COVENANT JOURNAL OF ENGINEERING TECHNOLOGY (CJET), VOL. 8, NO.1, JUNE 2024; DOI: XXX XXXX XXX

Enhancing Automated Face Recognition with Makeup Detection: A CNN-Based Approach

[1]Balogun, O. Monsurat, [2]Jimada-Ojuolape, Bilkisu, & [3]Odeniyi, A. Latifat.

[1]Department of Electrical and Computer Engineering, Kwara State University, Malete, Ilorin, Kwara State, Nigeria.
[2] Department of Electrical and Computer Engineering, Kwara State University, Malete, Ilorin, Kwara State, Nigeria..
[3]Department of Mathematical and Computing Scienced, Koladaisi University, Ibadan, Oyo State, Nigeria.
: Monsurat.balogun@kwasu.edu.ng ; +2348035196151

Received: 20.02.2024 **Revised**: 10.05.2024 **Accepted**: 19.05.2024 **Published**: 20.05.2024

Abstract:
This study delves into the complex issue posed by facial makeup, which has the potential to significantly alter the appearance of individuals, posing a challenge to automated face recognition systems, as well as age and beauty estimation methods. A model solution aimed at automatically detecting makeup in facial images to improve the accuracy of recognition systems was proposed in this work. The approach revolves around utilizing a sophisticated model that harnesses a feature vector encapsulating crucial aspects of facial attributes including shape, texture, and color. Employing an advanced Convolutional Neural Network (CNN) architecture, the model detects the presence of makeup by analyzing key facial landmarks such as eye distance, nose width, eye socket depth, cheekbones, jawline, and chin. Experiments were performed on a dataset consisting of 200 facial images to assess the effectiveness of the proposed method. The model achieved a validation accuracy of 100%, demonstrating its robustness in makeup face detection. Notably, the computational runtime for the validation process was 1 minute and 40 seconds, underscoring the efficiency of the proposed approach. Moreover, an innovative adaptive pre-processing strategy that capitalizes on makeup information to enhance the performance of facial recognition systems was developed. This strategy aims to optimize the recognition process by leveraging insights gained from makeup detection. By integrating this adaptive pre-processing step, further advancements in the accuracy and reliability of facial recognition technology, particularly in scenarios where makeup may confound traditional recognition methods, are envisioned.

Keywords: Accuracy, Appearance, Classification, Facial Makeup, Pre-processing, Recognition.

1. Introduction

Over the past few decades, significant efforts have been dedicated to enhancing the precision of automated face recognition systems [1]. Face recognition refers to the task of identifying a recognized object as a known or unknown face [2]. Challenges such as variations in posture, illumination, and appearance (PIE) have been addressed through advanced algorithms, which aim to enable unconstrained face recognition across diverse applications [3]. Despite these advancements, several factors continue to impede the performance of face recognition systems, including variations in human facial appearance due to factors such as lighting conditions, image noise, scale, pose, aging, plastic surgery, spoofing, and more [4].

It is important to distinguish between face recognition and face detection. While face recognition involves identifying a face "image" as that of a known or unknown individual, face detection focuses on determining the presence of a face in an image using various classification models [5]. The use of makeup as a means of altering facial appearance presents a significant challenge to biometric face recognition systems [6]. Being cost-effective, non-permanent, and socially acceptable in many cultures, makeup applicationan can effectively confound recognition systems [7-8]. Detecting makeup in facial images holds potential benefits for face recognition systems, enhancing recognition accuracy by enabling the application of makeup-specific preprocessing routines and bolstering security by identifying attempts at face spoofing or obfuscation [9]. Moreover, automated age estimation and aesthetic prediction methods can refine their outputs by incorporating knowledge about the presence of makeup [10].

The effectiveness of face recognition systems is notably hindered by challenges such as makeup, aging, plastic surgery, and spoofing, all of which significantly diminish algorithm performance and accuracy [11]. In critical scenarios such as pandemics where nose masks are essential, there is a heightened need for systems that can accurately detect and recognize faces for identification purposes while reducing unnecessary person-to-person contact [12]. Central to the operation of face recognition systems is the task of face detection [13-14], a concept that may initially appear perplexing to newcomers in the field. However, before face recognition can occur, reliable detection and landmarking of faces are imperative. This underscores the importance of effectively addressing segmentation problems, which have garnered increasing attention from researchers in recent years. The subsequent recognition process heavily relies on features extracted from these facial landmarks [15]. Face detection problems generally fall into two categories: detection in images and real-time detection. Therefore, this project endeavors to develop a system capable of detecting and recognizing faces, particularly those adorned with makeup, thereby addressing a crucial aspect of contemporary face recognition challenges.

A Convolutional Neural Network (CNN) is a specialized type of artificial neural network extensively utilized in image recognition and processing tasks [16]. Specifically tailored to handle pixel data, CNNs are renowned for their robustness, power, and speed [17]. This study employed a CNN (specifically, the Google network architecture) to train and recognize faces within images. The selection of CNN for this purpose is motivated by several key advantages:

- It provides a simple, non-permanent, and cost-effective means of confounding recognition systems.
- Compared to alternative facial recognition methods, CNN-based approaches offer reliability, affordability, and efficiency.
- Through unsupervised learning, CNNs autonomously identify significant features without requiring human intervention.

2. LITERATURE REVIEW

With the aid of advanced image processing and machine learning algorithms, researchers have made significant progress in discerning and analyzing crucial facial features, thereby distinguishing between natural and makeup-altered faces. This advancement facilitates biometric systems in adapting their recognition methodologies, ensuring consistent and dependable identification or verification regardless of facial alterations caused by makeup. This introductory review delves into the burgeoning field of makeup face recognition, seeking to elucidate its fundamental technologies, applications, challenges, and prospects. By unraveling the intricate interaction between AI algorithms and cosmetic enhancements, this review illuminates the transformative potential of makeup face recognition within the beauty industry and broader contexts. [18] introduced a system designed to detect makeup attacks within face-based biometric systems. [12], [19] conducted research investigating challenges in face recognition using machine learning algorithms, with a focus on makeup and occlusions as case studies.

