



A new acoustic sensing approach for predicting the percentage of filled rice grains based on the acoustic absorption spectrum using the Deep Spectra

Majid Fathi Ghalemiri¹ | Ali Maleki² | Majid Lashgari³ | Ali Loghmani⁴

1. Department of Mechanical Engineering of Biosystems, Faculty of Agriculture, Shahrekord University, Shahrekord, Iran.

Email: fathimajid1@gmail.com

2. Department of Mechanical Engineering of Biosystems, Faculty of Agriculture, Shahrekord University, Shahrekord, Iran.

Email: maleki@sku.ac.ir (Corresponding Author).

3. Department of Mechanical Engineering of Biosystems, Faculty of Agriculture, Arak University, Arak, Iran. Email: m-lashgari@araku.ac.ir

4. Department of Mechanical Engineering, Isfahan University of Technology, Isfahan, Iran. Email: a.loghmani@iut.ac.ir

Article Info

Article type: Research Article

Article history:

Received: May. 10, 2024

Revised: July. 17, 2024

Accepted: July. 29, 2024

Published online: Winter 2024

Keywords:

Acoustic Absorption Coefficient, Deep Spectra, Non-destructive Measurement, Rice Acoustic Sensor, Spectral Analysis

ABSTRACT

Rice is recognized as one of the main cereals in the world, feeding two-thirds of the global population, especially in Asian countries. Accurate assessment of the percentage of filled grains (PFG) is crucial for the efficiency and quality of rice harvesting. Traditional methods of measuring PFG are practical and based on personal judgment. This study introduces an innovative and non-destructive approach based on an acoustic sensor alongside deep learning models to predict PFG based on the acoustic spectrum of rice grains. Using an advanced deep learning architecture, the Deep Spectra, which works directly on raw spectral data, eliminates the need for preprocessing and enhances prediction accuracy. A modified impedance tube was used to measure the acoustic spectrum, which was then analyzed using the Deep Spectra model to predict PFG. Results indicated that this approach significantly improves the quantitative analysis of spectral data and provides a reliable prediction of rice grain filling. The prediction accuracy of the Deep Spectra model was significantly higher compared to traditional methods, with a low root mean square error of prediction (RMSEP) of 0.24 ± 0.05 and a coefficient of determination (R^2) of 0.95 ± 0.02 . This prediction is vital for assessing rice quality, breeding, and genetic research. This study introduces new perspectives and methods in the field of grain quality prediction and classification using acoustic spectrum analysis and deep learning, which could be beneficial for future research in this area.

Cite this article: Fathi Ghalemiri, M., Maleki, A., Lashgari, M., & Loghmani, A. (2024). A new acoustic sensing approach for predicting the percentage of filled rice grains based on the acoustic absorption spectrum using the Deep Spectra, *Iranian Journal of Biosystem Engineering*, 54 (4), 87-102. <https://doi.org/10.22059/ijbse.2024.376335.665548>

© The Author(s).

Publisher: The University of Tehran Press.

DOI: <https://doi.org/10.22059/ijbse.2024.376335.665548>



رویکرد جدید حسگر صوتی برای پیش‌بینی درصد دانه‌های پرشده برنج بر اساس طیف جذب صوتی به روش طیف عمیق

مجید فتحی قلعه میری^۱ | علی ملکی^۲ | مجید لشگری^۳ | علی لقمانی^۴

۱. گروه مهندسی مکانیک بیوسیستم، دانشکده کشاورزی، دانشگاه شهرکرد، شهرکرد، ایران. پست الکترونیک: fathimajid1@gmail.com

۲. گروه مهندسی مکانیک بیوسیستم، دانشکده کشاورزی، دانشگاه شهرکرد، شهرکرد، ایران، پست الکترونیک: maleki@sku.ac.ir

۳. گروه مهندسی مکانیک بیوسیستم، دانشکده کشاورزی، دانشگاه اراک، اراک، ایران. پست الکترونیک: m-lashgari@araku.ac.ir

۴. گروه مهندسی مکانیک دانشگاه صنعتی اصفهان، اصفهان، ایران پست الکترونیک: a.loghmani@iut.ac.ir

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی	برنج یکی از اصلی‌ترین غلات جهان شناخته می‌شود که دوسوم جمعیت جهان به‌ویژه در کشورهای آسیایی، از آن تغذیه می‌کنند. ارزیابی دقیق درصد دانه‌های پرشده (PFG) برای کارایی و کیفیت برداشت برنج حیاتی است. روش‌های سنتی اندازه‌گیری درصد دانه‌های پرشده کاربردی و مبتنی بر قضاوت شخصی است. این مطالعه رویکردی نوآورانه و غیرمخرب بر پایه حسگر صوتی در کنار مدل‌های یادگیری عمیق برای پیش‌بینی درصد دانه‌های پرشده بر اساس طیف صوتی دانه‌های برنج ارائه می‌دهد. با استفاده از معماری پیشرفته یادگیری عمیق، طیف عمیق که مستقیماً روی داده‌های طیفی خام کار می‌کند، نیاز به پیش‌پردازش حذف شد و دقت پیش‌بینی بهبود یافت. از لوله امیدانس تغییریافته، برای اندازه‌گیری طیف صوتی استفاده شد؛ سپس با استفاده از مدل طیف عمیق برای پیش‌بینی درصد دانه‌های پرشده، یافته‌ها تجزیه و تحلیل گردید. نتایج نشان داد که این رویکرد تجزیه و تحلیل داده‌های طیفی کمی را به‌طور قابل توجهی بهبود بخشیده و پیش‌بینی قابل اعتمادی از پرشدگی دانه‌های برنج ارائه می‌دهد. دقت پیش‌بینی مدل طیف عمیق در مقایسه با روش‌های سنتی به‌طور قابل توجهی بالاتر بود و خطای جذر میانگین مربعات پیش‌بینی (RMSEP) پایین (0.105 ± 0.024) و ضریب تعیین (R^2) (0.95 ± 0.02) بدست آمد. پیش‌بینی که برای ارزیابی کیفیت برنج، اصلاح نژاد و تحقیقات ژنتیکی برنج حیاتی است. این مطالعه دیدگاه و روش‌های جدیدی را در زمینه پیش‌بینی و طبقه‌بندی کیفیت دانه‌ها با استفاده از تحلیل طیف صوتی و یادگیری عمیق به حوزه تحقیقات کشاورزی ارائه می‌دهد که می‌تواند برای پژوهش‌های آتی در این زمینه مفید باشد.
تاریخ دریافت: ۱۴۰۳/۲/۲۱	
تاریخ بازنگری: ۱۴۰۳/۴/۲۷	
تاریخ پذیرش: ۱۴۰۳/۵/۸	
تاریخ انتشار: زمستان ۱۴۰۲	
واژه‌های کلیدی: اندازه‌گیری غیرمخرب، تحلیل طیفی، حسگر صوتی برنج، ضریب جذب صوتی، طیف عمیق	

استاد: فتحی قلعه میری، مجید، ملکی، علی، لشگری، مجید و لقمانی، علی. (۱۴۰۳). رویکرد جدید حسگر صوتی برای پیش‌بینی درصد دانه‌های پرشده برنج بر اساس طیف جذب صوتی به روش طیف عمیق، *مجله مهندسی بیوسیستم ایران*، ۵۴ (۴)، ۸۷-۱۰۲.



<https://doi.org/10.22059/ijbse.2024.376335.665548>

© نویسندگان.

ناشر: مؤسسه انتشارات دانشگاه تهران.

DOI: <https://doi.org/10.22059/ijbse.2024.376335.665548>

مقدمه

برنج از مهم‌ترین و اصلی‌ترین غلاتی است که حدود دوسوم جمعیت جهان از آن تغذیه می‌کند. برنج اصلی‌ترین منبع غذایی در کشورهای آسیایی است (Burlando & Cornara, 2014). بیشتر دانه‌های برنج به‌صورت پخته مصرف می‌شوند و از درصد کمی از محصول، به‌عنوان عنصری در فرآیند تولید مواد غذایی و غیرغذایی استفاده می‌گردد. موتایا و همکاران تولید سالانه برنج را تقریباً ۴۸۰ میلیون تن ارزیابی می‌کند. مردم منطقه آسیا و اقیانوسیه، بخش‌هایی از منطقه کارائیب و آمریکای لاتین برای تأمین کالری و تغذیه خود وابستگی بیشتری به برنج دارند (Muthayya *et al.*, 2014). با توجه به سلايق غذایی مردم ایران، برنج همیشه یکی از مهم‌ترین مواد غذایی مورد نیاز روزانه در سبد غذایی خانوارها بوده و در سه دهه گذشته، تقاضا برای این محصول به‌طور پیوسته افزایش یافته است؛ بنابراین، این محصول راهبردی یکی از کالاهای اساسی و ضروری در سبد خرید خانوارهای ایرانی است (Koulivand & Moniri).

