Evaluating Fourier Transform Features for Detection of AI-Generated Images

Sami Nourji, Tanay Subramanian, Everest Yang, Sujith Pakala

Project Motivation and Real-World Context

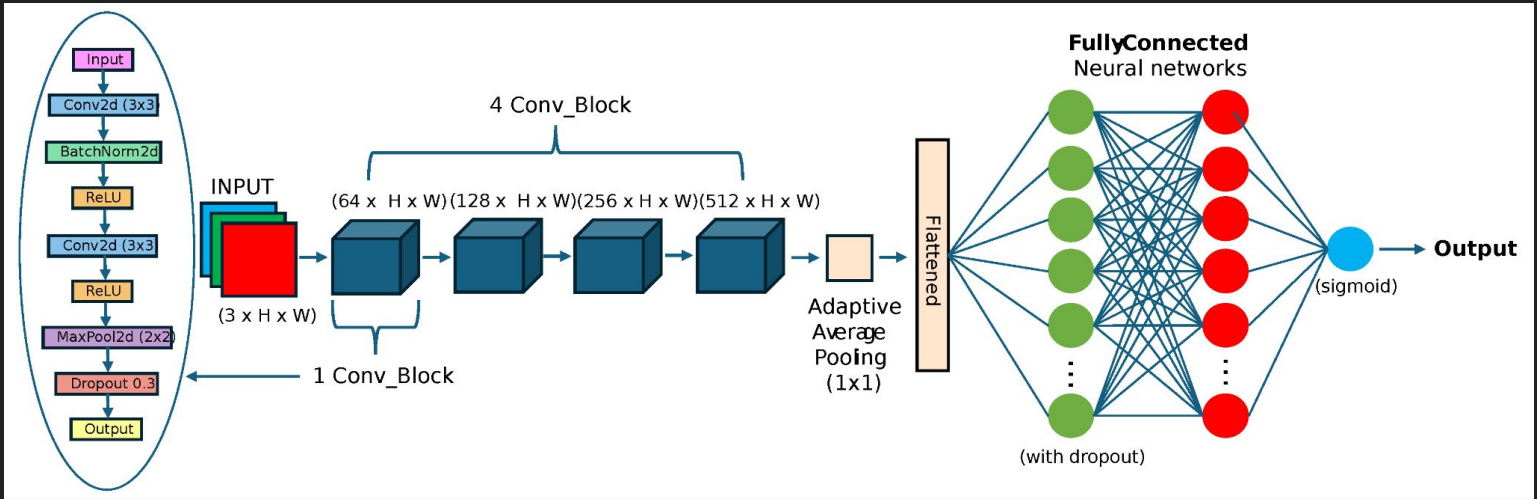
- Advances in AI Image Generation
- Misinformation & Privacy Risks
- Need for Accurate Detection Models

