

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ МИНИСТРЛІГІ

«Л.Н. ГУМИЛЕВ АТЫНДАҒЫ ЕУРАЗИЯ ҰЛТТЫҚ УНИВЕРСИТЕТІ» КЕАҚ

**Студенттер мен жас ғалымдардың
«GYLYM JÁNE BILIM - 2024»
XIX Халықаралық ғылыми конференциясының
БАЯНДАМАЛАР ЖИНАҒЫ**

**СБОРНИК МАТЕРИАЛОВ
XIX Международной научной конференции
студентов и молодых ученых
«GYLYM JÁNE BILIM - 2024»**

**PROCEEDINGS
of the XIX International Scientific Conference
for students and young scholars
«GYLYM JÁNE BILIM - 2024»**

**2024
Астана**

УДК 001

ББК 72

G99

«ǴYLYM JÁNE BILIM – 2024» студенттер мен жас ғалымдардың XIX Халықаралық ғылыми конференциясы = XIX Международная научная конференция студентов и молодых ученых «ǴYLYM JÁNE BILIM – 2024» = The XIX International Scientific Conference for students and young scholars «ǴYLYM JÁNE BILIM – 2024». – Астана: – 7478 б. - қазақша, орысша, ағылшынша.

ISBN 978-601-7697-07-5

Жинаққа студенттердің, магистранттардың, докторанттардың және жас ғалымдардың жаратылыстану-техникалық және гуманитарлық ғылымдардың өзекті мәселелері бойынша баяндамалары енгізілген.

The proceedings are the papers of students, undergraduates, doctoral students and young researchers on topical issues of natural and technical sciences and humanities.

В сборник вошли доклады студентов, магистрантов, докторантов и молодых ученых по актуальным вопросам естественно-технических и гуманитарных наук.

УДК 001

ББК 72

G99

ISBN 978-601-7697-07-5

**©Л.Н. Гумилев атындағы Еуразия
ұлттық университеті, 2024**

ӨКПЕ АУРУЛАРЫНЫҢ ЖІКТЕЛУІ ТУРАЛЫ МӘЛІМЕТТЕРДІ МАШИНАЛЫҚ ОҚЫТУ АРҚЫЛЫ ӨНДЕУ

Төлеген Бағыжан

b.tolegen02@gmail.com

Л.Н. Гумилев атындағы Еуразия ұлттық университетінің “Математикалық және компьютерлік моделдеу” мамандығының 4-курс студенті

Ғылыми жетекшісі – А.С. Жумаханова, Л.Н. Гумилев атындағы ЕҰУ, «Математикалық және компьютерлік модельдеу» кафедрасының аға оқытушысы

Бұл зерттеу кеуде қуысының медициналық рентгендік суреттері негізінде COVID-19 пневмония жағдайларын талдау және жіктеу үшін машиналық оқыту әдістерін қолдануды зерттейді. Пандемия ауқымын және ауруды жылдам және дәл диагностикалау қажеттілігін ескере отырып, автоматтандырылған талдау әдістері бұл процесте шешуші рөл атқара алады. Зерттеуде COVID-19 кеуде қуысының рентгендік деректер жинағы бастамасы және кеуде қуысының рентгендік суреттері (пневмония) сияқты қолжетімді рентгендік дерекқорлар пайдаланылады. Коронавирус тудырған пневмонияның белгілерін анықтау үшін бұл суреттерді талдау үшін машиналық оқыту қолданылады.

Зерттеуге негіз болған деректер қоры <https://github.com/ieee8023/covid-chestxray-dataset> және <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia> сайттарынан алынды.

Бұл зерттеу рентгендерінің медициналық кескіндерін төрт негізгі сыныпқа жіктеу үшін машиналық оқыту әдістерін қолдануға бағытталған:

0- COVID-19,

1- қалыпты кеуде қуысы,

2- вирустық пневмония және

3- бактериялық пневмония.

Осы мақсатта ашық дерекқорлар, соның ішінде COVID-19 кеуде қуысының рентгендік деректер жинағы бастамасы және кеуде қуысының рентгендік суреттері (пневмония) пайдаланылды.

Зерттеу деректерді алдын ала өңдеу процесін, сондай-ақ әр сыныпқа тән бірегей мүмкіндіктерді анықтау үшін конвульстік нейрондық желілерді қоса алғанда, әртүрлі машиналық оқыту алгоритмдерін пайдалануды қамтиды.

