IMAGE CLASSIFICATION USING CNN

P. Anusha[1], Dr. B. Santhosh Kumar[2], A.Rahul[3], B. Nanda Sai[4], B. Nandu[5] , G.Ashwitha[6]
A. Chandra Sekhar[7]
[1]Assistant Professor, [2]Professor and HOD, [3,4,5,6,7]UG Scholar
Department of CSE, Guru Nanak Institute of Technology, Hyderabad, Telangana, India
Email Id: palagatianushareddy@gmail.com[1], bsanthosh.csegnit@gniindia.org[2], anandapurahul9987@gmail.com[3], nandasaibalineni@gmail.com[4], nandubaanothunandu@gmail.com[5], ashwithagoudgodala@gmail.com[6]
chandraankolu@gmail.com[7].

Article Received: 28 April 2025 *Article Accepted: 29 April 2025* *Article Published: 30 April 2025*

Citation

P. Anusha, Dr. B. Santhosh Kumar, A.Rahul, B. Nanda Sai, B. Nandu, G.Ashwitha, A. Chandra Sekhar, "IMAGE CLASSIFICATION USING CNN", Journal of Next Generation Technology (ISSN: 2583-021X), vol. 5, no. 2, pp. 189-197, 2025. DOI: 10.5281/zenodo.15697543

ABSTRACT

A basic problem in computer vision is image classification, which entails grouping images into predetermined classifications. CNNs, or convolutional neural networks, have become an effective technique for their capacity to automatically extract hierarchical feature representations from unprocessed pixel data makes them ideal for this task. This study examines the use of CNNs for image classification, going into detail on their architecture, training procedure, and assessment criteria. There is discussion of the main elements of a CNN, including convolutional layers, pooling layers, and fully connected layers. By training models using benchmark datasets such as CIFAR-10 and ImageNet, the work shows the efficacy of CNNs by obtaining robust performance and high accuracy rates. The outcomes highlight CNNs' benefits. When capturing The findings highlight CNNs' benefits in recognizing spatial hierarchies and minimizing the requirement for human feature extraction, which makes them a popular option for picture classification jobs across a range of industries.

Keywords: Computer Vision, Feature Extraction, CNN, Image Classification.

I. INTRODUCTION

Image Classification and Convolutional Neural Networks (CNNs) Image classification, which involves classifying images into predefined classes, is a fundamental task in computer vision. It has a variety of applications, such as content recommendation systems, autonomous vehicles, medical diagnosis, and facial recognition. Conventional classification methods relied on manual feature extraction, which required extensive preprocessing and domain expertise. However, the rise of Convolutional Neural Networks (CNNs), which are deep learning models specifically made for visual data processing, has revolutionized this field by automating feature extraction and enabling hierarchical learning from raw pixel data. CNNs are deep learning models with convolutional layers, pooling layers, and fully connected layers, each of which has a specific function.CNNs' primary benefit is its capacity to learn and identify intricate patterns without the use of manually created features. CNNs are very effective for a variety of image classification applications because they use backpropagation to modify their parameters in order to reduce classification mistakes. CNN-based models have performed at the cutting edge of benchmark datasets like ImageNet, MNIST, and CIFAR-10. More complex CNN architectures, including as AlexNet, VGGNet, GoogLeNet, and ResNet, have been developed as a result of deep learning advancements. These models show that accuracy and generalization increase with network depth and complexity. The broad success of CNNs has also been facilitated by methods like data augmentation, transfer learning, and large-scale labeled datasets.

II. LITERATURE SURVEY

AlexNet et.al, a deep CNN that won the ILSVRC 2012 competition by a significant margin, was first presented in this groundbreaking study. Five convolutional layers, some followed by max-pooling layers, and three fully connected layers with a 1000-way softmax make up the design[1]. ReLU activation for quicker training, dropout to lessen overfitting, and data augmentation for improved generalization are some of the major advancements. This study opened the door for more deep learning research by showcasing the effectiveness of deep CNNs for large-scale picture classification[2]. The drawbacks of their study are:High processing power is needed for training, overfitting without dropout and data augmentation is likely, the size of the model makes it difficult to install on low-power devices, reliance on huge labeled datasets for efficient training, limited capacity to manage image rotation and scale changes.

