаар сүлжээг сургахад ашигладаг. Padding хийгдээгүй конволюци ашигласнаас болж гарцын зураг оролтын зургаас тогтмол ирмэгийн өргөнөөр жижиг болдог. GPU санахууныг хамгийн үр дүнтэй ашиглаж, илүү их өгөгдлийг боловсруулахын тулд бид том оролтын хэсгүүдийг илүүд үзэж, багцын хэмжээг нэг зураг болгон бууруулдаг. Үүний дагуу, өмнө нь үзсэн сургалтын жишээнүүдийн ихэнх нь одоогийн оновчлолын алхамд нөлөөлөхүйц байлгахын тулд өндөр инерци (0.99) хэрэглэдэг.

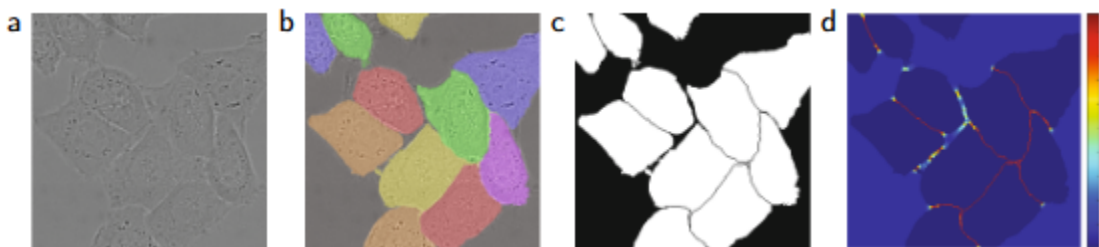
Энергийн функц нь эцсийн шинж чанарын зураг дээрх пиксел бүрийн soft-max ба кросс-энтропийн алдагдлын функцээр тооцогддог. Soft-max нь дараах байдлаар тодорхойлогдоно:

$$p_k(\mathbf{x}) = \frac{\exp(a_k(\mathbf{x}))}{\sum_{k'=1}^K \exp(a_{k'}(\mathbf{x}))}$$

энд  $a_k(\mathbf{x})$  нь  $\Omega \subset \mathbb{Z}^2$ -тэй  $\mathbf{x}$  байрлал дахь  $k$  шинж чанарын сувагт идэвхжсэн утгыг илэрхийлнэ.  $K$  нь ангийн тоо ба  $p_k(\mathbf{x})$  нь дээд утгыг ойролцоолсон функц юм. Өөрөөр хэлбэл,  $p_k(\mathbf{x}) \approx 1$  байх ба хамгийн их идэвхжилтэй  $k$ -д хамаарна  $a_k(\mathbf{x})$ , харин бусад бүх  $k$ -д  $p_k(\mathbf{x}) \approx 0$  байна. Кросс-энтропи нь тус бүрийн байрлалд  $p_{\ell(\mathbf{x})}(\mathbf{x})$ -ийн 1-ээс хазайлтыг дараах байдлаар шийтгэнэ:

$$E = \sum_{\mathbf{x} \in \Omega} w(\mathbf{x}) \log(p_{\ell(\mathbf{x})}(\mathbf{x}))$$

Энд  $\ell : \Omega \rightarrow \{1, \dots, K\}$  нь пиксел бүрийн үнэн шошгыг илэрхийлнэ,  $w : \Omega \rightarrow \mathbb{R}$  нь сургалтын үеэр тодорхой пикселд илүү ач холбогдол өгөхийн тулд ашиглагддаг жингийн зураг юм.



Зураг 3. DIC (дифференциал интерференцийн эсрэгцэл) микроскопоор бичигдсэн шилэн дээрх HeLa эсүүд. (a) түүхий зураг. (b) бодит сегментацийн давхцуулалт. Өөр өөр өнгөнүүд нь HeLa эсийн тусдаа жишээнүүдийг илэрхийлнэ. (c) үүсгэсэн сегментацийн маск (цагаан: урд талын хэсэг, хар: арын хэсэг). (d) хил залгаа пикселүүдийг сүлжээнд суралцуулахын тулд пиксел бүрийн алдагдлын жингийн зураг.

