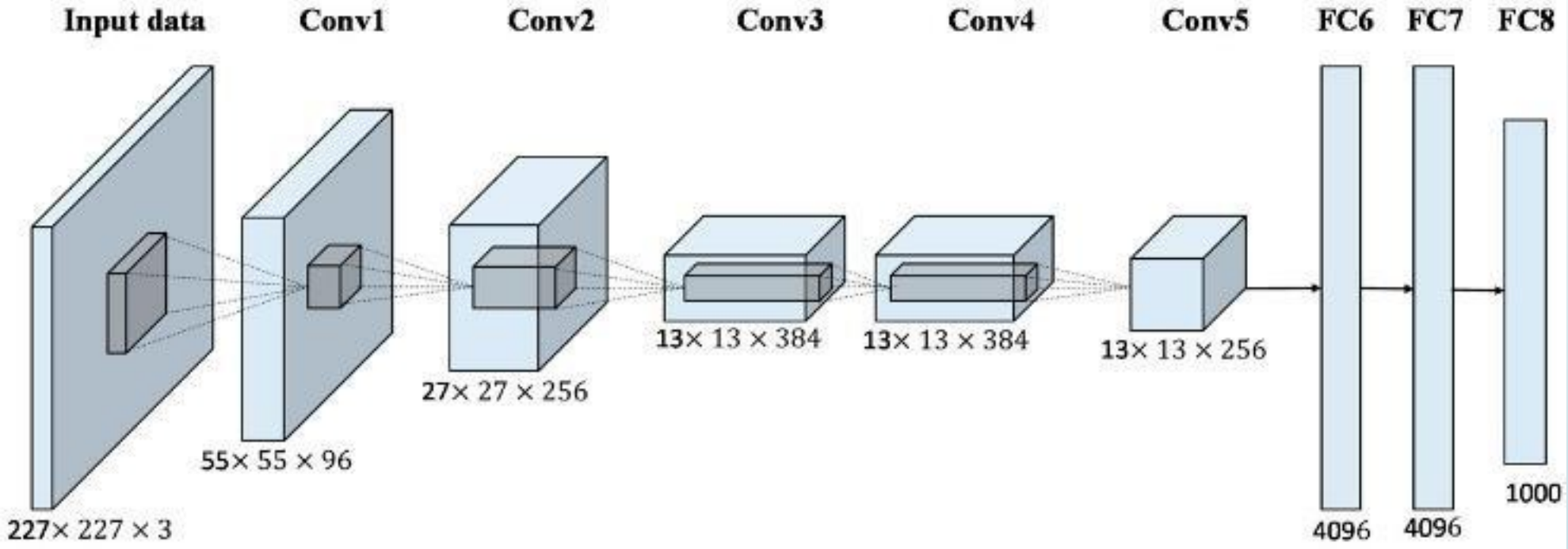
CNN’s Family

AlexNet (2012)

- Won the ImageNet competition in 2012, popularizing deep learning.
- Five convolutional layers followed by three fully connected layers, introduced **ReLU activation** for faster training.
- Used **dropout** for regularization.
- Trained on GPUs for the first time.

