■ Python Project Code Export

```
- [PY] class_count.py
- [PY] clean_pngs.py
- [PY] config.py
- [PY] convert_polygons.py
- [PY] dataset.py
- [PY] demo.py
- [PY] evaluate.py
- [PY] evaluation\metrics.py
- [PY] infer.py
- [PY] inference\engine.py
- [PY] models\dvx.py
- [PY] models\encoder.py
- [PY] models\extrusion.py
- [PY] models\heads.py
- [PY] models\model.py
- [PY] param.py
- [PY] setup.py
- [PY] train.py
- [PY] training\losses.py
- [PY] training\trainer.py
- [PY] utils\visualization.py
- [PY] vortex.py
```

■ File: class_count.py

==

```python
import cv2, numpy as np, glob
from collections import Counter

all_classes = set()
class_counts = Counter()

# Go through all mask images
for mask_file in glob.glob("./data/floorplans/train/*/mask.png"):
    mask = cv2.imread(mask_file, cv2.IMREAD_GRAYSCALE)
    unique, counts = np.unique(mask, return_counts=True)
    for u, c in zip(unique, counts):
        class_counts[u] += c
    all_classes.update(unique)

# Total pixels
total_pixels = sum(class_counts.values())

print("Classes found in dataset:", sorted(all_classes))
print("\nPixel distribution per class:")
for cls in sorted(class_counts.keys()):
    percentage = (class_counts[cls] / total_pixels) * 100
    print(f"Class {cls}: {class_counts[cls]} pixels ({percentage:.2f}%)")
```

■ File: clean_pngs.py

==

```python
from PIL import Image
import os
import shutil

data_dir = r"data/floorplans"

def safe_clean_image(path):
    """Safely clean image by only removing ICC profiles, preserving all pixel data"""
    try:
        # Create backup first (optional safety measure)
        backup_path = path + ".backup"

        with Image.open(path) as img:
            # Check if image is already clean
            if 'icc_profile' not in img.info:
                print(f"Already clean: {path}")
                return True

            # Create backup
            shutil.copy2(path, backup_path)

            # Method 1: Just strip ICC profile while preserving everything else
            img_data = img.copy()

            # Remove only the problematic ICC profile
            if 'icc_profile' in img_data.info:
                del img_data.info['icc_profile']

            # Save with same format and quality, just without ICC profile
            img_data.save(path, format="PNG", optimize=False)  # No optimization to preserve exact pixels

            # Remove backup if successful
            os.remove(backup_path)

        print(f"Cleaned ICC profile from: {path}")
        return True

    except Exception as e:
        # Restore backup if it exists
        backup_path = path + ".backup"
        if os.path.exists(backup_path):
            shutil.move(backup_path, path)
            print(f"Restored backup for: {path}")

```

```
        print(f"Failed {path}: {e}")
        return False

def verify_image_integrity(path):
    """Verify image can still be loaded properly after cleaning"""
    try:
        with Image.open(path) as img:
            # Try to access pixel data to ensure image is valid
            _ = img.size
            _ = img.mode
            # Try to load a small sample of pixel data
            _ = img.getpixel((0, 0))
        return True
    except Exception as e:
        print(f"WARNING: Image integrity check failed for {path}: {e}")
        return False

# Process only image.png files
processed_files = []
cleaned_count = 0
failed_count = 0
already_clean = 0

print("Starting safe ICC profile removal for dataset...")
print("This preserves all pixel data and only removes problematic metadata.")

for root, _, files in os.walk(data_dir):
    for f in files:
        if f == "image.png":  # Only process image.png files
            path = os.path.join(root, f)
            processed_files.append(path)

            # Check if already clean
            try:
                with Image.open(path) as img:
                    if 'icc_profile' not in img.info:
                        already_clean += 1
                        continue
            except:
                pass

            if safe_clean_image(path):
                # Verify integrity after cleaning
                if verify_image_integrity(path):
                    cleaned_count += 1
                else:
                    failed_count += 1
            else:
                failed_count += 1

print(f"\n" + "="*50)
print(f"DATASET CLEANING SUMMARY")
print(f"="*50)
print(f"Total image.png files found: {len(processed_files)}")
print(f"Already clean (no ICC profile): {already_clean}")
print(f"Successfully cleaned: {cleaned_count}")
print(f"Failed to clean: {failed_count}")
print(f"Total files processed: {already_clean + cleaned_count + failed_count}")

if failed_count > 0:
    print(f"\nWARNING: {failed_count} files couldn't be cleaned.")
    print(f"Check these files manually - they may be corrupted.")

print(f"\nDataset should now be ready for training without libpng warnings!")

# Optional: Test load a few random images to verify dataset integrity
print(f"\nTesting random samples for integrity...")
import random
test_files = random.sample(processed_files, min(5, len(processed_files)))
for test_path in test_files:
    if verify_image_integrity(test_path):
        print(f"? {test_path}")
    else:
```

```python
        print(f"? {test_path} - POTENTIAL ISSUE")
```

■ File: config.py

```python
"""
Configuration settings for the Neural-Geometric 3D Model Generator
Enhanced with dynamic curriculum and adaptive training strategies
"""
from dataclasses import dataclass
from typing import Tuple, Dict, Any, Optional, List
import torch


@dataclass
class DataConfig:
    """Data-related configuration"""
    data_dir: str = "./data/floorplans"
    image_size: Tuple[int, int] = (256, 256)   # keep full resolution for accuracy
    voxel_size: int = 64
    batch_size: int = 4                        # balance speed & memory
    num_workers: int = 8                       # faster dataloader (tune per CPU)
    augment: bool = True


@dataclass
class ModelConfig:
    """Model architecture configuration optimized for high accuracy"""
    input_channels: int = 3
    num_classes: int = 5
    feature_dim: int = 512     # reduced from 768 ? faster while keeping strong accuracy
    num_attributes: int = 6
    voxel_size: int = 64
    max_polygons: int = 20     # enough for complex layouts
    max_points: int = 50       # good detail without huge cost
    dropout: float = 0.05
    use_attention: bool = True
    use_deep_supervision: bool = True

    # Auxiliary heads for novel training strategies
    use_latent_consistency: bool = True
    use_graph_constraints: bool = True
    latent_embedding_dim: int = 256


@dataclass
class CurriculumConfig:
    """Dynamic curriculum learning configuration"""
    # Adaptive stage transitioning
    use_dynamic_curriculum: bool = True
    stage_switch_patience: int = 5
    min_improvement_threshold: float = 0.001
    plateau_detection_window: int = 3

    # GradNorm / gradient tracking
    gradient_norm_window: int = 100

    # Objectives for multi-objective optimization
    objectives: Optional[List[str]] = None

    # Topology-aware scheduling
    topology_schedule: str = "progressive"   # "progressive", "linear_ramp", "exponential"
    topology_start_weight: float = 0.1
    topology_end_weight: float = 1.0
    topology_ramp_epochs: int = 20

    # config.py (snippet ? add into the existing config class/dict)
    # Mixed precision and training conveniences
    use_mixed_precision = True            # enable AMP
    cache_in_memory = False               # set True if host RAM can hold dataset
    accumulation_steps = 1                # effective batch size multiplier
    dvx_step_freq = 1                     # run DVX refinement every N steps (1 = every step)
```

```
    persistent_workers = True            # DataLoader persistent workers
    prefetch_factor = 4                  # DataLoader prefetch
    num_workers = 8                      # default num workers for DataLoader (tune by CPU)
    # Progressive resolution settings (example)
    voxel_size_stage = { "stage1": 32, "stage2": 32, "stage3": 64 }  # voxel sizes per stage
    image_size_stage = { "stage1": (128,128), "stage2": (192,192), "stage3": (256,256)}


    # Loss component scheduling
    loss_schedule: Dict[str, str] = None

    # Multi-objective optimization (GradNorm)
    use_gradnorm: bool = True
    gradnorm_alpha: float = 0.12
    gradnorm_update_freq: int = 5

    # Graph constraint scheduling
    graph_weight_schedule: str = "delayed_ramp"
    graph_start_epoch: int = 15
    graph_end_weight: float = 0.5

    def __post_init__(self):
        # Provide default loss schedule if not set
        if self.loss_schedule is None:
            self.loss_schedule = {
                "segmentation": "static",
                "dice": "static",
                "sdf": "early_decay",
                "attributes": "static",
                "polygon": "staged_ramp",
                "voxel": "late_ramp",
                "topology": "progressive",
                "latent_consistency": "mid_ramp",
                "graph": "delayed_ramp",
            }

        # Default objectives used by GradNorm / trainer monitoring
        if self.objectives is None:
            self.objectives = [
                "segmentation",
                "dice",
                "sdf",
                "attributes",
                "polygon",
                "voxel",
                "topology",
                "latent_consistency",
                "graph",
            ]


@dataclass
class TrainingConfig:
    """Training configuration with adaptive strategies"""
    device: str = "cuda" if torch.cuda.is_available() else "cpu"

    # Dynamic epoch limits (maxima; curriculum may switch earlier)
    max_stage1_epochs: int = 40
    max_stage2_epochs: int = 25
    max_stage3_epochs: int = 60

    # Minimum epochs per stage (avoid switching too early)
    min_stage1_epochs: int = 8
    min_stage2_epochs: int = 5
    min_stage3_epochs: int = 12

    # Learning rates (per stage)
    stage1_lr: float = 1e-5  # was 3e-4
    stage1_weight_decay: float = 1e-5

    stage2_lr: float = 5e-6  # was 1e-4
    stage2_weight_decay: float = 1e-5

```

```python
    stage3_lr: float = 1e-6  # was 5e-5
    stage3_weight_decay: float = 1e-5

    # Advanced training techniques
    use_mixed_precision: bool = True
    use_cosine_restarts: bool = True
    warmup_epochs: int = 5
    grad_clip_norm: float = 0.5

    # Gradient monitoring for dynamic weighting
    track_gradient_norms: bool = True
    gradient_norm_window: int = 10  # rolling window for gradient tracking

    # Checkpointing
    checkpoint_freq: int = 1

    # Curriculum configuration
    curriculum: CurriculumConfig = None

    def __post_init__(self):
        if self.curriculum is None:
            self.curriculum = CurriculumConfig()


@dataclass
class LossConfig:
    """Loss function weights (will be dynamically adjusted during training)"""
    # Base weights (starting values)
    seg_weight: float = 1.0
    dice_weight: float = 1.0
    sdf_weight: float = 0.5
    attr_weight: float = 1.0
    polygon_weight: float = 1.0
    voxel_weight: float = 1.0
    topology_weight: float = 0.1  # start low, ramp up

    # New loss components
    latent_consistency_weight: float = 0.5
    graph_constraint_weight: float = 0.3

    # Dynamic weighting parameters
    enable_dynamic_weighting: bool = True
    weight_update_freq: int = 10
    weight_momentum: float = 0.9


@dataclass
class InferenceConfig:
    """Inference configuration"""
    model_path: str = "final_model.pth"
    test_images_dir: str = "./data/test_images"
    output_dir: str = "./outputs"
    export_intermediate: bool = True
    polygon_threshold: float = 0.5


# Curriculum stage transition logic
class StageTransitionCriteria:
    """Defines criteria for automatic stage transitions"""

    @staticmethod
    def should_transition_from_stage1(train_losses, val_losses, config: CurriculumConfig) -> bool:
        """Check if should transition from Stage 1 to Stage 2"""
        if len(val_losses) < config.plateau_detection_window:
            return False

        # Check for plateau in segmentation + dice losses
        recent_losses = val_losses[-config.plateau_detection_window:]
        if len(recent_losses) < 2:
            return False

        # Calculate improvement rate
        old_avg = sum(recent_losses[:len(recent_losses)//2]) / (len(recent_losses)//2)
```

```
        new_avg = sum(recent_losses[len(recent_losses)//2:]) / (len(recent_losses) - len(recent_losses)//2)

        improvement_rate = (old_avg - new_avg) / (old_avg + 1e-8)
        return improvement_rate < config.min_improvement_threshold

    @staticmethod
    def should_transition_from_stage2(polygon_losses, config: CurriculumConfig) -> bool:
        """Check if should transition from Stage 2 to Stage 3"""
        if len(polygon_losses) < config.plateau_detection_window:
            return False

        # Check polygon loss plateau
        recent_losses = polygon_losses[-config.plateau_detection_window:]
        if len(recent_losses) < 2:
            return False

        old_avg = sum(recent_losses[:len(recent_losses)//2]) / (len(recent_losses)//2)
        new_avg = sum(recent_losses[len(recent_losses)//2:]) / (len(recent_losses) - len(recent_losses)//2)

        improvement_rate = (old_avg - new_avg) / (old_avg + 1e-8)
        return improvement_rate < config.min_improvement_threshold


# Default configurations (import these in your trainer)
DEFAULT_DATA_CONFIG = DataConfig()
DEFAULT_MODEL_CONFIG = ModelConfig()
DEFAULT_TRAINING_CONFIG = TrainingConfig()
DEFAULT_LOSS_CONFIG = LossConfig()
DEFAULT_INFERENCE_CONFIG = InferenceConfig()
```

■ File: convert_polygons.py

==

```
import json
from pathlib import Path

def flatten_polygon_file(file_path: Path):
    with open(file_path, "r") as f:
        data = json.load(f)

    flat_list = []

    # If file already flat, skip
    if isinstance(data, list):
        print(f"[SKIP] Already flat: {file_path}")
        return

    # Otherwise, flatten categories
    for category in ["walls", "doors", "windows", "floors", "fixtures"]:
        if category in data and isinstance(data[category], list):
            for idx, poly in enumerate(data[category]):
                flat_list.append({
                    "id": idx,
                    "type": category[:-1],  # "walls" -> "wall"
                    "points": poly.get("points", []),
                    "area": poly.get("area", None)
                })

    # Save back in flat format
    with open(file_path, "w") as f:
        json.dump(flat_list, f, indent=2)

    print(f"[OK] Converted: {file_path}")


def batch_convert(root_dir="data/floorplans"):
    root = Path(root_dir)
    for polygon_file in root.rglob("polygon.json"):
        flatten_polygon_file(polygon_file)


if __name__ == "__main__":
```

```
    # Change path if needed
    batch_convert("data/floorplans")
```

■ File: dataset.py

==

```
"""
Dataset classes for the Neural-Geometric 3D Model Generator
Enhanced with in-memory caching for faster training
"""

import cv2
import json
import numpy as np
import torch
from torch.utils.data import Dataset
from pathlib import Path
from typing import Dict, List, Tuple, Optional, Union
import time

from config import DEFAULT_DATA_CONFIG


class AdvancedFloorPlanDataset(Dataset):
    """
    Research-grade dataset with complete ground truth:
    - Floorplan image + segmentation mask
    - Attribute dictionary (geometric parameters)
    - Ground-truth mesh + voxelized occupancy
    - Polygon outlines for vectorization supervision

    Enhanced with optional in-memory caching for performance
    """

    def __init__(
        self,
        data_dir: str = None,
        split: str = "train",
        image_size: Tuple[int, int] = None,
        voxel_size: int = None,
        augment: bool = None,
        config=None,
    ):
        # Use config if provided, otherwise defaults from config.py
        if config is None:
            config = DEFAULT_DATA_CONFIG

        self.data_dir = Path(data_dir or config.data_dir)
        self.split = split
        self.image_size = image_size or config.image_size
        self.voxel_size = voxel_size or config.voxel_size
        self.augment = (
            augment if augment is not None else config.augment
        ) and split == "train"

        # Collect all samples that contain every required file
        self.samples = self._find_complete_samples()
        print(f"Found {len(self.samples)} complete samples for {split}")

        # NEW: In-memory caching for performance
        self.cache_in_memory = getattr(config, "cache_in_memory", False)
        self._cache = None

        if self.cache_in_memory and len(self.samples) > 0:
            print(f"[DATA] Preloading {len(self.samples)} samples into RAM (cache_in_memory=True).")
            print("[DATA] This may take significant memory but will speed up training...")

            # Estimate memory usage
            estimated_mb = self._estimate_memory_usage()
            print(f"[DATA] Estimated memory usage: {estimated_mb:.1f} MB")

            start_time = time.time()
```

```
            self._preload_cache()
            load_time = time.time() - start_time
            print(f"[DATA] Cache preloading completed in {load_time:.2f}s")

    def _estimate_memory_usage(self):
        """Estimate memory usage for caching"""
        if not self.samples:
            return 0.0

        H, W = self.image_size
        n_samples = len(self.samples)

        # Rough estimates in bytes
        image_bytes = H * W * 3  # RGB uint8
        mask_bytes = H * W  # grayscale uint8
        voxel_bytes = self.voxel_size ** 3 * 4  # float32
        json_bytes = 1024  # rough estimate for params + polygons

        total_per_sample = image_bytes + mask_bytes + voxel_bytes + json_bytes
        total_mb = (total_per_sample * n_samples) / (1024 * 1024)

        return total_mb

    def _preload_cache(self):
        """Preload all samples into memory"""
        self._cache = []

        for i, sample in enumerate(self.samples):
            if i % 100 == 0:
                print(f"[DATA] Loading sample {i+1}/{len(self.samples)}")

            try:
                # Load image
                img = cv2.imread(str(sample["image"]))
                if img is None:
                    print(f"Warning: Could not load image {sample['image']}")
                    continue
                img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
                img = cv2.resize(img, self.image_size)  # (W, H) format for cv2.resize

                # Load mask
                mask = cv2.imread(str(sample["mask"]), cv2.IMREAD_GRAYSCALE)
                if mask is None:
                    print(f"Warning: Could not load mask {sample['mask']}")
                    continue
                mask = cv2.resize(mask, self.image_size, interpolation=cv2.INTER_NEAREST)

                # Load voxel data
                try:
                    voxel_data = np.load(sample["voxel"])
                    vox = voxel_data["voxels"]  # Keep as numpy array
                except Exception as e:
                    print(f"Warning: Could not load voxel data {sample['voxel']}: {e}")
                    # Create dummy voxel data
                    vox = np.zeros((self.voxel_size, self.voxel_size, self.voxel_size), dtype=np.float32)

                # Load parameters
                try:
                    with open(sample["params"], "r") as f:
                        params = json.load(f)
                except Exception as e:
                    print(f"Warning: Could not load params {sample['params']}: {e}")
                    params = self._get_default_attributes()

                # Load polygons
                try:
                    with open(sample["polygon"], "r") as f:
                        polygons = json.load(f)
                except Exception as e:
                    print(f"Warning: Could not load polygons {sample['polygon']}: {e}")
                    polygons = {"walls": []}

                self._cache.append({
```

```
                    "image": img,
                    "mask": mask,
                    "vox": vox,
                    "params": params,
                    "polygons": polygons,
                    "sample_id": sample["image"].parent.name,
                })

            except Exception as e:
                print(f"Error loading sample {i}: {e}")
                continue

    def _get_default_attributes(self):
        """Return default attributes for missing param files"""
        return {
            "wall_height": 2.6,
            "wall_thickness": 0.15,
            "window_base_height": 0.7,
            "window_height": 0.95,
            "door_height": 2.6,
            "pixel_scale": 0.02,
        }

    # ---------------------------------------------------------------------
    def _find_complete_samples(self):
        """Locate samples that contain all the expected files."""
        samples = []
        split_dir = self.data_dir / self.split

        if not split_dir.exists():
            print(f"Warning: Split directory {split_dir} does not exist")
            return samples

        for sample_dir in split_dir.iterdir():
            if not sample_dir.is_dir():
                continue

            required_files = {
                "image": sample_dir / "image.png",
                "mask": sample_dir / "mask.png",
                "params": sample_dir / "params.json",
                "mesh": sample_dir / "model.obj",
                "voxel": sample_dir / "voxel_GT.npz",
                "polygon": sample_dir / "polygon.json",
            }

            if all(f.exists() for f in required_files.values()):
                samples.append(required_files)

        return samples

    # ---------------------------------------------------------------------
    def __len__(self):
        return len(self._cache) if self._cache is not None else len(self.samples)

    # ---------------------------------------------------------------------
    def __getitem__(self, idx):
        # Use cached data if available
        if self._cache is not None:
            cached_sample = self._cache[idx]
            image = cached_sample['image']
            mask = cached_sample['mask']
            vox = cached_sample['vox']
            attributes = cached_sample['params']
            polygons_gt = cached_sample['polygons']
            sample_id = cached_sample['sample_id']
        else:
            # Fallback: load from disk on-the-fly
            sample = self.samples[idx]

            # Load image and mask
            image = cv2.imread(str(sample["image"]))
            image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
```

```
            image = cv2.resize(image, self.image_size)

            mask = cv2.imread(str(sample["mask"]), cv2.IMREAD_GRAYSCALE)
            mask = cv2.resize(mask, self.image_size, interpolation=cv2.INTER_NEAREST)

            # Load attributes
            with open(sample["params"], "r") as f:
                attributes = json.load(f)

            # Load voxel ground truth
            voxel_data = np.load(sample["voxel"])
            vox = voxel_data["voxels"]

            # Load polygon ground truth
            with open(sample["polygon"], "r") as f:
                polygons_gt = json.load(f)

            sample_id = sample["image"].parent.name

        # Normalize image to [0,1]
        image = image.astype(np.float32) / 255.0

        # Clean mask (remove class 5 if present)
        mask[mask == 5] = 0

        # Convert to tensors
        image_tensor = torch.from_numpy(image).float().permute(2, 0, 1)
        mask_tensor = torch.from_numpy(mask).long()
        voxels_tensor = torch.from_numpy(vox.astype(np.float32))

        attr_tensor = self._process_attributes(attributes)
        polygon_tensor = self._process_polygons(polygons_gt)

        # Apply augmentation if enabled
        if self.augment:
            image_tensor, mask_tensor = self._augment(image_tensor, mask_tensor)

        # Add validation before returning
        self._validate_sample_data(idx, image_tensor, mask_tensor, attr_tensor, voxels_tensor, polygon_tens

        return {
            "image": image_tensor,
            "mask": mask_tensor,
            "attributes": attr_tensor,
            "voxels_gt": voxels_tensor,
            "polygons_gt": polygon_tensor,
            "sample_id": sample_id,
        }

    # ----------------------------------------------------------------------
    def _validate_sample_data(self, idx, image, mask, attributes, voxels, polygons):
        """Validate sample data for NaN/Inf values"""
        tensors_to_check = [
            ("image", image),
            ("mask", mask),
            ("attributes", attributes),
            ("voxels", voxels),
            ("polygons", polygons["polygons"])
        ]

        corrupted_data = False

        for name, tensor in tensors_to_check:
            if torch.isnan(tensor).any():
                print(f"ERROR: {name} contains NaN values at sample {idx}")
                corrupted_data = True
            if torch.isinf(tensor).any():
                print(f"ERROR: {name} contains Inf values at sample {idx}")
                corrupted_data = True

        if corrupted_data:
            print(f"WARNING: Corrupted data detected in sample {idx}, replacing with safe fallback values")