Initial Observation

Project Introduction

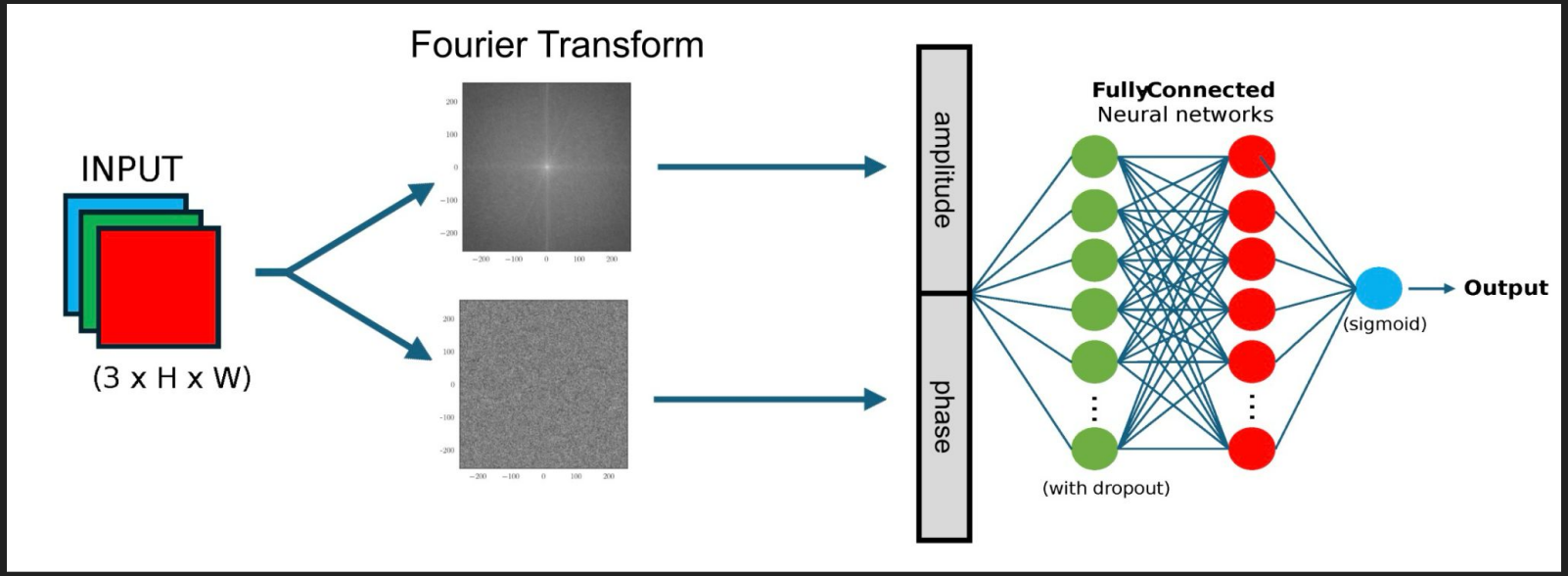
- **Hypothesis:**
  - Incorporating frequency information via Fourier transforms, in addition to spatial domain information, into a CNN can enhance the detection of AI-generated images
- **Dataset:**
  - CIFAKE dataset containing 120,000 total images (half AI-generated, half real)
- **Training:**
  - OSCAR GPUs