یکی از معیارهای سنجش کیفیت محصول برنج، درصد دانه‌های پر شده می‌باشد. به نسبت دانه‌های پر شده به کل دانه‌های موجود در خوشه، درصد دانه‌های پر شده (PFG) می‌گویند. این نسبت به عنوان یک معیار کلیدی برای ارزیابی عملکرد و کیفیت محصول برنج مورد استفاده قرار می‌گیرد. به عبارتی دیگر، هر چه درصد دانه‌های پر بیشتر باشد، نشان‌دهنده کیفیت بالاتر و عملکرد بهتر محصول است. این ویژگی کلیدی، بازده دانه را توصیف می‌کند و معیاری مهم برای توصیف مدیریت تولید برنج و تحقیقات علمی است. روش‌های کنونی اندازه‌گیری درصد دانه‌های پر شده (جداسازی به کمک آب، باد و شمارش دستی) روش‌هایی زمان‌بر و نیروبر هستند و نتایج آن‌ها مبتنی بر قضاوت شخصی است (Liu *et al.*, 2016). در حال حاضر، از برخی روش‌های نیمه خودکار برای اندازه‌گیری درصد دانه‌های پر شده استفاده می‌شود. از طریق این روش‌ها، کاربر معمولاً دانه‌های پر و ناپُر را به صورت دستی جدا می‌کند و سپس تعداد آن‌ها را با استفاده از تجهیزات شمارش دانه (دستگاه شمارش دانه خودکار، نرم‌افزار شمارش دانه مبتنی بر تصویر و سینی شمارش دانه) می‌شمارد.

دو شیوه مرسوم غیر مخرب، متکی بر بینایی ماشین و صوتی هستند. پژوهشگرانی روشی برای اندازه‌گیری درصد دانه‌های پر شده بر اساس سایه تصاویر با دقت قابل قبولی ارائه نمودند که به‌طور فزاینده‌ای از فناوری بینایی رایانه‌ای، برای تجزیه و تحلیل دانه‌های محصولات زراعی از جمله تقسیم بذر و شمارش استفاده می‌شود (Liu *et al.*, 2016). Lin *et al.* در مطالعه‌ای، به بررسی استفاده از بینایی ماشین برای تجزیه و تحلیل و شمارش دانه‌ها پرداختند (Lin *et al.*, 2014). Manickavasagan *et al.* از تصاویر تک‌رنگ، برای شناسایی گونه‌های مختلف گندم استفاده کردند (Manickavasagan *et al.*, 2008). برای رفع محدودیت‌های فناوری بینایی رایانه‌ای در تشخیص دانه‌های پر و ناپُر، Duan *et al.* فناوری پرتوی ایکس نرم را به سامانه بینایی رایانه‌ای افزودند. پرتوی ایکس نرم می‌تواند از پوسته دانه عبور کرده و توده داخلی آن را تحت تأثیر قرار دهد؛ این امر در تمایز قاطع بین دانه‌های پر و ناپُر مؤثر است؛ اما هزینه بالای تجهیزات پرتوی ایکس نرم، کاربرد آن را در کشاورزی محدود کرده است؛ بنابراین، کاهش هزینه‌های مرتبط با اندازه‌گیری پری دانه‌ها، همیشه مسئله‌ای مهم و اساسی بوده است (Duan *et al.*, 2011). Kumar *et al.* با استفاده از پیش گرم کردن نمونه‌ها و تصویربرداری حرارتی و RGB به دقت خوبی برای جداسازی دانه‌های پر و خالی رسیدند؛ روش آن‌ها مقرون به‌صرفه‌تر از روش پرتوی ایکس نرم بود (Kumar *et al.*, 2020). مطالعه‌ای دیگر درباره طبقه‌بندی ارقام گندم بر اساس طیف‌سنجی نزدیک به فروسرخ و با استفاده از رویکرد یادگیری عمیق مبتنی بر تصویر انجام شده است. نتایج این مطالعه، دقت بالایی را نشان می‌دهد و قابلیت گسترش برای طبقه‌بندی سایر انواع غلات را دارد (Ceyhan *et al.*, 2023). در مطالعه دیگر برای طبقه‌بندی دانه‌های برنج پر و خالی، روش جدیدی بر اساس تصویربرداری نور ساختاریافته سه‌بعدی و شبکه بهبودیافته یادگیری عمیق ارائه گردید که دقت نتایج در آن ۹۸/۵ درصد بود (Huang *et al.*, 2023).

پیشینه پژوهش

عمده تحقیقات صورت گرفته متکی بر روش تصویربرداری بوده‌اند. با این حال استفاده از روش‌های صوتی در فرایندهای آزمون و تشخیص غیر مخرب رو به گسترش است. استفاده از روش‌های صوتی برای تشخیص مواد دانه‌ای شامل چندین راهبرد است که اغلب از خصوصیات امواج صوتی در تعامل با مواد دانه‌ای، استفاده می‌کنند. به‌طور کلی روش‌های مبتنی بر صوت در مواد دانه‌ای را می‌توان به‌صورت زیر تقسیم‌بندی کرد:

تحلیل انتشار صوتی: این روش شامل نظارت بر امواج صوتی منتشر شده توسط مواد دانه‌ای در حالت تنش یا حرکت می‌شود.

تجزیه و تحلیل می‌تواند اطلاعاتی درباره ویژگی‌های ماده مانند اندازه دانه، چگالی و نقایص داخلی ارائه دهد (Nguyen *et al.*, 2022).
آزمایش فراصوت: از امواج صوتی با بسامد بالا برای نفوذ در مواد دانه‌ای استفاده می‌شود؛ صدای بازتاب شده برای شناسایی ویژگی‌ها و ناهنجاری‌های داخلی مورد تجزیه و تحلیل قرار می‌گیرد (Gheibi & Hedayat, 2018).

طیف‌سنجی صوتی رزونانس: روشی پیشرفته برای تحلیل خصوصیات فیزیکی و مکانیکی مواد دانه‌ای است. این روش بر اساس اصل رزونانس صوتی عمل می‌کند که در آن، بسامدهای خاصی از امواج صوتی با ماده تعامل دارند و باعث رزونانس یا تقویت امواج در آن بسامدهای خاص می‌شوند. این بسامدهای رزونانس به‌طور مستقیم با خصوصیات ماده مانند ساختار داخلی، چگالی، اندازه دانه و خواص الاستیک آن مرتبط هستند (Geimer *et al.*, 2022; Griffiths *et al.*, 2017; Ostrovsky *et al.*, 2001).

ضرب جذب صوتی: روش مهمی در تحلیل خصوصیات مواد است و در بسیاری از زمینه‌های علمی و صنعتی کاربرد دارد. این روش بر اساس اصل میزان جذب امواج صوتی توسط مواد مختلف عمل می‌کند. هنگامی که موج صوتی به ماده‌ای برخورد می‌کند، بخشی از انرژی آن جذب می‌شود و بخشی دیگر بازتاب یا عبور داده می‌شود. میزان جذب صوت توسط ماده می‌تواند اطلاعات مفیدی در مورد خصوصیات آن مانند چگالی، ترکیب، ساختار داخلی و خواص مکانیکی ارائه دهد (Gasso-Tortajada *et al.*, 2010; Boubel *et al.*, 2016; Tsuruha *et al.*, 2020).

ضرب جذب صوت و امپدانس سطح هر ماده، اغلب با استفاده از لوله امپدانس اندازه‌گیری می‌شود. اندازه‌گیری با استفاده از لوله امپدانس بسیار ساده‌تر از اندازه‌گیری انجام‌شده در اتاق بازتابش (reverberation chamber) است، اگرچه در این روش، فقط عوامل فرود عمودی مواد را می‌توان به دست آورد (ISO 10534-1). این عوامل در توصیف ماده و در بررسی اثر ترکیبی چندین ماده بسیار مفید هستند (Mosa *et al.*, 2019; Lee & Xu, 2009; Lee & Wang, 2006). علاوه بر روش استاندارد، اخیراً لوله امپدانس اصلاح‌شده یا روش لوله امپدانس چهار میکروفن برای اندازه‌گیری ضرب جذب و تلفات انتقال صدا ارائه شده است (Li *et al.*, 2020; Roozen & Piana, 2022). لئو و شی، الگوریتمی را برای محاسبه تلفات انتقال صدا در پوسته‌های ساندویچی استوانه‌ای با مواد متخلخل الاستیک و شکاف هوا، با استفاده از نظریه بایوت طراحی کردند. در مطالعه‌ای دیگر، از لوله کونت برای اندازه‌گیری تلفات انتقال و جذب صوتی به روش موج ایستا در مواد متخلخل استفاده شد و نتایج عملکرد تقریباً مشابه با روش چهار میکروفن بود (Caballol & Raposo, 2020). در مطالعه‌ای دیگر، طبقه‌بندی چند محصول کشاورزی بر اساس طیف ضرب جذب و تجزیه و تحلیل مؤلفه‌های اصلی (PCA) برای طبقه‌بندی دانه‌ها صورت گرفت. طیف ضرایب جذب صوتی انواع مختلف دانه‌ها، نشان می‌دهد که ساختار الگوها به اندازه و شکل دانه بستگی دارد. نتایج، رابطه عکس را بین اندازه ذرات دانه و کروی بودن آن‌ها با ضرب جذب نشان داد (Gasso-Tortajada *et al.*, 2010).

