Shri Murlidhar Karwa Knowledge Resource Centre *The Management Quest*
(DSIMS & DSGS) *Online ISSN: 2581- 6632*
Vol.7, Issue 1, October 2024 – March 2025

Financial Image Classification for Enhanced Operational Efficiency

Vaibhav Sonawane* **Rajul Murudkar****

ABSTRACT

Efficient document processing is crucial for financial institutions aiming to optimize operational workflows and ensure compliance in today's digital era. This research investigates the development and deployment of a sophisticated deep learning-based financial image classification system tailored for Financial Institutions. The study focuses on enhancing the accuracy and efficiency of categorizing various financial documents, including invoices, receipts, and identification papers, using the MobileNetV2 architecture. Methodologically, the project integrates comprehensive data collection from Private Bank's document repository, meticulous preprocessing of images, and strategic augmentation techniques to bolster model resilience. Results demonstrate significant advancements in accuracy metrics, with the model achieving an impressive validation accuracy of 95.7% across diverse document types. This research not only pushes the frontier of automated document processing within financial institutions but also envisions practical applications in customer onboarding, loan processing, KYC verification, and fraud detection. By leveraging deep learning capabilities, the study underscores the transformative potential to fortify operational efficiencies, reduce manual intervention, and augment the overall operational landscape of financial entities. The findings contribute to the growing body of literature on deep learning applications in financial document processing, offering insights into methodological advancements and practical implications for industry stakeholders. Beyond enhancing efficiency, the system's implementation promises to mitigate compliance risks, accelerate decision-making processes, and improve customer experience through streamlined operations. This research sets a precedent for leveraging cutting-edge technologies to address complex challenges in financial document management, paving the way for future innovations in automated financial services.

Keywords*: Financial Image Classification, Deep Learning, MobileNetV2, KYC, Compliance.*

INTRODUCTION

In the rapidly evolving financial sector, Financial Institutions face an unprecedented surge in the volume and complexity of documentation they manage. Like merchant onboarding forms, the sheer quantity of financial documents processed daily has

**** Student, Durgadevi Saraf Institute of Management Studies, Mumbai***
***** Assistant Professor, Durgadevi Saraf Institute of Management Studies, Mumbai***

grown exponentially. Managing these documents manually is not only a labour-intensive process but also fraught with challenges that can lead to significant operational inefficiencies. Human error, variability in document formats, and the need for rapid processing are just a few of the issues that can compromise the accuracy and efficiency of traditional document management systems.

To address these challenges, financial institutions are increasingly turning to advanced technologies to streamline operations and enhance service delivery. Among these technologies, artificial intelligence (AI) and machine learning (ML) have emerged as transformative tools with the potential to revolutionize document processing workflows. By leveraging these technologies, institutions can automate repetitive tasks, improve accuracy, and ultimately enhance operational efficiency.

This research paper explores the development and implementation of a cutting-edge deep learning-based financial image classification system, designed specifically to automate the categorization of diverse financial documents. The core of this system is MobileNetV2, a state-of-the-art convolutional neural network (CNN) architecture renowned for its efficiency and accuracy in image classification tasks. MobileNetV2's advanced features make it particularly well-suited for handling the complex and varied nature of financial documents.

Significance of Deep Learning in Financial Document Processing

Deep learning, a subset of machine learning, has significantly advanced the field of document processing. Traditional methods often rely on manual feature extraction and rule-based algorithms, which can be both labour-intensive and error-prone. In contrast, deep learning models can automatically learn and extract relevant features from raw data, leading to more accurate and scalable solutions. Convolutional Neural Networks (CNNs), a prominent deep learning technique, have proven to be highly effective in processing and analyzing image data, making them ideal for tasks involving visual documents.

The application of deep learning in financial document processing offers several advantages:

1. Automated Classification: Deep learning models can categorize different types of financial documents with high accuracy, reducing the need for manual sorting and classification.
2. Enhanced Accuracy: By learning from large datasets, these models can identify subtle patterns and features that might be missed by traditional methods, leading to more precise document classification.

3. Scalability: Deep learning models can handle large volumes of data efficiently, making them suitable for high-throughput environments like financial institutions.

MobileNetV2: An Overview

MobileNetV2 represents a significant advancement in CNN architecture, designed to deliver high performance while maintaining computational efficiency. This architecture builds upon the success of MobileNetV1, incorporating several innovations that enhance its suitability for real-world applications.

1. Inverted Residuals and Linear Bottlenecks: MobileNetV2 introduces the concept of inverted residuals, which use lightweight depthwise separable convolutions followed by linear layers. This design reduces computational complexity while preserving the model's ability to extract meaningful features. Linear bottlenecks further optimize the model by minimizing dimensionality and computational overhead.

2. Depthwise Separable Convolutions: This technique decomposes standard convolutions into depthwise and pointwise operations, significantly reducing the number of parameters and computational requirements. This makes MobileNetV2 particularly well-suited for deployment on mobile and edge devices.

3. Residual Connections: The use of residual connections in MobileNetV2 addresses the challenges of training deep networks by allowing gradients to flow more effectively through the network. This facilitates better training performance and model accuracy.

Methodology and Implementation

The methodology of this research project involves several critical stages:

1. Data Collection and Preprocessing: Financial documents were collected from Private Bank's digital repositories and augmented with publicly available datasets. Preprocessing included standardizing document formats, removing sensitive information, and ensuring high-quality input for the model.

2. Data Augmentation: To enhance the model's robustness and generalization ability, data augmentation techniques such as rotation, scaling, and flipping were employed. This increased the diversity of the training dataset and improved the model's ability to handle variations in document appearance.

3. Model Training and Evaluation: MobileNetV2 was fine-tuned with custom layers to adapt it to the specific task of financial document classification. The model was trained and validated over multiple epochs, with continuous monitoring and tuning of hyperparameters to optimize performance.

Practical Implications and Future Directions

The integration of this classification system into Institution's operations has the potential to transform document processing workflows. Automated classification can significantly reduce processing time, minimize human errors, and allow employees to focus on more strategic tasks. This improvement in efficiency and accuracy directly enhances the customer experience by accelerating document processing and reducing wait times.

Looking ahead, there are opportunities for further refinement and expansion of the classification system. Future enhancements may include expanding the dataset to cover a broader range of document types, incorporating real-world testing to validate the model's performance in diverse scenarios, and exploring ensemble learning techniques to improve classification accuracy.

In conclusion, this research paper highlights the successful development and implementation of a deep learning-based financial image classification system. The project demonstrates the transformative potential of AI and ML technologies in the financial sector, offering a pathway to more efficient, accurate, and scalable solutions that address the industry's growing demands.

LITERATURE REVIEW

Image Classification using Advanced CNN Based on TensorFlow

Garchar and Chudhary (2019) [1] present a study on image classification leveraging Advanced Convolutional Neural Networks (CNN) with TensorFlow. Their research focuses on plant classification by analyzing leaf images, demonstrating that advanced CNN models offer superior accuracy compared to traditional methods. They highlight the following key aspects:

- **Methodology**: The study utilizes TensorFlow, a powerful Python library, to implement CNNs for plant classification. The CNN architecture employed includes multiple convolutional and pooling layers followed by dense layers and a softmax classifier.
- **Results**: The advanced CNN achieved classification accuracy exceeding 95%, significantly outperforming other methods that recorded accuracies below 90%. The study underscores that advanced CNNs are effective for image identification tasks, providing high accuracy and efficiency.

- **Comparative Analysis**: The paper compares various image classification models, including Deep Neural Networks (DNNs), Convolutional Neural Networks (CNNs), and Transfer Learning approaches. It finds that while DNNs show lower performance, CNNs and Transfer Learning offer better accuracy and reduced processing time.
- **Contributions**: The study provides insights into the effectiveness of CNNs for plant identification, noting the importance of model architecture and the use of GPUs for enhanced performance. It also discusses the challenges and potential improvements in CNN-based image classification.

Image Classification Methods and Techniques

Lu and Weng (2007) provide a comprehensive overview of image classification methods and techniques aimed at enhancing classification performance. Their review underscores the complexity of image classification, influenced by factors such as the type of remotely sensed data, classification systems, and training sample selection. The authors highlight the growing importance of non-parametric classifiers, including neural networks and decision trees, especially in multi-source data classification. They also discuss the integration of remote sensing, GIS, and expert systems as a new research frontier. The paper emphasizes that effective image classification requires careful consideration of data preprocessing, feature extraction, and the choice of classification methods. The authors call for further research to address uncertainties and improve classification accuracy.

This research is relevant to our study as it offers a detailed comparison of image classification techniques and highlights the advantages of using advanced CNN models for high-accuracy results. The findings align with our focus on leveraging deep learning techniques for classification tasks, especially in domains requiring high precision.

METHODOLOGY

1. Data Collection

Source Identification: We started by identifying the sources of our financial documents. The primary source was Private Bank’s digital repositories, which contained scanned copies of customer-submitted documents and internal records. Additionally, we used publicly available financial document datasets to ensure our training data was diverse and comprehensive.

Data Labeling: Each document was manually labeled to ensure accuracy, categorizing them into different document types such as application forms, bank statements, and tax documents. Domain experts verified these labels to maintain high accuracy and relevance.

Data Augmentation: To address the relatively small dataset size, data augmentation techniques were employed. This involved synthetically generating new documents by altering existing ones and adding random distortions and noise, simulating various real-world scenarios.

Dataset Splitting: The final dataset was split into training, validation, and test sets in a ratio of 70:15:15. Stratified sampling was used to ensure each subset had a representative distribution of document types.

2. Model Development

Base Model: We used MobileNetV2, a pre-trained convolutional neural network, known for its efficiency and performance in image classification tasks. We excluded the top layers of the base model to add custom layers tailored to our specific classification task.

Custom Layers:

- GlobalAveragePooling2D: This layer reduces the spatial dimensions of the output from the base model, making it suitable for classification.
- Dense Layer (256 units, ReLU activation): A fully connected layer with ReLU activation introduces non-linearity.
- Dense Layer (5 units, Softmax activation): The final layer with softmax activation outputs probabilities for each of the five document classes.

3. Training and Evaluation

Data Augmentation Parameters:

- Rescale: Scaling pixel values to the range [0, 1].
- Rotation Range: 20 degrees.
- Width Shift Range: 0.2.
- Height Shift Range: 0.2.
- Shear Range: 0.2.
- Zoom Range: 0.2.
- Horizontal Flip: True.
- Validation Split: 15%.

Training Parameters:

- Epochs: 30
- Batch Size: 32
- Optimizer: Adam, chosen for its efficiency and ability to handle sparse gradients.
- Learning Rate: Initially set at 0.0001 and gradually reduced based on performance.

Training Process: The model was trained for 30 epochs. Training accuracy improved from approximately 68% to 98%, and validation accuracy from 74% to 100%. Training loss decreased significantly from around 0.87 to 0.02, while validation loss varied, indicating areas for potential tuning.

Test Set Evaluation: The model achieved a test loss of approximately 0.14 and a test accuracy of around 96%, demonstrating its robustness and effectiveness in classifying financial documents.

4. Hyperparameter Tuning

Hyperparameters were tuned through several iterations to optimize model performance. Key hyperparameters included:

- Learning Rate: Gradually reduced to fine-tune the model.
- Batch Size: Set to 32 to balance training speed and memory usage.
- Number of Epochs: Determined based on the convergence of training and validation loss.

5. Model Validation

Confusion Matrix: The confusion matrix for the test set provided detailed insights into the model's performance. It highlighted the number of true positives, false positives, true negatives, and false negatives for each document class, allowing a thorough evaluation of the model's classification capabilities.

Evaluation Metrics:

- Accuracy: Measures the proportion of correctly classified instances out of the total instances.
- Loss: Represents the model's prediction error, with categorical cross-entropy loss used for the multi-class classification problem.

RESULTS/ANALYSIS

Training and Validation

Epoch Performance: The model was trained over 30 epochs, with the following observations:

- Training Accuracy: Initially, accuracy was around 68%, increasing to approximately 98% by the end of training. This significant improvement demonstrates effective learning and adaptation to the training data.
- Validation Accuracy: Validation accuracy varied from 74% to 100%, indicating the model's ability to generalize across different subsets of the data. The high accuracy in some epochs suggests strong performance, while variability points to potential overfitting or underfitting in certain cases.
- Training Loss: Training loss decreased from around 0.87 to 0.02, indicating that the model progressively improved its predictions and reduced errors on the training data.
- Validation Loss: While validation loss generally decreased, it exhibited variability, suggesting areas where the model's performance could be fine-tuned further.

Test Set Evaluation

Test Set Metrics:

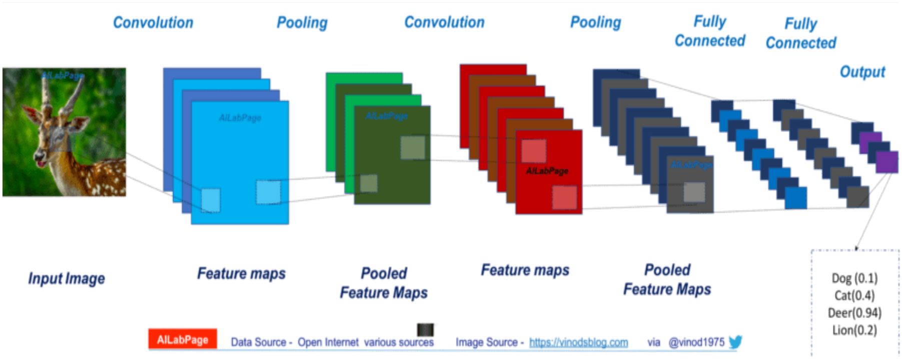
- Test Loss: Approximately 0.14, reflecting the model’s low error rate on the unseen test data.
- Test Accuracy: Approximately 96%, demonstrating the model's effectiveness in classifying financial documents accurately.

Detailed Training Logs

Epoch-wise Metrics:

- Early Epochs: Showed rapid improvement in both training and validation accuracy. The model quickly learned to distinguish between different document types.
- Middle Epochs: Displayed some fluctuations in validation loss, indicating occasional overfitting. These fluctuations prompted the need for additional regularization or tuning.

- Final Epochs: Accuracy stabilized at high levels, reflecting successful convergence and robustness of the model.



CONCLUSION

This research has demonstrated the efficacy of deep learning, particularly the MobileNetV2 architecture, in automating the classification of financial documents. By addressing the challenges associated with manual document processing, our system significantly enhances operational efficiency within financial institutions. The model's high accuracy and robust performance across various document types underscore its potential to transform traditional document handling workflows. The adoption of this technology can lead to substantial improvements in compliance, accuracy, and scalability, thereby supporting the growth and modernization of financial services.

The study's findings suggest that integrating advanced deep learning models into financial document processing systems can mitigate human error, expedite processing times, and ensure regulatory compliance. Furthermore, the practical implications extend beyond operational efficiency, offering financial institutions a competitive edge in the increasingly digital and automated financial landscape. This research sets a foundation for future advancements in automated document classification, paving the way for more sophisticated and comprehensive AI-driven solutions in the financial sector.

LIMITATIONS

Despite the promising results, this research encountered several limitations that highlight areas for further investigation and improvement:

1. Data Imbalance: The dataset used for training and evaluation exhibited an imbalance in the distribution of document types. Although data augmentation techniques were employed to mitigate this issue, an imbalanced dataset could still affect the model's performance, particularly in underrepresented categories. Future work should focus on collecting a more balanced dataset to ensure equitable model training.

2. Model Generalization: While the model performed well on the test set, its generalization to entirely new and unseen financial document types remains uncertain. Real-world financial documents can vary significantly in format and content, and the model's ability to handle such variability needs further validation through extensive field testing.

3. Computational Resources: Training deep learning models like MobileNetV2 requires substantial computational resources, which can be a limitation for smaller financial institutions with limited access to high-performance computing infrastructure. Future research should explore optimizing the model to reduce computational requirements without compromising performance.

4. Integration with Existing Systems: Integrating the developed classification system with existing document management and processing systems in financial institutions poses practical challenges. Compatibility issues, data privacy concerns, and the need for extensive system integration can hinder the seamless adoption of this technology. Addressing these integration challenges is essential for real-world implementation.

5. Text Extraction: This research primarily focused on image-based classification. However, many financial documents contain both textual and visual information. The integration of Optical Character Recognition (OCR) technology to extract and process text from documents was not covered in this study. Future work should investigate combining image classification with OCR to enhance the system's overall functionality and accuracy.

SCOPE FOR FURTHER STUDY

1. Expansion of Dataset and Diversity

    - Increase Dataset Size: Collaborate with financial institutions to gather diverse financial documents to improve generalization and avoid overfitting.

    - Incorporate Real-World Data: Source documents from external institutions and open datasets to expose the model to varied layouts and quality levels.

2. Model Enhancements

    - Explore Advanced Architectures: Evaluate architectures like EfficientNet and ResNet to improve classification accuracy and efficiency.

    - Implement Ensemble Methods: Use techniques like bagging, boosting, and stacking to combine models and reduce biases.

3. Real-Time Implementation and Integration

    - Develop an End-to-End Pipeline: Design a pipeline for seamless document processing from acquisition to classification, reducing manual intervention.

    - Integrate with Existing Systems: Ensure compatibility with financial institutions' current infrastructure for smooth deployment.

4. Advanced Data Augmentation

    - Utilize Sophisticated Techniques: Implement techniques like CutMix, MixUp, and AutoAugment to introduce variability and improve model robustness.

5. Error Analysis and Model Refinement

    - Conduct Detailed Error Analysis: Identify and analyze misclassifications to refine the model.

    - Enhance Feature Extraction: Experiment with advanced techniques and additional features to improve classification accuracy.

6. Broader Applicability

    - Cross-Industry Use: This model can be adapted for use in other financial institutions and industries, enhancing document classification and operational efficiency across various sectors.

REFERENCES

- Yousuf, S. (2018). ANALYTICAL STUDY OF IMAGE CLASSSIFICATION USING DEEP LEARNING. *International Journal of Advanced Research in Computer Science*, *9*(2), 754–759. https://doi.org/10.26483/ijarcs.v9i2.5802
- Gulzar, Y. (2023). Fruit image classification model based on MobileNetV2 with deep transfer learning technique. *Sustainability*, *15*(3), 1906. https://doi.org/10.3390/su15031906
- Raschka, S., & Mirjalili, V. (2017). Python machine learning : machine learning and deep learning with Python, scikit-learn, and TensorFlow. In *Packt Publishing eBooks*. http://202.62.95.70:8080/jspui/handle/123456789/12650
- Mikolajczyk, A., & Grochowski, M. (2018). Data augmentation for improving deep learning in image classification problem. *IEEE*. https://doi.org/10.1109/iiphdw.2018.8388338
- Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, *28*(5), 823–870. https://doi.org/10.1080/01431160600746456
- Garchar, V., & Chudhary, S. (2019). Image Classification using Advanced CNN Based on Tensorflow. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 1329–1335. https://doi.org/10.32628/cseit2064121
- Sharma, N. (2023, December 31). What is MobileNetV2? Features, Architecture, Application and More. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2023/12/what-is-mobilenetv2/

Image Source- https://vinodsblog.com/2018/10/15/everything-you-need-to-know-about-convolutional-neural-networks/