

یک روش کارآمد یادگیری عمیق برای تشخیص عفونت COVID-19 با استفاده از تصاویر رادیوگرافی قفسه سینه

چکیده

در سال‌های اخیر، جامعه پژوهشی علاقه چشمگیری به طراحی سامانه‌های خودکار برای تشخیص بیماری کووید-۱۹ با استفاده از رویکردهای یادگیری عمیق و تصاویر رادیوگرافی قفسه سینه نشان داده است. با این حال، روش‌های پیشرفتی یادگیری عمیق، بهویژه شبکه‌های عصبی کانولوشنی (CNN)، به تعداد زیادی پارامتر قابل یادگیری و حافظه بالا نیاز دارند؛ بنابراین ممکن است برای تشخیص بلادرنگ مناسب نباشند. از این‌رو، طراحی یک مدل CNN سبکوزن برای تشخیص سریع و دقیق کووید-۱۹ یک نیاز ضروری است.

در این پژوهش، یک مدل سبکوزن **LW-CORONet** با نام **CNN** پیشنهاد شده است که شامل دنباله‌ای از لایه‌های کانولوشن، تابع فعال ساز ReLU و لایه‌های pooling است و در ادامه دو لایه کاملاً متصل دارد. این مدل پیشنهادی با تنها پنج لایه قابل یادگیری، استخراج ویژگی‌های معنادار از تصاویر رادیوگرافی قفسه سینه (CXR) را تسهیل می‌کند.

این مدل با استفاده از دو مجموعه‌داده بزرگ CXR مورد ارزیابی قرار گرفته است (مجموعه‌داده ۱: شامل ۲۲۵۰ تصویر و مجموعه‌داده ۲: شامل ۱۵,۹۹۹ تصویر). دقت دسته‌بندی به دست آمده برای مجموعه‌داده اول در حالت چند کلاسه و دودویی به ترتیب ۹۸,۶۷٪ و ۹۹,۰۰٪، و برای مجموعه‌داده دوم ۹۵,۶۷٪ و ۹۶,۲۵٪ بوده است. نتایج با چهار مدل CNN از پیش‌آموزش‌دیده معاصر و مدل‌های پیشرفتی دیگر مقایسه شده‌اند. همچنین تأثیر چند هایپرپارامتر از جمله نوع بهینه‌ساز، اندازه بج، و نرخ یادگیری بررسی شده است.

مدل پیشنهادی به پارامترهای کمتر و فضای حافظه کمتری نیاز دارد؛ بنابراین برای تشخیص کووید-۱۹ مؤثر است و می‌تواند به عنوان یک ابزار کمکی برای پشتیبانی از رادیولوژیست‌ها در فرآیند تشخیص به کار گرفته شود.

کلیدواژه‌ها:

COVID-19، LW-CORONet، CNN، شبکه عصبی کانولوشنی، تصاویر رادیوگرافی، انتقال یادگیری، قفسه سینه

۱ مقدمه

همه‌گیری کووید-۱۹ در مدت زمانی بسیار کوتاه موجب بحران‌های شدید در نظام سلامت سراسر جهان شد. این پاندمی در اوایل دسامبر ۲۰۱۹ در ووهان چین آغاز شد و بعدها به عنوان یک همه‌گیری جهانی اعلام شد. پژوهشگران گزارش کردند که کووید-۱۹ یک بیماری بهشت و اگیردار است که سیستم تنفسی را به طور جدی تحت تأثیر قرار می‌دهد و دارای علائمی مانند سرفه خشک، درد عضلانی، تب، سردرد، درد قفسه سینه و گلودرد است.

روش‌های تشخیصی پزشکی موجود، هم از نظر دارو و هم منابع بیمارستانی، با کمبودهایی برای درمان عفونت کووید-۱۹ مواجه‌اند. روش RT-PCR که یک ابزار دستی، زمان‌بر و پرهزینه است، همچنان رایج‌ترین روش تشخیصی است و خطر ابتلای کارکنان پزشکی را افزایش می‌دهد. پاندمی همچنان ادامه دارد و ظهور انواع جدید ویروس موجب افزایش نرخ مرگ‌ومیر در بسیاری از کشورها شده است. بنابراین، نیاز شدیدی به یک روش ایمن و کارآمد برای تشخیص زودهنگام عفونت کووید-۱۹ وجود دارد.

اهمیت دو مدل‌الیته تصویربرداری شامل تصاویر رادیوگرافی قفسه سینه (CXR) و سی‌تی‌اسکن (CT) در تشخیص کووید-۱۹ مورد بررسی قرار گرفته است. با این حال، بررسی دستی و بصری این تصاویر برای متخصصان زمان‌بر، خسته‌کننده و مستعد خطا است. در سال‌های اخیر، تلاش‌های گسترده‌ای برای توسعه مدل‌های خودکار مبتنی بر هوش مصنوعی با هدف افزایش دقت تشخیص و کاهش بار کاری رادیولوژیست‌ها انجام شده است. الگوریتم‌های یادگیری عمیق، بهویژه شبکه‌های عصبی کانولوشنی (CNN)، راه حل‌های مؤثری برای تشخیص پنومونی در تصاویر CXR ارائه کرده‌اند و پس از آغاز پاندمی، مدل‌های بسیاری برای تشخیص کووید-۱۹ معرفی شده‌اند.

مدل‌های مختلفی مانند SqueezeNet، ResNet-50، COVIDX-Net، DarkCovidNet و ResNet50 و ترکیبات چندمدلی ارائه شده‌اند که دقت‌های متفاوتی در سناریوهای دودویی و چندکلاسه به دست آورده‌اند. برخی از این مدل‌ها بر پایه تصاویر CXR و برخی دیگر بر مبنای CT طراحی شده‌اند. همچنین مدل‌هایی با استفاده از روش‌های مبتنی بر ترکیب شبکه‌ها، تحلیل ویژگی‌ها، یادگیری چندمقیاسه و استفاده از الگوریتم‌های تکمیلی نیز معرفی شده‌اند.

با وجود این تلاش‌ها، اکثر روش‌های موجود تنها با مجموعه‌داده‌های کوچک ارزیابی شده‌اند و معمولاً به تعداد بسیار زیادی پارامتر و حافظه زیاد نیاز دارند. از این‌رو، طراحی یک روش یادگیری عمیق سبک‌وزن اهمیت ویژه‌ای دارد. علاوه‌بر این، عملکرد مدل‌های یادگیری عمیق به هایپرپارامترهایی مانند نرخ یادگیری، اندازه بچ، نوع بهینه‌ساز و تعداد اپوک‌ها وابسته است و تعداد کمی از مطالعات به تحلیل جامع این پارامترها پرداخته‌اند.

بررسی‌های اخیر نشان می‌دهد که بسیاری از مدل‌های مرسوم تنها در سناریوهای دودویی ارزیابی شده‌اند، در حالی که دسته‌بندی چندکلاسه—با وجود چالش‌برانگیز بودن—بسیار مورد نیاز است. برای رفع این محدودیت‌ها، در مطالعه حاضر یک CNN سبک وزن برای تشخیص خودکار عفونت کووید-۱۹ ارائه می‌شود. مهم‌ترین نوآوری‌های این پژوهش عبارت‌اند از:

- ارائه یک مدل CNN سبک وزن برای تشخیص کووید-۱۹ در تصاویر CXR که به هزینه محاسباتی و حافظه کم نیاز دارد و برای تشخیص بلادرنگ مناسب است.
- تحلیل تأثیر هایپرپارامترهایی مانند نوع بهینه‌ساز، تعداد اپوک‌ها، اندازه بج و نرخ یادگیری.
- بررسی عملکرد مدل در سناریوهای دودویی (عادی و کووید-۱۹) و چندکلاسه (عادی، کووید-۱۹ و پنومونی).
- انجام ارزیابی عملکرد با استفاده از دو مجموعه‌داده بزرگ.
- مقایسه مدل پیشنهادی با معماری‌های پیش‌آموزش‌دیده مانند ResNet-101، VGG-19، Xception و DenseNet-121 از نظر تعداد پارامتر، نیاز حافظه و دقت. همچنین مقایسه با چند روش پیشرفته دیگر نیز انجام شده است.

ساختار مقاله به این صورت است که در بخش ۲ مجموعه‌داده‌ها و روش پیشنهادی معرفی می‌شوند، بخش ۳ شامل تنظیمات آزمایش و نتایج است و در بخش ۴ جمع‌بندی ارائه می‌شود.

۲. مواد و روش‌ها

در این بخش، مجموعه‌داده‌های به کاررفته در این پژوهش و روش پیشنهادی برای دسته‌بندی بیماری‌های مختلف ریه به صورت کامل توضیح داده می‌شوند. همچنین چند معماری CNN متداول که برای مقایسه استفاده شده‌اند معرفی می‌شوند.