روش‌های کالیبراسیون غیرخطی مانند شبکه عصبی مصنوعی (ANN) می‌توانند ویژگی‌های خطی و غیرخطی طیف‌ها را در برگیرند. با این حال، اغلب این روش‌ها، به دلیل ساختار پیچیده مدل، خطر بیش‌برازش و نیاز به تعداد نمونه‌های بزرگ، مورد انتقاد قرار می‌گیرند (Ni *et al.*, 2014). بیشتر تحلیل‌های کالیبراسیون غیرخطی موجود بر روی داده‌های طیفی خام کار نمی‌کنند و اغلب نیاز به تجزیه مؤلفه‌های اصلی برای کاهش ابعاد ماتریس داده‌های طیفی دارند (Sun *et al.*, 2017; Kamboj *et al.*, 2021; Allegrini & Olivieri, 2016). با این حال ممکن است کاهش ابعاد داده‌ها، الگوهای اصلی سیگنال طیفی را تغییر دهد و اطلاعات مفید از دست برود. بنابراین، بهبود دقت و استحکام مدل نیازمند رویکردی یکپارچه برای حذف درستی‌های مختلف و استخراج اطلاعات مرتبط با ویژگی‌های محصول از داده‌های طیف‌سنجی است. (Passos & Mishra, 2022; Liu *et al.*, 2024a)

پیشرفت‌های اخیر در حوزه یادگیری ماشین نشان‌دهنده این واقعیت است که روش‌های یادگیری عمیق که بر اساس داده‌ها عمل می‌کنند، قادر به شناسایی ساختارهای پیچیده در داده‌هایی با ابعاد بزرگ هستند. این امر منجر به کاهش نیاز به دانش قبلی و کار انسانی در زمینه طراحی و تعریف ویژگی‌های مورد نیاز برای تحلیل داده‌ها می‌شود. به بیان ساده‌تر، به کمک یادگیری عمیق می‌توان بدون تلاش زیاد برای فهمیدن و پیش‌پردازش داده‌ها، از عهده شناسایی الگوهای موجود در داده‌های پیچیده و بزرگ برآمد. معماری یادگیری عمیق، از چندین لایه تشکیل شده است که هر کدام، از ماژول‌های ساده اما غیرخطی ساخته شده‌اند. این ماژول‌ها نمایش داده‌ها در یک سطح (شروع با ورودی خام) را به نمایشی در سطح بالاتر و کمی منفردتر تبدیل می‌کنند (Liu *et al.*, 2024b; Lecun *et al.*, 2015).

یادگیری عمیق روش پیشرفته یادگیری ماشینی است که با پردازش داده‌های ورودی لایه‌به‌لایه از طریق شبکه‌ای عمیق، ویژگی‌های عمیق و انتزاعی را استخراج می‌کند؛ اخیراً استفاده از روش‌های مبتنی بر یادگیری عمیق در پردازش داده‌های طیفی رو به رشد است

(Zhang et al., 2021b,a, 2019).

شبکه‌های عصبی پیچشی یا هم‌گشتی رویکردی محبوب در یادگیری عمیق هستند که در تحلیل داده‌های دو و سه‌بعدی، از جمله شناسایی تصویر، تجزیه و تحلیل ویدیو، کشف الگو به‌خوبی به کار می‌روند. رویکرد شبکه عصبی پیچشی، از اتصالات محلی پراکنده برای مطالعه الگوهای محلی و داده‌های خام بهره می‌برد و با اتخاذ اشتراک‌گذاری وزن‌ها، خطر بیش‌برازش را کاهش می‌دهد (Liu et al., 2024b). معماری شبکه‌های عصبی پیچشی شامل چندین لایه از جمله لایه‌های پیچشی، لایه‌های فعال‌سازی، لایه‌های استخراج، لایه تسطیح و لایه‌های کاملاً متصل^۳ می‌شود (Zhang et al., 2019; Passos & Mishra, 2022). بخش آغازین^۴ که در شبکه عصبی گوگل نت توسعه یافته، با افزایش همزمان عرض و عمق شبکه عصبی، دستاورد جدیدی در عرصه یادگیری عمیق به ارمغان آورده است (Szegedy et al., 2015).

مدل طیف عمیق^۵ که در پیش‌تر توسط شیائولی ژانگ و همکاران در مقاله تحت عنوان «رویکرد یادگیری عمیق انتها به انتها برای تحلیل کمی طیف‌سنجی» معرفی شده، معماری پیشرفته‌ای برای تحلیل داده‌های طیفی است که مستقیماً بر روی داده‌های خام کار می‌کند و به پیش‌پردازش نیازی ندارد. این مدل با استفاده از بخش آغازین که اجازه می‌دهد از معماری‌های عمیق‌تر و عریض‌تر بهره ببرد، قادر است ویژگی‌های مختلف داده‌ها را در سطوح مختلف انتزاعی کشف کند؛ این مدل در مقایسه با روش‌های کالیبراسیون خطی و غیرخطی سنتی مانند حداقل مربع جزئی، شبکه عصبی مصنوعی و ماشین بردار پشتیبانی، عملکرد بهتری را در تحلیل‌های طیفی نشان داده است (Zhang et al., 2019; Passos & Mishra, 2022).

این مطالعه با هدف توسعه رویکرد حسگر غیر مخرب و درجا جدید بر اساس پدیده جذب صوتی برای پیش‌بینی درصد پرشدگی با مدل طیف عمیق است. در این مطالعه، از لوله امیدانس تغییر یافته جهت اندازه‌گیری طیف صوتی جذبی توده با پایش اولیه استفاده شد و با مدل طیف عمیق پیش‌بینی درصد پرشدگی صورت پذیرفت. نتایج نشان داد، رویکرد طیف عمیق، تجزیه و تحلیل کمی داده‌های طیف‌سنجی را بدون نیاز به پیش‌پردازش داده‌ها بهبود می‌بخشد و پیش‌بینی مؤثری از درصد پرشدگی توده برنج ارائه می‌کند که شاخص مهم محصول برنج و دیگر غلات برای بررسی کیفیت برنج، اصلاح نژاد برنج و تحقیقات ژنتیکی است، این مطالعه در پی ارائه روش و دیدگاه متفاوتی برای پیش‌بینی و طبقه‌بندی با استفاده از طیف جذب صوتی است؛ هدف آن خلق ایده برای پژوهش‌گران جهت تحقیقات آینده در بررسی‌های کیفی و کمی توده دانه‌های غلات است که با استفاده از طیف جذب صوتی توده و با تحلیل روش‌های نوین یادگیری عمیق انجام می‌شود. این موارد مکملی بر روش‌هایی پیشنهادی مقالات پیشین خواهد بود.

روش‌شناسی پژوهش

آماده‌سازی نمونه

در این مطالعه، پنج گونه برنج با پوست (چمپا، طارم، شیروودی، فجر و لنجان) که اغلب در ایران کشت می‌شوند و یک گونه برنج وارداتی (باسماتی)، از پنج مزرعه برای هر رقم، مورد بررسی قرار گرفتند. جهت جداسازی دانه‌های پر و خالی برنج، از روش دستی مبتنی بر آب استفاده شد (Yang et al., 2000). در این روش دانه‌های پر به ته ظرف و دانه‌های خالی روی آب قرار می‌گیرند. سپس نمونه‌ها به مدت ۴۸ ساعت در گرم‌خانه قرار گرفته و پس از برچسب‌گذاری و جهت اطمینان از عدم تغییر رطوبت در بسته‌بندی‌های زیپ‌دار پلاستیکی ریخته شدند. سپس با دانه‌های پر و خالی از هر گونه برنج، ترکیبی شش‌تایی از درصد پرشدگی به صورت «۱۰۰، ۹۵، ۹۰، ۸۵، ۸۰، ۷۵، ۷۰، ۶۵، ۶۰، ۵۰» درصد درست شد (شکل ۱) بدین ترتیب کل نمونه‌ها ۳۶۰ عدد شد.

ضریب جذب صوتی

اندازه‌گیری ضریب جذب صوتی بر اساس روش تابع انتقال در لوله امیدانس با استاندارد ISO 10534-2 انجام شد (ISO 10534-2:2023). به‌منظور اندازه‌گیری ضریب جذب و با توجه به اینکه برخلاف مواد جاذب نمی‌توان نمونه را افقی قرار داد؛ زیرا هدف اندازه‌گیری ضریب

1. Convolutional Neural Networks
2. pooling
3. Flatten Layer
4. Inception module
5. GoogleNet
6. DeepSpectra

جذب، یک توده برنج است و امکان قرارگیری توده برنج به صورت افقی در چیدمان استاندارد نبود، چیدمان سفارشی به همراه نگهداره نمونه طراحی گردید (شکل ۲). برای اندازه گیری صدا از دو میکروفن مدل ECM8000 از برند بهرینگر^۱ و کارت BSWA DAQ MC3522 استفاده شد. موج صوتی صفحه ای تصادفی با زمان بارگذاری یک ثانیه توسط بلندگوی JBL 2408H-1 تولید می شد. سپس تابع انتقال صوتی پیچیده به دست آمده از سیگنال های هر دو میکروفن محاسبه گردید. این تابع در تعیین ضریب جذب طبیعی نمونه های آزمایش ضروری است. تولید سیگنال صوتی، اندازه گیری و تحلیل داده با توسعه اسکریپت A-Lab در نرم افزار Matlab 2023b صورت گرفت.



شکل ۱ - ترکیب بندی نمونه های برنج مورد استفاده. ستون عمودی درصد پرشدگی و ستون افقی گونه برنج است.

بر اساس استاندارد ISO 10534-2، امپدانس صوتی مخصوص را می توان با رابطه رابطه (۱) محاسبه کرد:

$$\frac{Z}{\rho \hat{A}.c} = \frac{R}{\rho \hat{A}.c} + \frac{ZX}{\rho \hat{A}.c} j = \frac{(1+r)}{(1-r)} \quad \text{رابطه (۱)}$$

در رابطه رابطه (۱)، R قسمت حقیقی امپدانس، X قسمت موهومی امپدانس، $\rho \hat{A}.c$ امپدانس مشخصه (چگالی ضرب در سرعت صوت)، r ضریب بازتاب، Z ، امپدانس صوتی است.

ضریب بازتاب با رابطه رابطه (۲) تعیین می شود:

$$r = \frac{H_{12} - H_1}{H_R - H_{12}} \hat{A}.e^{2jk_0x_1} \quad \text{رابطه (۲)}$$

و ضریب جذب صدا با رابطه رابطه (۳) به دست می آید:

$$\alpha = 1 - |r|^2 \quad \text{رابطه (۳)}$$

H_{12} تابع انتقال بین میکروفن ۱ و ۲ (شکل ۲)، H_R قسمت حقیقی H_{12} ، H_i قسمت موهومی H_{12} ، k_0 عدد موج، x_1 فاصله بین

نمونه و اولین میکروفن است.

