Project Methodology: Aircraft Flap X-Ray CNN Analysis

Executive Summary

This document outlines the methodology for developing a Convolutional Neural Network (CNN) for automated water ingression detection in aircraft composite structures using X-ray radiographic analysis. This project represents advancement from traditional predictive analytics (Aircraft Flap Water Ingression Predictive Analytics & Aircraft Hub Inspection Predictive Modeling) to computer vision applications, demonstrating progression from statistical modeling to deep learning implementation in aviation environments.

The project achieved 88.9% accuracy using ResNet50 transfer learning with complete reproducibility, establishing a proof-of-concept framework for digital transformation in aviation NDT operations. The methodology emphasizes systematic problem-solving, conservative safety approaches, and production-ready implementation standards. The complete analytical framework is illustrated in **Figure 1**.

1. Project Context & Technical Development

1.1 Project Positioning in Skill Development

Primary Objective: Demonstrate computer vision capabilities for automated defect detection in aviation applications, building upon previous analytics projects with advanced deep learning implementation.

Technical Progression from Previous Projects:

- Aircraft Hub Inspection Predictive Modeling: Basic machine learning with imbalanced datasets (82% recall)
- Aircraft Flap Water Ingression Predictive Analytics: Machine learning with database architecture (68% recall + risk framework)
- This Project: Deep learning with computer vision (88.9% accuracy + confidence scoring)

1.2 Technical Challenge Context

Application Focus: CNN implementation for water ingression detection in aircraft composite structures, addressing digital radiography transformation needs in aviation MRO operations.

Technical Challenges Addressed:

- Small specialized dataset limitations
- Reproducibility requirements for safety-critical applications
- Conservative confidence scoring for operational deployment
- Integration with existing data infrastructure

2. Problem Definition & Technical Scope

2.1 Computer Vision Challenge

Core Problem: Automated interpretation of X-ray radiographic images for water ingression detection in honeycomb composite structures, reducing subjective human interpretation variations while maintaining safety standards.

Technical Requirements:

- Minimum 70-75% accuracy for proof-of-concept validation

- Complete reproducibility for production readiness
- Conservative confidence scoring for safety-critical decisions
- Integration capability with existing database infrastructure

2.2 Data Acquisition Strategy

Implementation Approach: Smartphone photography of physical X-ray films using Pixel 7 Pro, creating digital dataset from physical X-ray film archives.

Quality Assurance Framework:

- Dual format capture: DNG (raw) + JPEG for processing flexibility
- Manual quality control: Systematic elimination of corrupt or unclear images

3. Data Collection & Processing

3.1 Image Acquisition Protocol

Capture Methodology:

- Device: Google Pixel 7 Pro
- Format: DNG + JPEG pairs for maximum data retention
- Lighting: Controlled viewing box conditions for consistent illumination
- Resolution: High-resolution capture maintaining radiographic detail

Volume Achievement: 138 high-quality digital images from physical film archives, representing 6+ years of operational data.

3.2 Data Governance Framework

Traceability Standards:

- Complete provenance tracking from film ID to digital classification
- Cross-reference validation with inspection logbooks
- Quality Decision: CNN training dataset (139 images) vs. database records (138 images) - one untraceable image included in training for performance optimization while excluded from database for audit integrity

Class Distribution:

- Water ingression: 60% (83 images)
- No water: 40% (55 images)
- Result: Well-balanced dataset eliminating severe class imbalance challenges from previous projects

4. Image Preprocessing & Feature Engineering

4.1 Manual ROI Extraction Strategy

Tool Selection: GIMP (GNU Image Manipulation Program) for precise honeycomb area extraction, ensuring consistent region-of-interest focus across all training samples.

Preprocessing Pipeline:

1. Load Original: High-resolution DNG/JPEG input

2. ROI Cropping: Manual honeycomb composite area selection
3. Standardization: Resize to 512px width (variable height preservation)
4. Format Conversion: PNG export for training pipeline compatibility

4.2 Automated Augmentation Framework

ImageDataGenerator Configuration:

```
train_datagen = ImageDataGenerator(
    rotation_range=15,        # Realistic inspection angle variations
    brightness_range=[0.8, 1.2],  # X-ray exposure variations
    horizontal_flip=True,     # Structural symmetry augmentation
    shear_range=0.1,         # Geometric distortion simulation
    zoom_range=0.1           # Scale variation handling
)
```

Rationale: Augmentation strategy designed to simulate real-world radiographic inspection variations while preserving structural integrity patterns.