Technical Implementation Details



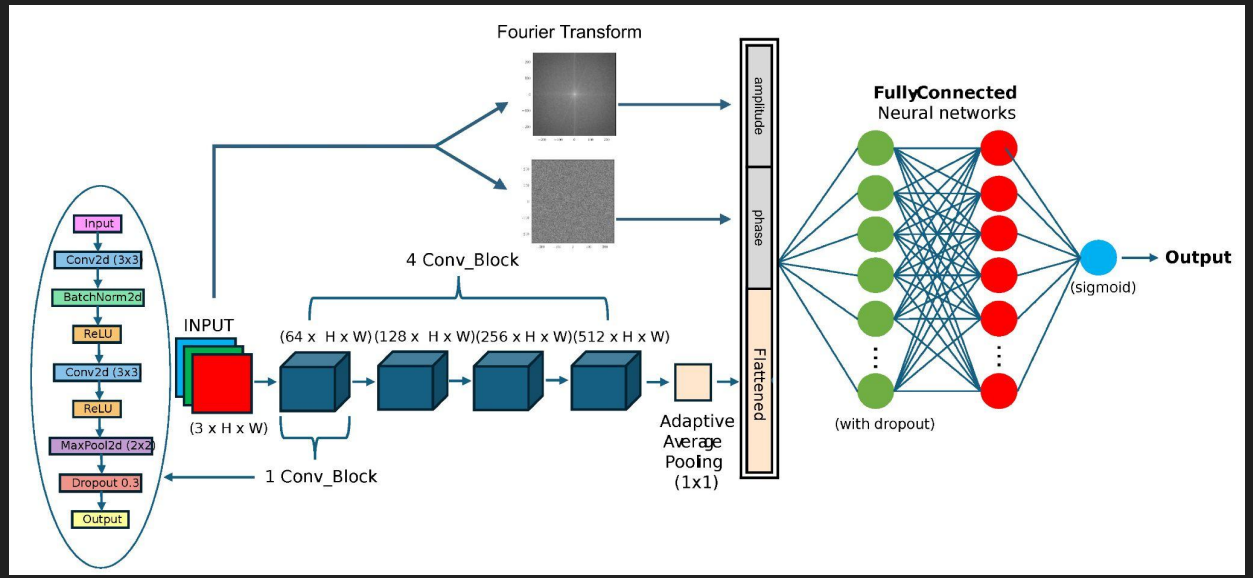
Original Architecture - 98.58%

Technical Implementation Details (Cont.)



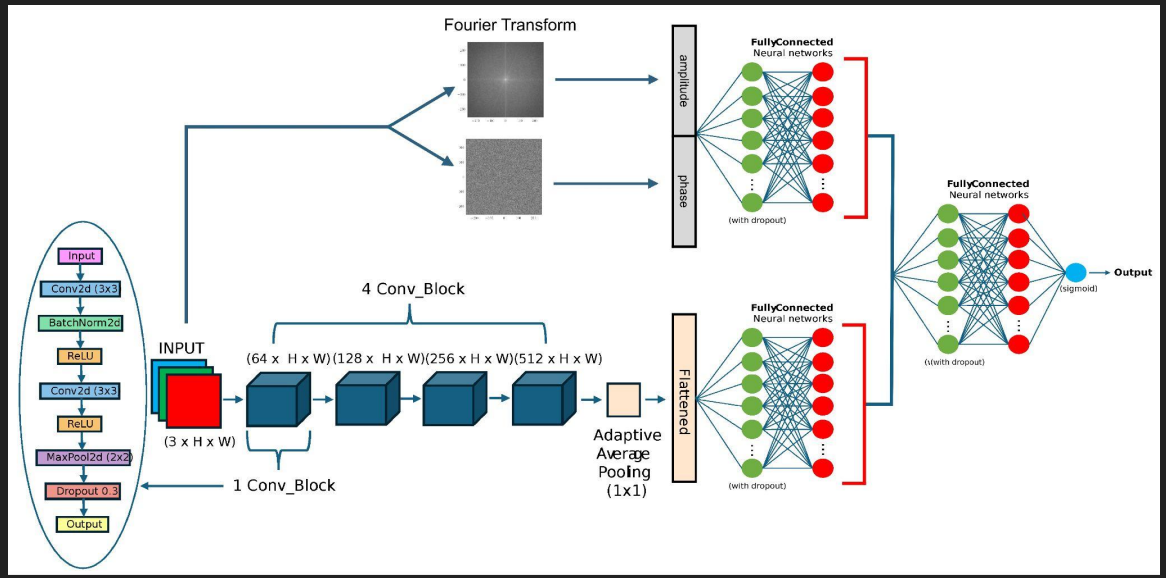
Fourier Transform - 82.59%

Technical Implementation Details (Cont.)



Experiment 1 (concatenation)
Actual Fourier: 95.36%
Random Noise: 50.35%

Technical Implementation Details (Cont.)



Experiment 2 (combination)
Actual Fourier: 98.5%
Random Noise: 98.5%

Results and Conclusion

- Frequency representation of images contain information that helps discern between real and AI-generated images
- However, this information is not complementary to the one extracted by a CNN
- Therefore, Fourier Transforms do not provide significant added value to AI-Generated Image Detection

Thank you!

References

- Sources:
  - Detection of AI-Generated Synthetic Images with a Lightweight CNN
  - Deepfake Detection with Deep Learning: Convolutional Neural Networks versus Transformers
  - Harnessing Machine Learning for Discerning AI-Generated Synthetic Images
  - Fourier Transform Layer: A proof of work in different training scenarios
  - CIFAKE: Image Classification and Explainable Identification of AI-Generated Synthetic Images
  - Faster Than Lies: Real-time Deepfake Detection using Binary Neural Networks