International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: https://doi.org/10.22214/ijraset.2026.82986

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Data-Driven Clinical Decision Support Framework for Automated Pneumonia Detection

Tejonidhi M R[1], Nandan H S[2], Poorab N G[3], Kushal H S[4], Poornachandra N[5]
Department of Computer Science and Engineering Malnad College of Engineering, Karnataka, India

Abstract: Pneumonia remains one of the leading respiratory diseases responsible for high mortality worldwide, particularly among children, elderly individuals, and immunocompromised patients. Chest X-ray imaging is widely used for diagnosis because of its affordability and accessibility; however, manual interpretation is time-consuming and highly dependent on radiological expertise. Recent advancements in Artificial Intelligence (AI) and deep learning have enabled automated pneumonia detection systems capable of assisting clinicians in medical image analysis.

This paper presents a comprehensive survey and a proposed Data-Driven Clinical Decision Support Framework for auto-mated pneumonia detection using chest X-ray images. Existing Convolutional Neural Network (CNN) architectures demonstrate strong local feature extraction capabilities but often fail to capture long-range contextual dependencies. Transformer-based architectures improve contextual understanding but introduce high computational complexity and increased dataset requirements. Recent hybrid CNN–Transformer frameworks attempt to combine both local and global feature learning for improved diagnostic performance.

Based on the identified research gaps, a hybrid framework integrating DenseNet, Swin Transformer, and CBAM attention mechanisms is proposed. The framework also incorporates Ex-plainable Artificial Intelligence (XAI) techniques such as Grad-CAM and LIME to improve interpretability and physician trust. Furthermore, a Large Language Model (LLM)-based reporting module is introduced to generate automated clinical summaries and diagnostic explanations. The proposed framework aims to improve prediction accuracy and contextual understanding while enhancing interpretability and clinical assistance. The work focuses on integrating hybrid deep learning, Explainable AI, and LLM-based clinical reasoning into a unified clinical decision support framework.

Index Terms: Pneumonia Detection, Chest X-ray Analysis, Clinical Decision Support System, DenseNet, Swin Transformer, Explainable AI, Grad-CAM, LIME, Hybrid Deep Learning, Medical Imaging.

I. INTRODUCTION

Pneumonia is a severe respiratory infection that affects the lungs and remains one of the leading causes of mortality worldwide. The disease primarily affects children, elderly individuals, and patients with weakened immune systems.

Common symptoms include fever, cough, breathing difficulty, chest pain, and inflammation of lung tissues. Delayed diag-nosis may lead to severe complications including respiratory failure and increased mortality [6]. Therefore, early and ac-curate diagnosis plays an important role in improving patient outcomes and reducing healthcare burden.

Chest X-ray imaging is one of the most widely used diagnostic techniques for pneumonia detection because of its affordability, accessibility, and effectiveness in identifying pulmonary abnormalities. Radiologists generally analyze chest radiographs for signs such as infiltrates, opacities, and con-solidation regions. However, manual interpretation is highly dependent on clinical expertise and may result in diagnostic inconsistencies. In many healthcare environments, especially in rural and resource-constrained regions, the shortage of ex-perienced radiologists further increases diagnostic challenges. Recent advancements in Artificial Intelligence (AI), Ma-chine Learning (ML), and Deep Learning (DL) have sig-nificantly improved automated medical image analysis. AI-assisted systems can support physicians by improving diag-nostic speed, consistency, and accuracy. Among deep learning approaches, Convolutional Neural Networks (CNNs) have demonstrated remarkable success in medical image classifi-cation tasks because of their ability to automatically learn hierarchical spatial representations from images.

CNN architectures such as VGG16, EfficientNet, ResNet, and DenseNet have achieved strong performance in chest X-ray analysis. DenseNet architectures improve feature reuse and gradient propagation through dense connectivity between layers, making them highly suitable for pneumonia classifi-cation [1]. However, CNN-based systems mainly focus on local spatial features and often struggle to capture long-range contextual relationships across the image.

To overcome these limitations, researchers introduced transformer-based architectures capable of modeling contex-tual relationships using self-attention mechanisms. Vision Transformers divide images into patches and process them using self-attention operations capable of learning global image relationships [7]. Swin Transformer architectures fur-ther improve efficiency by introducing shifted-window self-attention mechanisms that reduce computational complexity while preserving contextual understanding [8].

