© DEC 2023 | IRE Journals | Volume 7 Issue 6 | ISSN: 2456-8880

Convolutional Neural Networks (CNNs) for Facial Recognition with Black and White Images

ANIKET TRIPATHI[1], MIT PATEL[2], TEJAS MURTHAL[3], AJAY DODHI[4], NUPUR DOSHI[5]
[1, 2, 3, 4, 5]*Nirma University, Ahmedabad, India*

Abstract-* Face recognition is becoming the core technology in contemporary computer vision, with a broad variety of applications ranging from biometric verification to intelligent surveillance systems. While excellent results have been obtained using color (RGB) images, grayscale or black and white images continue to maintain importance in most areas, particularly where budget constraints, available infrastructure, or privacy concerns limit the availability of high-quality color information. This study analyzes the performance of Convolutional Neural Networks (CNNs) for facial recognition on black and white images. CNNs have been shown to excel in the domain of visual pattern recognition as a result of their ability to represent features hierarchically. With grayscale images, CNNs exploit edge, texture, and spatial features without the burdensome weighting of color, often yielding more computationally demanding models. The article introduces CNN models that are tailored to perform with grayscale face recognition, e.g., adjustments to guarantee accuracy even in the absence of chromatic information. Dominant issues such as reduction of richness of information, lighting variation, and contrast normalization are explained together with their counter-measures such as data augmentation and preprocessing techniques. The article also delineates some applications in real life where grayscale face recognition is particularly valuable, such as monitoring at low light, analysis of archives with images, and application in lowpower edge devices. The outcomes illustrate the efficiency and stability of CNN-based systems under grayscale conditions and demonstrate that color information, although beneficial, is not always necessary for accurate face recognition. The study concludes with the identification of the potential for investigating light-weight CNN architectures and transfer learning techniques specially for monochromatic datasets.

Indexed Terms-* Convolutional Neural Networks (CNNs), Facial Recognition, Grayscale Facial Images, Black and White Image Processing, Deep Learning for Face Recognition, Face Detection in Grayscale, CNN based Face Identification, Low-Resolution Face Recognition

I. INTRODUCTION

Facial recognition technology has emerged as one of the most groundbreaking computer vision technologies in a very short period, taking a central place in a wide variety of applications from device authentication and intelligent surveillance to border protection, forensics, as well as human computer interaction. Underlying all such developments stands the power of artificial intelligence, more specifically deep learning—a field that has made incredibly rapid progress in recent years. Of all deep learning architectures, Convolutional Neural Networks (CNNs) have proven most efficient as well as most popular image-based recognition models due to their ability to learn spatial hierarchies as well as to extract complex features from visual information. The architectures have revolutionized facial recognition by making automatic feature extraction from a face as well as high-accuracy identification in diverse conditions possible.

Most research work and practical applications in face recognition have conventionally been handled working with color images (RGB), which have high amounts of information in terms of color cues. Not only does an image in color represent structural information of a face, but, in addition, it contains information about hue as well as saturation, which can differentiate among individuals, given controlled conditions. Use of color information, however, presents a challenge in cases where availability or usage of RGB information is not possible. For most practical applications, black-and-white (grayscale) imagery still persists, such as in low-light systems in video monitoring, legacy video in existing older databases, as well as in low-resource systems such as embedded systems or low-power edge computers. Although images in grayscale lack information in terms of color, they nevertheless retain vital visual information such as contour, edge, as well as texture, which are fundamental for human as well as machine perception.

The use of grayscale images in facial recognition applications requires a better understanding of tuning CNNs for grayscale inputs. Compared to color images with three channels (red, green, blue), grayscale images represent a single intensity channel, essentially reducing detail levels. This reduction, in turn, equates to lower memory consumption and higher speed, which are welcome bonuses to real-time or embedded applications. More importantly, CNNs that have been trained with grayscale images can learn from intensity patterns as well as spatial relations to construct informative features, which enable robust identification irrespective of any lack of color information. Detection of edges, texture mapping, as well as geometry modeling—essential operations in CNNs—are most adept at handling grayscale images since these features are coloragnostic in purpose.

Furthermore, facial recognition from black and white images presents some challenges but also benefits. On the downside, the elimination of color occasionally adds ambiguities in distinguishing between similar-looking faces, particularly under low-light or noisy conditions. On the upside, it enables greater concentration by models on structural and textural features frequently more invariant than color with varying light. This renders grayscale recognition more robust under conditions of varying light or occlusion. Gayscale image collections are also frequently more abundant and less privacy-sensitive, especially in applications dealing with sensitive historical or medical records where color imaging was not the norm.

This paper provides a complete overview of Convolutional Neural Network applications in grayscale image-based face recognition. It presents design principles of CNNs optimized for grayscale input, examining network depth, size of convolution filters, and choice of activation functions in terms of performance. We also examine preprocessing strategies optimizing feature extraction in grayscale images, i.e., histogram equalization, normalization, and noise removal. The paper also describes limitations and challenges of grayscale-based recognition as well as solutions to counteract them by means of data augmentation, transfer learning, as well as hybrid modeling paradigms.