Karen Simonyan, Andrew Zisserman, Compared to AlexNet, VGGNet, a deep CNN developed by Simonyan and Zisserman, greatly enhanced network depth [3]. Its use of tiny 3x3 convolution filters allowed for richer structures while maintaining controllable computing demands. VGGNet achieved the best results on the ImageNet dataset, proving that depth boosts performance[4]. Later architectures were influenced by this work, which highlighted the significance of network depth and simplicity in CNN design.The disadvantages of their study are : High processing costs because to the deep architecture, many options that result in higher memory use, training time is slower than with shallower networks, extremely deep versions are susceptible to vanishing gradient problems, the high inference time makes it ineffective for real-time applications[5]-[7].

III. SCOPE OF THE PROJECT

The goal of the Convolutional Neural Networks (CNNs) image classification project is to create an advanced system that can reliably classify images into predetermined groups. This undertaking comprises gathering and preprocessing a variety of datasets, creating and putting into practice efficient CNN architectures, and improving model performance via training and hyperparameter adjustment.

To guarantee the model's resilience and capacity for generalization, a thorough assessment will be carried out utilizing criteria such as accuracy, precision, and recall. To show the trained model's practical usefulness, the project also includes using it in real-world scenarios like object detection and medical picture analysis. Thorough reporting and documentation will facilitate future research and development by offering in-depth insights into the project's methods and results. Furthermore, the study can investigate cutting-edge methods like real-time applications and transfer learning, expanding

To increase its impact and usefulness, the research might also investigate cutting-edge methods like transfer learning and real-time applications. By taking care of these elements, the project hopes to greatly to the field of computer vision, demonstrating how well CNNs automate and improve image classification tasks.

IV. PROBLEM STATEMENT

Accurately classifying photos into predetermined classifications is the challenge of image classification, which is crucial for applications like medical diagnosis, autonomous driving, and facial recognition. When dealing with huge and complicated datasets, traditional methods that rely on manual feature extraction are ineffective and frequently insufficient. By utilizing Convolutional Neural Networks (CNNs), which are capable of autonomously learning and extracting hierarchical information from photos, this study seeks to overcome these issues. The main goal is to create, deploy, and enhance a CNN model that can reliably and accurately identify images from a variety of datasets. To guarantee the model's efficacy, the project will entail data gathering, preprocessing, model training, and thorough evaluation. Furthermore, the model's practical implementation will show how applicable it is in the real world. The research aims to improve picture categorization jobs' efficiency and accuracy by resolving this issue, advancing artificial intelligence and computer vision.

V. EXISTING SYSTEM

Current CNN-Based Image Classification Systems Convolutional Neural Networks (CNNs), which provide notable advancements over conventional techniques, have emerged as the mainstay of image classification jobs in recent years. Numerous CNN architectures have been created and extensively used in both commercial and scholarly applications, each of which has made a distinct contribution to the field's progress.

VI. PROPOSED SYSTEM

For every employment opportunity, recruiters look for the most qualified candidates for a particular role. Useful resume information including education, skills, accomplishments, and experience should be automatically generated by technology. This program seeks to lessen bias and recruitment time. during the hiring process by assessing applicants according to a number of resume-related factors.

5.1. Proposed System Advantages

1. **Automatic Feature Extraction:** CNNs do not require human feature extraction; instead, they automatically extract hierarchical features from raw pixel data. This feature guarantees optimal extraction of pertinent features while minimizing human labor.
2. **Capacity to Acquire Complex Pattern Knowledge:** CNNs are excellent in identifying complex patterns and spatial relationships in images, including forms,

edges, and textures. Because of this, they can more accurately differentiate between classes, even in datasets that are complex.

3. **Representation of Hierarchical Features:** CNN architectures produce a hierarchical feature representation through the use of layers such as convolutional and pooling layers. The model can gradually comprehend both low-level and high-level features thanks to this hierarchical method, which raises classification accuracy overall.
4. **Flexibility and Adaptability:** By modifying network designs, incorporating regularization strategies, or employing transfer learning from pretrained models, CNNs can be tailored and improved for various image classification applications. Because of their adaptability, CNNs can be used in a variety of fields and applications.
5. **Cutting Edge Performance:** CNNs have continuously surpassed conventional techniques and attained cutting-edge results on benchmark datasets like as ImageNet. Their capacity to utilize extensive datasets and they can push the limits of accuracy in picture classification jobs thanks to complicated architectures.
6. **Real-time Processing Capabilities:** CNNs can process pictures in real-time because to developments in hardware acceleration (such as GPUs and TPUs), which makes them appropriate for applications that need to make decisions quickly, like medical diagnostics and driverless cars.