Despite these advancements, transformer architectures re-quire large-scale datasets and substantial computational re-sources. Consequently, recent studies increasingly focus on hybrid CNN–Transformer architectures that combine local feature extraction from CNNs with contextual reasoning from transformers. The proposed work focuses on integrating ex-isting deep learning, explainability, and clinical reasoning techniques into a unified clinical decision support framework. Attention mechanisms such as CBAM and SE modules have also been integrated into deep learning architectures to improve feature refinement and disease localization [2], [3]. Furthermore, Explainable Artificial Intelligence (XAI) techniques including Grad-CAM and LIME have become important for improving transparency and physician trust in healthcare AI systems [4], [5].

However, most existing systems focus primarily on classifi-cation accuracy and fail to provide integrated clinical reason-ing or reporting support. Motivated by these limitations, this paper presents a comprehensive survey and proposed frame-work for a hybrid clinical decision support system integrating DenseNet, Swin Transformer, CBAM attention mechanisms, Explainable AI techniques, and LLM-based report generation for automated pneumonia detection.

II. LITERATURE SURVEY

Automated pneumonia detection using chest X-ray images has evolved significantly with advancements in deep learn-ing and medical image analysis. Earlier systems relied on handcrafted feature extraction and traditional machine learning classifiers, whereas recent approaches increasingly integrate CNNs, transformers, attention mechanisms, and Explainable AI techniques.

A. CNN-Based Pneumonia Detection

Bundea and Danciu [1] investigated DenseNet architec-tures including DenseNet121, DenseNet169, and DenseNet201 for pneumonia classification. Their study demonstrated that DenseNet architectures improve feature reuse and gradient propagation while achieving strong classification performance. Sharma and Guleria [11] proposed a VGG16-based trans-fer learning framework for automated pneumonia detection.

Their work demonstrated the effectiveness of pretrained CNN models in healthcare applications.

Darapaneni et al. [9] explored transfer learning using Ef-ficientNet and Xception architectures for chest X-ray classi-fication. Their work highlighted the importance of pretrained feature extraction in improving performance on limited med-ical imaging datasets.

Although CNN architectures achieved strong classification accuracy, they mainly focus on local receptive fields and often fail to capture global contextual relationships.

B. Attention-Based Deep Learning Approaches

Rashid et al. [2] integrated DenseNet121 with CBAM at-tention mechanisms for pneumonia detection. Their framework improved focus on infected pulmonary regions and achieved better classification performance compared to baseline CNN architectures.

Potharaju et al. [3] incorporated CBAM and Squeeze-and-Excitation (SE) attention mechanisms into deep learning models for chest X-ray analysis. Their work improved feature discrimination and enhanced detection of subtle pulmonary abnormalities.

Slimi et al. [10] proposed an attention-guided deep learning framework for trustworthy pneumonia detection. Their work highlighted the importance of attention mechanisms in improv-ing robustness and interpretability.

C. Vision Transformer-Based Architectures

Dosovitskiy et al. [7] introduced the Vision Transformer (ViT), which processes image patches using transformer-based self-attention operations. Unlike CNNs, ViTs effectively cap-ture global contextual relationships.

Liu et al. [8] proposed the Swin Transformer architecture using hierarchical shifted-window attention mechanisms. Swin Transformers significantly reduced computational complexity while preserving contextual learning capabilities.

Singh et al. [13] applied Vision Transformers for pneumonia classification using chest X-ray images and demonstrated superior contextual learning compared to traditional CNN architectures.

D. Hybrid CNN–Transformer Architectures

Mustapha et al. [14] proposed a hybrid CNN-Swin Trans-former framework for pneumonia detection. Their architecture combined

CNN-based local feature learning with transformer-based contextual reasoning and achieved improved classifica-tion performance.
Qezelbash-Chamak and Hicklin [12] introduced a ConvNeXt-Swin Transformer fusion framework capable of capturing both local texture information and contextual dependencies.
Alhumaid and Fayoumi [15] developed a hybrid CNN-Swin Transformer framework integrated with Explainable AI techniques. Their work improved both interpretability and diagnostic performance through Grad-CAM visualizations.

E. Explainable AI in Pneumonia Detection

Ifty et al. [4] proposed an explainable lung disease classifi-cation framework integrating deep learning models with Grad-CAM and LIME techniques.
Khadidos et al. [5] implemented Explainable AI techniques for pediatric pneumonia detection using chest X-ray images. Their framework improved physician trust by highlighting infected pulmonary regions responsible for predictions.

F. Research Gaps Identified

The literature survey reveals several important limitations:
1) CNN architectures lack efficient global contextual under-standing.
2) Transformer models require high computational re-sources.
3) Many systems lack integrated Explainable AI frame-works.
4) Most approaches focus only on classification accuracy.
5) Automated clinical reporting support remains limited.