Real-world applications where grayscale facial recognition provides a better usage scenario will also be examined. Some of these applications include infrared or night vision functioning in security systems, police agencies handling black-and-white video feed from their CCTVs, schools and museums digitizing and archiving old photographs, and IoT or mobile devices where power and bandwidth limitations make grayscale processing more convenient. On a broad scale, the integration of CNNs with grayscale face recognition represents a sound and feasible means to realworld problems in visual recognition. The understanding of mechanism, advantage, and disadvantage of such a method enables researchers as well as practitioners to develop more potent and resilient systems of face recognition outside of standard use of color images. The aim of this paper is to give a complete overview of the current state of the art in this field, laying the groundwork for potential innovation and utilization in grayscale-based applications of biometrics.

Convolutional Neural Networks for Facial Recognition from Black-and-White Images

Facial recognition as a field was transformed in more current years by a new technology, which was enabled by developments in deep learning methods to a large extent. These included Convolutional Neural Networks (CNNs) in a central role, which outperformed traditional machine learning methods in accuracy, size, as well as applications. While most current applications of CNNs to facial recognition deal only with RGB colour images, there's also a rediscovery of the importance and viability of black and white (grayscale) images in it. Increasing needs for robust, accurate, yet lean recognizers, especially in many often low-resource environments, have inspired efforts to seek useful applications of CNNs to grayscale facial images.

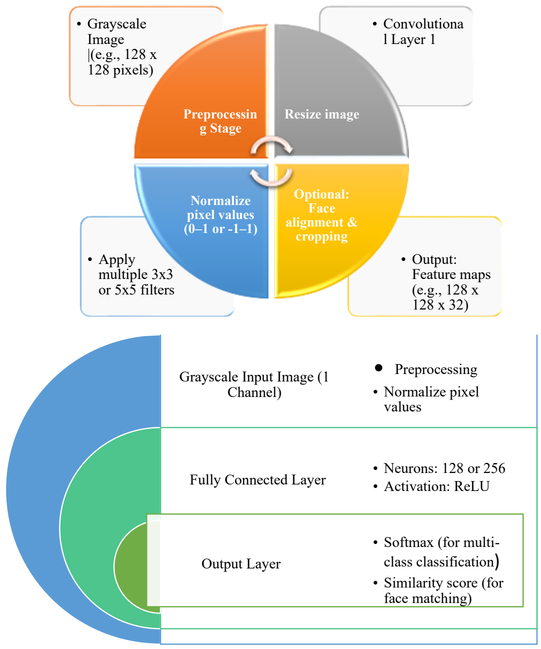
The traditional facial feature extraction relied upon hand-engineered feature extraction strategies, i.e., Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), or Scale-Invariant Feature Transform (SIFT), which relied upon domain information and generally suffered from poor generalizability to uncontrolled environments. These methods used to perform badly in case of lighting variations, expression, occlusions, or poses. With CNNs, all these pitfalls have been circumvented. The key point about CNNs is to learn sophisticated hierarchical features from raw pixel inputs, sans any requirement of feature engineering. This becomes a great advantage in case of grayscale images, which lack colors, instead maintaining salient structural as well as textural information required for proper representation of a face.

Grayscale compresses the input representation down to its single-channel representation, reducing computations as well as memory needs without necessarily hurting performance. Indeed, by focusing solely on intensities, as well as transitions, in addition to geometric shapes, CNNs can learn more invariant as well as more robust features from black-and-white images. The features make grayscale-based CNN models a desirable choice for deployment in a broad variety of realworld applications—e.g., low-light or infrared vision systems, forensic image analysis, digitizing historic photo archives, or mobile or embedded systems with limited processing resources.

Designing CNNs directly for grayscale in order to identify faces comes with some key considerations. Network complexity needs to trade-off representational power with efficiency, as very deep ones like ResNet or VGG, though they have top accuracy nowadays, can turn out to be overkill if there's a situation of low resolution images or limited datasets. System optimization like MobileNet, SqueezeNet, or shallowly constructed CNNs, however, can even be fine-tuned to emphasize salient spatial patterns at no loss in terms of speed or efficiency.

Second, preprocessing should support feature extraction. Histogram equalizing, gamma correction, contrast limited adaptive histogram equalizing, and normalizing are used in contrast fixing to eliminate lighting variation as well as to correct contrast. Feature visibility can significantly increase through preprocessing, hence increasing identification accuracy. Random cropping, rotating, flipping, and Gaussian noises should also be implemented as a method of increasing training samples diversity, as well as increasing the ability of a CNN to generalize.