۲.۱. مجموعه‌داده‌های مورد استفاده

مجموعه‌داده اول (Dataset-1) شامل تصاویر CXR در سه کلاس عادی، کووید-۱۹ و پنومونی است که از مخزن Figshare گردآوری شده‌اند. تعداد تصاویر رادیوگرافی مربوط به بیماران مبتلا به کووید-۱۹ هنوز محدود است و مسئله عدم توازن داده همچنان یک چالش مهم محسوب می‌شود. بنابراین در این پژوهش، در مجموع ۲۲۵۰ نمونه CXR برای این سه کلاس انتخاب شد که برای هر کلاس ۷۵۰ نمونه در نظر گرفته شده

است Dataset-1 شامل موارد کووید-۱۹ استخراج شده از چند منبع باز است که به صورت دورهای روزرسانی می‌شوند.

برای بررسی بیشتر کارایی مدل پیشنهادی، از یک مجموعه داده دیگر با نام COVIDx-V7A که در این پژوهش Dataset-2 نام‌گذاری شده (نیز استفاده شد. این مجموعه شامل ۱۵,۹۹۹ تصویر CXR از ۱۲۲ بیمار در ۵۱ کشور است و در حال حاضر یکی از بزرگ‌ترین مجموعه داده‌های عمومی موجود محسوب می‌شود. این تصاویر از پنج مخزن داده مختلف گردآوری شده‌اند.

جدول ۱ مشخصات دقیق هر دو مجموعه داده را نشان می‌دهد. نمونه‌ای از تصاویر CXR نمای قدامی در کلاس‌های مختلف مربوط به Dataset-1 در شکل ۱ ارائه شده است.

Table 1. Description of the datasets.

Dataset	Number of CXR Images			Total
	Normal	Pneumonia	COVID-19	
Dataset-1	750	750	750	2250
Dataset-2	8066	5575	2358	15,999

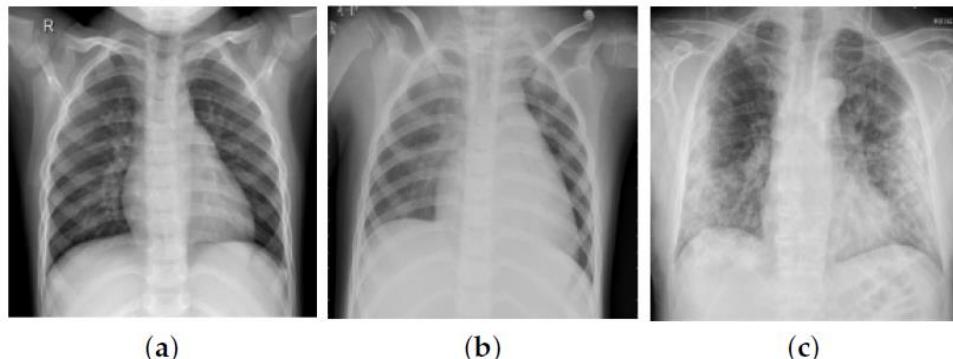


Figure 1. Frontal-view CXR samples of three categories: (a) Normal, (b) Pneumonia, and (c) COVID-19.

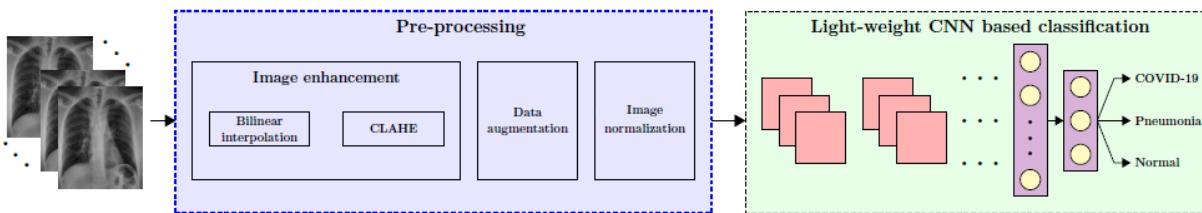
۲.۲ روش‌شناسی پیشنهادی

چارچوب خودکار پیشنهادی برای طبقه‌بندی موارد عفونت COVID-19 از تصاویر CXR و تشخیص آن‌ها از حالت‌های نرمال و پنومونی در شکل ۲ نشان داده شده است. این چارچوب از دو مرحله تشکیل شده است:

۱. پیش‌پردازش که شامل بهبود کیفیت تصویر، افزایش داده‌ها (Data Augmentation) و نرمال‌سازی تصویر است.

۲. طبقه‌بندی با استفاده از مدل سبک و کم حجم CNN پیشنهادی با نام LW-CORONet.

توضیح کامل هر مرحله در ادامه ارائه شده است.



۲.۱ پیش‌پردازش

این مرحله یکی از مهم‌ترین بخش‌های رویکرد پیشنهادی است. در این بخش، هر تکنیکی که در این مرحله استفاده شده توضیح داده می‌شود.

بهبود تصویر: بهبود کیفیت تصویر در حوزه پردازش تصاویر پزشکی به‌طور گسترده برای ارتقای کیفیت تصاویر به کار می‌رود. در این پژوهش، فرآیند بهبود تصویر با استفاده از روش درون‌یابی دوچمله‌ای (Bilinear Interpolation - BI) و پس از آن تکنیک «برابرسازی هیستوگرام تطبیقی با محدودسازی کنترast» (CLAHE) انجام شده است.

در ابتدا، روش BI برای جایگزینی پیکسل‌های از دست‌رفته به کار رفت؛ این روش با محاسبه میانگین وزنی نزدیک‌ترین پیکسل‌های مرزی، مقدار پیکسل جدید را تعیین می‌کند. سپس، تکنیک CLAHE برای افزایش کنترast تصویر استفاده شد. این روش از افزایش بیش از حد نویز جلوگیری کرده و در نتیجه به بهبود فرایند تشخیص کمک می‌کند.

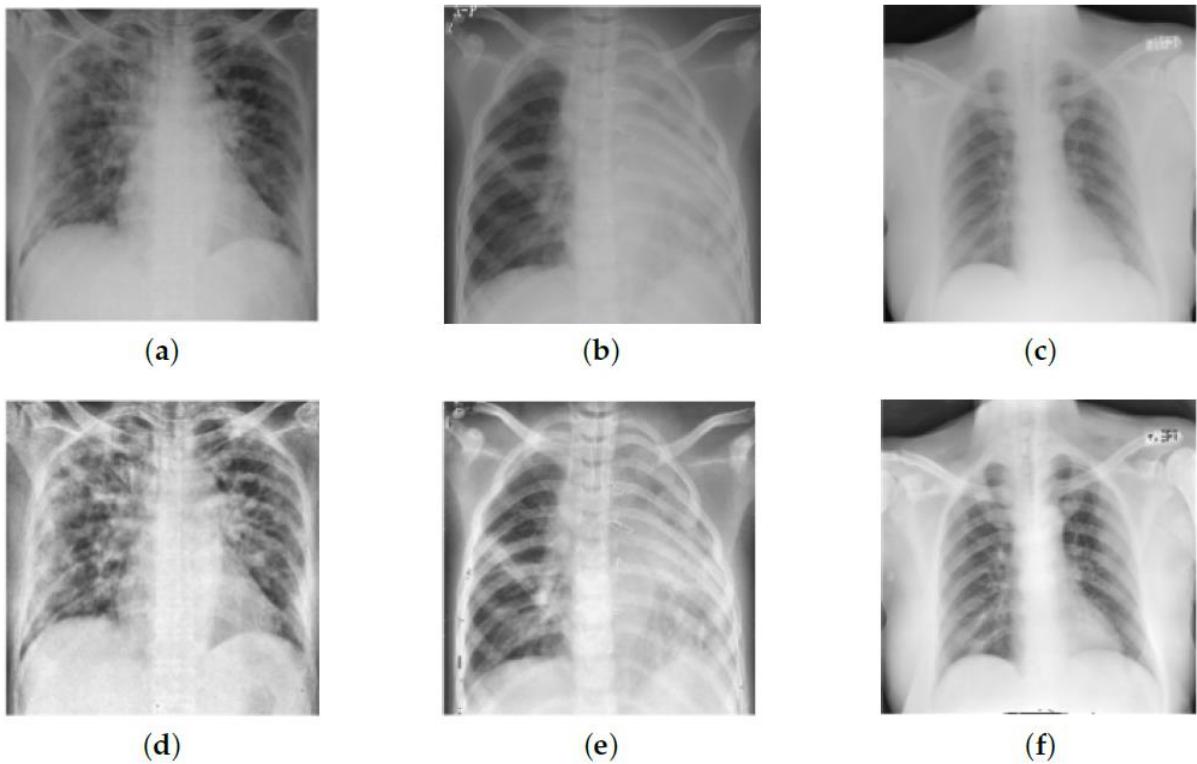


Figure 3. Results of image enhancement (BI and CLAHE): (a–c) indicate original images, (d–f) indicate enhanced images.

