



Archives available at journals.mriindia.com

International Journal of Advanced Scientific Research and Engineering Trends

ISSN: 2456-0774

Volume 10 Issue 05, 2026

Machine Learning Methods for Image Understanding: From Basics to Advanced Techniques

¹S. Z. Mullani, ²V. B. Waghmare

¹Assistant Professor, Vivekanand College, Kolhapur, Maharashtra, India

²Assistant Professor, Vivekanand College, Kolhapur, Maharashtra, India

Email: ¹szm.rs.csd@unishivaji.ac.in

Peer Review Information	Abstract
<i>Submission: 17 April 2026</i>	<p>: In today's web world, images are all around World Wide Web every second more and more uses image data – from social networking sites to medical reports, from CCTV cameras to satellite images. It is a huge task to make computers understand and analyze these images. Machine learning, particularly deep learning, has provided us with effective solutions to this problem. This paper provides a brief and comprehensive overview of various machine learning approaches used for image analysis. We begin with the traditional approaches and progress to the most recent deep learning approaches. We describe how these approaches work, their pros and cons, and how they are applied in practical situations. This paper will help students, researchers, and professionals understand the evolution of image analysis approaches and select the appropriate approach according to their requirements.</p>
<i>Revision: 30 April 2026</i>	
<i>Acceptance: 11 May 2026</i>	
Keywords	
<i>Image Analysis, Machine learning, Deep Learning, CNN, Computer Vision, Object Detection, India Applications.</i>	

Introduction

Today, images are used everywhere—from mobile phone cameras to medical scans and satellite images. Helping computers understand these images is a big challenge, and researchers have been working on this problem for many years. In the early days, computers were very simple. They could store images, but they could not understand what was inside them. Humans had to manually tell the computer what to look for, such as edges, corners, or shapes. This process took a lot of time and did not work well when lighting, angle, or background changed [1]. The situation changed with the introduction of deep learning. In deep learning, computers learn features on their own by studying thousands of example images. This is similar to how children learn to identify animals by seeing many pictures. Because of this ability, computers became much better at understanding images.

Image analysis is very important for India. Hospitals can use it to read X-ray images and detect diseases at an early stage. Farmers can check crop health using pictures taken from mobile phones. Cities can manage traffic better by analyzing camera footage. Security systems can identify faces and detect suspicious activities. Machine learning approaches are also widely used for facial expression and emotion understanding. Emotion recognition is important because emotions play a major role in human communication, behavior, and decision-making. With the rapid growth of machine learning (ML) and deep learning (DL), recognizing human emotions from facial images has become more accurate and efficient. Due to all these applications, image analysis has become a key technology for the development of our country [2]. This paper explains these techniques in a simple way so that researchers can understand

them easily and apply them to solve local and real-world problems.

Traditional Methods (Before Deep Learning)

Before deep learning became popular, image analysis mainly used traditional methods with two main steps. The first step was feature extraction, where useful information was taken from images using mathematical formulas designed by humans. For example, the SIFT algorithm identified important points in an image that remained the same even if the image was rotated or resized [3]. Another method, called HOG, analyzed the direction of edges in small parts of an image and was commonly used for detecting people in images [4]. These methods were like giving the computer a fixed checklist of what to look for in every image.

After feature extraction, simple machine learning algorithms were used for classification. Support Vector Machine (SVM) separated different categories by drawing boundaries between them, such as separating cat images from dog images [5]. Random Forest used many decision trees together, where each tree gave a result and the final output was decided by majority voting. These methods worked reasonably well for simple tasks but failed when images had complex backgrounds, different lighting conditions, or varied angles. They required a lot of human effort and were not suitable for real-world image complexity.

Deep Learning Revolution

A major breakthrough happened in 2012 when a deep learning model called AlexNet won an image recognition competition by a large margin [6]. Unlike traditional methods, AlexNet did not rely on manually designed features. Instead, it automatically learned features from millions of images. It used Convolutional Neural Networks (CNNs), which work like a series of intelligent filters. The first layers detect simple features such as edges, the next layers detect corners and shapes, and the final layers recognize complete objects like faces or vehicles. This automatic learning of features greatly improved image understanding.

After AlexNet, many advanced deep learning models were developed. VGGNet used small 3×3 filters throughout the network, which made the design simple but required high computing power [7]. ResNet solved the problem of performance degradation in very deep networks by introducing skip connections that allow information to pass directly across layers [8]. MobileNet was designed for mobile and low-power devices, using efficient operations to reduce computation while maintaining good

accuracy. These improvements made deep learning practical for real-world applications on regular computers and smartphones.

