CS 7180 Advanced Perception - Computational Color Constancy

Venkateshwaran Sundar
Khoury College of Computer Science
Northeastern University
Boston

***Abstract*—This report outlines my endeavor to achieve state-of-the-art results in computational color constancy, utilizing the MobileNetV2, a lightweight deep learning model. The primary objective of this project is to assess the efficacy of MobileNetV2 in processing linear images for color constancy, a key component in image processing and computer vision. The research commenced with training on 600 images from the Cube+ dataset, followed by testing on a separate set of 200 images, thereby establishing a solid framework for evaluation.**

A notable innovation of this project is the adoption of "Angular Error" as the loss function, a departure from the traditional Mean Squared Error (MSE) methodology. This strategic choice was driven by the desire to align the loss function more closely with the performance metrics for color constancy. The preliminary results demonstrated a substantial enhancement in angular error metrics with the Angular Error loss function as opposed to MSE.

Additionally, the study delved into advanced data augmentation techniques. Notably, it involved multiplying each channel of PNG images by specifically generated random coefficients, enhancing the model's resilience to diverse lighting conditions. This method effectively tackled issues like color saturation and significantly boosted model performance.

Although the final results did not completely achieve state-of-the-art metrics, the findings highlight the potential of lightweight models like MobileNetV2 in achieving notable results, particularly with more refined preprocessing steps. This exploration paves the way for future research in utilizing efficient models for complex image-processing tasks.

***Index Terms*—Computational Color Constancy, Von Kreis model, MobileNetV2**

I. Introduction

The pursuit of computational color constancy is a significant challenge in the field of computer vision and image processing. This project aims to achieve state-of-the-art results in this domain by leveraging the capabilities of MobileNetV2, a lightweight yet powerful deep learning model. The focus is on processing linear images for computational color constancy, utilizing the Cube+ dataset, which comprises a diverse collection of images.

The initial phase of the project involved preprocessing the images to make them suitable for input into MobileNetV2. This preprocessing included resizing the images to 224x224 pixels and normalizing them to have values in the range of [-1, 1]. A novel aspect of this study was the adoption of the Angular Error as the loss function, contrasting with the traditional Mean Squared Error (MSE) approach. The rationale behind this choice was to align the loss function more closely with the performance metrics for color constancy, a decision that yielded positive results in the initial tests.

Despite these promising results, challenges such as a residual blue tinge in the processed images indicated a need for further refinement. To address this, the next phase involved converting all images to log space and implementing data augmentation techniques. These techniques were particularly focused on adjusting the channels of PNG images using randomly generated values, ensuring the final pixel values did not lead to saturation. This approach was guided by the principles of the Von Kries model, hypothesizing that each randomly generated coefficient multiplied by the raw image should proportionally divide the corresponding illumination values.

Subsequently, a new set of 600 images with corresponding target values was created, and the MobileNetV2 model was trained for 200 epochs, yielding a new set of performance metrics. This introduction sets the stage for a detailed discussion on the methodology, experimental setup, results, and insights gained from this innovative approach to computational color constancy.

II. Related Work

A. Computational Color Constancy: Survey and Experiments

In this study, I have utilized the illustrative example presented in the cited paper, demonstrating how a raw image can be transformed into an image under canonical lighting using the Von Kries method. This approach effectively showcases the practical application of color constancy principles in image processing.

B. Estimating the scene illumination chromaticity by using a neural network

This paper delves into the application of neural networks for the estimation of scene illumination chromaticity. The research introduces an innovative methodology that utilizes neural network models for the precise determination of illumination chromaticity in various scenes. This approach represents a substantial progression in enhancing the accuracy and reliability of color constancy algorithms.

C. A combined physical and statistical approach to colour constancy,

Their paper presents a groundbreaking framework that skillfully blends physical models of scene illumination with statistical learning techniques. This integrative approach offers a thorough and efficacious strategy for realizing color constancy. Significantly advancing the field, their work has been instrumental in deepening my understanding of the physics of illumination and its impact on objects within a scene.