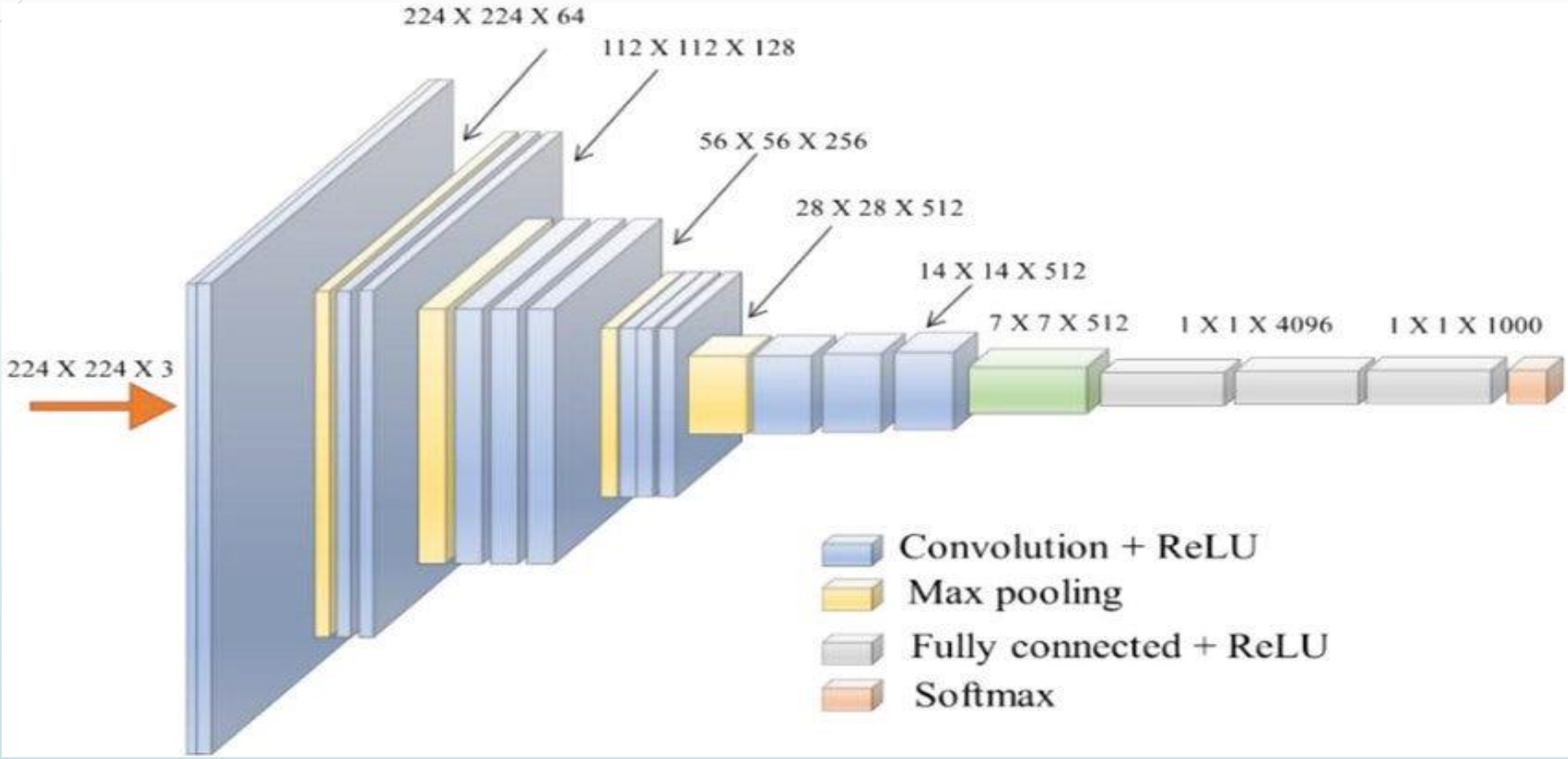
AlexNet Architecture



VGGNet (2014)

- Developed by the Visual Geometry Group at Oxford.
- Used small 3x3 filters throughout the network.
- Came in two variants: VGG-16 and VGG-19 (16 and 19 layers, respectively).
- Showed that depth improves performance.

VGGNet Architecture



https://ai.plainenglish.io/vggnet-with-tensorflow-transfer-learning-with-vgg16-included-7e5f6fa9479a

ResNet (2015)

- Introduced **residual connections** to enable training of very deep networks
- Used **skip connections** to bypass layers, addressing the vanishing gradient problem.
- Came in variants like ResNet-50, ResNet-101, and ResNet-152.
- Won the ImageNet competition in 2015 with a top-5 error rate of 3.57%

Advanced Architectures

1. **MobileNet**
2. **ShuffleNet**
3. **NASNet**
4. **Transformer-based Vision Models.**

Data Augmentation

What is Data Augmentation

- Data augmentation is a technique used to artificially increase the size and diversity of a training dataset by applying various transformations to the existing data.
- It helps improve the generalization of machine learning models, especially in scenarios where labeled data is scarce.

Why is it important

- Prevents overfitting: The model learns to generalize better than memorizing the training data by introducing variations.
- Improves model robustness: The model becomes more invariant to changes in input data (e.g., rotation, scaling, noise).

Data Augmentation Techniques

1. **Geometric Transformations**
2. **Color Space Transformations**
3. **Noise Injection**

Geometric Transformations

- **Rotation**: Rotate the image by a certain angle (e.g., 90°, 180°)
- **Translation**: Shift the image horizontally or vertically.
- **Scaling**: Zoom in or out of the image.
- **Flipping**: Flip the image horizontally or vertically.
- **Cropping**: Randomly crop a portion of the image.