Sudhamani and Kavya S. N. [18] presented a detailed study on facial expression recognition using machine learning techniques. Their work discussed traditional classifiers such as Support Vector Machines (SVM), k-Nearest Neighbour (KNN), and Decision Trees, along with feature extraction methods. They emphasized the importance of preprocessing steps like face detection, image normalization, and noise removal to improve accuracy. Their results showed that traditional machine learning methods perform well for basic facial expressions such as happy, sad, anger, fear, surprise, and neutral, especially when using controlled datasets.

With further advances in deep learning, transfer learning became a popular approach for image understanding. Raja Sekaran et al. [19] studied facial emotion recognition using transfer learning with AlexNet, a pre-trained CNN model. Their research showed that transfer learning improves accuracy while reducing training time and computational cost. This method is especially useful when only small datasets are available, which is common in facial emotion recognition tasks. Their study clearly shows the shift from traditional machine learning to advanced deep learning-based approaches.

Emotion recognition is also strongly connected to psychological theories. Lerner et al. [20] explained how emotions influence human thinking, judgement, and decision-making. Their work highlights the importance of accurately identifying emotional states. This psychological background supports the need for reliable image-based emotion recognition systems, especially in applications such as human-computer interaction, healthcare monitoring, sentiment analysis, and behavioral studies.

Further research by Galal Ahmed Ghaleb Abdo, Almaghrebi M., Mohammad Reza Sharifi, and Research Scholar Jagadeesh M. [21] focused on combining machine learning and deep learning techniques to improve facial emotion recognition systems. Their work emphasized better efficiency, robustness, adaptability, and real-time performance.

In summary, the literature clearly shows the evolution from traditional feature-based machine learning methods to deep learning and transfer learning approaches for facial emotion recognition. Facial emotion recognition is a strong example of how image understanding has progressed into intelligent, data-driven systems and provides a solid foundation for the research presented in this paper.

Table 1: Image Analysis Techniques - Traditional vs Deep Learning Approaches

Aspect	Traditional Methods	Deep Learning Methods	Best For
Approach	Handcrafted features + simple classifiers	Automatic feature learning + complex networks	Complex real-world images
Feature Extraction	Humans design features (SIFT, HOG, LBP)	Network learns features automatically from data	When you don't know what features to look for
Accuracy	Low to medium (60-80% for simple tasks)	High (90-99% for many tasks)	Critical applications like medical diagnosis
Data Needed	Can work with smaller datasets	Needs large labeled datasets (thousands of images)	When you have lots of labeled data
Computing Power	Works on normal computers/laptops	Needs GPUs for training, can run on CPUs for inference	Organizations with good computing resources
Training Time	Minutes to hours	Hours to days/weeks	When time for training is not critical
Interpretability	Easy to understand (features are human-designed)	Hard to understand (black box problem)	When you need to explain decisions
Adaptability	Need to redesign features for new problems	Can adapt to new problems with fine-tuning	Multiple related applications

Table2: Different Deep Learning Architectures

Architecture	Key Idea	Advantages	Limitations	Indian Context Suitability
AlexNet	First successful deep CNN	Started deep learning revolution	Old, not efficient now	Learning/teaching purposes
VGGNet	Simple 3×3 filters throughout	Easy to understand and implement	Needs lot of computation	Research projects with good GPUs
ResNet	Skip connections for deep networks	Can train very deep networks (100+ layers)	More complex to implement	Advanced applications needing high accuracy
MobileNet	Depthwise separable convolutions	Lightweight, good for mobile phones	Slightly lower accuracy	Mobile apps for Indian farmers/doctors
YOLO	Single pass for object detection	Very fast, real-time processing	Less accurate for small objects	Traffic monitoring, security cameras
U-Net	U-shaped for segmentation	Excellent for medical image segmentation	Mainly for segmentation tasks	Medical imaging in Indian hospitals
Vision Transformer	Treats image as sequence of patches	State-of-art for some tasks	Needs huge data, computing power	Large Indian companies/research labs

For simple classification tasks (that is, finding **what is in the image**), standard machine learning methods like CNNs work well. The computer looks at the full image and gives only one label, such as “cat” or “dog.” However, for object detection tasks (that is, **where the objects are in the image**), more advanced methods are required. YOLO (You Only Look Once) is very

popular because it is fast. It looks at the image only once and detects all objects in real time, which is very useful for video analysis [9]. R-CNN based methods are more accurate but slower, so they are suitable for applications where accuracy is more important than speed.