```

```
            # Replace corrupted tensors with safe fallback values
            for name, tensor in tensors_to_check:
                if torch.isnan(tensor).any() or torch.isinf(tensor).any():
                    if name == "image":
                        # Replace with zeros (black image)
                        image.data = torch.zeros_like(image)
                    elif name == "mask":
                        # Replace with zeros (background class)
                        mask.data = torch.zeros_like(mask).long()
                    elif name == "attributes":
                        # Replace with reasonable default values (0.5 normalized)
                        attributes.data = torch.ones_like(attributes) * 0.5
                    elif name == "voxels":
                        # Replace with empty voxel grid
                        voxels.data = torch.zeros_like(voxels)
                    elif name == "polygons":
                        # Replace polygons with zeros
                        polygons["polygons"].data = torch.zeros_like(polygons["polygons"])

    # ----------------------------------------------------------------------
    def _process_attributes(self, attributes):
        """Convert attribute dictionary to a normalized tensor."""
        # Normalize common architectural parameters into [0,1]
        attr_list = [
            attributes.get("wall_height", 2.6) / 5.0,
            attributes.get("wall_thickness", 0.15) / 0.5,
            attributes.get("window_base_height", 0.7) / 3.0,
            attributes.get("window_height", 0.95) / 2.0,
            attributes.get("door_height", 2.6) / 5.0,
            attributes.get("pixel_scale", 0.01) / 0.02,
        ]

        # Ensure no NaN/Inf values in attribute processing
        safe_attr_list = []
        for val in attr_list:
            if np.isnan(val) or np.isinf(val):
                safe_attr_list.append(0.5)  # Default normalized value
            else:
                safe_attr_list.append(max(0.0, min(1.0, val)))  # Clamp to [0,1]

        return torch.tensor(safe_attr_list, dtype=torch.float32)

    # ----------------------------------------------------------------------
    def _process_polygons(self, polygons_gt):
        """Convert polygon ground truth into a fixed tensor representation.
        Handles both formats:
        1. Nested dict: { "walls": [...], "doors": [...], ... }
        2. Flat list:   [ {"type": "wall", "points": [...]}, ... ]
        """
        max_polygons = 30   # number of polygons per sample
        max_points = 100     # max points per polygon

        processed = torch.zeros(max_polygons, max_points, 2)
        valid_mask = torch.zeros(max_polygons, dtype=torch.bool)

        poly_idx = 0

        try:
            # --- Case 1: dict format ---
            if isinstance(polygons_gt, dict):
                for class_name, polygon_list in polygons_gt.items():
                    if not isinstance(polygon_list, list):
                        continue
                    for polygon in polygon_list:
                        if poly_idx >= max_polygons:
                            break
                        if "points" not in polygon:
                            continue

                        points = np.array(polygon["points"])
                        if len(points) > max_points:
                            # Subsample evenly if too many points
                            indices = np.linspace(0, len(points) - 1, max_points, dtype=int)
```

```
                            points = points[indices]

                        # Check for NaN/Inf in points
                        if np.any(np.isnan(points)) or np.any(np.isinf(points)):
                            print(f"Warning: Invalid polygon points detected, skipping polygon")
                            continue

                        # Normalize to [0,1] relative to image size
                        points = points / np.array(self.image_size)
                        # Clamp to valid range
                        points = np.clip(points, 0.0, 1.0)

                        processed[poly_idx, : len(points)] = torch.from_numpy(points).float()
                        valid_mask[poly_idx] = True
                        poly_idx += 1

            # --- Case 2: list format ---
            elif isinstance(polygons_gt, list):
                for polygon in polygons_gt:
                    if poly_idx >= max_polygons:
                        break
                    if "points" not in polygon:
                        continue

                    points = np.array(polygon["points"])
                    if len(points) > max_points:
                        indices = np.linspace(0, len(points) - 1, max_points, dtype=int)
                        points = points[indices]

                    # Check for NaN/Inf in points
                    if np.any(np.isnan(points)) or np.any(np.isinf(points)):
                        print(f"Warning: Invalid polygon points detected, skipping polygon")
                        continue

                    points = points / np.array(self.image_size)
                    points = np.clip(points, 0.0, 1.0)

                    processed[poly_idx, : len(points)] = torch.from_numpy(points).float()
                    valid_mask[poly_idx] = True
                    poly_idx += 1

        except Exception as e:
            print(f"Warning: Error processing polygons: {e}")
            # Return safe empty polygon data
            processed = torch.zeros(max_polygons, max_points, 2)
            valid_mask = torch.zeros(max_polygons, dtype=torch.bool)

        return {"polygons": processed, "valid_mask": valid_mask}

    # ----------------------------------------------------------------------
    def _augment(self, image, mask):
        """Enhanced data augmentation with rotations, flips, and intensity changes."""
        # Random rotation (multiples of 90° only for architectural data)
        if torch.rand(1) < 0.5:
            k = torch.randint(1, 4, (1,)).item()
            image = torch.rot90(image, k, dims=[1, 2])
            mask = torch.rot90(mask, k, dims=[0, 1])

        # Random horizontal flip
        if torch.rand(1) < 0.5:
            image = torch.flip(image, dims=[2])
            mask = torch.flip(mask, dims=[1])

        # Random vertical flip
        if torch.rand(1) < 0.5:
            image = torch.flip(image, dims=[1])
            mask = torch.flip(mask, dims=[0])

        # Slight brightness/contrast adjustment with safety checks
        if torch.rand(1) < 0.3:
            brightness = torch.rand(1) * 0.2 - 0.1  # ±0.1
            contrast = torch.rand(1) * 0.2 + 0.9     # 0.9-1.1
            image = torch.clamp(image * contrast + brightness, 0, 1)
```

```python

            # Additional safety check for augmented image
            if torch.isnan(image).any() or torch.isinf(image).any():
                print("Warning: Augmentation produced invalid values, reverting to original")
                # Revert to safe values
                image = torch.clamp(image, 0, 1)
                image = torch.where(torch.isnan(image) | torch.isinf(image),
                                  torch.zeros_like(image), image)

        return image, mask

    # ----------------------------------------------------------------------
    def get_cache_info(self):
        """Return information about caching status"""
        return {
            "cache_enabled": self.cache_in_memory,
            "cache_loaded": self._cache is not None,
            "cached_samples": len(self._cache) if self._cache else 0,
            "total_samples": len(self.samples),
            "estimated_memory_mb": self._estimate_memory_usage() if self.cache_in_memory else 0
        }

    def disable_cache(self):
        """Disable caching and free memory"""
        if self._cache is not None:
            print(f"[DATA] Disabling cache and freeing memory for {len(self._cache)} samples")
            self._cache = None
            self.cache_in_memory = False

    def enable_cache(self):
        """Enable caching if not already enabled"""
        if not self.cache_in_memory and self.samples:
            self.cache_in_memory = True
            print("[DATA] Enabling cache...")
            self._preload_cache()


# ======================================================================
# Synthetic sample generator for testing without dataset
# ======================================================================
def create_synthetic_data_sample():
    """Generate a synthetic floorplan with attributes, voxels, and polygons."""
    image = np.ones((256, 256, 3), dtype=np.uint8) * 255
    mask = np.zeros((256, 256), dtype=np.uint8)

    # Simple square room
    room_points = np.array([[50, 50], [200, 50], [200, 200], [50, 200]])
    cv2.fillPoly(mask, [room_points], 1)  # Room = class 1
    cv2.polylines(image, [room_points], True, (0, 0, 0), 3)

    # Add door
    cv2.rectangle(mask, (90, 50), (110, 70), 2)  # Door = class 2
    cv2.rectangle(image, (90, 50), (110, 70), (255, 0, 0), -1)

    # Attributes
    attributes = {
        "wall_height": 2.6,
        "wall_thickness": 0.15,
        "window_base_height": 0.7,
        "window_height": 0.95,
        "door_height": 2.6,
        "pixel_scale": 0.02,
    }

    # Simple voxel GT
    voxels = np.zeros((64, 64, 64), dtype=bool)
    voxels[:20, 10:50, 10:50] = True

    # Polygon GT
    polygons = {"walls": [{"points": room_points.tolist()}]}

    return image, mask, attributes, voxels, polygons

```

```python

class SyntheticFloorPlanDataset(Dataset):
    """
    Synthetic dataset for testing and development when real data is not available
    """

    def __init__(self, num_samples=1000, image_size=(256, 256), voxel_size=64):
        self.num_samples = num_samples
        self.image_size = image_size
        self.voxel_size = voxel_size

    def __len__(self):
        return self.num_samples

    def __getitem__(self, idx):
        # Generate deterministic synthetic data based on index
        np.random.seed(idx)
        torch.manual_seed(idx)

        image, mask, attributes, voxels, polygons_gt = create_synthetic_data_sample()

        # Convert to tensors
        image_tensor = torch.from_numpy(image.astype(np.float32) / 255.0).permute(2, 0, 1)
        mask_tensor = torch.from_numpy(mask).long()
        voxels_tensor = torch.from_numpy(voxels.astype(np.float32))

        # Process attributes and polygons using same methods as main dataset
        dataset = AdvancedFloorPlanDataset.__new__(AdvancedFloorPlanDataset)
        dataset.image_size = self.image_size

        attr_tensor = dataset._process_attributes(attributes)
        polygon_tensor = dataset._process_polygons(polygons_gt)

        return {
            "image": image_tensor,
            "mask": mask_tensor,
            "attributes": attr_tensor,
            "voxels_gt": voxels_tensor,
            "polygons_gt": polygon_tensor,
            "sample_id": f"synthetic_{idx:06d}",
        }
```

■ File: demo.py

==

```python
"""
Demo script for the Neural-Geometric 3D Model Generator
"""

import torch
import cv2
import numpy as np
from pathlib import Path

from models.model import NeuralGeometric3DGenerator
from dataset import create_synthetic_data_sample
from utils.visualization import visualize_predictions, create_model_summary_report


def demo_pipeline():
    """Demonstrate the complete pipeline with synthetic data"""
    print("Neural-Geometric 3D Model Generator Demo")
    print("=" * 50)

    # Create output directory
    demo_dir = Path("./demo_outputs")
    demo_dir.mkdir(exist_ok=True)

    # Create synthetic sample
    print("Creating synthetic data sample...")
    image, mask, attributes, voxels, polygons = create_synthetic_data_sample()

```

```python
    # Save synthetic data
    cv2.imwrite(str(demo_dir / "demo_input.png"), image)
    cv2.imwrite(str(demo_dir / "demo_mask.png"), mask * 50)

    # Create model (random weights for demo)
    print("Initializing model...")
    model = NeuralGeometric3DGenerator()
    model.eval()

    # Convert to tensors
    image_tensor = torch.from_numpy(image / 255.0).float().permute(2, 0, 1).unsqueeze(0)

    # Create model summary
    create_model_summary_report(model, image_tensor, str(demo_dir / "model_summary.txt"))

    # Forward pass
    print("Running forward pass...")
    with torch.no_grad():
        predictions = model(image_tensor)

    print("Forward pass completed")
    print(f"Segmentation shape: {predictions['segmentation'].shape}")
    print(f"Attributes shape: {predictions['attributes'].shape}")
    print(f"Polygons shape: {predictions['polygons'].shape}")
    print(f"Voxels shape: {predictions['voxels_pred'].shape}")

    # Extract and save results
    seg_pred = torch.argmax(predictions["segmentation"], dim=1).squeeze().numpy()
    attr_pred = predictions["attributes"].squeeze().numpy()

    cv2.imwrite(str(demo_dir / "demo_seg_pred.png"), seg_pred * 50)

    print(f"Predicted attributes: {attr_pred}")

    # Create visualization
    print("Creating visualizations...")

    # Create targets for visualization
    targets = {
        "mask": torch.from_numpy(mask).unsqueeze(0),
        "attributes": torch.from_numpy(np.array([
            attributes["wall_height"] / 5.0,
            attributes["wall_thickness"] / 0.5,
            attributes["window_base_height"] / 3.0,
            attributes["window_height"] / 2.0,
            attributes["door_height"] / 5.0,
            attributes["pixel_scale"] / 0.02,
        ])).float().unsqueeze(0)
    }

    visualize_predictions(
        image_tensor,
        predictions,
        targets,
        save_path=str(demo_dir / "demo_predictions.png")
    )

    print(f"Demo completed successfully! Results saved to {demo_dir}")


def demo_with_pretrained(model_path, input_image_path=None):
    """Demo with a pretrained model"""
    if not Path(model_path).exists():
        print(f"Model file {model_path} not found!")
        return

    print(f"Running demo with pretrained model: {model_path}")

    # Load model
    model = NeuralGeometric3DGenerator()
    checkpoint = torch.load(model_path, map_location="cuda")
    model.load_state_dict(checkpoint["model_state_dict"])
    model.cuda()
```

```
    model.eval()

    # Use provided image or create synthetic
    if input_image_path and Path(input_image_path).exists():
        image = cv2.imread(input_image_path)
        image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
        image = cv2.resize(image, (256, 256))
        image_tensor = torch.from_numpy(image / 255.0).float().permute(2, 0, 1).unsqueeze(0).cuda()
        print(f"Using input image: {input_image_path}")
    else:
        print("Using synthetic data...")
        image, _, _, _, _ = create_synthetic_data_sample()
        image_tensor = torch.from_numpy(image / 255.0).float().permute(2, 0, 1).unsqueeze(0).cuda()

    # Run inference
    with torch.no_grad():
        predictions = model(image_tensor)

    # Visualize results
    demo_dir = Path("./demo_outputs")
    demo_dir.mkdir(exist_ok=True)

    visualize_predictions(
        image_tensor,
        predictions,
        save_path=str(demo_dir / "pretrained_demo.png")
    )

    print(f"Pretrained demo completed! Results saved to {demo_dir}")


if __name__ == "__main__":
    import argparse

    parser = argparse.ArgumentParser(description="Demo Neural-Geometric 3D Model Generator")
    parser.add_argument("--model_path", type=str, default=None,
                       help="Path to pretrained model (optional)")
    parser.add_argument("--input_image", type=str, default=None,
                       help="Input image path (optional)")

    args = parser.parse_args()

    if args.model_path:
        demo_with_pretrained(args.model_path, args.input_image)
    else:
        demo_pipeline()
```