در چیدمان سفارشی لوله امپدانس ساخته شده، فاصله دو میکروفن ۵ سانتی متر و فاصله دورترین میکروفن تا نمونه، ۱۳ سانتی متر و قطر لوله امپدانس، ۱۰ سانتی متر در نظر گرفته شد. با توجه به مقادیر و بر اساس محدودیت‌های اندازه‌گیری در روش استاندارد، محدوده بسامد پایین ۳۵۰ هرتز و بسامد بالای ۱۸۹۵ هرتز به دست می‌آید (رابطه (۴)).

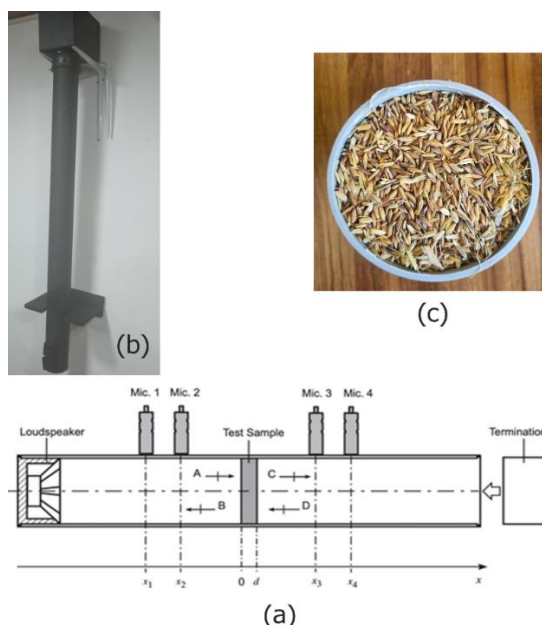
$$f_{\min}=0.05 \frac{c}{s}, f_{\max}=k \frac{c}{d}, s_{\max}=0.45 \frac{c}{f_{\max}} \quad (\text{رابطه } 4)$$

در رابطه (۴)، f_{\min} محدوده بسامد پایین، f_{\max} محدوده بسامد بالا، d قطر لوله امپدانس، s فاصله بین دو میکروفن، c سرعت صوت، k در سطح مقطع دایره ۰/۵ است.

پالودن داده‌های خام و کاهش نوفه

هدف از پالودن داده‌های طیفی، حذف نوفه، اعوجاج و عدم قطعیت‌ها از طیف‌هاست (Rinnan *et al.*, 2009). انتخاب روش مناسب پیش پردازش هنوز به صورت آزمایش و خطا انجام می‌شود و نیازمند محاسباتی قابل توجهی است (Zhang *et al.*, 2019). به طور متداول در اندازه‌گیری‌های α ، برخی نمونه‌های غیرعادی مشاهده می‌شوند که شامل کاهش‌های ناگهانی و شدید در ضریب جذب در محدوده‌های خاصی از بسامدها بود. این موارد نه تنها از نظر فیزیکی قابل تفسیر نبودند بلکه باعث ایجاد توزیع‌های داده‌ای با دنباله‌های بلند، می‌شدند که چنین توزیع‌هایی، چالش‌های عمده‌ای را برای روش‌های یادگیری ماشینی ایجاد می‌کنند. به همین دلیل، این نقاط دور از معیار، از طریق روش هموارسازی میانه متحرک از منحنی‌های اندازه‌گیری حذف شدند تا داده‌هایی یک‌دست‌تر و قابل تحلیل‌تر برای الگوریتم‌های یادگیری ماشینی فراهم شود (Stender *et al.*, 2021). چند فیلتر، قبل از ورود داده‌ها به الگوریتم یادگیری اعمال گردید. در آغاز، در برخی از موارد ضرایب جذب ناشی از نوفه، منفی بود که با مقادیر صفر جایگزین شدند. سپس اقدام به اعمال فیلتر میانه متحرک با اندازه پنجره ۲۰ و حداکثر متحرک با اندازه پنجره ۵ گردید.

هموارسازی میانه متحرک، روشی آماری است که برای صاف کردن یک سری داده‌های زمانی یا هر نوع داده متوالی استفاده می‌شود. این روش به ویژه برای حذف نوفه یا داده‌های پرت از مجموعه داده‌ها مفید است. در این روش، به جای استفاده از میانگین، از میانه داده‌های داخل یک پنجره متحرک استفاده می‌شود. هموارسازی میانه متحرک، به خصوص برای داده‌هایی که دارای نوسانات یا پرت‌های ناگهانی هستند، مفید است و به حفظ ساختار کلی داده‌ها کمک می‌کند؛ این روش نوفه و داده‌های نامطلوب را حذف یا کاهش می‌دهد (Hastie *et al.*, 2009).



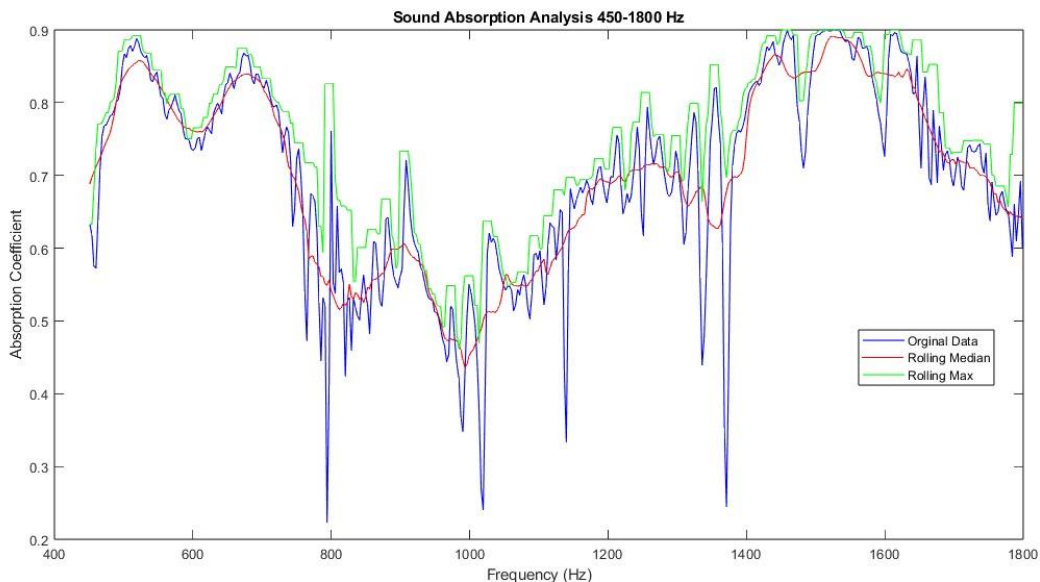
شکل ۲- (a) طرح لوله امپدانس مرسوم. (b) لوله امپدانس سفارشی‌سازی شده برای اندازه‌گیری ضریب جذب صوتی. (c) نمونه برنج با پوست به همراه نگهدارنده نمونه.

$$y_smooth = \text{median}([y_f, y_{f-1}, \dots, y_{f-(n-1)}]) \quad \text{رابطه ۵}$$

y_smooth مقدار صاف شده در بسامد f ، مقدار واقعی در بسامد f و n طول پنجره صاف کننده است. تابع میانه (median):

$$\text{median}(y) = \begin{cases} y\left[\frac{n+1}{2}\right] & \text{فرد } n \\ \frac{y\left[\frac{n}{2}\right] + y\left[\frac{n}{2}+1\right]}{2} & \text{زوج } n \end{cases} \quad \text{رابطه ۶}$$

حداکثر متحرک، روشی در پردازش داده‌ها و تحلیل سری‌های زمانی است. مشابه با میانه متحرک، این روش شامل حرکت دادن یک پنجره با اندازه مشخص بر روی مجموعه‌ای از نقاط داده و محاسبه بیشترین مقدار داخل آن پنجره برای هر موقعیت است. در فرآیند فیلتر کردن داده‌ها، داده‌های جدیدی به دست آمده، به خوبی ساختار اصلی داده‌های خام را دنبال می‌کنند. باین‌حال، این داده‌های جدید فاقد کاهش‌های ناگهانی هستند که در داده‌های اصلی مشاهده می‌شدند و از نظر فیزیکی قابل توجیه نبودند. این نوع کاهش‌ها معمولاً نشانه‌ای از درست‌نماهای اندازه‌گیری یا نوفه هستند. برای نشان دادن این فرآیند فیلترها و بهبود داده‌ها، شکل ۳، مثالی از داده‌های تمیز شده را در کنار کلیه داده‌های پردازش شده و بهبودیافته ارائه می‌دهد. این تصویر به خوبی نشان می‌دهد که چگونه داده‌های فیلتر شده می‌توانند ضمن حفظ ویژگی‌های اصلی، عناصر نامطلوب یا درست‌نماهای داده‌های خام را حذف یا کم اثر کند.