VII. SYSTEM ARCHITECTURE

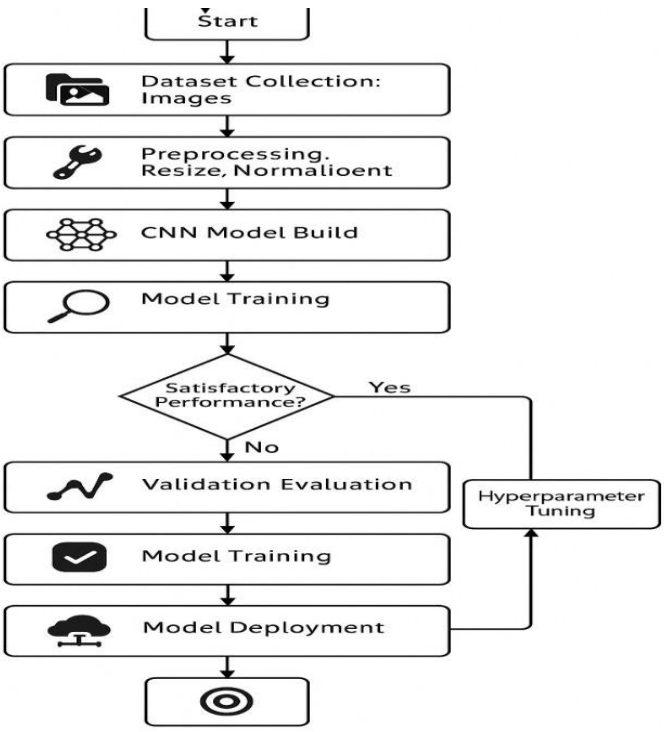

Fig 1: System Architecture

- The entire process of creating and implementing a Convolutional Neural Network (CNN) for picture classification tasks, like identifying age and gender, is depicted in the flowchart. In order to train the model, the first step in the procedure is the **collection of a labeled image dataset**.

- **Preprocessing** is the next phase, when images are scaled, standardized, and enhanced to increase the generalization and robustness of the model. After preprocessing, layers such as convolutional, pooling, and dense layers are used to create a **CNN architecture** that aids in the extraction and interpretation of information from the images. The training dataset is then used to **train** the model. The model's performance is evaluated following initial training to see if it satisfies acceptable accuracy and loss standards
- .If the model is successful, it moves on to the **deployment step**, where it is prepared for practical application. On the other hand, the model is **validated evaluation** using a different validation dataset if the performance is not adequate. In light of these findings, **hyperparameter tuning** is done to modify batch sizes, learning rates, and other parameters. These new parameters are then used to retrain the model, and it is subsequently reassessed.
- Until the best possible performance is attained, this cycle keeps going. Ultimately, the effective model is **deployed** in an appropriate setting, such a web or mobile application, enabling real-world applications. The methodical and iterative process of creating a CNN-based image classification system is well-represented in this flowchart, guaranteeing accuracy and dependability in practical situations.

7.1 CNN ARCHITECTURE:

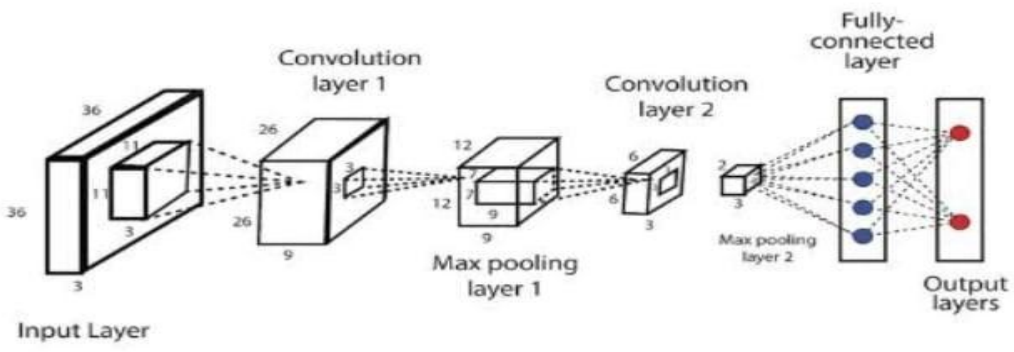

Fig 2: CNN Architecture

The following are components of the CNN system architecture for dog and cat image classification: Data collection and preprocessing (resizing, normalization, augmentation), Convolutional, pooling, and fully linked layers make up the model design. Training (forward evaluation (accuracy, precision, recall), deployment (integration into applications, real-time inference, monitoring), and propagation, loss calculation, and backpropagation).