For medical images, where we need to identify the exact area affected by a disease,

segmentation techniques are used. U-Net is specially designed for medical imaging and can accurately outline tumors or infections in X-rays and scans [10]. For generating new images from text descriptions (for example, “generate an image of a temple at sunset”), GANs and diffusion models are used. These models can create realistic images that do not exist in real life and are useful for design, media, and entertainment. Different image-related tasks need different techniques, and researchers have developed many specialized tools to handle these requirements.

Nowadays, the sharing of photos and videos on digital platforms is increasing very fast. Because of this, it has become important to study human emotions and feelings from such shared content. Image Sentiment Analysis (ISA) is a research area that focuses on identifying and classifying emotions directly from digital images. Emotion represents a person’s mood and is a psychophysiological state that arises due to interaction with internal thoughts and external environmental factors [17].

Real Applications for India

In Indian healthcare, image analysis technologies are helping doctors in many ways. AI systems can analyze chest X-rays to detect diseases such as TB and COVID-19, which is very useful in rural areas where expert radiologists are limited [11]. Eye scan analysis helps in early detection of diabetic retinopathy and prevents vision loss. Mobile applications can analyze skin disease images and suggest whether a patient should consult a doctor. These tools reduce the burden on hospitals and help in early diagnosis. For effective performance, such systems must be trained using Indian patient data.

Facial expression and emotion analysis is also gaining importance in healthcare. Emotion recognition systems can help in monitoring mental health conditions such as depression, stress, and anxiety by analyzing facial expressions during patient interactions. In telemedicine, doctors can better understand a patient’s emotional state through video consultations using facial emotion analysis. This is especially helpful for elderly patients and people living in remote areas.

In Indian education, facial expression analysis can be used in online learning platforms to understand student engagement. By analyzing facial emotions such as confusion, interest, or boredom, teachers can improve teaching methods and provide personalized support. During online examinations, facial expression

analysis can also assist in proctoring systems to detect unusual behavior and maintain exam integrity.

Image sentiment analysis is widely used in social media and digital platforms. In India, people share a large number of photos on platforms like Instagram, Facebook, and WhatsApp. Image sentiment analysis can be used to understand public mood and opinions during events such as elections, festivals, or public movements. Government agencies and organizations can use this information to study public response and plan better communication strategies.

In customer service and marketing, facial expression analysis is used to understand customer reactions. Retail stores and shopping malls can analyze customer facial expressions to measure satisfaction with products or services. Advertisers can test how people emotionally respond to advertisements, posters, or product designs. This helps companies improve customer experience and design better marketing campaigns.

In Indian agriculture, farmers are using mobile apps that analyze images of crops. Farmers can take photos of diseased leaves and receive instant advice on pesticide usage [12]. Image sentiment analysis can also be applied in agricultural surveys by studying farmer-shared images to understand crop conditions and farmer confidence levels. Satellite image analysis is used to monitor crop health and predict yields across large farming regions.

In smart cities, facial expression analysis can be used in public safety and security systems. Cameras installed in public places such as railway stations and airports can analyze facial expressions to detect fear, stress, or suspicious behavior, helping security personnel respond quickly. Traffic monitoring systems can also use image analysis to detect driver emotions like fatigue or distraction, reducing accident risks.

In e-commerce platforms, image-based sentiment analysis helps understand customer preferences. Users can upload images of products they like, and systems can recommend similar items. Facial expression analysis during product trials or virtual try-ons can help platforms understand user satisfaction levels and improve recommendations.

Overall, image sentiment analysis and facial expression recognition have wide applications in healthcare, education, security, agriculture, marketing, and smart city development in India [23]. These technologies support better decision-making, improve public services, and contribute to the country’s digital growth.

Table 3: Applications for India

Sector	Problem	Solution	Benefits for India	Challenges in India
Healthcare	Doctor shortage in rural areas	AI diagnosis from images	Reach remote villages, early detection	Need Indian medical image datasets
Agriculture	Crop disease identification	Mobile app with image analysis	Increase farmer income, reduce waste	Internet connectivity in villages
Transportation	Traffic congestion	Camera-based traffic analysis	Better traffic flow, reduce pollution	Chaotic Indian traffic conditions
Security	Face recognition	CCTV + AI analysis	Better security at airports/stations	Privacy concerns, diverse faces
Education	Teaching computer vision	Simple tools + local examples	Skill development for youth	Limited college resources
Retail	Visual product search	Upload image to find products	Help small businesses go online	Product images quality varies

Challenges We Face in India

The biggest challenge for India is the lack of good-quality labeled data. Deep learning models require thousands of labeled images to learn effectively, but for many India-specific problems, such datasets are not available [13]. For example, to detect diseases in Indian crops, we need images of infected plants that are carefully labeled by experts. Creating such datasets is costly, time-consuming, and requires domain knowledge. Another issue is that many models are trained using Western datasets, which often do not perform well for Indian faces, skin tones, clothing styles, road conditions, and environments. Therefore, it is very important to build our own datasets that truly represent the diversity of India.