Бодит сегментац бүрийн жингийн зургийг урьдчилан тооцоолж, сургалтын өгөгдлийн багц дахь тодорхой ангийн пикселүүдийн ялгаатай давтамжийг нөхөхийн зэрэгцээ хүрэлцсэн эсүүдийн хоорондын жижиг тусгаарлах ирмэгүүдийг сүлжээгээр суралцуулах зорилготой (Зураг 3с ба d-г үзнэ үү). Тусгаарлах ирмэгийг морфологийн үйлдлүүд ашиглан тооцдог. Жингийн зургийг дараах байдлаар тооцоолно:

$$w(\mathbf{x}) = w_c(\mathbf{x}) + w_0 \cdot \exp\left(-\frac{(d_1(\mathbf{x}) + d_2(\mathbf{x}))^2}{2\sigma^2}\right)$$

энд  $w_c : \Omega \rightarrow \mathbb{R} : \Omega \rightarrow \mathbb{R}$   $w_c : \Omega \rightarrow \mathbb{R}$  нь ангийн давтамжийг тэнцвэржүүлэх жингийн зураг,  $d_1 : \Omega \rightarrow \mathbb{R}$  нь хамгийн ойрын эсийн хил хүртэлх зайг,  $d_2 : \Omega \rightarrow \mathbb{R}$   $d_2 : \Omega \rightarrow \mathbb{R}$  нь хоёр дахь хамгийн ойрын эсийн хил хүртэлх зайг илэрхийлнэ. Туршилтууддаа бид  $w_0 = 10$  ба  $\sigma \approx 5$  пикселээр тохируулсан.

Гүн сүлжээ нь олон конволюцийн давхарга болон сүлжээг дамжин өнгөрөх өөр замуудтай байдаг тул жингийн зөв эхлэлт нь маш чухал. Үгүй бол сүлжээний зарим хэсэг хэт идэвхжил үүсгэх бол, бусад хэсэг нь огт хувь нэмэр оруулахгүй байх магадлалтай. Идеал тохиолдолд эхлэлийн жингүүд нь сүлжээний шинж чанарын зураг тус бүрийн хувьд ойролцоогоор нэгжийн дисперситай байх ёстой. Манай архитектуртай сүлжээнд (конволюци болон ReLU давхаргууд ээлжилсэн) энэ нь

эхлэлийн жингүүдийг стандарт хазайлтыг  $\sqrt{2/N}$  -тэй Гауссын тархалтаас сонгох замаар биелэгдэнэ, энд  $N$  нь нэг нейрон руу орж буй зангилаануудын тоог илэрхийлнэ [5]. Жишээ нь, өмнөх давхаргад  $3 \times 3$  конволюци ба 64 шинж чанарын сувагтай бол  $N = 9 \cdot 64 = 576$  байна.

### 3.1 Data Augmentation

Өгөгдлийн нэмэгдүүлэлт нь цөөн сургалтын жишээтэй нөхцөлд сүлжээг хүссэн үл хувирал болон бат бөх чанарын шинж чанарт сургахын тулд зайлшгүй шаардлагатай. Микроскопийн зургуудын хувьд бидэнд голчлон шилжилт болон эргэлтийн үл хувиралт, мөн деформаци болон саарал утгын өөрчлөлтөд бат бөх чанар шаардлагатай. Ялангуяа сургалтын жишээнүүдэд санамсаргүй уян деформацийг хийх нь цөөн аннотлогдсон зургаар сегментацийн сүлжээ сургалт хийх гол ойлголт болдог. Бид  $3 \times 3$  бүдүүн торон дээр санамсаргүй шилжилтийн векторууд ашиглан зөөлөн деформацуудыг үүсгэдэг.

Rank	Group name	Warping Error	Rand Error	Pixel Error
	<b>** human values **</b>	<b>0.000005</b>	<b>0.0021</b>	<b>0.0010</b>
1.	u-net	<b>0.000353</b>	0.0382	0.0611
2.	DIVE-SCI	0.000355	0.0305	0.0584
3.	IDSIA [2]	0.000420	0.0504	0.0613
4.	DIVE	0.000430	0.0545	<b>0.0582</b>
⋮				
10.	IDSIA-SCI	0.000653	<b>0.0189</b>	0.1027

Шилжилтүүдийг 10 пикселийн стандарт хазайлтаар Гауссын тархалтаас сонгож авдаг. Дараа нь нэг пиксел тутмын шилжилтийг бикубик интерполяци ашиглан тооцоолдог. Агшилтын шугамын төгсгөлд байрлах drop-out давхаргууд нь нэмэлт далд өгөгдлийн нэмэгдүүлэлт хийдэг.