Нейрондық желіні пайдалана отырып рентгендік кескіндерін жіктеу тапсырмасын орындау үшін қажетті кітапханалар мен модульдерді импорттаймыз.

Google Colab ортасында Google Drive орнату үшін келесі бағдарламалық кодты пайдаланамыз.

```
from google.colab import drive
drive.mount('/content/drive')
```

Деректердің каталогын көрсетеміз

```
X_Ray_Directory = '/content/drive/MyDrive/Lung Disease Classification Dataset (100+ images)/Operations Department/Dataset'
os.listdir(X_Ray_Directory)
```

```
['0', '1', '2', '3']
```

Біз тензорлық кескін деректерін жасау және оларды қалыпқа келтіру үшін кескін генераторын қолданамыз Кросс-тексеру үшін деректердің 20% пайдаланамыз:

```
image_generator = ImageDataGenerator(rescale=1./255, validation_split=0.2)
```

40 суреттен тұратын пакет жасаймыз. Жалпы суреттер саны - $133 \times 4 = 532$ сурет

Оқылуы - 428 (80%), тексеру - 104 (20%) суреттердің өлшемін ауыстыруды орындаймыз. Біз 40 сурет пен жазулар пакетін жасаймыз: `train_images, train_labels = next(train_generator)`

Деректердің өлшемі туралы ақпарат (суреттер саны , ені, биіктігі, каналдар саны)

```
train_images.shape  
(40, 256, 256, 3)
```

Массивтің өлшемі

```
train_labels.shape  
(40, 4)
```

Әрбір `train_images` үлгісі үшін класс белгілерін қамтитын бір өлшемді массив

```
train_labels
```

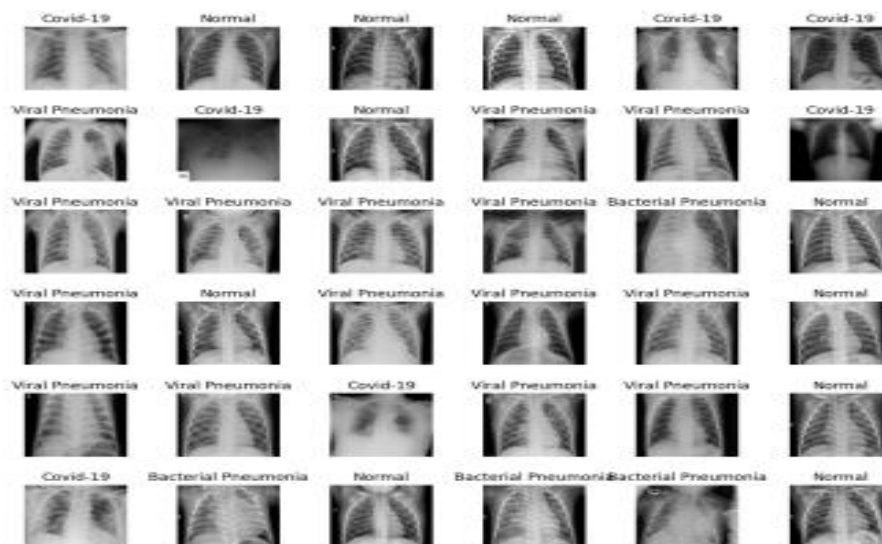
```
array([[1., 0., 0., 0.],  
       [0., 1., 0., 0.],  
       [0., 1., 0., 0.],  
       [0., 1., 0., 0.],  
       ....
```

```
       [0., 1., 0., 0.],  
       [0., 0., 0., 1.],  
       [0., 0., 0., 1.],  
       [0., 0., 0., 1.],  
       [0., 1., 0., 0.]], dtype=float32)
```

```
label_names = { 0: 'Covid-19', 1: 'Normal', 2: 'Viral Pneumonia', 3: 'Bacterial Pneumonia'}
```

Деректерді визуализациялау

Python тілінде Matplotlib және NumPy кітапханаларын қолданып жаттығу деректер жинағынан (`train_images`) кескіндерді сәйкес белгілермен (`train_labels`) визуализациялау үшін 6×6 ($L \times W$) кескіндер торы болып табылады, мұнда әрбір тор ұяшығы оқу деректер жинағынан бір суретті және сол кескіннің белгісі бар тақырыпты қамтиды. Егер сурет белгісі белгісіз болса (яғни жапсырма индексі `label_names` тізімінен тыс болса), онда белгі атауының орнына "Unknown Label" блды.