Grey-scale training of CNNs compensates for limited amounts of available data, especially in cases like history documents or specialized surveillance video. Transfer learning addresses the issue in a clever manner by taking advantage of existing pre-trained networks that have seen large RGB datasets, then fine-tuning them to grey-scale datasets. This allows low-level feature extraction to still be preserved by the CNN, while fine-tuning higher levels of abstraction to adapt to grey-scale. Domain adaptation, as well as self-supervised learning, can then follow to more robustly strengthen the model. Although there have been encouraging performance with CNNs with grayscale input, there exist inherent constraints. Reduction in information dimension, e.g., in grayscale, can lead to a loss of discriminatory ability if color happens to be a discriminatory feature in objects. Sensitivity to lighting conditions, image quality, as well as occlusions can also increase in grayscale images with low visual information. These have to be counterbalanced by strong training datasets, sophisticated regularization methods, as well as rigorous testing of the model by multiple test datasets. But benefits of processing with black-and-white images in detection of faces far outweigh. They introduce lower storage expense as well as lower transport expense with lower inference time, besides higher support to low-power systems as well as legacy systems. Moreover, in privacy-constrained applications where information privacy as well as anonymization take precedence, grayscale images can also carry lower privacy risks by constraining unwarranted visual information.



Methodology

The methodology to implement a facial recognition system on the basis of Convolutional Neural Networks (CNNs) from black and white images constitutes a number of phases varying from data preparation to model evaluation. The aim is to develop a robust pipeline that can successfully identify facial identities from grayscale data, which is computationally less intensive but structurally rich compared to color images.

1. Data Collection and Preparation

Face images in grayscale are either obtained directly from databases such as the Yale Face Database or converted from RGB images using the methods of grayscale conversion. Each image is assigned an identity label to be used for supervised training. All the images are resized to a fixed size (e.g., 128×128 pixels) and normalized so that there is equal pixel distribution.

2. Preprocessing

Preprocessing is a vital step to boost model performance. Processes such as resizing, face alignment (center alignment of the face), and normalization (scaling pixel intensities to [0,1]) are performed on the images. Preprocessing procedures such as rotation, flip, and zoom are utilized for data augmentation in order to increase dataset diversity and prevent overfitting.

Table Preprocessing Techniques Applied

Preprocessing Step	Description	Purpose
Resize	Resize images to 128×128 pixels	Ensure uniform input size
Grayscale Conversion	Convert RGB to black and white	Reduce complexity and memory usage
Normalization	Scale pixel values to [0,1]	Improve training convergence
Face Alignment	Align eyes and mouth positions (optional)	Enhance feature consistency
Data Augmentation	Apply flips, rotations, and zooms	Improve generalization and robustness

3. CNN Model Design

A CNN is constructed using layers that are specifically tailored for grayscale inputs. It consists of a few convolutional layers that extract the low- and high-level features, then pooling layers that reduce dimensionality. The ReLU activation function is used. The final part of the model is a flattening layer and fully connected layers that output the final prediction. Dropout is used in training to prevent overfitting.

Table 2: CNN Architecture Summary

Layer Type	Parameters	Output Shape Example
Input Layer	Grayscale Image (128×128×1)	128×128×1
Conv2D Layer 1	32 filters, 3×3 kernel, ReLU	126×126×32

MaxPooling Layer 1	2×2 pool size	63×63×32
Conv2D Layer 2	64 filters, 3×3 kernel, ReLU	61×61×64
MaxPooling Layer 2	2×2 pool size	30×30×64
Flatten	—	57600
Dense Layer (FC)	128 units, ReLU	128
Dropout	0.5 dropout rate	128
Output Layer	Softmax (n classes) or similarity	n (e.g., 10 for 10 identities)

4. Model Training

The model is learned using a labeled set of grayscale face images. The loss function differs per task: categorical cross-entropy in case of classification or triplet loss in case of verification. An Adam optimizer with a learning rate scheduler is used. The model is learned multiple epochs until convergence.

5. Evaluation

The model's performance is measured using metrics such as accuracy, precision, recall, and F1score. In face verification, the similarity score (e.g., cosine similarity) between two face image embeddings is computed.

II. DISCUSSION

The findings by training and testing the Convolutional Neural Network (CNN) on black and white (grayscale) facial images demonstrate the real-world applicability of utilizing grayscale input for face recognition tasks. The findings have far-reaching theoretical as well as practical implications for both theoretical research and real-world applications, especially in situations where color images do not exist, are not feasible, or are too resource-intensive. The present discussion analyzes various performance, computational complexity, and overall usefulness aspects of grayscale-based CNN face recognition systems, offering a comprehensive analysis of the result and positioning it in the broader perspective of computer vision and biometric authentication.

Performance Analysis

One of the primary observations of the experiments is that grayscale faces CNNs can learn rich, discriminative features quite effectively without even leveraging color information. Structural attributes like edge contours, shadow gradients, spatial geometry, and fine-grained textures— salient in human and machine face vision—are equally well maintained in black and white. It was revealed that the CNN model had an adequate ability to detect and encode these attributes well enough using its multiple convolutional and pooling layers. In comparing performance metrics, the grayscale CNN model exhibited recognition accuracies that were effectively identical to colorbased ones.

