

УДК 519.62

МРНТИ 27.35.21; 27.35.47; 27.41.19

<https://doi.org/10.55452/1998-6688-2026-23-1-147-162>

^{1*}**Исахов А.А.,**

профессор, ORCID ID: 0000-0002-1937-8615,

e-mail: alibek.issakhov@gmail.com

¹**Абылкасымова А.Б.,**

ассоциированный профессор, ORCID ID: 0000-0002-5967-6959,

e-mail: abylkassymova.aizhan@gmail.com

¹**Бревнов Т.,**

бакалавр, ORCID ID: 0009-0004-5395-4569,

e-mail: t_brevnov@kbtu.kz

¹Казахстанско-Британский технический университет,
г. Алматы, Казахстан

АНАЛИЗ РЕНТГЕНОВСКИХ СНИМКОВ ЛЕГОЧНОЙ СИСТЕМЫ ЧЕЛОВЕКА МЕТОДОМ МАШИННОГО ОБУЧЕНИЯ С ЦЕЛЬЮ ПОСТАНОВКИ ПРЕДВАРИТЕЛЬНОГО ДИАГНОЗА

Аннотация

В статье представлено комплексное исследование применения методов машинного обучения для автоматизированного анализа рентгенографических изображений органов дыхательной системы с целью раннего выявления патологических изменений. Предложена и реализована методика классификации легочных заболеваний на основе ансамбля глубоких сверточных нейронных сетей, включающего архитектуры DenseNet121, MobileNetV2, EfficientNetB0, SENet и ShuffleNetV2. В рамках исследования проведен сравнительный анализ эффективности различных методов предобработки изображений, включая использование исходных черно-белых рентгеновских снимков без дополнительной обработки, применение метода CLAHE (Contrast Limited Adaptive Histogram Equalization) в сочетании с цветовой фильтрацией, а также использование нейросетевого денойзера DynamicCNN для подавления шумов. Экспериментальные результаты показали, что ансамблевый подход с применением стратегии мягкого голосования (soft voting) обеспечивает статистически значимое повышение точности классификации по сравнению с отдельными моделями. Полученные результаты подтверждают высокую эффективность предложенного подхода и демонстрируют перспективность использования ансамблевых моделей глубокого обучения в задачах медицинской диагностики и поддержки принятия клинических решений.

Ключевые слова: методы машинного обучения, ансамбль глубоких нейронных сетей, рентгеновские снимки, денойзер DynamicCNN.

Введение

В современной медицине наблюдается стремительное развитие технологий искусственного интеллекта, которые революционно трансформируют процессы диагностики и лечения заболеваний. Особенно перспективным направлением является применение методов машинного обучения и компьютерного зрения в анализе медицинских изображений, в частности рентгеновских снимков легких.

Актуальность исследования обусловлена несколькими критическими факторами:

1. Глобальное бремя легочных заболеваний: согласно данным Всемирной организации здравоохранения, респираторные заболевания входят в число ведущих причин смертности в мире. Ежегодно от пневмонии, туберкулеза, рака легких и других патологий легочной системы погибают миллионы человек. Своевременная и точная диагностика является ключевым фактором эффективного лечения.

2. Ограничения традиционной диагностики: классические методы рентгенологического исследования имеют ряд существенных недостатков:

- ◆ высокая зависимость от квалификации и субъективного опыта врача-рентгенолога;
- ◆ риск человеческой ошибки при интерпретации снимков;
- ◆ задержка в постановке диагноза;
- ◆ сложность выявления ранних и латентных форм заболеваний.

3. Технологический прорыв в области искусственного интеллекта: развитие глубоких нейронных сетей, методов компьютерного зрения и машинного обучения открывает принципиально новые возможности в медицинской диагностике:

- ◆ высокая скорость анализа медицинских изображений;
- ◆ объективность оценки;
- ◆ возможность выявления тонких патологических изменений;
- ◆ потенциал масштабирования и применения в различных клинических условиях.

Исследователи продвинулись в области классификации заболеваний, разработав модели глубокого обучения, адаптированные к задачам медицинской визуализации. Основное внимание уделялось созданию архитектур, которые повышают точность, вычислительную эффективность и адаптивность для клинических приложений. Сяои Лю и др. в своей работе [1] выбрали 5 предварительно обученных моделей, которые впоследствии использовались для тестирования: SqueezeNet, DenseNet, MobileNetV2, ResNet18 и VGG11. Среди выбранных моделей MobileNetV2 показала наилучшие результаты с лучшей точностью. Затем авторы решили провести тонкую настройку модели, разморозив все слои, и она оказалась более точной по сравнению с точностью MobileNetV2. Затем к модели с размороженными слоями после первого сверточного слоя был добавлен блок сжатия и возбуждения, что снова немного улучшило точность. В целом интеграция механизмов внимания и тонкая настройка значительно улучшили производительность моделей, продемонстрировав их эффективность для задач классификации медицинских изображений. Шамрат и др. [2] провели комплексное исследование производительности модели с использованием настройки гиперпараметров и различных методов предварительной обработки, включая шумоподавление, удаление аннотаций, улучшение изображений и увеличение изображений. Denoising Autoencoder (DAE) и EnsNet использовались для шумоподавления и удаления аннотаций изображений соответственно с последующим применением метода улучшения изображений CLAHE. Авторы решили проблему дисбаланса данных путем повышения частоты дискретизации классов с меньшим количеством точек данных. Еще раз был оценен ряд предварительно обученных моделей, включая AlexNet, InceptionV3, MobileNetV2, VGG16, ResNet50, DenseNet121 и EfficientNetB7, с целью определения наиболее эффективной модели. VGG16 показала самые высокие показатели среди вышеупомянутых моделей, что побудило авторов изменить ее архитектуру и протестировать новую модель LungNet22. Это привело к превосходным результатам по сравнению с исходной VGG16. Исследование де Соузы и др. [3] представило еще одну архитектуру глубокой нейронной сети, CNN-COVID, которая продемонстрировала превосходную производительность при классификации случаев коронавирусного заболевания (Covid-19). В своем исследовании Ашраф и др. [4] использовали не только сверточные нейронные сети (CNN), но также визуальные преобразователи и гибридные модели. Конкретными используемыми моделями были CoAtNet, MaxViT, Swin V2, VOLO, DenseNet и ConvNeXt V2. Также применялись такие методы, как начальный поиск скорости обучения и эксперименты со стекированием и ансамблированием взвешенных средних. Среди всех моделей наибольшую эффективность продемонстрировала CoAtNet, за ней следует ConvNeXTV2. Кроме того, среди архитектур ансамбля ансамбль взвешенных средних продемонстрировал превосходную производительность. В работах [5, 6] исследователи улучшили MobileNet V2, включив глобальный средний слой объединения и полностью связанный слой с сигмоидальной функцией активации для классификации по нескольким меткам. Кроме того, они реализовали методы выборки данных и дополнения изображений для оптимизации производительности модели.

