# Fully Complex Valued Neural Network for Keypoints Detection

Kollamoram Karthik (2430322) under Prof. Koteswar Rao Jerripothula

Indian Institute of Technology Kanpur SURGE 2024

## Abstract

Keypoint detection is a critical task in computer vision, with applications in facial landmark detection and pose estimation. Despite the advancements in CNNs, , there exists an underexplored area in utilizing complex-valued operations for this purpose. This project addresses the challenge by integrating **fully complex-valued operations and inputs** inspired by concepts from Fully Complex-valued Convolutional Networks **(FCCNs)**. Our approach involves adapting these principles to enhance the accuracy and robustness of keypoint detection systems. While our model takes complex-valued images as inputs and performs operations solely in the complex domain, the loss function remains conventional. This work sets a foundation for future advancements in handling complex visual scenarios effectively.

## Theoretical Framework

This project introduces complex-valued convolutional neural networks for keypoint detection, integrating complex operations such as **1x1 convolutions** throughout the network. Unlike traditional CNNs, this approach aims to maintain complex-valued information flow from input to output, potentially enhancing accuracy and robustness in complex visual tasks like facial landmark detection and pose estimation. The model utilizes layers like **ComplexConv2D** and **ComplexBatchNormalization**, optimized with Adam optimizer and trained with Mean Squared Error loss.

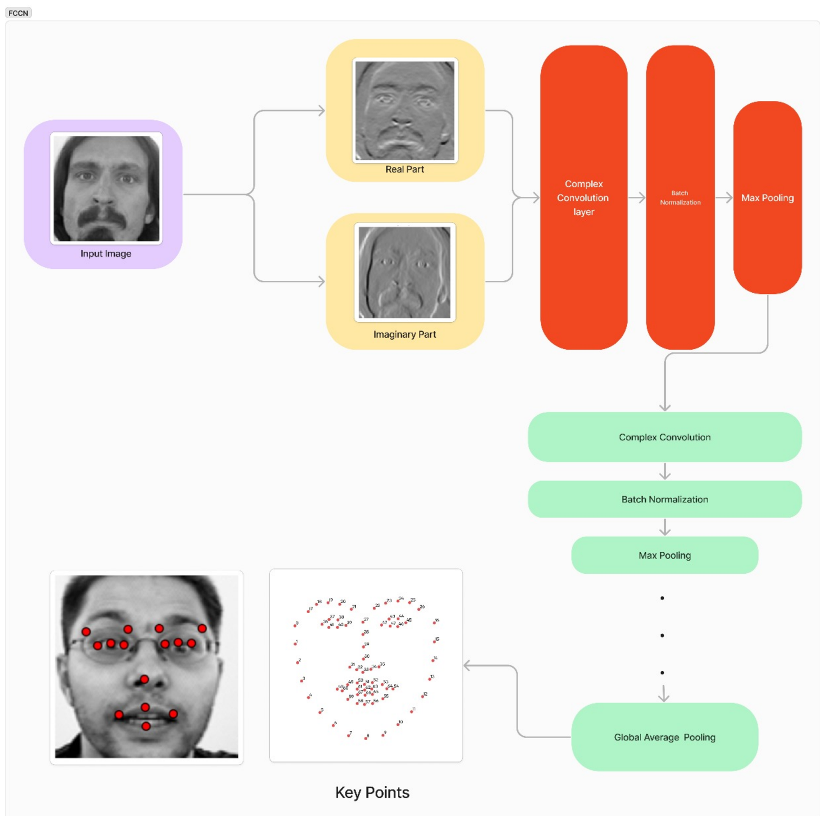



Figure 1. FCCN Architecture

## Research objectives

The present study investigates the following objectives:

- Improvise **complex information flow** using FCCN-inspired concepts in keypoint detection.
- Compare accuracy and performance results with a traditional CNN of similar parameters size.

## Study methodology

The present study adopted the following step-by-step methodology to achieve the research objectives.