7.2 ALGORITHM: (CNN MODEL)

- **Configuration and Data Gathering**: Extract the dataset to the working directory after downloading it from Kaggle. Use `image_dataset_from_directory()` to load the training and validation datasets. Resize photos to 256x256 pixels and set the batch scale to 32. Scale the pixel values between 0 and 1 to normalize the photos.

- **Explain the CNN Model**: Set up a sequential model from scratch. Include convolutional layers that are activated by ReLU. 32 filters with a 3x3 kernel make up the first layer, which is followed by max pooling and batch normalization. 64 filters with a 3x3 kernel, batch normalization, and max pooling are included in the second layer. The third layer features batch normalization, max pooling, and 128 filters with a 3x3 kernel. Add fully connected layers and flatten the feature maps. 64 neurons with ReLU activation and 0.1 dropout make up the second dense layer. One neuron with sigmoid activity for binary classification is present in the last dense layer. Utilizing the binary cross entropy loss function, the Adam optimizer, and accuracy as a metric, compile the model.
- **Teach the Model**: Use the training dataset to train the CNN for ten epochs, and then use the validation dataset to verify the results.
- **Graph of Plot Accuracy**: Plot the training and validation accuracy throughout the epochs using Matplotlib.
- **Test Model Using the New Picture**: OpenCV, load a test picture. It should be resized to 256 by 256 pixels and reshaped to conform to the model's input format. Determine if the image depicts a dog or a cat using the trained model.

VIII. RESULTS AND DISUSSIONS

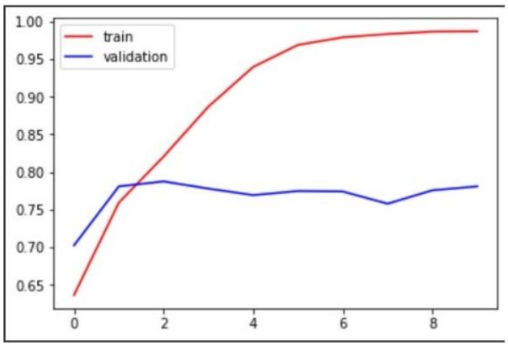

Fig. 3. Overfitting in CNN Model: Training vs. Validation Accuracy

A Convolutional Neural Network (CNN) model's training and validation accuracy over several epochs is depicted in this graphic. Training accuracy is shown by the red line, while validation accuracy is shown by the blue line.

- **Training Accuracy (Red Line):** Over a few epochs, the training accuracy rises quickly to nearly 100%. This suggests that the training dataset is helping the model learn.
- **Validation Accuracy (Blue Line)** - The validation accuracy rises at first, but then varies and does not become much better. According to this, the model does well on training data but struggles to generalize to new data.
- **Possible Overfitting -** Overfitting is indicated by a significant discrepancy between training and validation accuracy. The model performs poorly on unseen data, despite having acquired patterns unique to the training set.
- **Possible Remedies**- Increase the amount of training data or use data augmentation. Use regularization strategies such as L2 regularization or dropout. To avoid overtraining past the point of generalization, try early quitting.

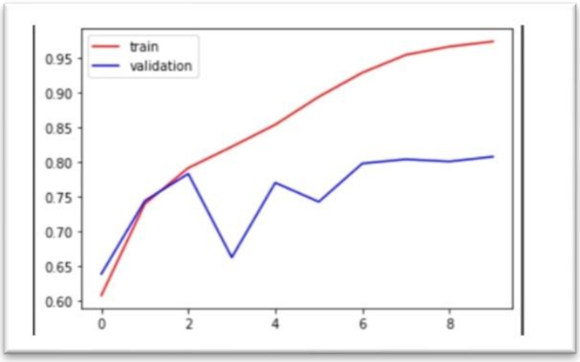

Fig 4: Accuracy Comparison :Training vs Validation

An image classification model's training and validation accuracy over several epochs is depicted in this graph. Here is a thorough explanation:

- **The X-Axis (Epochs):** indicates how many training iterations the model has undergone.
  The model picks up patterns in the dataset as the number of epochs increases.
- **Accuracy of the Y-Axis:** shows the model's accuracy across the training and validation datasets. Better categorization performance is indicated by higher values.
- **Training Accuracy Red Line:** steadily rises, suggesting that the model is picking up new information and accurately fitting the training set. gets close to 100%, indicating that the model is learning the training data.
- **Validation Accuracy, or Blue Line:** increases at first, indicating that the model is picking up practical aspects. shows possible overfitting when it eventually varies and stops getting better.