Целью настоящей работы является разработка комплексной методики автоматизированного анализа рентгеновских снимков легких с использованием ансамблевых методов машинного обучения для повышения точности и достоверности первичной диагностики.

Для достижения поставленной цели необходимо решить следующие задачи:

- ◆ сформировать репрезентативный датасет рентгеновских снимков легких;
- ◆ разработать методику предварительной обработки медицинских изображений;
- ◆ исследовать эффективность различных архитектур глубоких нейронных сетей;
- ◆ создать ансамблевую модель с использованием техник soft-voting и hard-voting;
- ◆ провести сравнительный анализ точности и информативности различных подходов машинного обучения;
- ◆ оценить клиническую применимость разработанной методики.

Научная новизна исследования заключается в:

- ◆ комплексном подходе к решению задачи медицинской диагностики с использованием мультимодельного ансамбля;
- ◆ оригинальной методике предобработки рентгеновских снимков, включающей денойзинг и цветовую коррекцию;
- ◆ экспериментальном сравнении различных архитектур нейронных сетей в задаче классификации легочных патологий;
- ◆ демонстрации эффективности ансамблевых методов машинного обучения для повышения точности диагностики;

Практическая значимость работы определяется:

- ◆ созданием инструмента экспресс-диагностики легочных заболеваний;
- ◆ возможностью существенного сокращения времени первичного обследования;
- ◆ потенциалом снижения нагрузки на врачей-рентгенологов;
- ◆ перспективой применения разработанной методики в клинической практике для скрининга и ранней диагностики.

Представленное исследование вносит существенный вклад в развитие технологий искусственного интеллекта в медицине и демонстрирует высокий потенциал машинного обучения для решения сложных диагностических задач.

Материалы и методы

Данные для исследования были взяты из разных источников, находящихся в общем доступе, таких как Kaggle, GitHub и др. [7–22]. Собранный датасет содержит рентгеновские снимки разных легочных заболеваний, таких как пневмония (4237 снимков), туберкулез (3494 снимка), фиброз (1686 снимков) и Covid-19 (3570 снимков), а также изображения здоровых легких (10191 снимок) (рисунок 1). Данные содержат как снимки мужчин, так и женщин. Данное исследование проводилось с целью выявления вышеперечисленных заболеваний вне зависимости от пола или возраста пациента.

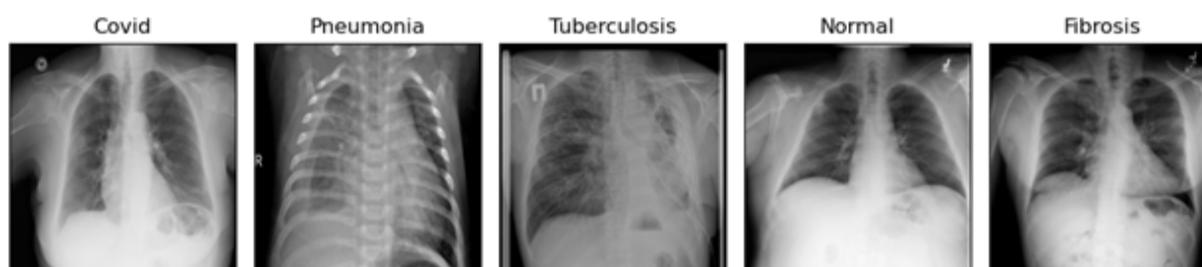


Рисунок 1 – Пример изображений для каждого класса

Для анализа был предложен план провести исследование в 3 вариациях:

- 1) использовать данные, не применяя к ним никаких методов предобработок кроме изменения разрешения изображений и нормализации;
- 2) использование CLANE и цветового фильтра, а также изменение разрешения изображений и нормализации;
- 3) использование денойзера, CLANE, цветового фильтра, изменение расширения и нормализация.

Модель DynamicCNN

В качестве денойзера была выбрана модель DynamicCNN.

DynamicCNN состоит из слоев трех типов:

- 1) Conv+ReLU: для первого слоя, 64 фильтра размером $3 \times 3 \times c$ используются для создания 64 карт признаков. $c = 1$ для черно-белых изображений и $c = 3$ для цветных.
- 2) Conv+BN+ReLU: для слоев с 2 по $(D - 1)$, используются 64 фильтра размером $3 \times 3 \times 64$, и между сверткой (Convolution) и активацией (ReLU) применяется пакетная нормализация (Batch Normalization).
- 3) Conv: для последнего слоя, c фильтров размером $3 \times 3 \times 64$ используются, чтобы восстановить выходные данные.

Принцип работы DynamicCNN состоит в том, что денойзинг восстанавливает чистое изображение X из зашумленного наблюдения Y (рисунок 2):

$$Y = X + v,$$

где v – это шум, обычно моделируемый как аддитивный белый гауссовский шум (AWGN).

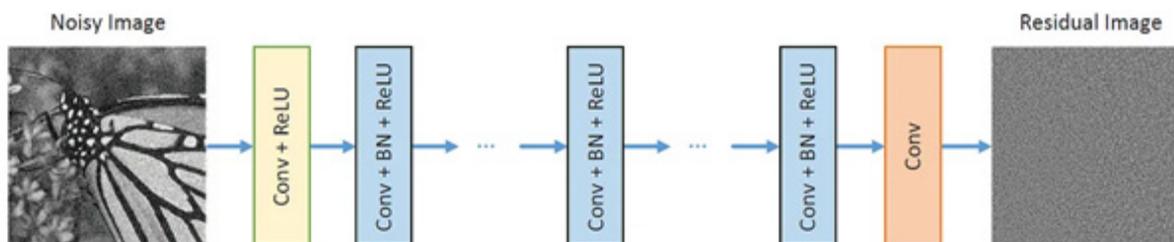


Рисунок 2 – Архитектура DynamicCNN

DynamicCNN упрощает задачу, обучаясь предсказывать шум, а не само чистое изображение. Это снижает сложность задачи, так как шум проще моделировать.

Вместо прямого предсказания чистого изображения $F(Y) = X$ сеть обучается остаточному отображению $R(Y)$, которое оценивает шум v . Чистое изображение затем вычисляется как:

$$X = Y - R(Y).$$

Этот подход позволяет сети сосредоточиться на извлечении шума, игнорируя детали исходного изображения.

Модель минимизирует среднеквадратичную ошибку между истинным шумом и предсказанным:

$$L(\theta) = \frac{1}{2N} \sum_{i=1}^N \|R(Y_i; \theta) - (Y_i - X_i)\|_F^2,$$

где Y_i – зашумленное изображение, X_i – истинное чистое изображение, $R(Y_i; \theta)$ – предсказанный шум, θ – параметры сети.

Пример применения модели отображаются на рисунке 3. Более подробно принципы действия и архитектура DynamicCNN описаны в работе [6].