Similarly, [20] evaluated challenges in face recognition employing machine learning algorithms. [21]developed rule-based facial makeup recommendation systems, while [22]created a smart mirror intelligent makeup recommendation and synthesis system. [23] proposed an examples-rules guided deep neural network for makeup recognition, whereas [24]developed a style and latent guided generative adversarial network for desirable makeup transfer and removal. The pioneering work of [25] examined the influence of facial makeup on the accuracy of face recognition systems, revealing a significant decline in biometric performance with makeup application to either reference or probe images. Subsequent studies, including [26-27] corroborated these findings. Furthermore, [28]demonstrated that heavy makeup notably impairs humans' ability to recognize faces.

Utilizing web-collected databases containing pairs of face images captured before and after applying predominantly light makeup, [29]categorized them as genuine representations. [30] devised a method to automatically categorize potential presentation attacks (PAs) into semantic sub-groups without supervision, termed DTN. Their methodology evaluated using a face database containing Multiple Presentation Attacks (M-PAs), including concealment and impersonation, revealed detection error rates (D-EER) of approximately 50% for concealment attacks and 10% for impersonation attacks. [31] and [32] developed systems that were distinguished by addressing various types of M-PAs within a broader context, highlighting challenges in detecting M-PAs using contemporary software-based face Presentation Attack Detection (PAD) methods. Unlike traditional PAD techniques optimized for specific artifacts resulting from Presentation Attack Instruments (PAIs), M-PAs exhibit distinct characteristics posing difficulties for existing detection methodologies [33-34].

Despite significant progress in makeup face recognition systems, challenges persist, which include the diversity and complexity of makeup styles, variations in lighting conditions, and the necessity for comprehensive datasets. This work aims to overcome these challenges by enhancing the robustness and accuracy of face makeup identification within biometric systems.

3. METHODOLOGY

This study initially devised a methodology to identify facial makeup within diverse and uncontrolled face images. The developed method extracts a comprehensive set of features for any given face image, including shape, colour, and texture. Subsequently, these feature sets are utilized by a Convolutional Neural Network (CNN) to determine the presence or absence of makeup in the input facial images. Experiments were rigorously conducted on a dataset comprising two hundred (200) challenging and uncontrolled images, spanning both female and male subjects. These datasets encompassed a wide array of variations including facial poses, lighting conditions, facial expressions, and image resolutions. Additionally, the output of the makeup detector underwent selective pre-processing before being matched against non-makeup images. The performance of this proposed approach was thoroughly evaluated to assess its recognition accuracy, thereby demonstrating its efficacy in accurately detecting and distinguishing facial makeup within diverse and uncontrolled image datasets. Shown in Figure 1 is the block diagram of the developed facial recognition system. The figure consists of seven blocks: face image input, face detection, face alignment/landmark localizaion, normalization, detection of region of interest, feature extraction and recognition/Classification. These stages collectively form the pipeline of a facial recognition system, with each stage playing a crucial role in achieving accurate and reliable recognition results.

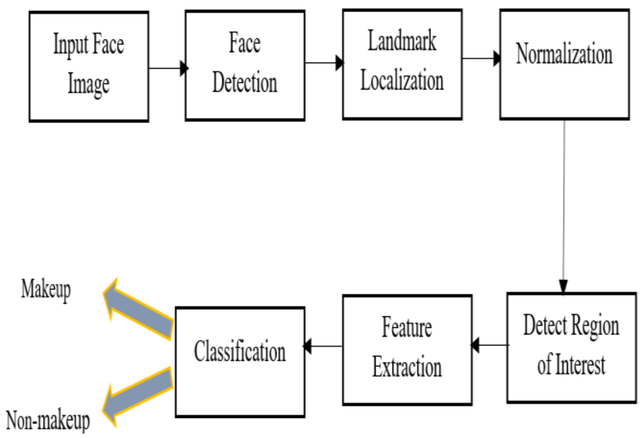

Figure 1: Block Diagram of the Developed Facial Recognition System

Firstly, the Viola-Jones algorithm was utilized for facial detection, yielding suboptimal accuracy due to its reliance on human supervision [35]. Nonetheless, enhancements were made by integrating a layer that facilitated precise facial localization within images, subsequently enhancing the algorithm's accuracy. Real-time face detection involves identifying faces within a sequence of frames captured by a video device [36]. Despite the heightened hardware requirements, real-time face detection is computationally simpler than static image detection from a computer vision perspective. This is attributed to the constant movement of humans, contrasting with the relatively static surrounding environment by walking, fidgeting, blinking, gesturing, and more [37].

Real-time facial recognition employs the vision cascade object detector introduced by the Viola-Jones algorithm. In real-time face detection, the system processes a series of frames to detect faces by employing spatial-temporal filtering, discerning alterations between consecutive frames, and identifying the modified regions for character detection [38]. In this study, precise facial locations were easily identified based on two assumptions:

- The head appears as a small blob above a larger blob (representing the body).
- Head movements exhibit continuity and relatively slow pace, avoiding erratic bouncing.

Consequently, real-time face detection emerges as a comparatively straightforward task, even in unstructured or dynamic environments, thanks to these simple reasoning rules and image processing techniques, thus facilitating the application of convolutional neural networks (CNNs) in this work.

In this study, the CNN algorithm serves as a classifier, and its utilization is demonstrated within the framework of googLeNet. This model employs various strategies, including 1×1 convolutions positioned within the middle of the architecture and global average pooling, as depicted in Figure 2. Additionally, Figure 3 illustrates the Inception module incorporating dimension reductions.