III. PROPOSED FRAMEWORK

The proposed framework integrates DenseNet, Swin Trans-former, CBAM attention mechanisms, Explainable AI tech-niques, and LLM-based report generation into a unified clinical decision support system.

A. System Overview

The framework consists of the following stages:
- Chest X-ray Image Acquisition
- Image Preprocessing
- Hybrid CNN–Transformer Classification
- Explainability Generation
- LLM-Based Clinical Report Generation
- Final Diagnosis Output

The workflow begins with chest X-ray image acquisition followed by preprocessing operations such as resizing, nor-malization, contrast enhancement, and data augmentation. The processed images are then passed through a hybrid CNN–Transformer architecture consisting of DenseNet, Swin Trans-former, and CBAM attention modules.
DenseNet extracts local pulmonary features while Swin Transformer captures global contextual relationships across the image. DenseNet was selected because of its efficient feature reuse and strong performance in medical image analysis tasks. CBAM attention modules refine feature representations by emphasizing diagnostically important regions.
The extracted features are fused and passed through fully connected layers for final classification. Explainable AI tech-niques such as Grad-CAM and LIME generate visual explana-tions highlighting infected pulmonary regions responsible for predictions.
Finally, an LLM-based reporting module generates auto-mated clinical summaries and supportive diagnostic explanations.

B. Explainable AI Integration

The proposed framework integrates Grad-CAM and LIME techniques to improve interpretability.
Grad-CAM generates heatmaps highlighting pulmonary re-gions responsible for predictions, while LIME explains local feature contributions toward classification outputs.
Explainable AI improves physician trust, transparency, and reliability in healthcare AI systems.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

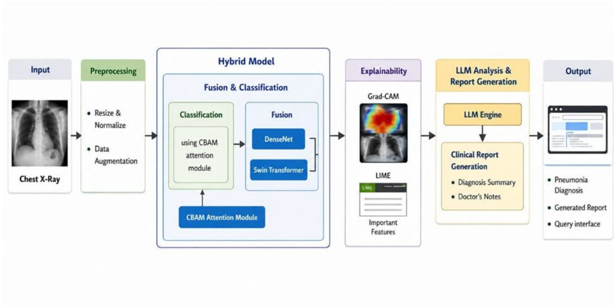


Fig. 1. Proposed Hybrid CNN–Transformer Clinical Decision Support Frame-work

C. LLM-Based Clinical Reporting

An LLM-based reporting module generates automated clin-ical summaries using classification outputs and explainability information.

The module generates:

- Diagnostic summaries
- Clinical explanations
- Physician-supportive observations
- Automated radiology-style reports

IV. SYSTEM ARCHITECTURE

The proposed architecture integrates preprocessing, feature extraction, contextual learning, attention refinement, explain-ability, and automated reporting into a unified pipeline. DenseNet extracts local pulmonary features while Swin Transformer captures contextual relationships. CBAM atten-tion modules refine feature representations before classification.

Grad-CAM and LIME provide visual interpretability, and the LLM module generates clinical summaries and supportive diagnostic explanations.

V. COMPARATIVE ANALYSIS

Comparative analysis helps evaluate the strengths and limi-tations of existing pneumonia detection systems with respect to contextual learning, interpretability, computational complexity, and clinical usability.

The proposed framework combines local feature extraction, contextual learning, explainability, and automated reporting within a unified clinical decision support system.

VI. DISCUSSION

Recent advancements in deep learning have significantly im-proved automated pneumonia detection using chest X-ray im-ages. CNN architectures demonstrated strong performance in extracting local pulmonary features, while transformer-based architectures improved contextual understanding through self-attention mechanisms.

Hybrid CNN–Transformer architectures emerged as highly effective because they combine complementary feature learn-ing capabilities. DenseNet architectures improve local feature

TABLE I
COMPARATIVE ANALYSIS OF EXISTING PNEUMONIA DETECTION APPROACHES

Approach	Local Feature Learning	Global Contextual Learning	Explainability Support	Computational Complexity	Clinical Decision Sup-port Capability
Deep CNN Models	Strong spatial feature extraction using con-volutional operations	Limited contextual understanding due to local receptive fields	Minimal interpretability support	Moderate	Primarily focused on classification prediction
Attention-Based CNN Models	Improved pulmonary feature refinement us-ing	Moderate contextual learning capability	Partial explainabil-ity through attention	Moderate	Limited clinical inter-pretation support