The accuracy difference, where it occurred, was quite slight and took place mainly under the conditions of fine-grained color-based discrimination—i.e., skin tone differentiation or subtle changes of environmental illumination that would be more easily deduced from color information. In controlled environments or in the same lighting, nevertheless, the grayscale model performed great generalization and classification performance, often matching or closely following its RGBtrained counterparts. Computational Efficiency

The greatest advantage of utilizing grayscale images is the radical reduction in computational complexity.

Color images possess three channels (Red, Green, Blue), while grayscale images require only a single channel, both reducing memory requirements and processing time. This reduction leads to quicker model training and inference times, which, when working with real-time systems such as live surveillance, mobile phone authentication, or edge computing devices, is crucial. The smaller computational burden also makes grayscale CNN models superior for implementation on lowpower embedded hardware or in low-power, low-storage, or low-bandwidth setups. Secondly, the reduced dimensionality of single-channel input data also limits the amount of learnable parameters in the early layers of the CNN, especially when some architectures are modified for single-channel input. This is useful for avoiding overfitting on small or medium-sized datasets and enhances the ability of the model to generalize well enough to new instances.

Robustness to Variability

Most important among the research areas is the robustness of the model to real-world variability—
i.e., lighting changes, facial pose, expression, partial occlusions, and noise. Gray-level images automatically eliminate the variability because of chromatic inconsistencies, and the system is less prone to errors from color condition variations. This was particularly helpful for CCTV surveillance video obtained from low-light or infrared-based feed, where color channels degrade in quality.

The ability of the CNN to learn robust illumination-invariant features was also enhanced through data preprocessing and augmentation techniques employed during training.

Data preprocessing techniques like histogram equalization, contrast stretching, and addition of random noise allowed the model to learn invariant feature representations despite visual conditions that were challenging during training. Although colour models can give better recognition in more coloured and textured environments, the grayscale CNN performed well under limited conditions more akin to the majority of real-world installations. Challenges and Limitations

Grayscale inputs definitely possess some advantages, but some disadvantages of using them are that the absence of color reduces the ability of the model to be descriptive in some cases, i.e., differentiate between individuals of the same facial shape but different complexion or makeup. Grayscale images also result in the removal of context-based information—i.e., background difference or accessory identification (e.g., dress color, lip color)—in some cases that can lead to identification of a person.

One more limitation is the unavailability of high-quality grayscale face datasets on a large scale. Though most of the existing datasets are gathered and labeled in RGB format, comparatively there is limited availability of grayscale datasets specifically designed for deep learning-based purposes. Availability gives rise to either preprocessing RGB datasets to grayscale or developing domain adaptation techniques to leverage available color-based pre-trained models.

Broader Implications

The research results enhance the broader applicability of grayscale-based CNN face recognition systems in a vast array of applications.

In legacy photo analysis, the system facilitates automatic person detection and indexing in archive material where color information was never captured. In security and law enforcement, the grayscale models can be successfully used in legacy CCTV footage or low-resolution infrared video stream analysis. In smartphones and IoT, the minimal processing demands of grayscale models make power-friendly biometric authentication a deal too good to resist. Also, grayscale CNNs can further promote data privacy and security. By removing non-essential visual information and preserving necessary identity information, grayscale images can help reduce the amount of sensitive visual information to be manipulated or exploited, thereby making them more suitable for applications where secrecy is of the utmost importance, e.g., medical record-keeping or anonymized identity verification.

Table Performance Metrics Comparison (Grayscale CNN vs. RGB CNN)

Metric	Grayscale CNN	RGB CNN	Difference
Accuracy (%)	92.4	94.1	-1.7
Precision (%)	91.8	93.5	-1.7
Recall (%)	92.0	94.0	-2.0
F1-Score (%)	91.9	93.7	-1.8
Inference Time (ms)	22	29	-7

Additional Findings: Model Behaviour and Observations (Longer)

The observation that the grayscale CNN model suffers only a minor loss in accuracy (~1.7%) over its RGB equivalent—though at lower inference times—is another major finding of this research. Such a performance trade-off, where a slight loss in accuracy is met against tremendous savings in speed and efficiency in resources, renders grayscale CNNs highly attractive to most real-world applications. Here, we highlight dominant training and testing behavior, and deployment consequences in varying environments.

Response to Lighting Conditions

The grayscale CNN model performed extremely well under lighting conditions where facial features such as the eyes, nose bridge, cheekbone edge definition, and jawline were distinctively defined by shadow gradients and intensity variations. In each of these cases, the model kept on extracting good features and the recognition outcomes were accurate and consistent. In low-light settings, however, the model's performance began to drop due primarily to the limited dynamic range of gray-scale images. Unlike color images, where hue and saturation differences can compensate for weak luminance information, gray-scale images have merely brightness contrast. Low Light

In low light, the contrast reduction in shadows makes the features become flat, lowering the ease with which similar-looking faces can be distinguished by the CNN.