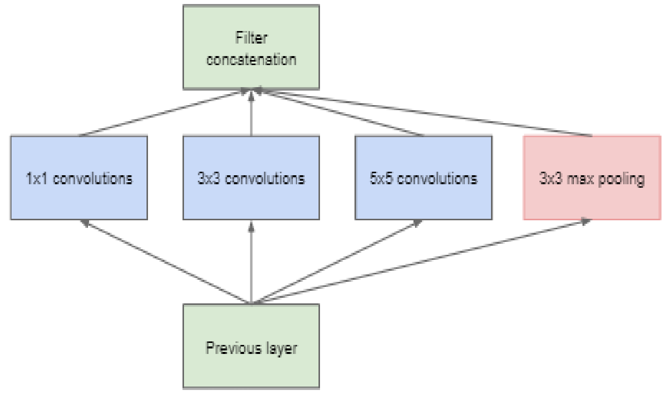

Figure 2 : Inception Module, Naïve Version

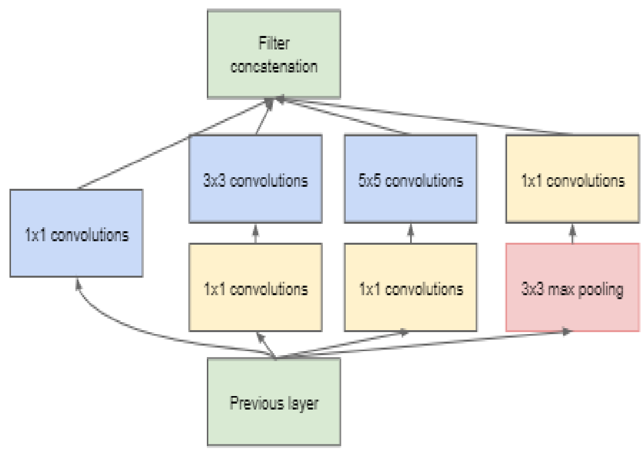

Figure 3: Inception Module with Dimension Reductions

FACE RECOGNITION USING GOOGLENET

The GoogLeNet is an exclusive architecture from previous state-of-the-art architectures. It uses many different contemporary techniques such as 1×1 convolution and global average pooling, allowing it to create a deeper structure as shown in Figure 4. The architecture consists of the following steps:

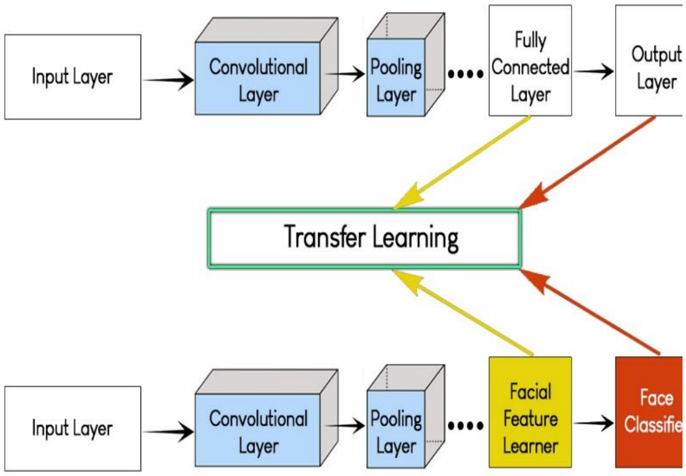

Figure 4: Googlenet Algorithm Block Diagram

1×1 Convolution: Within the inception architecture, 1×1 convolutions are employed to reduce the number of parameters (weights and biases), consequently deepening the architecture. This reduction in parameters enhances the architectural depth.

Global Average Pooling: Traditionally, fully connected layers in various architectures contribute to increased computational costs due to numerous parameters. However, in the GoogLeNet architecture, a method known as global average pooling is implemented at the network's conclusion. This technique averages a 7×7 feature map down to 1×1, effectively reducing the trainable parameters to zero while boosting the top-1 accuracy by 0.6%.

Inception Module: This study adopts a fixed convolution size for each layer. Within the Inception module, parallel operations of 1×1, 3×3, 5×5 convolutions, and 3×3 max pooling are conducted at the input. The outputs of these operations are then stacked together to produce the final output. The rationale behind using convolution filters of various sizes is to enhance the detection of objects at multiple scales.

Auxiliary Classifier for Training: Inception architecture incorporates intermediate classifier branches within the architecture, utilized solely during training. These branches feature a 5×5 average pooling layer with a stride of 3, followed by 1×1 convolutions with 128 filters, two fully connected layers yielding 1024 and 1000 outputs respectively, and a softmax classification layer. The loss generated by these layers is added to the total loss with a weight of 0.3. These auxiliary layers address the gradient vanishing problem and provide regularization. Figure 5 illustrates the detailed layering of the inception architecture as implemented in GoogLeNet.

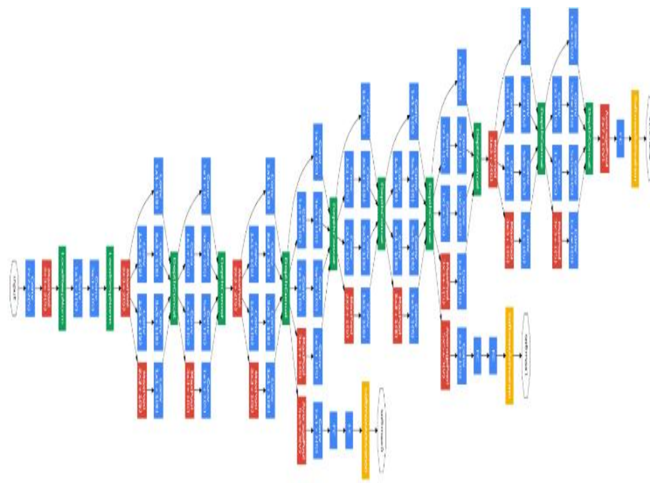

Figure 5: Inception Architecture

The complete architecture comprises 22 layers of significant depth. Engineered with computational efficiency as a priority, it is compatible with devices possessing limited computational resources. Additionally, the architecture integrates two auxiliary classifier layers linked to the outputs of Inception (4a) and Inception (4d) layers, as depicted in Figures 6 and 7.