Модель CLAHE

CLAHE (Contrast Limited Adaptive Histogram Equalization) – это метод улучшения контраста изображений, который широко используется в области обработки изображений. Использование CLAHE оправданно в ситуациях, когда необходимо улучшить видимость деталей в изображениях с низким контрастом, особенно в областях, где точность и качество изображения имеют критическое значение. Его адаптивный подход и контроль за контрастом делают его мощным инструментом в арсенале методов обработки изображений. CLAHE используется в различных методах медицинской визуализации, включая ультразвук, МРТ и КТ [23, 24].

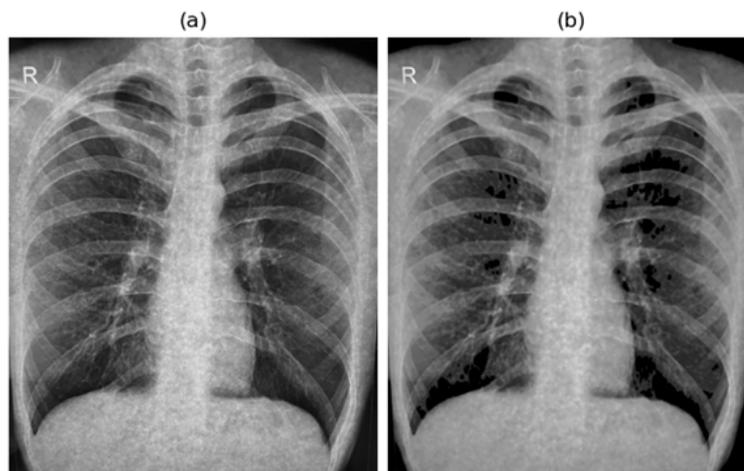


Рисунок 3 – Изображение до применения DynamicCNN (a) и после применения DynamicCNN (b)

Принцип работы CLAHE состоит в том, что сначала изображение разделяется на несколько неперекрывающихся блоков одинакового размера. Затем для каждой области рассчитывается гистограмма и определяется желаемое ограничивающее значение (clip limit) для обрезки гистограмм. Гистограммы перераспределяются таким образом, чтобы их высоты не превышали это ограничивающее значение. Этот процесс перераспределения может потребовать нескольких итераций в зависимости от коэффициента обрезки (clipping factor). Наконец, все соседние плитки объединяются с использованием билинейной интерполяции, и их значения уровня серого изменяются в соответствии с модифицированными гистограммами (рисунок 4).

Алгоритм CLAHE:

1. Разделить исходное изображение на неперекрывающиеся контекстные области. Общее количество областей равно .
2. Рассчитать гистограмму каждой контекстной области на основе уровней яркости, присутствующих в исходном массиве изображений.
3. Вычислить ограниченную адаптивную гистограмму контраста для каждой контекстной области, используя ограничивающее значение (Clip Limit):

$$N_{average} = (NrX \times NrY) / N_{levels}$$

$$N_{CLIP} = N_{clip} \times N_{gray}$$

Среднее количество перераспределенных пикселей вычисляется как:

$$N_{avggray} = N_{total} / N_{levels}$$

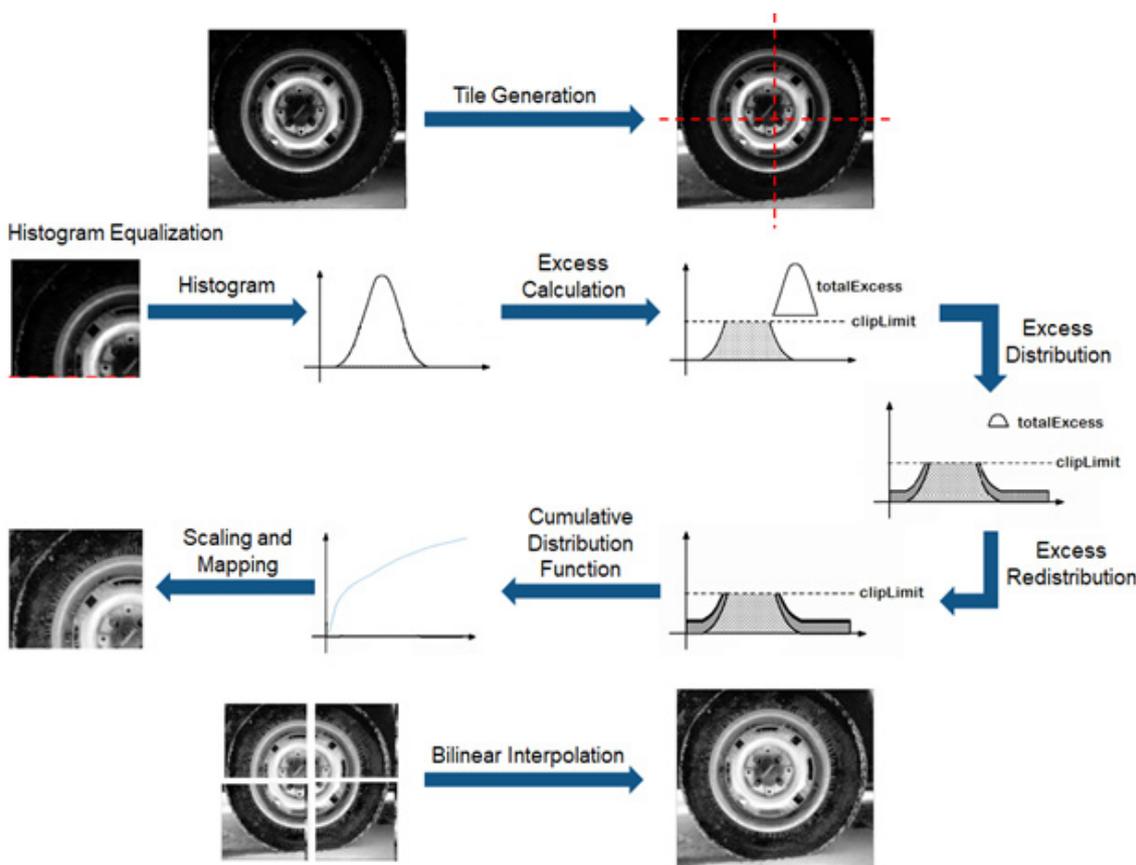


Рисунок 4 – Визуальное пошаговое представление алгоритма CLAHE

Среднее количество перераспределенных пикселей вычисляется как:

4. Для обрезки пикселей выполнить следующую процедуру:

```

procedure Clip
    if  $H(i) > N_{CLIP}$  then
         $H_{clip} = N_{CLIP}$ 
    else
        if  $(H(i) + N_{avggray}) > N_{CLIP}$  then
             $H_{clip} = N_{CLIP}$ 
        else
             $H_{clip}(i) = N_{CLIP} + H(i)$ 
        end if
    end if
end procedure
    
```

5. Перераспределить оставшиеся пиксели, пока они полностью не распределятся. Шаг перераспределения вычисляется как:

$$\frac{N_{levels}}{N_{remain}}$$

6. Вычислить новое распределение уровней яркости для пикселей внутри каждой контекстной области с использованием интерполяции.