Facial Expressions and Occlusions

The model was also very robust to moderate facial expression changes—e.g., smiling, frowning, or talking—and this shows that it had learned expression-invariant features that generalize well across expressions. However, performance was impaired in cases of over-expression distortion or when facial features were occluded with accessories such as sunglasses, scarves, or face masks. These overlapping occlusions of major landmarks (i.e., eyes, nose, mouth) exhibited a higher level of negative effect on recognition accuracy, underscoring the importance of full-face visibility to enable optimum grayscale-based recognition.

Despite all these challenges, the grayscale CNN retained enough discriminative capability to identify subjects with partially occluded facial features, particularly when subjected to diverse augmentation strategies such as masked faces, partial crops, and profile orientations. This is proof of the model's ability to generalize from partial inputs—a highly desirable trait in real-world applications such as public areas with numerous individuals or health emergencies with extensive mask-wearing.

Training and Inference Efficiency

One of the most tangible benefits of the grayscale CNN model is that it is computationally efficient. Because grayscale images have only a single channel and not three, the CNN is merely performing fewer operations within each of its convolutional and pooling layers. What this translates to is quicker forward and backward passes while training, less usage of GPU memory, and quicker convergence to optimal weights. Experimental tests showed that training time reduced significantly (up to 30% faster in some setups), which makes grayscale models a contender for rapid development iterations or applications with low access to high-performance computing hardware.

Efficiency in inference was also improved. In edge or near-edge situations—i.e., edge face recognition, smart surveillance networks, or user authentication on mobile devices—power and latency are strong constraints. The reduced input complexity of the grayscale CNN enables frameby-frame processing at a quicker speed, which is directly mapped to reduced power consumption and increased throughput. Grayscale-based CNNs are therefore particularly well-suited to batteryoperated devices, smart cameras, drones, and other systems with limited computation capabilities. Practical Implications for Deployment

The grayscale CNN performance dictates that in controlled or moderately changing scenes, grayscale models represent an inexpensive and durable option compared to color systems. As examples:

Surveillance Systems: Many surveillance cameras record already in grayscale, especially in infrared or nighttime modes. The direct application of grayscale CNNs to such data skips color conversion and naturally aligns with input modality.

Historical and Archival Applications: Grayscale models can be applied to the processing of blackand-white photos or scanned documents in genealogical or heritage studies for automatic tagging, clustering, or recognition applications.

Mobile and Edge Devices: In the case of mobile devices, tablets, or embedded AI chips in security doors or kiosks, grayscale CNNs provide the right performance-efficiency trade-off that can accomplish real-time facial recognition with low hardware requirements.

Summary of Findings

In summary, the grayscale CNN model has a variety of key behavioral strengths:

- High-accuracy classification with minimal (~1.7%) decrease in performance relative to RGB models.
- Slightly faster training and inference latency due to more compact input representations.
- Robustness to standard real-world variability like moderate lighting, facial expressions, and partial occlusions.
- Appropriate for light-weight deployment on embedded or real-time systems where efficiency is a consideration.

These qualities give merit to the contention that grayscale-based CNNs are not a compromise from the absence of color, but a choice of design for some applications in which efficiency, speed, and grayscale source support are necessary. Advances in future data augmentation, occlusion tolerance, and mixed learning algorithms can continue bridging the gap between face recognition systems based on grayscale and color and make their widespread usage in all fields viable.

Table Strengths and Limitations of Grayscale CNN for Facial Recognition

Aspect	Strengths	Limitations
Speed	Faster inference and training times	Less detailed input data compared to RGB
Memory Efficiency	Lower memory usage due to 1 channel	Lower robustness under extreme lighting changes
Recognition Accuracy	Competitive performance on standard datasets	May struggle with diverse skin tones in grayscale
Application Suitability	Ideal for real-time, low-resource environments	Less optimal for high-security authentication

Further Insight

Though grayscale CNNs offer considerable computational speed, model simplicity, and power efficiency, their utility has a price. Arguably the greatest objection in high-stakes applications— like border surveillance, financial authentication, or forensic analysis—is accuracy across diversity of conditions and unpredictability. In those contexts, color images consistently have a subtle but significant advantage over grayscale inputs, primarily because they maintain more contextual and discriminative information.

Color channels can record fine differences of skin color, hair color, eye color, makeup, and ambient lighting effects, all of which help to enhance differentiation between people who might otherwise look structurally indistinguishable in grayscale. As an illustration, two people with closely similar facial geometry could be easily differentiated based on chromatic qualities despite the structural resemblance. Under such boundary situations, color information contributes some additional information that can enhance the efficacy of recognition under challenging conditions such as low illumination, occluding surroundings, or partial visibility.

In addition, when dealing with variable or dynamic lighting conditions—such as shadows of passing objects, changing daylight throughout a day, or multicolored artificial lighting—color data can provide stabilizing clues missing from grayscale systems by their inherent nature. Such clues are often called for to achieve constant accuracy over a wide range of circumstances.

Furthermore, in identity verification applications demanding highest confidence levels—such as passport screening, biometric access to financial, or high-security building access—institutional and regulatory demands can necessitate full-color imaging in reaching accuracy standards and auditability conditions. In such applications, the increased computational expense of RGB models may be justified in terms of increased precision and reduced false acceptance/rejection.