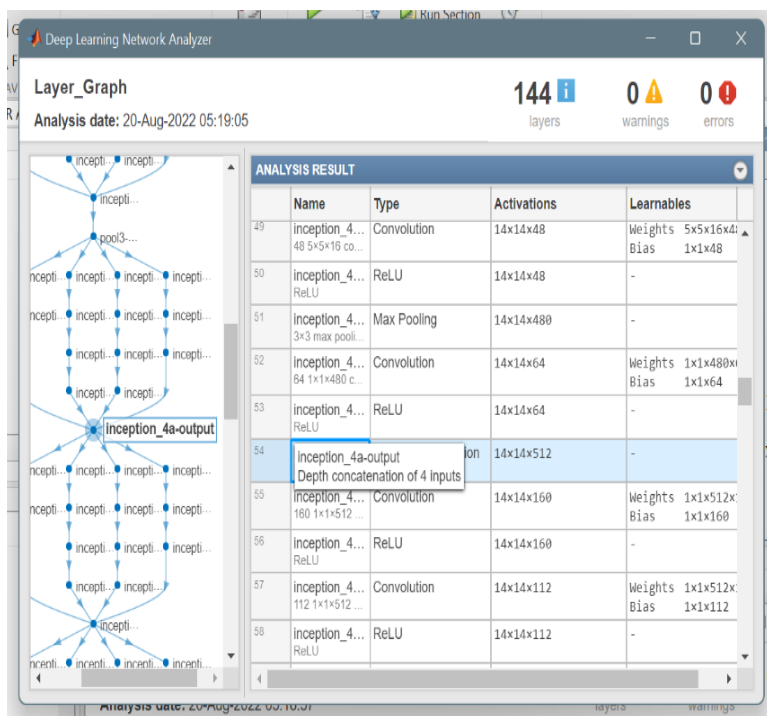

Figure 6: Inception 4a Output layer graph

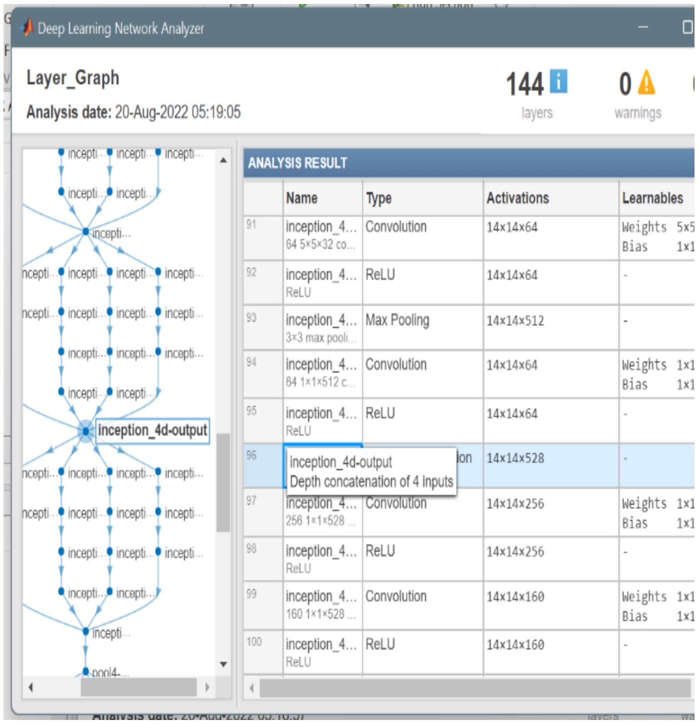

Figure 7: Inception 4d Output layer graph

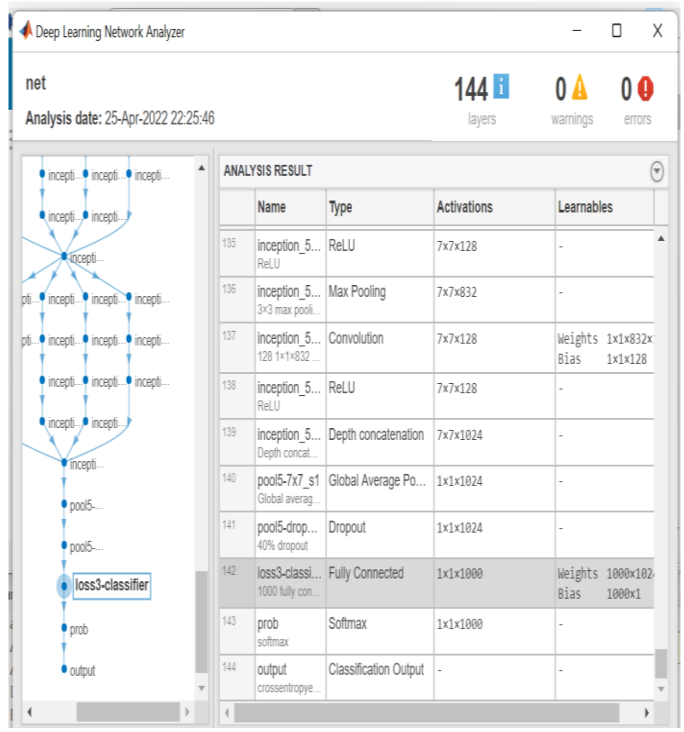

Figure 8: Fully Connected Net Layer with loss classifier

The auxiliary architecture was meticulously crafted, incorporating the following specifications:

- An average pooling layer with a filter size of 5×5 and a stride of 3.
- A dimension-reducing 1×1 convolutional layer with 128 filters, utilizing ReLU activation.
- A fully connected layer producing 1025 outputs, employing ReLU activation.
- Dropout regularization with a dropout ratio set to 0.7 to mitigate overfitting.
- A softmax classifier generating a 1000-class output, akin to the primary softmax classifier.

These components were intricately designed to enhance the performance and robustness of the auxiliary classifier within the architecture.

This architecture is tailored to process images of dimensions 224 x 224 with RGB colour channels. Throughout the architecture, Rectified Linear Units (ReLU) serve as the activation functions for all convolutions, enhancing computational efficiency and feature extraction. At the onset, the image input layer stands as the first element within the Layers property of the network. Specifically designed for the GoogLeNet network, this layer necessitates input images sized 224-by-224-by-3, where 3 signifies the number of colour channels. The final stages of this architecture encompass the extraction of features crucial for image classification. The last learnable layer and the ultimate classification layer, denoted as 'loss3-classifier' and 'output' respectively, are derived from the convolutional layers of the network. These layers encapsulate information regarding the amalgamation of extracted features into class probabilities, loss values, and predicted labels.

To adapt the pre-trained network for classifying new images, these layers are replaced with new counterparts tailored to the new dataset, and the layer graph is extracted from the trained network using transfer learning techniques, as illustrated in Figure 8. Regarding the detection process, when integrated into a video surveillance system, the recognition system meticulously scans the field of view captured by a video camera for facial features. Upon detecting a face within the view, the system swiftly identifies it within a fraction of a second. To optimize efficiency, a multi-scale algorithm is employed to search for faces in low resolution initially. Subsequently, the system transitions to a high-resolution search only after discerning a head-like shape, effectively minimizing computational overhead while maintaining accuracy.