■ File: evaluate.py

==

```
"""
evaluate.py
Comprehensive evaluation CLI for Neural-Geometric 3D Model Generator.

Usage examples:
  python evaluate.py --model_path checkpoints/final_model.pth --data_dir ./data/floorplans
  python evaluate.py --model_path checkpoints/final_model.pth --data_dir ./data/floorplans --visualize --sa
"""

import argparse
import json
from pathlib import Path
from typing import List, Dict

import torch
from torch.utils.data import DataLoader

# Project imports (match your repo layout)
from dataset import AdvancedFloorPlanDataset
from evaluation.metrics import ModelEvaluator
from inference.engine import ResearchInferenceEngine
from utils.visualization import visualize_predictions
```

```


def save_json(obj, path: Path):
    path.parent.mkdir(parents=True, exist_ok=True)
    with open(path, "w") as f:
        json.dump(obj, f, indent=2)
    print(f"[?] Saved JSON -> {path}")


def gather_per_sample_metrics(
    evaluator: ModelEvaluator,
    dataset: AdvancedFloorPlanDataset,
    device: str,
    max_samples: int = None,
) -> List[Dict]:
    """
    Re-run evaluation loop sample-by-sample and collect per-sample metrics.
    We use evaluator._evaluate_* helper methods (present in evaluation/metrics.py)
    so metrics match the overall evaluation.
    """
    loader = DataLoader(dataset, batch_size=1, shuffle=False)
    per_sample = []

    with torch.no_grad():
        for idx, batch in enumerate(loader):
            if max_samples is not None and idx >= max_samples:
                break

            # Move tensors to device where applicable
            batch_for_model = {}
            for k, v in batch.items():
                if torch.is_tensor(v):
                    batch_for_model[k] = v.to(device)
                else:
                    batch_for_model[k] = v  # dicts/strings stay as-is

            # Forward
            preds = evaluator.model(batch_for_model["image"])

            # segmentation
            seg_res = evaluator._evaluate_segmentation(preds["segmentation"], batch_for_model["mask"])
            # attributes
            attr_res = evaluator._evaluate_attributes(preds["attributes"], batch_for_model["attributes"].to
            # voxels
            voxel_res = evaluator._evaluate_voxels(preds["voxels_pred"], batch_for_model["voxels_gt"].to(dev
            # polygons ? evaluator._evaluate_polygons expects format used in metrics.py
            # batch["polygons_gt"] is a dict with "polygons" and "valid_mask"
            poly_res = evaluator._evaluate_polygons(preds["polygons"], preds.get("polygon_validity", preds.

            sample_id = batch["sample_id"][0] if isinstance(batch["sample_id"], (list, tuple)) else batch["
            sample_metrics = {
                "sample_id": str(sample_id),
                "segmentation": seg_res,
                "attributes": attr_res,
                "voxels": voxel_res,
                "polygons": poly_res,
            }
            per_sample.append(sample_metrics)

            if (idx + 1) % 10 == 0:
                print(f"[INFO] Collected per-sample metrics for {idx+1}/{len(loader)} samples")

    return per_sample


def run_visualization_and_exports(
    engine: ResearchInferenceEngine,
    dataset: AdvancedFloorPlanDataset,
    output_dir: Path,
    device: str,
    num_viz: int = 10,
    max_export: int = 5,
):
```

```
    """
    For the first `num_viz` samples, create visualizations using the model and optionally
    run deterministic 3D export to save intermediate results and a .obj.
    """
    output_dir.mkdir(parents=True, exist_ok=True)
    loader = DataLoader(dataset, batch_size=1, shuffle=False)

    viz_count = 0
    export_count = 0

    with torch.no_grad():
        for idx, batch in enumerate(loader):
            sample_id = batch["sample_id"][0] if isinstance(batch["sample_id"], (list, tuple)) else batch["
            sample_dir = Path(dataset.data_dir) / "test" / str(sample_id)

            # Prepare tensors
            image_tensor = batch["image"].to(device)
            target_mask = batch["mask"].unsqueeze(0) if torch.is_tensor(batch["mask"]) else None
            target_attrs = batch["attributes"].unsqueeze(0) if torch.is_tensor(batch["attributes"]) else No

            # Model predictions using engine.model (same underlying model)
            preds = engine.model(image_tensor)

            # Visualization
            if viz_count < num_viz:
                vis_path = output_dir / f"viz_{sample_id}.png"
                try:
                    visualize_predictions(
                        image_tensor,
                        preds,
                        {"mask": target_mask, "attributes": target_attrs},
                        save_path=str(vis_path),
                    )
                    print(f"[?] Saved visualization for sample {sample_id} -> {vis_path}")
                except Exception as e:
                    print(f"[!] Visualization failed for {sample_id}: {e}")
                viz_count += 1

            # Export deterministic 3D (uses the image file path)
            if export_count < max_export:
                image_file = sample_dir / "image.png"
                out_obj = output_dir / f"{sample_id}_predicted_model.obj"
                try:
                    success = engine.generate_3d_model(str(image_file), str(out_obj), export_intermediate=T
                    if success:
                        print(f"[?] Exported deterministic 3D model for {sample_id} -> {out_obj}")
                    else:
                        print(f"[!] 3D export returned False for {sample_id}")
                except Exception as e:
                    print(f"[!] 3D export failed for {sample_id}: {e}")
                export_count += 1

            if viz_count >= num_viz and export_count >= max_export:
                break


def main():
    parser = argparse.ArgumentParser(description="Evaluate Neural-Geometric 3D model")
    parser.add_argument("--model_path", "-m", required=True, help="Path to model checkpoint (checkpoint.pth
    parser.add_argument("--data_dir", "-d", default="./data/floorplans", help="Dataset root with train/val/
    parser.add_argument("--device", default=None, help="Device to use (cuda or cpu). Auto-detect if omitted
    parser.add_argument("--visualize", action="store_true", help="Save visual comparison images (pred vs GT
    parser.add_argument("--save_outputs", action="store_true", help="Run deterministic 3D export for some s
    parser.add_argument("--output_dir", default="./evaluation_outputs", help="Where to save reports/visuals
    parser.add_argument("--num_viz", type=int, default=10, help="How many visualizations to produce (defaul
    parser.add_argument("--max_exports", type=int, default=3, help="How many deterministic 3D exports to ru
    parser.add_argument("--per_sample_json", action="store_true", help="Save per-sample metrics JSON (may b
    parser.add_argument("--limit_samples", type=int, default=None, help="If set, limit evaluation to first 

    args = parser.parse_args()

    # Determine device
    device = args.device or ("cuda" if torch.cuda.is_available() else "cpu")
```

```python
    print(f"[INFO] Using device: {device}")

    model_path = Path(args.model_path)
    if not model_path.exists():
        print(f"[ERROR] Model not found at: {model_path}")
        return

    # Load test dataset
    dataset = AdvancedFloorPlanDataset(data_dir=args.data_dir, split="test")
    if len(dataset) == 0:
        print("[ERROR] No test samples found (dataset may be empty or data_dir incorrect).")
        return

    # If user asked for a limited quick run, slice dataset.samples accordingly.
    if args.limit_samples is not None:
        # Create a shallow copy dataset pointing to first N samples
        dataset.samples = dataset.samples[: args.limit_samples]
        print(f"[INFO] Limiting evaluation to first {len(dataset)} samples")

    # Create evaluator and run full evaluation
    evaluator = ModelEvaluator(str(model_path), device=device)
    summary = evaluator.evaluate_dataset(dataset)
    evaluator.print_evaluation_results(summary)

    # Save summary JSON
    out_dir = Path(args.output_dir)
    out_dir.mkdir(parents=True, exist_ok=True)
    save_json(summary, out_dir / f"{model_path.stem}_summary.json")

    # Optionally collect detailed per-sample metrics
    if args.per_sample_json:
        print("[INFO] Collecting per-sample metrics (this re-runs model inference sample-by-sample)...")
        per_sample = gather_per_sample_metrics(evaluator, dataset, device, max_samples=None)
        save_json(per_sample, out_dir / f"{model_path.stem}_per_sample_metrics.json")

    # Visualization and/or exports
    if args.visualize or args.save_outputs:
        print("[INFO] Initializing inference engine for visualizations/exports...")
        engine = ResearchInferenceEngine(model_path=str(model_path), device=device)
        run_visualization_and_exports(
            engine,
            dataset,
            out_dir,
            device,
            num_viz=args.num_viz,
            max_export=args.max_exports,
        )

    print("[?] Evaluation finished.")


if __name__ == "__main__":
    main()
```

■ File: evaluation\metrics.py

==

```python
"""
Evaluation metrics and utilities for the Neural-Geometric 3D Model Generator
"""

import torch
import numpy as np
from torch.utils.data import DataLoader

from models.model import NeuralGeometric3DGenerator
from dataset import AdvancedFloorPlanDataset


def compute_iou(pred, target):
    """Compute IoU for segmentation"""
    intersection = (pred & target).float().sum()
```

```
    union = (pred | target).float().sum()
    return (intersection / (union + 1e-6)).item()


def compute_3d_iou(pred, target):
    """Compute 3D IoU for voxel grids"""
    pred_bool = pred.bool()
    target_bool = target.bool()

    intersection = (pred_bool & target_bool).float().sum()
    union = (pred_bool | target_bool).float().sum()

    return (intersection / (union + 1e-6)).item()


def compute_polygon_metrics(pred_polygons, gt_polygons, validity_pred, validity_gt):
    """Compute metrics for polygon prediction"""
    # Chamfer distance between polygon sets
    valid_pred = pred_polygons[validity_pred > 0.5]
    valid_gt = gt_polygons[validity_gt]

    if len(valid_pred) == 0 or len(valid_gt) == 0:
        return {"chamfer_distance": float('inf'), "validity_accuracy": 0.0}

    # Simplified chamfer distance computation
    chamfer_dist = 0.0
    for pred_poly in valid_pred:
        min_dist = float('inf')
        for gt_poly in valid_gt:
            dist = torch.norm(pred_poly - gt_poly, dim=-1).min().item()
            min_dist = min(min_dist, dist)
        chamfer_dist += min_dist

    chamfer_dist /= len(valid_pred)

    # Validity accuracy
    validity_acc = ((validity_pred > 0.5) == validity_gt).float().mean().item()

    return {
        "chamfer_distance": chamfer_dist,
        "validity_accuracy": validity_acc
    }


class ModelEvaluator:
    """Comprehensive model evaluation"""

    def __init__(self, model_path, device="cuda"):
        self.device = device
        self.model = NeuralGeometric3DGenerator()

        # Load model
        checkpoint = torch.load(model_path, map_location=device)
        self.model.load_state_dict(checkpoint["model_state_dict"])
        self.model.to(device)
        self.model.eval()

        print(f"Loaded model from {model_path}")

    def evaluate_dataset(self, test_dataset):
        """Comprehensive evaluation on test dataset"""
        test_loader = DataLoader(test_dataset, batch_size=1, shuffle=False)

        # Metrics storage
        metrics = {
            "segmentation": {"ious": [], "class_ious": []},
            "attributes": {"maes": [], "mses": []},
            "voxels": {"ious": [], "dice_scores": []},
            "polygons": {"chamfer_distances": [], "validity_accs": []},
        }

        with torch.no_grad():
            for batch_idx, batch in enumerate(test_loader):
```

```
                batch = {k: v.to(self.device) if torch.is_tensor(v) else v
                        for k, v in batch.items()}

                predictions = self.model(batch["image"])

                # Evaluate segmentation
                seg_metrics = self._evaluate_segmentation(
                    predictions["segmentation"], batch["mask"]
                )
                metrics["segmentation"]["ious"].append(seg_metrics["iou"])
                metrics["segmentation"]["class_ious"].append(seg_metrics["class_ious"])

                # Evaluate attributes
                attr_metrics = self._evaluate_attributes(
                    predictions["attributes"], batch["attributes"]
                )
                metrics["attributes"]["maes"].append(attr_metrics["mae"])
                metrics["attributes"]["mses"].append(attr_metrics["mse"])

                # Evaluate voxels
                voxel_metrics = self._evaluate_voxels(
                    predictions["voxels_pred"], batch["voxels_gt"]
                )
                metrics["voxels"]["ious"].append(voxel_metrics["iou"])
                metrics["voxels"]["dice_scores"].append(voxel_metrics["dice"])

                # Evaluate polygons
                poly_metrics = self._evaluate_polygons(
                    predictions["polygons"],
                    predictions["polygon_validity"],
                    batch["polygons_gt"]
                )
                metrics["polygons"]["chamfer_distances"].append(poly_metrics["chamfer_distance"])
                metrics["polygons"]["validity_accs"].append(poly_metrics["validity_accuracy"])

                if (batch_idx + 1) % 10 == 0:
                    print(f"Evaluated {batch_idx + 1}/{len(test_loader)} samples")

        return self._compute_summary_metrics(metrics)

    def _evaluate_segmentation(self, pred_seg, target_mask):
        """Evaluate segmentation performance"""
        seg_pred = torch.argmax(pred_seg, dim=1)

        # Overall IoU
        overall_iou = compute_iou(seg_pred, target_mask)

        # Per-class IoU
        num_classes = pred_seg.shape[1]
        class_ious = []

        for c in range(num_classes):
            pred_c = (seg_pred == c)
            target_c = (target_mask == c)

            if target_c.sum() > 0:  # Only compute if class exists in ground truth
                iou_c = compute_iou(pred_c, target_c)
                class_ious.append(iou_c)

        return {
            "iou": overall_iou,
            "class_ious": class_ious
        }

    def _evaluate_attributes(self, pred_attrs, target_attrs):
        """Evaluate attribute prediction"""
        mae = torch.mean(torch.abs(pred_attrs - target_attrs)).item()
        mse = torch.mean((pred_attrs - target_attrs) ** 2).item()

        return {"mae": mae, "mse": mse}

    def _evaluate_voxels(self, pred_voxels, target_voxels):
        """Evaluate 3D voxel prediction"""
```

```
        pred_binary = (torch.sigmoid(pred_voxels) > 0.5).float()
        target_float = target_voxels.float()

        # 3D IoU
        iou_3d = compute_3d_iou(pred_binary, target_float)

        # 3D Dice score
        intersection = (pred_binary * target_float).sum()
        dice = (2 * intersection) / (pred_binary.sum() + target_float.sum() + 1e-6)

        return {
            "iou": iou_3d,
            "dice": dice.item()
        }

    def _evaluate_polygons(self, pred_polygons, pred_validity, gt_polygons):
        """Evaluate polygon prediction"""
        return compute_polygon_metrics(
            pred_polygons[0],
            gt_polygons["polygons"][0],
            pred_validity[0],
            gt_polygons["valid_mask"][0]
        )

    def _compute_summary_metrics(self, metrics):
        """Compute summary statistics"""
        summary = {}

        # Segmentation
        summary["segmentation_mIoU"] = np.mean(metrics["segmentation"]["ious"])
        summary["segmentation_std"] = np.std(metrics["segmentation"]["ious"])

        # Attributes
        summary["attribute_MAE"] = np.mean(metrics["attributes"]["maes"])
        summary["attribute_MAE_std"] = np.std(metrics["attributes"]["maes"])

        # Voxels
        summary["voxel_mIoU"] = np.mean(metrics["voxels"]["ious"])
        summary["voxel_mIoU_std"] = np.std(metrics["voxels"]["ious"])
        summary["voxel_dice"] = np.mean(metrics["voxels"]["dice_scores"])

        # Polygons
        valid_chamfer = [d for d in metrics["polygons"]["chamfer_distances"] if d != float('inf')]
        if valid_chamfer:
            summary["polygon_chamfer"] = np.mean(valid_chamfer)
            summary["polygon_chamfer_std"] = np.std(valid_chamfer)
        else:
            summary["polygon_chamfer"] = float('inf')
            summary["polygon_chamfer_std"] = 0.0

        summary["polygon_validity_acc"] = np.mean(metrics["polygons"]["validity_accs"])

        return summary

    def print_evaluation_results(self, summary):
        """Print formatted evaluation results"""
        print("=" * 60)
        print("COMPREHENSIVE EVALUATION RESULTS")
        print("=" * 60)

        print(f"Segmentation mIoU: {summary['segmentation_mIoU']:.4f} ± {summary['segmentation_std']:.4f}")
        print(f"Attribute MAE: {summary['attribute_MAE']:.4f} ± {summary['attribute_MAE_std']:.4f}")
        print(f"Voxel 3D mIoU: {summary['voxel_mIoU']:.4f} ± {summary['voxel_mIoU_std']:.4f}")
        print(f"Voxel Dice Score: {summary['voxel_dice']:.4f}")

        if summary['polygon_chamfer'] != float('inf'):
            print(f"Polygon Chamfer Distance: {summary['polygon_chamfer']:.4f} ± {summary['polygon_chamfer_
        else:
            print("Polygon Chamfer Distance: No valid polygons")

        print(f"Polygon Validity Accuracy: {summary['polygon_validity_acc']:.4f}")
        print("=" * 60)
```

```python

def evaluate_model(model_path, data_dir="./data/floorplans"):
    """Standalone evaluation function"""
    # Load test dataset
    test_dataset = AdvancedFloorPlanDataset(data_dir, split="test")

    if len(test_dataset) == 0:
        print("No test samples found!")
        return None

    # Create evaluator
    evaluator = ModelEvaluator(model_path)

    # Run evaluation
    summary = evaluator.evaluate_dataset(test_dataset)

    # Print results
    evaluator.print_evaluation_results(summary)

    return summary
```

■ File: infer.py

```python
"""
Main inference script for generating 3D models from 2D floorplans
"""

import argparse
from pathlib import Path

from inference.engine import ResearchInferenceEngine
from config import DEFAULT_INFERENCE_CONFIG


def main():
    parser = argparse.ArgumentParser(description="Generate 3D models from 2D floorplans")
    parser.add_argument("--model_path", type=str, default="final_model.pth",
                        help="Path to trained model")
    parser.add_argument("--input", type=str, required=True,
                        help="Input image path or directory")
    parser.add_argument("--output", type=str, required=True,
                        help="Output path or directory")
    parser.add_argument("--device", type=str, default="cuda",
                        help="Inference device")
    parser.add_argument("--export_intermediate", action="store_true",
                        help="Export intermediate results")
    parser.add_argument("--polygon_threshold", type=float, default=0.5,
                        help="Threshold for polygon validity")

    args = parser.parse_args()

    # Initialize inference engine
    print(f"Initializing inference engine...")
    engine = ResearchInferenceEngine(
        model_path=args.model_path,
        device=args.device
    )

    input_path = Path(args.input)
    output_path = Path(args.output)

    if input_path.is_file():
        # Single image inference
        print(f"Processing single image: {input_path}")

        if not output_path.suffix:
            output_path = output_path / f"{input_path.stem}_model.obj"

        success = engine.generate_3d_model(
            str(input_path),
            str(output_path),
```

```python
            export_intermediate=args.export_intermediate
        )

        if success:
            print(f"? Successfully generated: {output_path}")
        else:
            print(f"? Failed to generate model for: {input_path}")

    elif input_path.is_dir():
        # Batch processing
        print(f"Processing directory: {input_path}")

        # Find all image files
        image_extensions = {'.png', '.jpg', '.jpeg', '.bmp', '.tiff'}
        image_files = [
            f for f in input_path.iterdir()
            if f.suffix.lower() in image_extensions
        ]

        if not image_files:
            print("No image files found in input directory!")
            return

        print(f"Found {len(image_files)} image files")

        # Create output directory
        output_path.mkdir(exist_ok=True)

        # Process batch
        results = engine.process_batch(image_files, output_path)

        # Print summary
        successful = sum(1 for r in results if r["success"])
        print(f"\nBatch processing completed:")
        print(f"? Successful: {successful}/{len(results)}")
        print(f"? Failed: {len(results) - successful}/{len(results)}")

        # List failed cases
        failed_cases = [r for r in results if not r["success"]]
        if failed_cases:
            print("\nFailed cases:")
            for case in failed_cases:
                error_msg = case.get("error", "Unknown error")
                print(f"  - {Path(case['input']).name}: {error_msg}")

    else:
        print(f"Error: Input path {input_path} does not exist!")


if __name__ == "__main__":
    main()
```

