

Research Article

A CNN–LSTM Based Differential Protection Scheme for Early Detection and Classification of Internal Winding Faults in Power Transformers

Ehsan Akbari¹, *Assistant professor*

¹ Department of Electrical Engineering, Mazandaran University of Science and Technology, Babol, Iran.
e.akbari@ustmb.ac.ir

Abstract:

Power transformers are critical assets in transmission and distribution networks, and internal faults can cause substantial equipment damage and compromise power system stability. The rapid detection of winding faults and their accurate discrimination from transient phenomena remain key challenges in transformer differential protection. This paper proposes an intelligent hybrid model based on a convolutional neural network and long short-term memory network (CNN–LSTM) for the early detection and classification of internal transformer faults. Three-phase differential current signals are used as inputs to the proposed model. The CNN layers extract salient local and spatial features from the current waveforms, while the LSTM layer captures the temporal dependencies embedded in the transient signals. Various operating scenarios are simulated, including normal operating conditions, magnetizing inrush current, and different types of internal faults, namely phase-to-ground, phase-to-phase, and three-phase faults. The evaluation results demonstrate that the proposed hybrid architecture achieves an accuracy of 98.8%, outperforming standalone CNN and LSTM models. With its fast response and high accuracy in discriminating internal faults from magnetizing inrush currents, the proposed model shows strong potential for implementation in next-generation intelligent protective relays.

Keywords: CNN–LSTM, differential protection, early fault detection, internal fault, long short-term memory, power transformer.

Received: 1 Feb. 2026

Revised: 9 Apr. 2026

Accepted: 14 Apr. 2026

*** Corresponding Author:** Dr. Ehsan Akbari

Citation: Ehsan Akbari, “A CNN–LSTM Based Differential Protection Scheme for Early Detection and Classification of Internal Winding Faults in Power Transformers”, Journal of Novel Researches on Smart Power Systems, vol. 15, no. 1, pp. 1-14, June 2026 (in Persian).

مقاله پژوهشی

ارائه طرح حفاظتی دیفرانسیلی مبتنی بر مدل ترکیبی $CNN-LSTM$ برای تشخیص زودهنگام و طبقه‌بندی خطاهای داخلی سیم‌پیچی در ترانسفورماتورهای قدرت

احسان اکبری^۱، استادیار

۱- دانشکده مهندسی برق، دانشگاه علوم و فنون مازندران، بابل، ایران، e.akbari@ustmb.ac.ir

چکیده: ترانسفورماتورهای قدرت از تجهیزات حیاتی در شبکه‌های انتقال و توزیع هستند که بروز خطاهای داخلی در آنها می‌تواند منجر به خسارت‌های سنگین و کاهش پایداری شبکه شود. تشخیص سریع و تفکیک دقیق خطاهای سیم‌پیچی از پدیده‌های گذرا، چالش اصلی در حفاظت دیفرانسیل است. در این مقاله، یک مدل هوشمند ترکیبی مبتنی بر شبکه‌های عصبی کانولوشنی و حافظه طولانی‌مدت کوتاه‌مدت ($CNN-LSTM$) جهت تشخیص زودهنگام و طبقه‌بندی خطاهای داخلی ارائه شده است. در روش پیشنهادی، از سیگنال‌های جریان تفاضلی سه‌فاز به‌عنوان ورودی استفاده می‌شود؛ به‌طوری‌که لایه‌های CNN وظیفه استخراج ویژگی‌های مکانی و محلی از شکل‌موج‌ها را بر عهده داشته و لایه $LSTM$ وابستگی‌های زمانی موجود در سیگنال‌های گذرا را فرا می‌گیرد. سناریوهای مختلفی شامل وضعیت عادی، جریان هجومی و انواع خطاهای داخلی (فاز به زمین، فاز به فاز و سه‌فاز) شبیه‌سازی شده است. نتایج حاصل از ارزیابی مدل نشان می‌دهد که ساختار ترکیبی پیشنهادی با دستیابی به دقت ۹۸٫۴ درصد، عملکرد بهتری نسبت به مدل‌های تک‌ساختار CNN و $LSTM$ نشان می‌دهد. این مدل با سرعت پاسخ‌دهی بالا و دقت در تفکیک جریان هجومی، پتانسیل بالایی جهت به‌کارگیری در نسل جدید رله‌های حفاظتی هوشمند دارد.

کلمات کلیدی: ترانسفورماتور قدرت، خطای داخلی، حفاظت دیفرانسیل، شبکه عصبی کانولوشنی، $LSTM$ ، تشخیص زودهنگام خطا.

تاریخ دریافت: ۱۴۰۴/۱۱/۱۲

تاریخ بازنگری: ۱۴۰۵/۱/۲۰

تاریخ پذیرش: ۱۴۰۵/۱/۲۵

* نام نویسنده مسئول: دکتر احسان اکبری

نشانی نویسنده مسئول: بابل، دانشگاه علوم و فنون مازندران، دانشکده مهندسی برق.

۱- مقدمه

شبکه‌های ترانسفورماتورهای قدرت از مهم‌ترین تجهیزات الکتریکی در شبکه‌های انتقال و توزیع انرژی الکتریکی به شمار می‌روند و نقش اساسی در انتقال مطمئن و پایدار توان بین سطوح مختلف ولتاژ دارند [۱]. با توجه به اهمیت بالای این تجهیز، بروز هرگونه خطا در آن می‌تواند منجر به اختلال در بهره‌برداری، افزایش هزینه‌های تعمیر و نگهداری، کاهش قابلیت اطمینان شبکه و حتی وقوع خاموشی‌های گسترده شود [۲]. در میان انواع خطاهای ترانسفورماتور، خطاهای داخلی سیم‌پیچی از اهمیت ویژه‌ای برخوردارند؛ زیرا این خطاها معمولاً با سرعت بالا پیشرفت کرده و در صورت عدم تشخیص به‌موقع، می‌توانند خسارات جدی و جبران‌ناپذیری به ترانسفورماتور وارد کنند [۳].

در حال حاضر، رله دیفرانسیلیکی از اصلی‌ترین روش‌های حفاظت ترانسفورماتور در برابر خطاهای داخلی محسوب می‌شود. با وجود عملکرد مناسب این رله در شرایط عادی، تشخیص صحیح خطاهای داخلی در برخی وضعیت‌های گذرا همچون جریان هجومی مغناطیس‌کنندگی، اشباع ترانس جریان، تغییرات ناگهانی بار و خطاهای خارجی همراه با اعوجاج سیگنال، همچنان با چالش‌هایی همراه است [۴]. این شرایط می‌توانند موجب عملکرد نادرست رله یا کاهش حساسیت و پایداری آن شوند. از این‌رو، استفاده از روش‌های هوشمند و داده‌محور برای افزایش دقت و سرعت تشخیص خطا، به‌عنوان یک رویکرد مکمل در کنار حفاظت‌های سنتی، مورد توجه پژوهشگران قرار گرفته است [۵].

