**Documentation Report**

**Gender Classification using Convolutional Neural Networks (CNN)**

**Task Overview:**

The primary objective of this project was to implement Convolutional Neural Networks (CNN) for gender classification based on facial images. The dataset contained images representing both male and female faces. The training dataset consisted of 1,000 samples for each gender category, while the validation dataset included 400 samples for each category.

**Data Collection and Preparation**

The dataset was sourced from Kaggle and was carefully balanced to ensure an equal number of male and female samples. To optimize computational efficiency, all images were resized to a consistent 150x150 pixel resolution. This preprocessing step, while essential for computational efficiency, contributed significantly to the data preparation phase's duration.

**Model Selection and Architecture**

The project's core focus was the utilization of Convolutional Neural Networks (CNNs) for gender classification. Key aspects of the model architecture included:

**Input Layer**

The input layer was configured to accept 150x150-pixel images with three color channels (RGB).

**Convolutional Layers**

Multiple convolutional layers were employed to extract essential image features. The number of filters and kernel sizes were tuned through experimentation.

**Pooling Layers**

Max-pooling layers were utilized to down sample feature maps and reduce spatial dimensions.

**Fully Connected Layers**

Flattened feature maps were passed through fully connected layers with ReLU activation functions.

**Output Layer**

The output layer consisted of a single neuron with a sigmoid activation function, generating binary gender classification output.

**Model Training and Evaluation**

The CNN model underwent training over 10 epochs. Below are the training accuracy and validation accuracy for each epoch:

**Epoch 1/10**

Training Loss: 0.6775

Training Accuracy: 0.5597

Validation Loss: 0.5276

Validation Accuracy: 0.7680

**Epoch 2/10**

Training Loss: 0.6152

Training Accuracy: 0.6860

Validation Loss: 0.5839

Validation Accuracy: 0.7509

**Epoch 3/10**

Training Loss: 0.5904

Training Accuracy: 0.7040

Validation Loss: 0.5068

Validation Accuracy: 0.7705

**Epoch 4/10**

Training Loss: 0.5943

Training Accuracy: 0.6880

Validation Loss: 0.4980

Validation Accuracy: 0.7790

**Epoch 5/10**

Training Loss: 0.5949

Training Accuracy: 0.6914

Validation Loss: 0.4567

Validation Accuracy: 0.8205

**Epoch 6/10**

Training Loss: 0.5646

Training Accuracy: 0.7250

Validation Loss: 0.4501

Validation Accuracy: 0.8156

**Epoch 7/10**

Training Loss: 0.5687

Training Accuracy: 0.7148

Validation Loss: 0.5050

Validation Accuracy: 0.7521

**Epoch 8/10**

Training Loss: 0.5572

Training Accuracy: 0.7275

Validation Loss: 0.4550

Validation Accuracy: 0.7998

**Epoch 9/10**

Training Loss: 0.5512

Training Accuracy: 0.7289

Validation Loss: 0.4334

Validation Accuracy: 0.8254

**Epoch 10/10**

Training Loss: 0.5232

Training Accuracy: 0.7455

Validation Loss: 0.4174

Validation Accuracy: 0.8095

**Model Evaluation and Results**

After rigorous experimentation and evaluation, the CNN model demonstrated promising results on the validation dataset:

**Final Accuracy:** The CNN model achieved an accuracy of 0.8095, affirming its capability to classify gender based on facial images.

**F1 Score:** The F1 score, a balance between precision and recall, was calculated at 0.49.

**Precision:** The precision score was 0.49, indicating the model's proficiency in gender classification.

These results underscore the potential for further optimization and potentially a more complex CNN architecture to enhance performance.

**Challenges and Limitations**

Several challenges and limitations were encountered during the project:

**Complexity of Image Data:** Working with image data, inherently high-dimensional, posed computational and memory challenges. The model's performance significantly depends on feature extraction, and image-based tasks can be resource-intensive.

**Accuracy vs. Data Size:** While the CNN exhibited potential, achieving higher accuracy may necessitate a larger dataset and more advanced techniques, such as further tuning of CNN architectures.

**Image Resizing and Computational Time:** Preprocessing steps like image resizing contributed to increased data preparation time. Striking a balance between computational efficiency and model accuracy remains an ongoing challenge.

**Data Reduction Impact:** Data reduction, while essential for computational efficiency, may have affected the model's generalization and accuracy.

Conclusion

In conclusion, this project highlighted the effectiveness of Convolutional Neural Networks (CNNs) for gender classification based on facial images. The CNN model achieved a commendable accuracy of 0.8095 on the validation dataset, demonstrating its potential in gender classification tasks.

However, it's essential to acknowledge the complexities associated with image data and the challenges in achieving higher accuracy.

Future work in this field may involve exploring more advanced CNN architectures, optimizing image preprocessing steps, and considering the trade-offs between data reduction for computational efficiency and enhancing model accuracy.