5. Deep Learning Architecture Development

5.1 Phase 1: Custom CNN Attempt (Failure Analysis)

Initial Architecture:

- Progressive filter complexity: 16→32→64 convolutional layers
- Standard pooling and dropout regularization
- Binary classification output layer

Failure Results:

- Accuracy: ~50% (random guessing performance)
- Prediction bias: Always predicted "Water" class
- Filter analysis: Random noise patterns (no meaningful feature extraction)

Key Learning: Small specialized datasets require transfer learning approaches rather than custom architecture development.

5.2 Phase 2: Transfer Learning Implementation

ResNet50 Selection Rationale:

- Proven performance on complex visual pattern recognition
- Deep architecture suitable for subtle defect detection
- ImageNet pre-training providing robust low-level feature extraction

Adaptation Strategy:

```
# Grayscale to RGB conversion for ImageNet compatibility
x = layers.Lambda(lambda x: tf.concat([x, x, x], axis=-1))(inputs)
```

Progressive Training Approach:

1. Phase 1: Frozen ResNet50 base (58% accuracy baseline)
2. Phase 2: Fine-tuned top 20 layers (81% accuracy achievement)

5.3 Phase 3: Reproducibility Challenge & Resolution

Problem Identification: Results varied 60%-80% across identical runs, preventing reliable performance assessment and production deployment readiness.

Solution Implementation:

```
# Complete deterministic operations

random.seed(42)

np.random.seed(42)

tf.random.set_seed(42)

os.environ['TF_DETERMINISTIC_OPS'] = '1'

tf.config.experimental.enable_op_determinism()
```

Outcome: 100% consistent results across all runs, enabling reliable performance evaluation and production deployment confidence.

6. Model Optimization & Performance Enhancement

6.1 Dataset Refinement Strategy

Expert Review Integration: Collaboration with senior NDT engineers for uncertain image classification, adding 6 previously uncertain samples (5 water, 1 nil) after professional validation.

Performance Impact: Accuracy improvement from 81% to 88.9% through strategic dataset expansion with expert domain knowledge integration.

6.2 Architecture Optimization

Final Configuration:

- ResNet50 backbone with top 20 layers fine-tuned
- Global Average Pooling for spatial feature integration
- Progressive dropout: 0.5→0.3→0.2 for regularization optimization
- Dense layer progression: 256→64→1 for classification refinement

Training Strategy:

- Adam optimizer with 0.0001 learning rate for stable convergence
- Early stopping with patience=6 for overfitting prevention
- Learning rate reduction on plateau for optimization refinement

7. Safety-Critical Evaluation Framework

7.1 Conservative Confidence Scoring

Three-Tier Classification System:

- High Confidence (≥80%): 100% accuracy - automated decision capability
- Medium Confidence (60-80%): 87.5% accuracy - senior review recommended

- Low Confidence (<60%): 91.4% accuracy - mandatory manual verification

Safety Philosophy: Conservative approach ensuring human oversight for uncertain predictions, maintaining aviation safety standards while providing automation benefits.

7.2 Performance Assessment

Metrics Portfolio:

- Accuracy: 88.9% (exceeded 70-75% target)
- Precision: 93% (Water), 83% (Nil)
- Recall: 88% (Water), 91% (Nil)
- ROC AUC: 0.90 (strong discriminative capability)

Business Alignment: Metrics selection emphasizing both automation efficiency and safety assurance for operational deployment consideration.