Limited computing power is another major challenge. Training deep learning models needs high-performance GPUs, which are expensive and consume a large amount of electricity [14]. Many Indian educational institutions and startups cannot afford such infrastructure. In addition, most people in India use mobile phones with limited processing capacity. Hence, there is a strong need for lightweight models that can run efficiently on mobile devices, even without continuous internet access. Privacy is also a serious concern, especially when dealing with medical and personal images. Such data must be stored and processed securely. Overall, the challenge is to develop solutions that respect these limitations while still providing accurate and practical results for Indian applications.

Table 4: Challenges & Solutions for Indian

Challenge	Traditional Methods	Deep Learning Methods	Practical Solution for India
Limited Data	Can work with less data	Needs lots of data	Use transfer learning + data augmentation
Computing Power	Works on normal PCs	Needs expensive GPUs	Use cloud services or mobile-optimized models
Diverse Population	Can adjust features manually	May not recognize all Indian faces	Train on diverse Indian datasets
Multiple Languages	Separate systems for each language	Can integrate with NLP models	Multilingual AI systems
Cost Constraints	Lower implementation cost	Higher initial cost	Use open-source tools + shared resources
Internet Connectivity	Can work offline	Often needs cloud connection	Edge computing on mobile devices
Skill Gap	Easier to understand and implement	Needs specialized knowledge	Simplified tools + training programs
Real-time Needs	Faster processing	Can be slow if not optimized	Use efficient models like MobileNet

Future Directions for Indian Research

Indian researchers should focus on developing efficient image analysis models that can work on low-cost and low-power devices. Since a large part of the Indian population uses affordable smartphones and basic hardware, heavy deep learning models are not always practical. Techniques such as model compression and quantization can reduce the size of large networks without losing much accuracy [15]. Transfer learning is also very useful, as it allows researchers to fine-tune pre-trained models for Indian-specific problems using limited data. This is especially helpful for facial expression and emotion recognition, where collecting large labeled datasets is difficult.

There is a strong need to create open and publicly available datasets for Indian problems, especially for image sentiment analysis and facial expression recognition. Such datasets should include Indian faces with different skin tones, age groups, genders, facial structures, and cultural expressions. Emotions in India are often expressed differently due to cultural and social factors, so Indian-specific emotion datasets are very important. Collaboration between engineering colleges, medical institutions, psychology departments, and agricultural universities can help create high-quality datasets and realistic solutions.

For facial expression analysis, future research should focus on recognizing subtle and mixed emotions, which are common in real-life situations. Most current systems recognize only basic emotions like happy or sad, but real human emotions are more complex. Research should also address challenges such as facial occlusions caused by masks, spectacles, beards, or traditional clothing, which are common in India. Robust models that work well under different lighting conditions and camera qualities are also required.

In image sentiment analysis, researchers should explore understanding emotions from both facial expressions and surrounding context in images. For example, an image of a person may look neutral, but the background, gestures, or event context may reflect strong emotions [22]. Combining facial features with scene understanding and object detection can improve sentiment analysis accuracy. Multimodal approaches that combine images with text, such as captions or social media comments, are another important future direction.

Ethical considerations must be given high priority. Facial expression and emotion recognition systems should work fairly across all Indian communities and must avoid bias related to skin tone, gender, or region [16]. Privacy is

especially important in emotion analysis; as emotional data is sensitive in nature. Clear consent, secure data handling, and responsible use of emotion recognition systems are necessary, particularly in healthcare, education, and surveillance applications.

Explainable AI is another key research direction. Emotion recognition and sentiment analysis systems should clearly explain how they reached a particular decision [23]. Providing explanations in simple Hindi or regional languages will help doctors, teachers, farmers, and policymakers trust and effectively use these systems. Medical emotion analysis tools, such as those used for mental health monitoring, must be clinically validated on Indian populations before real-world deployment.

By focusing on efficient models, Indian-specific datasets, ethical practices, and advanced image sentiment and facial expression analysis techniques, Indian researchers can address local challenges and also develop solutions that are useful for other developing countries with similar social and economic conditions.