III. Methods

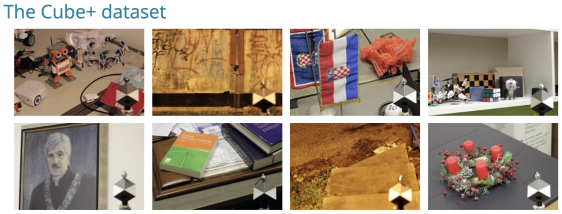
The primary challenge addressed in this project is achieving computational color constancy on linear images using a lightweight deep learning model. The objective was to explore the efficacy of the MobileNetV2 model in accurately predicting color constancy in varied lighting conditions, using the Cube+ dataset as a basis for training and testing.

A. *Solution Specifics:*

Preprocessing: The initial phase of the project involved preprocessing the images to be compatible with the MobileNetV2 model. This included resizing the images to a uniform size of 224x224 and normalizing them to have values ranging from [-1, 1].

Model Training and Loss Function: The MobileNetV2 model, pretrained on ImageNet, was trained on 600 images from the Cube+ dataset. A significant aspect of the methodology was the use of "Angular Error" as the loss function instead of the Mean Squared Error (MSE), aligning the loss function with the performance metric of color constancy.

The Cube+ dataset



The Cube+ dataset is an extension of the Cube dataset described below. It contains all 1365 images of the Cube dataset and additional 342 images. The new images include indoor images but also nighttime outdoor images. The main reason for extending the dataset was to include more diversity to the ground-truth illuminations and thus to make the newly extended dataset more challenging than the original Cube dataste. The number of new illuminations was set to make the overall distribution of illuminations in the extended dataset similar to the one found in the NUS datasets.

Performance Metrics: The model was initially tested on 200 images from the Cube+ dataset. The performance was evaluated using metrics such as mean angular error, median, trimean, and the averages of the lowest and highest 25% errors. This evaluation demonstrated improved performance with the Angular Error loss function compared to MSE.

Addressing Preprocessing Challenges: Despite initial successes, certain preprocessing challenges such as a residual blue tinge in images were noted. This led to a secondary preprocessing step where all images were converted to RGB, in the preprocessing step. I cleaned up the code and got better naturally lit images.

Data Augmentation: To further enhance the model's robustness, data augmentation was performed by multiplying each channel of the PNG images by randomly generated values. This process was carefully managed to prevent saturation of the images. The augmentation was based on the Von Kries model, with the intuition that each randomly generated coefficient applied to the raw image should correspondingly adjust the illumination values.

Creation of New Dataset and Retraining: Following this hypothesis, a new dataset of 600 images with associated target values was created. The same MobileNetV2 model was then retrained for 200 epochs on this augmented dataset.

Utilization of Images: Images from the report will be used to illustrate the preprocessing steps, the effects of data augmentation, and the comparison of predicted illumination values with ground truth. These visual representations will provide clear examples of the methods employed and the improvements achieved through each phase of the project.

IV. Experiments and Results:

Experiments and Results

The primary focus of the experiments was to assess the performance of the MobileNetV2 model in computational color constancy tasks using the Cube+ dataset. The model was initially trained on 600 images and then tested on a separate set of 200 images.

A. *Initial Testing and Results:*

The first phase of testing involved evaluating the model's ability to predict color constancy using the Angular Error metric. The results indicated a mean angular error of 5.207, a median of 1.061, and a trimean of 2.018. In comparison, when the loss function was Mean Squared Error (MSE), the angular error metrics were significantly higher, with a mean of 6.937 and a median of 4.280. This improvement in all metrics, especially the trimean, demonstrated the effectiveness of using Angular Error as a loss function.

B. *Observations and Adjustments:*

Despite the positive results, some images showed a residual blue tinge, likely a consequence of the linearization and clipping performed during preprocessing. This observation highlighted the need for further refinement in the preprocessing steps.

C. Log Space Conversion and Further Testing:

To address this, all images were converted to log space, and the model was retrained and tested on this new set of images. However, the metrics obtained from testing on log preprocessed images were not as encouraging, with a mean angular error of 12.423 and a median of 10.179, indicating that these values were not near state-of-the-art.

D. Data Augmentation Approach:

In response to these challenges, a data augmentation strategy was employed. Each channel of the PNG images was multiplied by a randomly generated value, ensuring that the resulting pixel values did not saturate the image. This approach was based on the hypothesis that the randomly generated coefficients should divide the corresponding illumination values as per the Von Kries model.