■ File: inference\engine.py

==

```python
"""
Research-grade inference engine for 2D to 3D floorplan generation
"""

import torch
import cv2
import numpy as np
import json
import trimesh
from pathlib import Path

from models.model import NeuralGeometric3DGenerator
from config import DEFAULT_INFERENCE_CONFIG


class ResearchInferenceEngine:
    """
```

```
    Complete inference system that converts 2D floorplans to 3D models
    following the deterministic export pipeline
    """

    def __init__(self, model_path=None, device="cuda", config=None):
        if config is None:
            config = DEFAULT_INFERENCE_CONFIG

        self.device = device
        self.config = config
        self.model = NeuralGeometric3DGenerator()

        # Load trained model
        model_path = model_path or config.model_path
        checkpoint = torch.load(model_path, map_location=device)
        self.model.load_state_dict(checkpoint["model_state_dict"])
        self.model.to(device)
        self.model.eval()

        print(f"Loaded trained model from {model_path}")

    def generate_3d_model(
        self,
        image_path: str,
        output_path: str,
        export_intermediate: bool = None
    ):
        """
        Complete pipeline: Image -> Segmentation -> Polygons -> 3D Model
        """
        export_intermediate = export_intermediate or self.config.export_intermediate

        # Load and preprocess image
        image = self._load_image(image_path)

        with torch.no_grad():
            # Neural network inference
            predictions = self.model(image)

            # Extract predictions
            segmentation = predictions["segmentation"]
            attributes = predictions["attributes"]
            polygons = predictions["polygons"]
            validity = predictions["polygon_validity"]

            print("Neural network inference complete")

            # Convert to deterministic representations
            mask_np = self._extract_mask(segmentation)
            attributes_dict = self._extract_attributes(attributes)
            polygons_list = self._extract_polygons(polygons, validity)

            print(f"Extracted: {len(polygons_list)} valid polygons")

            # Export intermediate results if requested
            if export_intermediate:
                self._export_intermediates(
                    mask_np, attributes_dict, polygons_list, Path(output_path).parent
                )

            # Generate 3D model using deterministic pipeline
            success = self._generate_deterministic_3d(
                mask_np, attributes_dict, polygons_list, output_path
            )

            return success

    def _load_image(self, image_path):
        """Load and preprocess input image"""
        image = cv2.imread(image_path)
        if image is None:
            raise ValueError(f"Could not load image from {image_path}")

```

```
        image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
        image = cv2.resize(image, (256, 256))
        image = torch.from_numpy(image / 255.0).float()
        image = image.permute(2, 0, 1).unsqueeze(0)
        return image.to(self.device)

    def _extract_mask(self, segmentation):
        """Convert soft segmentation to hard mask"""
        seg_pred = torch.argmax(segmentation, dim=1)
        mask_np = seg_pred.squeeze().cpu().numpy().astype(np.uint8)
        return mask_np

    def _extract_attributes(self, attributes):
        """Convert normalized attributes back to physical values"""
        attr_np = attributes.squeeze().cpu().numpy()

        # Denormalize (reverse of normalization in dataset)
        attributes_dict = {
            "wall_height": float(attr_np[0] * 5.0),
            "wall_thickness": float(attr_np[1] * 0.5),
            "window_base_height": float(attr_np[2] * 3.0),
            "window_height": float(attr_np[3] * 2.0),
            "door_height": float(attr_np[4] * 5.0),
            "pixel_scale": float(attr_np[5] * 0.02),
        }

        return attributes_dict

    def _extract_polygons(self, polygons, validity, threshold=None):
        """Extract valid polygons from network predictions"""
        threshold = threshold or self.config.polygon_threshold
        batch_size, num_polys, num_points, _ = polygons.shape

        polygons_list = []

        for poly_idx in range(num_polys):
            if validity[0, poly_idx] > threshold:  # Only valid polygons
                poly_points = polygons[0, poly_idx].cpu().numpy()

                # Remove zero-padded points
                valid_points = poly_points[poly_points.sum(axis=1) > 0]

                if len(valid_points) >= 3:  # Minimum for a polygon
                    # Convert to image coordinates (assuming 256x256 input)
                    valid_points = valid_points * 256
                    polygons_list.append(
                        {
                            "points": valid_points.tolist(),
                            "class": "wall",  # Simplified - in practice classify polygon type
                        }
                    )

        return polygons_list

    def _export_intermediates(self, mask, attributes, polygons, output_dir):
        """Export intermediate results for debugging/analysis"""
        output_dir = Path(output_dir)
        output_dir.mkdir(exist_ok=True)

        # Export mask
        cv2.imwrite(str(output_dir / "predicted_mask.png"), mask * 50)

        # Export attributes
        with open(output_dir / "predicted_attributes.json", "w") as f:
            json.dump(attributes, f, indent=2)

        # Export polygons
        with open(output_dir / "predicted_polygons.json", "w") as f:
            json.dump(polygons, f, indent=2)

        # Visualize polygons on mask
        vis_img = np.zeros((256, 256, 3), dtype=np.uint8)
        vis_img[:, :, 0] = mask * 50  # Background
```

```python

        for poly in polygons:
            points = np.array(poly["points"], dtype=np.int32)
            cv2.polylines(vis_img, [points], True, (0, 255, 0), 2)

        cv2.imwrite(str(output_dir / "polygon_visualization.png"), vis_img)

        print(f"Intermediate results exported to {output_dir}")

    def _generate_deterministic_3d(self, mask, attributes, polygons, output_path):
        """Generate 3D model using deterministic geometric operations"""
        try:
            # Initialize mesh components
            vertices = []
            faces = []
            vertex_count = 0

            # Extract geometric parameters
            wall_height = attributes.get("wall_height", 2.6)
            wall_thickness = attributes.get("wall_thickness", 0.15)
            pixel_scale = attributes.get("pixel_scale", 0.01)

            print(
                f"Generating 3D model with wall_height={wall_height:.2f}m, thickness={wall_thickness:.2f}m"
            )

            # Process each polygon (walls, rooms, etc.)
            for poly_idx, polygon in enumerate(polygons):
                poly_vertices, poly_faces = self._extrude_polygon_3d(
                    polygon["points"],
                    wall_height,
                    wall_thickness,
                    pixel_scale,
                    vertex_count,
                )

                vertices.extend(poly_vertices)
                faces.extend(poly_faces)
                vertex_count += len(poly_vertices)

            # Add floor and ceiling
            floor_verts, floor_faces = self._generate_floor_ceiling(
                mask, pixel_scale, wall_height, vertex_count
            )
            vertices.extend(floor_verts)
            faces.extend(floor_faces)

            if len(vertices) == 0:
                print("No geometry generated")
                return False

            # Create mesh
            mesh = trimesh.Trimesh(vertices=np.array(vertices), faces=np.array(faces))

            # Clean up mesh
            mesh.remove_duplicate_faces()
            mesh.remove_unreferenced_vertices()
            mesh.fix_normals()

            # Export
            mesh.export(output_path)
            print(f"3D model exported to {output_path}")
            print(
                f"Mesh statistics: {len(mesh.vertices)} vertices, {len(mesh.faces)} faces"
            )

            return True

        except Exception as e:
            print(f"Error generating 3D model: {str(e)}")
            return False

    def _extrude_polygon_3d(self, points, height, thickness, scale, vertex_offset):
```

```
        """Extrude a 2D polygon to create 3D wall geometry"""
        vertices = []
        faces = []

        # Convert points to 3D coordinates
        points_3d = []
        for point in points:
            x = (point[0] - 128) * scale  # Center and scale
            z = (128 - point[1]) * scale  # Flip Y and scale
            points_3d.append([x, 0, z])

        # Create bottom vertices (y=0)
        bottom_outer = points_3d
        bottom_inner = self._inset_polygon(points_3d, thickness)

        # Create top vertices (y=height)
        top_outer = [[p[0], height, p[2]] for p in bottom_outer]
        top_inner = [[p[0], height, p[2]] for p in bottom_inner]

        # Combine all vertices
        all_vertices = bottom_outer + bottom_inner + top_outer + top_inner
        vertices.extend(all_vertices)

        n_points = len(points_3d)

        # Generate faces for walls
        for i in range(n_points):
            next_i = (i + 1) % n_points

            # Outer wall faces
            v1 = vertex_offset + i  # bottom outer
            v2 = vertex_offset + next_i  # bottom outer next
            v3 = vertex_offset + 2 * n_points + next_i  # top outer next
            v4 = vertex_offset + 2 * n_points + i  # top outer

            faces.extend([[v1, v2, v3], [v1, v3, v4]])

            # Inner wall faces (reverse winding)
            v1 = vertex_offset + n_points + i  # bottom inner
            v2 = vertex_offset + n_points + next_i  # bottom inner next
            v3 = vertex_offset + 3 * n_points + next_i  # top inner next
            v4 = vertex_offset + 3 * n_points + i  # top inner

            faces.extend([[v1, v3, v2], [v1, v4, v3]])

        # Top cap (between outer and inner)
        for i in range(n_points):
            next_i = (i + 1) % n_points

            v1 = vertex_offset + 2 * n_points + i  # top outer
            v2 = vertex_offset + 2 * n_points + next_i  # top outer next
            v3 = vertex_offset + 3 * n_points + next_i  # top inner next
            v4 = vertex_offset + 3 * n_points + i  # top inner

            faces.extend([[v1, v2, v3], [v1, v3, v4]])

        # Bottom cap (between outer and inner)
        for i in range(n_points):
            next_i = (i + 1) % n_points

            v1 = vertex_offset + i  # bottom outer
            v2 = vertex_offset + next_i  # bottom outer next
            v3 = vertex_offset + n_points + next_i  # bottom inner next
            v4 = vertex_offset + n_points + i  # bottom inner

            faces.extend([[v1, v3, v2], [v1, v4, v3]])

        return vertices, faces

    def _inset_polygon(self, points, inset_distance):
        """Create inset polygon for wall thickness"""
        if len(points) < 3:
            return points
```

```
        # Simple inset by moving each point inward along angle bisector
        inset_points = []
        n = len(points)

        for i in range(n):
            prev_i = (i - 1) % n
            next_i = (i + 1) % n

            p_prev = np.array(points[prev_i])
            p_curr = np.array(points[i])
            p_next = np.array(points[next_i])

            # Vectors to adjacent points
            v1 = p_curr - p_prev
            v2 = p_next - p_curr

            # Normalize vectors (in XZ plane, ignore Y)
            v1_norm = np.array([v1[0], 0, v1[2]])
            v2_norm = np.array([v2[0], 0, v2[2]])

            v1_len = np.linalg.norm(v1_norm)
            v2_len = np.linalg.norm(v2_norm)

            if v1_len > 1e-6:
                v1_norm /= v1_len
            if v2_len > 1e-6:
                v2_norm /= v2_len

            # Angle bisector
            bisector = v1_norm + v2_norm
            bisector_len = np.linalg.norm(bisector)

            if bisector_len > 1e-6:
                bisector /= bisector_len

                # Move point inward
                inset_point = p_curr - bisector * inset_distance
                inset_points.append([inset_point[0], inset_point[1], inset_point[2]])
            else:
                inset_points.append(points[i])

        return inset_points

    def _generate_floor_ceiling(self, mask, scale, wall_height, vertex_offset):
        """Generate floor and ceiling geometry from segmentation mask"""
        vertices = []
        faces = []

        # Find floor regions (assuming class 0 = floor/room)
        floor_mask = (mask == 0).astype(np.uint8)

        # Find contours
        contours, _ = cv2.findContours(
            floor_mask, cv2.RETR_EXTERNAL, cv2.CHAIN_APPROX_SIMPLE
        )

        for contour in contours:
            if cv2.contourArea(contour) < 100:  # Skip small regions
                continue

            # Simplify contour
            epsilon = 0.02 * cv2.arcLength(contour, True)
            approx = cv2.approxPolyDP(contour, epsilon, True)

            if len(approx) < 3:
                continue

            # Convert to 3D coordinates
            floor_points = []
            for point in approx.reshape(-1, 2):
                x = (point[0] - 128) * scale
                z = (128 - point[1]) * scale
```

```
                floor_points.append([x, 0, z])  # Floor at y=0

            ceiling_points = []
            for point in approx.reshape(-1, 2):
                x = (point[0] - 128) * scale
                z = (128 - point[1]) * scale
                ceiling_points.append([x, wall_height, z])  # Ceiling at y=wall_height

            # Add vertices
            n_points = len(floor_points)
            vertices.extend(floor_points)
            vertices.extend(ceiling_points)

            # Triangulate floor
            if n_points >= 3:
                for i in range(1, n_points - 1):
                    faces.append(
                        [vertex_offset, vertex_offset + i + 1, vertex_offset + i]
                    )

                # Triangulate ceiling (reverse winding)
                for i in range(1, n_points - 1):
                    faces.append(
                        [
                            vertex_offset + n_points,
                            vertex_offset + n_points + i,
                            vertex_offset + n_points + i + 1,
                        ]
                    )

            vertex_offset += 2 * n_points

        return vertices, faces

    def process_batch(self, image_paths, output_dir):
        """Process multiple images in batch"""
        output_dir = Path(output_dir)
        output_dir.mkdir(exist_ok=True)

        results = []

        for img_path in image_paths:
            img_path = Path(img_path)
            print(f"Processing: {img_path.name}")

            output_path = output_dir / f"{img_path.stem}_model.obj"

            try:
                success = self.generate_3d_model(
                    str(img_path), str(output_path), export_intermediate=True
                )

                results.append({
                    "input": str(img_path),
                    "output": str(output_path),
                    "success": success
                })

                if success:
                    print(f"? Generated: {output_path}")
                else:
                    print(f"? Failed: {img_path.name}")

            except Exception as e:
                print(f"? Error processing {img_path.name}: {str(e)}")
                results.append({
                    "input": str(img_path),
                    "output": str(output_path),
                    "success": False,
                    "error": str(e)
                })

        return results
```