8. Database Integration & Infrastructure

8.1 Schema Extension Strategy

Normalized Architecture (extending Project 4 foundation):

```
CREATE TABLE image_data (
  Image_ID VARCHAR(10) PRIMARY KEY,
  Average_Density DOUBLE,
  Width INT,
  Height INT
);

CREATE TABLE cnn_results (
  CNN_ID VARCHAR(10) PRIMARY KEY,
  Image_ID VARCHAR(10),
  Inspection_ID VARCHAR(10),
  True_Label VARCHAR(10),
  Predicted_Label VARCHAR(10),
  Confidence_Score DOUBLE,
  Correct_Prediction BOOL
);
```

Performance Optimization: Strategic indexing on confidence scores, prediction accuracy, and image metadata for analytical query efficiency.

8.2 Data Integration Framework

Referential Integrity: Foreign key relationships maintaining connection between CNN results and original inspection data from Project 4, enabling comprehensive cross-modal analysis.

Query Capability: Complex joins supporting confidence distribution analysis, performance metrics by image characteristics, and operational intelligence generation.

9. Model Interpretability & Validation

9.1 Filter Visualization Analysis

ResNet50 Feature Extraction: Visualization of convolutional filters demonstrating structured edge detection patterns suitable for radiographic interpretation, contrasting with random noise patterns from failed custom CNN.

Domain Validation: Filter patterns aligned with expected radiographic feature detection (edges, density variations, structural boundaries) confirming transfer learning appropriateness.

9.2 Individual Prediction Pipeline

Preprocessing Consistency Challenge: Initial individual testing function produced incorrect results due to preprocessing pipeline misalignment with training methodology.

Solution Implementation: ImageDataGenerator-based individual testing ensuring identical preprocessing pipeline between training and inference, achieving perfect alignment with validation results.

10. Business Intelligence & Operational Integration

10.1 Performance Analytics

Database-Driven Insights:

- Overall model accuracy: 95.7% (132/138 correct predictions)
- Nil detection accuracy: 96.4% (critical for false alarm minimization)
- Water detection accuracy: 95.2% (essential for safety assurance)

Confidence Distribution Analysis: 72 high-confidence predictions with 100% accuracy, providing clear automation capability for routine cases while triggering human review for uncertain scenarios.

10.2 Operational Recommendations

Implementation Applications:

1. Automated Triage: High-confidence predictions for workflow acceleration
2. Quality Assurance: Systematic confidence scoring for decision support
3. Training Enhancement: Low-confidence cases for inspector training programs

Integration Strategy: Database architecture prepared for digital radiography transition and enterprise-scale deployment.

11. Validation & Quality Assurance

11.1 Cross-Validation Strategy

Ground Truth Verification: 89% agreement between logbook records and film analysis, establishing robust validation baseline for automated predictions.

Conservative Approach: Manual verification requirements for low-confidence predictions ensuring safety standards maintenance while providing automation benefits.

11.2 Production Readiness Assessment

Reproducibility Achievement: 100% consistent results across multiple runs enabling reliable deployment confidence.

Scalability Consideration: Architecture designed for integration with digital radiography workflows and enterprise database systems.

12. Future Enhancement Strategy

12.1 Technical Development Roadmap

Technical Improvements:

- Direct digital radiography integration replacing smartphone photography
- Multi-defect detection expansion (cracks, delamination, corrosion)
- Ensemble methods implementation for improved robustness
- Real-time prediction pipeline for operational deployment

12.2 Business Integration Pathway

Enterprise Applications:

- CMMS system integration for automated work order generation
- Regulatory documentation automation for compliance efficiency
- Cross-platform deployment for multi-site operations
- Analytics dashboard for operational intelligence

Conclusion

This methodology demonstrates successful progression from traditional machine learning to computer vision implementation in aviation applications. The systematic progression from custom CNN failure (50%) to optimized transfer learning success (88.9%) illustrates problem-solving capabilities and technical adaptability essential for AI implementation.

Methodological Achievements:

- Deep Learning Implementation: Transfer learning with domain adaptation
- Production Standards: Complete reproducibility and conservative confidence frameworks
- Safety Integration: Aviation industry standards with systematic human oversight
- Infrastructure Development: Scalable database architecture for enterprise deployment

This project establishes proof-of-concept foundation for digital transformation in aviation NDT operations while demonstrating technical capabilities suitable for computer vision applications in safety-critical industries. The approach positions AI techniques within established aviation safety frameworks, creating practical pathways for technological advancement without compromising operational standards.