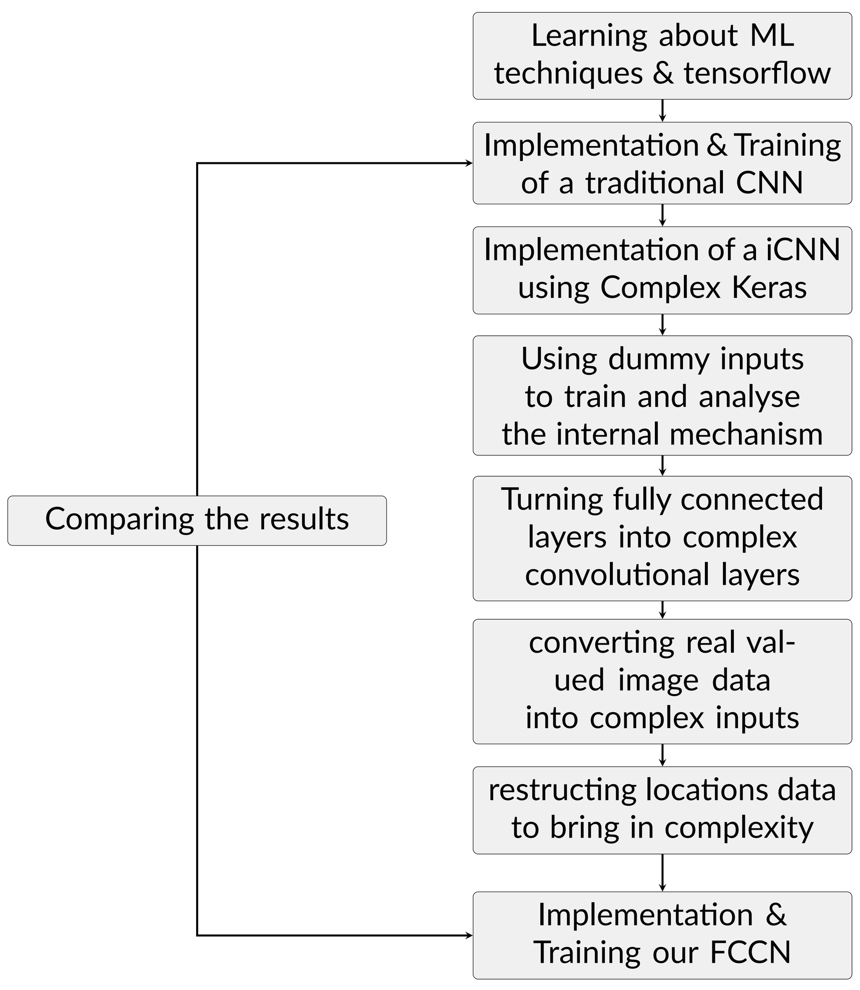



Figure 2. PROJECT FLOWCHART

## How we converted the real image data into complex inputs?

- The Complex-valued Local Binary Pattern (iLBP) function processes image data to generate complex representations. It shifts the image in eight directions, calculates pixel-wise differences, and multiplies these by The Complex-valued Local Binary Pattern (iLBP) function processes image data to generate complex representations. It shifts the image in eight directions, calculates pixel-wise differences, and multiplies these by specific complex numbers. The combined results are split into real (FCR) and imaginary (FCI) sub-images. We used this function to convert our image data into complex inputs.