در سال‌های اخیر، کاربرد روش‌های یادگیری ماشین و یادگیری عمیق در حوزه حفاظت و پایش تجهیزات قدرت رشد قابل توجهی داشته است [۶]. در میان این روش‌ها، شبکه‌های عصبی کانولوشنی (CNN) به دلیل توانایی بالا در استخراج ویژگی‌های موضعی از داده‌های ورودی [۷]، و شبکه‌های حافظه کوتاه‌مدت بلندمدت (LSTM) به دلیل قدرت در مدل‌سازی وابستگی‌های زمانی [۸]، گزینه‌های مناسبی برای تحلیل سیگنال‌های گذرای الکتریکی هستند. ترکیب این دو شبکه در قالب یک ساختار هیبریدی CNN-LSTM می‌تواند هم‌زمان ویژگی‌های مکانی و زمانی سیگنال را یاد گرفته و در نتیجه، دقت تشخیص را نسبت به مدل‌های منفرد افزایش دهد [۹]. این ویژگی به‌ویژه برای تحلیل سیگنال‌های جریان دیفرانسیل ترانسفورماتور که دارای رفتارهای گذرا و غیرخطی هستند، بسیار مفید است. برخلاف برخی پژوهش‌های پیشین که بر تحلیل گازهای محلول در روغن [۱۰] یا داده‌های حرارتی [۱۱] برای تشخیص وضعیت ترانسفورماتور تمرکز داشته‌اند، در این مقاله از سیگنال‌های جریان دیفرانسیل سه‌فاز به‌عنوان ورودی مدل استفاده می‌شود تا تشخیص خطا در بازه زمانی کوتاه و با هدف کاربرد حفاظتی انجام گیرد. به این ترتیب، مسئله از پایش وضعیت عمومی به سمت تشخیص سریع و دقیق خطاهای داخلی سیم‌پیچی برای بهبود عملکرد رله دیفرانسیل سوق داده می‌شود [۱۲]. این رویکرد می‌تواند علاوه بر افزایش سرعت تشخیص، حساسیت سیستم حفاظتی را در برابر خطاهای داخلی افزایش داده و از عملکرد نادرست در شرایط گذرا جلوگیری کند [۱۳]. در این مقاله، یک مدل ترکیبی CNN-LSTM برای تشخیص زودهنگام و طبقه‌بندی خطاهای داخلی ترانسفورماتور قدرت ارائه می‌شود. در این مدل، پس از پیش‌پردازش و نرمال‌سازی داده‌ها، سیگنال‌های جریان در قالب پنجره‌های زمانی کوتاه به شبکه ورودی داده می‌شوند. سپس لایه‌های CNN ویژگی‌های مهم سیگنال را استخراج کرده و لایه‌های LSTM وابستگی‌های زمانی را برای تصمیم‌گیری نهایی مدل می‌کنند. عملکرد مدل پیشنهادی در سناریوهای مختلف شامل حالت عادی، جریان هجومی و انواع خطاهای داخلی مورد ارزیابی قرار گرفته و با مدل‌های کلاسیک ماشین بردار پشتیبان (SVM) و شبکه پرسپترون چند لایه (MLP) مقایسه می‌شود. نتایج نشان می‌دهد که ساختار هیبریدی پیشنهادی می‌تواند به‌عنوان یک ابزار هوشمند و کارآمد در بهبود عملکرد حفاظت دیفرانسیل ترانسفورماتورهای قدرت مورد استفاده قرار گیرد.

نوآوری‌های اصلی مقاله به شرح زیر است:

- ارائه یک مدل ترکیبی CNN-LSTM برای تشخیص زودهنگام خطاهای داخلی ترانسفورماتور قدرت.
- استفاده از سیگنال‌های جریان دیفرانسیل به‌عنوان ورودی مدل به‌جای داده‌های تحلیل گازهای محلول (DGA) یا حرارتی.
- طبقه‌بندی چند حالت عملکردی شامل حالت عادی، جریان هجومی و انواع خطاهای داخلی.
- مقایسه عملکرد مدل پیشنهادی با مدل‌های کلاسیک SVM و MLP و نشان دادن برتری ساختار ترکیبی.

ادامه مقاله به صورت زیر سازمان‌دهی شده است: در بخش دوم، مبانی نظری و مفاهیم مرتبط با حفاظت دیفرانسیل ترانسفورماتور و خطاهای داخلی بررسی می‌شود. در بخش سوم، مدل پیشنهادی مبتنی بر شبکه عصبی ترکیبی CNN-LSTM و مراحل پیش‌پردازش داده‌ها ارائه می‌گردد. در بخش چهارم، نتایج شبیه‌سازی و ارزیابی عملکرد روش پیشنهادی ارائه و با روش‌های دیگر مقایسه می‌شود. نهایتاً، در بخش پنجم نتیجه‌گیری و پیشنهادهایی برای تحقیقات آینده بیان می‌شود.

۲- مبانی نظری حفاظت دیفرانسیل و خطاهای داخلی ترانسفورماتور

در این حفاظت دیفرانسیل یکی از رایج‌ترین و مؤثرترین روش‌های حفاظت ترانسفورماتور در برابر خطاهای داخلی است. اساس عملکرد این حفاظت بر مقایسه جریان‌های دو طرف ترانسفورماتور در محدوده حفاظتی استوار است [۱۴]. در شرایط بهره‌برداری عادی، جریان ورودی به ترانسفورماتور و جریان خروجی از آن، پس از اعمال اصلاحات مربوط به نسبت تبدیل، گروه برداری و نسبت ترانسفورماتورهای جریان، تقریباً با یکدیگر برابر هستند. بنابراین اختلاف جریان دو طرف مقدار ناچیزی دارد. اما در صورت وقوع خطای داخلی، بخشی از جریان در محل خطا جاری شده و تعادل بین جریان‌های دو سمت ترانسفورماتور از بین می‌رود. در نتیجه، جریان دیفرانسیل افزایش یافته و می‌تواند به‌عنوان شاخصی برای تشخیص خطای داخلی مورد استفاده قرار گیرد. جریان دیفرانسیل هر فاز به‌صورت کلی از رابطه زیر محاسبه می‌شود:

$$I_{diff,k} = |I_{p,k} - I'_{s,k}| \quad (1)$$

که در آن، $I_{p,k}$ جریان سمت اولیه، $I'_{s,k}$ جریان سمت ثانویه ارجاع داده شده به سمت اولیه و k نشان‌دهنده فازهای a ، b و c است. جریان سمت ثانویه پس از در نظر گرفتن نسبت تبدیل ترانسفورماتور، نسبت ترانسفورماتورهای جریان و سایر ضرایب اصلاحی، به سمت اولیه ارجاع داده می‌شود. این جریان را می‌توان به‌صورت زیر بیان کرد:

$$I'_{s,k} = KI_{s,k} \quad (2)$$

در این رابطه، K ضریب اصلاحی برای هم‌مقیاس‌سازی جریان‌های دو سمت ترانسفورماتور است. در حالت ایده‌آل و در نبود خطای داخلی، مقدار I_{diff} بسیار کوچک خواهد بود. با این حال، در شرایط واقعی، عواملی مانند خطای اندازه‌گیری، اختلاف مشخصات ترانسفورماتورهای جریان و شرایط گذرا می‌توانند باعث ایجاد جریان دیفرانسیل غیرصفر شوند. برای افزایش پایداری حفاظت دیفرانسیل در برابر شرایط غیرخطا، معمولاً از جریان مهارکننده یا جریان بایاس استفاده می‌شود. این جریان معیاری از شدت جریان عبوری از ترانسفورماتور است و به رله کمک می‌کند تا بین خطاهای داخلی و شرایطی مانند خطاهای خارجی یا خطاهای ناشی از اشباع ترانسفورماتور جریان تمایز قائل شود [۱۵]. جریان بایاس برای هر فاز به‌صورت زیر تعریف می‌شود:

$$I_{bias,k} = \frac{|I_{p,k}| + |I'_{s,k}|}{2} \quad (3)$$

در رله‌های دیفرانسیل متداول، تصمیم‌گیری حفاظتی بر اساس مقایسه جریان دیفرانسیل با یک مقدار آستانه وابسته به جریان بایاس انجام می‌شود [۱۶]. به‌طور ساده، شرط عملکرد رله را می‌توان به‌صورت زیر نوشت:

$$I_{diff,k} > I_{set} + SI_{bias,k} \quad (4)$$

که در آن، I_{set} جریان تنظیمی اولیه رله و S شیب مشخصه دیفرانسیل است. در صورتی که جریان دیفرانسیل از مقدار مجاز تعیین‌شده بیشتر شود، رله وقوع خطای داخلی را تشخیص داده و فرمان قطع صادر می‌کند. با وجود سادگی و کارایی این روش، عملکرد آن در برخی شرایط گذرا ممکن است با چالش مواجه شود.