The process of facial identification comprises several intricate stages, each crucial for accurately recognizing and categorizing faces:

- **Alignment:** Upon detecting a face within an image or video frame, the system meticulously analyzes the head's position, size, and pose. This stage ensures that the face is correctly oriented towards the camera, requiring a minimum turn of at least 35 degrees for proper registration.
- **Normalization:** Following alignment, the captured image of the head undergoes scaling and rotation processes to standardize its size and orientation. This normalization step is essential for mapping the facial features into a consistent size and pose, irrespective of the head's location or distance from the camera. Notably, variations in lighting conditions do not impact the normalization process.
- **Representation:** The facial data is translated into a unique code or representation once normalized. This coding process facilitates easier comparison of the newly acquired facial data with previously stored facial data. Each face is thus encoded into a format suitable for computational analysis and comparison.
- **Matching:** In the matching stage, the newly acquired facial data is compared to the stored data, with the aim of linking it to at least one stored facial representation. This comparison process results in the creation of a "faceprint," a distinctive numerical code specific to each face. Once a faceprint is generated, the system can efficiently compare it against the thousands or even millions of faceprints stored in its database. Each faceprint is stored as an 84-byte file, ensuring efficient storage and retrieval.

Additionally, the network undergoes training to adapt to images of various sizes present in the image dataset. This adaptation involves resizing images to meet the required size (required input of 224-by-224-by-3 pixels). To achieve this, an augmented image data store is utilized, automatically resizing training images while also performing additional augmentation operations. These operations may include random flipping along the vertical axis, random translations of up to 30 pixels, and horizontal and vertical scaling of up to 10%. These augmentation techniques are crucial for preventing the model from overfitting, ensuring robust performance, and preventing memorization of exact training image details, as depicted in Figure 9. This comprehensive approach enhances the system's ability to effectively process and identify faces across diverse conditions and datasets.

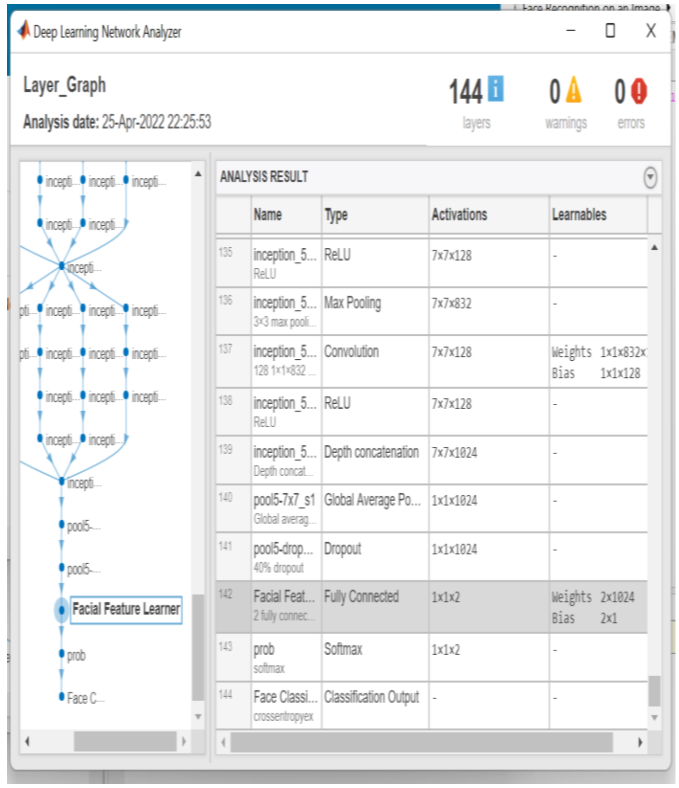

Figure 9: Fully Connected Net Layer with facial feature learner

The implementation of fully automated facial makeup detection and recognition, specifically focusing on faces in frontal view, is achieved through the utilization of GoogLeNet as the Convolutional Neural Network (CNN) algorithm. This choice is made due to the CNN's capability to leverage image invariants inherent to human faces, enabling robust and accurate facial analysis. Unlike traditional methods, CNNs do not necessitate an extensive training dataset and boast low computational time, making them ideal for real-time applications. Within the system, nodal points critical for facial analysis are meticulously measured. These include the distance between the eyes, the nose's width, the eye socket's depth, the cheekbones' prominence, the jawline's contour, and the chin's structure. These measurements serve as fundamental features for facial makeup detection and recognition.

The system operates within a controlled environment, leveraging MATLAB R2020b 64-bit version as the primary platform. The operating system utilized is Windows 11 Home Single Language, 64-bit edition. Hardware specifications include 8.00 GB DDR4 RAM, coupled with an 11th Gen Intel(R) Core (TM) i5-1135G7 processor clocked at 2.40GHz with a turbo boost up to 2.42 GHz. Additionally, Intel(R) Iris(R) Xe Graphics are utilized to facilitate graphical processing tasks. This setup ensures a stable and efficient environment for conducting facial makeup detection and recognition tasks. The combination of robust hardware capabilities and sophisticated software algorithms enables accurate and reliable analysis of frontal view faces, contributing to advancements in automated facial recognition technology [39].

4. Results and Discussion

The results and discussion section of the study provides a comprehensive overview of the analysis conducted on a dataset comprising 200 facial images. This dataset encompasses individuals with varying appearances, including both natural facial features and those enhanced with makeup. Notably, the dataset features a diverse representation of both female and male subjects, acknowledging the transformative effects of makeup on facial appearance irrespective of gender. This deliberate inclusivity ensures a thorough evaluation of the capabilities of facial makeup detection and recognition systems across different demographics in line with [40]. Figure 10 serves as a visual representation of the validation accuracy results derived from the retraining of GoogLeNet using the provided image dataset. Remarkably, the retrained GoogLeNet model achieved a validation accuracy rate of 100%, highlighting its efficacy in accurately classifying facial images. This achievement is particularly noteworthy considering the varying appearances and makeup styles present in the dataset[41]. Moreover, the retraining process was completed with impressive efficiency, taking only 1 minute and 40 seconds. This expedited process can be attributed to the utilization of a deep Convolutional Neural Network (CNN) architecture, known for its ability to extract intricate features crucial for accurate classification tasks, as supported by [42].

Furthermore, the incorporation of additional layers within the CNN architecture, exemplified by GoogLeNet's inception module consisting of 22 layers, plays a significant role in enhancing accuracy. These additional layers facilitate the extraction of nuanced features from facial images, thereby improving the model's discriminative capabilities. However, it is important to note that while the accuracy is commendable, further computational resources and time investments could yield even higher accuracy levels. This underscores the inherent trade-off between computational efficiency and model performance in deep learning tasks. Overall, the primary objective of GoogLeNet architecture is to optimize accuracy in facial makeup detection and recognition tasks. The inception of GoogLeNet was driven by the aspiration to develop a deeper CNN architecture capable of effectively capturing complex facial features. The successful validation of the retrained GoogLeNet model underscores its efficacy in achieving this objective, heralding advancements in automated facial analysis technologies and their widespread application across various domains [43].