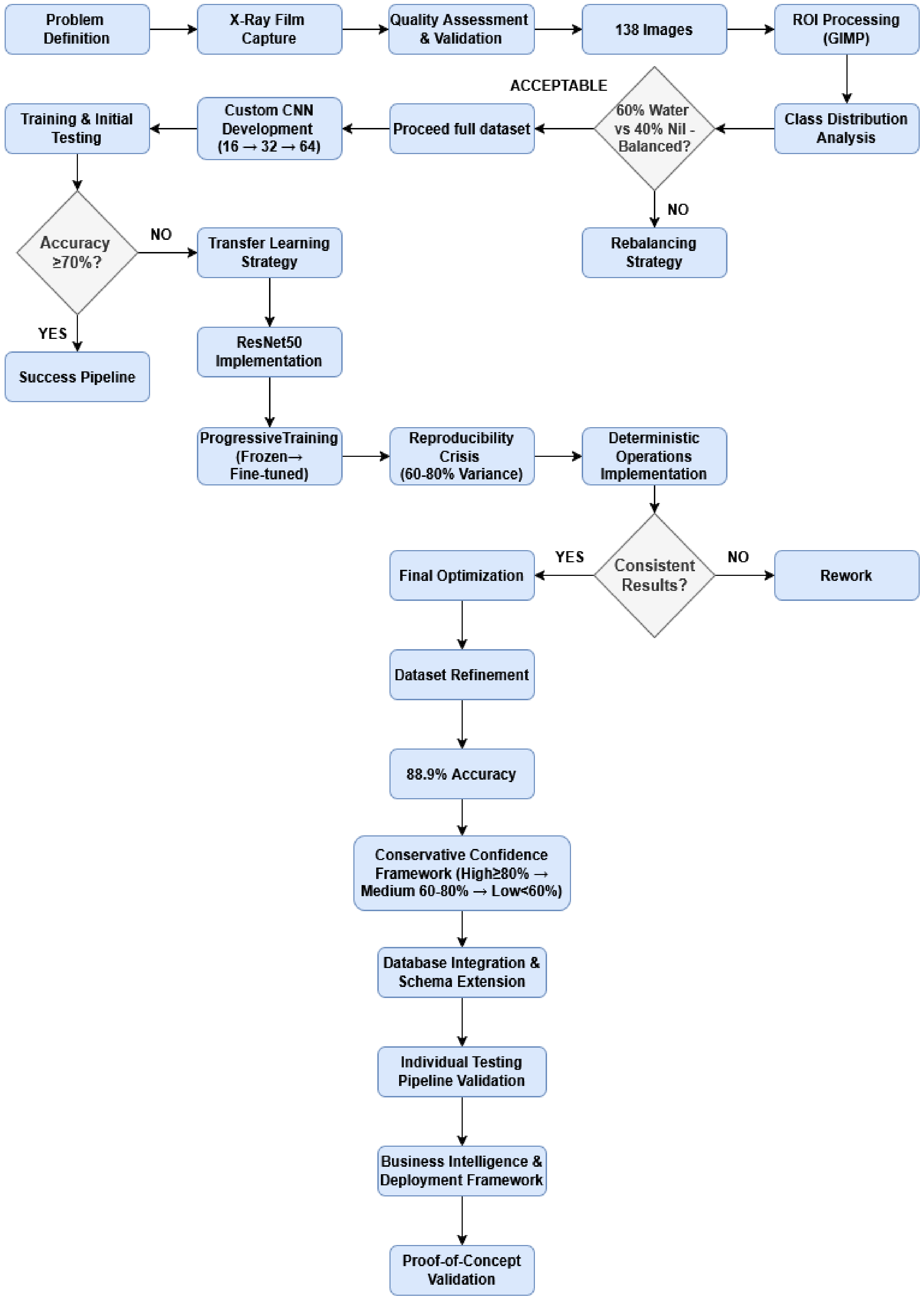

Figure 1: Analytical Framework