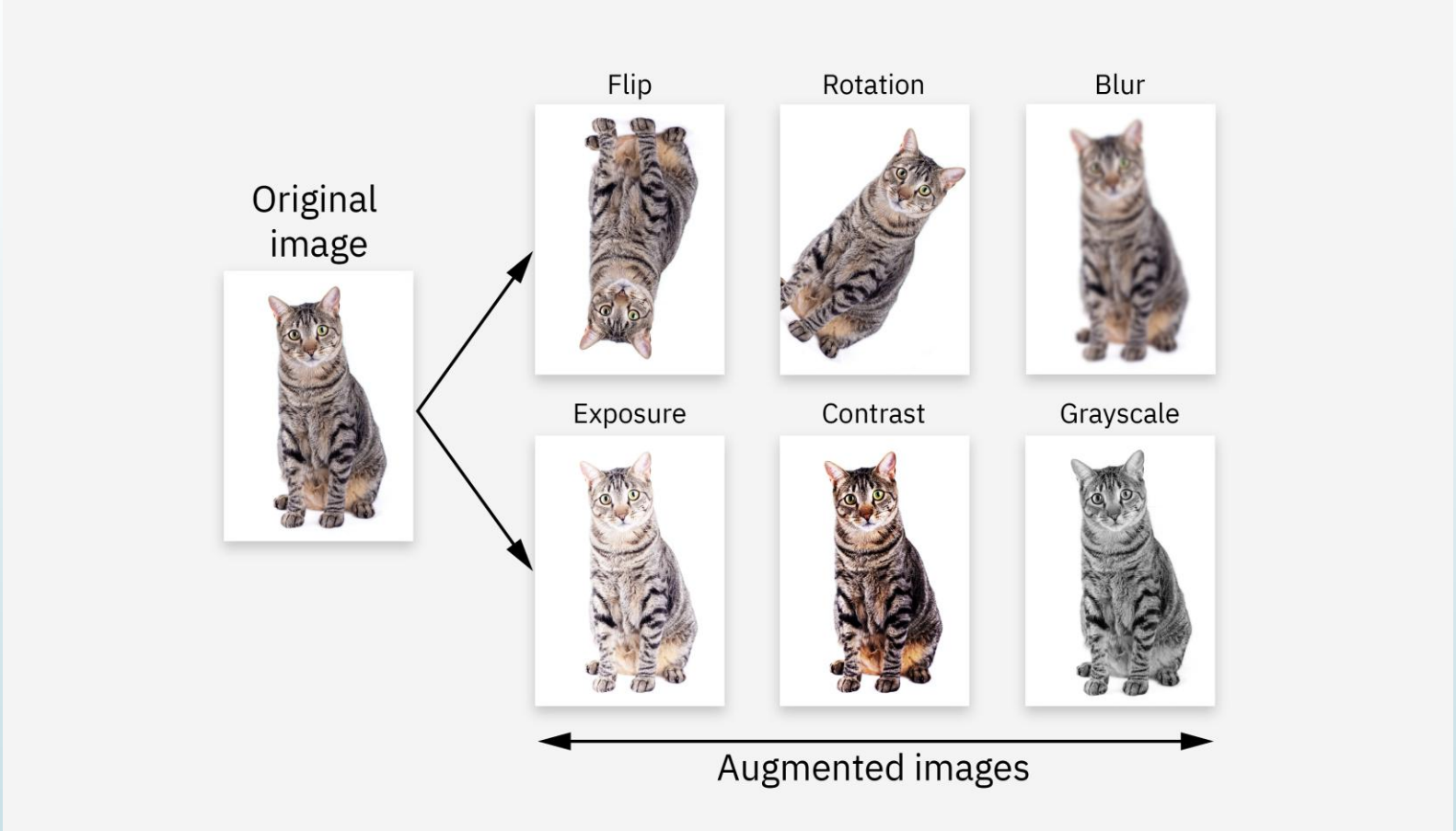
Geometric Transformations

https://www.researchgate.net/publication/364764514_Improved_Classification_Approach_for_Fruits_and_Vegetables_Freshness_Based_on_Deep_Learning/figures?lo=1&utm_source=google&utm_medium=organic

Color Space Transformations

- **Brightness Adjustment**: Increases or decreases the brightness.
- **Contrast Adjustment**: Modify the contrast of the image.
- **Color Jittering**: Randomly change the color balance.
- **Grayscale Conversion**: Convert the image to grayscale.