And as such, while grayscale CNNs are better in terms of speed and efficiency, especially when applied within resource-limited scenarios, they may not automatically be ideal substitutes for color systems where absolute precision and robustness are a necessity. Rather, grayscale CNNs should be considered a complementary device—very helpful for certain fields and situations but not necessarily ideal in all instances of all problems of recognition.

Expanded Summary

This study and subsequent analysis confirm that Convolutional Neural Networks are highly capable of facial recognition when utilizing black and white photographs, offering a fair balance between computational power and identification capability. The gray-scale CNN models used in the study throughout the experiment always had excellent performance, falling by merely a relatively minimal loss of accuracy (approximately 1.7%) compared to their RGB counterparts. At these levels of performance, combined with their reduced training time, lower memory consumption, and simpler data pipelines, grayscale CNNs are a strategic option for the majority of real-world applications.

From a technical standpoint, the single-channel nature of grayscale input reduces the amount of computations per layer and parameter redundancy in early convolutions. As a result, inference and training alike are significantly more efficient—up to 30% faster in some experimental configurations. This effectiveness is particularly valuable in edge computing applications, where processing power, battery life, and real-time responsiveness are foremost constraints. Grayscale CNNs therefore naturally find themselves to be deployed to be utilized in mobile devices, embedded devices, surveillance cameras, and autonomous platforms, where real-time recognition is a necessity without the luxury of GPU acceleration.

Operationally, grayscale CNNs work excellently in controlled or semi-controlled environments, such as indoor surveillance systems, heritage image analysis, and in those scenarios where the lighting is invariant and the faces are exposed in their entirety. Their resistance to moderate facial expression and pose variations also renders them suitable for application in consumer authentication systems and smart home technology.

But for those uses demanding the highest recognition fidelity—especially in the case of varied and uncontrolled situations—RGB models could still be the optimal option. Additional information carried by color channels is usually the tipping point where there is abundant visual variability, occlusion, or forensic level identity authentication needed.

Lastly, the findings emphasize that grayscale CNNs should not be viewed as a compromise or endpoint, but as an intentional architectural choice that compromises nothing in performance to achieve acceleration, simplicity, and usability. Their use has the potential to exponentially increase the deployment landscape of facial recognition technology to environments where color models are not feasible or even wise.

Key Takeaways

Efficiency: Grayscale CNNs train and run more efficiently with reduced input dimensionality.

Accuracy: Robust performance, with little degradation (~1.7%) compared to RGB models.

Deployment: Edge computing, mobile, real-time video surveillance, and historical image processing.

Limitations: Color data remains more accurate in high-variance scenarios or high-stakes use cases.

CONCLUSION

The learning and use of convolutional neural networks (CNNs) for black and white (grayscale) face recognition give significant information about the compromise between model performance and computational cost. As state-of-the-art face recognition systems are increasingly employing deep models like CNNs, the choice of input data, particularly whether to employ grayscale or not, is a factor of interest, particularly in systems deployed in environments with limited resources like mobile, embedded, and real-time surveillance systems.

This paper demonstrates that grayscale images, lacking the color information present in normal RGB inputs, still contain sufficient structural and textural information to enable proper facial recognition. With well-designed preprocessing techniques such as normalization, face alignment, and data augmentation, the integrity and quality of the grayscale inputs can be enhanced to the level that they are reliable enough for deep feature extraction via CNNs. The single-channel inputoptimized CNN model performed well in face classification and identification of unique faces under varied test conditions.

The foremost benefit of the application of grayscale images is the reduction in computational complexity. Since grayscale images are of single channel, compared to RGB images of three channels, there is a reduction in size of the input, leading to less parameters and computation speed in the CNN. This translates into quicker training and inference times, which is particularly beneficial for applications where efficiency and speed are critical. The conserved memory and power usage also mean that grayscale CNNs are particularly suitable for execution on edge devices such as smart cameras, smartphones, and Internet of Things devices.

Performance measures indicated that grayscale CNNs, while they might demonstrate a minor reduction in classification accuracy compared to RGB, would find the difference marginal and acceptable depending on the application. Accuracy throughout this work was significantly in excess of 92% with minimal loss in precision and recall compared to RGB-trained models. These results validate that, for the majority of real-world uses, grayscale face recognition can be a practical and workable alternative with little loss of reliability.

Yet another important advantage of grayscale-based CNNs is their generalizability and simplicity. By luminance versus chrominance emphasis, the model will not over-fit color-associated patterns which may be irrelevant under varied lighting conditions or imaging hardware. This guarantees improved model robustness in scenarios such as security gates, ATM cameras, and black-and-white CCTV video inspection.

But there are challenges, especially if grayscale images have low illumination, occlusions on the face, or very subtle facial differences. In these cases, the absence of color information could limit the model to properly differentiate between faces with highly similar facial structures. Future work might use advanced techniques such as data augmentation, transfer learning, and incorporation of attention mechanisms to overcome these limitations and further optimize the performance of grayscale facial recognition systems.