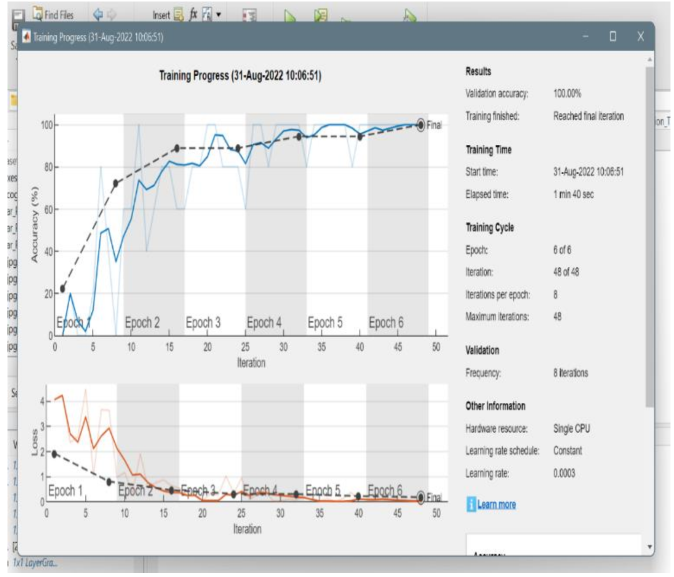

Figure 10: Validation Accuracy for the GoogleNet training progress

The neural networks utilized in this study have undergone extensive training on a diverse array of images, enabling them to learn distinct feature representations suited for various tasks. Specifically, both pretrained networks have been trained with images having an input size of 224-by-224 pixels. However, for the specific task of facial image classification and identification, one of these networks underwent retraining using a technique known as transfer learning in line with [44]. Transfer learning involves repurposing a pretrained neural network for a new task by fine-tuning its parameters. In this case, the network was retrained to excel in facial image classification and identification. This retraining process involves adjusting the network's parameters and architecture to better suit the requirements of the new task. By leveraging transfer learning, the model benefits from the knowledge and feature representations learned during its initial training on a vast dataset, such as ImageNet, leading to higher accuracies and improved performance in the new task [45].

The GoogLeNet-M network was trained using the ImageNet dataset, which serves as a rich source of diverse images spanning numerous object categories to facilitate transfer learning. This dataset provides a solid foundation for transfer learning, allowing the network to effectively generalize its learned features to facial image classification and identification tasks. Furthermore, to evaluate the performance of the retrained network, both training and validation datasets were utilized. The dataset was prepared into different partition has presented in Table 1. It was discovered that the of 70% to 30%, produced the highest recognition accuracy and average accuracy was 95.41%. This partitioning ensures that the model's performance is assessed on both seen and unseen data, providing insights into its generalization capabilities and overall accuracy.

Table 1: Accuracy Table for varying Training Data

TRIAL	TRAINING DATA%	VALIDATION DATA%	ACCURACY (%)
1	50	10	92.86
2	55	22	93.33
3	60	25	93.55
4	70	28	97.32
5	70	30	100
	Average		**95.41**

In summary, this study has leveraged transfer learning to retrain the GoogLeNet-M network for facial image classification and identification. The model achieves enhanced accuracy and performance in the targeted task by fine-tuning the network's parameters and utilizing the knowledge gained from pretraining on the ImageNet dataset. The evaluation of the model's performance on both training and validation datasets further validates its effectiveness in accurately classifying facial images.

5. CONCLUSION

The study introduces and validates an innovative automated makeup detection and recognition system tailored to handle unconstrained facial images, accommodating both makeup-enhanced and natural faces. By harnessing shape, texture, and color features extracted from the entire facial region and specific facial subregions, the detector demonstrates remarkable proficiency in accurately discerning the presence of makeup. Experimental investigations conducted on a dataset comprising 200 unconstrained facial images showcase the system's robust capabilities, achieving makeup detection training rates of up to 100%. This achievement translates into an impressive overall validation accuracy and average classification rates of up to 95.41%. Such high accuracy underscores the effectiveness of the system in accurately distinguishing between makeup-adorned and natural facial appearances, showcasing its potential for real-world applications.

Moreover, the output of the makeup detector is strategically utilized to perform adaptive pre-processing within the context of face recognition. The experimental findings elucidate the efficacy of this pre-processing routine in enhancing the accuracy of face recognition tasks, particularly when validating makeup images against non-makeup counterparts. This adaptive pre-processing approach showcases the system's adaptability and its potential to contribute to improved performance in face recognition systems. Future research endeavors are poised to focus on refining the performance of the makeup detector without the need for extensive pretraining. Additionally, efforts can be directed towards exploring methodologies to mitigate artifacts introduced by makeup or facial alterations, addressing challenges such as spoofing or obfuscation in attendance biodata or biometric systems. A significant area of interest lies in developing techniques to quantify the degree of makeup applied to the face, as this poses a critical challenge with profound implications for various applications. Furthermore, there is potential to extend the capabilities of the presented model to serve as an automatic age estimation and beauty assessment system, considering the influence of makeup on these metrics. Such advancements aim to broaden the applicability of automated makeup detection, offering valuable insights into its utility across diverse domains within the realm of facial recognition and biometrics. Ultimately, this research underscores the transformative potential of automated makeup detection in advancing the efficacy and reliability of facial analysis technologies, paving the way for enhanced performance, and expanded applications in real-world scenarios. By addressing key challenges and exploring new avenues of research, this study contributes to the ongoing evolution of automated makeup detection and its impact on facial analysis technologies.

REFERENCES

[1] A. Mondal and S. Roy, "Design and development of fusion-based expert system for multimodal biometric recognition with facemask authentication," *International Journal of System of Systems Engineering*, vol. 13, no. 2, 2023, doi: 10.1504/IJSSE.2023.131239.

[2] A. A. Mohamad Alshiha, M. W. Al-Neama, and A. R. Qubaa, "Biometric face recognition method using graphics processing unit system," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 30, no. 1, 2023, doi: 10.11591/ijeecs.v30.i1.pp183-191.