تقویت داده‌ها (Data Augmentation)

مدل‌های CNN زمانی عملکرد خوبی دارند که با مجموعه داده‌های بزرگ آموزش ببینند. با این حال، بیشتر مجموعه داده‌های تصاویر پزشکی از نظر حجم داده محدود هستند. بنابراین، تقویت داده‌ها عمدتاً برای حل این مشکل به کار می‌رود؛ این کار تنوع تصاویر را افزایش داده و نقش یک تنظیم‌کننده (**regularizer**) را دارد.

در این مطالعه، تقویت داده‌ها بر روی مجموعه داده آموزشی با چهار تبدیل انجام شد:

- برگرداندن افقی تصویر (Horizontal Flipping)
- چرخش تصویر به اندازه ۱۵ درجه به صورت ساعت‌گرد
- تغییر مقیاس (Scaling) به میزان ۱۰٪
- افزودن نویز گاوسی با میانگین صفر و واریانس ۰،۲۵

نرمال‌سازی تصویر (Image Normalization)

برای حفظ پایداری عددی، نرمال‌سازی در مدل‌های CNN به‌طور گسترده استفاده می‌شود. این کار همچنین باعث یادگیری سریع‌تر و افزایش ثبات روش نزول گرادیان می‌شود. در این مطالعه، هر پیکسل تصویر با ضرب در ۲۵۵/۱ نرمال‌سازی شد تا مقدار آن در بازه ۰ تا ۱ قرار گیرد.

۲،۲،۲ مدل پیشنهادی (Proposed Model)

مدل‌های یادگیری عمیق به دلیل قابلیت خودیادگیری و عملکرد برتر در مسائل دسته‌بندی دودویی و چندکلاسه، تحول بزرگی در حوزه هوش مصنوعی ایجاد کرده‌اند. با این حال، در زمینه تحلیل تصاویر پزشکی، اغلب مطالعات اخیر یا مدل‌های CNN را از صفر پیاده‌سازی کرده‌اند یا از یادگیری انتقالی (Transfer Learning) برای رسیدن به بهترین عملکرد استفاده کرده‌اند. هر دو روش معمولاً با هزینه محاسباتی و نیاز حافظه بالا همراه هستند. بنابراین، طراحی یک مدل CNN سبک‌وزن برای تشخیص عفونت COVID-19 از تصاویر CXR اهمیت بالایی دارد.

مدل پیشنهادی شامل سه بلوک CBR است که هر کدام شامل لایه‌های کانولوشن (CONV)، نرمال‌سازی (Batch Normalization - BN) و تابع فعال‌سازی ReLU است. هر بلوک CBR به یک لایه pooling متصل است و در انتهای دو لایه کاملاً متصل (FC) قرار دارند. توضیح کوتاه هر لایه به شرح زیر است:

- **لایه CONV:** این لایه با اعمال فیلترهای مختلف بر روی ورودی، نقشه‌های ویژگی (Feature Maps) تولید می‌کند. هایپرپارامترهای آن شامل تعداد فیلترها، اندازه فیلتر، padding و stride هستند.
- **لایه BN:** این لایه نرمال‌سازی دسته‌ای سرعت آموزش را افزایش داده و فعال‌سازی‌های لایه قبل را روی یک mini-batch نرمال می‌کند. همچنین به عنوان یک regularizer عمل کرده و از overfitting جلوگیری می‌کند. معمولاً بین لایه‌های CONV و ReLU قرار می‌گیرد.
- **لایه ReLU:** این لایه غیرخطی بودن مدل را ایجاد می‌کند. به دلیل کارایی محاسباتی بالا، به‌طور گسترده در مدل‌های DL استفاده می‌شود. مقادیر منفی ورودی را صفر و مقادیر مثبت را دست‌نخورده نگه می‌دارد. تابع ReLU به صورت زیر تعریف می‌شود:

$$f(v) = \max(0, v)$$

- **لایه Pooling:** این لایه اندازه فضایی نقشه ویژگی را کاهش داده و معمولاً بین لایه‌های CONV متوالی قرار می‌گیرد. این کار تعداد پارامترهای قابل یادگیری و هزینه محاسباتی را کاهش داده و به

مقابله با overfitting کمک می‌کند. انواع متداول pooling شامل average ، max pooling و sum pooling هستند که در این مطالعه از **max pooling** استفاده شده است.

- **لایه FC:** ویژگی‌های استخراج شده از لایه‌های CONV ابتدا به صورت یک بردار مسطح (flatten) درمی‌آیند و سپس به لایه‌های FC ارسال می‌شوند. این لایه‌ها در انتهای مدل قرار دارند و مشابه لایه‌های شبکه‌های عصبی سنتی عمل می‌کنند. برای مقابله با overfitting، بین لایه‌های FC از **Softmax** استفاده شده است. لایه FC آخر پس از یک لایه **dropout layer** قرار دارد که احتمال هر کلاس را برای دسته‌بندی خروجی می‌دهد.

۲،۲،۳ مدل پیشنهادی LW-CORONet

مدل پیشنهادی **LW-CORONet** در شکل ۴ نشان داده شده است. این مدل شامل سه بلوک CBR است که هر بلوک یک دنباله از لایه‌های **CONV**، **ReLU** و **BN** را شامل می‌شود و پس از آن یک لایه **max pooling** و در نهایت دو لایه **FC** متوالی قرار دارد. برای جلوگیری از overfitting ۵۰٪ **Softmax** بر روی هر یک از لایه‌های FC اعمال شده است. در انتهای، یک لایه **dropout** قرار می‌گیرد.

هدف اصلی از طراحی این مدل سبک وزن سفارشی، کاهش هزینه محاسباتی و تعداد پارامترهای قابل یادگیری است که باعث افزایش سرعت یادگیری آن در مقایسه با مدل‌های CNN پیشرفته موجود می‌شود.

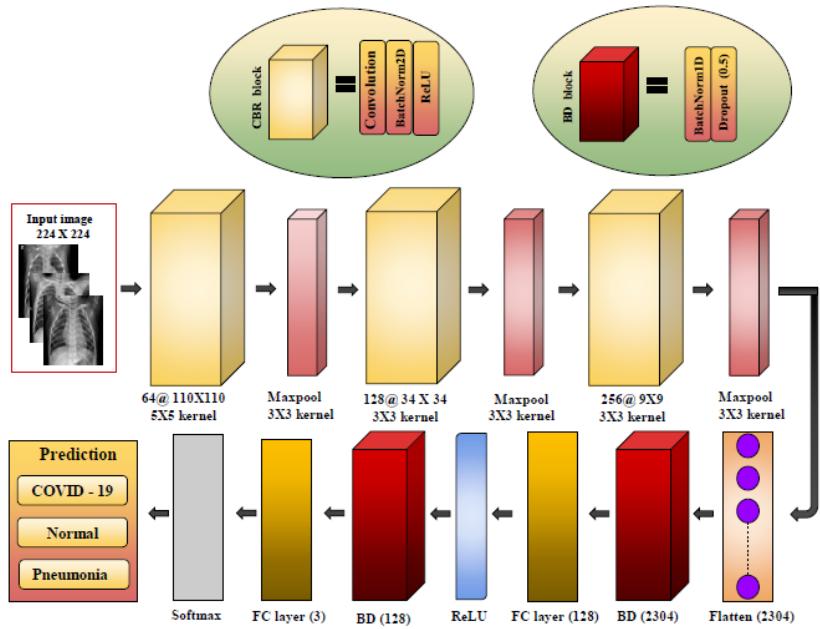


Figure 4. Illustration of proposed LW-CORONet model.

لایه اول CONV یک تصویر CXR با اندازه 224×224 دریافت می‌کند و با استفاده از ۶۴ فیلتر 5×5 و stride برابر ۲ آن را کانولوشن می‌کند، که یک حجم خروجی با اندازه $64 \times 110 \times 110$ تولید می‌کند.

پس از اولین بلوک CBR، یک عملیات max pooling با اندازه فیلتر 3×3 و stride برابر ۳ اعمال می‌شود که حجم خروجی را به $64 \times 36 \times 36$ کاهش می‌دهد.

این خروجی سپس به دومین بلوک CBR داده می‌شود، جایی که ۱۲۸ کرنل 3×3 استفاده شده و حجم خروجی $128 \times 34 \times 34$ به دست می‌آید. یک عملیات max pooling دیگر اعمال می‌شود تا حجم خروجی به $128 \times 11 \times 11$ برسد.

به همین ترتیب، در بلوک سوم $256 \times 3 \times 3$ فیلتر 3×3 اعمال می‌شوند و پس از آن یک عملیات max pooling انجام می‌شود تا حجم خروجی $256 \times 3 \times 3$ حاصل شود.