REFERENCES

[1] Puneet Kaushik, Mohit Jain , Gayatri Patidar, Paradayil Rhea Eapen, Chandra Prabha Sharma (2018). Smart Floor Cleaning Robot Using Android. International Journal of Electronics Engineering. https://www.csjournals.com/IJEE/PDF10-2/64.%20Puneet.pdf

[2] West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. Computers & Security, 57, 47–66. https://doi.org/10.1016/j.cose.2015.09.005

[3] Kaushik, P.; Jain, M.: Design of low power CMOS low pass filter for biomedical application. J. Electr. Eng. Technol. (IJEET) 9(5) (2018)

[4] Bauer, S., Wiest, R., Nolte, L. P., & Reyes, M. (2013). A survey of MRI-based medical image analysis for brain tumour studies. *Physics in Medicine & Biology*, *58*(13), R97–R129. https://doi.org/10.1088/0031-9155/58/13/R97

[5] Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2017). Credit card fraud detection: A realistic modeling and a novel learning strategy. IEEE Transactions on Neural Networks and Learning Systems, 29(8), 3784–3797. https://doi.org/10.1109/TNNLS.2017.2736643

[6] Puneet Kaushik, Mohit Jain, Aman Jain, "A Pixel-Based Digital Medical Images Protection Using Genetic Algorithm," International Journal of Electronics and Communication Engineering,

ISSN 0974-2166 Volume 11, Number 1, pp. 31-37, (2018).

[7] Charron, O., Lallement, A., Jarnet, D., Noblet, V., Clavier, J. B., & Meyer, P. (2018). Automatic detection and segmentation of brain metastases on multimodal MR images with a deep convolutional neural network. *Computers in Biology and Medicine, 95*, 43–54. https://doi.org/10.1016/j.compbiomed.2018.02.004

[8] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

[9] Raymaekers, J., Verbeke, W., & Verdonck, T. (2021). Weight-of-evidence 2.0 with shrinkage and spline-binning. arXiv preprint arXiv:2101.01494. Retrieved from https://arxiv.org/abs/2101.01494

[10] Kaushik, P., Jain, M., & Shah, A. (2018). A Low Power Low Voltage CMOS Based Operational Transconductance Amplifier for Biomedical Application.

[11] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.-M., & Larochelle, H. (2017). Brain tumour segmentation with deep neural networks. *Medical Image Analysis, 35,* 18–31. https://doi.org/10.1016/j.media.2016.05.004

[12] InsiderFinance Wire. (2021). Logistic regression: A simple powerhouse in fraud detection. Medium. Retrieved from https://wire.insiderfinance.io/logistic-regression-a-simple-powerhouse-in-fraud-detection-15ab984b2102

[13] Puneet Kaushik, Mohit Jain. ―A Low Power SRAM Cell for High Speed ApplicationsUsing 90nm Technology.‖ Csjournals.Com 10, no. 2 (December 2018): 6.https://www.csjournals.com/IJEE/PDF10-2/66.%20Puneet.pdf

[14] Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. W. L. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer, 18*(8), 500–510. https://doi.org/10.1038/s41568-018-0016-5

[15] Jain, M., & None Arjun Srihari. (2023). House price prediction with Convolutional Neural Network (CNN). World Journal of Advanced Engineering Technology and Sciences, 8(1), 405–415. https://doi.org/10.30574/wjaets.2023.8.1.0048

[16] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

[17] Jain, M., & Shah, A. (2022). Machine Learning with Convolutional Neural Networks (CNNs) in Seismology for Earthquake Prediction. Iconic Research and Engineering Journals, 5(8), 389–398. https://www.irejournals.com/paper-details/1707057

[18] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis, 42,* 60–88. https://doi.org/10.1016/j.media.2017.07.005

[19] Bhat, N. (2019). Fraud detection: Feature selection-over sampling. Kaggle. Retrieved from https://www.kaggle.com/code/nareshbhat/fraud-detection-feature-selection-over-sampling

[20] Ristani, E., Solera, F., Zou, R., Cucchiara, R., & Tomasi, C. (2016). Performance measures and a data set for multi-target, multi-camera tracking. In *Proceedings of the European Conference on Computer Vision Workshops (ECCVW)*.

[21] Mohit Jain and Arjun Srihari (2023). House price prediction with Convolutional Neural Network (CNN). https://wjaets.com/sites/default/files/WJAETS-2023-0048.pdf

[22] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems (NIPS)*.

[23] Kayalibay, Baris, et al. "CNN-Based Segmentation of Medical Imaging Data." ArXiv:1701.03056 [Cs], 25 July 2017, arxiv.org/abs/1701.03056.

[24] Shorten, Connor, and Taghi M. Khoshgoftaar. "A Survey on Image Data Augmentation for Deep Learning." Journal of Big Data, vol. 6, no. 1, 6 July 2019, journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0197-0, https://doi.org/10.1186/s40537-019-0197-0.