یکی از مهم‌ترین چالش‌های حفاظت دیفرانسیل، جریان هجومی مغناطیس‌کنندگی در زمان برق‌دار شدن ترانسفورماتور است [۱۷]. این جریان می‌تواند دامنه‌ای بزرگ داشته باشد و از نظر مقدار، شبیه جریان خطای داخلی ظاهر شود؛ در حالی که ناشی از خطا نیست و نباید موجب عملکرد رله شود. همچنین، در هنگام وقوع خطاهای خارجی شدید، ترانسفورماتورهای جریان ممکن است وارد ناحیه اشباع شوند. اشباع ترانسفورماتور جریان باعث اعوجاج در سیگنال اندازه‌گیری شده و ممکن است

جریان دیفرانسیل کاذب ایجاد کند [۱۸]. علاوه بر این، تغییرات ناگهانی بار، نویز اندازه‌گیری، عدم تطابق کامل ترانسفورماتورهای جریان و تغییرات پارامتری نیز می‌توانند تشخیص دقیق خطا را دشوار سازند [۱۹]. با توجه به این چالش‌ها، استفاده صرف از دامنه جریان دیفرانسیل همیشه برای تشخیص دقیق خطاهای داخلی کافی نیست. شکل موج جریان، نحوه تغییرات آن در زمان، الگوهای گذرا و ارتباط میان نمونه‌های متوالی سیگنال می‌توانند اطلاعات ارزشمندی درباره وضعیت ترانسفورماتور فراهم کنند. از این رو، روش‌های مبتنی بر یادگیری عمیق می‌توانند به‌عنوان ابزاری مناسب برای استخراج ویژگی‌های پنهان از سیگنال‌های جریان و افزایش دقت تشخیص خطا مورد استفاده قرار گیرند [۲۰]. در این مقاله، از سیگنال‌های جریان دیفرانسیل سه‌فاز به‌عنوان ورودی مدل پیشنهادی استفاده می‌شود تا شبکه بتواند الگوهای مربوط به حالت عادی، جریان هجومی و خطاهای داخلی را یاد بگیرد. برای تشکیل ورودی شبکه، جریان دیفرانسیل سه‌فاز در هر لحظه نمونه‌برداری به‌صورت یک بردار داده در نظر گرفته می‌شود:

$$X(n) = [I_{diff,a}(n) \quad I_{diff,b}(n) \quad I_{diff,c}(n)]^T \quad (5)$$

با در نظر گرفتن یک پنجره زمانی شامل N نمونه متوالی، ماتریس ورودی شبکه به‌صورت زیر تشکیل می‌شود:

$$X = \begin{bmatrix} x(1) \\ x(2) \\ \vdots \\ x(N) \end{bmatrix} \in \mathbb{R}^{N \times 3} \quad (6)$$

در این ماتریس، هر سطر بیانگر نمونه‌ای از جریان‌های دیفرانسیل سه‌فاز در یک لحظه زمانی و هر ستون مربوط به یکی از فازها است. استفاده از پنجره زمانی این امکان را فراهم می‌کند که مدل پیشنهادی علاوه بر مقدار لحظه‌ای جریان، روند تغییرات آن در بازه‌ای کوتاه را نیز بررسی کند. این ویژگی برای تشخیص زودهنگام خطاهای داخلی اهمیت زیادی دارد، زیرا بسیاری از خطاها در لحظات اولیه دارای الگوهای گذرای مشخصی در شکل موج جریان هستند. پیش از اعمال داده‌ها به شبکه عصبی، سیگنال‌ها نرمال‌سازی می‌شوند تا اثر اختلاف مقیاس داده‌ها کاهش یافته و فرآیند آموزش پایدارتر شود. نرمال‌سازی داده‌ها می‌تواند به‌صورت زیر انجام گیرد:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (7)$$

که در آن، X مقدار اصلی سیگنال و X_{min} و X_{max} به‌ترتیب کمینه و بیشینه مقادیر داده هستند. پس از نرمال‌سازی، داده‌ها به دو مجموعه آموزش و آزمون تقسیم شده و برای هر پنجره زمانی، برچسب متناظر با وضعیت ترانسفورماتور اختصاص داده می‌شود. این برچسب‌ها شامل حالت عادی، جریان هجومی مغناطیس‌کنندگی و انواع خطاهای داخلی مورد مطالعه هستند. در بخش بعد، ساختار مدل ترکیبی CNN-LSTM و نحوه استفاده از این داده‌ها برای تشخیص و طبقه‌بندی خطاها ارائه می‌شود.

۳- مدل پیشنهادی CNN-LSTM و روش آموزش آن

۳-۱- ساختار کلی مدل پیشنهادی

به‌منظور تشخیص سریع خطاهای داخلی سیم‌پیچی ترانسفورماتور، در این مقاله یک مدل ترکیبی مبتنی بر شبکه‌های عصبی کانولوشنی و حافظه بلندمدت کوتاه‌مدت پیشنهاد می‌شود. ایده اصلی مدل آن است که ابتدا ویژگی‌های محلی و الگوهای متمایزکننده موجود در شکل موج جریان دیفرانسیل توسط لایه‌های CNN استخراج شوند و سپس وابستگی زمانی این ویژگی‌ها در پنجره نمونه‌برداری شده توسط لایه LSTM تحلیل گردد. ترکیب این دو ساختار باعث می‌شود مدل پیشنهادی بتواند هم اعوجاج‌ها و تغییرات موضعی شکل موج و هم روند زمانی وقوع خطا را مورد بررسی قرار دهد.

ورودی مدل پیشنهادی، ماتریس جریان‌های دیفرانسیل سه‌فاز با ابعاد $N \times 3$ است که در آن N تعداد نمونه‌های موجود در پنجره زمانی و سه ستون ماتریس مربوط به فازهای a ، b و c هستند. این ساختار به شبکه اجازه می‌دهد اطلاعات سه‌فاز را

به صورت همزمان دریافت کند و ارتباط میان تغییرات جریان در فازهای مختلف را نیز در فرآیند تشخیص در نظر بگیرد. پس از دریافت داده ورودی، نمونه‌ها از مراحل استخراج ویژگی، تحلیل زمانی و طبقه‌بندی عبور می‌کنند. خروجی نهایی مدل، کلاس متناظر با وضعیت بهره‌برداری ترانسفورماتور است. شکل (۱) ساختار کلی مدل چندلایه پیشنهادی CNN-LSTM را نشان می‌دهد. در این ساختار، ماتریس جریان‌های دیفرانسیل سه فاز پس از پیش‌پردازش و نرمال‌سازی به لایه‌های کانولوشنیک بعدی اعمال می‌شود. لایه‌های CNN ویژگی‌های محلی و الگوهای متمایزکننده موجود در شکل موج جریان را استخراج می‌کنند. سپس ویژگی‌های به دست آمده به لایه‌های LSTM منتقل می‌شوند تا وابستگی‌های زمانی موجود در پنجره نمونه‌برداری تحلیل شوند. در پایان، ویژگی‌های استخراج شده توسط لایه کاملاً متصل پردازش شده و لایه Softmax وضعیت ترانسفورماتور را در یکی از کلاس‌های تعریف شده طبقه‌بندی می‌کند.