Conclusion

Image analysis technology has developed a lot over the years. It has moved from simple rule-based systems to advanced deep learning models that can understand images almost like humans. For India, this technology provides many opportunities to improve important areas such as healthcare, agriculture, security, social media, advertising, and image sentiment analysis. These applications are already helping people in daily life and have the potential to make services faster, smarter, and more accessible. Although India faces challenges such as limited labeled data and lack of high computing resources, these problems can be solved through careful planning, efficient models, and focused research.

The future of image analysis research in India is very promising. Indian students and researchers have many real-world problems to work on, and support from government initiatives and industry is also increasing. By developing solutions that suit Indian conditions, we can not only solve local problems but also create technologies that can be useful for other developing countries. The best approach is to begin with simple ideas, use available tools and pre-trained models, and slowly build strong expertise. Image analysis is no longer limited to big companies or laboratories—it is a powerful tool for every Indian engineer who wants to use technology to solve real problems and contribute to the growth of the country.

References

- Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep learning," *Nature*, 2015. (Explains why deep learning is better than old methods)
- A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep CNNs," *NeurIPS*, 2012. (The paper that started deep learning revolution)
- D. Lowe, "Distinctive image features from scale-invariant key points," *IJCV*, 2004. (Original SIFT paper)
- N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *CVPR*, 2005. (HOG method for detecting people)
- C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, 1995. (SVM classifier basics)
- A. Krizhevsky et al., "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, 2017. (Details about AlexNet success)
- K. Simonyan and A. Zisserman, "Very deep convolutional networks," *arXiv*, 2014. (VGGNet architecture)
- K. He et al., "Deep residual learning for image recognition," *CVPR*, 2016. (ResNet with skip connections)
- J. Redmon et al., "You only look once: Unified object detection," *CVPR*, 2016. (YOLO for fast detection)
- O. Ronneberger et al., "U-Net: Biomedical image segmentation," *MICCAI*, 2015. (Medical segmentation network)
- G. Litjens et al., "Survey on deep learning in medical image analysis," *Medical Image Analysis*, 2017. (Medical applications review)
- A. Kamilaris and F. Prenafeta-Boldú, "Deep learning in agriculture: Survey," *Computers and Electronics in Agriculture*, 2018. (Farming applications)
- J. Buolamwini and T. Gebru, "Gender shades: Accuracy disparities in commercial gender classification," *FAT*, 2018. (About bias in AI systems)
- E. Strubell et al., "Energy and policy considerations for deep learning in NLP," *arXiv*, 2019. (Computing cost problems)
- S. Han et al., "Deep compression: Compressing neural networks," *arXiv*, 2015. (Making models smaller)
- R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks," *ICCV*, 2017. (Making AI decisions understandable)
- M. Aslan, "CNN based efficient approach for emotion recognition," *Journal of King Saud University – Computer and Information Sciences*, vol. 34, pp. 7335–7346, Aug. 2021.
- Sudhamani, D., and S. N. Kavya, "Facial Expression Detection Using Machine Learning Techniques," *International Journal of Scientific Research & Engineering Trends*, vol. 11, no. 1, Jan–Feb 2025, ISSN (Online): 2395-566X.
- Raja Sekaran, S. A. P. C., Lee, P., and Lim, K. M., "Facial Emotion Recognition Using Transfer Learning of AlexNet," *Proceedings of the 9th International Conference on Information and Communication Technology (ICoICT 2021)*, pp. 170–174, 2021. <https://doi.org/10.1109/ICoICT52021.2021.9527512>.
- Lerner, J. S., Li, Y., Valdesolo, P., and Kassam, K. S., "Emotion and Decision Making," *Annual Review of Psychology*, vol. 66, pp. 799–823, 2015. <https://doi.org/10.1146/annurev-psych-010213-115043>.
- Abdo, G. A. G., Almaghrebi, M., Sharifi, M. R., and Jagadeesh, M., "Advanced Machine Learning Approaches for Facial Emotion Recognition," *International Journal / Conference Proceedings*, 2023–2024.
- Ms. S. Z. Mullani and Dr. R. D. Kumbhar, "IMAGE SENTIMENT ANALYSIS: A CRITICAL REVIEW," *Madhya Pradesh Journal of Social Sciences*, vol. 29, no. 2(ii), Dec. 2024.
- S. Z. Mullani and R. D. Kumbhar, "Advancements in Image Sentiment Analysis: Techniques, Challenges, and Future Directions," *Proceedings of the National Conference on Computer Science & Information Technology*, pp. 1–6, Feb. 15–16, 2024, ISBN: 978-81-967894-0-4.