شکل ۳- نمونه اعمال فیلترسازی اولیه در یک داده جذب صوتی

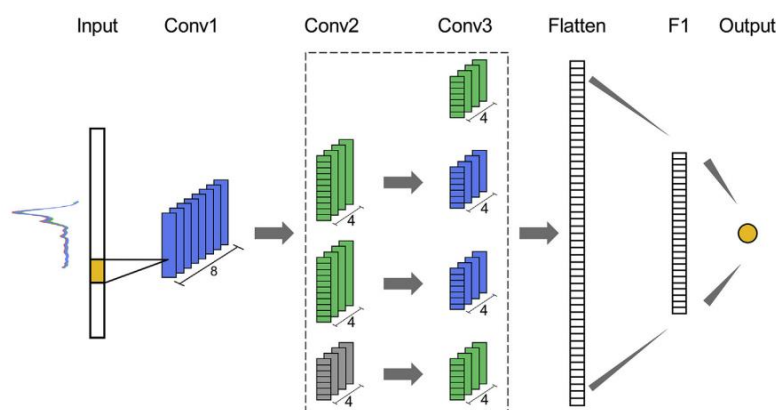
معماری مدل طیف عمیق

مدل طیف عمیق، ساختاری نوآورانه در زمینه یادگیری عمیق است که برای تجزیه و تحلیل کمی داده‌های طیف‌سنجی طراحی شده است (Zhang *et al.*, 2019; Passos & Mishra, 2022). این مدل بدون نیاز به پیش‌پردازش داده‌های طیفی که گاهی اوقات می‌تواند منجر به از دست رفتن یا تحریف اطلاعات شود، به کار گرفته می‌شود. معماری مدل بر پایه شبکه‌های عصبی پیچشی بنا شده است که به واسطه استفاده از وزن‌های مشترک و اتصالات محلی، تعداد عوامل قابل آموزش را کاهش می‌دهد. این ویژگی‌ها به کارایی مدل کمک کرده و خطر بیش‌برازش را کاهش می‌دهد.

معماری مدل طیف عمیق در شکل ۴ به تصویر کشیده شده است. این مدل شامل سه لایه پیچشی (Conv1، Conv2 و Conv3) است که هر کدام ویژگی‌های مختلفی از داده‌های ورودی را استخراج می‌کنند. لایه تخت‌کننده^۱ و لایه کاملاً متصل (F1)، به ترتیب وظیفه تبدیل خروجی‌های چندبعدي لایه‌های پیچشی به برداری یک‌بعدي و ترکیب ویژگی‌های استخراج شده را برای تولید خروجی نهایی بر عهده دارند. این ساختار بهینه‌سازی شده با استفاده از اتصالات مستقیم و ماژول‌های پیچشی ۱×۱ در لایه‌های Conv2 و Conv3،

پیچیدگی محاسباتی را کاهش داده و کارایی مدل را افزایش می‌دهد.

ماژول‌های آبی در شکل ۴ نمایانگر لایه پیچشی عمومی، ماژول‌های سبز نمایانگر پیچشی 1×1 و ماژول‌های خاکستری نمایانگر عملیات ادغام است که به ترتیب به کاهش ابعاد ویژگی‌ها و کاهش ریسک بیش‌برازش کمک می‌کنند. اتصال کوتاه در لایه Conv3 که مستقیماً به لایه Conv1 متصل شده، راه‌برد مهمی برای حفظ اطلاعات طی فرآیند آموزش و کاهش مشکل ناپدید شدن گرادینان است. این معماری همچنین به مدل امکان می‌دهد تا از داده‌های خام بدون نیاز به مراحل پیش‌پردازش پیچیده استفاده کند؛ این امر به حفظ همه اطلاعات مفید موجود در داده‌های ورودی و بهبود دقت تحلیل کمی کمک می‌کند.



شکل ۴ - معماری مدل طیف عمیق شامل سه لایه پیچشی، لایه تخت‌کننده، لایه کاملاً متصل و لایه خروجی، همراه با استفاده از ماژول‌های ادغام و پیچشی 1×1 (Zhang et al., 2019)

لایه Conv1 در معماری مدل طیف عمیق، لایه پیچشی اولیه است که نقش بسیار مهمی در استخراج ویژگی‌های ابتدایی از داده‌های طیفی ورودی ایفا می‌کند. این لایه مجهز به هشت فیلتر است که همگی اندازه یکسانی دارند. در این لایه، هر فیلتر به‌طور مستقل روی داده‌های ورودی اعمال می‌شود تا ویژگی‌های خاصی را در سطوح مختلف تشخیص دهد. هنگامی که داده‌ها از این فیلترها عبور می‌کنند، طیف‌های یک‌بعدی به نقشه‌های ویژگی که در واقع نمایانگر پاسخ‌های فیلتر به داده‌ها هستند، تبدیل می‌شوند.

اندازه فیلتر و گام در این لایه بر اساس ویژگی‌های داده‌های ورودی و اندازه نمونه‌های آموزشی تنظیم می‌شوند. این تنظیمات بهینه به مدل کمک می‌کنند تا با کارایی بالا، ویژگی‌های مهم را شناسایی و استخراج کند، زیرا تعیین اندازه صحیح فیلتر و گام می‌تواند تأثیر بسزایی در قابلیت مدل برای تشخیص الگوهای مفید در داده‌ها داشته باشد. این لایه به‌عنوان پایه و اساس برای لایه‌های پیچشی بعدی عمل می‌کند، که هر کدام ویژگی‌های بیشتری را در سطوح پیچیده‌تر استخراج می‌کنند (جدول ۱).

جدول ۱- ابرعوامل استفاده‌شده در مدل‌سازی طیف عمیق

پارامتر	مقدار
اندازه هسته ۱	۷
اندازه هسته ۲	۳
اندازه هسته ۳	۵
گام ۱	۳
گام ۲	۲
تعداد نورون‌های مخفی	۳۲
اندازه دسته	۱۲۸
نرخ حذف	۰/۱
ضریب تنظیم	۰/۰۱

لایه تخت‌کننده در معماری مدل طیف عمیق، عملیات مهمی را انجام می‌دهد؛ این لایه خروجی‌های موازی را از لایه Conv3 جمع‌آوری کرده و آن‌ها را به برداری یک‌بعدی تبدیل می‌کند. این بردار یک‌بعدی سپس به لایه کاملاً متصل F1 است.

منتقل می‌شود. یکی از ویژگی‌های کلیدی لایه تخت‌کننده این است که هیچ عامل قابل آموزشی ندارد، بنابراین تنها وظیفه‌اش، تغییر شکل داده‌هاست. برای مقابله با مشکل بیش‌برازش در مدل، عملکرد حذف تصادفی^۱ در معماری، مورد استفاده قرار گرفته است. این روش با حذف تصادفی تعدادی از نورون‌ها در طول فرآیند آموزش، به کاهش پیچیدگی مدل کمک کرده و در نتیجه به افزایش کارایی محاسباتی و پیشگیری از بیش‌برازش منجر می‌شود. این رویکرد به مدل اجازه می‌دهد فقط با ویژگی‌های مهم‌تر کار کند؛ با این رویکرد عملکرد کلی مدل بهبود می‌یابد.

لایه F1 لایه‌ای تراکمی^۲ است که به‌طور کامل به لایه تخت‌کننده متصل می‌شود. تعداد نورون‌ها در لایه F1، کمتر از تعداد نورون‌های موجود در لایه تخت‌کننده است. این کاهش تعداد نورون‌ها به مدل کمک می‌کند تا از داده‌های فشرده‌تری برای تصمیم‌گیری استفاده کند و به این ترتیب از پیچیدگی بیش‌ازحد مدل جلوگیری می‌کند. لایه خروجی نیز به‌طور کامل به لایه F1 متصل شده و عملکرد حذف تصادفی در آن گنجانده شده است تا احتمال بیش‌برازش را کاهش دهد.

لایه Output، که لایه خروجی مدل است، فقط یک نورون دارد که وظیفه آن ارائه خروجی نهایی یا هدف تخمین زده شده است. این نورون بر اساس اطلاعات دریافتی از لایه‌های قبلی، مقداری را تولید می‌کند که نشان‌دهنده پاسخ نهایی مدل به داده‌های ورودی است. این ساختار به مدل امکان می‌دهد تا از نمایش میانی و فشرده برای محاسبه خروجی نهایی استفاده کند، که چنین امری کارایی تخمین را بهبود می‌بخشد.

در تابع هدف، از خطای میانگین مربعات (MSE) و تنظیم نرم L2 استفاده می‌کنیم تا مجموع خطاهای مربعی را کاهش دهیم و از بیش‌برازش ممانعت به عمل آوریم (رابطه رابطه ۷).

$$\text{Loss} = \frac{1}{N} \sum_{n=1}^N [(y_n - \hat{y}_n)^2] + \lambda \sum_k w_k^2 \quad (\text{رابطه ۷})$$

در رابطه (رابطه ۷)، Loss تابع هزینه، N تعداد نمونه‌ها در داده‌های آموزشی، y_n مقادیر واقعی، \hat{y}_n مقادیر پیش‌بینی شده توسط مدل و λ ضریب تنظیم است که میزان تأثیر تنظیم نرم L2 را کنترل می‌کند؛ w_k^2 وزن‌های مدل هستند که در جریمه کردن برای جلوگیری از بیش‌برازش استفاده می‌شوند. رابطه (رابطه ۷) کمک می‌کند تا با کاهش خطای پیش‌بینی، از پدیده بیش‌برازش مدل جلوگیری شود. در مدل‌سازی طیف عمیق، برای لایه‌های پیچشی و لایه‌های کاملاً متصل از تابع فعال‌سازی خطی Leaky ReLU استفاده می‌شود (Maas et al., 2013). به‌منظور افزایش سرعت آموزش و بهبود دقت، پس از لایه‌های تخت‌کننده و F1، از نرمال‌سازی دسته‌ای استفاده شده است. آموزش مدل با استفاده از روش پس‌زنی و بهینه‌ساز Adam انجام می‌گیرد تا کمینه محلی تابع هدف یافت شود. برای پیشگیری از بیش‌برازش، راه‌برد توقف زودهنگام به کار برده می‌شود. علاوه بر این، برای جلوگیری از مشکلات ناپدید شدن یا انفجار گرادیان، مقیاس‌بندی واریانس به‌عنوان روشی برای مقداردهی اولیه به وزن‌ها انتخاب شده است. این رویکرد به حفظ پایداری عملکرد مدل کمک می‌کند و اطمینان می‌دهد که وزن‌ها در مقادیر مناسبی شروع به یادگیری می‌کنند (Kingma & Ba, 2014; Ioffe & Szegedy, 2015).