۲-۳- استخراج ویژگی با شبکه کانولوشنی

لایه کانولوشنیک بعدی برای استخراج ویژگی از سیگنال‌های ترتیبی مناسب است، زیرا می‌تواند الگوهای محلی موجود در نمونه‌های متوالی را شناسایی کند. در این مقاله، کانولوشن در راستای محور زمان انجام می‌شود و هر فیلتر، بخشی از شکل موج جریان را در یک بازه کوتاه بررسی می‌کند. خروجی فیلتر J در موقعیت زمانی n از رابطه زیر به دست می‌آید:

$$Z_j(n) = \sum_{r=0}^{K-1} \sum_{c=1}^3 w_j(r, c) x(n+r, c) + b_j \quad (8)$$

که در آن، $x(n, c)$ مقدار جریان مربوط به نمونه n در کانال c ، K طول فیلتر، $w_j(r, c)$ وزن فیلتر، و b_j بایاس مربوط به فیلتر J است. در این رابطه، تعداد کانال‌های ورودی برابر با سه بوده و متناظر با سه جریان دیفرانسیل فازهای ترانسفورماتور است. پس از انجام عملیات کانولوشن، تابع فعال‌سازی غیرخطی ReLU به کار گرفته می‌شود. این تابع به صورت زیر تعریف می‌شود:

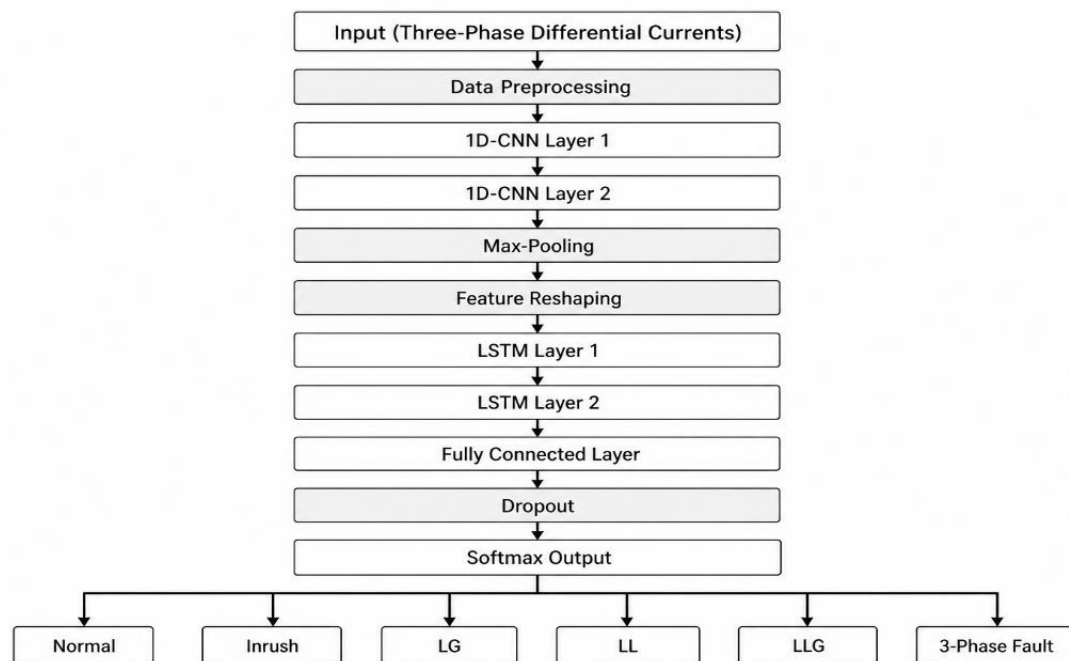
$$f(z) = \max(0, z) \quad (9)$$

استفاده از تابع ReLU موجب می‌شود شبکه بتواند روابط غیرخطی میان ویژگی‌های جریان را یاد بگیرد و در عین حال، از پیچیدگی محاسباتی بالای توابع فعال‌سازی اشباع‌شونده جلوگیری شود. خروجی لایه فعال‌سازی را می‌توان به صورت زیر نمایش داد:

$$h_{CNN}(n) = \text{ReLU}(W_{CNN} * X(n) + b_{CNN}) \quad (10)$$

که در آن، $x(n)$ بیانگر داده‌های ورودی شامل نمونه‌های جریان تفاضلی در گام زمانی n ، W_{CNN} مجموعه کرنل‌های قابل آموزش لایه کانولوشنی و b_{CNN} بردار بایاس متناظر است. همچنین، نماد $*$ عملیات کانولوشن را نشان می‌دهد و $h_{CNN}(n)$ نگاشت ویژگی استخراج شده پس از اعمال تابع فعال‌سازی ReLU است. تابع ReLU به صورت $\text{ReLU}(z) = \max(0, z)$ تعریف شده و به صورت درایه‌به‌درایه اعمال می‌شود. فیلترهای ابتدایی معمولاً تغییرات سریع، قله‌های جریان، اعوجاج‌ها و ناپیوستگی‌های موجود در شکل موج را شناسایی می‌کنند. در لایه‌های عمیق‌تر، این ویژگی‌های اولیه با یکدیگر ترکیب شده و الگوهای پیچیده‌تری مانند تفاوت میان جریان هجومی و جریان خطای داخلی استخراج می‌شوند. برای کاهش ابعاد داده و افزایش مقاومت مدل در برابر تغییرات جزئی، پس از لایه‌های کانولوشن از عملیات Max-Pooling استفاده می‌گردد. در این عملیات، بیشترین مقدار موجود در یک ناحیه مشخص از ویژگی‌ها انتخاب می‌شود:

$$p_j(m) = \max_{n \in R_m} h_j(n) \quad (11)$$



شکل (۱): ساختار کلی مدل چند لایه CNN-LSTM برای تشخیص خطاهای داخلی ترانسفورماتور

Figure (1): General structure of CNN-LSTM multilayer model for transformer internal fault detection

که در آن، $\mathcal{R}_m$ ناحیه مربوط به پنجره pooling و $p_j(m)$ خروجی کاهش یافته ویژگی است. این عملیات ضمن کاهش حجم محاسبات، ویژگی‌های مهم‌تر را حفظ کرده و حساسیت شبکه را نسبت به جابه‌جایی‌های جزئی در محل وقوع الگو کاهش می‌دهد.

۳-۳- تحلیل وابستگی‌های زمانی با لایه LSTM

اگرچه لایه‌های CNN در استخراج ویژگی‌های محلی شکل موج عملکرد مناسبی دارند، اما به‌تنهایی نمی‌توانند وابستگی‌های زمانی بلندمدت میان نمونه‌های متوالی را به‌طور کامل مدل‌سازی کنند [۲۱]. برای رفع این محدودیت، خروجی توالی لایه‌های کانولوشنی به یک لایه LSTM داده می‌شود. شبکه LSTM نوعی شبکه عصبی بازگشتی است که با استفاده از ساختار حافظه داخلی خود می‌تواند اطلاعات مهم مربوط به نمونه‌های قبلی را حفظ کرده و اثر اطلاعات کم اهمیت یا نامرتبط را کاهش دهد [۲۲]. در گام زمانی n ، روابط اصلی لایه LSTM به‌صورت زیر تعریف می‌شوند:

$$i_n = \sigma(W_i u_n + U_i h_{n-1} + b_i) \quad (12)$$

$$f_n = \sigma(W_f u_n + U_f h_{n-1} + b_f) \quad (13)$$

$$o_n = \sigma(W_o u_n + U_o h_{n-1} + b_o) \quad (14)$$

$$\tilde{c}_n = \tanh(W_c u_n + U_c h_{n-1} + b_c) \quad (15)$$

$$c_n = f_n \odot c_{n-1} + i_n \odot \tilde{c}_n \quad (16)$$

$$h_n = o_n \odot \tanh(c_n) \quad (17)$$

در این روابط، u_n ورودی لایه LSTM در گام زمانی n ، h_n بردار حالت پنهان، c_n بردار وضعیت سلول و f_n ، i_n و o_n به‌ترتیب درگاه ورودی، درگاه فراموشی و درگاه خروجی هستند. همچنین، σ تابع سیگموید، $\tanh$ تابع تانژانت هایپربولیک و $\odot$ عملگر ضرب عضو به عضو است.

درگاه فراموشی تعیین می‌کند که چه مقدار از اطلاعات مربوط به وضعیت قبلی حفظ شود. درگاه ورودی اطلاعات جدید و مهم را به حافظه اضافه می‌کند و درگاه خروجی نیز مشخص می‌سازد که چه بخشی از وضعیت داخلی سلول در خروجی هر گام زمانی ظاهر شود. این ویژگی باعث می‌شود شبکه LSTM بتواند تفاوت میان تغییرات گذرای جریان هجومی و روند تغییرات جریان ناشی از خطای داخلی را بهتر شناسایی کند.