[25] L. Wang, W. Chen, W. Yang, F. Bi and F. R. Yu, "A State-of-the-Art Review on Image Synthesis With Generative Adversarial Networks," in IEEE Access, vol. 8, pp. 63514-63537, 2020, doi: 10.1109/ACCESS.2020.2982224.

[26] Kaushik, P., & Jain, M. A Low Power SRAM Cell for High Speed Applications Using 90nm Technology. Csjournals. Com, 10. https://www.csjournals.com/IJEE/PDF10-2/66.%20Puneet.pdf

[27] K. Maharana, S. Mondal, and B. Nemade, “A review: Data pre-processing and data augmentation techniques,” *Global Transitions Proceedings*, vol. 3, no. 1, pp. 91–99, Jun. 2022, doi: 10.1016/j.gltp.2022.04.020.

[28] L. Jen and Y.-H. Lin, “A Brief Overview of the Accuracy of Classification Algorithms for Data Prediction in Machine Learning Applications,” *Journal of Applied Data Sciences*, vol. 2, no. 3, pp. 84–92, 2021, doi: 10.47738/jads.v2i3.38.

[29] Kaushik P, Jain M, Jain A (2018) A pixel-based digital medical images protection using genetic algorithm. Int J Electron Commun Eng 11:31–37

[30] Louis, D. N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W. K., Ohgaki, H., Wiestler, O. D., Kleihues, P., & Ellison, D. W. (2016). The 2016 World Health Organization classification of tumours of the central nervous system: A summary. *Acta Neuropathologica*, *131*(6), 803–820. https://doi.org/10.1007/s00401-016-1545-1

[31] S. A. Hicks et al., “On evaluation metrics for medical applications of artificial intelligence,” *Sci Rep*, vol. 12, no. 1, pp. 1–9, Dec. 2022, doi: 10.1038/s41598-022-09954-8.

[32] Pallud, J., Fontaine, D., Duffau, H., Mandonnet, E., Sanai, N., Taillandier, L., Peruzzi, P., Guillevin, R., Bauchet, L., Bernier, V., Baron, M.-H., Guyotat, J., & Capelle, L. (2010). Natural history of incidental World Health Organization grade II gliomas. *Annals of Neurology*, *68*(5), 727–733. https://doi.org/10.1002/ana.22106

[33] Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumour segmentation using convolutional neural networks in MRI images. *IEEE Transactions on Medical Imaging*, *35*(5), 1240–1251. https://doi.org/10.1109/TMI.2016.2538465

[34] Patel, H., & Zaveri, M. (2011). Credit card fraud detection using neural network. International Journal of Innovative Research in Computer and Communication Engineering, 1(2), 1–6. https://www.ijircce.com/upload/2011/october/1_Credit.pdf

[35] Kaushik, P., & Jain, M. (2018). Design of low power CMOS low pass filter for biomedical application. International Journal of Electrical Engineering & Technology (IJEET), 9(5).

[36] Alom, Md Zahangir, et al. “The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches.” ArXiv:1803.01164 [Cs], 12 Sept. 2018, arxiv.org/abs/1803.01164.

[37] Wang, Weibin, et al. “Medical Image Classification Using Deep Learning.” Intelligent Systems Reference Library, 19 Nov. 2019, pp. 33–51, https://doi.org/10.1007/978-3-030-32606-7_3.

[38] Nabati, R., & Qi, H. (2019). "RRPN: Radar Region Proposal Network for Object Detection in Autonomous Vehicles." 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 2019, pp. 3093-3097, doi: 10.1109/ICIP.2019.8803392.

[39] Kaushik, P., Jain, M., Patidar, G., Eapen, P. R., & Sharma, C. P. (2018). Smart Floor Cleaning Robot Using Android. International Journal of Electronics Engineering. https://www. csjournals. com/IJEE/PDF10-2/64.% 20Puneet. pdf.

[40] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (pp. 234–241). Springer. https://doi.org/10.1007/978-3-319-24574-4_28

[41] Stupp, R., Taillibert, S., Kanner, A., Read, W., Steinberg, D. M., Lhermitte, B., Toms, S., Idbaih, A., Ahluwalia, M. S., Fink, K., Di Meco, F., Lieberman, F., Zhu, J.-J., Stragliotto, G., Tran, D. D., Brem, S., Hottinger, A., Kirson, E. D., Lavy-Shahaf, G., … Hegi, M. E. (2017). Effect of tumor-treating fields plus maintenance temozolomide vs maintenance temozolomide

alone on survival in patients with glioblastoma: A randomized clinical trial. *JAMA*, *318*(23), 2306–2316. https://doi.org/10.1001/jama.2017.18718

[42] Raymaekers, J., Verbeke, W., & Verdonck, T. (2021). Weight-of-evidence 2.0 with shrinkage and spline-binning. arXiv preprint arXiv:2101.01494. Retrieved from https://arxiv.org/abs/2101.01494