1 сурет - Рентген суреттері

Бұл мысал ImageNet деректер жинағында алдын ала дайындалған ResNet50 архитектурасын пайдаланып негізгі үлгіні құрады. Модель жоғарғы қабаттарсыз конфигурацияланған (`include_top = False` параметрі), бұл оны еркін өлшемдегі кескіндерден мүмкіндіктерді шығару үшін пайдалануға мүмкіндік береді. Бұл жағдайда кіріс кескіндерінің өлшемі 3 түсті арнамен (RGB) 256×256 пиксельге орнатылады.

Машиналық оқыту шеңберінен `fit_generator` әдісі.

Бұл әдіс, әдетте, әр оқу дәуірінде оқыту деректер пакеттерін ұсынатын итераторлар болып табылатын деректер генераторларын қолдана отырып, модельді оқыту үшін қолданылады. Берілген код Машиналық оқыту моделін оқыту деректерінің генераторын қолдана отырып үйретеді, модельді жеке тексеру деректер жиынтығында тексереді және өнімділік жақсармаса, модельдің ең жақсы салмағын сақтау және оқуды мерзімінен бұрын тоқтату үшін кері қоңырауларды қолданады.

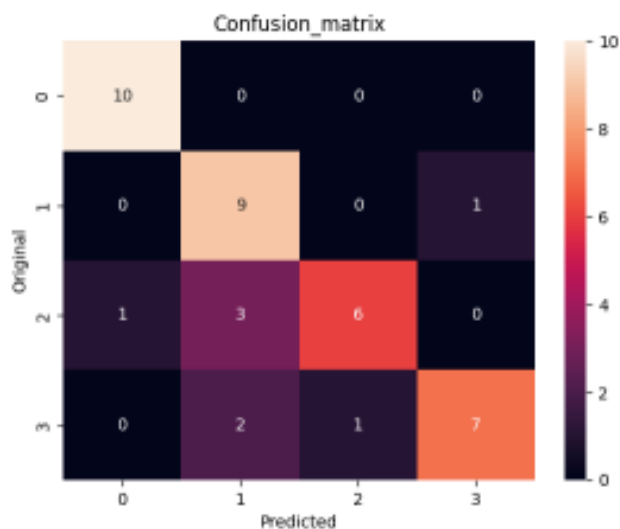
Бұл жолды тест суреттерін жүктеу және модельіңіздің жаңа, көрінбейтін деректерге қаншалықты жалпыланғанын бағалау үшін пайдалана аламыз.

Assurasy_score функциясы арқылы тест дәлдігі үшін шынайы және болжамды белгілер арасындағы дәлдікті есептейді және нәтижені басып шығарады. Бұл сіздің модельіңіздің көрінбейтін сынақ жиынтығында қаншалықты жақсы жұмыс істейтіндігінің сандық көрсеткішін ұсынады.

Бізде мұнда тесттің дәлдігі 0,8-ге тең, бұл сіздің модель сіздің тестілік деректер жиынтығыңыздағы даналардың 80% - ы үшін белгілерді дұрыс болжағанын білдіреді. Бұл жіктеу модельдерін бағалау үшін қолданылатын жалпы көрсеткіш және жоғары дәлдік әдетте жақсы өнімділікті көрсетеді.

Қысқаша айтқанда, бұл код сынақ деректеріне негізделген модельді болжау үшін шатасу матрицасын жасайды және оны жылу картасы арқылы көрсетеді. Жылу картасы модельдің жұмысындағы заңдылықтарды тез анықтауға мүмкіндік береді, мысалы, қай класстар бір-бірімен жиі шатастырылады. Бұл визуализация сіздің жіктеу моделіңіздің күшті және әлсіз жақтарын түсіну үшін пайдалы болуы мүмкін.

Мұнда алғашқы жолда "original" және "prediction" массивтер арқылы берілген мәліметтердің орындалған және болжалған деректері кездесетін және оның көрсеткіштерін анықтау үшін Confusion Matrix түрінде қолданылады.



Қателік матрицасын визуализациялау машиналық оқыту моделінің өнімділігін талдауға көмектеседі, Бұл модельәсіресе жіктеу есептерінде қай сыныптарда жиі қателесетінін анықтауды жеңілдетеді.

УДК 519.813.7

ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ ДИНАМИЧЕСКОЙ ЗАДАЧИ «ГОНКА ВООРУЖЕНИЙ»

Тулькибаев Чингис Куанышбаевич

2020