■ File: models\dvx.py

==

```
"""
Robust Differentiable Vectorization (DVX) module.

Improvements vs naive DVX:
- Projects backbone feature maps to `feature_dim` if channels don't match via 1x1 conv.
- Multi-step iterative refinement (improves final polygon accuracy).
- Safe guards for shapes, device handling, and grid-sampling.
- Returns init_polygons, final polygons, per-step displacements, and validity scores.

Usage:
- features: dict of feature maps (e.g. "p2", "p4"), each tensor (B, C, H, W).
- segmentation: (B, 1, H_img, W_img) or similar ? only used for optional initialization logic.
"""

from typing import Dict, Any, Optional, Tuple
import torch
import torch.nn as nn
import torch.nn.functional as F


class DifferentiableVectorization(nn.Module):
    def __init__(
        self,
        max_polygons: int = 20,
        max_points: int = 50,
        feature_dim: int = 256,
        displacement_scale: float = 0.12,
        num_refinement_steps: int = 3,
        align_corners: bool = False,
        padding_mode: str = "border",  # options for grid_sample
        use_proj_conv: bool = True,
    ):
        """
        Args:
            max_polygons: maximum polygons to predict per image
            max_points: number of control points per polygon
            feature_dim: number of channels the DVX expects (will project backbone features to this)
            displacement_scale: multiplier for predicted displacement (tanh output)
            num_refinement_steps: how many iterative refinement steps to apply (>=1)
            align_corners: align_corners for F.grid_sample
            padding_mode: padding_mode for F.grid_sample
            use_proj_conv: whether to use 1x1 conv to project backbone features to feature_dim (recommended
        """
        super().__init__()
        assert max_points > 2, "max_points must be > 2"
        assert num_refinement_steps >= 1

        self.max_polygons = int(max_polygons)
        self.max_points = int(max_points)
        self.feature_dim = int(feature_dim)
        self.displacement_scale = float(displacement_scale)
        self.num_refinement_steps = int(num_refinement_steps)
        self.align_corners = bool(align_corners)
        self.padding_mode = padding_mode
        self.use_proj_conv = bool(use_proj_conv)

        # init_net: from pooled p4 -> flattened -> produce normalized coords in [0,1]
        # AdaptiveAvgPool2d(8) -> (B, C, 8, 8) -> flatten -> Linear(C*8*8 -> hidden)
        hidden = max(512, feature_dim * 2)
        self.init_pool = nn.AdaptiveAvgPool2d(8)

        # we'll create a projector conv for p4/p2 channels if necessary at runtime
        # but also create an MLP init_net that assumes feature_dim channels after pooling
        self.init_mlp = nn.Sequential(
            nn.Flatten(),
            nn.Linear(self.feature_dim * 8 * 8, hidden),
            nn.ReLU(inplace=True),
            nn.Linear(hidden, 1024),
            nn.ReLU(inplace=True),
            nn.Linear(1024, self.max_polygons * self.max_points * 2),
            nn.Sigmoid(),
```

```
        )

        # refinement network: maps (feature_dim + 2) -> displacement in [-1,1]
        self.refine_net = nn.Sequential(
            nn.Linear(self.feature_dim + 2, 256),
            nn.ReLU(inplace=True),
            nn.Linear(256, 128),
            nn.ReLU(inplace=True),
            nn.Linear(128, 2),
            nn.Tanh(),
        )

        # validity net (reads flattened coords only)
        self.validity_net = nn.Sequential(
            nn.Linear(self.max_points * 2, 128),
            nn.ReLU(inplace=True),
            nn.Linear(128, 1),
            nn.Sigmoid(),
        )

        # projector convs (create lazily when first seen a feature channel mismatch)
        # stored per-key: e.g., self._proj_convs['p2'] = nn.Conv2d(in_ch, feature_dim, 1)
        self._proj_convs = nn.ModuleDict()
        self._proj_created = set()

    def _ensure_projector(self, key: str, in_channels: int):
        """
        Ensure a 1x1 conv exists that projects `in_channels` -> self.feature_dim for feature map `key`.
        """
        if not self.use_proj_conv:
            return None
        if key in self._proj_created:
            return self._proj_convs[key]

        if in_channels != self.feature_dim:
            conv = nn.Conv2d(in_channels, self.feature_dim, kernel_size=1, stride=1, padding=0)
            # initialize conv: kaiming
            nn.init.kaiming_normal_(conv.weight, a=0.2)
            if conv.bias is not None:
                nn.init.zeros_(conv.bias)
            self._proj_convs[key] = conv
        else:
            # identity mapping using 1x1 conv with weights = identity-like is tricky
            # Instead simply keep no conv; we'll pass feature as-is
            self._proj_convs[key] = nn.Identity()
        self._proj_created.add(key)
        return self._proj_convs[key]

    def _project_feature(self, key: str, feat: torch.Tensor) -> torch.Tensor:
        """
        Project or verify feature map to have self.feature_dim channels.
        If projector conv wasn't present and channels == feature_dim, returns feat unchanged.
        """
        in_ch = feat.shape[1]
        proj = self._ensure_projector(key, in_ch)
        if proj is None:
            # projection not desired; assert channels match
            if in_ch != self.feature_dim:
                raise RuntimeError(
                    f"Feature '{key}' channels ({in_ch}) != feature_dim ({self.feature_dim}) "
                    "and projection disabled."
                )
            return feat
        # if proj is Identity, apply it still (fast path)
        return proj(feat)

    def forward(
        self,
        features: Dict[str, torch.Tensor],
        segmentation: Optional[torch.Tensor] = None,
        return_all_steps: bool = False,
    ) -> Dict[str, Any]:
        """
```

```
        features: dict with keys like "p2", "p4" containing tensors (B, C, H, W)
        segmentation: optional (B, 1, H_img, W_img) or similar (not strictly required)
        return_all_steps: if True returns per-step intermediate polygons & displacements
        """
        # pick features for init and refinement
        p4 = features.get("p4", None)
        p2 = features.get("p2", None)

        if p4 is None and p2 is None:
            raise ValueError("At least one of 'p4' or 'p2' must be present in features.")

        # prefer p4 for init; fallback to p2 if not present
        init_feat = p4 if p4 is not None else p2
        refine_feat = p2 if p2 is not None else p4

        B = init_feat.shape[0]

        # Project features to feature_dim (if needed)
        init_feat = self._project_feature("p4_init", init_feat)
        refine_feat = self._project_feature("p2_refine", refine_feat)

        # -- Initialize polygons --
        # Pool then MLP; ensure init_mlp expects feature_dim channels
        pooled = self.init_pool(init_feat)  # [B, C', 8, 8]
        if pooled.shape[1] != self.feature_dim:
            # If the projector returned Identity but pooled channels mismatch, try to apply a runtime proje
            pooled = self._project_feature("p4_init_postpool", pooled)

        init_logits = self.init_mlp(pooled)  # [B, max_polygons * max_points * 2]
        init_polygons = init_logits.view(B, self.max_polygons, self.max_points, 2)  # normalized [0,1]

        # Iterative refinement
        polygons = init_polygons.clone()
        per_step_displacements = []
        for step in range(self.num_refinement_steps):
            # sample features at the polygon control-point locations
            displ = self._single_refine_step(polygons, refine_feat)
            per_step_displacements.append(displ)
            polygons = torch.clamp(polygons + displ, 0.0, 1.0)

        # final validity
        validity = self._predict_validity(polygons)

        out: Dict[str, Any] = {
            "polygons": polygons,  # [B, P, N, 2]
            "validity": validity,  # [B, P]
            "init_polygons": init_polygons,
            "refinement_displacements": per_step_displacements,  # list of [B, P, N, 2]
        }

        if return_all_steps:
            out["all_step_polygons"] = [
                torch.clamp(init_polygons + sum(per_step_displacements[:i + 1]), 0.0, 1.0)
                for i in range(len(per_step_displacements))
            ]

        return out

    def _single_refine_step(self, polygons: torch.Tensor, feature_map: torch.Tensor) -> torch.Tensor:
        """
        One refinement step: sample features at polygon points, predict displacement (scaled), return displ
        polygons: [B, P, N, 2] in [0,1]
        feature_map: [B, C, H, W] with C == feature_dim (or projected)
        returns displacement: [B, P, N, 2] in [-displacement_scale, displacement_scale]
        """
        B, P, N, _ = polygons.shape
        # flatten pts to sample
        coords = polygons.view(B, -1, 2)  # [B, P*N, 2], coords in [0,1]
        grid = coords * 2.0 - 1.0  # to [-1,1]
        # grid_sample expects (B, H_out, W_out, 2); use W_out=1
        grid_sample = grid.view(B, -1, 1, 2)
        sampled = F.grid_sample(
            feature_map,
```

```
            grid_sample,
            mode="bilinear",
            padding_mode=self.padding_mode,
            align_corners=self.align_corners,
        )  # [B, C, P*N, 1]
        sampled = sampled.squeeze(-1).permute(0, 2, 1).contiguous()  # [B, P*N, C]

        # combine sampled features and coords (coords in [0,1])
        input_feats = torch.cat([sampled, coords], dim=-1)  # [B, P*N, C+2]
        # predict displacements in [-1,1] via tanh on last layer
        disp = self.refine_net(input_feats)  # [B, P*N, 2], values ~[-1,1]
        disp = disp.view(B, P, N, 2)
        disp = disp * self.displacement_scale  # scale
        return disp

    def _predict_validity(self, polygons: torch.Tensor) -> torch.Tensor:
        B, P, N, _ = polygons.shape
        if N != self.max_points:
            # If someone truncated or padded points, adapt: flatten to last dim whatever it is
            poly_flat = polygons.view(B * P, -1)
        else:
            poly_flat = polygons.view(B * P, -1)
        validity = self.validity_net(poly_flat)  # [B*P, 1]
        validity = validity.view(B, P)
        return validity


# ------------------ quick unit test / smoke test ------------------
def _smoke_test():
    torch.manual_seed(0)
    B = 2
    C1 = 384  # different from feature_dim to test projector conv
    C2 = 128
    H2, W2 = 64, 64
    H4, W4 = 16, 16

    # create dummy backbone features with different channels
    p2 = torch.randn(B, C1, H2, W2)
    p4 = torch.randn(B, C2, H4, W4)
    seg = torch.rand(B, 1, H2 * 4, W2 * 4)  # just a placeholder

    dvx = DifferentiableVectorization(
        max_polygons=4,
        max_points=16,
        feature_dim=256,
        displacement_scale=0.08,
        num_refinement_steps=3,
        align_corners=False,
        padding_mode="border",
        use_proj_conv=True,
    )

    # ensure module moves projector convs to device when dvx.to(device) called
    dvx = dvx.eval()  # inference mode ok
    # Forward pass
    out = dvx({"p2": p2, "p4": p4}, seg, return_all_steps=True)
    print("polygons shape:", out["polygons"].shape)  # expected [B, P, N, 2]
    print("validity shape:", out["validity"].shape)  # expected [B, P]
    print("init shape:", out["init_polygons"].shape)
    print("refinement steps:", len(out["refinement_displacements"]))
    # check ranges
    assert out["polygons"].min().item() >= 0.0 - 1e-6
    assert out["polygons"].max().item() <= 1.0 + 1e-6
    print("smoke test passed")


if __name__ == "__main__":
    _smoke_test()
```

■ File: models\encoder.py

==

```
"""
Encoder architecture for multi-scale feature extraction
"""

import torch
import torch.nn as nn
import torch.nn.functional as F


class ResidualBlock(nn.Module):
    """Basic residual block for the encoder"""

    def __init__(self, in_channels, out_channels, stride=1):
        super().__init__()

        self.conv1 = nn.Conv2d(in_channels, out_channels, 3, stride, 1, bias=False)
        self.bn1 = nn.BatchNorm2d(out_channels)
        self.conv2 = nn.Conv2d(out_channels, out_channels, 3, 1, 1, bias=False)
        self.bn2 = nn.BatchNorm2d(out_channels)

        self.shortcut = nn.Sequential()
        if stride != 1 or in_channels != out_channels:
            self.shortcut = nn.Sequential(
                nn.Conv2d(in_channels, out_channels, 1, stride, bias=False),
                nn.BatchNorm2d(out_channels),
            )

    def forward(self, x):
        out = F.relu(self.bn1(self.conv1(x)))
        out = self.bn2(self.conv2(out))
        out += self.shortcut(x)
        return F.relu(out)


class MultiScaleEncoder(nn.Module):
    """
    Advanced encoder with skip connections and multi-scale feature extraction
    Based on ResNet architecture with Feature Pyramid Network (FPN)
    """

    def __init__(self, input_channels=3, feature_dim=512):
        super().__init__()

        # Stem
        self.stem = nn.Sequential(
            nn.Conv2d(input_channels, 64, 7, 2, 3, bias=False),
            nn.BatchNorm2d(64),
            nn.ReLU(inplace=True),
            nn.MaxPool2d(3, 2, 1),
        )

        # ResNet blocks
        self.layer1 = self._make_layer(64, 64, 2, stride=1)  # 64x64
        self.layer2 = self._make_layer(64, 128, 2, stride=2)  # 32x32
        self.layer3 = self._make_layer(128, 256, 2, stride=2)  # 16x16
        self.layer4 = self._make_layer(256, 512, 2, stride=2)  # 8x8

        # FPN lateral connections
        self.lateral4 = nn.Conv2d(512, feature_dim, 1)
        self.lateral3 = nn.Conv2d(256, feature_dim, 1)
        self.lateral2 = nn.Conv2d(128, feature_dim, 1)
        self.lateral1 = nn.Conv2d(64, feature_dim, 1)

        # FPN output layers
        self.smooth4 = nn.Conv2d(feature_dim, feature_dim, 3, 1, 1)
        self.smooth3 = nn.Conv2d(feature_dim, feature_dim, 3, 1, 1)
        self.smooth2 = nn.Conv2d(feature_dim, feature_dim, 3, 1, 1)
        self.smooth1 = nn.Conv2d(feature_dim, feature_dim, 3, 1, 1)

        # Global context
        self.global_pool = nn.AdaptiveAvgPool2d(1)
```

```python
        self.global_fc = nn.Sequential(
            nn.Linear(512, feature_dim),
            nn.ReLU(),
            nn.Linear(feature_dim, feature_dim)
        )

    def _make_layer(self, in_channels, out_channels, blocks, stride=1):
        layers = []
        layers.append(ResidualBlock(in_channels, out_channels, stride))
        for _ in range(1, blocks):
            layers.append(ResidualBlock(out_channels, out_channels))
        return nn.Sequential(*layers)

    def forward(self, x):
        # Bottom-up pathway
        x = self.stem(x)  # 64x64

        c1 = self.layer1(x)  # 64x64
        c2 = self.layer2(c1)  # 32x32
        c3 = self.layer3(c2)  # 16x16
        c4 = self.layer4(c3)  # 8x8

        # Global context
        global_feat = self.global_pool(c4).flatten(1)
        global_feat = self.global_fc(global_feat)

        # Top-down pathway (FPN)
        p4 = self.lateral4(c4)
        p3 = self.lateral3(c3) + F.interpolate(p4, scale_factor=2)
        p2 = self.lateral2(c2) + F.interpolate(p3, scale_factor=2)
        p1 = self.lateral1(c1) + F.interpolate(p2, scale_factor=2)

        # Smooth
        p4 = self.smooth4(p4)
        p3 = self.smooth3(p3)
        p2 = self.smooth2(p2)
        p1 = self.smooth1(p1)

        return {
            "p1": p1,  # 64x64
            "p2": p2,  # 32x32
            "p3": p3,  # 16x16
            "p4": p4,  # 8x8
            "global": global_feat,
        }
```

■ File: models\extrusion.py

==

```python
"""
Vectorized Differentiable 3D extrusion module for converting polygons to 3D occupancy
Optimized version with GPU-accelerated vectorized operations
"""

import torch
import torch.nn as nn
import torch.nn.functional as F
import math
import logging


# ---------------------------------------------------------------------------
# Logging and sanitization helper
# ---------------------------------------------------------------------------
logger = logging.getLogger(__name__)
if not logger.handlers:
    handler = logging.StreamHandler()
    handler.setFormatter(logging.Formatter("%(asctime)s | %(levelname)s | %(message)s"))
    logger.addHandler(handler)
logger.setLevel(logging.INFO)


```

```
def _sanitize_normalized_height(value, sample_id=None, default=0.6):
    """
    Ensure normalized height value is finite and in [0,1].
    Returns a float in [0,1].

    Args:
        value: torch scalar tensor or float
        sample_id: optional identifier for logging (string or int)
        default: fallback normalized height
    """
    try:
        if isinstance(value, torch.Tensor):
            raw = float(value.item())
        else:
            raw = float(value)
    except Exception:
        raw = float("nan")

    # Build label for logging
    sid = f"[sample={sample_id}]" if sample_id is not None else ""

    # Check finite
    if not math.isfinite(raw):
        logger.warning(f"{sid} Invalid wall height value (not finite): {raw}; using default {default}")
        raw = default

    # Clamp to [0,1]
    if raw < 0.0 or raw > 1.0:
        logger.warning(f"{sid} Wall height normalized {raw} out of [0,1]; clamping.")
        raw = max(0.0, min(1.0, raw))

    return raw


def _sanitize_tensor(tensor, default_value=0.0, name="tensor"):
    """
    Sanitize an entire tensor by replacing NaN/Inf values with default.

    Args:
        tensor: Input tensor
        default_value: Value to replace invalid entries with
        name: Name for logging

    Returns:
        Sanitized tensor
    """
    if tensor.numel() == 0:
        return tensor

    # Check for any invalid values
    invalid_mask = ~torch.isfinite(tensor)
    num_invalid = invalid_mask.sum().item()

    if num_invalid > 0:
        logger.warning(f"Found {num_invalid} invalid values in {name}, replacing with {default_value}")
        tensor = tensor.clone()
        tensor[invalid_mask] = default_value

    return tensor


# ---------------------------------------------------------------------------
# Main extrusion module
# ---------------------------------------------------------------------------
class DifferentiableExtrusion(nn.Module):
    """
    Vectorized Differentiable 3D extrusion module
    Converts polygons + attributes to soft 3D occupancy grids
    """

    def __init__(self, voxel_size: int = 64):
        super().__init__()
        self.voxel_size = int(voxel_size)
```

```
        self.register_buffer("_coords", None)

    def _ensure_coords(self, device):
        """Initialize or update coordinate grid if needed"""
        if (self._coords is None or
            self._coords.device != device or
            self._coords.shape[0] != (self.voxel_size * self.voxel_size)):

            H = W = self.voxel_size
            y, x = torch.meshgrid(
                torch.arange(H, device=device),
                torch.arange(W, device=device),
                indexing="ij"
            )
            coords = torch.stack([x.flatten().float(), y.flatten().float()], dim=1)  # [H*W, 2]
            coords = coords / float(self.voxel_size - 1)
            self.register_buffer("_coords", coords)

    def polygon_sdf(self, polygon_xy):
        """
        Compute signed distance field for a polygon using vectorized operations.
        """
        device = polygon_xy.device
        self._ensure_coords(device)
        pts = self._coords  # [M, 2]
        P = polygon_xy.shape[0]

        if P < 2:
            return torch.full((pts.shape[0],), 1.0, device=device)

        # Sanitize polygon coordinates
        polygon_xy = _sanitize_tensor(polygon_xy, default_value=0.0, name="polygon_xy")

        v0 = polygon_xy.unsqueeze(1)
        v1 = torch.roll(polygon_xy, shifts=-1, dims=0).unsqueeze(1)
        pts_exp = pts.unsqueeze(0)

        e = v1 - v0
        v = pts_exp - v0
        e_norm_sq = (e**2).sum(dim=2, keepdim=True) + 1e-8
        t = (v * e).sum(dim=2, keepdim=True) / e_norm_sq
        t_clamped = t.clamp(0.0, 1.0)

        proj = v0 + t_clamped * e
        diff = pts_exp - proj
        dists = torch.norm(diff, dim=2)

        # Sanitize distances before min operation
        dists = _sanitize_tensor(dists, default_value=1.0, name="distances")
        min_dist_per_point, _ = dists.min(dim=0)

        x_pts = pts[:, 0].unsqueeze(0)
        y_pts = pts[:, 1].unsqueeze(0)
        x0, y0 = v0[..., 0], v0[..., 1]
        x1, y1 = v1[..., 0], v1[..., 1]

        y_crosses = ((y0 <= y_pts) & (y1 > y_pts)) | ((y1 <= y_pts) & (y0 > y_pts))
        inter_x = x0 + (x1 - x0) * ((y_pts - y0) / (y1 - y0 + 1e-8))
        crossings = (inter_x > x_pts) & y_crosses
        crossing_count = crossings.sum(dim=0)
        inside = (crossing_count % 2 == 1)

        sdf = min_dist_per_point.clone()
        sdf[inside] = -sdf[inside]

        # Final sanitization of SDF output
        sdf = _sanitize_tensor(sdf, default_value=1.0, name="sdf")
        return sdf

    def forward(self, polygons, attributes, validity_scores, sample_ids=None):
        """
        Convert polygons to 3D voxel occupancy.
        sample_ids: optional list/array of identifiers (e.g., filenames or dataset indices)
```

```python
        """
        device = polygons.device
        B, N, P, _ = polygons.shape
        D = H = W = self.voxel_size

        # Sanitize input tensors
        polygons = _sanitize_tensor(polygons, default_value=0.0, name="input_polygons")
        attributes = _sanitize_tensor(attributes, default_value=0.6, name="input_attributes")
        validity_scores = _sanitize_tensor(validity_scores, default_value=0.0, name="input_validity_scores"

        voxels = torch.zeros((B, D, H, W), device=device)

        for b in range(B):
            # pick identifier if available
            sid = sample_ids[b] if sample_ids is not None else b

            # Sanitize height with logging
            wall_height_normalized = attributes[b, 0]
            sanitized_norm = _sanitize_normalized_height(
                wall_height_normalized, sample_id=sid, default=0.6
            )

            wall_height_m = sanitized_norm * 5.0
            height_frac = wall_height_m / 5.0
            height_voxels = int(round(height_frac * D))
            height_voxels = max(1, min(D, height_voxels))

            # Process each polygon for this batch
            validity_mask = validity_scores[b] >= 0.5
            if not validity_mask.any():
                continue

            combined_mask = torch.zeros((H, W), device=device)
            sharpness = 100.0

            for n in range(N):
                if not validity_mask[n]:
                    continue

                polygon = polygons[b, n]  # [P, 2]

                # Filter out zero-padded vertices
                vertex_mask = (polygon.sum(dim=1) != 0.0)
                if vertex_mask.sum().item() < 3:
                    continue

                valid_polygon = polygon[vertex_mask]

                # Compute SDF for this polygon
                sdf = self.polygon_sdf(valid_polygon)
                mask = torch.sigmoid(-sdf * sharpness)
                mask_2d = mask.view(H, W)

                # Sanitize mask before combining
                mask_2d = _sanitize_tensor(mask_2d, default_value=0.0, name=f"mask_2d_b{b}_n{n}")
                combined_mask = torch.maximum(combined_mask, mask_2d)

            # Create 3D mask by extruding to the computed height
            mask_3d = combined_mask.unsqueeze(0).expand(height_voxels, -1, -1)

            # Sanitize final mask before assignment
            mask_3d = _sanitize_tensor(mask_3d, default_value=0.0, name=f"final_mask_3d_b{b}")
            voxels[b, :height_voxels] = mask_3d

        # Final sanitization of output
        voxels = _sanitize_tensor(voxels, default_value=0.0, name="output_voxels")
        return voxels


# ---------------------------------------------------------------------------
# Fast extrusion module
# ---------------------------------------------------------------------------
class DifferentiableExtrusionFast(nn.Module):
```

```python
    """
    Optimized version that batches polygon processing.
    """

    def __init__(self, voxel_size: int = 64):
        super().__init__()
        self.voxel_size = int(voxel_size)
        self.register_buffer("_coords", None)

    def _ensure_coords(self, device):
        if (self._coords is None or
            self._coords.device != device or
            self._coords.shape[0] != (self.voxel_size * self.voxel_size)):

            H = W = self.voxel_size
            y, x = torch.meshgrid(
                torch.arange(H, device=device),
                torch.arange(W, device=device),
                indexing="ij"
            )
            coords = torch.stack([x.flatten().float(), y.flatten().float()], dim=1)
            coords = coords / float(self.voxel_size - 1)
            self.register_buffer("_coords", coords)

    def batch_polygon_sdf(self, polygons_batch, validity_mask):
        device = polygons_batch.device
        self._ensure_coords(device)

        N, P, _ = polygons_batch.shape
        M = self._coords.shape[0]
        sdfs = torch.full((N, M), 1.0, device=device)

        valid_indices = torch.where(validity_mask)[0]
        if len(valid_indices) == 0:
            return sdfs

        valid_polygons = polygons_batch[valid_indices]
        for i, poly_idx in enumerate(valid_indices):
            poly = valid_polygons[i]
            vertex_mask = (poly.sum(dim=1) != 0.0)
            if vertex_mask.sum().item() >= 3:
                valid_poly = poly[vertex_mask]
                sdf = self.polygon_sdf(valid_poly)
                sdfs[poly_idx] = sdf

        return sdfs

    def polygon_sdf(self, polygon_xy):
        device = polygon_xy.device
        self._ensure_coords(device)
        pts = self._coords
        P = polygon_xy.shape[0]

        if P < 2:
            return torch.full((pts.shape[0],), 1.0, device=device)

        # Sanitize polygon coordinates
        polygon_xy = _sanitize_tensor(polygon_xy, default_value=0.0, name="polygon_xy_fast")

        v0 = polygon_xy.unsqueeze(1)
        v1 = torch.roll(polygon_xy, shifts=-1, dims=0).unsqueeze(1)
        pts_exp = pts.unsqueeze(0)

        e = v1 - v0
        v = pts_exp - v0
        e_norm_sq = (e**2).sum(dim=2, keepdim=True) + 1e-8
        t = (v * e).sum(dim=2, keepdim=True) / e_norm_sq
        t_clamped = t.clamp(0.0, 1.0)

        proj = v0 + t_clamped * e
        diff = pts_exp - proj
        dists = torch.norm(diff, dim=2)