این حجم خروجی سپس مسطح (flatten) شده و به یک لایه Dense با ۱۲۸ نورون ارسال می‌شود. قبل از این لایه، لایه‌های BN و dropout قرار دارند و پس از آن نیز لایه‌های ReLU و BN قرار می‌گیرند. در نهایت، یک لایه Softmax در انتهای قرار دارد که نتایج دسته‌بندی را تولید می‌کند.

تعداد پارامترهای قابل پادگیری در هر لایه و جزئیات کامل معماری در جدول ۲ ارائه شده است.

Table 2. Detail configuration of the proposed LW-CORONet.

Layer	Activation Size	Parameters	Stride	Kernel Size	Filter
Input	(3, 224, 224)	0	-	-	-
CONV	(64, 110, 110)	4864	2	5 × 5	64
BN	(64, 110, 110)	128	-	-	-
ReLU	(64, 110, 110)	0	-	-	-
MaxPool	(64, 36, 36)	0	3	3 × 3	-
CONV	(128, 34, 34)	73,856	1	3 × 3	128
BN	(128, 34, 34)	256	-	-	-
ReLU	(128, 34, 34)	0	-	-	-
MaxPool	(128, 11, 11)	0	3	3 × 3	-
CONV	(256, 9, 9)	295,168	1	3 × 3	256
BN	(256, 9, 9)	512	-	-	-
ReLU	(256, 9, 9)	0	-	-	-

Table 2. Cont.

Layer	Activation Size	Parameters	Stride	Kernel Size	Filter
MaxPool	(256, 3, 3)	0	3	3 × 3	-
BN	2304	4068	-	-	-
Dropout (0.5)	2304	0	-	-	-
FC	128	295,040	-	-	-
ReLU	128	0	-	-	-
BN	128	256	-	-	-
Dropout (0.5)	128	0	-	-	-
Softmax	3	387	-	-	-
Total \cong 0.68 million					

۴/۲/۴ مدل‌های CNN پیش‌آموزش‌دیده با یادگیری انتقالی

در این بخش، مروری کوتاه بر چهار معماری CNN پیش‌آموزش‌دیده که در این مطالعه مورد استفاده قرار گرفته‌اند، همراه با مفهوم یادگیری انتقالی (Transfer Learning - TL) ارائه می‌شود.

با استفاده از TL، مدل‌های پیش‌آموزش‌دیده که قبلاً روی مجموعه داده‌های بزرگ مانند ImageNet آموزش دیده‌اند، برای یادگیری یک وظیفه مشابه مجددآموزش داده می‌شوند. این روش باعث می‌شود فرآیند یادگیری سریع‌تر از آموزش مدل از صفر باشد و در نبود داده‌های آموزشی گسترده نیز عملکرد مناسبی داشته باشد.

به دلیل محدود بودن تعداد نمونه‌های کووید-۱۹، تلاش‌های زیادی با استفاده از TL برای تشخیص این بیماری انجام شده است. در این مطالعه، تنها لایه نهایی مدل‌ها دوباره آموزش داده شد و وزن‌های پیش‌آموزش‌دیده دیگر دست‌نخورده باقی ماندند.

چهار معماری مورد بررسی عبارت‌اند از **VGG-19**، **ResNet-101**، **DenseNet-121** و **Xception** که اثر آن‌ها با مدل پیشنهادی **LW-CORONet** مقایسه شد. جدول ۳ نمای کلی معماری هر مدل پیش‌آموزش‌دیده همراه با طرح پیشنهادی ارائه می‌دهد.

Table 3. Architectural details of pre-trained CNN architecture.

Architecture	Layers	Input Layer Size	Output Layer Size	Parameters
ResNet-101	101	(224,224,3)	(3,1)	44.1×10^6
VGG-19	19	(224,224,3)	(3,1)	143×10^6
DenseNet-121	121	(224,224,3)	(3,1)	7.5×10^6
Xception	71	(299,299,3)	(3,1)	22.3×10^6
Proposed LW-CORONet	5	(224,224,3)	(3,1)	0.68×10^6

۳. تنظیمات آزمایش و نتایج

در این بخش، تنظیمات آزمایش و نتایج به منظور بررسی اثربخشی مدل پیشنهادی ارائه می‌شود. مجموعه‌ای از آزمایش‌ها با استفاده از دو مجموعه‌داده انجام شد:

- Dataset-1: شامل ۲۲۵۰ تصویر CXR نمای قدامی (پنومونی: ۷۵۰، کووید-۱۹: ۷۵۰ و نرمال: ۷۵۰)

- Dataset-2: شامل ۱۵,۹۹۹ تصویر CXR پنومونی: ۵۵۷۵، کووید-۱۹: ۲۳۵۸ و نرمال: ۸۰۶۶

ابتدا تمام تصاویر به اندازه 224×224 تغییر اندازه داده شدند. تمامی مدل‌های CNN با استفاده از ابزار **GPU NVIDIA Tesla** توسعه یافتند و تمام آزمایش‌ها روی پلتفرم **Google Colab** با **T4** و حافظه ۱۶ گیگابایت اجرا شدند.

عملکرد مدل پیشنهادی و معماری‌های پیش‌آموزش‌دیده با استفاده از اعتبارسنجی متقابل ۱۰ بخشی-**(10-fold cross-validation)** برای Dataset-1 ارزیابی شد، به طوری که در هر تکرار یک بخش برای تست و بقیه برای آموزش استفاده شد. از مجموعه آموزشی، ۲۰٪ نمونه‌ها به صورت تصادفی برای اعتبارسنجی انتخاب شدند. برای Dataset-2، نتایج بر اساس تقسیم‌بندی آموزش و تست ارائه شده توسط Wang و همکاران ارزیابی شد، به طوری که مجموعه تست شامل ۲۰۰ تصویر CXR از هر کلاس بود و بقیه نمونه‌ها برای آموزش مدل استفاده شدند که از آن‌ها ۲۰٪ برای اعتبارسنجی اختصاص داده شد.

مجموعه‌ای از معیارهای عملکرد شامل حساسیت، اختصاریت، دقت، صحت (**Accuracy**) و **F1-score** برای ارزیابی هر مدل استفاده شد. علاوه بر این، نتایج **Grad-CAM** با استفاده از **Heap-Map**

(Gradient-weighted Class Activation Mapping) محاسبه شد تا با برجسته‌سازی مناطق اثربخشی مدل به صورت بصری تفسیر شود.

عملکرد مدل در دو سناریو ارزیابی شد:

۱. دسته‌بندی سه کلاسه: پنومونی، کووید-۱۹ و نرمال
۲. دسته‌بندی دودویی: نرمال و کووید-۱۹

هایپرپارامترهای استفاده شده در این مطالعه به صورت تجربی تعیین شده و در جدول ۴ ارائه شده است. علاوه بر این، تحلیل مقایسه‌ای با معماری‌های پیش‌آموزش دیده شامل **DenseNet-101**, **ResNet-101**, **VGG-19** و **Xception 121** انجام شد.

Table 4. Illustration of hyperparameter setting used in this study.

Hyperparameter	Value
Batch size	32
Learning rate	0.0005
Epochs	100
Optimizer	Adam
Loss function	Categorical cross-entropy

۳،۱ نتایج مدل پیشنهادی

نتایج دسته‌بندی به دست آمده توسط مدل پیشنهادی برای هر دو مجموعه داده در این بخش ارائه می‌شود.

۳،۱،۱ نتایج روی Dataset-1

نتایج هر Fold در قالب دقت، دقت پیش‌بینی (Precision)، حساسیت (Sensitivity) و F1-score برای Dataset-1 در جدول ۵ نشان داده شده است. مدل پیشنهادی همچنین روی وظایف دسته‌بندی دودویی آزمایش شد و نتایج در جدول ۶ ارائه شده است. شایان ذکر است که نتایج ارائه شده در هر Fold، میانگین نتایج سه کلاس محاسبه شده‌اند.

- دقت میانگین به دست آمده: ۹۸,۶۷٪
- دسته‌بندی چند کلاسه

◦ دسته‌بندی دودویی: ۹۹,۰۰٪

ماتریس‌های سردرگمی (Confusion Matrices) در هر اجرای ۱۰-Fold CV برای سناریوهای چند کلاسه و دودویی به ترتیب در شکل‌های ۵ و ۶ نمایش داده شده‌اند.

شکل ۷ منحنی‌های خطای آموزش و اعتبارسنجی را نشان می‌دهد. مشاهده می‌شود که مدل پیشنهادی در طی ۱۰۰ اپوک همگرا می‌شود. نمودار ارائه شده مربوط به یک اجرای ۱۰-Fold CV است.