Все используемые обозначения выглядят в таком виде, как $N_{average}$ – среднее количество пикселей, N_{levels} – количество уровней серого в контекстных областях, N_{rX} – количество пикселей по оси X в контекстной области, N_{rY} – количество пикселей по оси Y в контекстной области, N_{CLIP} – фактическое ограничивающее значение (clip limit), N_{clip} – нормализованное ограничивающее значение в диапазоне [0, 1], N_{total} – общее количество обрезанных пикселей, $N_{avggray}$ – среднее количество пикселей, подлежащих перераспределению, N_{remain} – количество пикселей, подлежащих перераспределению, $H(i)$ – исходная гистограмма каждой области на i-м уровне серого, $H_{clip}(i)$ – обрезанная гистограмма каждой области на i-м уровне серого. Пример применения модели отображается на рисунке 5.

Цветовой фильтр

Цветовые фильтры могут быть использованы для улучшения контраста в черно-белых изображениях, что позволяет лучше различать детали. В нашем случае для удобства был выбран фильтр COLORMAP_JET из библиотеки OpenCV [25]. Пример применения модели отображается на рисунке 6.

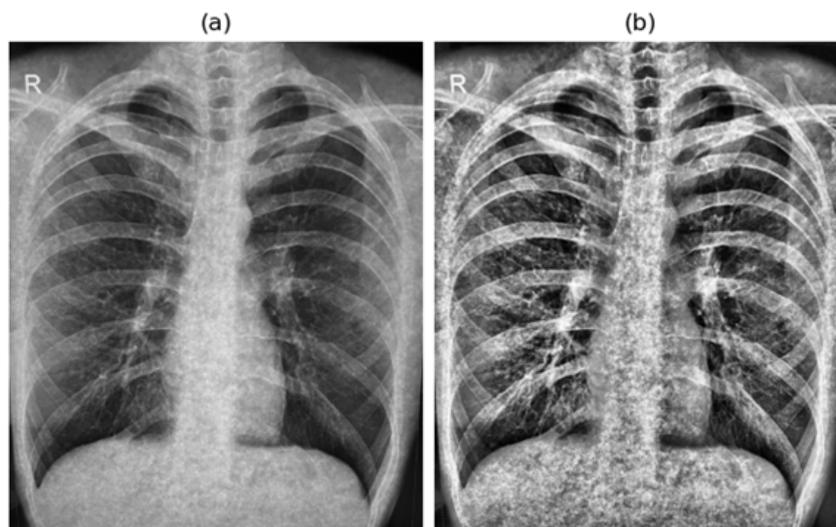


Рисунок 5 – До использования CLAHE (a)
и после использования CLAHE (b)

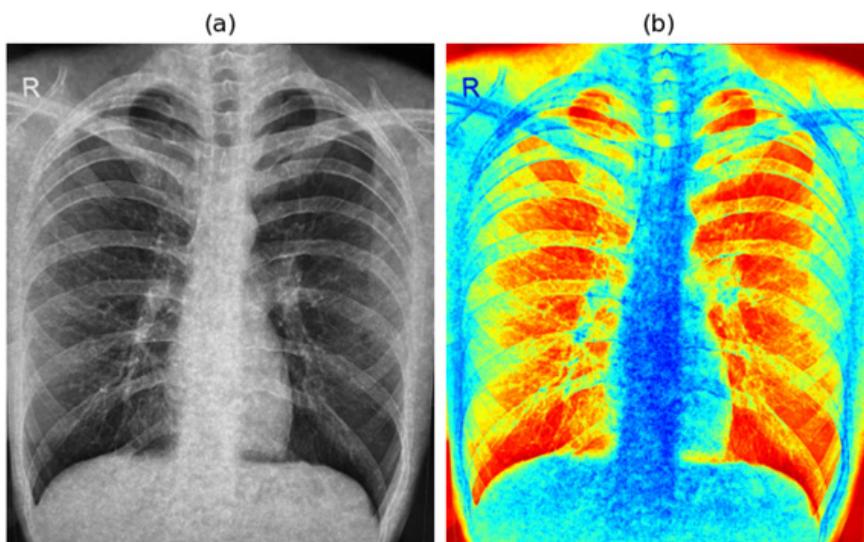


Рисунок 6 – до окрашивания (a) и после окрашивания (b)

Использование различных моделей архитектуры

1) Архитектура Densenet121, полное описание использования архитектуры Densenet121 подробно дано в работе [26] и отображено на рисунке 7. Более подробно с описанием модели можно ознакомиться в работе [26].

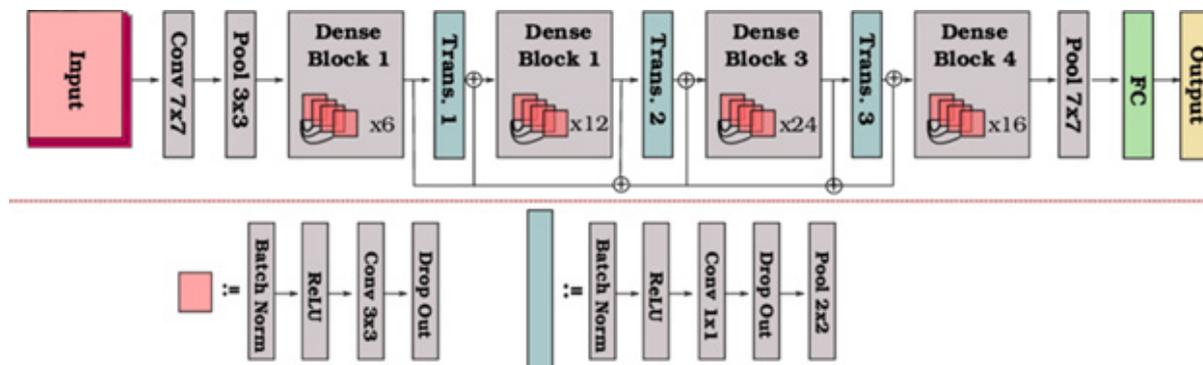


Рисунок 7 – Архитектура Densenet121 [26]

2) Архитектура MobileNetV2, полное описание использования архитектуры MobileNetV2 дано в работе [27] и отображено на рисунке 8. Более подробно с описанием модели можно ознакомиться в работе [27].