```

```
        # Sanitize distances before min operation
        dists = _sanitize_tensor(dists, default_value=1.0, name="distances_fast")
        min_dist_per_point, _ = dists.min(dim=0)

        x_pts = pts[:, 0].unsqueeze(0)
        y_pts = pts[:, 1].unsqueeze(0)
        x0, y0 = v0[..., 0], v0[..., 1]
        x1, y1 = v1[..., 0], v1[..., 1]

        y_crosses = ((y0 <= y_pts) & (y1 > y_pts)) | ((y1 <= y_pts) & (y0 > y_pts))
        inter_x = x0 + (x1 - x0) * ((y_pts - y0) / (y1 - y0 + 1e-8))
        crossings = (inter_x > x_pts) & y_crosses
        crossing_count = crossings.sum(dim=0)
        inside = (crossing_count % 2 == 1)

        sdf = min_dist_per_point.clone()
        sdf[inside] = -sdf[inside]

        # Final sanitization of SDF output
        sdf = _sanitize_tensor(sdf, default_value=1.0, name="sdf_fast")
        return sdf

    def forward(self, polygons: torch.Tensor, attributes: torch.Tensor, validity_scores: torch.Tensor) -> t
        device = polygons.device
        B, N, P, _ = polygons.shape
        D = H = W = self.voxel_size

        # Sanitize input tensors
        polygons = _sanitize_tensor(polygons, default_value=0.0, name="input_polygons_fast")
        attributes = _sanitize_tensor(attributes, default_value=0.6, name="input_attributes_fast")
        validity_scores = _sanitize_tensor(validity_scores, default_value=0.0, name="input_validity_scores_

        voxels = torch.zeros((B, D, H, W), device=device)

        for b in range(B):
            validity_mask = validity_scores[b] >= 0.5
            if not validity_mask.any():
                continue

            sdfs = self.batch_polygon_sdf(polygons[b], validity_mask)

            # Sanitize SDFs before sigmoid
            sdfs = _sanitize_tensor(sdfs, default_value=1.0, name=f"batch_sdfs_b{b}")

            sharpness = 100.0
            masks = torch.sigmoid(-sdfs * sharpness)
            masks_2d = masks.view(N, H, W)

            # Sanitize masks
            masks_2d = _sanitize_tensor(masks_2d, default_value=0.0, name=f"masks_2d_b{b}")

            # Sanitize height
            wall_height_normalized = attributes[b, 0]
            sanitized_norm = _sanitize_normalized_height(wall_height_normalized, sample_id=b, default=0.6)
            wall_height_m = sanitized_norm * 5.0
            height_frac = wall_height_m / 5.0
            height_voxels = int(round(height_frac * D))
            height_voxels = max(1, min(D, height_voxels))

            combined_mask = torch.zeros((H, W), device=device)
            for n in range(N):
                if validity_mask[n]:
                    combined_mask = torch.maximum(combined_mask, masks_2d[n])

            mask_3d = combined_mask.unsqueeze(0).expand(height_voxels, -1, -1)

            # Sanitize final mask before assignment
            mask_3d = _sanitize_tensor(mask_3d, default_value=0.0, name=f"final_mask_3d_fast_b{b}")
            voxels[b, :height_voxels] = mask_3d

        # Final sanitization of output
        voxels = _sanitize_tensor(voxels, default_value=0.0, name="output_voxels_fast")
        return voxels
```

■ File: models\heads.py

==

```python
"""
Multi-task prediction heads for the Neural-Geometric 3D Model Generator
"""

import torch
import torch.nn as nn
import torch.nn.functional as F


class SegmentationHead(nn.Module):
    """Semantic segmentation head with multi-scale fusion"""

    def __init__(self, feature_dim=512, num_classes=5, dropout=0.1):
        super().__init__()

        # Multi-scale fusion
        self.fusion = nn.Sequential(
            nn.Conv2d(feature_dim * 4, feature_dim, 3, 1, 1),
            nn.BatchNorm2d(feature_dim),
            nn.ReLU(),
            nn.Dropout2d(dropout),
        )

        # Segmentation decoder
        self.decoder = nn.Sequential(
            nn.Conv2d(feature_dim, feature_dim // 2, 3, 1, 1),
            nn.BatchNorm2d(feature_dim // 2),
            nn.ReLU(),
            nn.Conv2d(feature_dim // 2, feature_dim // 4, 3, 1, 1),
            nn.BatchNorm2d(feature_dim // 4),
            nn.ReLU(),
            nn.Conv2d(feature_dim // 4, num_classes, 1),
        )

    def forward(self, features):
        # Fuse multi-scale features
        p1, p2, p3, p4 = features["p1"], features["p2"], features["p3"], features["p4"]

        # Upsample all to p1 resolution
        p2_up = F.interpolate(
            p2, size=p1.shape[-2:], mode="bilinear", align_corners=False
        )
        p3_up = F.interpolate(
            p3, size=p1.shape[-2:], mode="bilinear", align_corners=False
        )
        p4_up = F.interpolate(
            p4, size=p1.shape[-2:], mode="bilinear", align_corners=False
        )

        fused = torch.cat([p1, p2_up, p3_up, p4_up], dim=1)
        fused = self.fusion(fused)

        # Final segmentation
        seg = self.decoder(fused)
        return F.interpolate(seg, scale_factor=4, mode="bilinear", align_corners=False)


class AttributeHead(nn.Module):
    """Attribute regression head for geometric parameters"""

    def __init__(self, feature_dim=512, num_attributes=6, dropout=0.2):
        super().__init__()

        self.regressor = nn.Sequential(
            nn.Linear(feature_dim, feature_dim),
            nn.ReLU(),
            nn.Dropout(dropout),
            nn.Linear(feature_dim, feature_dim // 2),
            nn.ReLU(),
            nn.Dropout(dropout),
            nn.Linear(feature_dim // 2, num_attributes),
```

```python
            nn.Sigmoid(),  # Output in [0,1] range
        )

    def forward(self, global_features):
        return self.regressor(global_features)


class SDFHead(nn.Module):
    """Signed Distance Field prediction for sharp boundaries"""

    def __init__(self, feature_dim=512, dropout=0.1):
        super().__init__()

        self.sdf_decoder = nn.Sequential(
            nn.Conv2d(feature_dim, feature_dim // 2, 3, 1, 1),
            nn.BatchNorm2d(feature_dim // 2),
            nn.ReLU(),
            nn.Dropout2d(dropout),
            nn.Conv2d(feature_dim // 2, feature_dim // 4, 3, 1, 1),
            nn.BatchNorm2d(feature_dim // 4),
            nn.ReLU(),
            nn.Conv2d(feature_dim // 4, 1, 1),
            nn.Tanh(),  # SDF in [-1, 1]
        )

    def forward(self, features):
        # Use highest resolution features
        p1 = features["p1"]
        sdf = self.sdf_decoder(p1)
        return F.interpolate(sdf, scale_factor=4, mode="bilinear", align_corners=False)
```

■ File: models\model.py

==

```python
"""
Advanced loss functions for multi-task training with dynamic weighting
Enhanced with cross-modal consistency, graph constraints, and GradNorm
Modified to support conditional geometric losses via run_full_geometric flag
"""
import torch
import torch.nn as nn
import torch.nn.functional as F
from .encoder import MultiScaleEncoder
from .heads import SegmentationHead, AttributeHead, SDFHead
from .dvx import DifferentiableVectorization
from .extrusion import DifferentiableExtrusion


class L2Normalize(nn.Module):
    """L2 normalization layer"""

    def __init__(self, dim=1):
        super().__init__()
        self.dim = dim

    def forward(self, x):
        return F.normalize(x, p=2, dim=self.dim)


class LatentEmbeddingHead(nn.Module):
    """Auxiliary head for cross-modal latent consistency"""

    def __init__(self, feature_dim: int, embedding_dim: int = 256):
        super().__init__()
        self.embedding_dim = embedding_dim

        # 2D embedding path
        self.embedding_2d = nn.Sequential(
            nn.AdaptiveAvgPool2d((1, 1)),
            nn.Flatten(),
            nn.Linear(feature_dim, embedding_dim * 2),
            nn.ReLU(),
```

```
            nn.Dropout(0.1),
            nn.Linear(embedding_dim * 2, embedding_dim),
            L2Normalize(dim=1),  # L2 normalize for cosine similarity
        )

        # 3D embedding path (from voxel features)
        self.embedding_3d = nn.Sequential(
            nn.AdaptiveAvgPool3d((1, 1, 1)),
            nn.Flatten(),
            nn.Linear(feature_dim, embedding_dim * 2),
            nn.ReLU(),
            nn.Dropout(0.1),
            nn.Linear(embedding_dim * 2, embedding_dim),
            L2Normalize(dim=1),
        )

    def forward(
        self, features_2d: torch.Tensor, features_3d: torch.Tensor = None
    ) -> tuple:
        """
        Generate 2D and 3D embeddings for consistency loss

        Args:
            features_2d: [B, C, H, W] - 2D feature maps
            features_3d: [B, C, D, H, W] - 3D feature maps (optional)

        Returns:
            tuple: (embedding_2d, embedding_3d)
        """
        # 2D embedding
        emb_2d = self.embedding_2d(features_2d)

        # 3D embedding (if available, otherwise use 2D features reshaped)
        if features_3d is not None:
            emb_3d = self.embedding_3d(features_3d)
        else:
            # Create pseudo-3D from 2D features
            B, C, H, W = features_2d.shape
            pseudo_3d = features_2d.unsqueeze(2).expand(
                B, C, 4, H, W
            )  # Duplicate across depth
            emb_3d = self.embedding_3d(pseudo_3d)

        return emb_2d, emb_3d


class GraphStructureHead(nn.Module):
    """Head for predicting graph structure (room connectivity)"""

    def __init__(self, feature_dim: int, max_rooms: int = 10):
        super().__init__()
        self.max_rooms = max_rooms

        # Room detection branch
        self.room_detector = nn.Sequential(
            nn.Conv2d(feature_dim, feature_dim // 2, 3, padding=1),
            nn.ReLU(),
            nn.Conv2d(feature_dim // 2, max_rooms, 3, padding=1),
            nn.Sigmoid(),  # Room probability maps
        )

        # Room feature extractor
        self.room_features = nn.Sequential(
            nn.AdaptiveAvgPool2d((8, 8)),  # Pool to fixed size
            nn.Flatten(),
            nn.Linear(feature_dim * 64, 256),
            nn.ReLU(),
            nn.Linear(256, 128),  # Room feature vectors
        )

        # Adjacency predictor
        self.adjacency_net = nn.Sequential(
            nn.Linear(128 * 2, 64),  # Pairwise room features
```

```
            nn.ReLU(),
            nn.Linear(64, 32),
            nn.ReLU(),
            nn.Linear(32, 1),
            nn.Sigmoid(),  # Adjacency probability
        )

    def forward(self, features: torch.Tensor) -> dict:
        """
        Predict room graph structure

        Args:
            features: [B, C, H, W] - Feature maps

        Returns:
            dict with 'room_maps', 'room_features', 'adjacency_matrix'
        """
        B = features.shape[0]

        # Detect room probability maps
        room_maps = self.room_detector(features)  # [B, max_rooms, H, W]

        # Extract room features
        room_feats = self.room_features(features)  # [B, 128]

        # Create adjacency matrix for all room pairs
        adjacency_matrices = []

        for b in range(B):
            # Get room features for this batch item
            feat_b = room_feats[b : b + 1]  # [1, 128]

            # Create pairwise combinations
            adj_matrix = torch.zeros(
                (self.max_rooms, self.max_rooms), device=features.device
            )

            for i in range(self.max_rooms):
                for j in range(i + 1, self.max_rooms):
                    # Concatenate features for room pair
                    pair_feat = torch.cat([feat_b, feat_b], dim=1)  # [1, 256]

                    # Predict adjacency
                    adj_prob = self.adjacency_net(pair_feat)  # [1, 1]

                    # Fill symmetric matrix
                    adj_matrix[i, j] = adj_prob.squeeze()
                    adj_matrix[j, i] = adj_prob.squeeze()

            adjacency_matrices.append(adj_matrix)

        return {
            "room_maps": room_maps,
            "room_features": room_feats,
            "adjacency_matrices": torch.stack(adjacency_matrices),
        }


class NeuralGeometric3DGenerator(nn.Module):
    """
    Enhanced neural-geometric system with auxiliary heads for novel training strategies:
    - Cross-modal latent consistency
    - Graph structure prediction
    - Multi-view embeddings for dynamic curriculum
    - Conditional geometric computation via run_full_geometric flag
    """

    def __init__(
        self,
        input_channels=3,
        num_classes=5,
        feature_dim=512,
        num_attributes=6,
```

```python
        voxel_size=64,
        max_polygons=20,
        max_points=50,
        use_latent_consistency=True,
        use_graph_constraints=True,
        latent_embedding_dim=256,
    ):
        super().__init__()

        # Store configuration
        self.use_latent_consistency = use_latent_consistency
        self.use_graph_constraints = use_graph_constraints
        self.feature_dim = feature_dim

        # Core components
        self.encoder = MultiScaleEncoder(input_channels, feature_dim)
        self.seg_head = SegmentationHead(feature_dim, num_classes)
        self.attr_head = AttributeHead(feature_dim, num_attributes)
        self.sdf_head = SDFHead(feature_dim)
        self.dvx = DifferentiableVectorization(max_polygons, max_points, feature_dim)
        self.extrusion = DifferentiableExtrusion(voxel_size)

        # NEW: Auxiliary heads for novel training strategies
        if use_latent_consistency:
            self.latent_head = LatentEmbeddingHead(feature_dim, latent_embedding_dim)

        if use_graph_constraints:
            self.graph_head = GraphStructureHead(feature_dim)

        # Enhanced feature processing for multi-stage training
        self.feature_enhancer = nn.Sequential(
            nn.Conv2d(feature_dim, feature_dim, 3, padding=1),
            nn.GroupNorm(32, feature_dim),
            nn.ReLU(),
            nn.Conv2d(feature_dim, feature_dim, 3, padding=1),
            nn.GroupNorm(32, feature_dim),
        )

        # lazy-created 3d voxel processor will be attached on first use
        self._voxel_processor = None

    def _select_spatial_feature(self, features):
        """
        Given encoder output (dict or tensor), select a spatial 4-D feature map
        Prefer high-resolution feature maps (p1) and avoid selecting 'global' vector.
        """
        # If encoder returned a tensor already, make sure it's 4D
        if not isinstance(features, dict):
            if features.dim() == 4:
                return features
            else:
                raise ValueError(
                    f"Encoder returned a tensor with shape {tuple(features.shape)}; "
                    "expected a 4D feature map [B, C, H, W]."
                )

        # Encoder returned dict: prefer p1,p2,p3,p4,high_res,out,main but NEVER 'global'
        preferred_keys = ["p1", "p2", "p3", "p4", "high_res", "out", "main"]
        for k in preferred_keys:
            if k in features:
                candidate = features[k]
                if isinstance(candidate, torch.Tensor) and candidate.dim() == 4:
                    return candidate

        # As a last resort, scan dict values for the first 4D tensor that isn't 'global'
        for k, v in features.items():
            if k == "global":
                continue
            if isinstance(v, torch.Tensor) and v.dim() == 4:
                return v

        # If nothing found, raise informative error rather than silently picking wrong shape
        raise RuntimeError(
```

```python
            "No spatial 4D feature map found in encoder output. Encoder returned keys: "
            f"{list(features.keys())}. Ensure encoder provides at least one [B,C,H,W] tensor "
            "under keys like 'p1','p2','p3','p4','out', or 'high_res'."
        )

    def forward(self, image, run_full_geometric=True, return_aux=True):
        """
        Enhanced forward pass with auxiliary outputs and conditional geometric computation

        Args:
            image: [B, C, H, W] input images
            run_full_geometric: Whether to run heavy DVX and extrusion computations
            return_aux: Whether to compute auxiliary outputs

        Returns:
            dict with predictions, conditionally including geometric outputs
        """
        # Multi-scale feature extraction
        features = self.encoder(image)

        # Enhance features
        spatial_feat = self._select_spatial_feature(features)
        enhanced_features = self.feature_enhancer(spatial_feat)

        # keep structured features dict for heads that expect multi-scale inputs
        if isinstance(features, dict):
            features["enhanced"] = enhanced_features
            main_features = enhanced_features
        else:
            features = {"main": enhanced_features, "enhanced": enhanced_features}
            main_features = enhanced_features

        # Core predictions (always computed - these are fast)
        segmentation = self.seg_head(features)
        attributes = self.attr_head(
            features.get("global")
            if isinstance(features, dict) and "global" in features
            else main_features.mean(dim=[2, 3])
        )
        sdf = self.sdf_head(features)

        # Base outputs
        outputs = {
            "segmentation": segmentation,
            "attributes": attributes,
            "sdf": sdf,
            "features": features,
        }

        # Conditional geometric computation (heavy operations)
        if run_full_geometric:
            # DVX polygon fitting
            dvx_output = self.dvx(features, segmentation)
            polygons = dvx_output.get("polygons", None)
            validity = dvx_output.get("validity", None)

            # 3D extrusion (defensive: ensure inputs exist)
            try:
                voxels_pred = self.extrusion(polygons, attributes, validity)
            except Exception as e:
                # Log or print a helpful message for debugging; avoid crashing training
                # (Replace print with logger if you have one)
                print(f"[Warning] extrusion failed: {e}")
                voxels_pred = None

            # Add geometric outputs
            outputs.update({
                "polygons": polygons,
                "polygon_validity": validity,
                "voxels_pred": voxels_pred,
            })

            # NEW: Auxiliary outputs for novel training strategies (only when geometric is enabled)
```

```
            if return_aux:
                # Cross-modal consistency embeddings
                if self.use_latent_consistency:
                    if voxels_pred is not None:
                        voxel_features = self._create_3d_features_from_voxels(voxels_pred)
                        latent_2d, latent_3d = self.latent_head(main_features, voxel_features)
                    else:
                        # Fall back to pseudo-3D built from 2D features if voxels not available
                        latent_2d, latent_3d = self.latent_head(main_features, None)
                    outputs["latent_2d_embedding"] = latent_2d
                    outputs["latent_3d_embedding"] = latent_3d
        else:
            # Geometric path explicitly skipped for this stage
            outputs.update({
                "polygons": None,
                "polygon_validity": None,
                "voxels_pred": None,
            })

            # Still compute some auxiliary outputs that don't depend on geometry
            if return_aux and self.use_latent_consistency:
                # Use pseudo-3D features for 2D-only consistency inside latent head
                latent_2d, latent_3d = self.latent_head(main_features, None)
                outputs["latent_2d_embedding"] = latent_2d
                outputs["latent_3d_embedding"] = latent_3d

        # Graph structure predictions (independent of geometric computation)
        if return_aux and self.use_graph_constraints:
            graph_output = self.graph_head(main_features)
            outputs.update(graph_output)

        return outputs

    def get_latent_embeddings(self, image):
        """
        Convenience method to get just the latent embeddings
        Used by trainer for consistency loss
        """
        if not self.use_latent_consistency:
            return None, None

        with torch.no_grad():
            features = self.encoder(image)
            spatial_feat = self._select_spatial_feature(features)
            main_features = self.feature_enhancer(spatial_feat)

            # Quick forward to get segmentation/attributes
            segmentation = self.seg_head(features)
            attributes = self.attr_head(
                features.get("global")
                if isinstance(features, dict) and "global" in features
                else main_features.mean(dim=[2, 3])
            )

            # Attempt DVX + extrusion, but be defensive (may be expensive)
            dvx_output = self.dvx(features, segmentation)
            polygons = dvx_output.get("polygons", None)
            validity = dvx_output.get("validity", None)

            try:
                voxels_pred = self.extrusion(polygons, attributes, validity)
            except Exception as e:
                print(f"[Warning] get_latent_embeddings: extrusion failed: {e}")
                voxels_pred = None

            # If voxels not available, latent_head will fall back to pseudo-3D
            if voxels_pred is not None:
                voxel_features = self._create_3d_features_from_voxels(voxels_pred)
            else:
                voxel_features = None

        return self.latent_head(main_features, voxel_features)