◦ شاخص‌های عملکرد تکمیلی:

◦ ۰,۹۷۳ Mathew Correlation Coefficient (MCC):

◦ ۰,۹۸۶ Kappa Score:

این مقادیر نشان‌دهنده نتایج پیش‌بینی بهتر برای هر کلاس هستند. نمایش تصویری Kappa و MCC همراه با دقت دسته‌بندی چند کلاسه در شکل ۸ ارائه شده است. این مقادیر در اپوک‌های مختلف ثبت شده و برای یک Fold محاسبه شده‌اند.

شکل ۹ نمودارهای ROC را برای سناریوهای دسته‌بندی دودویی به دست آمده توسط مدل پیشنهادی و مدل‌های پیش‌آموزش‌دیده نشان می‌دهد.

در نهایت، Heat Map‌ها با استفاده از Grad-CAM ارائه شدند تا قابلیت تفسیر بصری مدل بررسی شود. شکل ۱۰ نتایج Heat Map چند نمونه از تصاویر CXR نرمال و کووید-۱۹ را نشان می‌دهد. مدل پیشنهادی قادر است مناطق مشکوک را شناسایی کند که نشان‌دهنده قابلیت تفسیر بهتر نتایج دسته‌بندی است و می‌تواند به رادیولوژیست‌ها در تشخیص کمک کند. شکل ۱۰ Heat Map یک نمونه نرمال را نشان می‌دهد که در آن مناطق مشکوک مشخص نشده‌اند.

Table 5. Fold-wise multi-class classification performance (in %) on Dataset-1.

Fold	Accuracy	Precision	Sensitivity	Specificity	F1-Score	COVID-19 Class Accuracy
1	97.92	96.89	96.89	98.44	96.88	98.66
2	99.41	99.12	99.11	99.56	99.11	99.56
3	97.97	97.15	97.03	98.49	96.99	99.11
4	98.82	98.65	97.33	99.33	97.88	99.56
5	97.63	96.45	96.44	98.22	96.44	97.33
6	97.35	96.02	96.02	98.01	96.01	97.78
7	99.11	98.69	98.67	99.33	98.67	99.11
8	99.41	99.12	99.11	99.56	99.11	99.56
9	99.41	99.13	99.11	99.56	99.11	99.11
10	99.70	99.56	99.56	99.78	99.56	100.00
Average	98.67	98.08	97.93	99.03	97.98	98.98

Table 6. Fold-wise binary classification performance (in %) on Dataset-1.

Fold	Accuracy	Precision	Sensitivity	Specificity	F1-Score	COVID-19 Class Accuracy
1	99.33	100.00	98.67	100.00	99.33	98.67
2	98.67	97.40	100.00	97.33	98.68	100.00
3	99.33	98.68	100.00	98.67	99.33	100.00
4	98.67	98.67	98.67	98.67	98.67	98.67
5	100.00	100.00	100.00	100.00	100.00	100.00
6	98.00	98.65	97.33	98.67	97.99	97.33
7	99.33	98.68	100.00	98.67	99.34	100.00
8	98.67	97.40	100.00	97.33	98.68	100.00
9	99.33	100.00	98.67	100.00	99.33	98.67
10	98.67	98.67	98.67	98.67	98.67	98.67
Average	99.00	98.82	99.20	98.80	99.00	99.20

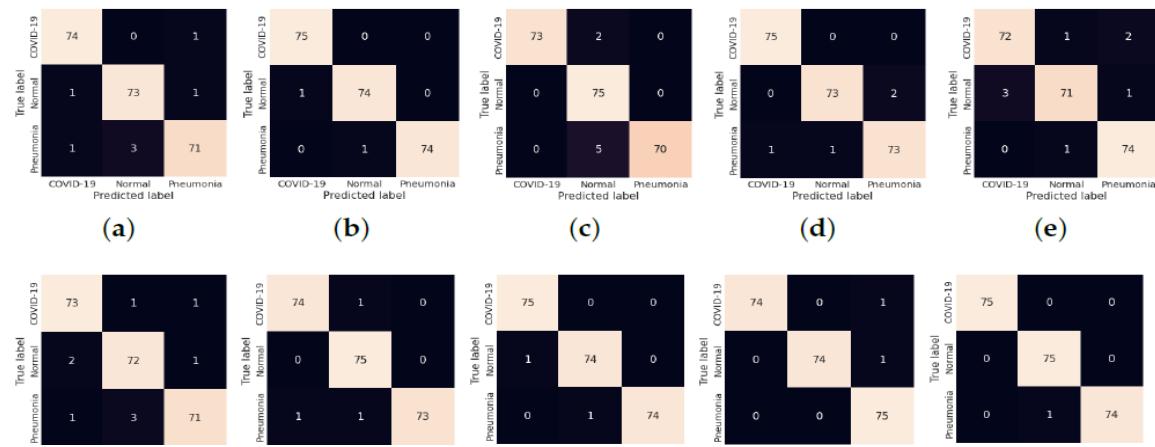


Figure 5. (a–j) Confusion matrix obtained by LW-CORONet fold-wise (fold 1 to fold 10) for multi-class classification case.

	True label COVID-19	Normal	Predicted label
Normal	74	1	0
COVID-19	0	75	75
Predicted label			

(a) (b) (c) (d) (e)

	True label COVID-19	Normal	Predicted label
Normal	75	0	2
COVID-19	2	73	73
Predicted label			

(f) (g) (h) (i) (j)

Figure 6. (a–j) Confusion matrix obtained by LW-CORONet fold-wise (fold 1 to fold 10) for binary classification case.

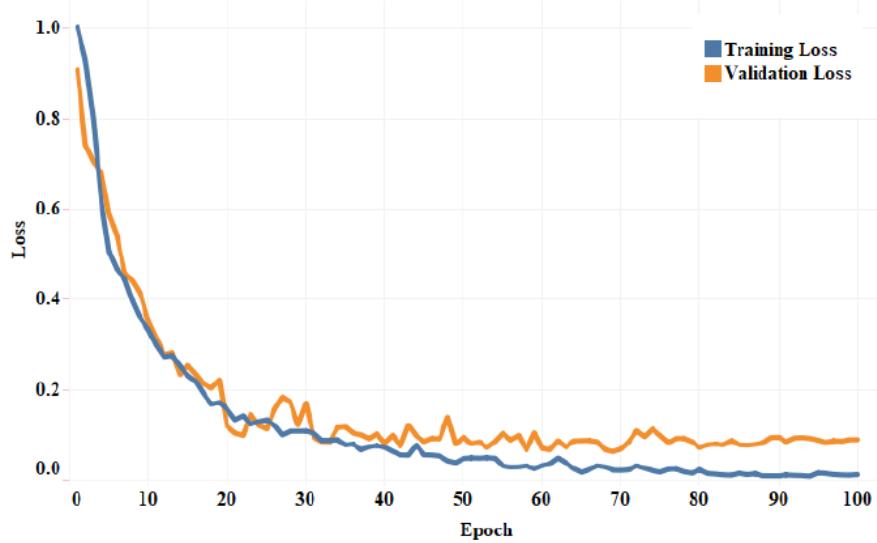


Figure 7. Illustration of the loss convergence plot for the first run of 10-fold CV.

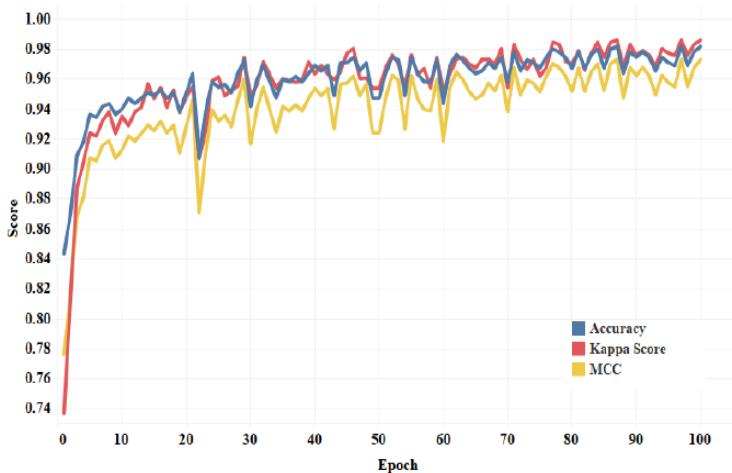


Figure 8. Graphical plot of Mathews correlation, kappa score, and accuracy for LW-CORONet model.

