**Deep Learning Models for IBCS Dashboard Analysis**

**Introduction**

The International Business Communication Standards (IBCS) ensure clear data visualizations (IBCS Association, 2023). Rule B, comprising SI 2.2 (Avoid Decorative Colors) and UN 4.1 (Unify Scenario Analyses), requires meaningful colors (e.g., green for positive variances, red for negative, blue for neutral) and prohibits decorative use. This project develops an AI solution to classify the cropped graph images, such as bar charts, from dashboards as compliant or non-compliant with Rule B, using data augmentation (e.g., rotating graphs) to expand the dataset. The initial goal is binary classification, with future expansion to whole dashboards. This report evaluates deep learning models and parameters for IBCS dashboard analysis, addressing the question, *Which models and parameters are suitable for IBCS dashboard analysis?*, and concludes with project alignment.

**Limitations of Traditional Machine Learning**

Traditional models like Support Vector Machines (SVM), Principal Component Analysis (PCA), and Random Forest are unsuitable for graph analysis. They require manual feature engineering, which is impractical for diverse visualizations. SVM struggles with raw images, PCA misses color semantics, and Random Forest is ineffective for visual tasks, necessitating deep learning (Goodfellow et al., 2016).

**Advantages of Deep Learning**

Deep learning models automatically extract features (e.g., colors), understand color meanings, adapt to varied graph types, and scale with augmented data. These strengths suit Rule B compliance detection, identifying decorative colors or incorrect variance color coding (LeCun et al., 2015).

**Suitable Deep Learning Models**

**Convolutional Neural Networks (CNNs)**  
*Description*: CNNs detect patterns like colors for classification (Krizhevsky et al., 2017).  
*Suitability*: Models like ResNet50 or EfficientNet classify graphs as compliant or non-compliant, fitting the project’s detection goal.  
*Advantages*: Simple, efficient for small datasets.

*Limitations*: Whole-image analysis only.

**Vision Transformers (ViT)**  
*Description*: ViT uses transformers to process images, capturing complex patterns (Dosovitskiy et al., 2020).  
*Suitability*: Effective for tricky graphs, using small-dataset techniques (Lee et al., 2021).  
*Advantages*: Handles complex designs.  
*Limitations*: Needs stronger computers.

**Object Detection Models**  
*Description*: Models like YOLO identify graph elements (e.g., bars) (Redmon et al., 2016).  
*Suitability*: Classify graphs by checking element colors, aiding detection.  
*Advantages*: Precise analysis.  
*Limitations*: Requires labeled element data.

**Multimodal Models**  
*Description*: Combine image analysis with text via Optical Character Recognition (OCR) and natural language processing (NLP).  
*Suitability*: OCR extracts labels (e.g., “Actual vs. Plan”) to understand color context.  
*Advantages*: Enhances accuracy.  
*Limitations*: Complex, better for later phases.

**Key Parameters**

**CNNs and ViT**

* *Learning Rate*: 0.0001–0.001 for stable training.
* *Batch Size*: 16–64 graphs per training step.
* *Epochs*: 20–100, with early stopping if no improvement.
* *Optimizer*: Adam for smooth adjustments.
* *Loss Function*: Cross-entropy for yes/no classification.
* *Data Augmentation*: Rotate, flip, scale graphs to increase data.

**Object Detection**

* *Backbone*: ResNet50 for feature detection.
* *IoU (Intersection over Union) Threshold*: 0.5 to filter overlapping detections.
* *Batch Size*: 8–16.

**Multimodal**

* *OCR*: Tesseract for text extraction.
* *NLP*: BERT for text understanding.

**Model Summary**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **Task** | **Key Parameters** | **Advantages** |
| CNNs | Classification | Learning rate: 0.0001–0.001, Batch size: 16–64 | Simple, efficient |
| ViT | Classification | Patch size: 16x16, Layers: 12–24 | Handles complex graphs |
| YOLO | Detection | IoU: 0.5, Batch size: 8–16 | Precise element analysis |
| Multimodal | Contextual Analysis | OCR settings, BERT tuning | Adds color context |

**Project Alignment**

The project analyzes around 200 graphs to determine Rule B compliance, using data augmentation (e.g., rotating graphs) to increase training data. The focus is classifying graphs as compliant or non-compliant, with plans to analyze whole dashboards later. CNNs (e.g., ResNet50) are ideal for their simplicity, fine-tuning on the dataset (80% training, 20% validation). ViT suits complex graphs, object detection checks specific elements, and multimodal models support future phases by reading labels for color context. Parameters like learning rate (0.0001–0.001) ensure effective training.

**Conclusion**

CNNs, ViT, object detection, and multimodal models are suitable for IBCS dashboard analysis, addressing Rule B’s color requirements. CNNs offer simplicity, ViT handles complexity, object detection targets elements, and multimodal models add context. Parameters like learning rate and batch size are critical. For the project’s 200 graphs, CNNs are recommended for binary classification with data augmentation, with ViT and object detection as alternatives, and multimodal models for future dashboard analysis, aligning with the phased approach.

**References**

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszoreit, J., & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv*. <https://arxiv.org/abs/2010.11929>

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. <https://www.deeplearningbook.org/>

IBCS Association. (2023). *IBCS standards version 1.2*. <https://www.ibcs.com/standards/>

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84–90. <https://doi.org/10.1145/3065386>

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. <https://doi.org/10.1038/nature14539>

Lee, S.-H., Kim, S., & Lee, S. (2021). Vision transformer for small-size datasets. *arXiv*. <https://arxiv.org/abs/2112.13492>

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *arXiv*. <https://arxiv.org/abs/1506.02640>