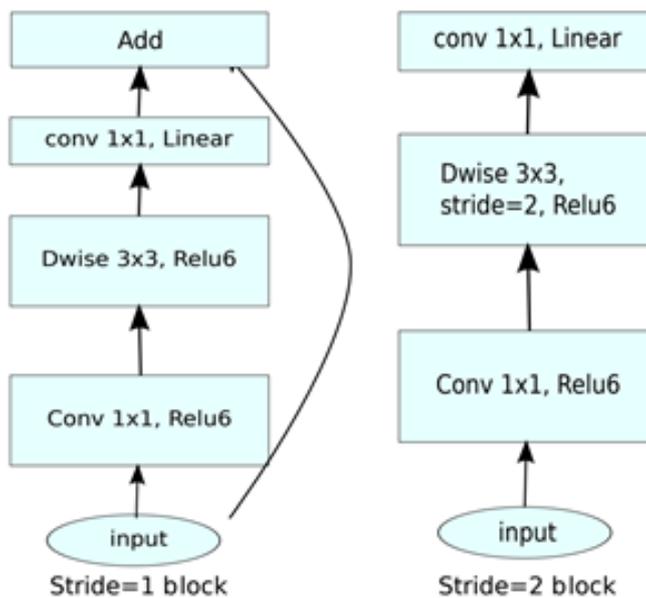


Рисунок 8 – Архитектура MobileNetV2 [27]

3) Архитектура EfficientNetb0, полное описание использования архитектуры EfficientNetb0 дано в работе [28] и отображено на рисунке 9. Более подробно с описанием модели можно ознакомиться в работе [28].

4) Архитектура SENet, полное описание использования архитектуры SENet отображены на рисунке 10.

5) Архитектура SE-ResNet, полное описание использования архитектуры SE-ResNet дано в работе [29] и отображено на рисунке 11. Более подробно с описанием модели можно ознакомиться в работе [29].

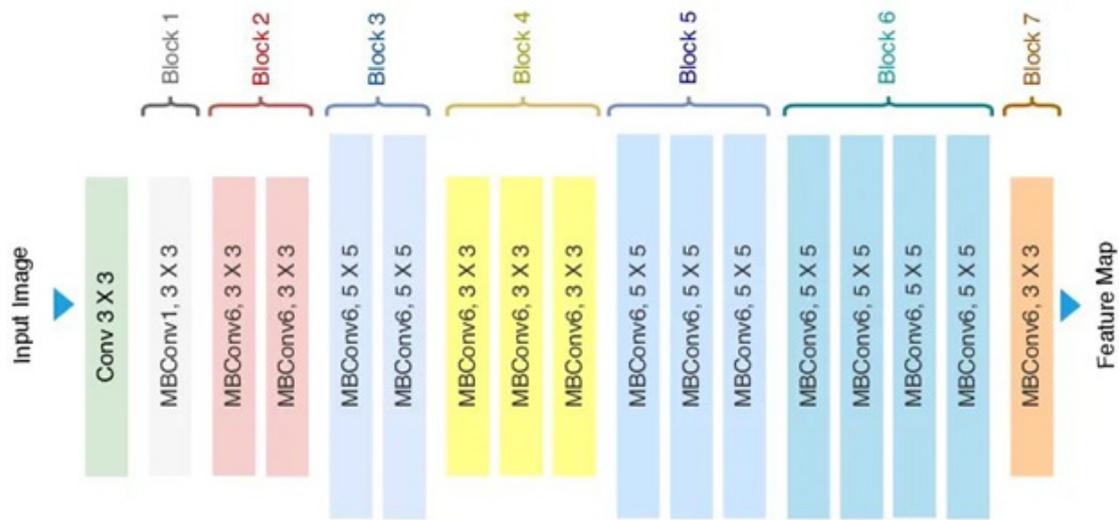


Рисунок 9 – Архитектура EfficientNetb0 [28]

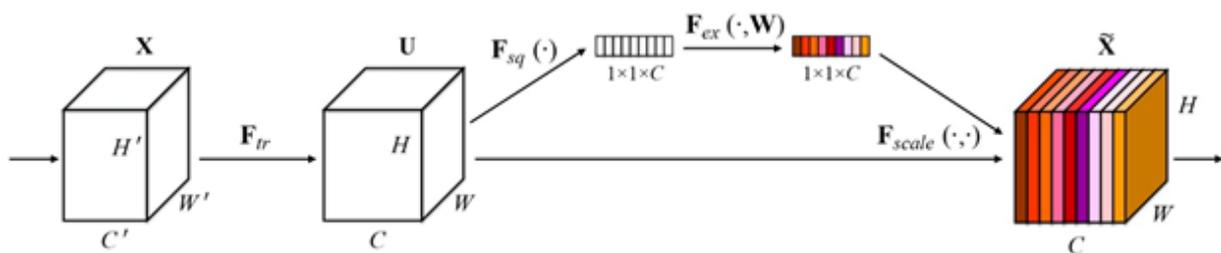


Рисунок 10 – Архитектура Squeeze-and-Excitation блока

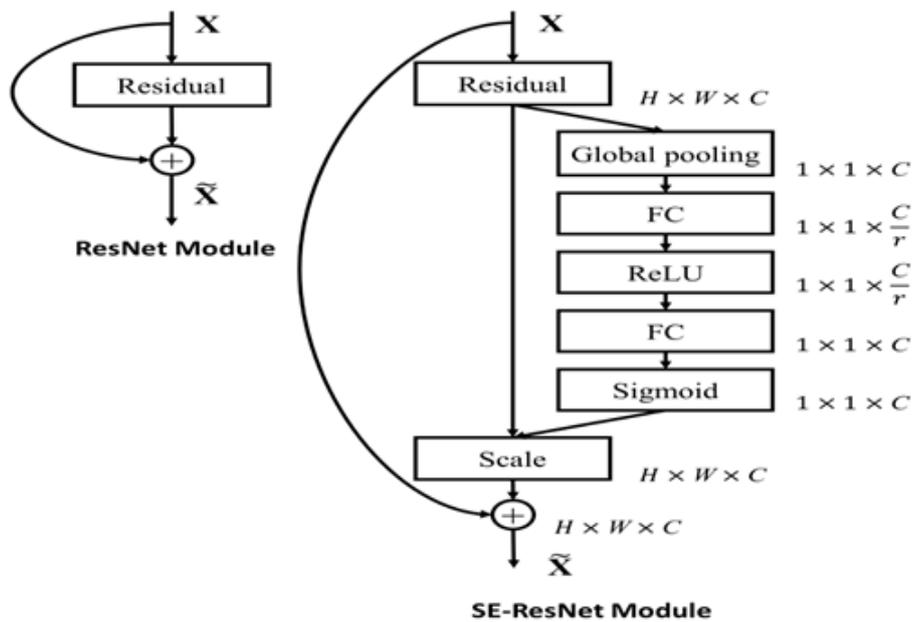


Рисунок 11 – Архитектура SE-ResNet [29]

6) Архитектура ShuffleNetV2, полное описание использования архитектуры ShuffleNetV2 дано в работе [30] и отображено на рисунке 12. Более подробно с описанием модели можно ознакомиться в работе [30].