مشخصات سخت‌افزاری مورد استفاده برای اجرای این الگوریتم و چارچوب مورد استفاده در الگوریتم، در جدول ۲ نشان داده شده است.

جدول ۲- مشخصات سخت‌افزاری مورد استفاده و عوامل شبکه

مشخصات فنی	بخش سخت‌افزار
Intel Core i7-12700K	واحد پردازش مرکزی
۳۲ گیگابایت	رم
Nvidia RTX 3070 TI	کارت گرافیک
ویندوز ۱۰	سیستم‌عامل
پایتون ۳/۶	زبان برنامه‌نویسی
Scikit-learn و Keras	چارچوب
۸۵ درصد آموزش، ۱۵ درصد آزمون	تقسیم‌بندی داده‌ها
۸ دقیقه	زمان یادگیری

1 .dropout
2 .dense

یافته‌های پژوهشی و بحث

ارزیابی مدل طیف عمیق از طریق معیارهای استاندارد انجام شده، به ما امکان می‌دهد تا دقت و قابلیت تعمیم مدل را درک کنیم. این معیارها شامل خطای جذر میانگین مربعات پیش‌بینی (RMSEP) و ضریب تعیین (R²) هستند، که هر دو برای ارزیابی کیفیت پیش‌بینی‌های مدل بسیار مهم هستند.

خطای جذر میانگین مربعات پیش‌بینی یکی از رایج‌ترین معیارهای ارزیابی در مدل‌سازی پیش‌بینی و بیانگر انحراف استاندارد خطاهای پیش‌بینی است. این معیار به‌طور مستقیم خطاهای پیش‌بینی مدل را نشان می‌دهد و با ارائه یک عدد کلی، تصویر واضحی از عملکرد مدل در مقیاس واقعی فراهم می‌آورد. خطای جذر میانگین مربعات پیش‌بینی پایین نشان‌دهنده دقت بالاتر مدل در پیش‌بینی داده‌هاست و بنابراین، مطلوبیت مدل را تأیید می‌کند.

ضریب تعیین نیز معیاری برای سنجش میزان توانایی مدل در توضیح تغییرات داده‌های وابسته بر اساس متغیرهای مستقل است. از این ضریب معمولاً برای ارزیابی کیفیت مدل‌های رگرسیونی استفاده می‌شود؛ مقدار نزدیک به یک آن نشان‌دهنده این است که مدل تقریباً همه تغییرات متغیر وابسته را توضیح می‌دهد. ضریب تعیین کمتر از یک، نشان می‌دهد که بخشی از داده‌ها توسط مدل توضیح داده نشده‌اند که ممکن است به دلیل وجود نوفه، تعداد ناکافی ویژگی‌ها، یا محدودیت‌های دیگر مدل باشد.

$$RMSEP = \sqrt{\frac{1}{N} \sum_{n=1}^N (y_n - \hat{y}_n)^2} \tag{۸}$$

$$R^2 = 1 - \frac{\sum_{n=1}^N (y_n - \hat{y}_n)^2}{\sum_{n=1}^N (y_n - \bar{y})^2} \tag{۹}$$

در رابطه (۸) y_n مقادیر، \hat{y}_n مقادیر پیش‌بینی شده، \bar{y} میانگین y و N تعداد نمونه آموزش است.

ضریب تغییرات (CV، Coefficient of Variation) در مطالعات کالیبراسیون و ارزیابی مدل‌های پیش‌بینی برای سنجش تکرارپذیری و ثبات مدل بین داده‌های مختلف به کار می‌رود.

$$CV = \frac{\sqrt{\frac{1}{R} \sum_{n=1}^R (RMSEP_n - RMSEP_{avg})^2}}{RMSEP_{avg}} \tag{۱۰}$$

R تعداد تکرار نمونه است که در مطالعه ما این تعداد، ۲۵ است. $RMSEP_n$ ، خطای میانگین مربعات پیش‌بینی برای تکرار n ام و $RMSEP_{avg}$ ، میانگین خطای میانگین مربعات پیش‌بینی برای ۲۵ بار تکرار است.

مجموعه داده‌های مورد استفاده در این مدل‌سازی، ۳۶۰ نمونه است که طیفی از ۴۰۰ تا ۱۸۰۰ هرتز با فاصله ۱ هرتز (معادل ۱۲۰۰ کانال) را پوشش می‌دهد. از این میان، ۲۸۸ نمونه برای آموزش و ۷۲ نمونه برای آزمون استفاده شده‌اند (جدول ۳).

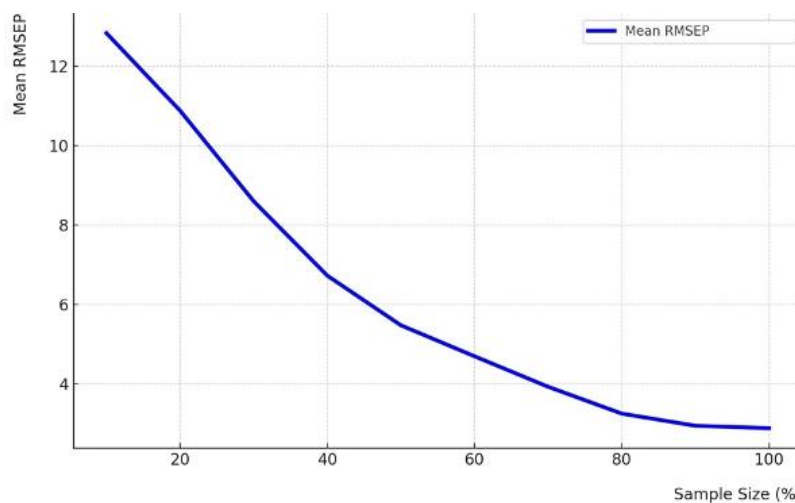
جدول ۳- مجموعه داده‌های توده برنج و نتایج مدل طیف عمیق

نوع داده	پرشدهی توده برنج درصد
تعداد کل نمونه‌ها	۳۶۰
تعداد نمونه‌های آموزش	۲۸۸
تعداد نمونه‌های آزمون	۷۲
تعداد ویژگی‌ها	۱۲۰۰
حداقل	۰
حداکثر	۱۰۰
میانگین	۷۲
انحراف معیار	۲۶/۹۴
RMSEP (Mean ± Std)	۰/۲۴ ± ۰/۰۵
R2 (Mean ± Std)	۰/۹۵ ± ۰/۰۲
CV	۵/۰۵

نتایج مدل طیف عمیق ضریب تعیین، نزدیک به ۱ و خطای جذر میانگین مربعات پیش‌بینی نسبتاً پایین بود که نشان از عملکرد درست مدل دارد. مدل طیف عمیق به دلیل ساختار منحصر به فرد شبکه عصبی آن در لایه‌های Conv2 و Conv3، قابل توجه است. این ساختار امکان استفاده از هم‌گشت‌های موازی با اندازه‌های فیلتر مختلف در لایه Conv3 را فراهم می‌کند که به مدل اجازه می‌دهد تا هم ویژگی‌های محلی و هم ویژگی‌های با سطح انتزاعی بالا را در یک لایه شناسایی کند. هم‌گشت‌های ۱ در ۱ و عملیات ادغام در لایه Conv2 تعداد متغیرها را کاهش می‌دهد تا مشکلات محاسباتی را کمتر کند. این ویژگی‌های منحصر به فرد ساختار شبکه به جمع‌آوری ویژگی‌های مهم از داده‌های ورودی کمک می‌کند و از بیش‌برازش احتمالی جلوگیری می‌کند.