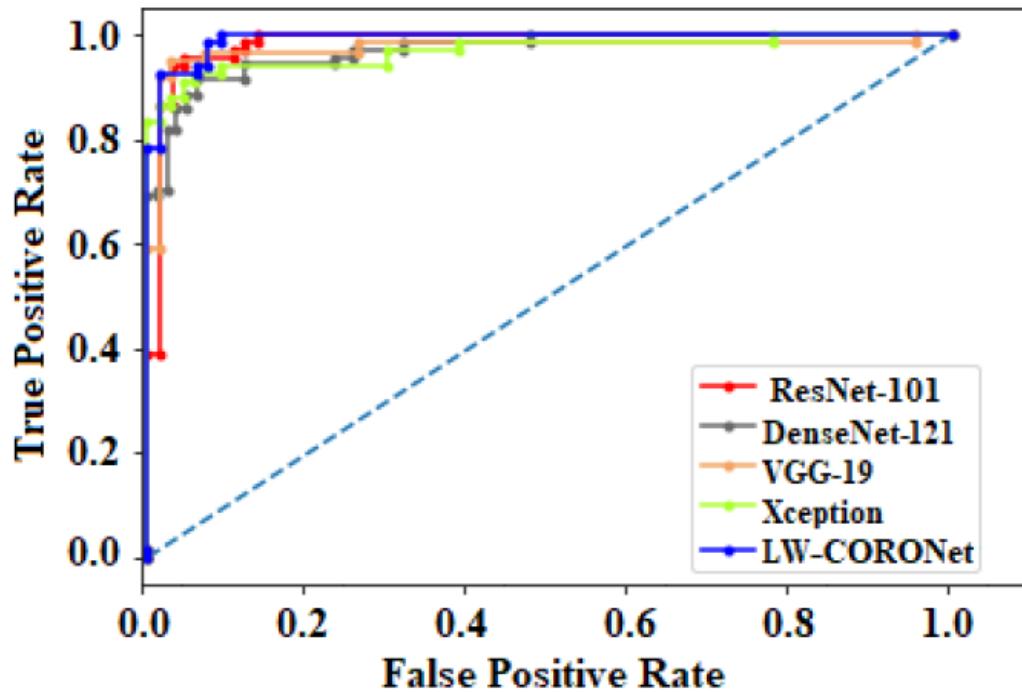


Figure 9. ROC curves of LW-CORONet and the pre-trained CNNs.

هزینه محاسباتی مدل پیشنهادی

هزینه محاسباتی مدل پیشنهادی بر اساس زمان (بر حسب ثانیه) ارزیابی شد.

- مرحله آموزش : مدل پیشنهادی حدود 8.5×10^3 ثانیه برای همگرایی نیاز دارد، که نسبت به آموزش معماری‌های پیشرفته CNN سرعت بیشتری دارد.
- مرحله تست : دسته‌بندی یک تصویر با مدل پیشنهادی تقریباً ۰.۳۵ ثانیه طول می‌کشد.

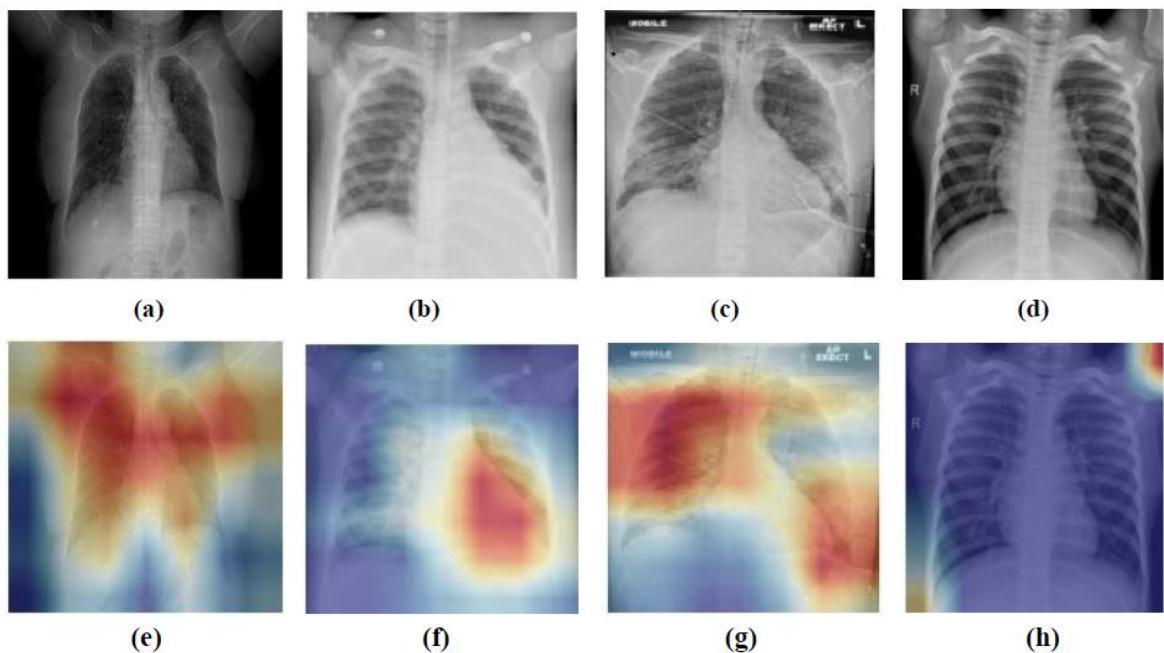


Figure 10. Grad-CAM visualization results of the proposed model on a few sample CXR images. (a–d) The top row indicates the original images. (e–h) The bottom row indicates the heat maps obtained by LW-CORONet.

۳.۱.۲ نتایج روی Dataset-2

عملکرد مدل پیشنهادی بر روی **Dataset-2** رزیابی شد.

- جدول ۷ نتایج دسته‌بندی را برای هر دو سناریوی دسته‌بندی چندکلاسه و دودویی نشان می‌دهد.
- دقت میانگین به دست آمده:
 - چندکلاسه: ۹۵,۶۷٪
 - دودویی: ۹۶,۲۵٪

شکل ۱۱ ماتریس‌های سردرگمی (Confusion Matrices) را برای هر دو سناریو نمایش می‌دهد.

Table 7. Classification performance (in %) of LW-CORONet on Dataset-2.

Task	Accuracy	Precision	Sensitivity	Specificity	F1-Score	COVID-19 Class Accuracy
Multi-class	95.67	93.51	93.50	96.75	93.50	95.00
Binary	96.25	95.57	97.00	95.50	96.28	97.00

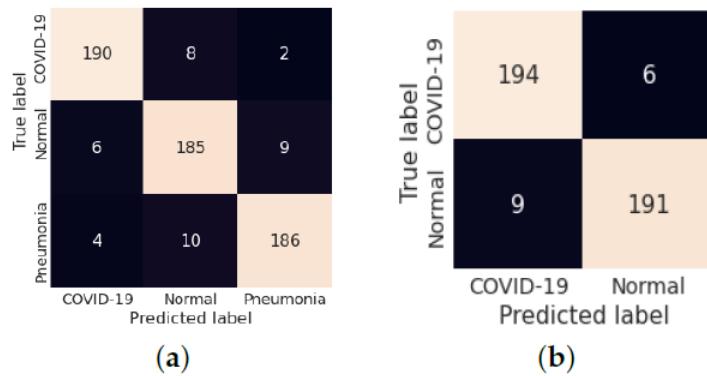


Figure 11. (a,b) Confusion matrix obtained for both cases on Dataset-2.

۳،۲ آزمایش روی هایپرپارامترهای مختلف

در این بخش، تأثیر هایپرپارامترهای مختلف شامل تعداد اپوکها، اندازه Batch، نرخ یادگیری، بهینهسازهای مختلف و غیره بررسی شده است تا بهترین عملکرد دسته‌بندی شناسایی شود. شایان ذکر است که تمام آزمایش‌ها با استفاده از **Dataset-1** انجام شدند.

۳،۲،۱ تأثیر تکنیک‌های بهینه‌سازی

در این آزمایش، چند تکنیک بهینه‌سازی مختلف بررسی شدند، از جمله:

- SGD •
- Adam •
- RMSProp •
- AdaGrad •

هدف از این آزمایش، تعیین بهترین عملکرد دسته‌بندی بود.

جدول ۸ نتایج دقیق دسته‌بندی با استفاده از بهینه‌سازهای مختلف را نشان می‌دهد. مشاهده شد که مدل **LW-CORONet** با بهینه‌ساز **Adam** عملکرد بسیار امیدوارکننده‌ای در مقایسه با سایر بهینه‌سازها دارد. بنابراین، تمام آزمایش‌های باقی‌مانده در این مطالعه با استفاده از **Adam** انجام شدند.

Table 8. Classification results (in %) obtained by various optimizers.

Optimizer	Accuracy	Precision	Sensitivity	Specificity	F1-Score
SGD	98.18	97.55	97.42	98.61	97.39
RMSProp	97.94	97.20	97.05	98.43	97.03
AdaGrad	98.55	98.01	97.91	98.89	97.90
Adam	98.67	98.08	97.93	99.03	97.98

۳.۲.۲ تأثیر نرخ یادگیری

در این آزمایش، تأثیر نرخ یادگیری‌های مختلف بررسی شد و نرخ یادگیری که کمترین خطای اعتبارسنجی داشت، به عنوان نرخ بهینه انتخاب شد.