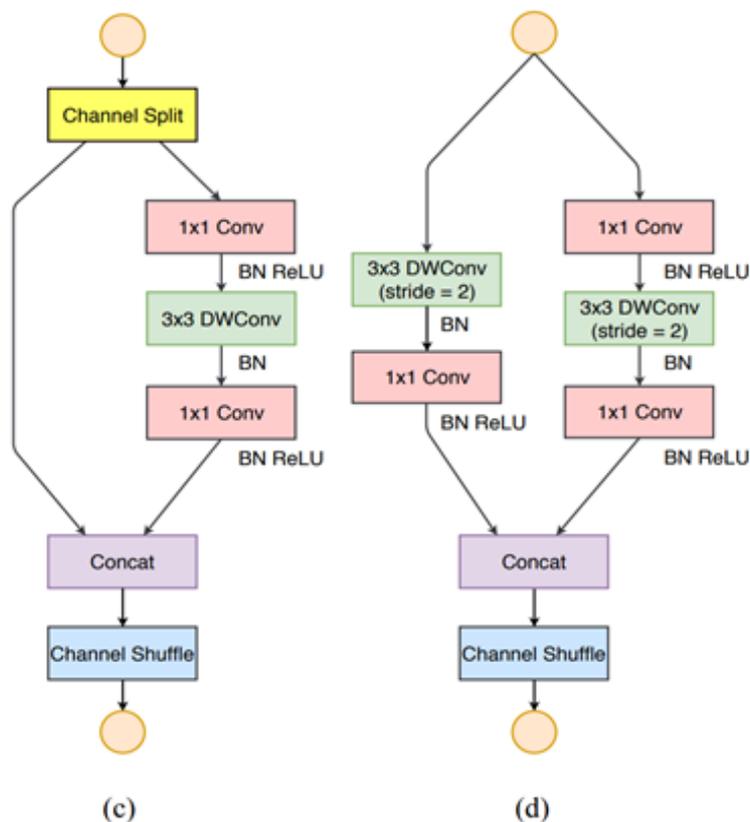


Рисунок 12 – Базовая единица ShuffleNet (c); единица для пространственного понижение разрешения (d) [30].

Результаты и обсуждение

Все модели сначала обучались, а затем тестировались на тестовых данных. Для оптимизации модели был использован оптимизатор Adam [31], в качестве функции потерь (loss function) – перекрестная энтропия (cross entropy), размер пакетов (batch size) – 64, темп обучения (learning rate) – 0.001.

В ходе обучения были применены такие методы, как ранняя остановка и планировщик скорости обучения. Ранняя остановка позволяет остановить обучение модели, когда ее производительность на валидационном наборе данных начинает ухудшаться, что помогает избежать переобучения. Сокращает время и вычислительные ресурсы, необходимые для обучения.

Планировщик скорости обучения позволяет модели быстрее достигать минимума функции потерь, начиная с высокой скорости обучения и постепенно снижая ее. Помогает избежать колебаний и нестабильности в процессе оптимизации, что может происходить при использовании слишком высокой скорости обучения. Позволяет модели лучше изучать сложные паттерны, изменяя скорость обучения в зависимости от этапа обучения.

Затем комбинация этих моделей была использована в ансамблевых классификаторах двух типов: soft-voting и hard-voting. По полученным результатам лучше себя проявляет soft-voting ансамбль.

Можно будет обозначить основное различие, которое отображается на рисунке 13. Для типа Hard-voting принятия окончательного предсказания используется простое большинство голосов, основанное на наиболее частом классе, предсказанном отдельными моделями. А для Soft-voting учитываются вероятностные оценки каждого класса, предсказанные отдельными моделями, и они усредняются для получения более точного окончательного предсказания. При работе с несбалансированными наборами данных мягкое голосование может помочь смягчить предвзятость в сторону большинства, принимая во внимание вероятности всех классов.

Soft-voting ансамбль показывает результаты лучше, чем каждая отдельно взятая модель, что говорит в пользу использования ансамблевых моделей.

Были выбраны 3 варианта предобработки данных для их последующего сравнения.

A – никаких предобработок кроме изменения расширения до и нормализации.

B – CLANE с последующим применением цветового фильтра, также изменение расширения до и нормализация.

C – денойзер с последующим применением CLANE и цветового фильтра, также изменение расширения до и нормализация.

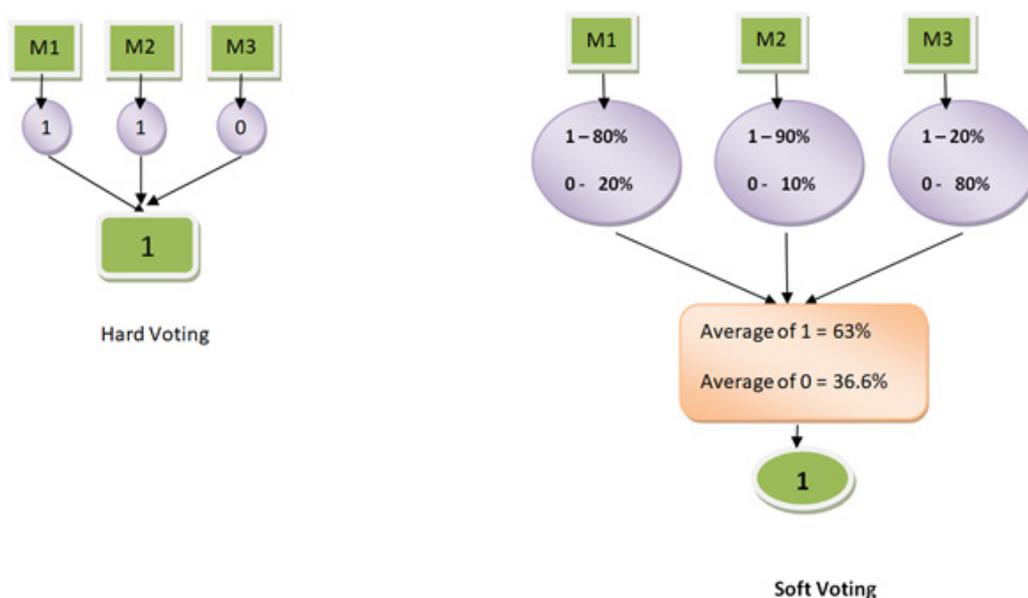


Рисунок 13 – Разница между Hard-Voting и Soft-Voting

Для проведения оценки полученных результатов в работе были использованы различные метрики. При работе с несбалансированными данными Accuracy не самая репрезентативная метрика, поэтому упор был сделан на Precision, Recall, F1-score метрики (таблица 1).

Precision отражает надежность положительных предсказаний модели. В контексте несбалансированных данных эта метрика приобретает особую значимость, поскольку позволяет удостовериться, что объекты, классифицированные как принадлежащие к классу с меньшим количеством данных, действительно к нему относятся. Высокий показатель precision свидетельствует о минимизации ложноположительных результатов, что критически важно в ситуациях, когда цена ошибочной классификации высока.

Recall демонстрирует способность модели обнаруживать все истинно положительные случаи. Данная метрика имеет первостепенное значение для предотвращения пропуска экземпляров класса с меньшим количеством данных. Высокое значение recall указывает на то, что модель эффективно идентифицирует большинство действительно положительных случаев,

что особенно важно в ситуациях, когда пропуск даже небольшой доли таких случаев может иметь серьезные последствия.