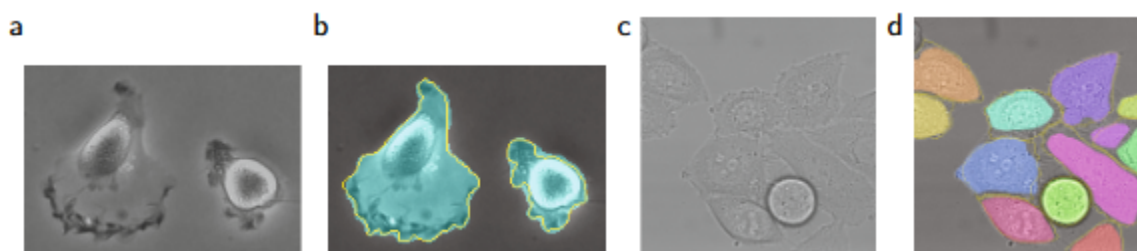
## 4 Experiments

Бид u-net-ийг гурван өөр сегментацийн даалгаварт ашиглаж буйг харууллаа. Эхний даалгавар нь электрон микроскопийн бичлэгт мэдрэлийн бүтцийн сегментац хийх явдал юм. Өгөгдлийн багц болон бидний гаргасан сегментацийн жишээг Зураг 2-т харуулсан. Бүтэн үр дүнг Нэмэлт материалд оруулсан болно. Уг өгөгдлийн багцыг ISBI 2012 онд эхэлсэн бөгөөд шинэ оролтуудыг хүлээн авдаг EM сегментацийн уралдаан [14,1] өгсөн. Сургалтын өгөгдөл нь *Drosophila*-ийн анхны үе шатны авгалдайн хэвлийн мэдрэлийн шугам (VNC)-ын цуврал дамжуулах электрон микроскопийн 30 зураг (512x512 пиксел)-аас бүрдэнэ. Тус бүрийн зураг нь эсүүд (цагаан) болон мембрануудын (хар) бүрэн аннотлогдсон бодит сегментацийн зургаар хамт ирдэг. Туршилтын багц олон нийтэд нээлттэй боловч сегментацийн зургууд нь нууцлагдсан байдаг. Үнэлгээг таамагласан мембраны магадлалын зургийг зохион байгуулагчдад илгээснээр авах боломжтой. Үнэлгээг зурагт 10 өөр түвшинд босго тогтоох, “warping алдаа,” “Rand алдаа” болон “пикселийн алдаа” [14]-г тооцоолох замаар гүйцэтгэдэг.

U-net нь (оролтын өгөгдлийн 7 эргүүлсэн хувилбараар дундажлан) ямар нэгэн нэмэлт урьдчилсан эсвэл дараах боловсруулалтгүйгээр 0.0003529 warping алдаа (шинэ хамгийн сайн оноо, 1-р хүснэгтийг үзнэ үү) болон 0.0382 Rand алдааг гаргасан.

Энэ нь Ciresan нарын [2] гулсах-цонх конволюцийн сүлжээний үр дүнгээс эрс дээр бөгөөд тэдний хамгийн сайн оролтын warping алдаа 0.000420, Rand алдаа 0.0504 байсан. Rand алдааны хувьд энэ өгөгдлийн багц дээр илүү сайн гүйцэтгэлтэй цорын ганц алгоритмууд нь Ciresan нарын [2] магадлалын зураг дээр өндөр өгөгдөлд онцлог тохирсон дараах боловсруулалтын аргуудыг ашигласан байдаг.

Энэхүү алгоритмын зохиогчид энэ үр дүнд хүрэхийн тулд 78 өөр шийдэл ирүүлсэн байна.



Зураг 4. ISBI эсийн мөрдөлтийн уралдааны үр дүн. (a) "PhC-U373" өгөгдлийн багцын оролтын зургийн хэсэг. (b) Сегментацийн үр дүн (хөх маск) ба гараар тэмдэглэсэн бодит шошго (шар хүрээ). (c) "DIC-HeLa" өгөгдлийн багцын оролтын зураг. (d) Сегментацийн үр дүн (санамсаргүй өнгөт маскууд) ба гараар тэмдэглэсэн бодит шошго (шар хүрээ).

2-р хүснэгт. 2015 оны ISBI эсийн мөрдөлтийн уралдааны сегментацийн үр дүн (IOU буюу Intersection over Union).



Name	PhC-U373	DIC-HeLa
IMCB-SG (2014)	0.2669	0.2935
KTH-SE (2014)	0.7953	0.4607
HOUS-US (2014)	0.5323	-
second-best 2015	0.83	0.46
u-net (2015)	<b>0.9203</b>	<b>0.7756</b>