Color Space Transformations

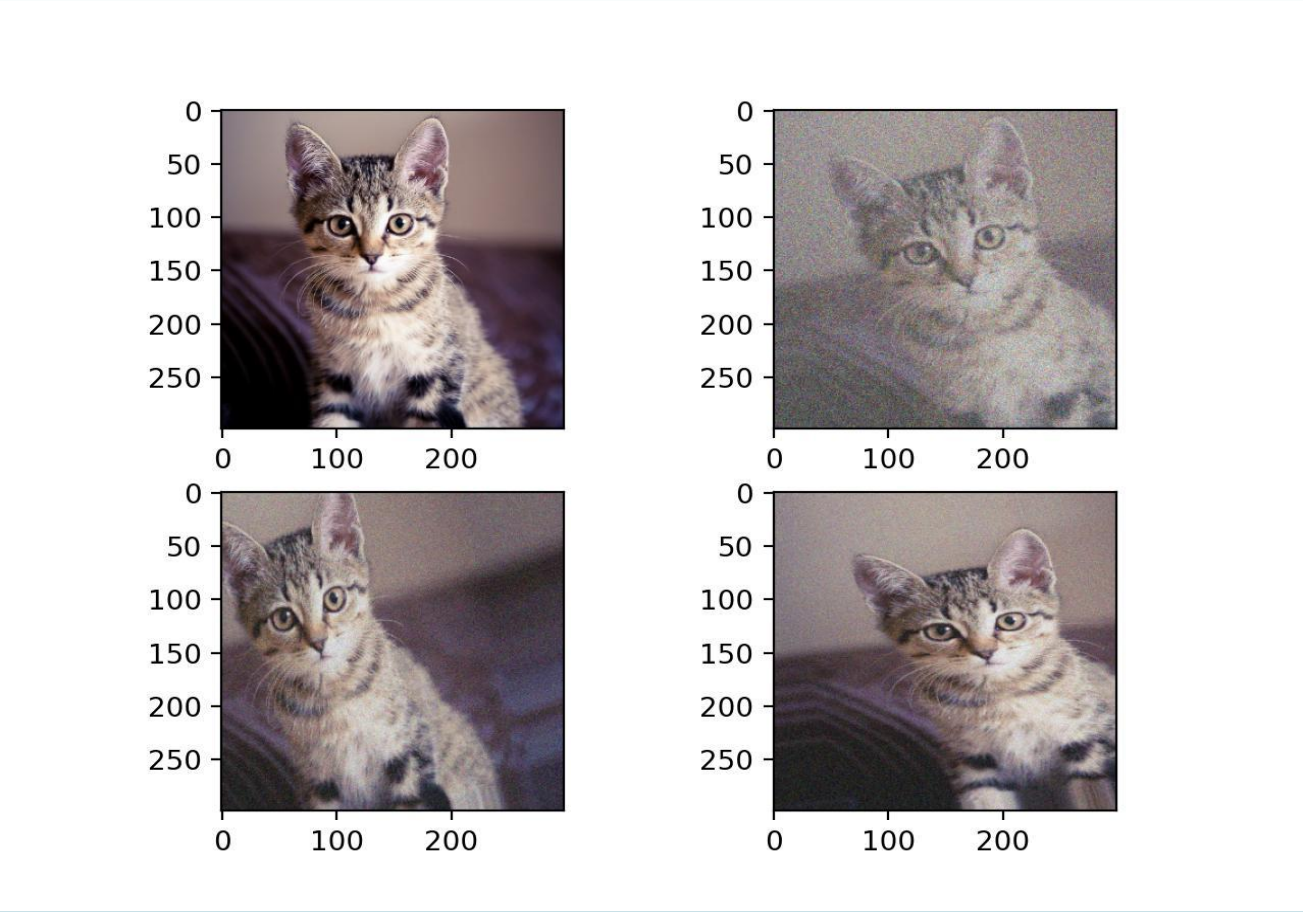


https://www.ibm.com/think/topics/data-augmentation

Noise Injection

- Add random noise (e.g., Gaussian noise) to the image to make the model robust to imperfections.

Noise Injection



https://stackoverflow.com/questions/43382045/keras-realtime-augmentation-adding-noise-and-contrast

Best Practices for Data Augmentation

- Preserve Label Integrity
- Avoid Over-augmentation
- Domain-Specific Augmentation

Challenges and Limitations

- Computational Cost
- Loss of Information
- Task-Specific Constraints

Applications of Data Augmentation

- **Computer Vision**: Image classification, object detection, segmentation.
- **Natural Language Processing**: Text classification, machine translation.
- **Speech Recognition**: Improving robustness to noise and accents.
- **Healthcare**: Augmenting medical images for better diagnosis.