شکل ۱۲ نموداری از نرخ یادگیری در مقابل خطای اعتبارسنجی را نشان می‌دهد. نقطه مشخص شده با نقطه قرمز در شکل، نرخ یادگیری بهینه را نشان می‌دهد که از آنجا خطا به طور قابل توجهی کاهش یافته و در نتیجه عملکرد بهتری در دسته‌بندی حاصل شده است.

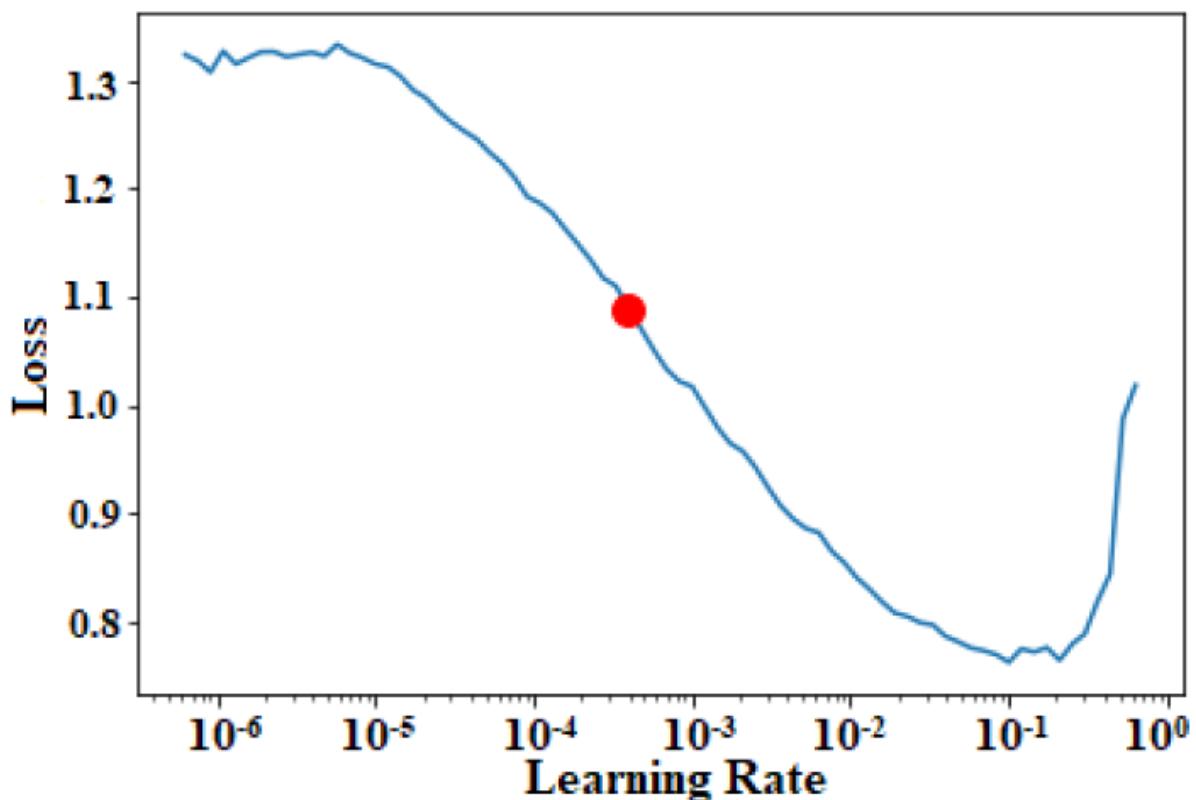


Figure 12. Plot between the learning rate and validation loss.

۳،۲،۳ تأثیر اندازه‌های مختلف Batch

در این آزمایش، اثر اندازه‌های مختلف **Batch** بررسی شد. عملکرد مدل LW-CORONet هنگام آموزش با **Batch size**‌های ۳۲، ۱۶ و ۸ در جدول ۹ ارائه شده است.

نتایج نشان داد که مدل پیشنهادی با **Batch size** برابر ۳۲، عملکرد تست پایدارتر و بالاتری ارائه می‌دهد.

Table 9. Accuracy of LW-CORONet with different batch sizes.

Model	Batch Size		
	32	16	8
Proposed LW-CORONet	98.67	98.03	97.26

۳،۳ تحلیل نتایج نادرست طبقه‌بندی

شکل ۱۳ چند نمونه از نتایج نادرست طبقه‌بندی (misclassification) مدل LW-CORONet روی Dataset-1 را نشان می‌دهد.

این خطاهای احتمالاً به دلیل ویژگی‌های بصری مشابه بین تصاویر CXR سه کلاس ایجاد شده‌اند.

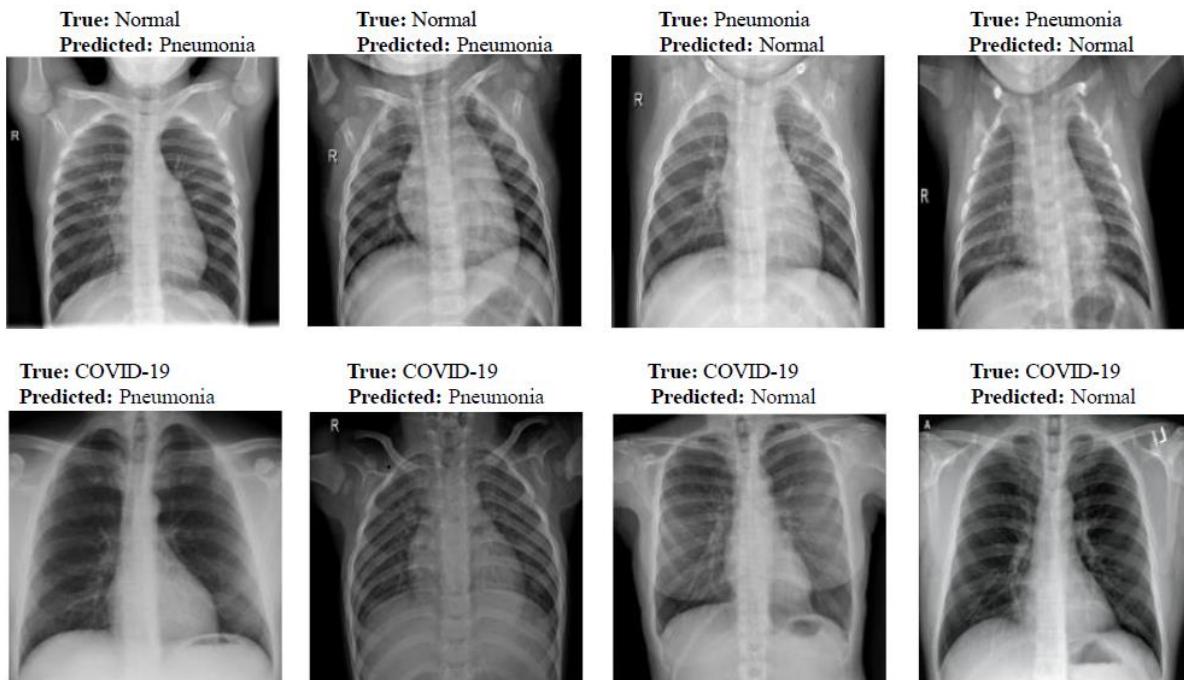


Figure 13. Sample misclassification outcomes of LW-CORONet.

۳،۴ تحلیل مقایسه‌ای معماری‌های CNN پیش‌آموزش‌دیده و مدل پیشنهادی

یک تحلیل مقایسه‌ای بین معماری‌های پیش‌آموزش‌دیده شامل **DenseNet**, **ResNet-101**, **VGG-19** و **Xception** و مدل پیشنهادی انجام شد.

- جدول ۱۰ عملکرد دسته‌بندی دقیق این مدل‌ها روی Dataset-1 را نشان می‌دهد.
- شایان ذکر است که این مدل‌ها روی همان تصاویر مدل پیشنهادی اعتبارسنجی شدند و تنظیمات آزمایش نیز یکسان بود.

نتایج نشان داد که مدل پیشنهادی تعداد پارامترها و حافظه کمتری نیاز دارد، اما عملکردی مقایسه‌پذیر یا بهتر نسبت به معماری‌های قدرتمندی مانند ResNet-101 ارائه می‌دهد.

تمام این مدل‌های پیش‌آموزش دیده در ابتدا روی مجموعه‌داده **ImageNet** طراحی شده بودند، اما پس از آن با تصاویر CXR برای انجام دسته‌بندی چندکلاسه تنظیم شدند. در این مطالعه، فقط لایه آخر این مدل‌ها مجددآموزش داده شد و سایر لایه‌ها با وزن‌های پیش‌آموزش دیده اولیه مقداردهی شدند.