Бид u-net-ийг гэрэлтүүлэлтийн микроскопийн зургууд дээр эсийн сегментацийн даалгаварт мөн ашигласан. Энэ сегментацийн даалгавар нь 2014 болон 2015 оны ISBI эсийн мөрдөлтийн уралдааны нэг хэсэг юм [10,13]. Эхний өгөгдлийн багц болох "*PhC – U373*"<sup>2</sup> нь фазын эсрэгцлийн микроскопоор полиакриламид субстратад байрлах глиобластом-астроцитом U373 эсүүдийг агуулдаг (Зураг 4a,b болон нэмэлт материалд үзнэ үү). Энэ багц нь хэсэгчлэн аннотлогдсон 35 сургалтын зургийг агуулна. Энд бид 92%-ийн дундаж IOU ("intersection over union") үзүүлэлт гаргасан нь хоёр дахь шилдэг алгоритмын 83%-аас мэдэгдэхүйц өндөр байна (2-р хүснэгтийг үз). Хоёр дахь өгөгдлийн багц болох "*DIC – HeLa*"<sup>3</sup> нь дифференциал интерференцийн эсрэгцлийн (DIC) микроскопоор хавтгай шилэн дээр бичигдсэн HeLa эсүүдийг агуулдаг (Зураг 3, Зураг 4c,d болон нэмэлт материалд үзнэ үү). Энэ багц нь 20 хэсэгчлэн аннотлогдсон сургалтын зургийг агуулдаг. Энд бид 77.5%-ийн дундаж IOU үзүүлсэн нь хоёр дахь шилдэг алгоритмын 46%-аас хамаагүй илүү байна.

## Conclusion

U-net архитектур нь маш өөр төрлийн биомедикал сегментацийн хэрэглээнд маш сайн гүйцэтгэл үзүүлдэг. Уян деформацийг ашиглан өгөгдөл нэмэгдүүлэлт хийдэг тул цөөн тооны аннотлогдсон зураг шаардлагатай бөгөөд NVidia Titan GPU (6 GB)-д ердөө 10 цагийн сургалтын хугацаа хангалттай байдаг. Бид Caffe[6]-д суурилсан бүрэн хэрэгжүүлэлт болон сургагдсан сүлжээгээ<sup>4</sup> нийтэлсэн болно. U-net архитектурыг олон бусад даалгаварт хялбархан ашиглах боломжтой гэдэгт итгэлтэй байна.

<sup>2</sup> Өгөгдлийн багцыг Калифорнийн их сургуулийн Беркли дэх (АНУ) Биоинженерийн энхимийн доктор Санжай Кумар өгсөн.

<sup>3</sup> Өгөгдлийн багцыг Нидерландын Роттердам хот дахь Эразмус эмнэлгийн төвийн доктор Герт ван Каппеллен өгсөн.

<sup>4</sup> U-net-ийн хэрэгжүүлэлт, сургагдсан сүлжээнүүд болон нэмэлт материалыг дараах хаягаас авах боломжтой: <http://lmb.informatik.uni-freiburg.de/people/ronneber/u-net>.

Acknowledgements. Энэхүү судалгааг Германы Холбооны болон Муж улсын засгийн газрын "Шилдэг санаачилга" (EXC 294) болон BMBF (Fkz 0316185B) дэмжлэгтэйгээр хэрэгжүүлсэн болно.

## References

1. Cardona, A., et al.: An integrated micro- and macroarchitectural analysis of the drosophila brain by computer-assisted serial section electron microscopy. *PLoS Biol.* 8(10), e1000502 (2010)
2. Ciresan, D.C., Gambardella, L.M., Giusti, A., Schmidhuber, J.: Deep neural networks segment neuronal membranes in electron microscopy images. In: *NIPS*, pp. 2852–2860 (2012)
3. Dosovitskiy, A., Springenberg, J.T., Riedmiller, M., Brox, T.: Discriminative unsupervised feature learning with convolutional neural networks. In: *NIPS* (2014)
4. Hariharan, B., Arbel'aez, P., Girshick, R., Malik, J.: Hypercolumns for object segmentation and fine-grained localization (2014), arXiv:1411.5752 [cs.CV]
5. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human level performance on imagenet classification (2015), arXiv:1502.01852 [cs.CV]
6. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding (2014), arXiv:1408.5093 [cs.CV]
7. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: *NIPS*, pp. 1106–1114 (2012)
8. LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D.: Backpropagation applied to handwritten zip code recognition. *Neural Computation* 1(4), 541–551 (1989)
9. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation (2014), arXiv:1411.4038 [cs.CV]
10. Maska, M., et al.: A benchmark for comparison of cell tracking algorithms. *Bioinformatics* 30, 1609–1617 (2014)
11. Seyedhosseini, M., Sajjadi, M., Tasdizen, T.: Image segmentation with cascade hierarchical models and logistic disjunctive normal networks. In: *2013 IEEE International Conference on Computer Vision (ICCV)*, pp. 2168–2175 (2013)
12. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition (2014), arXiv:1409.1556 [cs.CV]
13. WWW: Web page of the cell tracking challenge, [http://www.codesolorzano.com/celltrackingchallenge/Cell Tracking Challenge/Welcome.html](http://www.codesolorzano.com/celltrackingchallenge/Cell%20Tracking%20Challenge/Welcome.html)
14. WWW: Web page of the em segmentation challenge, [http://brainiac2.mit.edu/isbi\\_challenge/](http://brainiac2.mit.edu/isbi_challenge/)