	CBAM and SE modules		visualization		
Vision Transformer Mod-els	Moderate local repre-sentation learning	Strong long-range contextual reasoning using self-attention	Limited interpretability support	High	Focused mainly on im-age classification tasks
Hybrid CNN–Transformer Models	Combined local fea-ture extraction with contextual reasoning	High contextual learning capability	Moderate explainability integration	High	Partial physician-supportive capability
Proposed Hybrid Clinical Decision Support Frame-work	DenseNet-based pul-monary feature ex-traction with CBAM refinement	Swin Transformer-based contextual reasoning using shifted-window attention	Integrated Grad-CAM and LIME visual explanations	Moderate to High	Integrated prediction, explainability, and clinical reporting support

reuse and gradient propagation, whereas Swin Transformers capture long-range contextual relationships efficiently.

Attention mechanisms such as CBAM improve feature refinement and disease localization by emphasizing diagnosti-cally important pulmonary regions. Explainable AI techniques further improve transparency and physician trust by generating visual explanations.

The integration of Large Language Models introduces au-tomated clinical reporting and diagnostic reasoning support, transforming the framework into a comprehensive clinical decision support system.

Despite these advancements, challenges related to computa-tional complexity, dataset limitations, and real-world deploy-ment remain important research areas.

Another important observation from the literature survey is that no single architecture independently satisfies all healthcare requirements. CNN-based systems provide strong spatial fea-ture extraction but often miss broader contextual relationships, while transformer-based models improve contextual reasoning at the cost of higher computational complexity. Hybrid frame-works therefore offer a more balanced solution by combining the strengths of both approaches.

The survey also highlights the growing importance of explainability in healthcare AI systems. Prediction accuracy alone is insufficient in clinical environments unless the system provides understandable reasoning and visual evidence. Tech-niques such as Grad-CAM and LIME improve transparency by helping clinicians verify whether the model focuses on medically relevant pulmonary regions during diagnosis.

VII. CONCLUSION

This paper presented a comprehensive survey and proposed framework for a Data-Driven Clinical Decision Support Sys-tem for automated pneumonia detection using chest X-ray images.

The survey analyzed CNN-based architectures, attention mechanisms, Vision Transformers, Explainable AI tech-niques, and hybrid CNN–Transformer models. Existing studies demonstrated that CNNs provide strong local feature ex-traction capabilities while transformer architectures improve contextual understanding.

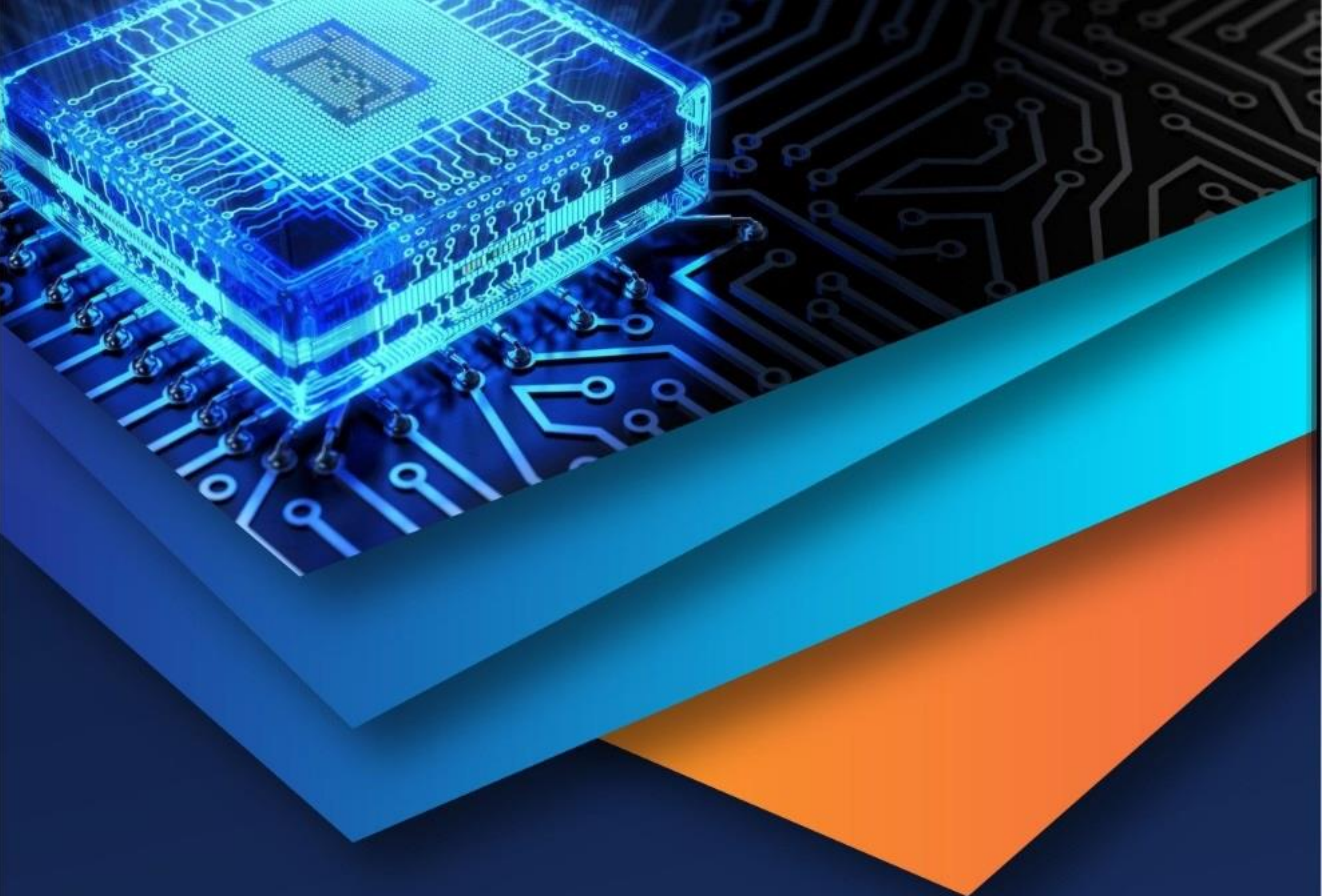
Based on the identified research gaps, a hybrid framework integrating DenseNet, Swin Transformer, CBAM attention mechanisms, Explainable AI techniques, and LLM-based re-porting was proposed.