[3] W. Ismaila, A. Adetunji, … A. F.-… J. of S. and, and undefined 2012, "A study of features extraction algorithms for human face recognition," *tjournal.org*, Accessed: Aug. 30, 2023. [Online]. Available: http://www.tjournal.org/celosni/july2012.pdf#page=17

[4] S. Hegde, N. M S, R. Samar, and R. H N, "Face Recognition Based Biometric Identification System with COVID-19 Detection," *Int J Res Appl Sci Eng Technol*, vol. 11, no. 2, 2023, doi: 10.22214/ijraset.2023.49319.

[5] S. Tharewal *et al.*, "Score-Level Fusion of 3D Face and 3D Ear for Multimodal Biometric Human Recognition," *Comput Intell Neurosci*, vol. 2022, 2022, doi: 10.1155/2022/3019194.

[6] R. Tripathi, A. J.-2021 3rd I. C. on, and undefined 2021, "Make-Up Invariant Face Recognition under Uncontrolled Environment," *ieeexplore.ieee.org*, Accessed: Feb. 12, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9451704/

[7] V. Albiero, K. Zhang, … M. K.-I. T. on, and undefined 2021, "Gendered differences in face recognition accuracy explained by hairstyles, makeup, and facial morphology," *ieeexplore.ieee.org*, Accessed: Feb. 12, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9650887/

[8] C. Rathgeb, P. Drozdowski, C. B.-I. Access, and undefined 2020, "Makeup presentation attacks: Review and detection performance benchmark," *ieeexplore.ieee.org*, Accessed: Feb. 12, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9293285/

[9] C. Rathgeb, C. Busch, C. Rathgeb, P. Drozdowski, D. Fischer, and C. Busch, "Vulnerability assessment and detection of makeup presentation attacks," *ieeexplore.ieee.org*, 2020, doi: 10.1109/IWBF49977.2020.9107961.

[10] C. Rathgeb, A. Dantcheva, and C. Busch, "Impact and Detection of Facial Beautification in Face Recognition: An Overview," *IEEE Access*, vol. 7, 2019, doi: 10.1109/ACCESS.2019.2948526.

[11] C. Rathgeb, … P. D.-… on P. R., and undefined 2021, "Detection of makeup presentation attacks based on deep face representations," *ieeexplore.ieee.org*, Accessed: Feb. 12, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9413347/

[12] N. Selitskaya, S. Sielicki, and N. Christou, "Challenges in Face Recognition Using Machine Learning Algorithms: Case of Makeup and Occlusions," *Advances in Intelligent Systems and Computing*, vol. 1251 AISC, pp. 86–102, 2021, doi: 10.1007/978-3-030-55187-2_9.

[13] M. Azimi, A. P.-C. & E. Engineering, and undefined 2020, "Investigation into the reliability of facial recognition systems under the simultaneous influences of mood variation and makeup," *Elsevier*, Accessed: Feb. 12, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045790620305176

[14] U. Saeed, K. Masood, H. D.-C. & E. Engineering, and undefined 2021, "Illumination normalization techniques for makeup-invariant face recognition," *Elsevier*, Accessed: Feb. 12, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045790620307734

[15] S. Y. Jhong *et al.*, "An automated biometric identification system using CNN-based palm vein recognition," in *International Conference on Advanced Robotics and Intelligent Systems, ARIS*, 2020. doi: 10.1109/ARIS50834.2020.9205778.

[16] M. Arab, P. A. Moghadam, … M. H.-P. of the 28th, and undefined 2020, "Revealing true identity: Detecting makeup attacks in face-based biometric systems," *dl.acm.orgMA Arab, P Azadi Moghadam, M Hussein, W Abd-Almageed, M HefeedaProceedings of the 28th ACM International Conference on Multimedia, 2020•dl.acm.org*, Accessed: Feb. 14, 2024. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/3394171.3413606

[17] M. A. Arab, P. Azadi Moghadam, M. Hussein, W. Abd-Almageed, and M. Hefeeda, "Revealing True Identity: Detecting Makeup Attacks in Face-based Biometric Systems," *MM 2020 - Proceedings of the 28th ACM International Conference on Multimedia*, pp. 3568–3576, Oct. 2020, doi: 10.1145/3394171.3413606.

[18] B. Alharbi and H. S. Alshanbari, "Face-voice based multimodal biometric authentication system via FaceNet and GMM," *PeerJ Comput Sci*, vol. 9, 2023, doi: 10.7717/peerj-cs.1468.

[19] T. Alashkar, S. Jiang, … Y. F. & gesture recognition (FG 2017, and undefined 2017, "Rule-based facial makeup recommendation system," *ieeexplore.ieee.orgT Alashkar, S Jiang, Y Fu2017 12th IEEE International conference on automatic face, 2017•ieeexplore.ieee.org*, 2017, doi: 10.1109/FG.2017.47.

[20] N. Selitskaya, S. Sielicki, … N. C. the 2020 I. S., and undefined 2021, "Challenges in face recognition using machine learning algorithms: case of makeup and occlusions," *Springer*, Accessed: Feb. 12, 2024. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-55187-2_9

[21] M. A. Arab, P. Azadi Moghadam, M. Hussein, W. Abd-Almageed, and M. Hefeeda, "Revealing True Identity: Detecting Makeup Attacks in Face-based Biometric Systems," *MM 2020 - Proceedings of the 28th ACM International Conference on Multimedia*, pp. 3568–3576, Oct. 2020, doi: 10.1145/3394171.3413606.

[22] T. V. Nguyen and L. Liu, "Smart mirror: Intelligent makeup recommendation and synthesis," *MM 2017 - Proceedings of the 2017 ACM Multimedia Conference*, pp. 1253–1254, Oct. 2017, doi: 10.1145/3123266.3127926.

[23] T. Alashkar, S. Jiang, S. Wang, … Y. F. the A. conference on artificial, and undefined 2017, "Examples-rules guided deep neural network for makeup recommendation," *ojs.aaai.orgT Alashkar, S Jiang, S Wang, Y FuProceedings of the AAAI conference on artificial intelligence, 2017•ojs.aaai.org*, Accessed: Feb. 12, 2024. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/10626

[24] D. Horita, K. A.-P. of the 4th A. International, and undefined 2022, "SLGAN: style-and latent-guided generative adversarial network for desirable makeup transfer and removal," *dl.acm.orgD Horita, K AizawaProceedings of the 4th ACM International Conference on Multimedia in Asia, 2022•dl.acm.org*, vol. 22, Dec. 2022, doi: 10.1145/3551626.3564967.

