

Project Design Phase-I

Solution Architecture

Date	21 September 2022
Team ID	PNT2022TMID591732
Project Name	Project - Image Caption Generation
Maximum Marks	4 Marks

Solution Architecture Simplified Steps:

1. Problem Statement:

- **Business Problem:** Make the platform more interesting and engaging by providing meaningful captions for images.

2. Requirements Gathering:

- **Understand Expectations:** Talk a lot with everyone involved to know what users want and what the platform aims to achieve.
- **Define Image Generator Needs:** Figure out how the image caption generator should work, considering things like clear captions, diverse language, and cultural awareness.

3. Data Collection and Preprocessing:

- **Get Diverse Data:** Collect lots of different pictures to cover various categories and cultural contexts.
- **Prepare Data:** Make sure pictures and captions are ready for the computer to understand, including making them the right size and organizing the words.

4. Technology Selection:

- **Pick Tools:** Choose computer tools like TensorFlow or PyTorch that work well with pre-trained models.
- **Choose Models:** Decide on pre-trained models, like VGG16, that are good at pulling important details from images.

5. Architecture Components:

- **Image Encoder:** Use a special computer program (CNN) to understand details in pictures.
- **Caption Generator:** Create a system using smart computer programs (RNN or transformers) to make captions that fit the picture.
- **Tokenizer:** Develop a tool to help the computer understand and process words in the captions.
- **Loss Function and Optimizer:** Use special methods to make the computer learn better from the data.
- **Post-Processing (Optional):** Add extra steps to make the captions even better, using language techniques.

6. Specification Documentation:

- **Write Down Plans:** Make detailed plans for each part of the computer system, explaining what each part does and how they work together.
- **Explain Expectations:** Write down what the computer should do and how it should act, making it clear for everyone.

7. Development Phases:

- **Step-by-Step Development:** Plan how to build the system, starting from getting data to making the computer smarter in stages.
- **Stay Flexible:** Use flexible ways of working to adjust to changes in what's needed and new technology.

8. User Interface :

- **Make It Easy:** Design an easy-to-use system for people to upload pictures and get cool captions.
- **Include Everyone:** Think about making it work for everyone, even those who might find it harder to use.

9. Monitoring and Maintenance:

- **Keep an Eye on Things:** Use tools to watch how well the computer system is doing in real-time.
- **Fix Issues:** Regularly check and fix any problems, update the system with new information, and keep it working well.

10. Delivery and Deployment:

- **Give It to Everyone:** Make sure the finished system is reliable and can be used by a lot of people.
- **Test It Well:** Check everything works correctly before letting people use it for real.

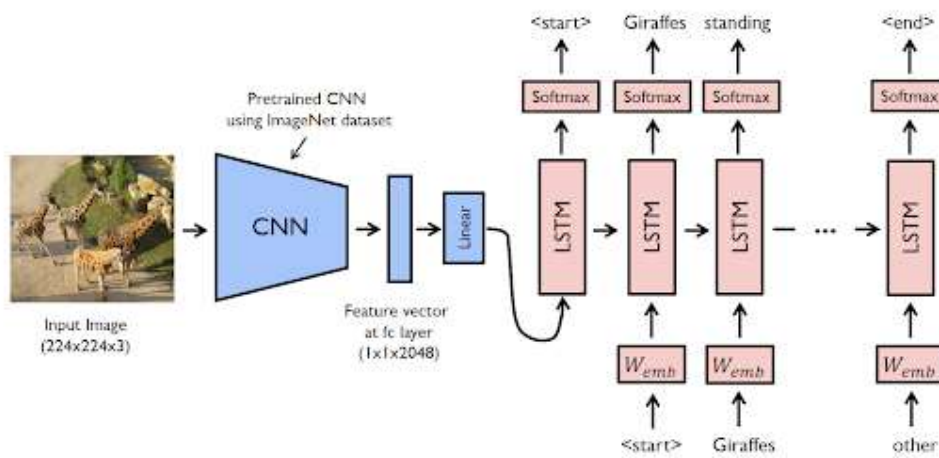
11. Stakeholder Communication:

- **Keep Talking:** Stay in touch with everyone involved, updating them on how things are going and solving any challenges.
- **Ask for Feedback:** Get thoughts and ideas from everyone to make things better.

12. Ethical Considerations:

- **Do the Right Thing:** Think about what's right and wrong, making sure the system doesn't treat anyone unfairly.
- **Be Open:** Make sure people know how the computer is making captions and deal with any concerns responsibly.

Example - Solution Architecture Diagram:



Reference: [Image Caption Generator Using Deep Learning \(image-caption-generator.blogspot.com\)](http://image-caption-generator.blogspot.com)