۳-۴- طبقه‌بندی وضعیت ترانسفورماتور

پس از پردازش توالی ویژگی‌ها توسط لایه LSTM، آخرین بردار حالت پنهان یا خروجی تجمیع‌شده آن به یک لایه کاملاً متصل منتقل می‌شود. این لایه، ویژگی‌های استخراج شده را ترکیب کرده و نمایش مناسبتی برای تصمیم‌گیری نهایی ایجاد می‌کند. برای جلوگیری از بیش‌برازش، در این بخش از روش Dropout استفاده می‌گردد. در روش Dropout، بخشی از نورون‌ها به صورت تصادفی در طول فرآیند آموزش غیرفعال می‌شوند و شبکه به جای وابستگی به تعداد محدودی از ویژگی‌ها، نمایش عمومی‌تری از داده‌ها یاد می‌گیرد. به فرض تعداد کلاس‌های مورد بررسی برابر با C باشد. خروجی لایه Softmax برای کلاس q از رابطه زیر محاسبه می‌شود:

$$\hat{y}_q = \frac{\exp(s_q)}{\sum_{r=1}^C \exp(s_r)} \quad (18)$$

که در آن، s_q مقدار خام خروجی مربوط به کلاس q و $\hat{y}_q$ احتمال پیش‌بینی‌شده برای آن کلاس است. مجموع احتمال‌های خروجی برابر با یک خواهد بود و کلاس نهایی با بیشترین احتمال به عنوان وضعیت تشخیص داده شده انتخاب می‌شود:

$$\hat{C} = \arg \max_q \hat{y}_q \quad q \in \{1, \dots, C\} \quad (19)$$

در این مقاله، کلاس‌های خروجی شامل وضعیت عادی، جریان هجومی، خطای فاز به زمین، خطای فاز به فاز، خطای دو فاز به زمین و خطای سه فاز هستند.

۳-۵- فرآیند آموزش مدل

آموزش مدل پیشنهادی با استفاده از پنجره‌های زمانی برچسب‌گذاری شده انجام می‌شود. هر نمونه آموزشی شامل ماتریس جریان‌های دیفرانسیل سه فاز و برچسب وضعیت متناظر با آن است. پس از نرمال‌سازی، داده‌ها به مجموعه‌های آموزش و اعتبارسنجی تقسیم می‌شوند. مجموعه آموزش برای تنظیم وزن‌ها، و مجموعه اعتبارسنجی برای بررسی عملکرد مدل در طول آموزش، انتخاب پارامترهای مناسب، و ارزیابی نهایی مدل مورد استفاده قرار می‌گیرد. برای یک مسئله چندکلاسه، تابع هزینه آنتروپی متقاطع به صورت زیر تعریف می‌شود:

$$\mathfrak{J} = -\frac{1}{M} \sum_{m=1}^M \sum_{q=1}^C y_{m,q} \log(\hat{y}_{m,q}) \quad (20)$$

که در آن، M تعداد نمونه‌های آموزشی، C تعداد کلاس‌ها، $y_{m,q}$ مقدار واقعی برچسب نمونه m برای کلاس q ، و $\hat{y}_{m,q}$ احتمال پیش‌بینی‌شده توسط مدل است. اگر نمونه m به کلاس q تعلق داشته باشد، مقدار $y_{m,q}$ برابر یک و در غیر این صورت برابر صفر خواهد بود.

وزن‌های شبکه با استفاده از الگوریتم بهینه‌سازی Adam به روزرسانی می‌شوند [۲۳]. این الگوریتم با استفاده از گرادینت تابع هزینه، نرخ یادگیری تطبیقی برای پارامترهای مختلف شبکه تعیین می‌کند. به روزرسانی کلی پارامترهای مدل را می‌توان به صورت زیر نمایش داد:

$$\theta_{t+1} = \theta_t - \eta \nabla_{\theta} \mathfrak{J}(\theta_t) \quad (21)$$

که در آن، مجموعه پارامترهای قابل آموزش شبکه، η نرخ یادگیری و $\nabla_{\theta} \mathcal{L}$ گرادیان تابع هزینه نسبت به پارامترهای مدل است. در فرآیند آموزش، داده ها در دسته های کوچک به شبکه ارائه می شوند و در پایان هر دوره، عملکرد مدل با استفاده از داده های اعتبارسنجی بررسی می گردد [۲۵].

به منظور جلوگیری از بیش برآزش، آموزش شبکه تا زمانی ادامه می یابد که خطای اعتبارسنجی کاهش یابد. در صورتی که خطای اعتبارسنجی در چند دوره متوالی افزایش پیدا کند، فرآیند آموزش متوقف شده و وزن های مربوط به بهترین عملکرد اعتبارسنجی انتخاب می شوند. همچنین، استفاده از Dropout، نرمال سازی داده ها و انتخاب مناسب طول پنجره زمانی می تواند به افزایش قابلیت تعمیم مدل کمک کند.

در نهایت، مدل آموزش دیده با استفاده از داده های آزمون ارزیابی می شود. برای سنجش عملکرد مدل، معیارهای دقت، صحت، فراخوانی، امتیاز F_1 و ماتریس درهم ریختگی مورد استفاده قرار می گیرند. دقت مدل از رابطه زیر محاسبه می شود:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (22)$$

که در آن، TP تعداد تشخیص های صحیح برای کلاس مثبت، TN تعداد تشخیص های صحیح برای کلاس منفی، FP تعداد موارد مثبت کاذب و FN تعداد موارد منفی کاذب است. همچنین، صحت و فراخوانی به ترتیب از روابط (۲۳) و (۲۴) به دست می آیند:

$$Precision = \frac{TP}{TP + FP} \quad (23)$$

$$Recall = \frac{TP}{TP + FN} \quad (24)$$

امتیاز F_1 نیز به صورت زیر تعریف می شود:

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (25)$$

این شاخص ها امکان ارزیابی جامع عملکرد مدل را فراهم می کنند، زیرا در مسئله تشخیص خطا، صرفاً دقت کلی کافی نیست و باید توانایی مدل در شناسایی صحیح کلاس های مختلف نیز مورد توجه قرار گیرد. به ویژه در مواردی که داده ها نامتوازن باشند یا برخی خطاها فراوانی کمتری داشته باشند، استفاده از شاخص های صحت و فراخوانی اهمیت بیشتری پیدا می کند.

۴- شبیه سازی و نتایج عددی

سیستم تحت مطالعه در محیط نرم افزاری PSCADTM/EMTDC مدل سازی شده تا داده های جریان دیفرانسیل با دقت بالا استخراج گردند. همچنین به منظور پیاده سازی الگوی تشخیص خطای پیشنهادی، از نرم افزار MATLAB 2020b استفاده شده است. ترانسفورماتور قدرت مورد مطالعه دارای ظرفیت و ولتاژ نامی مطابق جدول (۱) بوده و مدل سازی آن شامل جزئیات دقیق سیم پیچی، هسته مغناطیسی و ترانسفورماتورهای جریان CT جهت در نظر گرفتن پدیده های غیرخطی نظیر اشباع است [۲۴]. سناریوهای مختلف شامل شرایط عادی بهره برداری، وقوع انواع خطاهای داخلی (فاز به زمین، فاز به فاز و سه فاز) و شرایط گذرای ناشی از جریان هجومی مغناطیس کنندگی تولید شده اند. فرکانس نمونه برداری سیگنال ها برابر با ۱۶ کیلوهرتز انتخاب شده و پنجره های زمانی به طول یک سیکل از جریان دیفرانسیل به عنوان ورودی مدل استخراج گردید. داده های تولید شده پس از نرمال سازی در بازه $[0, 1]$ ، به سه دو داده های آموزش و اعتبارسنجی تقسیم شدند تا از بیش برآزش مدل جلوگیری شود. معماری CNN-LSTM با هدف استخراج ویژگی های سلسله مراتبی از سیگنال جریان طراحی شده است. بخش CNN

شامل لایه‌های کانولوشن یک‌بعدی است که مسئولیت شناسایی الگوهای محلی شکل موج را بر عهده دارند. پس از لایه‌های کانولوشن و لایه‌های تجمیع (Pooling)، داده‌ها به لایه‌های LSTM منتقل می‌شوند تا روابط زمانی میان نمونه‌های متوالی تحلیل گردد. نرخ یادگیری برابر با $\alpha=10^{-3}$ ، تعداد دوره‌های آموزش (Epochs) و اندازه دسته (Batch Size) به گونه‌ای بهینه شدند که تابع هزینه (Categorical Cross-Entropy) به حداقل مقدار خود برسد. تنظیمات دقیق هاپر پارامترها در جدول (۲) ارائه شده است.