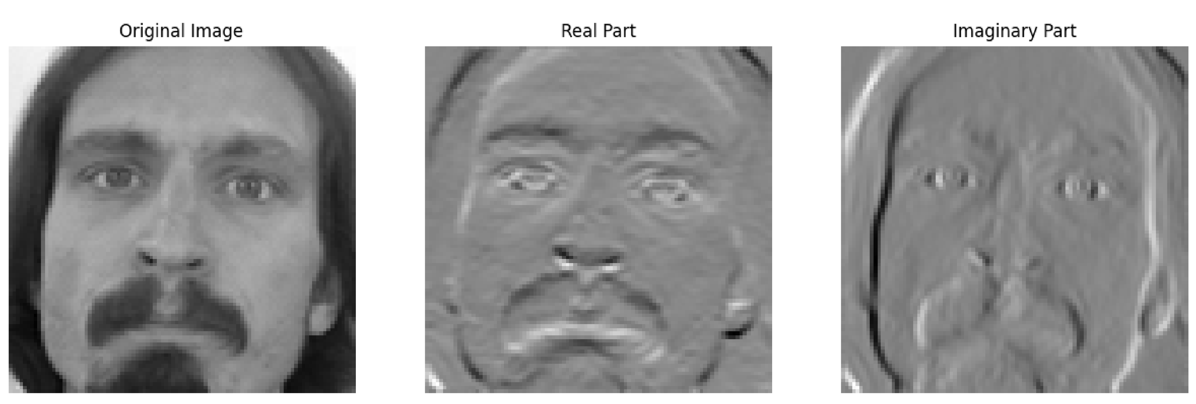


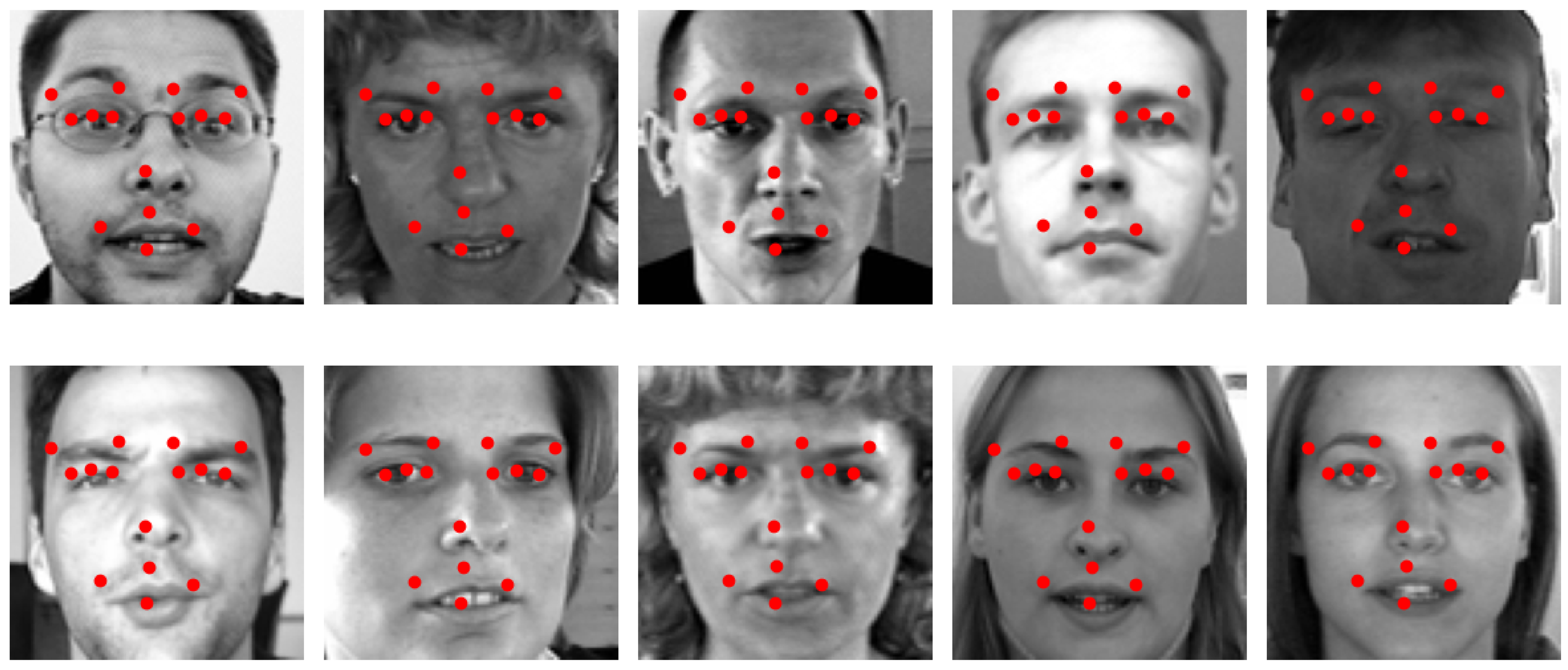

Figure 3. The original image and its real and imaginary parts obtained using the iLBP function