```

```python
    def _create_3d_features_from_voxels(self, voxels):
        """
        Create 3D feature representation from voxel predictions

        Args:
            voxels: [B, D, H, W] voxel predictions

        Returns:
            [B, C, D, H, W] 3D features
        """
        # Defensive check
        if voxels is None:
            raise ValueError(
                "Received voxels=None in _create_3d_features_from_voxels(). "
                "This indicates that the geometric pipeline was skipped or extrusion failed. "
                "Call this method only when voxels are available, or use latent_head(..., None) to "
                "compute pseudo-3D features from 2D."
            )

        # Ensure expected shape
        if voxels.dim() != 4:
            raise ValueError(f"voxels must have shape [B,D,H,W], got {tuple(voxels.shape)}")

        B, D, H, W = voxels.shape

        # Expand voxels to have feature channels
        # Simple approach: repeat voxel values across feature dimension
        rep_ch = max(1, self.feature_dim // 4)
        voxel_features = voxels.unsqueeze(1).expand(B, rep_ch, D, H, W).contiguous()

        # Add some learned 3D processing
        if self._voxel_processor is None:
            # Build with correct device
            device = voxels.device
            self._voxel_processor = nn.Sequential(
                nn.Conv3d(rep_ch, max(rep_ch, self.feature_dim // 2), 3, padding=1),
                nn.ReLU(),
                nn.Conv3d(max(rep_ch, self.feature_dim // 2), self.feature_dim, 3, padding=1),
            ).to(device)

        return self._voxel_processor(voxel_features)

    def get_stage_parameters(self, stage: int):
        """
        Get parameters for specific training stage
        Useful for stage-specific optimization
        """
        if stage == 1:
            # Stage 1: 2D components only
            params = []
            params.extend(list(self.encoder.parameters()))
            params.extend(list(self.seg_head.parameters()))
            params.extend(list(self.attr_head.parameters()))
            params.extend(list(self.sdf_head.parameters()))
            params.extend(list(self.feature_enhancer.parameters()))

            if self.use_latent_consistency:
                params.extend(list(self.latent_head.parameters()))

        elif stage == 2:
            # Stage 2: DVX components
            params = list(self.dvx.parameters())

        else:  # stage == 3
            # Stage 3: All parameters
            params = list(self.parameters())

        return params

    def freeze_stage_parameters(self, stages_to_freeze: list):
        """
        Freeze parameters for specific stages

```

```
        Args:
            stages_to_freeze: List of stage numbers to freeze
        """
        for stage in stages_to_freeze:
            stage_params = self.get_stage_parameters(stage)
            for param in stage_params:
                param.requires_grad = False

    def unfreeze_stage_parameters(self, stages_to_unfreeze: list):
        """
        Unfreeze parameters for specific stages

        Args:
            stages_to_unfreeze: List of stage numbers to unfreeze
        """
        for stage in stages_to_unfreeze:
            stage_params = self.get_stage_parameters(stage)
            for param in stage_params:
                param.requires_grad = True

    def get_curriculum_metrics(self):
        """
        Get metrics useful for curriculum learning decisions
        """
        metrics = {}

        # Parameter counts per stage
        for stage in [1, 2, 3]:
            stage_params = self.get_stage_parameters(stage)
            metrics[f"stage_{stage}_params"] = sum(p.numel() for p in stage_params)

        # Feature dimensions
        metrics["feature_dim"] = self.feature_dim
        metrics["has_latent_consistency"] = self.use_latent_consistency
        metrics["has_graph_constraints"] = self.use_graph_constraints

        return metrics
```

■ File: param.py

==

```
import json
import numpy as np
from pathlib import Path

# Adjust this to your dataset path
data_root = Path("./data/floorplans")

# Expected attributes with their default values
expected_keys = {
    "wall_height": 2.6,
    "wall_thickness": 0.15,
    "window_base_height": 0.7,
    "window_height": 0.95,
    "door_height": 2.6,
    "pixel_scale": 0.01
}

def is_valid_number(value):
    if value is None:
        return False
    if isinstance(value, (int, float)):
        return np.isfinite(value)
    return False

def check_params_file(params_file):
    invalid_entries = []
    try:
        with open(params_file, "r") as f:
            params = json.load(f)
    except Exception as e:
        invalid_entries.append(f"Could not load JSON: {e}")
```

```
        return invalid_entries

    for key in expected_keys.keys():
        val = params.get(key)
        if val is None:
            invalid_entries.append(f"missing '{key}'")
        elif not is_valid_number(val):
            invalid_entries.append(f"{key}={val} (invalid)")
    return invalid_entries

def check_split(split="train"):
    split_dir = data_root / split
    total_files = 0
    good_files = 0
    bad_files = 0

    print(f"\nChecking split: {split}")
    if not split_dir.exists():
        print(f"Warning: {split_dir} does not exist")
        return

    # Recursively find all params.json files
    for params_file in split_dir.rglob("params.json"):
        total_files += 1
        invalid_entries = check_params_file(params_file)

        if invalid_entries:
            print(f"[BAD] {params_file}")
            for entry in invalid_entries:
                print(f"    - {entry}")
            bad_files += 1
        else:
            good_files += 1

    print(f"\nSummary for split: {split}")
    print(f"Total files checked: {total_files}")
    print(f"Good files: {good_files}")
    print(f"Bad files: {bad_files}")

def main():
    for split in ["train", "val", "test"]:
        check_split(split)

if __name__ == "__main__":
    main()
```

■ File: setup.py

==

```
"""
Setup script for the Neural-Geometric 3D Model Generator project
"""

from pathlib import Path
import os

def create_project_structure():
    """Create the complete project directory structure"""

    # Define directory structure
    directories = [
        "models",
        "training",
        "inference",
        "evaluation",
        "utils",
        "data/floorplans/train",
        "data/floorplans/val",
        "data/floorplans/test",
        "data/test_images",
        "checkpoints",
        "outputs",
```

```
        "demo_outputs",
        "evaluation_results",
        "logs"
    ]

    # Create directories
    for directory in directories:
        Path(directory).mkdir(parents=True, exist_ok=True)
        print(f"Created directory: {directory}")

    # Create __init__.py files
    init_files = [
        "models/__init__.py",
        "training/__init__.py",
        "inference/__init__.py",
        "evaluation/__init__.py",
        "utils/__init__.py"
    ]

    init_content = {
        "models/__init__.py": '''"""
Model components for Neural-Geometric 3D Model Generator
"""

from .encoder import MultiScaleEncoder, ResidualBlock
from .heads import SegmentationHead, AttributeHead, SDFHead
from .dvx import DifferentiableVectorization
from .extrusion import DifferentiableExtrusion
from .model import NeuralGeometric3DGenerator

__all__ = [
    'MultiScaleEncoder',
    'ResidualBlock',
    'SegmentationHead',
    'AttributeHead',
    'SDFHead',
    'DifferentiableVectorization',
    'DifferentiableExtrusion',
    'NeuralGeometric3DGenerator'
]''',

        "training/__init__.py": '''"""
Training components for Neural-Geometric 3D Model Generator
"""

from .losses import ResearchGradeLoss
from .trainer import MultiStageTrainer

__all__ = [
    'ResearchGradeLoss',
    'MultiStageTrainer'
]''',

        "inference/__init__.py": '''"""
Inference components for Neural-Geometric 3D Model Generator
"""

from .engine import ResearchInferenceEngine

__all__ = [
    'ResearchInferenceEngine'
]''',

        "evaluation/__init__.py": '''"""
Evaluation components for Neural-Geometric 3D Model Generator
"""

from .metrics import ModelEvaluator, evaluate_model, compute_iou, compute_3d_iou

__all__ = [
    'ModelEvaluator',
    'evaluate_model',
    'compute_iou',
```

```
    'compute_3d_iou'
]''',

        "utils/__init__.py": '''"""
Utility functions for Neural-Geometric 3D Model Generator
"""

from .visualization import (
    plot_training_history,
    visualize_predictions,
    visualize_polygons,
    save_model_outputs,
    create_comparison_grid,
    create_3d_visualization
)

__all__ = [
    'plot_training_history',
    'visualize_predictions',
    'visualize_polygons',
    'save_model_outputs',
    'create_comparison_grid',
    'create_3d_visualization'
]'''
    }

    for file_path, content in init_content.items():
        with open(file_path, 'w') as f:
            f.write(content)
        print(f"Created: {file_path}")


def create_sample_config():
    """Create a sample configuration file for easy customization"""

    sample_config = '''"""
Sample configuration for quick customization
Copy this to config_custom.py and modify as needed
"""

from config import *

# Custom configuration example
CUSTOM_DATA_CONFIG = DataConfig(
    data_dir="./my_data/floorplans",
    batch_size=16,  # Larger batch if you have more GPU memory
    num_workers=8,  # More workers if you have more CPU cores
)

CUSTOM_TRAINING_CONFIG = TrainingConfig(
    stage1_epochs=30,  # More epochs for better 2D learning
    stage2_epochs=20,  # More DVX training
    stage3_epochs=50,  # Longer end-to-end training
    stage1_lr=2e-4,    # Higher learning rate
)

CUSTOM_MODEL_CONFIG = ModelConfig(
    feature_dim=768,   # Larger model
    voxel_size=128,    # Higher resolution 3D
    max_polygons=30,   # More polygons
)
'''

    with open("config_custom_example.py", "w") as f:
        f.write(sample_config)
    print("Created: config_custom_example.py")


def create_gitignore():
    """Create .gitignore file"""

    gitignore_content = '''# Python
__pycache__/
```

```
*.py[cod]
*$py.class
*.so
.Python
build/
develop-eggs/
dist/
downloads/
eggs/
.eggs/
lib/
lib64/
parts/
sdist/
var/
wheels/
*.egg-info/
.installed.cfg
*.egg

# PyTorch
*.pth
*.pt

# Data
data/
datasets/
*.npz
*.obj
*.off
*.ply

# Outputs
outputs/
results/
checkpoints/
logs/
demo_outputs/
evaluation_results/
training_progress/

# IDE
.vscode/
.idea/
*.swp
*.swo

# OS
.DS_Store
Thumbs.db

# Jupyter
.ipynb_checkpoints

# Environment
.env
.venv
env/
venv/

# Images and videos
*.png
*.jpg
*.jpeg
*.gif
*.mp4
*.avi

# Except sample images
!sample_images/
!docs/images/
'''

```

```python
    with open(".gitignore", "w") as f:
        f.write(gitignore_content)
    print("Created: .gitignore")


def main():
    """Main setup function"""
    print("Setting up Neural-Geometric 3D Model Generator project...")
    print("=" * 60)

    # Create directory structure
    create_project_structure()
    print()

    # Create sample config
    create_sample_config()
    print()

    # Create gitignore
    create_gitignore()
    print()

    print("Project setup completed!")
    print("=" * 60)
    print("IMPORTANT ACCURACY EXPECTATIONS:")
    print("- 90%+ accuracy in 2D-to-3D generation is extremely challenging")
    print("- Actual accuracy depends heavily on:")
    print("  * Dataset quality and size (need 10K+ samples)")
    print("  * Ground truth accuracy")
    print("  * Problem complexity (simple vs complex floorplans)")
    print("  * Evaluation metrics used")
    print("- Realistic expectations:")
    print("  * Segmentation: 75-85% mIoU with good data")
    print("  * 3D reconstruction: 60-75% IoU for architectural scenes")
    print("  * Polygon fitting: 70-80% accuracy")
    print("=" * 60)
    print("\nNext steps:")
    print("1. Install dependencies: pip install -r requirements.txt")
    print("2. Prepare high-quality dataset (critical for accuracy)")
    print("3. Run demo: python demo.py")
    print("4. Start training: python train.py")


if __name__ == "__main__":
    main()
```