Table (1): Power transformer technical specifications and simulation settings

جدول (۱): مشخصات فنی ترانسفورماتور قدرت و تنظیمات شبیه‌سازی

پارامتر	مقدار
توان نامی (MVA)	۱۰۰
ولتاژ نامی (فشارقوی / فشارضعیف) (kV)	۶۳ / ۲۳۰
فرکانس شبکه (Hz)	۵۰
گروه برداری	YNd11
نسبت تبدیل ترانسفورماتور جریان (CT)	۵ / ۴۰۰
فرکانس نمونه‌برداری (kHz)	۱۶
طول پنجره زمانی (ms)	۲۰

Table (2): Hyperparametric parameters of the proposed model

جدول (۲): پارامترهای هاپر پارامتریک مدل پیشنهادی

پارامتر	مقدار
تعداد لایه‌های CNN	۲ لایه (۳۲ و ۶۴ فیلتر)
تعداد نورون‌های LSTM	۶۴ نورون
نرخ یادگیری (Learning Rate)	10^{-3}
تعداد دوره‌های آموزش (Epochs)	۵۰
اندازه دسته (Batch Size)	۳۲
طول پنجره زمانی	۲۰ میلی ثانیه

در این مطالعه، برای هر حالت بهره‌برداری، یک پنجره زمانی کوتاه از سیگنال‌های جریان انتخاب شد تا امکان تشخیص سریع خطا فراهم شود. استفاده از پنجره‌های کوتاه زمانی از این جهت اهمیت دارد که در کاربردهای حفاظتی، تصمیم‌گیری باید در کمترین زمان ممکن انجام شود و هرگونه تأخیر می‌تواند منجر به آسیب بیشتر به تجهیزات شود. بنابراین، ساختار داده به گونه‌ای طراحی شد که مدل بتواند تنها بر اساس چند سیکل ابتدایی پس از وقوع رویداد، نوع وضعیت ترانسفورماتور را شناسایی کند.

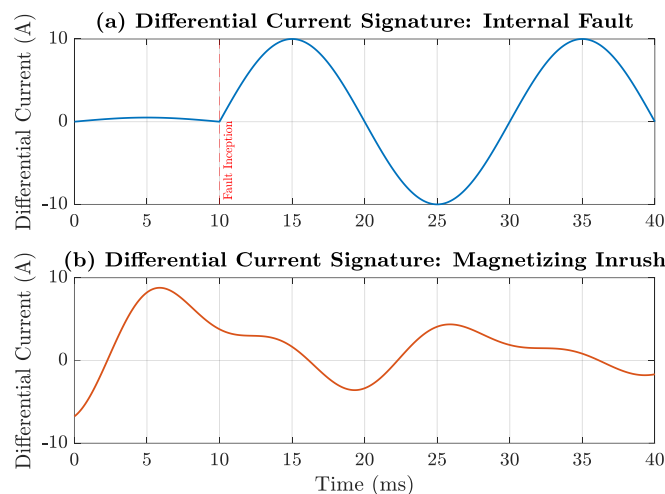
نتایج عددی حاصل از ارزیابی مدل در جدول (۳) خلاصه شده است. شاخص‌های اصلی ارزیابی شامل دقت (Accuracy)، صحت (Precision)، فراخوانی (Recall) و امتیاز F_1 هستند. طبق نتایج به دست آمده، مدل پیشنهادی موفق شده است خطاهای داخلی را با دقت بسیار بالایی از سایر شرایط متمایز نماید.

برای مقایسه عملکرد مدل، نتایج آن با روش‌های مرسوم مبتنی بر آستانه‌گذاری جریان دیفرانسیل و همچنین روش‌های یادگیری ماشین کلاسیک مانند شبکه عصبی پرسپترون چندلایه و ماشین بردار پشتیبان مقایسه شده است. مقایسه‌ها نشان می‌دهند که مدل پیشنهادی نه تنها از نظر دقت کلی برتری دارد، بلکه به دلیل ساختار ترکیبی خود، نسبت به نویز و اعوجاج‌های ناشی از اشباع CT نیز مقاوم‌تر عمل می‌کند.

شکل (۲)، نمونه‌ای از سیگنال‌های جریان تفاضلی ورودی به مدل پیشنهادی را در دو حالت خطای داخلی و جریان هجومی نشان می‌دهد. همان‌طور که ملاحظه می‌شود، جریان خطای داخلی (a) با افزایش ناگهانی دامنه و رفتار سینوسی متقارن همراه است، در حالی که جریان هجومی (b) دارای ماهیت غیر سینوسی، مؤلفه DC و میراث‌وندگی نمایشی است. این تفاوت‌های ظریف

در الگوهای زمانی، زیربنای استخراج ویژگی توسط لایه‌های کانولوشن CNN و تحلیل وابستگی‌های زمانی توسط لایه‌های LSTM را تشکیل می‌دهند تا تمایز دقیق بین این دو پدیده میسر شود.

شکل (۳) ماتریس درهم‌ریختگی مدل پیشنهادی CNN-LSTM را در طبقه‌بندی پنج وضعیت عملیاتی مختلف ترانسفورماتور نشان می‌دهد. همان‌گونه که در این ماتریس مشاهده می‌شود، مدل به دقت کلی ۹۸٫۴٪ بر روی داده‌های آزمون دست یافته است که نشان‌دهنده توانمندی بالای ساختار ترکیب‌در استخراج ویژگی‌های فضایی و تحلیل وابستگی‌های زمانی سیگنال‌های جریان تفاضلی است. بیشترین دقت تفکیک (۹۹٪) در کلاس‌های «عادی» و «خطای سه‌فاز» مشاهده می‌شود. در کلاس «جریان هجومی»، نرخ تشخیص صحیح ۹۷٪ ثبت شده است؛ تحلیل خطاها نشان می‌دهد که تعداد اندکی از نمونه‌های هجومی به دلیل شباهت شکلی در لحظات اولیه با خطاهای داخلی اشتباه شده‌اند، با این حال، میزان این خطا در مقایسه با کل نمونه‌ها بسیار ناچیز بوده و تأثیری بر پایداری سیستم حفاظت ترانسفورماتور نخواهد داشت.



شکل (۲): نمونه‌ای از سیگنال‌های جریان تفاضلی ورودی به مدل پیشنهادی در دو حالت (a) خطای داخلی و (b) جریان هجومی
Figure (2): Example of differential current signals input to the proposed model in two cases (a) internal fault and (b) inrush current

CNN-LSTM Model Confusion Matrix								
True Class	Three-Phase	99				1	99.0%	1.0%
	Normal	1	99				99.0%	1.0%
	Phase-to-Ground			99	1		99.0%	1.0%
	Phase-to-Phase	1	1		98		98.0%	2.0%
	Inrush	1		1	1	97	97.0%	3.0%
		97.1%	99.0%	99.0%	98.0%	99.0%		
		2.9%	1.0%	1.0%	2.0%	1.0%		
		Three-Phase	Normal	Phase-to-Ground	Phase-to-Phase	Inrush		
		Predicted Class						

شکل (۳): ماتریس درهم‌ریختگی حاصل از پیاده‌سازی مدل بر روی داده‌های آزمون
Figure (3): Confusion matrix resulting from implementing the model on the test data

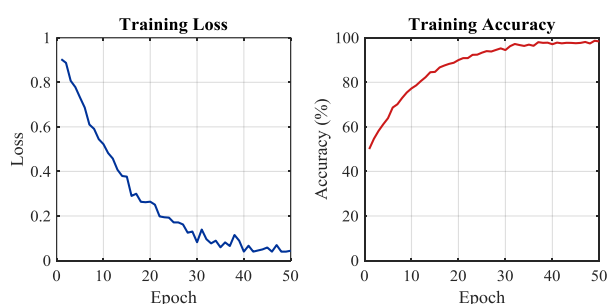
شکل (۴) روند همگرایی مدل پیشنهادی CNN-LSTM را در طول فرآیند آموزش نشان می‌دهد. همان‌گونه که مشاهده می‌شود، مقدار تابع هزینه در ابتدای آموزش در سطح بالاتری قرار داشته و با افزایش تعداد دوره‌ها به تدریج کاهش یافته است.