F1-score представляет собой гармоническое среднее между precision и recall, обеспечивая сбалансированную оценку эффективности модели. Высокий показатель F1-score говорит об оптимальном балансе между минимизацией ложноположительных и ложноотрицательных результатов, что является желательным результатом в случаях, когда оба типа ошибок имеют существенное значение.

Таблица 1 – Матрица ошибок

| | | | |
|------------------------|-------------------|------------|------------|
| | Реальные значения | | |
| Предсказанные значения | | Негативные | Позитивные |
| Негативные | | TN | FN |
| Позитивные | | FP | TP |

Для вычисления метрики Precision, Recall, F1-score были использованы формулы, которые имеют следующий вид:

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

Были выбраны 3 варианта предобработки данных для их последующего сравнения.

A – никаких предобработок кроме изменения расширения до и нормализации.

B – CLANE с последующим применением цветового фильтра, также изменение расширения до и нормализация.

C – денойзер с последующим применением CLANE и цветового фильтра, также изменение расширения до и нормализация.

По полученным результатам можно заметить, что все наилучшие значения достигаются при использовании метода предобработки с Soft-Voting ансамблем. Также видно, что Soft-Voting ансамбль показывает лучше результаты для каждого варианта A, B или C в сравнении с Hard-Voting ансамблем (таблица 2).

Таблица 2 – Полученные результаты

| | Hard-Voting | | | Soft-Voting | | |
|-----------|-------------|-------|-------|-------------|-------|-------|
| | A | B | C | A | B | C |
| Precision | 94.92 | 98.14 | 98.51 | 96.05 | 98.38 | 98.60 |
| Recall | 94.98 | 98.09 | 98.47 | 96.07 | 98.36 | 98.58 |
| F1-score | 94.85 | 98.10 | 98.48 | 96.02 | 98.37 | 98.58 |

Заклучение

В рамках данного исследования была разработана и протестирована методология классификации заболеваний легких с использованием глубоких сверточных нейронных сетей, включая ShuffleNet, EfficientNet, MobileNet, SENet и DenseNet. Несмотря на отсутствие сбалансированности исходного набора данных, предложенный подход продемонстрировал высокую эффективность в решении задачи классификации легочных заболеваний.

Полученные результаты подчеркивают значимость выбора архитектуры модели для задач медицинской диагностики и подтверждают потенциал глубокого обучения в анализе рентгеновских изображений. Однако отсутствие тонкой настройки гиперпараметров, таких как темп обучения (learning rate), размер пакетов (batch size), выбор оптимизаторов и функций потерь (loss function), может ограничивать достижение максимальной производительности моделей.

Будущие исследования могут быть направлены на:

- ♦ тщательный подбор гиперпараметров для улучшения качества обучения моделей;
- ♦ сравнение различных оптимизаторов и функций потерь;
- ♦ применение методов предобработки данных для улучшения качества изображений;
- ♦ использование методов аугментации данных для балансировки классов и повышения устойчивости моделей;
- ♦ использование новых моделей и подходов, таких как Visual Transformers, RNN, и других видов CNN;

Дальнейшее развитие данного направления, включая интеграцию предложенных улучшений, может существенно повысить надежность и точность диагностики, обеспечивая широкий спектр возможностей для внедрения в клиническую практику.

Информация о финансировании. Работа поддержана грантом Министерства науки и высшего образования Республики Казахстан (AP23488833).

ЛИТЕРАТУРА

1 Liu, X., Yu, Z., Tan, L. Deep learning for lung disease classification using transfer learning and a customized convolutional neural network architecture with attention. Proceedings of the IEEE 2nd International Conference on Sensors, Electronics and Computer Engineering (ICSECE), 2024, pp. 341–346.

2 Shamrat, F.M.J.M., Azam, S., Karim, A., Islam, R., Tasnim, Z., Ghosh, P., De Boer, F. LunGNeT22: A fine-tuned model for multiclass classification and prediction of lung disease using X-ray images. Journal of Personalized Medicine, 12, 680 (2022).

3 De Sousa, P.M., Carneiro, P.C., Oliveira, M.M., Pereira, G.M., Da Costa, C.A. Jr., De Moura, L.V., Mattjie, C., Da Silva, A.M.M., Patrocínio, A.C. COVID-19 classification in X-ray chest images using a new convolutional neural network: CNN-COVID. Research on Biomedical Engineering, 38, 87–97 (2021).

4 Ashraf, S.M.N., Mamun, Md.A., Abdullah, H.Md., Alam, Md.G.R. SynthEnsemble: A fusion of convolutional neural networks, vision transformers, and hybrid models for multi-label chest X-ray classification. Proceedings of the 26th International Conference on Computer and Information Technology (ICCIT), 2023, pp. 1–6.

5 Soud, A., Sakli, N., Sakli, H. Classification and prediction of lung diseases from chest X-ray images using MobileNetV2. Applied Sciences, 11, 2751 (2021).

6 Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L. Beyond a Gaussian denoiser: Residual learning of deep convolutional neural networks for image denoising. IEEE Transactions on Image Processing, 26, 3142–3155 (2017).

7 Chest X-ray COVID-19 Pneumonia Dataset. Kaggle. URL: <https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database> (accessed: 05.01.2022).

8 COVID Chest X-ray Image Dataset for Research. Kaggle (accessed 05 January 2022). URL: <https://www.kaggle.com/datasets/pranavraikokte/covid19-image-dataset>