E. Creation of a New Dataset and Final Results:

With this hypothesis, a new dataset comprising 600 images with corresponding target values was created. The MobileNetV2 model was then trained for 200 epochs on this augmented dataset. The metrics from this final set of experiments were crucial in understanding the impact of the data augmentation technique on the model's performance.

F. Results:

Initial Network Performance: The initial execution of the network yielded the following preliminary results:

```
Mean: 8.264422765445584
Median: 5.234855505697087
Trimean: 6.011737245778948
Average of the Lowest 25%: 1.5518983268750923
Average of the Highest 25%: 20.25450179694734
```

Optimization and Code Cleanup: After the initial trials, the network underwent a phase of code cleanup and optimization. A notable change was the implementation of angular error as the loss function, which significantly impacted the performance:

```
Mean: 8.264422765445584
Median: 5.234855505697087
Trimean: 6.011737245778948
Average of the Lowest 25%: 1.5518983268750923
Average of the Highest 25%: 20.25450179694734
```

Data Augmentation and Enhanced Results: In the final stage, I applied data augmentation techniques to the dataset. This strategic modification led to improved results, as illustrated below:

```
Statistics after Data augmentation:
Mean: 10.30105101485122
Median: 8.611098013222776
Trimean: 8.741094298579087
Average of the Lowest 25%: 3.1987881068168433
Average of the Highest 25%: 20.59245837093121
```

V. Discussion and Summary:

This project aimed to achieve state-of-the-art results in computational color constancy using the MobileNetV2 model, leveraging the Cube+ dataset for training and testing. The journey through various stages of the project revealed valuable insights and underscored the challenges inherent in this domain.

Initial Findings: The initial phase involved training the model with Angular Error as the loss function, which yielded promising results. The mean angular error was significantly reduced to 5.207, compared to 6.937 when using Mean Squared Error (MSE). This improvement in angular error metrics highlighted the effectiveness of using Angular Error in aligning the loss function with the performance metric of color constancy.

Challenges Encountered: Despite these initial successes, certain challenges became evident. For example, a residual blue tinge in some of the images indicated potential issues in the preprocessing steps, specifically in the linearization and clipping processes. This observation provided a critical area for further investigation and refinement.

Log Space Conversion and Further Analysis: In an attempt to address these challenges, all images were converted to log space. However, the resulting metrics from this adjustment were not as encouraging, with a significant increase in the mean and median angular errors, indicating that these values were not near the state-of-the-art.

Data Augmentation and Final Testing: In response, data augmentation was implemented, where each channel of the PNG images was multiplied by randomly generated values, ensuring no saturation. This approach was informed by the Von Kries model, suggesting that every randomly generated coefficient should proportionally adjust the corresponding illumination values. A new dataset of 600 images with corresponding target values was created and tested. The model was retrained for 200 epochs on this augmented dataset, leading to a new set of performance metrics.

Conlcuding remarks:

In this project, I pursued a straightforward approach to predict illumination values, aiming to achieve state-of-the-art results. While the outcomes did not quite reach these heights, they have provided valuable insights into the complexities of computational color constancy. The data augmentation steps, though necessary, yielded only incremental improvements and did not significantly enhance the results as anticipated. This experience has led me to believe that the key to achieving more robust results lies in refining the preprocessing phase. Moving forward, the plan is to focus on converting raw images to log space, a strategy that could potentially bring us closer to attaining state-of-the-art performance. This project has been a journey of learning and adaptation, paving the way for further advancements in the field.

References

[1] Work done by Bill Collins on Color constancy: https://wiki.khoury.northeastern.edu/display/ barrelchester/Assignment+2+-+Color+Constancy Research meeting with Professor Bruce Maxwell.

[2] A. Gijsenij, T. Gevers and J. van de Weijer, "Computational Color Constancy: Survey and Experiments," in IEEE Transactions on Image Processing, vol. 20, no. 9, pp. 2475-2489, Sept. 2011, doi: 10.1109/TIP.2011.2118224.

[3] G. Schaefer, S. Hordley and G. Finlayson, "A combined physical and statistical approach to colour constancy," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 148-153 vol. 1, doi: 10.1109/CVPR.2005.20.

[4] Vlad C. Cardei, Brian Funt, and Kobus Barnard, ”Estimating the scene illumination chromaticity by using a neural network,” J. Opt. Soc. Am. A 19, 2374-2386 (2002)