Table (3): Comparison of performance indicators of the proposed model with conventional methods

روش	دقت (% Accuracy)	صحت (% Precision)	فراخوانی (% Recall)	امتیاز (% F1)
آستانه‌گذاری سنتی	۸۷,۵	۸۶,۲	۸۵,۰	۸۵,۶
ماشین بردار پشتیبان (SVM)	۹۲,۸	۹۱,۵	۹۲,۰	۹۱,۷
پرسپترون چندلایه (MLP)	۹۴,۲	۹۳,۰	۹۳,۵	۹۳,۲
روش پیشنهادی (CNN-LSTM)	۹۸,۸	۹۸,۵	۹۸,۹	۹۸,۷

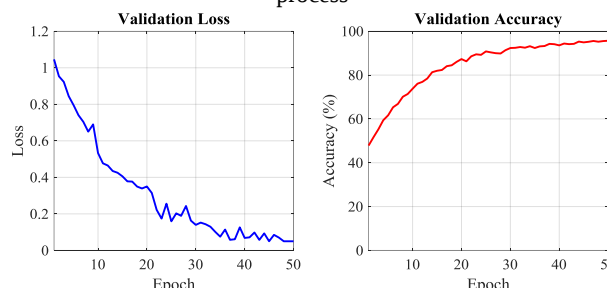
هم‌زمان، دقت مدل نیز روندی صعودی داشته و پس از چند دوره ابتدایی به مقادیر بالاتری رسیده است. این رفتار نشان می‌دهد که شبکه توانسته الگوهای مؤثر موجود در سیگنال‌های جریان تفاضلی را به‌خوبی فراگرفته و در مسیر آموزش به یک پاسخ پایدار و همگرا دست یابد.

شکل ۵ عملکرد مدل را بر روی داده‌های اعتبارسنجی نمایش می‌دهد و برای بررسی قابلیت تعمیم آن به کار می‌رود. مطابق نتایج این شکل، کاهش تابع هزینه و افزایش دقت در مجموعه اعتبارسنجی نیز با روندی مطلوب و نزدیک به نتایج آموزش مشاهده می‌شود. نزدیکی منحنی‌های آموزش و اعتبارسنجی بیانگر آن است که مدل پیشنهادی دچار بیش‌برازش قابل توجه نشده و توانایی مناسبی در تعمیم‌پذیری به داده‌های دیده‌نشده دارد. این موضوع، کارایی ساختار ترکیبی CNN-LSTM را در یادگیری ویژگی‌های متمایزکننده سیگنال‌های جریان تفاضلی تأیید می‌کند.



شکل (۴): روند تغییرات تابع هزینه و دقت مدل پیشنهادی CNN-LSTM در طول فرآیند آموزش

Figure (4): Trend of changes in the cost function and accuracy of the proposed CNN-LSTM model during the training process



شکل (۵): روند تغییرات تابع هزینه و دقت مدل پیشنهادی CNN-LSTM در طول فرآیند اعتبارسنجی

Figure (5): Trend of changes in the cost function and accuracy of the proposed CNN-LSTM model during the validation process

نتایج به‌دست‌آمده نشان می‌دهند که مدل پیشنهادی قادر است الگوهای گذرای مربوط به خطاهای داخلی را با دقت مناسبی از جریان هجومی و حالت عادی تفکیک کند. این موضوع ناشی از آن است که لایه‌های CNN ویژگی‌های محلی و ساختارهای لحظه‌ای موجود در شکل موج را استخراج می‌کنند و لایه LSTM نیز وابستگی زمانی بین این ویژگی‌ها را در طول پنجره نمونه‌برداری تحلیل می‌کند. به همین دلیل، مدل پیشنهادی تنها به مقدار لحظه‌ای جریان متکی نیست، بلکه روند تغییرات آن

در زمان را نیز در تصمیم‌گیری لحاظ می‌کند. این ویژگی موجب افزایش قابلیت تشخیص در شرایطی می‌شود که شکل موج خطا شباهت زیادی به سایر وضعیت‌های گذرا دارد.

۵- نتیجه‌گیری

در این پژوهش، یک مدل ترکیبی مبتنی بر شبکه‌های عصبی کانولوشنی یک‌بعدی و حافظه کوتاه مدت بلند مدت (CNN-LSTM) به منظور تشخیص زودهنگام خطاهای داخلی سیم‌پیچ ترانسفورماتور قدرت ارائه شد. در ساختار پیشنهادی، لایه‌های CNN با استخراج ویژگی‌های محلی و متمایزکننده از سیگنال جریان تفاضلی، اطلاعات مؤثر مربوط به شکل موج را شناسایی کردند و سپس لایه‌های LSTM با مدل‌سازی وابستگی‌های زمانی، امکان تفکیک دقیق‌تر حالت‌های مختلف بهره‌برداری و خطا را فراهم نمودند.

نتایج شبیه‌سازی نشان داد که مدل پیشنهادی در طول فرآیند آموزش از روند همگرایی مناسبی برخوردار بوده و با کاهش تابع هزینه و افزایش دقت، الگوهای مربوط به جریان‌های تفاضلی را به خوبی فراگرفته است. همچنین، عملکرد مطلوب مدل در داده‌های اعتبار سنجی نشان‌دهنده قابلیت تعمیم مناسب آن و نبود بیش‌برازش قابل توجه است. براساس نتایج به دست آمده، مدل CNN-LSTM توانست با دقتی در حدود ۹۸.۰۸ درصد و زمان پاسخ‌دهی تقریباً ۱۰ میلی‌ثانیه، خطاهای داخلی سیم‌پیچی را از جریان‌های هجومی و آثار ناشی از اشباع ترانسفورماتور جریان تفکیک کند. مقایسه با روش‌های SVM و MLP نیز بیانگر عملکرد بهتر مدل پیشنهادی از نظر دقت تشخیص و سرعت تصمیم‌گیری بود.

در مجموع، استفاده هم‌زمان از قابلیت استخراج ویژگی مکانی CNN و توانایی تحلیل توالی زمانی LSTM، موجب افزایش قابلیت اطمینان سامانه تشخیص خطا شده است. با این حال، ارزیابی مدل در برابر تغییرات گسترده‌تر پارامترهای شبکه، نویزهای اندازه‌گیری، خطاهای ترکیبی و داده‌های واقعی آزمایشگاهی می‌تواند در پژوهش‌های آینده مورد توجه قرار گیرد. همچنین، ترکیب سیگنال‌های جریان تفاضلی با اطلاعات تکمیلی، مانند داده‌های تحلیل گازهای محلول در روغن، و بررسی معماری‌های پیشرفته‌تری نظیر Transformer (برخلاف LSTM که نمونه‌های سیگنال را عمدتاً به صورت متوالی پردازش می‌کند)، Transformer می‌تواند وابستگی بین نمونه‌های دور از یکدیگر را مستقیم‌تر شناسایی کند و پردازش موازی‌تری داشته باشد. می‌تواند به بهبود دقت تشخیص و افزایش قابلیت استفاده عملی از مدل در سامانه‌های حفاظتی ترانسفورماتور منجر شود.