7.3. OUTPUT OF THE TESTED IMAGE:

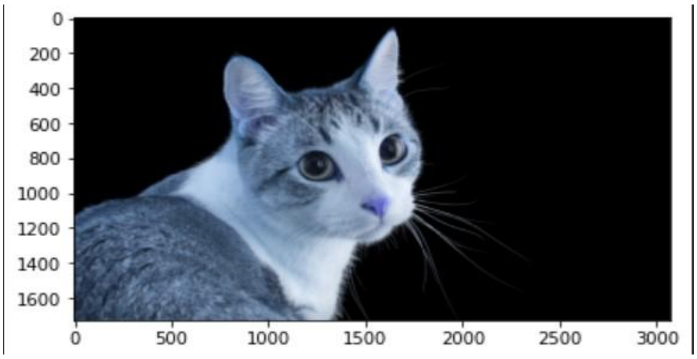

Fig 5 : Labeled Image

7.3 SUMMARY OF THE MODEL TESTING:

Shape of the original image: 1728, 3072, 3 → High-resolution color image. Prior to model input, it was resized to (256,256,3).Output for Prediction: array([[0.]], dtype=float32). In conclusion, this image is categorized as class 0 by the model.

The model uses cv2 and the other convolutional layers to scale the images and normalize the image pixels from (0,256) to (0,1) each. The CNN model recognizes the provided image by using the downloaded database that was used to train the model. A CAT is the outcome if the output is equal to zero. The outcome is a DOG if the output is equal to 1.

IX. CONCLUSION

Convolutional Neural Networks (CNNs) have revolutionized image categorization by providing a potent and effective method for analyzing visual data. CNNs are perfect for complicated image identification tasks because of their ability to identify patterns and characteristics in images, which is modeled after the human visual cortex. CNNs automatically learn spatial hierarchies using layers like convolutional, pooling, and fully connected layers, in contrast to conventional techniques that depend on manual feature extraction. Preparing the data, including preprocessing and augmentation to increase model resilience, is the first step in the development process. The CNN architecture is then constructed, usually consisting of dense layers for classification after several convolutional and pooling layers. The model is trained and verified to make sure it avoids overfitting and generalizes properly after being compiled with the proper loss functions and optimizers. CNNs are well suited for applications such as autonomous vehicles and medical imaging because of their great scalability and ability to process big datasets effectively, particularly when GPU acceleration is used. Notwithstanding obstacles like the requirement for sizable labeled datasets and high processing requirements, developments like transfer learning have lessened these restrictions. In conclusion, CNNs are a key component of contemporary AI because they provide automated feature extraction, scalability, and adaptability, all of which spur further advancements in computer vision.

References

[1]. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in Advances in Neural Information Processing Systems 25 (NeurIPS 2012), pp. 1097–1105, 2012.

[2]. P. Muthukumar, P. A, R. A. Karthika, K. Eswaramoorthy, V. U. Rani and R. V. K, "Handwritten Text Recognition from Image Using LSTM Integrated with Pixel Shifting Optimization Algorithm," 2024 International Conference on Advancement in Renewable Energy and Intelligent Systems (AREIS), Thrissur, India, 2024, pp. 1-6, doi: 10.1109/AREIS62559.2024.10893651.

[3]. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in Proc. Int. Conf. Learn. Representations (ICLR), 2015.

[4]. Ganesan, Vithya, V. Anjana Devi, and Pachipala Yellamma. “Object Detection using PYTHON Programming.” Journal of Next Generation Technology (ISSN: 2583-021X) 2.1 (2022).

[5]. G. Padmapriya, B. Santhosh Kumar, M.N. Kavitha, V. Vennila, "Deep convolutional neural network for real time object detection using tensor flow", Materials Today: Proceedings, 2021, ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2021.02.671.

[6]. P. Kadiri, P. Anusha, M. Prabhu, R. Asuncion, V. S. Pavan and J. V. Suman, "Morphed Picture Recognition using Machine Learning Algorithms," 2024 Second International

Conference on Advances in Information Technology (ICAIT), Chikkamagaluru, Karnataka, India, 2024, pp. 1-6, doi: 10.1109/ICAIT61638.2024.10690845.

[7]. Singh, A., Kumar, T. C. A., Mithun, T., Majji, S., Rajesh, M., & Anusha, P. (2021, December). Image Processing Approaches for Oral Cancer Detection in Color Images. In 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 817-821). IEEE.