Table 10. Classification performance comparison (in %) among pre-trained CNNs.

Model	Accuracy	Precision	Sensitivity	Specificity	F1-Score	Parameter	Memory
ResNet-101	98.19	97.29	97.29	98.64	97.29	44.1 M	168 MB
VGG-19	97.81	96.76	96.71	98.36	96.71	143 M	547 MB
DenseNet-121	97.45	96.16	96.29	98.10	96.20	7.5 M	31 MB
Xception	96.97	95.43	95.53	97.74	95.46	22.3 M	84 MB
LW-CORONet	98.67	98.08	97.93	99.03	97.98	0.68 M	6 MB

۳.۵ مقایسه با روش‌های موجود

مدل پیشنهادی با روش‌های مبتنی بر یادگیری عمیق موجود برای تشخیص خودکار کووید-۱۹ از تصاویر CXR مقایسه شد.

- جدول ۱۱ عملکرد به دست آمده را خلاصه می‌کند.
- مشاهده شد که مدل پیشنهادی نتایج امیدوارکننده‌ای نسبت به سایر روش‌ها روی Dataset-1 ارائه می‌دهد و حتی روی مجموعه داده بزرگ‌تر (**Dataset-2**) عملکردی مقایسه‌پذیر یا بهتر دارد.

مدل پیشنهادی با تعداد نمونه‌های CXR نسبتاً بیشتری نسبت به چندین مطالعه اخیر اعتبارسنجی شده است. اگرچه در برخی مطالعات مشابه، تعداد نمونه‌های کلاس کووید-۱۹ بسیار کم بوده است (مثلاً ۷۶ و ۱۸۰ نمونه) که باعث ایجاد مسئله عدم تعادل کلاس‌ها می‌شود.

علاوه بر این، بسیاری از مطالعات قبلی تنها روی دسته‌بندی دودویی یا چندکلاسه تمرکز داشتند، اما مطالعه حاضر هر دو سناریو را بررسی کرد و به دقت ۹۹,۰۰٪ برای دودویی و ۹۸,۶۷٪ برای چندکلاسه روی Dataset-1 رسید. مشابه آن، روی ۹۶,۲۵٪ نیز دقت ۹۵,۶۷٪ برای دودویی و به دست آمد.

Table 11. Comparison of proposed model with existing COVID-19 detection approaches.

Reference	Method	Number of CXR Samples	Accuracy (%)	
			Binary Class	Multi-Class
Hemdan et al. [18]	COVIDX-Net	50 (C ₁ : 25 and C ₂ : 25)	90.00	-
Narin et al. [19]	ResNet-50	100 (C ₁ : 50 and C ₂ : 50)	98.00	-
Ozturk et al. [17]	DarkCovidNet	1125 (C ₁ : 125, C ₂ : 500 and C ₃ : 500)	98.08	87.02
Ucar and Korkmaz [20]	Bayes-SqueezeNet	5949 (C ₁ : 76, C ₂ : 1583 and C ₃ : 4290)	-	98.30
Rahimzadeh and Attar [21]	Xception and ResNet50V2	15085 C ₁ : 180, C ₂ : 6054 and C ₃ : 8851	-	91.40
Toğuçar et al. [23]	SqueezeNet and MobileNetV2	458 (C ₁ : 295, C ₂ : 65 and C ₃ : 98)	-	98.25
Nayak et al. [32]	SVM	406	98.33	-
Toramana et al. [24]	ResNet-34	(C ₁ : 203 and C ₂ : 203)	-	-
Chaudhary and Pachori [28]	FBSED and Inception-ResNet-v2	3150 C ₁ : 1050, C ₂ : 1050 and C ₃ : 1050	97.24	84.22
Bhattacharyya et al. [30]	VGG-19 and BRISK	1446 C ₁ : 482, C ₂ : 482 and C ₃ : 482	-	93.06
Jyoti et al. [31]	ResNet-50 and MCA-TQWT	1030 C ₁ : 342, C ₂ : 341 and C ₃ : 347	-	96.60
Proposed model	LW-CORONet	5275 C ₁ : 2409 and C ₂ : 2866	94.64	-
		Dataset-1: 2250 (C ₁ : 750, C ₂ : 750 and C ₃ : 750)	99.00	98.67
		Dataset-2: 15,999 (C ₁ : 2358, C ₂ : 8066 and C ₃ : 5575)	96.25	95.67

C₁: COVID-19, C₂: Normal, C₃: Pneumonia

۳، بحث

یک مدل خودکار مبتنی بر یادگیری عمیق به نام **LW-CORONet** برای تشخیص مؤثر عفونت کروید-۱۹ با استفاده از تصاویر **CXR** ارائه شده است. اخیراً مطالعات زیادی در زمینه تشخیص کروید-۱۹ با استفاده از **DL** انجام شده است؛ اما اکثر این مطالعات محدود به مجموعه‌داده‌های کوچک بوده و نیازمند حافظه زیاد و هزینه محاسباتی بالا هستند. بنابراین، **LW-CORONet** با هدف رفع این محدودیت‌ها طراحی شد. این مدل قادر است ویژگی‌های متمایز را مستقیماً از تصاویر **CXR** بیاموزد در حالی که پارامترها و حافظه کمتری نیاز دارد.

برای بررسی اثربخشی مدل پیشنهادی، چندین آزمایش گسترده با استفاده از دو مجموعه‌داده بزرگ شامل ۲۲۵۰ و ۱۵,۹۹۹ تصویر **CXR** از سه دسته نرمال، پنومونی و کروید-۱۹ انجام شد. این مطالعه همچنین تأثیر هایپرپارامترهای مختلف شامل اندازه **Batch**، نرخ یادگیری و بهینه‌سازها را بررسی کرد. تحلیل مقایسه‌ای با چهار شبکه پیش‌آموزش‌دیده مدرن و روش‌های روز نیز انجام شد.

• دقت مدل پیشنهادی روی Dataset-1:

- چندکلاسه: ۹۸,۶۷٪
- دودویی: ۹۹,۰۰٪
- Dataset-2: دقت روی ۹۵,۶۷٪
- چندکلاسه: ۹۶,۲۵٪
- دودویی: ۹۶,۲۵٪

مدل پیشنهادی سبک، مؤثر و مناسب برای کمک به رادیولوژیست‌ها در تشخیص زودهنگام کووید-۱۹ و پنومونی است.

LW-CORONet: مزایای اصلی

- تنها دارای ۵ لایه قابل یادگیری است که یادگیری ویژگی‌های سطح بالا را از تصاویر CXR بهصورت خودکار ممکن می‌کند.
- برای هر دو سناریوی دسته‌بندی دودویی و چندکلاسه مناسب است و نیاز به استخراج ویژگی دستی ندارد.
- نسبت به سایر مدل‌های CNN، پارامترهای بسیار کمتری دارد و بنابراین هزینه محاسباتی پایینی دارد.
- سبک و کم حجم است و حافظه کمی استفاده می‌کند.

محدودیت اصلی این مطالعه: مدل پیشنهادی با داده‌های محدود کووید-۱۹ آموزش دیده است، زیرا مجموعه داده بزرگ و جامع در دسترس نبوده است.

۴. نتیجه‌گیری

در این مقاله، مدل LW-CORONet برای تشخیص زودهنگام و دقیق کووید-۱۹ از تصاویر CXR ارائه شد.

- LW-CORONet شامل ۵ لایه قابل یادگیری است و با مصرف کم حافظه و توان محاسباتی، ویژگی‌های دقیق تصاویر را استخراج می‌کند.
- برای اعتبارسنجی مدل، مجموعه‌ای گستردۀ از آزمایش‌ها روی دو مجموعه داده عمومی CXR با نمونه‌های زیاد کووید-۱۹ انجام شد.
- تأثیر هایپرپارامترهای مهم بررسی شد تا بهترین عملکرد تشخیص حاصل شود.

- مقایسه با مدل‌های پیش‌آموزش دیده و روش‌های موجود نشان داد که LW-CORONet در هر دو سناریوی چندکلاسی و دودویی عملکرد برتری دارد.
 - از نظر حافظه و هزینه محاسباتی نیز بهتر عمل می‌کند.
- در مجموع، مدل پیشنهادی سبک، مؤثر و مناسب برای تشخیص زنده کووید-۱۹ توسط پزشکان است.

پیشنهادات برای آینده:

- ارزیابی LW-CORONet روی مجموعه‌داده بزرگ‌تر و متنوع‌تر با نمونه‌های بیشتر کووید-۱۹.
- شناسایی مناطق آسیب‌دیده ریه برای هر دو حالت پنومونی و کووید-۱۹ همزمان با دسته‌بندی تصاویر CXR.