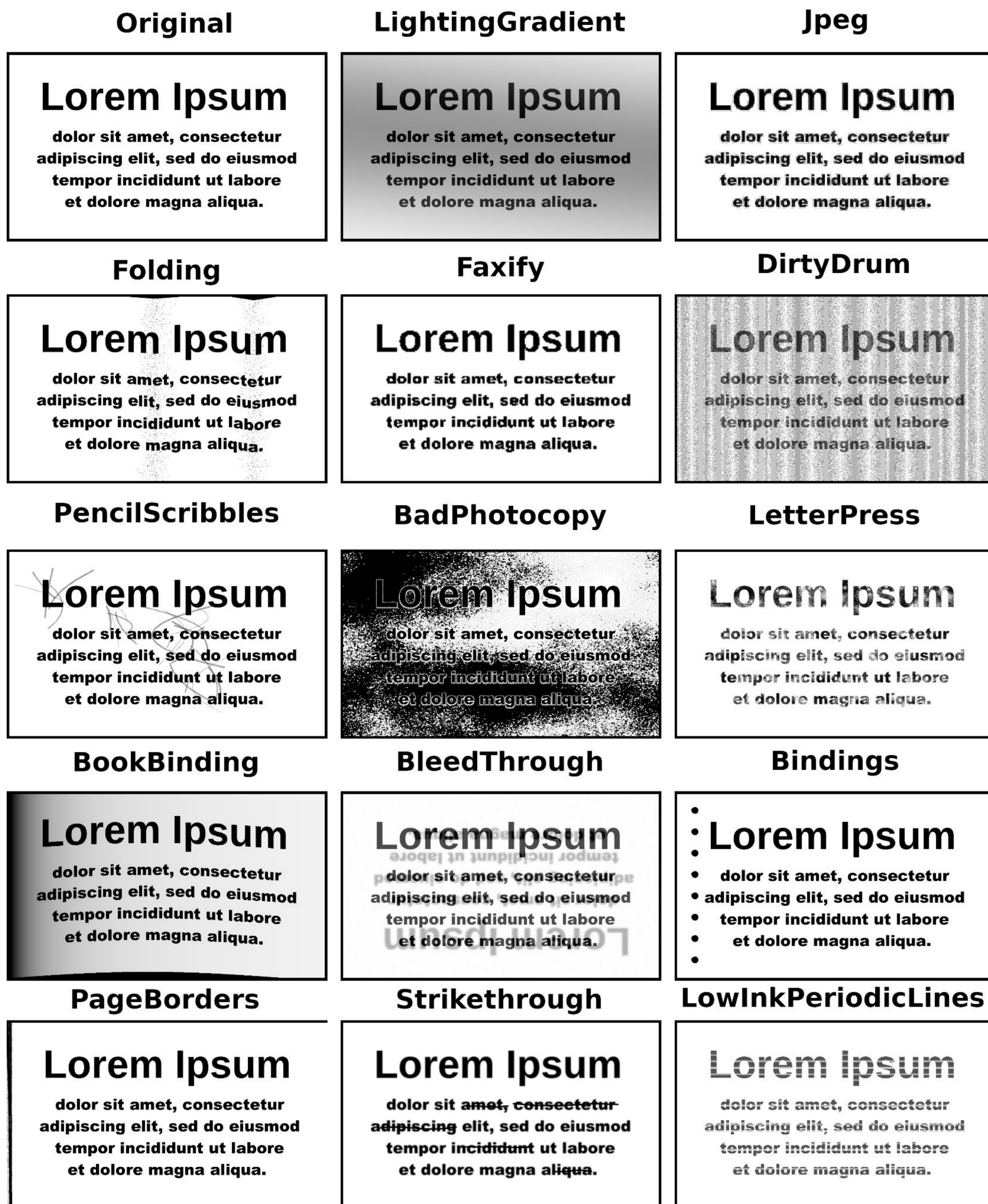


Augraphy: Data Augmentation for Document Images

Alexander Groleau, Stefan Larson, Kok Wei Chee, Jonathan Boarman



github.com/sparkfish/augraphy



This work introduces *Augraphy*, a Python library for data augmentation pipelines for document image analysis. Augraphy uses many different augmentation strategies to produce augmented versions of clean document images that appear as if they have been distorted due to noisy paper printing, faxing, scanning, or copy machine processes.

```
from augraphy import *; import cv2
pipeline = default_augraphy_pipeline()
im = cv2.imread('invoice.png')
data = pipeline.augment(im)
```