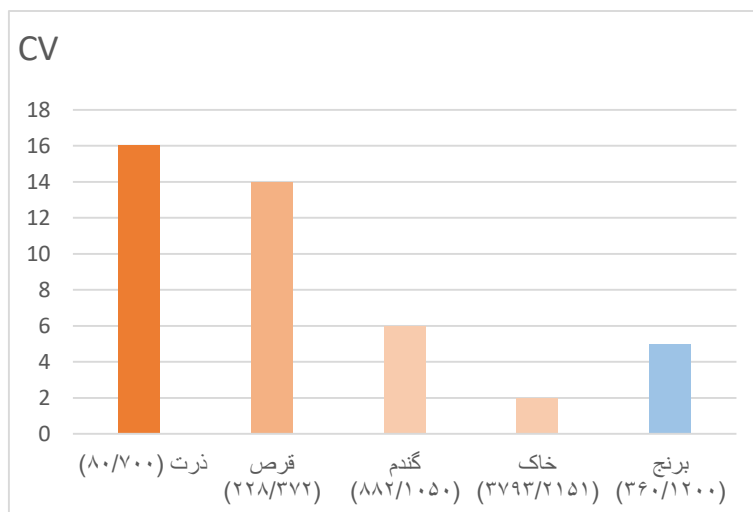
طراحی ساختار مدل نیازمند تعادل برای جلوگیری از بیش‌برازش و کم‌برازش است. در مقایسه با مدل‌های دیگر شبکه عصبی عمیق، طیف عمیق از هم‌گشت‌های ۱ در ۱ و عملیات استخراج در یک لایه واحد درون ماژول آغازین استفاده می‌کند، که این طراحی به کاهش از دست دادن اطلاعات ناشی از عملیات استخراج کمک کرده و در نتیجه اطلاعات مفید بیشتری را در تحلیل مدل، حفظ می‌کند (Zhang *et al.*, 2019; Szegedy *et al.*, 2015). این ساختار، کاهش اطلاعات را در حین عملیات استخراج به حداقل می‌رساند و با کم کردن وزن‌ها، از بیش‌برازش ممانعت به عمل می‌آورد؛ همین امر منجر به هماهنگی بهتر با چهار مجموعه داده طیفی می‌شود. نتایج این تحقیق مزیت ساختار آغازین را با افزایش عرض و عمق، نسبت به شبکه‌های عصبی عمیق سنتی در تحلیل طیفی یک‌بعدی، مخصوصاً در مجموعه داده‌هایی با تعداد نمونه و ویژگی‌های طیفی کم، مانند مجموعه داده‌های طیف صوتی توده برنج، به خوبی نمایان می‌کند. با این حال مدل طیف عمیق نسبت به دیگر مدل‌های شبکه عصبی عمیق و سنتی، زمان بیشتری برای یادگیری نیاز دارد که با توجه به پیشرفت‌های سخت‌افزاری و ارائه راهکارهای پردازش ابری قابل چشم‌پوشی است. شکل ۵ نشان‌دهنده عملکرد مدل DeepSpectra با تغییر اندازه نمونه‌ها است. در این نمودار، محور افقی نشان‌دهنده درصد نمونه‌های استفاده شده از مجموعه آموزشی اصلی و محور عمودی نشان‌دهنده RMSEP است. همان‌طور که مشاهده می‌شود، با افزایش تعداد نمونه‌ها، مقدار RMSEP کاهش یافته و مدل به تدریج پایدارتر می‌شود. پس از استفاده از حدود ۸۰ درصد نمونه‌ها، بهبود قابل توجهی در دقت مدل مشاهده نمی‌شود.

مدل طیف عمیق به دلیل استفاده از روش‌های بهینه‌سازی تصادفی، در هر بار اجرا بر روی یک مجموعه اطلاعات، نتایج متفاوتی ارائه می‌دهد. میانگین و انحراف معیار خطای جذر میانگین مربعات پیش‌بینی از ۲۵ بار اجرا، در جدول ۳ نشان داده شده است. با اینکه عوامل فرعی ثابت می‌مانند، ولی وزن‌های دانه‌های آموزش دیده ممکن است در اجراهای مختلف مدل به دلیل عوامل تصادفی مانند حذف تصادفی برخی وزن‌ها و فرایند بهینه‌سازی تصادفی تغییر کنند. استفاده از نمونه‌های بیشتر آموزشی باعث بهبود قابلیت تکرارپذیری این روش می‌شود و در نتیجه، ضریب تغییرات کمتری را نشان می‌دهد. به عنوان مثال، ضریب تغییرات خطای جذر میانگین مربعات پیش‌بینی برای مجموعه توده برنج با ۲۸۸ نمونه آموزشی، ۵/۰۵ درصد است، در حالی که در مطالعات دیگر با تعداد نمونه‌های آموزشی بیشتر، ۲/۳۷ درصد گزارش شده است (X. Zhang *et al.*, 2019).



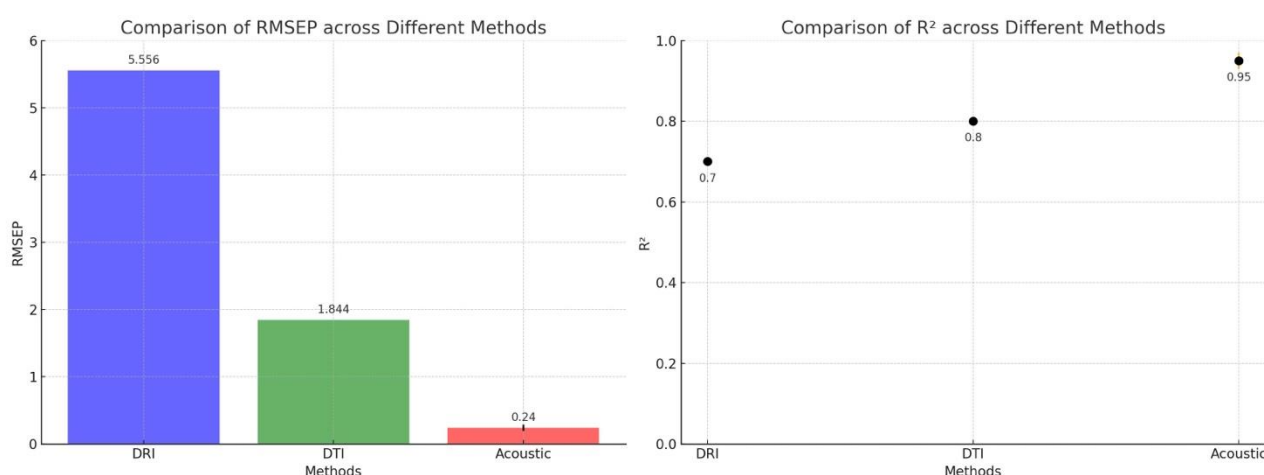
شکل ۵ - عملکرد مدل DeepSpectra با ۲۵ اجرای تکراری و اندازه نمونه‌ای بین ۱۰٪ تا ۱۰۰٪ که به‌طور تصادفی از مجموعه آموزشی اصلی انتخاب شده است.

شکل ۶ نشان‌دهنده ضریب تغییرات مدل DeepSpectra چهارمجموعه داده (ذرت، تپلت، گندم و خاک) (X. Zhang et al., 2019) و برنج (تحقیق حاضر) می‌باشد. همان طور که در شکل مشاهده می‌شود، ضریب تغییرات برای داده‌های خاک با ۳۷۹۳ نمونه و ۲۱۵۱ ویژگی کمترین مقدار (۲/۳۷ درصد) ضریب تغییرات و برای داده‌های ذرت با ۸۰ نمونه و ۷۰۰ ویژگی بیشترین مقدار ضریب تغییرات (۱۶/۶۷ درصد) را دارد. این نمودار نشان‌دهنده بهبود پایداری و تکرارپذیری مدل با افزایش تعداد نمونه‌ها است.



شکل ۶ - مقایسه ضریب تغییرات مدل DeepSpectra بر روی چهار مجموعه داده و تحقیق حاضر. (ویژگی /تعداد نمونه)

نتایج نشان می‌دهد که روش طیف عمیق مبتنی بر حسگر صوتی دارای دقت بالاتری نسبت به روش‌های DRI (تصاویر RGB) و DTI (تصاویر حرارتی) است (Kumar et al., 2020). در شکل ۷، روش DRI دارای مقدار ۵/۵۶، روش DTI دارای مقدار ۱/۸۴، و روش طیف عمیق مبتنی بر حسگر صوتی دارای مقدار ۰/۲۴ است. در نمودار R^2 ، روش DRI دارای مقدار ۰/۷، روش DTI دارای مقدار ۰/۸، و روش طیف عمیق مبتنی بر حسگر صوتی دارای مقدار ۰/۹۵ است. این اعداد نشان می‌دهند که روش طیف عمیق مبتنی بر حسگر صوتی دارای خطای کمتری (RMSEP) و دقت بالاتری (R^2) است و عملکرد بهتری در پیش‌بینی درصد دانه‌های پر شده برنج نسبت به روش‌های دیگر دارد.



شکل ۷ - مقایسه RMSEP (چپ) و R^2 (راست) بین روش‌های DRI، DTI، و طیف جذب صوتی برای پیش‌بینی درصد دانه‌های پر شده برنج.

نتیجه‌گیری

قیمت و کیفیت تجاری شلتوک برنج به طور کلی به اصالت نژادی و درصد پرشدگی دانه‌های آن وابسته است. درصد دانه‌های پر شده مرحله‌ای مهم برای انتخاب محصول با کیفیت و قیمت گذاری محصول است. روش‌های سنتی برای شناسایی نوع درصد دانه‌های پر شده



برنج زمان بر، گران قیمت و تخریبی هستند. اندازه‌گیری طیف ضریب جذب صوتی توده شلتوک به عنوان روشی سریع و غیرتخریبی جدید، به نظر می‌آید برای ارزیابی کیفیت دانه‌ها مناسب باشد. علاوه بر این، مدل طیف عمیق دارای قابلیت بهتری برای تمایز دقیق بین طیف‌های استخراج شده از دانه‌های انواع مختلف برنج است.

طیف ضریب جذب صوتی (α) انواع مختلف دانه‌های توده برنج ارتباط معنی‌داری با پرشدگی دانه‌ها دارد که با مدل‌سازی طیف عمیق این ارتباط پیش‌بینی‌پذیر گردید. نتایج نشان دادند که روش ارائه شده طیف ضریب جذب صوتی به همراه مدل تحلیل طیف عمیق، می‌تواند روشی خوب و سریع برای شناسایی درصد پرشدگی دانه‌های برنج به صورت دقیق و غیرتخریبی باشد. نتایج نشان داد که ساختار اجرایی تغییر یافته لوله امپدانس جهت اندازه‌گیری ضریب جذب توده برنج با پوست همراه با روش یادگیری طیف عمیق می‌تواند راهی قابل قبول برای شناسایی سریع، دقیق و غیر مخرب در صد دانه‌های پر شده در انواع برنج باشد. مطالعات آینده بر روی دیگر دانه‌های غلات نظیر گندم و جو با نمونه‌های بیشتر، می‌تواند باعث افزایش دقت و گستردگی این نتایج شوند.