[25] S. Ueda, T. K.- Perception, and undefined 2010, "Influence of make-up on facial recognition," *journals.sagepub.comS Ueda, T KoyamaPerception, 2010•journals.sagepub.com*, Accessed: Feb. 14, 2024. [Online]. Available: https://journals.sagepub.com/doi/abs/10.1068/p6634

[26] X. Chen and T. Yang, "Influence of the Emergence of Gait Recognition System on Human Society," in *Proceedings of the 2013 International Conference on Advances in Social Science, Humanities, and Management*, 2013. doi: 10.2991/asshm-13.2013.38.

[27] S. Ueda and T. Koyama, "Influence of make-up on facial recognition," *Perception*, vol. 39, no. 2, pp. 260–264, 2010, doi: 10.1068/P6634.

[28] A. Silnova, N. Brümmer, D. Garcia-Romero, D. Snyder, and L. Burget, "Fast variational Bayes for heavy-tailed PLDA applied to i-vectors and x-vectors," *arxiv.org*, Accessed: Feb. 15, 2024. [Online]. Available: https://arxiv.org/abs/1803.09153

[29] A. Dantcheva, C. Chen, A. R.-2012 I. F. international, and undefined 2012, "Can facial cosmetics affect the matching accuracy of face recognition systems?," *ieeexplore.ieee.orgA Dantcheva, C Chen, A Ross2012 IEEE Fifth international conference on biometrics: theory, 2012•ieeexplore.ieee.org*, doi: 10.1109/BTAS.2012.6374605.

[30] D. Singhal and A. Doegar, "Face-Iris multimodal biometric system using feedforward backpropagation neural network," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 8 Special Issue 3, 2019.

[31] A. Gona and M. Subramoniam, "Convolutional neural network with improved feature ranking for robust multi-modal biometric system," *Computers and Electrical Engineering*, vol. 101, 2022, doi: 10.1016/j.compeleceng.2022.108096.

[32] A. Gona, M. Subramoniam, and R. Swarnalatha, "Transfer learning convolutional neural network with modified Lion optimization for multimodal biometric system," *Computers and Electrical Engineering*, vol. 108, 2023, doi: 10.1016/j.compeleceng.2023.108664.

[33] D. Alita, Y. Fernando, and H. Sulistiani, "IMPLEMENTASI ALGORITMA MULTICLASS SVM PADA OPINI PUBLIK BERBAHASA INDONESIA DI TWITTER," *Jurnal Tekno Kompak*, vol. 14, no. 2, 2020, doi: 10.33365/jtk.v14i2.792.

[34] M. Azhari, Z. Situmorang, and R. Rosnelly, "Perbandingan Akurasi, Recall, dan Presisi Klasifikasi pada Algoritma C4.5, Random Forest, SVM dan Naive Bayes," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 2, 2021, doi: 10.30865/mib.v5i2.2937.

[35] J. Huang, Y. Shang, and H. Chen, "Improved Viola-Jones face detection algorithm based on HoloLens," *EURASIP J Image Video Process*, vol. 2019, no. 1, 2019, doi: 10.1186/s13640-019-0435-6.

[36] Y.-Q. Wang, "An Analysis of the Viola-Jones Face Detection Algorithm," *Image Processing On Line*, vol. 4, 2014, doi: 10.5201/ipol.2014.104.

[37] T. Wang, A. K.-2016 I. I. C. on, and undefined 2016, "Recognizing human faces under disguise and makeup," *ieeexplore.ieee.orgTY Wang, A Kumar2016 IEEE International Conference on Identity, Security and, 2016•ieeexplore.ieee.org*, Accessed: Feb. 14, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7477243/

[38] M. U. Ragashe, M. M. Goswami, and M. M. Raghuwanshi, "Approach towards real time face recognition in streaming video under partial occlusion," in *Proceedings of 2015 IEEE 9th International Conference on Intelligent Systems and Control, ISCO 2015*, 2015. doi: 10.1109/ISCO.2015.7282394.

[39] E. Abusham, B. Ibrahim, K. Zia, and M. Rehman, "Facial Image Encryption for Secure Face Recognition System," *Electronics (Switzerland)*, vol. 12, no. 3, 2023, doi: 10.3390/electronics12030774.

[40] M. O. Balogun, L. Adeola Odeniyi, E. Olusola Omidiora, S. O. Olabiyisi, and A. S. Falohun, "Optimized Negative Selection Algorithm for Image

Classification in Multimodal Biometric System," *Acta Informatica Pragensia*, Apr. 2022, doi: 10.18267/j.aip.186.

[41] S. Sarin, A. Mittal, A. Chugh, and S. Srivastava, "CNN-based Multimodal Touchless Biometric Recognition System using Gait and Speech," *Journal of Intelligent and Fuzzy Systems*, vol. 42, no. 2, 2022, doi: 10.3233/JIFS-189765.

[42] A. Baaloul, N. Benblidia, … F. R.-M. T. and, and undefined 2024, "An arabic visual speech recognition framework with CNN and vision transformers for lipreading," *Springer*, Accessed: Feb. 15, 2024. [Online]. Available: https://link.springer.com/article/10.1007/s11042-024-18237-5

[43] A. K. Gona and M. Subramoniam, "Multimodal Biometric Reorganization System using Deep Learning Convolutional Neural Network," in *International Conference on Edge Computing and Applications, ICECAA 2022 - Proceedings*, 2022. doi: 10.1109/ICECAA55415.2022.9936398.

[44] N. Yang, Z. Zhang, J. Yang, Z. Hong, and J. Shi, "A Convolutional Neural Network of GoogLeNet Applied in Mineral Prospectivity Prediction Based on Multi-source Geoinformation," *Natural Resources Research*, vol. 30, no. 6, pp. 3905–3923, Dec. 2021, doi: 10.1007/S11053-021-09934-1.

[45] Y. Ibrahim, H. Wang, K. A.-2020 I. C. on, and undefined 2020, "Analyzing the reliability of convolutional neural networks on gpus: Googlenet as a case study," *ieeexplore.ieee.org*, Accessed: Feb. 18, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9213804/