References

- [1] S. T. Zahra, S. K. Imdad, S. Khan, S. Khalid, and N. A. Baig, "Power transformer health index and life span assessment: A comprehensive review of conventional and machine learning based approaches," *Engineering Applications of Artificial Intelligence*, vol. 139, pt. A, Art. no. 109474, Jan. 2025, doi: 10.1016/j.engappai.2024.109474.
- [2] J. Amjad, A. Siddique, and W. Aslam, "Machine learning-based health index evaluation of power transformers using novel parameterization for predictive maintenance: Data-driven research on Pakistan's national grid regarding maintenance cost optimization," *Energies*, vol. 19, no. 15, Art. no. 3653, 2026, doi: 10.3390/en19153653.
- [3] A.-M. Aciu, C.-I. Nicola, M.-C. Nițu, and M. Nicola, "Machine learning-based fault diagnosis of power transformers using a Duval pentagon combined complex and a weighted probabilistic ensemble," *Machines*, vol. 14, no. 6, Art. no. 634, 2026, doi: 10.3390/machines14060634.
- [4] A. A. Nazari, S. A. Hosseini, and B. Taheri, "Improving the performance of differential relays in distinguishing between high second harmonic faults and inrush current," *Electric Power Systems Research*, vol. 223, Art. no. 109675, Oct. 2023, doi:10.1016/j.jestch.2025.102144.

- [5] M. Oztekin, S. Karagol, and O. Ozgonenel, "Low-level inter-turn fault detection algorithm for transformer differential protection," *Applied Sciences*, vol. 16, no. 14, Art. no. 7073, 2026, doi: 10.3390/app16147073.
- [6] O. Surucu, S. A. Gadsden, and J. Yawney, "Condition monitoring using machine learning: A review of theory, applications, and recent advances," *Expert Systems with Applications*, vol. 221, Art. no. 119738, Jul. 2023, doi:10.1016/j.eswa.2023.119738.
- [7] M. H. Reza, M. N. B. Sibly, S. G. Rabbani, S. H. Sadi, M. F. Ahamed, F. B. Shafi, R. Sarmun, and M. E. H. Chowdhury, "A comprehensive review of convolutional neural networks: Foundations, enhancements and applications," *Neural Computing and Applications*, vol. 38, Art. no. 56, 2026, doi:10.1007/s00521-025-11827-w.
- [8] M. Krichen and A. Mihoub, "Long short-term memory networks: A comprehensive survey," *AI*, vol. 6, no. 9, Art. no. 215, 2025, doi: 10.3390/ai6090215.
- [9] F. Hafeez, Z. A. Arfeen, T. A. Jumani, M. I. Masud, N. Alkhaldi, A. Azhar, M. Aman, and M. K. Azam, "Hybrid LSTM-CNN model with temporal feature engineering and genetic algorithm optimization for temperature forecasting," *Eng*, vol. 7, no. 5, Art. no. 224, 2026, doi: 10.3390/eng7050224.
- [10] L. L. Mogotsi, A. Rasool, E. Matlotse, S. Ali, and A. Ali, "Intelligent fault detection of MV/HV transformers using fuzzy logic based on DGA," *Eng*, vol. 6, no. 9, Art. no. 228, 2025, doi: 10.3390/eng6090228.
- [11] H. Al Issa, M. Z. Qawaqzeh, S. Kurashkin, and O. Miroshnyk, "Monitoring of power transformers using thermal model and permission time of overload," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 3, pp. 2323–2334, Jun. 2022, doi: 10.11591/ijece.v12i3.pp2323-2334.
- [12] S. M. Ngong, F. M. Guy-de-patience, and N. N. Stephane, "A comparative study of power transformers windings diagnostic techniques: Part I," *Discover Electronics*, vol. 2, Art. no. 14, 2025, doi:10.1016/j.meaeene.2026.100112.
- [13] O. Abdusalam, A. Ibrahim, F. Anayi, and M. Packianather, "New hybrid machine learning method for detecting faults in three-phase power transformers," *Energies*, vol. 15, no. 11, Art. no. 3905, 2022, doi: 10.3390/en15113905.
- [14] D. Bejmert, M. Kereit, F. Mieske, W. Rebizant, K. Solak, and A. Wiszniewski, "Power transformer differential protection with integral approach," *International Journal of Electrical Power & Energy Systems*, vol. 118, Art. no. 105859, Jun. 2020, doi: 10.1016/j.ijepes.2020.105859.
- [15] C.-W. Huang, C.-C. Fang, W.-T. Hsu, C.-C. Yang, and L.-T. Zhou, "Improving the protection of step-down transformers by utilizing percentage differential protection and scale-dependent intrinsic entropy," *Entropy*, vol. 27, no. 4, Art. no. 444, 2025, doi: 10.3390/e27040444.
- [16] C. Mo, T. Y. Ji, L. L. Zhang, and Q. H. Wu, "Equivalent statistics based inrush identification method for differential protection of power transformer," *Electric Power Systems Research*, vol. 203, Art. no. 107664, Feb. 2022, doi: 10.1016/j.epsr.2021.107664.

- [17] E. Ali, A. Helal, H. Desouki, K. Shebl, S. Abdelkader, and O. P. Malik, "Power transformer differential protection using current and voltage ratios," *Electric Power Systems Research*, vol. 154, pp. 140–150, Jan. 2018, doi: 10.1016/j.epsr.2017.08.026.
- [18] A. Feizbakhsh, H. Meyar-Naimi, and M. Vakilian, "An accurate method for detection and compensation of current transformer saturation using unscented Kalman filter," *Electric Power Systems Research*, vol. 249, Art. no. 112015, Dec. 2025, doi: 10.1016/j.epsr.2025.112015.
- [19] T. Liu, X. Zeng, H. Xie, W. Wu, and J. Xiao, "Method for identifying error states of distribution network current transformers based on current imbalance weights," *Sustainable Energy Research*, vol. 12, Art. no. 42, 2025, doi:10.1186/s40807-025-00185-9.
- [20] H. Cao, C. Zhou, Y. Meng, J. Shen, and X. Xie, "Advancement in transformer fault diagnosis technology," *Frontiers in Energy Research*, vol. 12, 2024, Art. no. 1437614, Jul. 22, 2024, doi: 10.3389/fenrg.2024.1437614.
- [21] B. Zhao, H. Lu, S. Chen, and D. Wu, "Convolutional neural networks for time series classification," *Journal of Systems Engineering and Electronics*, vol. 28, no. 1, pp. 162–169, Feb. 2017, doi: 10.21629/JSEE.2017.01.18.
- [22] J.-W. Kim, J.-H. Lee, D.-H. Son, S.-H. Choi, and K.-S. Park, "Comparative analysis of CNN and LSTM for bearing fault mode classification and causality through representation analysis," *Lubricants*, vol. 14, no. 1, Art. no. 12, 2026, doi: 10.3390/lubricants14010012.
- [23] Y. Wang, Z. Xiao, and G. Cao, "A convolutional neural network method based on Adam optimizer with power-exponential learning rate for bearing fault diagnosis," *Journal of Vibroengineering*, vol. 24, no. 4, Jun. 2022, doi: 10.21595/jve.2022.22271.
- [24] E. Akbari, "A Frequency-Domain Differential Protection Scheme Based on Discrete Laplace Transform Considering the Effects of Voltage and Current Transformers in Transmission Lines," *Journal of Novel Researches on Smart Power Systems*, vol. 14, no. 2, pp. 35-59, Summer 2025 (in Persian).
- [25] E. Akbari, "Fault Detection, Classification and Location in Compensated Transmission System Using Gabor Wavelet Transform and Traveling Wave Theory," *Journal of Novel Researches on Smart Power Systems*, vol. 13, no. 2, pp. 1-17, Summer 2024 (in Persian).

زیرنویس‌ها:

- ¹ Convolutional Neural Network
² Long Short-Term Memory
³ Support Vector Machines
⁴ Multi Layer Perceptron
⁵ Dissolved Gas Analysis
⁶ Current Transformer