## Results

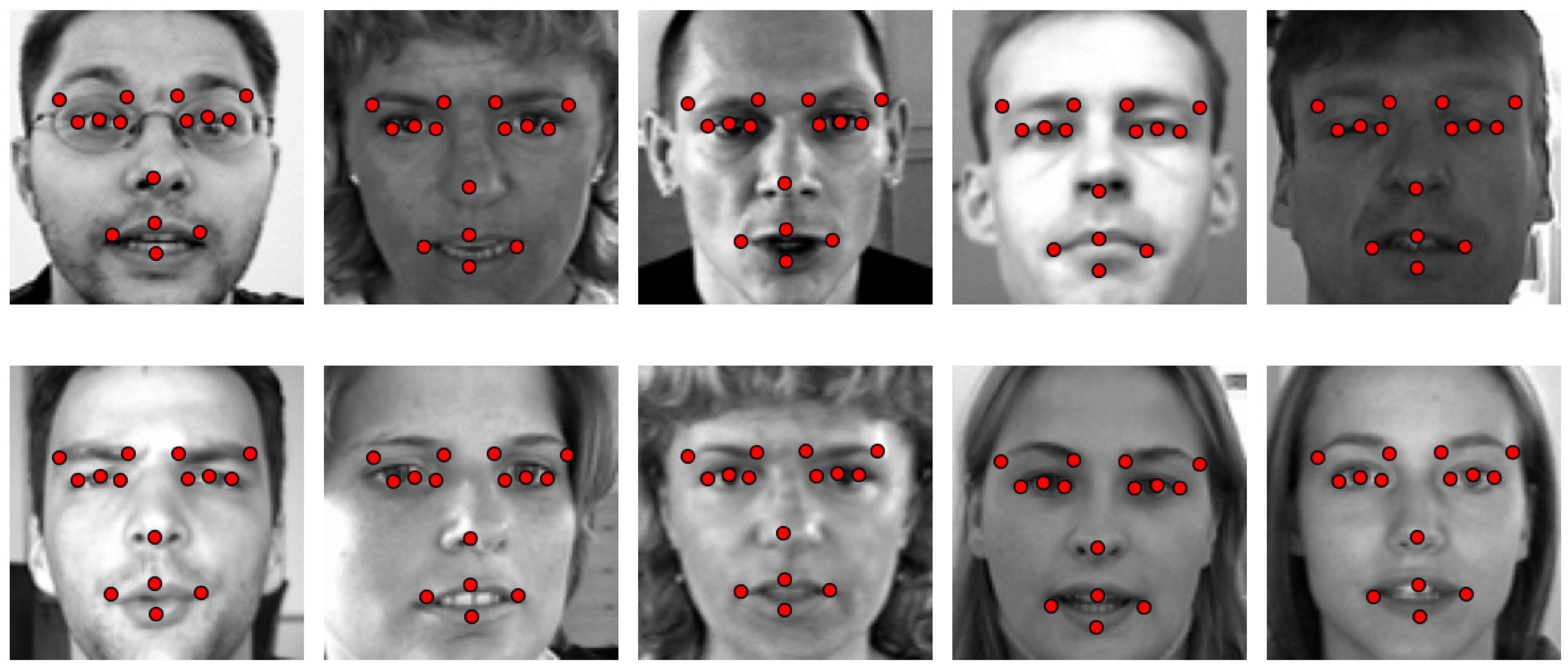
### Model results

Our model demonstrated improved performance compared to the traditional CNN. Despite the CNN having **more than twice** as many **parameters**, our implemented model achieved higher **validation accuracy, precision, recall, and F1 score metrics**. This highlights the effectiveness of leveraging complex-valued operations for achieving efficient and effective keypoint detection.

**Here are the keypoints detected by a traditional CNN compared to our model for a few test images.**



(a) Keypoints detected by CNN (Total params: 7268670 (27.73 MB))



(b) Keypoints detected by FCCN (Total params: 3,288,222 (12.54 MB))

We now tabulate the performances for both the models trained and implemented.

Table 1. Compiled Results

| Metrics | CNN | Our Model |
|---|---|---|
| Model size | 27.73 B | 12.5 4MB |
| F1 Score | 0.596 | 0.647 |
| Precision | 0.511 | 0.723 |
| Validation Accuracy | 0.715 | 0.734 |

## Conclusions

- **Our study highlights improved model accuracy by integrating complex-valued operations in keypoint detection. This underscores the efficiency gains from enhancing architectural flow for complex information, indicating promising directions for future research in complex-valued neural networks for computer vision.**

## Future Prospects

- Looking ahead, our model holds promise for further optimization through the implementation of custom-designed **complex loss functions and activation functions**, opening up a new dimension for research in the analysis of complex numbers in machine learning.
- Furthermore, as we achieve better models with less computational efforts, this approach holds potential for broad application across industries utilizing keypoint detection. Examples include **healthcare for patient monitoring, robotics for precise movement tracking, and augmented reality for immersive experiences**, illustrating its potential impact in many industrial applications.

## Experimental Setup

Epochs: 20; Training Dataset Size: 2140 images; Validation Split: 0.2 (i.e, 428 images); Optimizer: Adam (learning rate = 1e-3); Loss Function: Mean Squared Error; Metrics: MAE, Accuracy;

## References

• Saurabh Yadav and Koteswar Rao Jerripothula, "FCCNs: Fully Complex-valued Convolutional Networks using Complex-valued Color Model and Loss Function," in IEEE/CVF International Conference on Computer Vision (ICCV), 2023. **[pdf] [code]**

• For the dataset used, please refer to the Facial Keypoints Detection dataset on Kaggle: `https://www.kaggle.com/c/facial-keypoints-detection`.

• For the implementation using Complex Keras, refer to the following Documentation and GitHub repository : **[doc] [code]**.

• For the Training and implementation of our FCCN and CNN models please refer to the following GitHub repository : **[code]**.