The proposed framework aims to improve classification ac-curacy, contextual understanding, interpretability, and clinical usability within a unified healthcare-oriented system.

Future work will focus on model evaluation, lightweight deployment, multimodal integration, advanced Explainable AI methods, and real-world clinical validation.

REFERENCES

[1] M. Bundea and G. M. Danciu, "Pneumonia Image Classification Using DenseNet Architecture," *Information*, vol. 15, 2024.
[2] A. B. Rashid et al., "An Enhanced Deep Learning Framework for Pneumonia Detection in Chest X-rays," SN Computer Science, 2025.
[3] S. Potharaju et al., "Enhanced X-ray Image Classification for Pneumonia Detection using DL based CBAM and SE Mechanisms," Intelligence-Based Medicine, 2025.
[4] T. T. Ifty et al., "Explainable Lung Disease Classification from Chest X-Ray Images Utilizing Deep Learning and XAI," IEEE ICMI, 2024.
[5] A. O. Khadidos et al., "Explainable Deep Learning for Pediatric Pneu-monia Detection in Chest X-Ray Images," arXiv preprint, 2026.
[6] W. S. Lim, "Pneumonia: Overview," Elsevier Encyclopedia of Respira-tory Medicine, 2021.
[7] A. Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," ICLR, 2021.
[8] Z. Liu et al., "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," ICCV, 2021.
[9] N. Darapaneni et al., "Pneumonia Detection from Chest X-rays using Neural Networks," arXiv preprint, 2022.
[10] H. Slimi et al., "Trustworthy Pneumonia Detection in Chest X-ray Imaging through Attention-Guided Deep Learning," Scientific Reports, 2025.
[11] S. Sharma and K. Guleria, "A Deep Learning Based Model for Detection of Pneumonia from Chest X-ray images using VGG16 and Neural Networks," Procedia Computer Science, 2023.
[12] J. Qezelbash-Chamak and K. Hicklin, "A Hybrid Learnable Fusion of ConvNeXt and Swin Transformer for Optimized image Classification," IoT Journal, 2025.
[13] S. Singh et al., "Efficient Pneumonia Detection using Vision Transform-ers on Chest X-rays," Scientific Reports, 2024.
[14] B. Mustapha et al., "Enhanced Pneumonia Detection in Chest X-rays using Hybrid Convolutional and Vision Transformer Networks," Current Medical Imaging, 2025.
[15] M. Alhumaid and A. G. Fayoumi, "Hybrid CNN-Swin Transformer Model to Advance the Diagnosis of Maxillary Sinus Abnormalities on CT Images Using Explainable AI," Computers, 2025.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY