Physical Education, Health and Social Sciences

https://journal-of-social-education.org E-ISSN: 2958-5996

P-ISSN: 2958-5988

Cricket Legends: Exploring VGG-16 for Sports Figure Identification

Usama [1], Muhammad Usman [2], Talha Saleem [3], Muhammad Shahzad [4], Irfan Ullah [5], Riaz Ahmad [6], Muhammad Hamid [7], Muhammad Rayyan Amjad [8], Zafar Khan [9]

[1] Govt Degree College, Hayatabad, Peshawar, Pakistan. Email: usamak2004@gmail.com
[2] Govt Degree College, Hayatabad, Peshawar, Pakistan. Email: usmanakbar4444.com@gmail.com
[3] Department of Computer Science, Edinburgh Napier University, UK. Email:talhatasleem996@gmail.com
[4] Govt Degree College, Hayatabad, Peshawar, Pakistan. Email: muhammadshahzad8390@gmail.com
[5] Govt Degree College, Hayatabad, Peshawar, Pakistan. Email: irfanullahafridi046@gmail.com
[6] Higher Education Department, KP, Pakistan. Email: riazahmad@hed.gkp.pk
[7] Govt Degree College, Hayatabad, Peshawar, Pakistan. Email:Muhammadhamidkhan0306@gmail.com
[8] Govt Degree College, Hayatabad, Peshawar, Pakistan. Email:muhammadrayyandev@gmail.com
[9] Higher Education Department, KP, Pakistan. Email: zafar.khalil88@gmail.com

DOI: https://doi.org/10.63163/jpehss.v3i4.769

Abstract

Cricket is among the world's most popular sports, appreciated for its competitive spirit and for the legendary players who have shaped its history. This study presents a deep learning framework that automatically classifies thirty renowned cricket legends. A custom dataset with over 22,817 images was assembled and used to fine-tune a pre-trained VGG-16 convolutional neural network via transfer learning. To ensure accuracy, the model was tested using 5-fold stratified cross-validation, achieving an average accuracy of 94.37% (±0.39%) and consistent results across validation sets. These findings demonstrate the effectiveness of transfer learning for sports image classification and point to valuable applications in digital sports archiving, media analysis, and fan engagement platforms.

Keywords: Deep Learning, Computer Vision, Image Classification, Vgg-16, Transfer Learning, Convolutional Neural Networks, Cricket, Cricket Legends Sports Analytics, Data Augmentation.

Introduction

Sports have always played an essential role in society, bringing people of all ages, cultures, and backgrounds together. Whether they are children, teenagers, adults, or older people, millions of individuals share a passion for various sports. Across the world, many games are played and celebrated, among them cricket, which holds a significant position. Cricket is played in more than 100 countries and is currently ranked among the top 10 most popular sports worldwide. The sport has a massive following, with fans not only supporting their national teams but also admiring legendary players from across the globe. Cricket players, commonly called cricketers, take on different roles in the game. Some are known for their exceptional batting skills, while others are recognized as top-class bowlers. A few possess the unique ability to perform both roles effectively, earning them the title of all-rounders. Over the years, many players have achieved legendary status due to their remarkable performances, records, and contributions to the sport. These individuals are admired not only in their home countries but also by fans worldwide, making them global icons of the game.

In this study, we focus on thirty of the greatest cricket legends from different international teams. Images of these players were collected and compiled into a specialized dataset. The purpose of building this dataset is to train a deep learning model that can accurately identify these legendary figures from images. Using this dataset, the study aims to explore the potential of computer vision for recognizing sports personalities and to demonstrate how artificial intelligence can contribute to the digital archiving and analysis of cricket history.

To achieve this goal, we have developed a robust framework for image classification using the well-established VGG-16 convolutional neural network. Our methodology employs transfer learning, allowing us to fine-tune a pre-trained VGG-16 model on a custom dataset containing over 22,817 images. This dataset was meticulously curated to represent the 30 selected cricket legends, enabling the model to adapt its powerful feature-extraction capabilities to the unique characteristics of cricket imagery. By leveraging this approach, we ensure that the system learns both general visual features from the original ImageNet training and the specific nuances of our dataset. Furthermore, to ensure a reliable and unbiased evaluation of our model's performance, we use 5-fold cross-validation. This process systematically trains and tests the model across different subsets of the dataset, ensuring that the reported accuracy reflects the model's generalization rather than being biased toward any particular subgroup. Such rigor is critical when developing systems intended for large-scale or real-world applications, where robustness is vital.

The primary contribution of this research is the creation of a scalable, accurate framework for automated classification of cricket legends from images. Beyond its technical achievements, this system offers practical value in multiple domains. It has potential applications in sports media analytics, where automated player identification can enhance content tagging and retrieval; in fan engagement platforms, where personalized experiences can be enriched through accurate recognition of favorite players; and in digital library management, where the systematic archiving of sports history can be significantly streamlined.

By presenting a clear benchmark for cricket legend image classification, our work not only addresses a niche but also a significant problem in sports analysis and contributes to the broader field of computer vision applications in cultural and entertainment domains. Ultimately, this research highlights how artificial intelligence can bridge the gap between technology and sports, offering innovative solutions to preserve, analyze, and celebrate the legacies of cricket's greatest players.

Literature Review

The domain of cricket analytics and player recognition has grown rapidly in recent years, driven by advancements in computer vision, artificial intelligence, and immersive technologies. The review of ten relevant studies presents a comprehensive understanding of the current methods and technologies used in facial recognition, player identification, and performance evaluation. These works, though diverse in focus—from emotion recognition to immersive VR simulations—collectively contribute to the conceptual and technical foundation for building a comprehensive image dataset focused on cricket legends.

Lehuger et al. (2007), [1] Developed a robust player-detection method for sports imagery using CNNs and compared it with the Viola-Jones AdaBoost framework. Tested on over 6,000 frames from FIFA World Cup matches, CNNs achieved 91.08% accuracy, compared with AdaBoost's 85.41%, particularly excelling in scenarios involving motion blur. Their work demonstrates the superior accuracy of deep learning techniques over traditional methods for dynamic environments, making a strong case for integrating CNN-based detection in the cricket dataset pipeline.

Rehman et al. (2004), [2] Examined the influence of illumination on face verification using SVM and k-nearest neighbor (k-NN) classifiers. Experimental results showed that SVMs, particularly with linear kernels, outperformed k-NN in both controlled and uncontrolled lighting environments, achieving recognition rates of up to 90%. Their findings reinforce the importance of preprocessing and lighting normalization in real-world sports imagery, making SVMs a suitable candidate for preliminary classification stages in cricket image datasets.

Zahid Mahmood et al. (2012), [3] Developed an augmented reality sports broadcasting application for automatic player detection and recognition. Their system used Haar-like features and AdaBoost for player and face detection and employed LDA with nearest center classification for robust face recognition. Tested on baseball datasets, the system achieved 100% player detection and up to 99.13% face detection accuracy. Notably, it worked on low-resolution images (down to 5x5 pixels) but struggled under complete occlusion and poor image quality. This system is especially relevant for enhancing live match experiences via smart devices.

Mahmood et al. (2015), [4] Contributed another layer to the literature by combining AdaBoost-based player and face detection with LDA and nearest-neighbor classification for recognition. Their system achieved 100% player detection and up to 100% face recognition under ideal conditions. Despite being tested on baseball footage, the methodology is adaptable to cricket, especially when annotated datasets are sparse. Limitations included reliance on frontal or near-frontal facial images and sensitivity to occlusion. Nevertheless, the implementation achieved high system accuracy (~88.88%) and acceptable computational efficiency (~3 seconds per execution), suggesting its feasibility for semi-real-time cricket applications.

Chugh et al. (2012), [5] Investigated the psychological dimension of cricket players through the lens of Rotter's Locus of Control (LOC) scale. Data collected from 148 players revealed that international-level athletes, especially fast-bowling all-rounders, exhibited strong internal LOC—believing performance outcomes stemmed from their own actions rather than external factors. This psychological profile, while not technical, can inform metadata enrichment of a cricket legend's dataset by correlating visual data (e.g., focused facial expressions) with psychological metrics. Mahajan et al. (2020), [6] Proposed a music recommendation system based on facial emotion detection, using Haar cascades for face detection and SVMs for emotion classification. Although outside the sports domain, this system demonstrates how emotion recognition from facial cues can be applied to multimedia interaction. The methodology is adaptable to sports by detecting stress, excitement, and fatigue on players' faces, thereby enhancing the contextual annotation of cricket images.

Kumar et al. (2019), [7] Provided a comprehensive review of face detection techniques, comparing models such as PCA, SVMs, Haarcascades, and deep learning approaches across datasets including MIT, FERET, and SCface. While this work was not cricket-specific, it highlighted universal challenges in facial recognition, such as low resolution, occlusion, and non-uniform illumination—challenges prevalent in cricket match footage. The authors emphasized the trade-offs between real-time performance and detection accuracy, which is vital for selecting optimal models for the proposed dataset. Prakash and Singh (2020), [8] Conducted a comparative statistical study focusing on legendary Indian cricketers Sachin Tendulkar and Virat Kohli. Using match data from ODIs and Tests, the study evaluated performance across parameters such as average, strike rate, and match-winning ability. Although no machine learning models were employed, the study emphasized Kohli's superior performance under pressure and adaptability in modern cricket formats. This paper provides valuable insight into metadata attributes—like form, consistency, and adaptability—that could be associated with visual data in an image dataset.

Mahmood et al. (2024) [9] introduced an augmented reality-based system for real-time player face detection and recognition, designed explicitly for cricket games. The model integrates Haar-like features and AdaBoost for player and face detection, followed by a PAL (Preprocessing + AdaBoost-LDA) framework for face recognition. The system performed robustly across varying lighting, pose, and occlusion conditions, achieving face recognition accuracies of 67%-97%. However, recognition accuracy dropped with severe occlusion, and the system was computationally expensive on large datasets. The system demonstrated an ability to handle low-resolution inputs as small as 20x30 pixels, making it highly relevant for archival cricket footage where high-definition imagery might be unavailable. Runswick et al. (2023), [10] explored the application of Virtual Reality (VR) and 360-degree video in cricket training and testing. Thirty-nine players assessed the realism, interactivity, and visual clarity of both media. VR provided superior interactivity, physical effort simulation, and realism, whereas 360-degree video offered better spatial awareness and cost-effectiveness. These findings support integrating such technologies with facial recognition systems for immersive analysis or coaching, where datasets like the cricket legends project could provide foundational visuals for training simulations.

Methodology

The proposed framework for classifying cricket legends is grounded in deep learning, specifically, convolutional neural networks (CNNs). Our approach follows a transfer learning strategy, leveraging the VGG-16 architecture, which has been widely recognized for its ability to extract rich, discriminative image features. The methodology consists of five key stages: data collection and preprocessing, model selection, adaptation of the pre-trained network, training and validation, and evaluation.

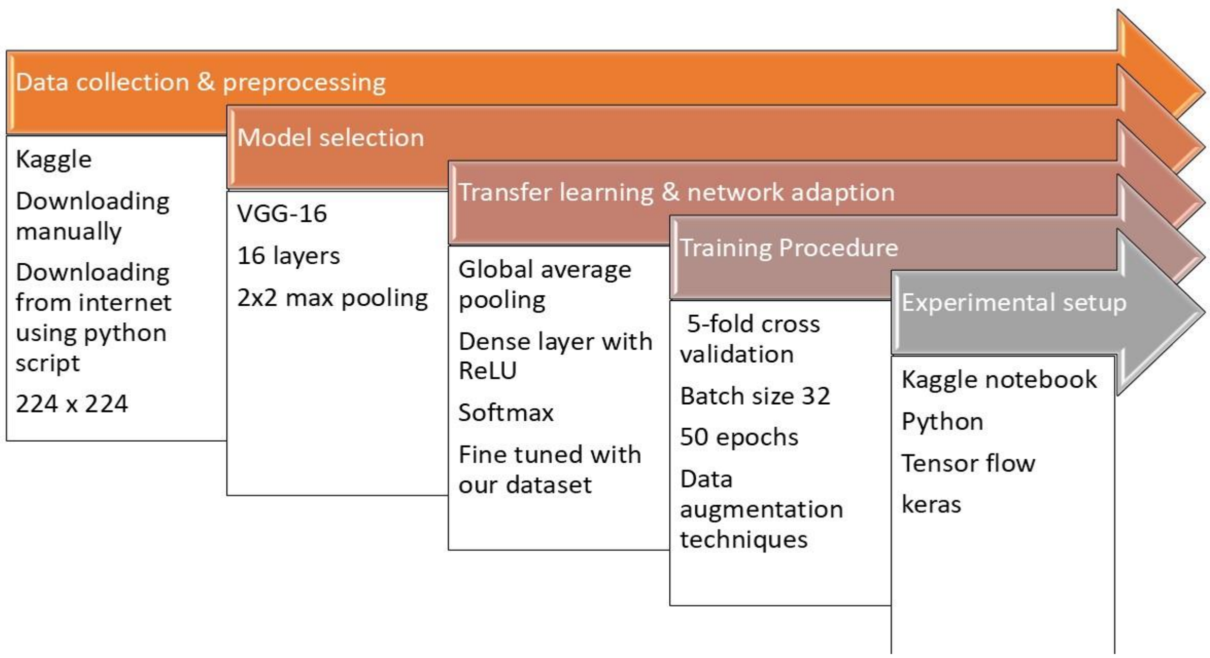

Figure 1: *overview of proposed Methodology*

Data Collection and Preprocessing

Building a reliable dataset is often one of the most challenging and time-consuming stages of any machine learning project. Unlike generic objects widely available online, creating a specialized dataset for a niche category such as cricket legends requires a deliberate, systematic approach. In our case, the task was to assemble an extensive collection of images of 30 legendary cricketers, ensuring sufficient variation in angles, lighting, and context to train a deep learning model effectively. Initially, we explored publicly available resources and identified a Kaggle dataset containing 30 cricketer classes. However, the number of images per class was low and insufficient for robust training. To address this limitation, we adopted multiple strategies to expand the dataset. At first, browser extensions were used to scrape images from the web. However, this approach yielded minimal results since there were no dedicated repositories or websites with a wide variety of cricketer images. Recognizing this challenge, we shifted towards a more automated, scalable solution by employing Python-based scripts that used image crawlers to download images directly from the internet via keyword queries. This approach proved highly effective, enabling us to increase the dataset size significantly.

While this method enriched the dataset, it also introduced new challenges. Automated downloads often capture irrelevant or noisy images that match the keywords superficially but do not belong to the intended class. For example, a search for a specific player might also return images of unrelated people, team logos, or cricket equipment. To ensure dataset quality, a manual inspection and filtering step was carried out to remove inconsistent, duplicate, and incorrect images. This process is a common yet essential part of preparing real-world datasets, where raw data rarely comes in a perfectly curated form.

Another important aspect was the diversity of image resolutions. Since images were sourced from different platforms, their dimensions varied widely. Deep learning models such as VGG-16 require uniform input sizes, specifically 224 × 224 pixels. Therefore, the entire dataset of 22,817 images was resized to this fixed dimension. Performing this task manually would have been impractical, so we developed a Python script to automate the resizing process across all class folders. This not only standardized the model's input but also saved considerable time and effort.

Through these steps, we successfully constructed a comprehensive dataset of cricket legends that balances both scale and quality. The final dataset is structured into 30 distinct classes, each representing an individual player, with enough images per class to support practical training and evaluation. By combining automated image collection, manual refinement, and preprocessing techniques such as resizing, we ensured that the dataset was both robust and ready for use in deep learning experiments.

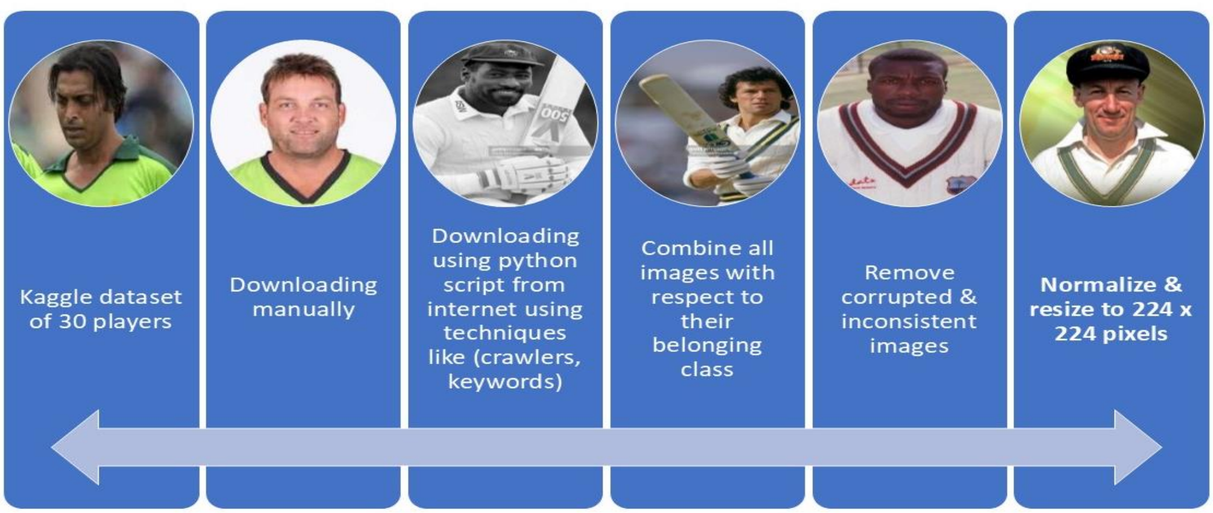

Figure 2: *Data collecting & preprocessing steps with some legend's images*

Model Selection

We selected VGG-16, developed by the Visual Geometry Group at the University of Oxford and introduced by Simonyan and Zisserman (2014), which is one of the most influential convolutional neural network architectures. VGG-16 is known for its depth and simplicity, with 16 layers that use uniform 3×3 convolutional kernels and 2×2 max pooling, as shown in Figure 3. A VGG-16 is originally trained on the ImageNet dataset over 1.2 million labeled images across 1,000 classes. Its strength lies in its ability to capture hierarchical image features, ranging from low-level edges and textures to high-level object representations. Given the complexity of facial and pose variations in cricketer images, VGG-16 provides a solid foundation for transfer learning. VGG is renowned for its simplicity and effectiveness as well as its ability to achieve strong performance on various computer vision tasks, including image classification and object recognition. This uniformity not only enhances feature extraction consistency but also simplifies model training. The convolutional operation can be mathematically expressed as:

$$Y(i,j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X(i+m, j+n) \cdot K(m,n)$$

Where X represents the input feature map, K is the convolutional kernel, and $Y(i,j)$Is the resulting activation at position?(i,j), the subsequent max-pooling layer performs spatial down-sampling according to:

$$P(i,j) = \max_{(m,n)\in R} Y(i+m, j+n)$$

Where R defines the local pooling region (typically 2×2These operations allow the network to progressively capture low-level features, such as edges and corners, in the early layers, and high-level semantic representations (such as facial structure or jersey patterns) in deeper layers.

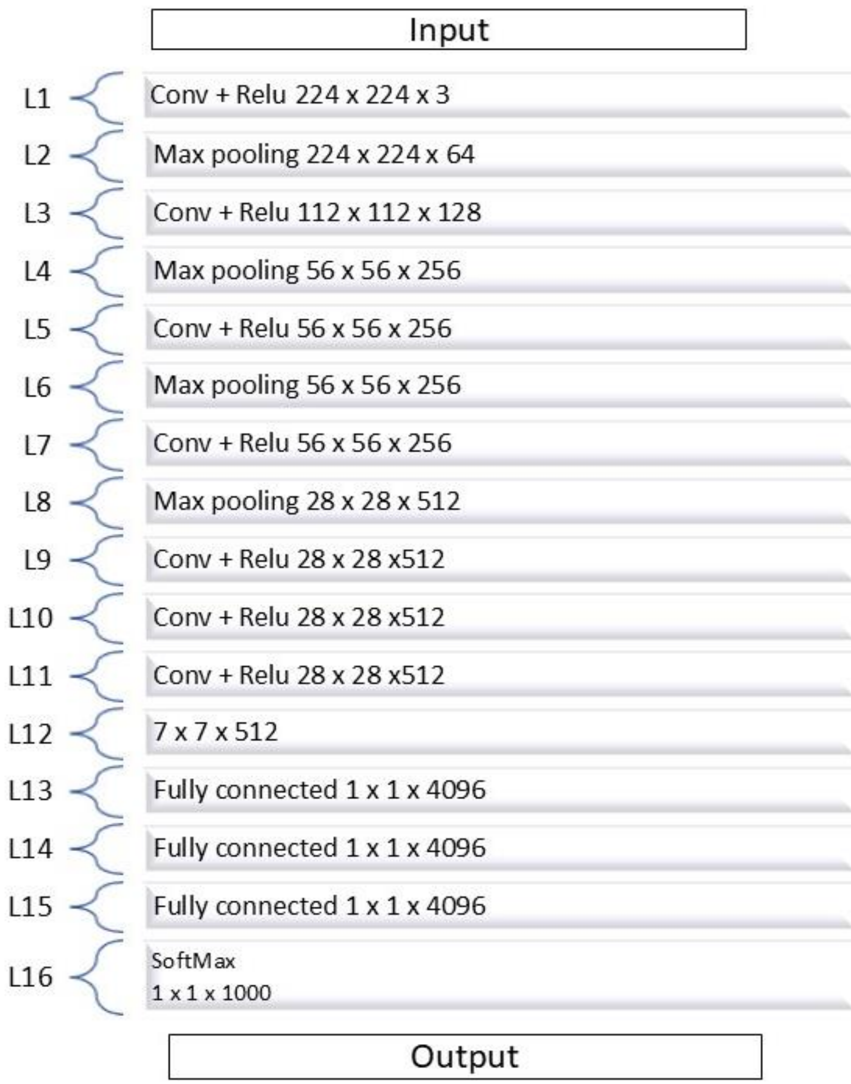

Figure 3: *VGG-16 Architecture*

Transfer Learning and Network Adaptation

Instead of training the network from scratch, which would require an impractically large dataset, we adopted a transfer learning approach. The lower convolutional layers of VGG-16 were retained to leverage their pre-trained feature extraction capabilities. The fully connected layers at the top were replaced with new layers tailored for our classification task. Specifically:

- A Global Average Pooling (GAP) layer was used to reduce the feature maps into a compact representation.
- A dense layer with ReLU activation followed this to introduce non-linearity.
- A final SoftMax layer was employed to output probabilities across the 30 classes, each corresponding to a cricket legend.

During training, the early convolutional layers were frozen to preserve the generic features learned from ImageNet, while the higher layers were fine-tuned on our dataset. This balance allowed the model to adapt effectively to cricket legend recognition without overfitting.

Training Procedure

The dataset of 22,817 images was split into five folds using 5-fold cross-validation, ensuring that each model iteration was trained and tested on different subsets. This process improves the robustness of results and provides a more reliable estimate of generalization performance. The following training configuration was adopted:

- Optimizer: Adam, with a learning rate of 0.0001.
- Batch Size: 32.
- Epochs: 50.
- Loss Function: Categorical cross-entropy.
- To further enhance performance, data augmentation techniques such as horizontal flipping, rotation, and zooming were applied during training. This helped the model generalize better to unseen variations of cricketer images.

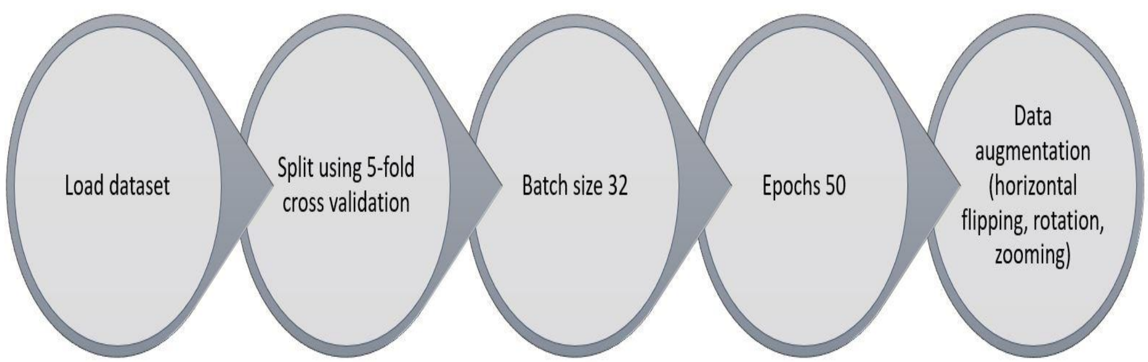

Figure 4: *Training process*

Experimental Environment

All experiments were conducted in Python using TensorFlow and Keras. Training was performed in a cloud-based GPU environment (Kaggle), which provided the computational resources needed to train deep CNNs efficiently. The environment setup ensured reproducibility and consistency across experiments.

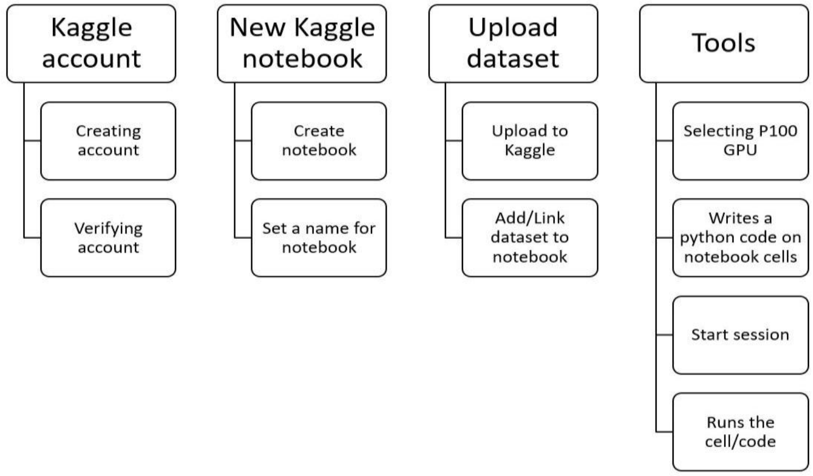

Figure 5: *Steps for setting the environment for model training*

cronym	Definition	Acronym	Definition
A_G	Adam_Gilchrist	K_S	Kumar_Sangakkara
A_C	Alastair_Cook	M_J	Mahela_Jayawardene
A_D	Allan_Donald	M_D	Ms_Dhoni
B_L	Brian_Lara	M_M	Muttiah_Muralitharan
C_G	Chris_Gayle	R_D	Rahul_Dravid
C_A	Curtly_Ambrose	R_H	Richard_Hadlee
D_S	Dale_Steyn	R_P	Ricky_Ponting
D_B	Don_Bradman	S_T	Sachin_Tendulkar
G_S	Garfield_Sobers	S_W	Shane_Warne
G_M	Glenn_Mcgrath	S_A	Shoaib_Akhtar
I_B	Ian_Botham	S_W	Steve_Waugh
I_M	Imran_Khan	S_G	Sunil_Gavaskar
J_A	James_Anderson	V_K	Virat_Kohli
J_K	Jaques_Kallis	V_R	Viv_Richards
K_D	Kapil_Dev	W_A	Wasim_Akram

Table 1 *List of abbreviations for Sportsmen*

Results

The performance of our VGG-16-based image classification model was rigorously evaluated using a 5-fold stratified cross-validation approach. This methodology ensures that the reported results are robust and representative of the model's performance across different subsets of the dataset.

The model achieved an outstanding average classification accuracy of 94.37% across all five folds, with a low standard deviation of ±0.39%. This minimal variance indicates that the model's performance is highly consistent and stable, demonstrating its reliability on different data splits. The highest accuracy observed in any single fold was 94.76%, while the lowest was 93.97%. To provide a more detailed analysis, a composite confusion matrix was generated by averaging the results from all five folds, as shown in the accompanying Figure 6. The matrix visually confirms the model's strong performance, with the highest values along the diagonal, indicating correct classifications.

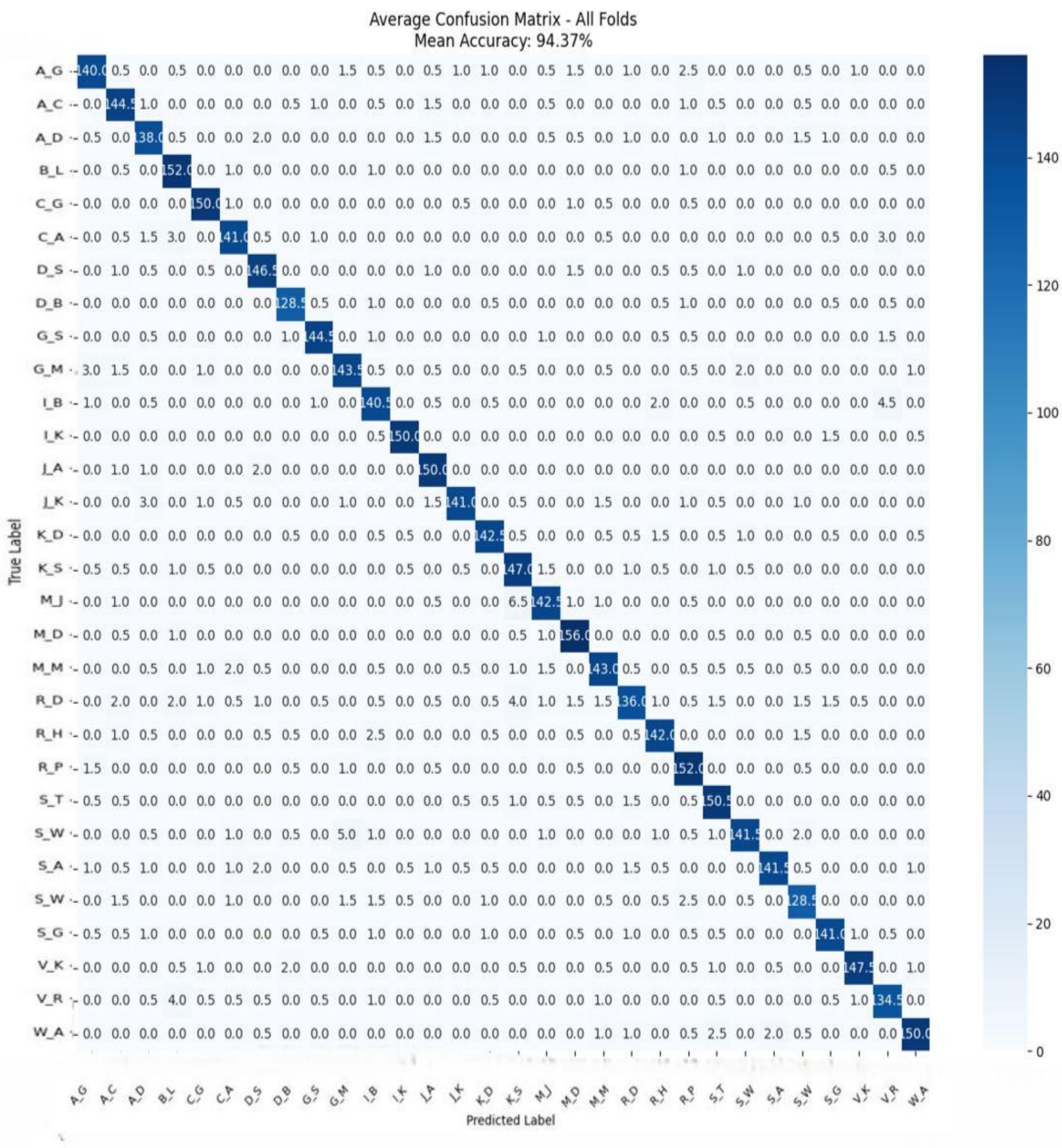

Figure 6: *Average Confusion Matrix*

The per-class performance metrics reveal the model's effectiveness in identifying individual players. For instance, the model demonstrated near-perfect performance in classifying images of Imran Khan, with average precision, recall, and F1-score over 98% (Figure 7). Similarly, players like Chris Gayle and Don Bradman were also identified with very high accuracy. While the model performed exceptionally well overall, some classes, such as Rahul Dravid (with an average recall of 85.5%), showed slightly lower performance, indicating minor confusion with other players. This is likely due to subtle visual similarities or variations in image quality across specific subjects.

Per Class Performance

In Figure 7, the performance of each class is shown, including its class name and precision, recall, and F1 score, demonstrating the model's learning effectiveness. The model shows high precision, recall, and F1 scores of 98% for Imran Khan and James Anderson. Moreover, the performance of the overall classes is shown in Figure 7 below.

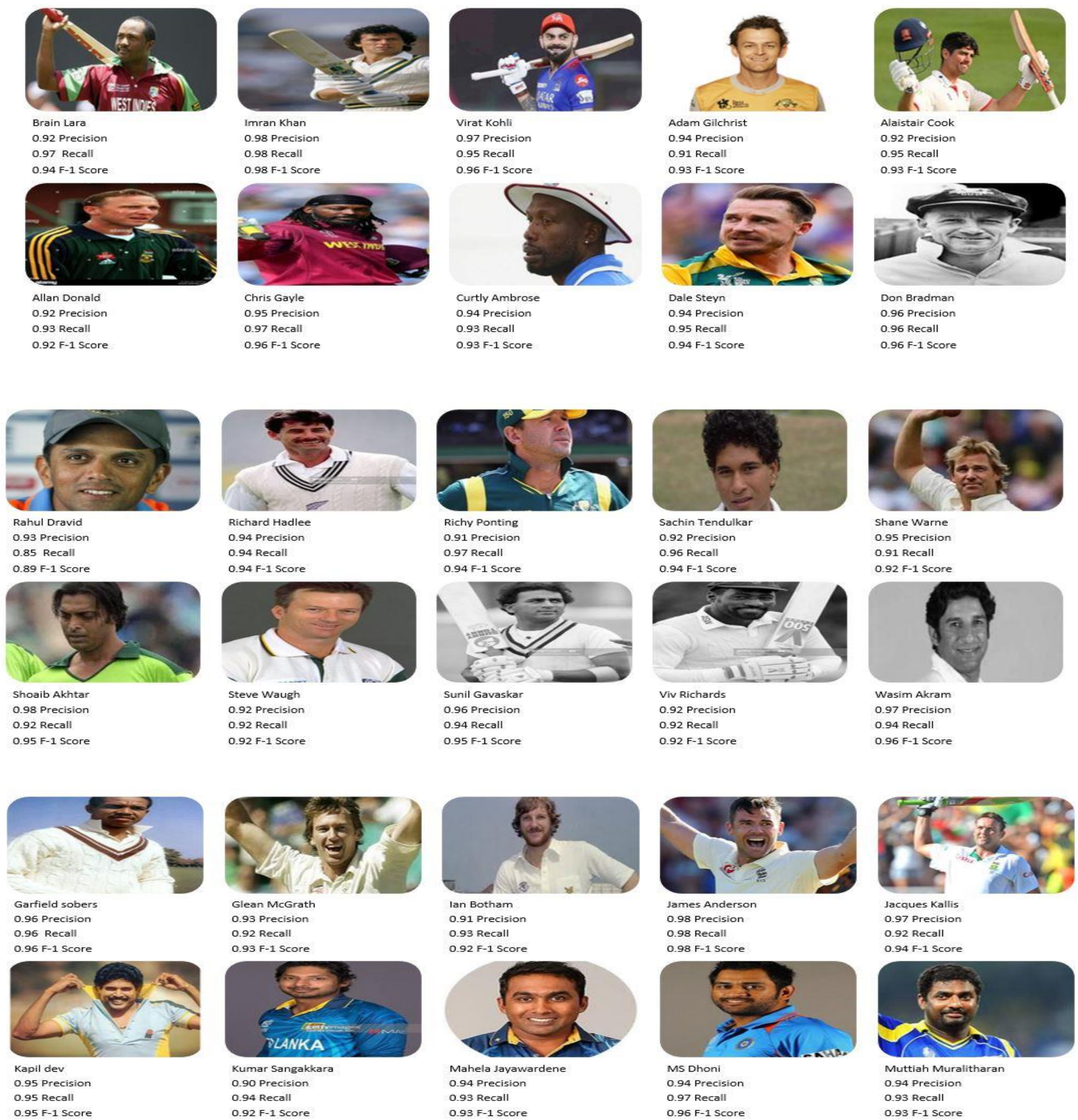

***Figure** 7 Average Per-Class Performance*

Cross-Validation Performance Analysis

The model's performance was systematically evaluated using a 5-fold cross-validation procedure to ensure robustness and generalizability of the results. Each fold was carefully analyzed for its training dynamics, convergence behavior, and classification effectiveness. The evaluation relied on the progression of accuracy and loss metrics across epochs, supported by confusion matrices that provide insight into the model's ability to distinguish among multiple classes. Training in Fold 1 spanned 50 epochs in two phases. Phase 1 has 15 epochs, and phase 2 has 35, demonstrating a stable, consistent learning process. The training accuracy steadily improved from approximately 0.74 to around 0.83 by the final epoch. Validation accuracy showed stronger performance, stabilizing early at 0.88-0.92, suggesting that the model quickly captured features with strong generalization potential. The loss curves aligned with these observations: training loss decreased gradually from 1.0 to 0.7, while validation loss started lower at about 0.65 and fell further to roughly 0.5. The persistent gap between the training and validation losses—despite excellent validation accuracy—suggests mild underfitting to the training data. Nevertheless, the high validation accuracy indicates effective learning. The confusion matrix displayed a strong diagonal structure, confirming high accuracy across most classes, with only minor off-diagonal misclassifications typical of multi-class problems.

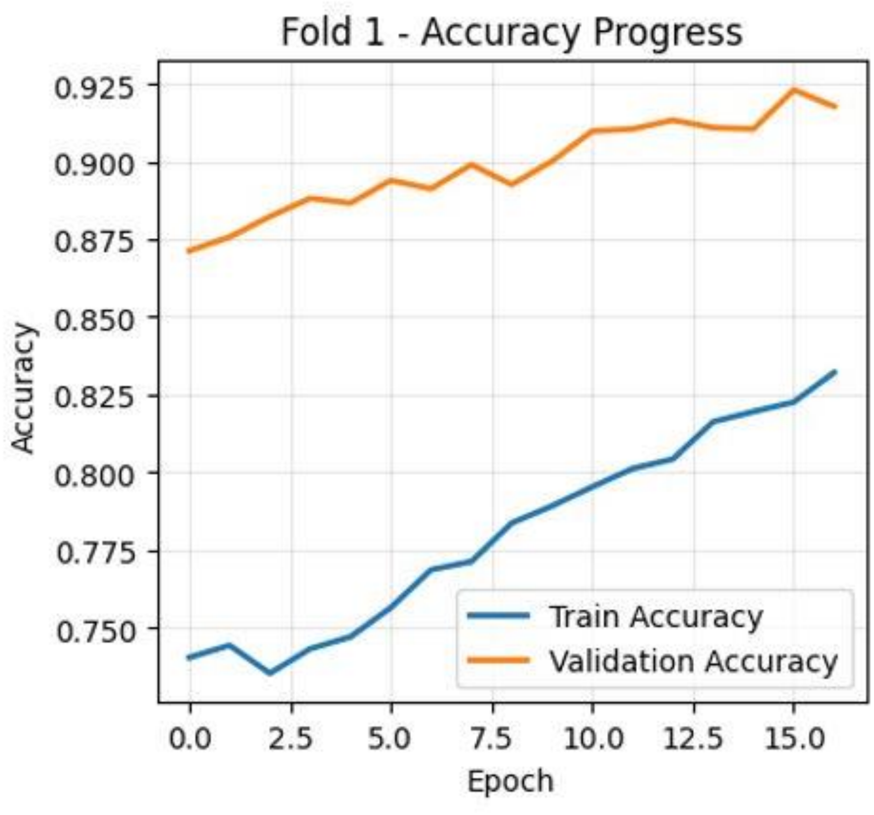

Figure 8: *Fold 1 Accuracy Progress diagram*

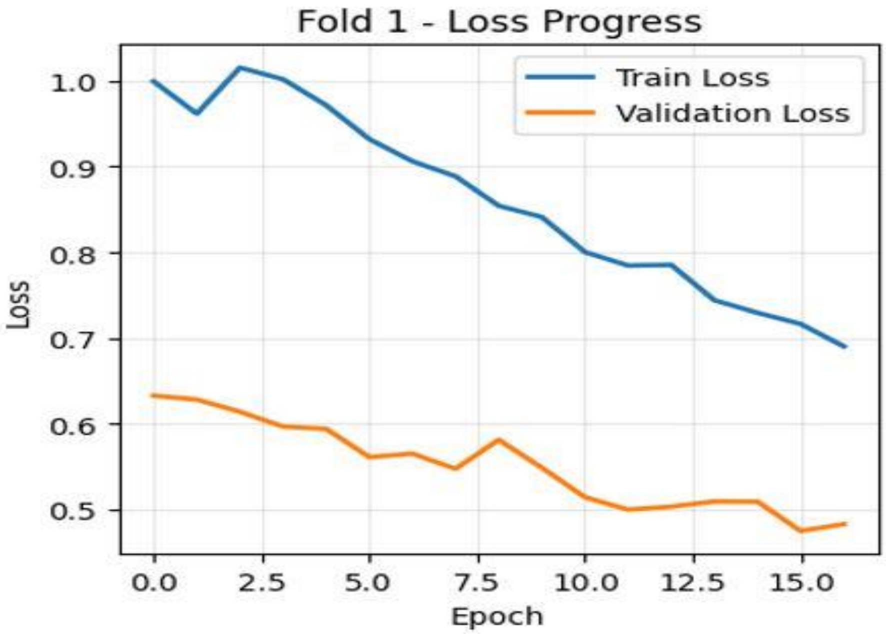

Figure 9: *Fold 1 Loss Progress diagram*

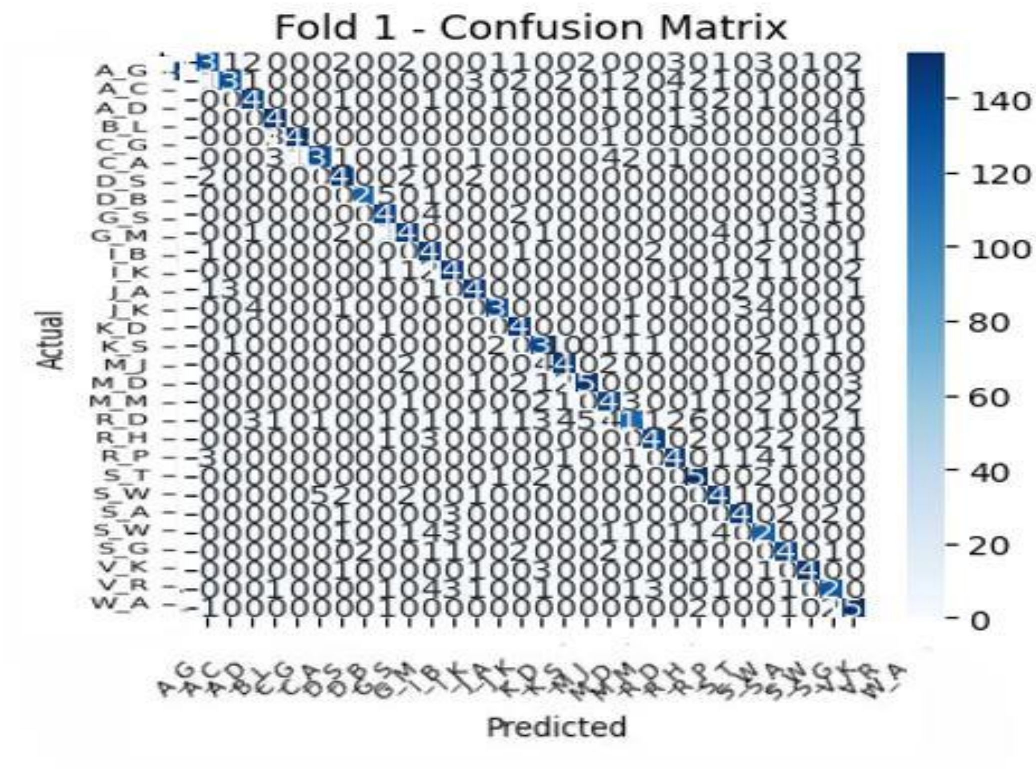

Figure 10: *Fold one confusion Matrix*

Fold 2 showed a well-balanced, effective training trajectory over 30 epochs. Both training and validation accuracies increased rapidly from initial values of around 0.15–0.20 to above 0.90 by the end, as shown in Figure 11. Interestingly, the validation accuracy remained slightly higher than the training accuracy during the early epochs, then converged closely around epoch 20, demonstrating excellent generalization without overfitting. The loss value reflected this pattern: both training and validation losses dropped sharply from above 3.0 to below 1.0, with minimal divergence throughout training. This indicates a stable learning rate and appropriate model complexity for the data in this fold. The confusion matrix reinforced these findings, with a clear diagonal pattern reflecting strong predictive accuracy and low error rates across all classes.

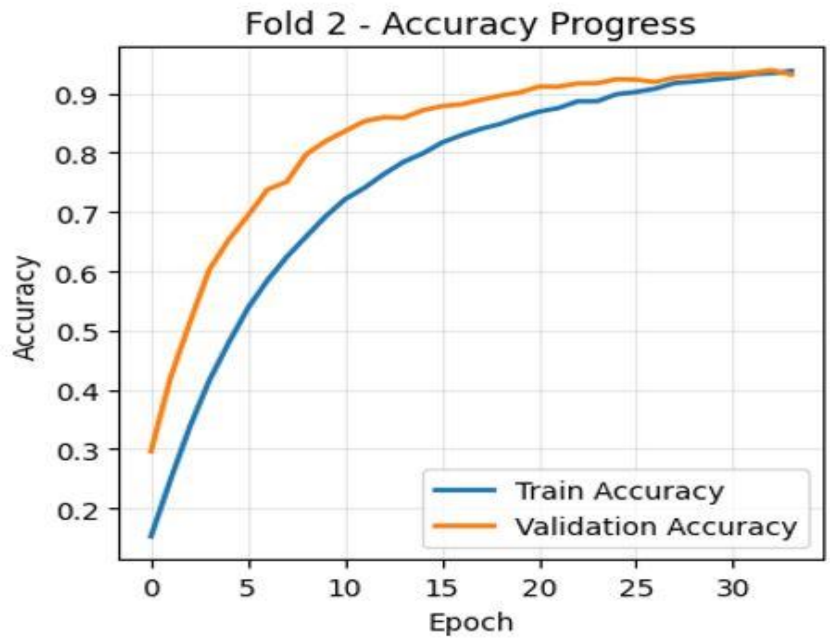

Figure 11: *Fold 2 Accuracy Progress diagram*

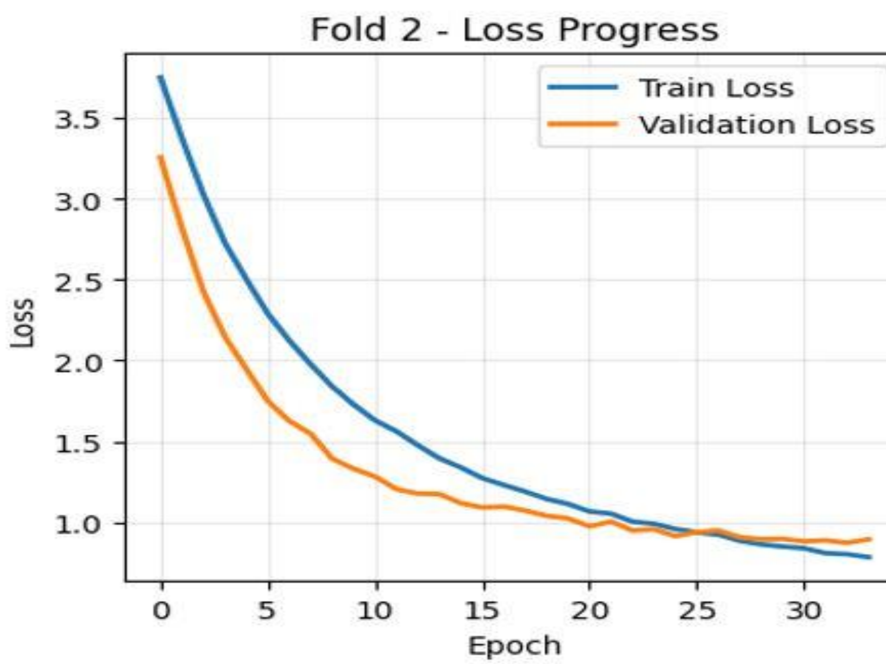

Figure 12: *Fold 2 Loss Progress diagram*

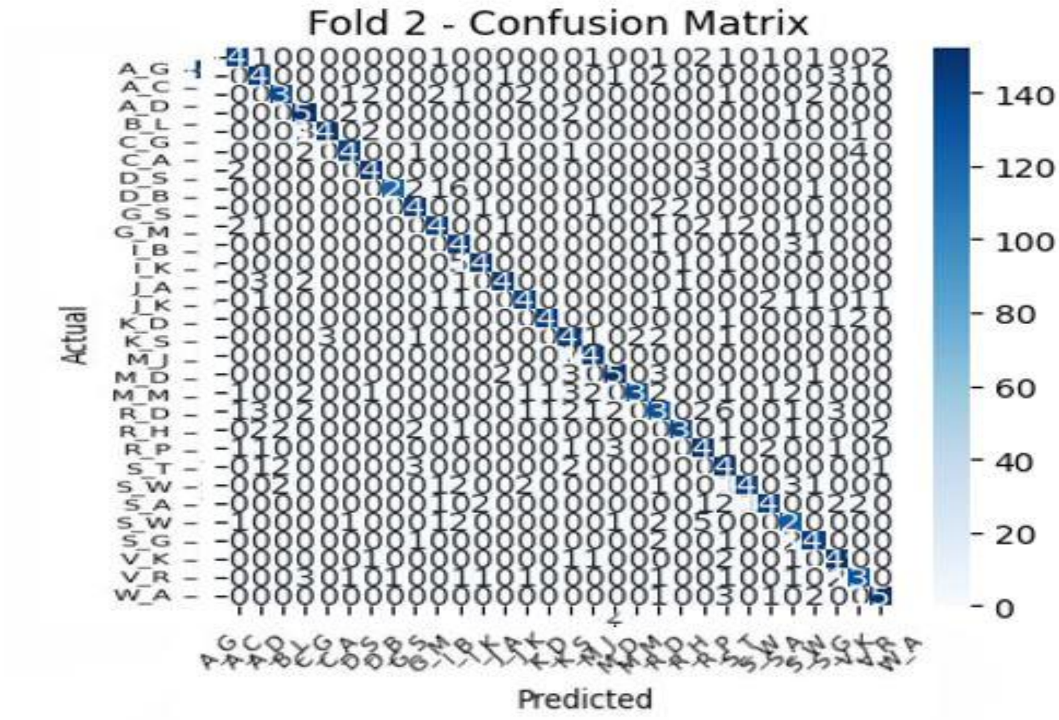

Figure 13: *Fold two confusion matrix*

The training dynamics of Fold 3 were comparatively volatile in the initial epochs but stabilized quickly. Validation accuracy began high at approximately 0.936, dipped slightly to 0.932, and then recovered to a peak of 0.938. At the same time, training accuracy started around 0.929 and gradually increased, surpassing validation accuracy by the final epoch. Loss curves followed a complementary trend: validation loss remained relatively steady, decreasing slightly from 0.8 to 0.79, whereas training loss declined from 0.72 to 0.69. This convergence indicates that the model progressively optimizing its weight while maintaining a balanced generalization capacity. The confusion matrix for this fold revealed a sharp diagonal structure, confirming that the model achieved high precision and recall despite the shorter training cycle.

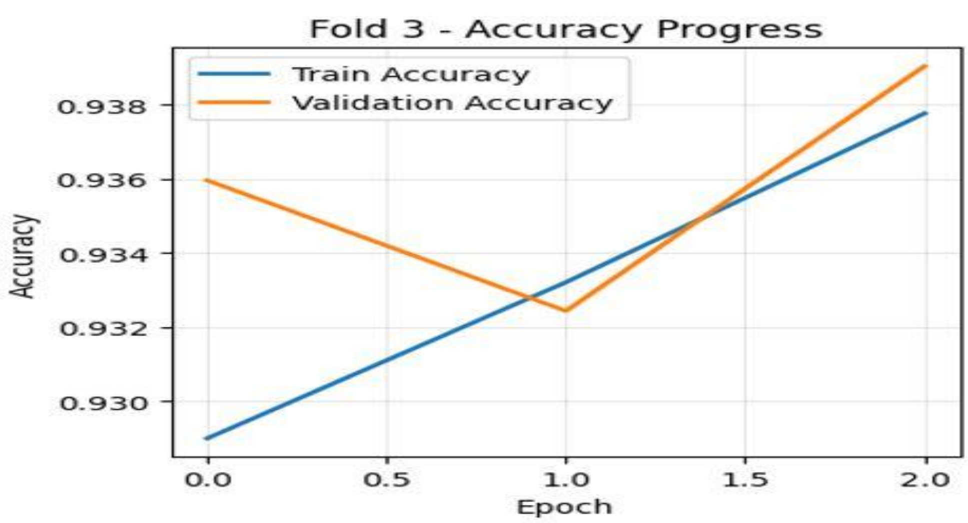

Figure 14: *Fold 3 Accuracy Progress diagram*

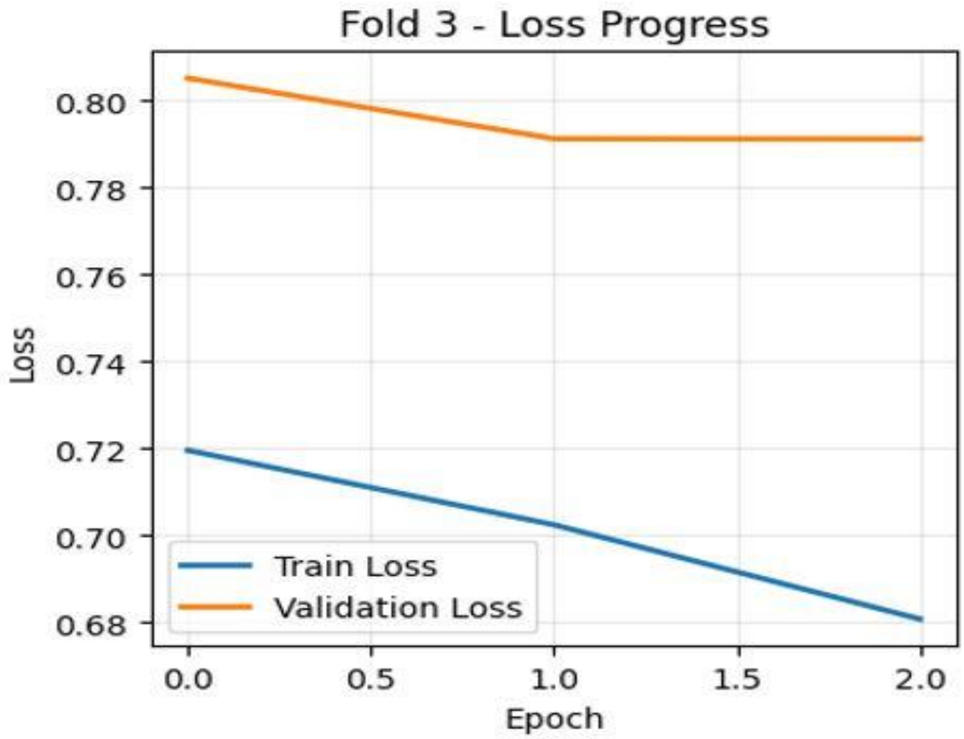

Figure 15: *Fold three loss Progress diagram*

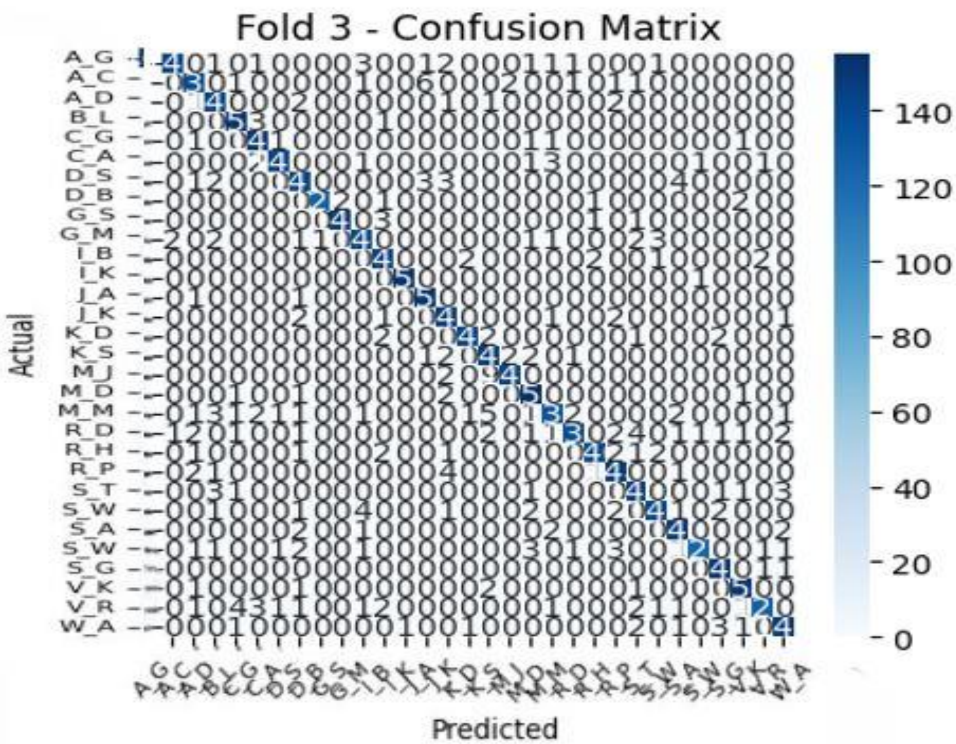

Figure 16: *Fold three confusion matrix*

Training performance in Fold 4 was conducted over fifty epochs. In Figure 17, the results after four episodes showed steady, moderate improvement. The training accuracy increased from approximately 0.925 to 0.935, while validation accuracy remained consistently higher, starting at 0.943 and peaking near 0.950. Although it is uncommon for validation accuracy to exceed training accuracy, this suggests that the validation subset contained examples that the model found easier to classify. Both training and validation losses showed a closely aligned downward trend, with training loss decreasing from 0.85 to 0.795 and validation loss fluctuating slightly around 0.85 before settling near 0.84. The near-parallel nature of these curves indicates balanced learning without overfitting. The confusion matrix showed strong diagonal dominance, confirming the model's reliability in distinguishing between classes in this subset.

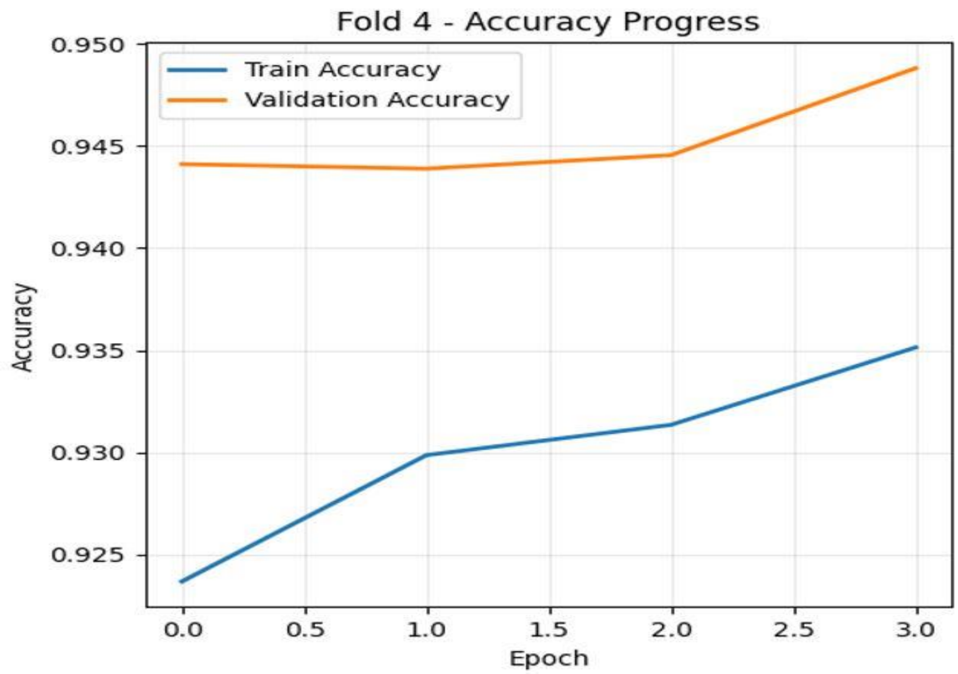

Figure 17: *Fold 4 Accuracy Progress diagram*

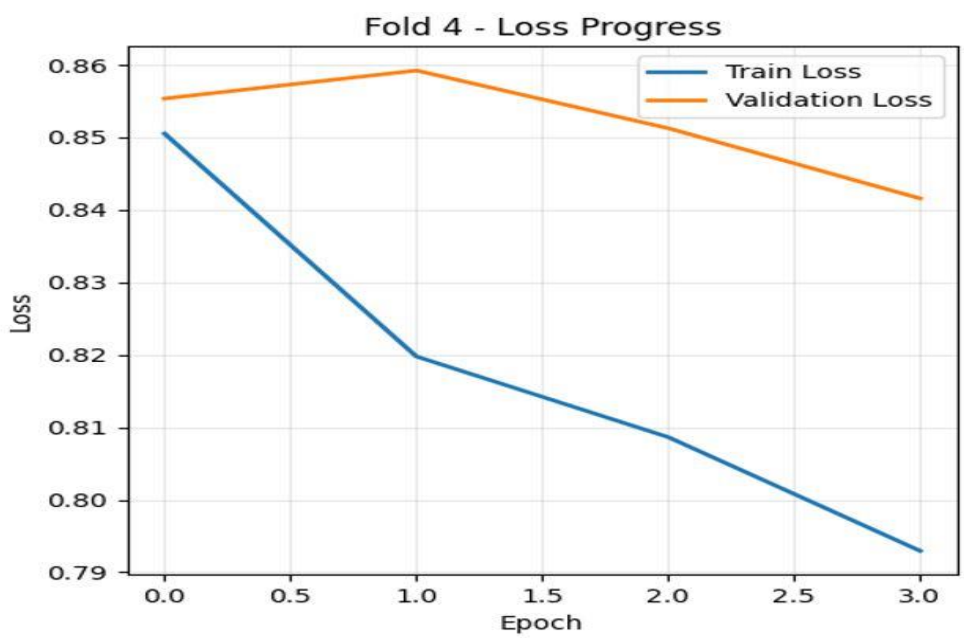

Figure 18: *Fold four loss Progress diagram*

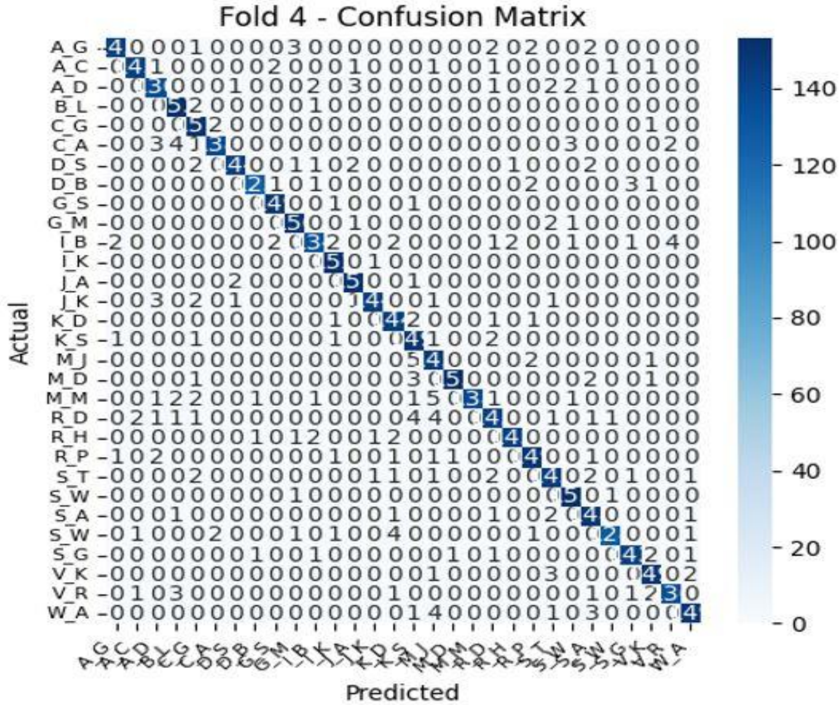

Figure 19: *Fold four confusion matrix*

Fold 5, trained for 50 epochs with 2 phases, where phase 1 has 15 epochs and phase 2 has 35 epochs. In Figure 20, fold 5 displayed the most classical and stable learning behavior. Both training and validation accuracy started below 0.10 and rose steadily in tandem, eventually stabilizing above 0.85. The consistent alignment between the two curves throughout training confirmed that the model achieved strong generalization while avoiding overfitting. Loss progression supported this observation: training and validation losses began at approximately 4.0 and 3.5, respectively, and decayed exponentially before converging near 0.8. This reflects efficient parameter optimization and well-calibrated training dynamics. The confusion matrix for Fold 5, with a large number of classes, showed a clear diagonal pattern with minimal noise, confirming the model's ability to discriminate among numerous distinct categories accurately.