نتایج این مطالعه می‌تواند به عنوان مرجع و پایه نظری برای توسعه سامانه بازرسی در زمان واقعی جهت آزمون اصالت و کیفیت شلتوک برنج به صورت سریع، دقیق و غیرتخریبی باشد.

منابع

- A|Lab - File Exchange - MATLAB Central <https://www.mathworks.com/matlabcentral/fileexchange/93180-a-lab>.
- Allegrini, F. & Olivieri, A. C. (2016). *Anal Chem* 88, 7807–7812.
- Boubel, A., Boussine, S., Garoum, M., Abdelmajid, B. & Laaroussi, N. (2016).
- Burlando, B. & Cornara, L. (2014). *Trends Food Sci Technol* 40, 82–98.
- Caballol, D. & Raposo, Á. P. (2020). *Constr Build Mater* 256, 119297.
- Ceyhan, M., Kartal, Y., Özkan, K. & Seke, E. (2023). *Multimed Tools Appl* 1–23.
- Duan, L., Yang, W., Bi, K., Chen, S., Luo, Q. & Liu, Q. (2011). *Comput Electron Agric* 75, 196–203.
- Gasso-Tortajada, V., Ward, A. J., Mansur, H., Bröchner, T., Sørensen, C. G. & Green, O. (2010). *Sensors (Basel)* 10, 10027–10039.
- Geimer, P. R., Ulrich, T. J., Luke, J., Beardslee, B., Hayne, M. L., Remillieux, M. C., Saleh, T. A., Freibert, F. J., Ulrich, T. J. & Beardslee, L. B. (2022). *J Acoust Soc Am* 151, 3633–3640.
- Gheibi, A. & Hedayat, A. (2018). *Ultrasonics* 87, 112–125.
- Griffiths, S., Nennig, B. & Job, S. (2017). *J Acoust Soc Am* 141, 254–264.
- Hastie, T., Tibshirani, R. & Friedman, J. (2009). <https://doi.org/10.1007/978-0-387-84858-7>.
- Huang, S., Lu, Z., Shi, Y., Dong, J., Hu, L., Yang, W., Huang, C., Huang, S., Lu, Z., Shi, Y., Dong, J., Hu, L., Yang, W. & Huang, C. (2023). *Sensors* 2023, Vol. 23, Page 6331 23, 6331.
- Ioffe, S. & Szegedy, C. (2015). 32nd International Conference on Machine Learning, ICML 2015 1, 448–456.
- ISO 10534-1:1996 - Acoustics — Determination of sound absorption coefficient and impedance in impedance tubes — Part 1: Method using standing wave ratio <https://www.iso.org/standard/18603.html>.
- ISO 10534-2:2023 - Acoustics — Determination of acoustic properties in impedance tubes — Part 2: Two-microphone technique for normal sound absorption coefficient and normal surface impedance <https://www.iso.org/standard/81294.html>.
- Kamboj, U., Guha, P. & Mishra, S. (2021). *Mater Today Proc* 48, 576–582.
- Kingma, D. P. & Ba, J. L. (2014). 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.
- Koulivand, K. & Moniri, | Kamel.
- Kumar, A., Taparia, M., Madapu, A., Rajalakshmi, P., Marathi, B. & Desai, U. B. (2020). *J Cereal Sci* 95, 103037.
- Lecun, Y., Bengio, Y. & Hinton, G. (2015). *Nature* 2015 521:7553 521, 436–444.
- Lee, C. M. & Wang, Y. S. (2006). *J Sound Vib* 298, 350–365.
- Lee, C. M. & Xu, Y. (2009). *J Sound Vib* 326, 290–301.
- Li, Z., Chen, S., Wu, Z. & Yang, L. (2020). <https://doi.org/10.1177/1077546320926905> 27, 332–342.
- Lin, P., Chen, Y. M., He, Y. & Hu, G. W. (2014). *Comput Electron Agric* 109, 124–133.
- Liu, T., Wu, W., Chen, W., Sun, C., Chen, C., Wang, R., Zhu, X. & Guo, W. (2016). *Biosyst Eng* 150, 79–88.
- Liu, X., An, H., Cai, W. & Shao, X. (2024a). *TrAC Trends in Analytical Chemistry* 172, 117612.
- Liu, X., An, H., Cai, W. & Shao, X. (2024b). *TrAC Trends in Analytical Chemistry* 172, 117612.

- Maas, A. L., Hannun, A. Y. & Ng, A. Y. (2013).
- Manickavasagan, A., Sathya, G., Jayas, D. S. & White, N. D. G. (2008). *J Cereal Sci* 47, 518–527.
- Mosa, A. I., Putra, A., Ramlan, R., Prasetyo, I. & Esraa, A. A. (2019). *Applied Acoustics* 146, 409–419.
- Muthayya, S., Sugimoto, J. D., Montgomery, S. & Maberly, G. F. (2014). *Ann NY Acad Sci* 1324, 7–14.
- Nguyên, T. T. T., Doanh, T., Bot, A. Le & Dalmas, D. (2022). *Scientific Reports* 2022 12:1 12, 1–9.
- Ni, W., Nørgaard, L. & Mørup, M. (2014). *Anal Chim Acta* 813, 1–14.
- Ostrovsky, L., Lebedev, A., Matveyev, A., Potapov, A., Sutin, A., Soustova, I. & Johnson, P. (2001). *J Acoust Soc Am* 110, 1770–1777.
- Passos, D. & Mishra, P. (2022). *Chemometrics and Intelligent Laboratory Systems* 223, 104520.
- Rinnan, Å., Berg, F. van den & Engelsen, S. B. (2009). *TrAC Trends in Analytical Chemistry* 28, 1201–1222.
- Roozen, N. B. & Piana, E. A. (2022). *Applied Acoustics* 198, 109002.
- Stender, M., Adams, C., Wedler, M., Grebel, A. & Hoffmann, N. (2021). *J Acoust Soc Am* 149, 1932–1945.
- Sun, M., Zhang, D., Liu, L. & Wang, Z. (2017). *Food Chem* 218, 413–421.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. & Rabinovich, A. (2015). 1–9.
- Tsuruha, T., Yamada, Y., Otani, M. & Takano, Y. (2020). *J Acoust Soc Am* 147, 3418–3428.
- Yang, J., Peng, S., Visperas, R. M., Sanico, A. L., Zhu, Q. & Gu, S. (2000). *Plant Growth Regul* 30, 261–270.
- Zhang, J., Dai, L. & Cheng, F. (2021a). *Journal of Food Measurement and Characterization* 15, 484–494.
- Zhang, J., Dai, L. & Cheng, F. (2021b). *Food Anal Methods* 14, 389.
- Zhang, X., Lin, T., Xu, J., Luo, X. & Ying, Y. (2019). *Anal Chim Acta* 1058, 48–57.



New acoustic sensing approach for predicting the percentage of filled rice grains based on the acoustic absorption spectrum using the Deep Spectra

EXTENDED ABSTRACT

Introduction

Rice serves as a staple food for a substantial portion of the global population, particularly in Asian countries. Traditional methods of determining the percentage of filled grains (PFG) in rice are labor-intensive and often rely on subjective judgments. This study introduces an innovative and non-destructive approach that leverages a new acoustic sensing technology combined with deep learning models. The primary innovation of this approach is its use of the acoustic absorption spectrum to predict the PFG without the need for pre-processing the spectral data. This method holds significant promise for enhancing the efficiency and accuracy of rice quality assessments, crucial for agricultural productivity and research in rice breeding and genetics. By employing advanced deep learning architectures, this study aims to overcome the limitations of traditional grain analysis methods, offering a more reliable and scalable solution.

Material and Methods

The research utilizes an altered impedance tube to capture the acoustic spectrum of rice grains. This method measures how sound waves are absorbed by the grains, providing a detailed spectral profile. The acoustic data collected is then analyzed using a deep learning model, specifically designed for this study. The model, referred to as DeepSpectra, operates on the raw spectral data directly, eliminating the need for the usual pre-processing steps that can degrade data quality. This model employs a convolutional neural network (CNN) architecture, which is highly effective in pattern recognition tasks. CNNs utilize layers of filters to process spatial hierarchies in data, making them ideal for handling complex spectral data. This approach not only enhances the prediction accuracy but also significantly speeds up the analysis process by leveraging the model's ability to learn and generalize from the data presented.

Results and Discussion

The results of this study were highly promising, demonstrating that the deep learning model, DeepSpectra, significantly outperformed traditional methods in predicting the percentage of filled grains in rice. The model achieved high accuracy, showcasing its capability to handle the raw spectral data effectively without the need for manual preprocessing. This is a considerable advancement over traditional grain analysis methods, which are slower and often less reliable. The discussion also highlights how the model's deep learning architecture, which includes multiple convolutional and pooling layers, contributes to its robust performance by effectively extracting and learning important features from complex datasets. The study further discusses the potential applications of this technology in the field of agricultural research and grain quality assessment, suggesting that it could be a valuable tool for rice breeders and geneticists looking to improve crop yields and grain quality.

Conclusion

The study concludes that the novel use of acoustic sensing technology combined with a deep learning analysis model offers a powerful and non-destructive method for accurately predicting the percentage of filled grains in rice. The acoustic absorption spectrum, as analyzed by the DeepSpectra model, provides a reliable indicator of grain quality, which is essential for rice breeding and genetic research. The implications of these findings are significant, as they suggest a shift towards more automated, precise, and efficient methods of grain analysis. This could lead to improvements in rice production strategies and better food security globally. Additionally, the study underscores the potential for expanding this technology to other grains and agricultural products, which could broadly revolutionize quality assessment in agriculture. The researchers highlight the need for further studies to refine the technology and fully realize its potential across different environments and grain types.