- 9 Chest X-ray Pneumonia, COVID-19, Tuberculosis Dataset. Kaggle. URL: <https://www.kaggle.com/datasets/amanullahasraf/chest-xray-pneumonia-covid19-tuberculosis> (accessed: 05.01.2022).
- 10 COVID-19 Chest X-ray Dataset. GitHub. URL: <https://github.com/ieee8023/covid-chestxray-dataset> (accessed: 03.01.2022).
- 11 COVIDx CXR-2 Dataset. Kaggle. URL: <https://www.kaggle.com/datasets/andyczao/covidx-cxr2> (accessed: 04.01.2022).
- 12 Tuberculosis Chest X-ray Dataset. Kaggle. URL: <https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset> (accessed: 04.01.2022).
- 13 Tuberculosis Chest X-rays Shenzhen Dataset. Kaggle. URL: <https://www.kaggle.com/datasets/raddar/tuberculosis-chest-xrays-shenzhen> (accessed: 05.01.2022).
- 14 COVID-19 Detection Dataset. Kaggle. URL: <https://www.kaggle.com/datasets/darshan1504/covid19-detection-xray-dataset> (accessed: 05.01.2022).
- 15 Pneumothorax Binary Classification Dataset. Kaggle. URL: <https://www.kaggle.com/datasets/vbookshelf/pneumothorax-chest-xray-images> (accessed: 05.01.2022).
- 16 National Institutes of Health Chest X-ray Dataset (Sample). Kaggle. URL: <https://www.kaggle.com/datasets/nih-chest-xrays/sample> (accessed: 06.01.2022).
- 17 COVID Normal Viral Opacity Dataset Version 2. Kaggle. URL: <https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database> (accessed: 03.01.2022).
- 18 X-ray Report Dataset. Kaggle. URL: <https://www.kaggle.com/datasets/raddar/xray-report-dataset> (accessed: 04.01.2022).
- 19 ChestX-ray14 Dataset. National Institutes of Health. URL: <https://nihcc.app.box.com/v/ChestXray-NIHCC> (accessed: 03.01.2022).
- 20 Worldwide Chest X-ray Datasets. Kaggle. URL: <https://www.kaggle.com/datasets/nikhilpandey360/chest-xray-multiple-datasets> (accessed: 03.01.2022).
- 21 COVID-19 Radiography Database. Kaggle. URL: <https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database> (accessed: 06.01.2022).
- 22 Tuberculosis Chest X-rays Shenzhen Dataset. Kaggle. URL: <https://www.kaggle.com/datasets/raddar/tuberculosis-chest-xrays-shenzhen> (accessed: 06.01.2022).
- 23 Hussein, F., Mughaid, A., AlZu'bi, S., El-Salhi, S.M., Abuhaija, B., Abualigah, L., Gandomi, A.H. Hybrid contrast limited adaptive histogram equalization and convolutional neural network deep models for lung disease classification from X-ray images. *Electronics*, 11, 3075 (2022).
- 24 Sonali, N., Sahu, S., Singh, A.K., Ghrera, S.P., Elhoseny, M. An approach for denoising and contrast enhancement of retinal fundus images using contrast limited adaptive histogram equalization. *Optics and Laser Technology*, 110, 87–98 (2018).
- 25 Das, H., Rout, J.K., Moharana, S.C., Dey, N. *Applied intelligent decision making in machine learning*. CRC Press, Boca Raton, 320 p. (2020).
- 26 Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q. Densely connected convolutional networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4700–4708.
- 27 Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C. MobileNetV2: Inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510–4520.
- 28 Tan, M., Le, Q.V. EfficientNet: Rethinking model scaling for convolutional neural networks. *arXiv preprint*, arXiv:1905.11946 (2019).
- 29 Hu, J., Shen, L., Albanie, S., Sun, G., Wu, E. Squeeze-and-excitation networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42, 2011–2023 (2019).
- 30 Ma, N., Zhang, X., Zheng, H.-T., Sun, J. ShuffleNet V2: Practical guidelines for efficient convolutional neural network architecture design. *Lecture Notes in Computer Science*, 11218, 122–138 (2018).
- 31 Kingma, D.P., Ba, J.L. Adam: A method for stochastic optimization. *arXiv preprint*, arXiv:1412.6980 (2014).

¹Исахов А.,

профессор, ORCID ID: 0000-0002-1937-8615,
e-mail: alibek.issakhov@gmail.com

¹Абылкасымова А.Б.,

кауымдастырылған профессор, ORCID ID: 0000-0002-5967-6959,
e-mail: abylkassymova.aizhan@gmail.com

¹Бревнов Т.

бакалавр, ORCID ID: 0009-0004-5395-4569, e-mail: t_brevnov@kbtu.kz

¹Қазақстан-Британ техникалық университет,
Алматы қ., Қазақстан

АДАМНЫҢ ӨКПЕ ЖҮЙЕСІНІҢ РЕНТГЕНДІК КЕСКІНДЕРІН МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІ АРҚЫЛЫ ТАЛДАУ ЖӘНЕ АЛДЫН АЛА ДИАГНОЗ ҚОЮ

Аңдатпа

Бұл мақалада патологиялық өзгерістерді ерте анықтау үшін тыныс алу жолдарының радиографиялық кескіндерін автоматтандырылған талдауға арналған машиналық оқыту әдістерінің кешенді зерттеуі ұсынылған. Өкпе ауруларын жіктеу әдісі DenseNet121, MobileNetV2, EfficientNetB0, SENet және ShuffleNetV2 архитектураларын қамтитын терең конволюциялық нейрондық желілер ансамблін пайдаланып ұсынылды және енгізілді. Зерттеуде сонымен қатар әртүрлі кескінді алдын ала өңдеу әдістерінің тиімділігі, соның ішінде қосымша өңдеусіз шикі қара-ақ рентгенограммаларды пайдалану, түсті сүзумен біріктірілген CLAHE (Contrast Limited Adaptive Histogram Equalization) әдісі және шуды басу үшін DynamicCNN нейрондық желілік шуды басу құралы салыстырылды. Тәжірибе нәтижелері жұмсақ дауыс беру стратегиясын қолданатын ансамбльдік тәсіл жеке модельдермен салыстырғанда жіктеу дәлдігін статистикалық тұрғыдан айтарлықтай жақсартатынын көрсетті. Нәтижелер ұсынылған тәсілдің жоғары тиімділігін растайды және медициналық диагностика жмен клиникалық шешім қабылдауды қолдау үшін ансамбльдік терең оқыту модельдерінің әлеуетін көрсетеді.

Тірек сөздер: машиналық оқыту әдістері, терең нейрондық желі ансамбльдері, рентгендік кескіндер, DynamicCNN денойзері.

¹Issakhov A.A.,

Professor, ORCID ID: 0000-0002-1937-8615,
e-mail: alibek.issakhov@gmail.com

¹Abylkassymova A.B.,

Associate Professor, ORCID ID: 0000-0002-5967-6959,
e-mail: abylkassymova.aizhan@gmail.com

¹Brevnov T.

Bachelor's student, ORCID ID: 0009-0004-5395-4569,
e-mail: t_brevnov@kbtu.kz

¹Kazakh-British technical university,
Almaty, Kazakhsan

MACHINE LEARNING ANALYSIS OF HUMAN LUNG X-RAY IMAGES TO MAKE A PRELIMINARY DIAGNOSIS

Abstract

The article presents a comprehensive study of the application of machine learning methods for automated analysis of radiographic images of the respiratory system for the early detection of pathological changes. A method

for classifying pulmonary diseases based on an ensemble of deep convolutional neural networks, including the DenseNet121, MobileNetV2, EfficientNetB0, SEnet, and ShuffleNetV2 architectures, is proposed and implemented. The study included a comparative analysis of the effectiveness of various image preprocessing methods, including the use of raw black-and-white X-ray images without additional processing, the use of the CLAHE (Contrast Limited Adaptive Histogram Equalization) method in combination with color filtering, and the use of the DynamicCNN neural network denoiser for noise suppression. Experimental results showed that the ensemble approach using the soft voting strategy provides a statistically significant improvement in classification accuracy compared to individual models. The obtained results confirm the high efficiency of the proposed approach and demonstrate the potential of using ensemble deep learning models in medical diagnostics and clinical decision support tasks.

Keywords: machine learning methods, deep neural network ensembles, X-ray images, DynamicCNN denoiser.

Received: January 16, 2025; accepted: February 17, 2026.