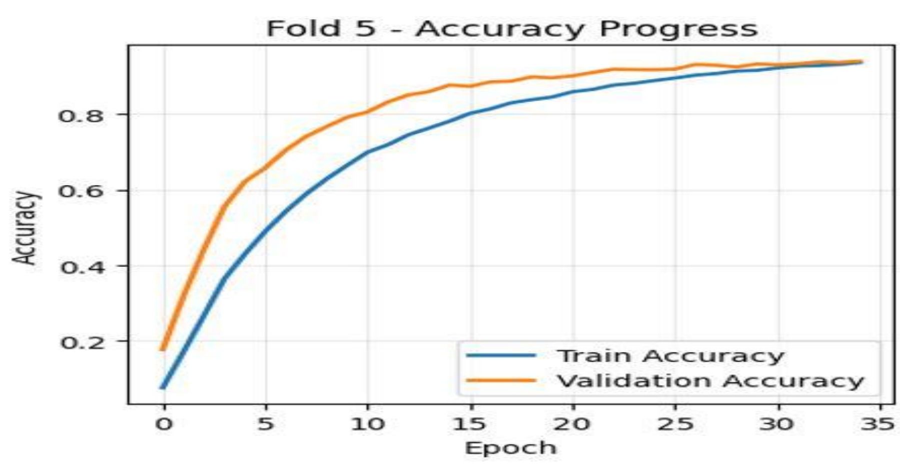

Figure 20: *Fold 5 Accuracy Progress diagram*

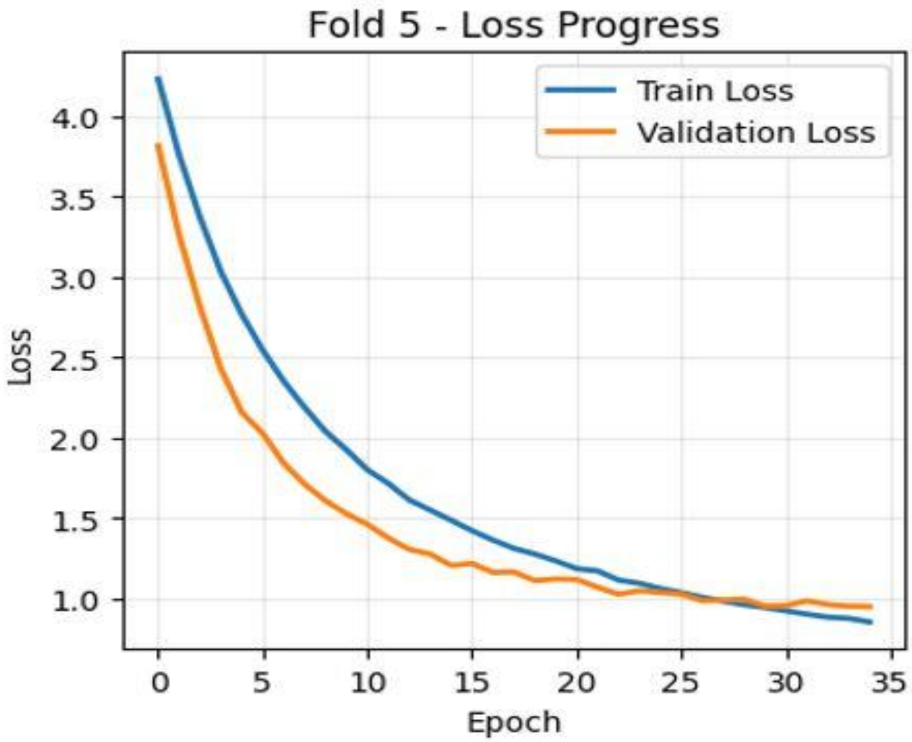

***Figure 21:** Fold five loss Progress diagram*

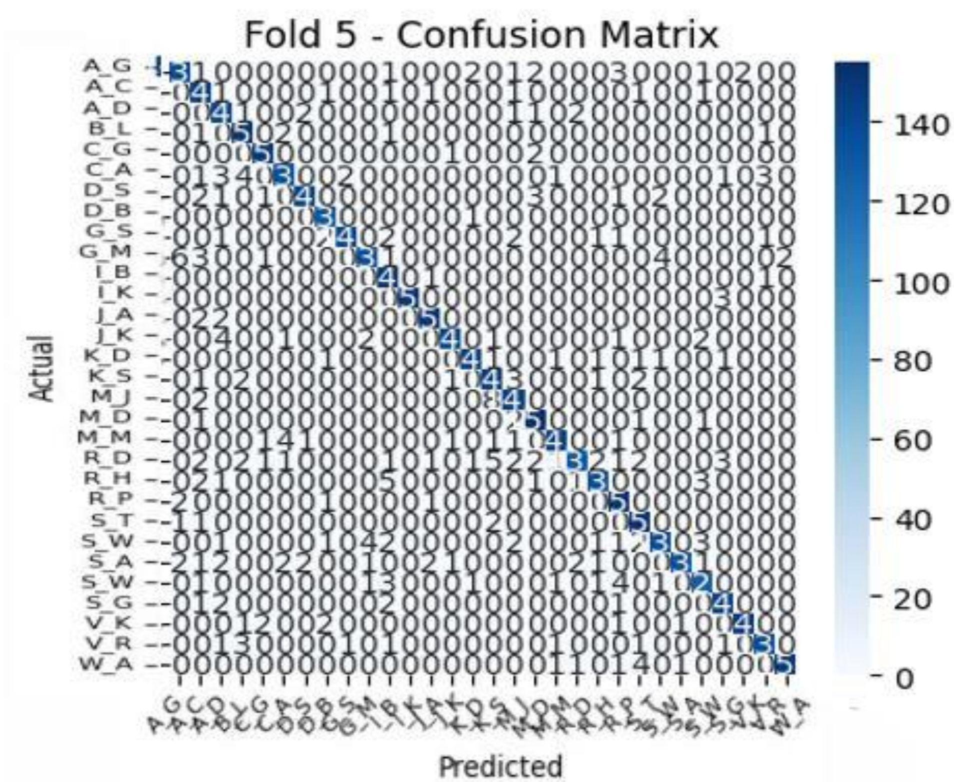

***Figure 22:** Fold five confusion matrix*

Discussion

The high accuracy achieved in this study underscores the effectiveness of our methodology for identifying and classifying cricket legends from images. The use of a transfer learning approach with the pre-trained VGG-16 model was a critical factor in this success. By leveraging the feature-extraction capabilities of a network trained on a massive dataset such as ImageNet, our model quickly and effectively learned the distinguishing visual characteristics of each cricketer. The robust data augmentation techniques employed during training were also instrumental in preventing overfitting and enhancing the model's ability to generalize to new, unseen images. By artificially increasing the diversity of the training data through rotations, shifts, and flips, the model learned to recognize players regardless of pose or minor image variations.

The results demonstrate the practical viability of deep learning for automating complex visual tasks in sports. The system's high accuracy and reliability suggest its potential for real-world applications, such as populating digital archives, enhancing fan engagement platforms with player recognition features, or supporting media analytics. While the overall performance is highly encouraging, the minor performance dip observed for certain players highlights potential areas for future research. A deeper analysis of the specific images causing misclassifications could lead to targeted improvements, such as adding more varied training data for those individuals. Nevertheless, this work provides a strong foundation and a clear benchmark for computer vision-based sports analysis, demonstrating that even with a challenging, large-scale dataset, highly accurate classification is achievable.

Conclusion

In this study, we successfully developed and implemented a robust, highly accurate deep learning framework for automated classification of legendary cricketers from a large-scale image dataset. Addressing the inherent challenges of manually archiving and tagging historical sports imagery, our work provides a scalable, reliable solution grounded in modern computer vision. The core of our approach was to leverage transfer learning with the VGG-16 convolutional neural network. By fine-tuning this robust, pre-trained architecture on a meticulously curated dataset of over 22,817 images spanning 30 iconic players, we enabled the model to learn the specific visual nuances required for this complex task. The rigor of our evaluation was ensured through a 5-fold cross-validation, which provided a comprehensive and unbiased assessment of the model's performance.

The results unequivocally demonstrate the effectiveness of this system, which achieved an outstanding average classification accuracy of 94.37% with minimal variance. This high level of performance not only validates our methodology but also underscores the immense potential of artificial intelligence to transform how we manage and interact with sports history. The framework we have presented offers tangible value for applications in sports media analytics, enriching fan experiences, and creating efficient, searchable digital archives. Ultimately, this research serves as a significant contribution to the field of computer vision and its application in cultural and entertainment domains, bridging the gap between technological innovation and the celebration of sporting greatness.

References

1.Lehuger, A., S. Duffner, and C. Garcia, *A robust method for automatic player detection in sport videos.* Orange Labs, 2007. **4**.

2.Conde, C., A. Ruiz, and E. Cabello, *Face verification using SVM: influence of illumination*, in *Applications and Science in Soft Computing*. 2004, Springer. p. 301-308.

3.Zahid, M., T. Ali, and S. Khattak, *Automatic player detection and recognition in images using AdaBoost.* Electrical Engineering, Mathematics & Computer Science faculty, Twente University, 2012.

4.Mahmood, Z., et al., *Automatic player detection and identification for sports entertainment applications.* Pattern Analysis and Applications, 2015. **18**(4): p. 971-982.

5.Chugh, G., R. Kalw, and P. Jha, *Locus of control of cricket players.* SPORT-Science & Practice, 2012. **2**(5): p. 29-38.
6.Mahajan, P., P. Khurad, and P. Chaudhari. *Face Player: Facial Emotion-Based Music Player*. 2020.
7.Kumar, A., A. Kaur, and M. Kumar, *Face detection techniques: a review.* Artificial Intelligence Review, 2019. **52**(2): p. 927-948.
8.KOHLI, T.V., *A COMPARATIVE STUDY BETWEEN THE CRICKET LEGENDS OF INDIA.* 2020.
9.Haq, M.U., et al., *Automatic player face detection and recognition for players in cricket games.* IEEE Access, 2024. **12**: p. 41219-41233.
10.Runswick, O.R., *Player perceptions of face validity and fidelity in 360-video and virtual reality cricket.* Journal of Sport and Exercise Psychology, 2023. **45**(6): p. 347-354.
11.*Tammina, S. (2019). Transfer learning using vgg-16 with deep convolutional neural network for classifying images. International Journal of Scientific and Research.*
12.*Shaha, M., & Pawar, M. (2018). Transfer learning for image classification. 2018 Second International Conference.*
13.*Jin, X., Du, X., & Sun, H. (2021). VGG-S: An Improved small sample image recognition model based on VGG16. 2021 3rd International Conference.*
14.*Tiwari, V., Pandey, C., Dwivedi, A., et al. (2020). Image classification using a deep neural network. 2020 2nd International.*
15.*Reja, S. A., & Rahman, M. M. (2021). Sports recognition using a convolutional neural network with optimization techniques from images and live streams. International Journal.*
16.*Das, S., Singha, R., Das, K., et al. (2025). Sports Image Classification Using Deep Learning Models. Conference on Quantum.*
17.*Farhad, M. Y., Hossain, S., Tanvir, M. D. R. K., et al. (2020). Sports-net18: Various sports classification using transfer learning. For Industry 4.0 (STI).*
18.*Liu, J., Carr, P., Collins, R. T., & Liu, Y. (2013). Tracking sports players with context-conditioned motion models. Proceedings of the IEEE.*
19.*Breen, C., Khan, L., & Ponnusamy, A. (2002). Image classification using neural networks and ontologies. 13th International Workshop.*
20.*Lu, K., Chen, J., Little, J. J., & He, H. (2018). Lightweight convolutional neural networks for player detection and classification. Computer Vision and Image Understanding. Elsevier.*
21.*Direkoglu, C., Sah, M., & O'Connor, N. E. (2018). Player detection in field sports. Machine Vision and Applications. Springer.*