■ File: train.py

==

```python
"""
Enhanced training script for the Neural-Geometric 3D Model Generator
Implements novel training strategies: dynamic curriculum, adaptive weighting, cross-modal consistency
"""

import argparse
import torch
from torch.utils.data import DataLoader
from pathlib import Path
import json
import matplotlib.pyplot as plt
import numpy as np

from dataset import AdvancedFloorPlanDataset
from models.model import NeuralGeometric3DGenerator
from training.trainer import AdaptiveMultiStageTrainer, MultiStageTrainer
from utils.visualization import plot_training_history, plot_curriculum_analysis
from config import (
    DEFAULT_DATA_CONFIG,
    DEFAULT_MODEL_CONFIG,
    DEFAULT_TRAINING_CONFIG,
    DEFAULT_LOSS_CONFIG,
    TrainingConfig,
```

```
    CurriculumConfig
)


def create_enhanced_config(args):
    """Create enhanced training configuration with novel strategies"""
    config = TrainingConfig()

    # Basic settings
    config.device = args.device or ("cuda" if torch.cuda.is_available() else "cpu")

    # Dynamic curriculum settings
    if args.dynamic_curriculum:
        config.curriculum = CurriculumConfig()
        config.curriculum.use_dynamic_curriculum = True
        config.curriculum.stage_switch_patience = args.patience
        config.curriculum.min_improvement_threshold = args.min_improvement

        # Adjust epoch limits for dynamic training
        config.max_stage1_epochs = args.max_stage1_epochs
        config.max_stage2_epochs = args.max_stage2_epochs
        config.max_stage3_epochs = args.max_stage3_epochs

        print("Dynamic curriculum learning enabled")
        print(f"  Stage switch patience: {config.curriculum.stage_switch_patience}")
        print(f"  Min improvement threshold: {config.curriculum.min_improvement_threshold}")
    else:
        # Disable dynamic curriculum for traditional training
        config.curriculum.use_dynamic_curriculum = False
        print("Using traditional fixed-epoch training")

    # GradNorm dynamic weighting
    if args.gradnorm:
        config.curriculum.use_gradnorm = True
        config.curriculum.gradnorm_alpha = args.gradnorm_alpha
        config.curriculum.gradnorm_update_freq = args.gradnorm_freq
        print(f"GradNorm dynamic weighting enabled (alpha={args.gradnorm_alpha})")

    # Topology-aware scheduling
    if args.topology_schedule != "static":
        config.curriculum.topology_schedule = args.topology_schedule
        config.curriculum.topology_start_weight = args.topology_start
        config.curriculum.topology_end_weight = args.topology_end
        print(f"Topology-aware scheduling: {args.topology_schedule}")
        print(f"  Weights: {args.topology_start} -> {args.topology_end}")

    return config


def create_enhanced_model(args):
    """Create enhanced model with auxiliary heads"""
    model = NeuralGeometric3DGenerator(
        input_channels=args.input_channels,
        num_classes=args.num_classes,
        feature_dim=args.feature_dim,
        num_attributes=args.num_attributes,
        voxel_size=args.voxel_size,
        max_polygons=args.max_polygons,
        max_points=args.max_points,
        use_latent_consistency=args.latent_consistency,
        use_graph_constraints=args.graph_constraints,
        latent_embedding_dim=args.embedding_dim
    )

    print(f"Enhanced model created:")
    print(f"  Feature dim: {args.feature_dim}")
    print(f"  Latent consistency: {args.latent_consistency}")
    print(f"  Graph constraints: {args.graph_constraints}")

    # Print parameter counts
    total_params = sum(p.numel() for p in model.parameters())
    trainable_params = sum(p.numel() for p in model.parameters() if p.requires_grad)
    print(f"  Total parameters: {total_params:,}")
```

```python
    print(f"  Trainable parameters: {trainable_params:,}")

    return model


def visualize_training_results(history, output_dir):
    """Create comprehensive training visualizations"""
    output_dir = Path(output_dir)
    output_dir.mkdir(exist_ok=True)

    # Traditional loss curves
    plot_training_history(history, save_path=str(output_dir / "training_history.png"))

    # Novel curriculum analysis plots
    if "stage_transitions" in history and history["stage_transitions"]:
        plot_curriculum_analysis(history, save_path=str(output_dir / "curriculum_analysis.png"))

    # Dynamic weight evolution
    if "dynamic_weights" in history and history["dynamic_weights"]:
        plt.figure(figsize=(12, 8))

        # Extract weight evolution data
        epochs = [entry["epoch"] for entry in history["dynamic_weights"]]
        weight_names = list(history["dynamic_weights"][0]["weights"].keys())

        for weight_name in weight_names:
            weights = [entry["weights"].get(weight_name, 0) for entry in history["dynamic_weights"]]
            if any(w > 0.001 for w in weights):  # Only plot significant weights
                plt.plot(epochs, weights, label=weight_name, linewidth=2)

        plt.xlabel("Global Epoch")
        plt.ylabel("Loss Weight")
        plt.title("Dynamic Loss Weight Evolution")
        plt.legend()
        plt.grid(True, alpha=0.3)
        plt.tight_layout()
        plt.savefig(output_dir / "weight_evolution.png", dpi=300)
        plt.close()

    # Component loss breakdown
    fig, axes = plt.subplots(1, 3, figsize=(18, 5))
    stage_names = ["stage1", "stage2", "stage3"]

    for idx, stage_name in enumerate(stage_names):
        if stage_name in history and "component_losses" in history[stage_name]:
            component_data = history[stage_name]["component_losses"]
            if component_data:
                # Get component names from first entry
                component_names = list(component_data[0].keys())

                for comp_name in component_names:
                    if comp_name in ['seg', 'dice', 'polygon', 'voxel', 'topology',
                                   'latent_consistency', 'graph']:
                        values = [entry.get(comp_name, 0) for entry in component_data]
                        if any(v > 0.001 for v in values):  # Only plot significant losses
                            axes[idx].plot(values, label=comp_name, linewidth=2)

                axes[idx].set_title(f"{stage_name.upper()} Component Losses")
                axes[idx].set_xlabel("Epoch")
                axes[idx].set_ylabel("Loss Value")
                axes[idx].legend()
                axes[idx].grid(True, alpha=0.3)

    plt.tight_layout()
    plt.savefig(output_dir / "component_losses.png", dpi=300)
    plt.close()

    print(f"Training visualizations saved to {output_dir}")


def save_training_summary(history, config, output_dir):
    """Save comprehensive training summary"""
    output_dir = Path(output_dir)
```

```python

    summary = {
        "training_config": {
            "dynamic_curriculum": config.curriculum.use_dynamic_curriculum,
            "gradnorm_enabled": config.curriculum.use_gradnorm,
            "topology_schedule": config.curriculum.topology_schedule,
            "max_epochs": [config.max_stage1_epochs, config.max_stage2_epochs, config.max_stage3_epochs]
        },
        "training_results": {},
        "novel_strategies_summary": {}
    }

    # Training results
    for stage_name, data in history.items():
        if isinstance(data, dict) and "val_loss" in data and data["val_loss"]:
            summary["training_results"][stage_name] = {
                "final_val_loss": data["val_loss"][-1],
                "best_val_loss": min(data["val_loss"]),
                "epochs_trained": len(data["val_loss"])
            }

    # Novel strategies summary
    if "stage_transitions" in history:
        summary["novel_strategies_summary"]["adaptive_transitions"] = len(history["stage_transitions"])

    if "dynamic_weights" in history:
        summary["novel_strategies_summary"]["weight_updates"] = len(history["dynamic_weights"])

    if "curriculum_events" in history:
        summary["novel_strategies_summary"]["curriculum_events"] = len(history["curriculum_events"])

    # Save as JSON
    with open(output_dir / "training_summary.json", 'w') as f:
        json.dump(summary, f, indent=2)

    print(f"Training summary saved to {output_dir / 'training_summary.json'}")


def main():
    parser = argparse.ArgumentParser(description="Enhanced Neural-Geometric 3D Model Generator Training")

    # Basic arguments
    parser.add_argument("--data_dir", type=str, default="./data/floorplans",
                       help="Path to dataset directory")
    parser.add_argument("--batch_size", type=int, default=2, help="Batch size")
    parser.add_argument("--num_workers", type=int, default=4, help="Number of data workers")
    parser.add_argument("--device", type=str, default=None, help="Training device")
    parser.add_argument("--resume", type=str, default=None, help="Resume from checkpoint")
    parser.add_argument("--output_dir", type=str, default="./checkpoints",
                       help="Output directory for checkpoints")

    # Training mode selection
    parser.add_argument("--training_mode", type=str, choices=["traditional", "adaptive"],
                       default="adaptive", help="Training mode (traditional fixed epochs vs adaptive)")
    parser.add_argument("--stage", type=str, choices=["1", "2", "3", "all"], default="all",
                       help="Training stage to run (only for traditional mode)")

    # Novel training strategies
    parser.add_argument("--dynamic-curriculum", action="store_true", default=True,
                       help="Enable adaptive stage transitioning")
    parser.add_argument("--patience", type=int, default=5,
                       help="Epochs without improvement before stage transition")
    parser.add_argument("--min-improvement", type=float, default=0.001,
                       help="Minimum relative improvement threshold")

    parser.add_argument("--gradnorm", action="store_true", default=True,
                       help="Enable GradNorm dynamic loss weighting")
    parser.add_argument("--gradnorm-alpha", type=float, default=0.12,
                       help="GradNorm restoring force parameter")
    parser.add_argument("--gradnorm-freq", type=int, default=5,
                       help="GradNorm update frequency (batches)")

    parser.add_argument("--topology-schedule", type=str,
```

```
                         choices=["static", "progressive", "linear_ramp"],
                         default="progressive", help="Topology loss scheduling strategy")
    parser.add_argument("--topology-start", type=float, default=0.1,
                         help="Starting weight for topology loss")
    parser.add_argument("--topology-end", type=float, default=1.0,
                         help="Ending weight for topology loss")

    # Model enhancements
    parser.add_argument("--latent-consistency", action="store_true", default=True,
                         help="Enable cross-modal latent consistency loss")
    parser.add_argument("--graph-constraints", action="store_true", default=True,
                         help="Enable graph-based topology constraints")
    parser.add_argument("--embedding-dim", type=int, default=256,
                         help="Latent embedding dimension")

    # Model architecture
    parser.add_argument("--input_channels", type=int, default=3, help="Input image channels")
    parser.add_argument("--num_classes", type=int, default=5, help="Number of segmentation classes")
    parser.add_argument("--feature_dim", type=int, default=768, help="Feature dimension")
    parser.add_argument("--num_attributes", type=int, default=6, help="Number of attribute predictions")
    parser.add_argument("--voxel_size", type=int, default=64, help="3D voxel grid size")
    parser.add_argument("--max_polygons", type=int, default=30, help="Maximum number of polygons")
    parser.add_argument("--max_points", type=int, default=100, help="Maximum points per polygon")

    # Dynamic epoch limits
    parser.add_argument("--max-stage1-epochs", type=int, default=50, help="Max epochs for Stage 1")
    parser.add_argument("--max-stage2-epochs", type=int, default=35, help="Max epochs for Stage 2")
    parser.add_argument("--max-stage3-epochs", type=int, default=100, help="Max epochs for Stage 3")

    parser.add_argument("--persistent_workers",action="store_true",default=False,help="Keep DataLoader worke

    parser.add_argument("--prefetch_factor",type=int,default=2,help="Number of batches preloaded by each wo


    args = parser.parse_args()

    # Setup device
    device = args.device or ("cuda" if torch.cuda.is_available() else "cpu")
    print(f"Using device: {device}")

    import torch.backends.cudnn as cudnn
    if device == "cuda":
        cudnn.benchmark = True

    # Create output directory
    output_dir = Path(args.output_dir)
    output_dir.mkdir(exist_ok=True)

    # Create enhanced configuration
    config = create_enhanced_config(args)

    print("\n" + "="*80)
    print("NEURAL-GEOMETRIC 3D MODEL GENERATOR - ENHANCED TRAINING")
    print("="*80)
    print("Novel Training Strategies Enabled:")
    if config.curriculum.use_dynamic_curriculum:
        print("? Adaptive Stage Transitioning (Dynamic Curriculum)")
    if config.curriculum.use_gradnorm:
        print("? Multi-objective Optimization with GradNorm")
    if config.curriculum.topology_schedule != "static":
        print("? Topology-aware Loss Scheduling")
    if args.latent_consistency:
        print("? Cross-modal Latent Consistency Learning")
    if args.graph_constraints:
        print("? Graph-based Topology Constraints")
    print("="*80)

    # Create datasets
    print("\nLoading datasets...")
    train_dataset = AdvancedFloorPlanDataset(
        args.data_dir, split="train", augment=True
    )
    val_dataset = AdvancedFloorPlanDataset(
```

```python
        args.data_dir, split="val", augment=False
    )

    print(f"Train samples: {len(train_dataset)}")
    print(f"Validation samples: {len(val_dataset)}")

    if len(train_dataset) == 0:
        print("Error: No training samples found!")
        return

    # Create data loaders
    train_loader = DataLoader(
        train_dataset,
        batch_size=args.batch_size,
        shuffle=True,
        num_workers=args.num_workers,
        pin_memory=True,
        drop_last=True,
        persistent_workers=args.persistent_workers if args.num_workers > 0 else False,
        prefetch_factor=args.prefetch_factor if args.num_workers > 0 else None
    )

    val_loader = DataLoader(
        val_dataset,
        batch_size=max(1, args.batch_size),
        shuffle=False,
        num_workers=max(1, args.num_workers // 2),
        pin_memory=True,
        drop_last=False,
        persistent_workers=args.persistent_workers if args.num_workers > 0 else False,
        prefetch_factor=args.prefetch_factor if args.num_workers > 0 else None
    )

    # Create enhanced model
    print("\nInitializing enhanced model...")
    model = create_enhanced_model(args)

    # Create appropriate trainer
    if args.training_mode == "adaptive":
        print("\nUsing Adaptive Multi-Stage Trainer with Novel Strategies")
        trainer = AdaptiveMultiStageTrainer(
            model=model,
            train_loader=train_loader,
            val_loader=val_loader,
            device=device,
            config=config
        )
    else:
        print("\nUsing Traditional Multi-Stage Trainer")
        trainer = MultiStageTrainer(
            model=model,
            train_loader=train_loader,
            val_loader=val_loader,
            device=device,
            config=config
        )

    # Resume from checkpoint if specified
    if args.resume:
        print(f"Resuming from checkpoint: {args.resume}")
        trainer.load_checkpoint(args.resume)

    # Run training
    if args.training_mode == "adaptive" or args.stage == "all":
        print("\nStarting adaptive multi-stage training with novel strategies...")
        history = trainer.train_all_stages()
    else:
        # Traditional single-stage training
        stage_num = int(args.stage)
        print(f"Training Stage {stage_num} only...")
        if stage_num == 1:
            trainer.train_stage1()
        elif stage_num == 2:
```

```
            trainer.train_stage2()
        elif stage_num == 3:
            trainer.train_stage3()
        history = trainer.history

    # Save final model
    final_model_path = output_dir / "final_enhanced_model.pth"
    if hasattr(trainer, '_save_checkpoint'):
        trainer._save_checkpoint(str(final_model_path))
    print(f"Final model saved to: {final_model_path}")

    # Create comprehensive visualizations
    print("\nGenerating training analysis...")
    visualize_training_results(history, output_dir)

    # Save training summary
    save_training_summary(history, config, output_dir)

    print(f"\n? Enhanced training completed successfully!")
    print(f"? Results saved to: {output_dir}")
    print("\nNovel contributions implemented:")
    print("- Dynamic curriculum learning with adaptive stage transitions")
    print("- Multi-objective optimization with gradient-based reweighting")
    print("- Topology-aware progressive constraint injection")
    print("- Cross-modal latent consistency learning")
    print("- Graph-based architectural constraint learning")


if __name__ == "__main__":
    main()
```

■ File: training\losses.py

==

```
"""
Advanced loss functions for multi-task training with dynamic weighting
Enhanced with cross-modal consistency, graph constraints, and GradNorm
Modified to support conditional geometric losses via run_full_geometric flag
"""

import torch
import torch.nn as nn
import torch.nn.functional as F
import cv2
import numpy as np
from typing import Dict, Optional, Tuple, List
import networkx as nx


class DynamicLossWeighter:
    """Implements GradNorm and other dynamic weighting strategies"""

    def __init__(self, loss_names: List[str], alpha: float = 0.12, device: str = 'cuda'):
        self.loss_names = loss_names
        self.alpha = alpha
        self.device = device

        # Initialize weights
        self.weights = {name: 1.0 for name in loss_names}  # FIX: keep floats, easier logging
        self.initial_task_losses = None
        self.running_mean_losses = {name: 0.0 for name in loss_names}
        self.update_count = 0

    def update_weights(self, task_losses: Dict[str, torch.Tensor],
                       shared_parameters, update_freq: int = 10):
        """Update loss weights using GradNorm algorithm with enhanced stability"""
        if self.update_count % update_freq != 0:
            self.update_count += 1
            return self.weights

        # Store initial losses on first update
        if self.initial_task_losses is None:
```

```
            self.initial_task_losses = {name: loss.item() for name, loss in task_losses.items()}

        # Update running mean
        for name, loss in task_losses.items():
            self.running_mean_losses[name] = (0.9 * self.running_mean_losses[name] +
                                            0.1 * loss.item())

        # Calculate relative decrease rates
        loss_ratios = {}
        for name in self.loss_names:
            if name in task_losses:
                current_loss = self.running_mean_losses[name]
                initial_loss = self.initial_task_losses[name]
                loss_ratios[name] = current_loss / (initial_loss + 1e-8)

        # Calculate average relative decrease
        if not loss_ratios:
            self.update_count += 1
            return self.weights
        avg_loss_ratio = sum(loss_ratios.values()) / len(loss_ratios)

        # Calculate gradient norms with enhanced safety checks
        grad_norms = {}
        for name in self.loss_names:
            if name in task_losses:
                # Check if loss is finite before gradient computation
                if not torch.isfinite(task_losses[name]):
                    print(f"Warning: Non-finite loss for {name}, skipping gradient computation")
                    continue

                try:
                    grads = torch.autograd.grad(
                        task_losses[name], shared_parameters,
                        retain_graph=True, create_graph=False, allow_unused=True
                    )
                    grad_norm = 0.0
                    for grad in grads:
                        if grad is not None and torch.isfinite(grad).all():
                            grad_norm += grad.norm().item() ** 2
                        elif grad is not None:
                            print(f"Warning: Non-finite gradient detected for {name}")
                            grad_norm = 0.0
                            break

                    if grad_norm > 0:
                        grad_norms[name] = grad_norm ** 0.5
                except Exception as e:
                    print(f"Warning: Gradient computation failed for {name}: {e}")
                    continue

        if not grad_norms:
            self.update_count += 1
            return self.weights

        avg_grad_norm = sum(grad_norms.values()) / len(grad_norms)

        # Enhanced safety parameters
        epsilon = 1e-6  # Increased minimum threshold
        max_target_grad = 10.0  # Reduced maximum allowed gradient
        max_weight_update = 2.0  # Reduced maximum weight change

        for name in self.loss_names:
            if name in grad_norms and name in loss_ratios:
                target_grad = avg_grad_norm * (loss_ratios[name] ** self.alpha)
                target_grad = min(target_grad, max_target_grad)

                safe_grad_norm = max(grad_norms[name], epsilon)
                if safe_grad_norm == epsilon:
                    print(f"? Small gradient norm for {name}, clamped to {epsilon}")

                weight_update = target_grad / safe_grad_norm
                weight_update = min(weight_update, max_weight_update)
                if weight_update == max_weight_update:
```

```
                    print(f"? Weight update for {name} clamped to {max_weight_update}")

                # More conservative momentum update
                new_w = 0.95 * self.weights[name] + 0.05 * float(weight_update)  # Slower adaptation
                self.weights[name] = float(np.clip(new_w, 0.01, 3.0))  # Tighter bounds

        self.update_count += 1
        return self.weights

class GraphTopologyExtractor:
    """Extracts graph structure from segmentation for topology constraints"""

    @staticmethod
    def extract_room_graph(segmentation: torch.Tensor) -> Dict[str, torch.Tensor]:
        """Extract room connectivity graph from segmentation mask"""
        B, C, H, W = segmentation.shape
        device = segmentation.device

        # Get room predictions (assume classes: 0=bg, 1=wall, 2=door, 3=window, 4=room)
        room_probs = F.softmax(segmentation, dim=1)
        room_mask = room_probs[:, 4] if C > 4 else torch.zeros((B, H, W), device=device)
        wall_mask = room_probs[:, 1] if C > 1 else torch.zeros((B, H, W), device=device)

        # Simple connectivity: rooms connected if they share wall boundary
        adjacency_matrices = []
        room_features = []

        for b in range(B):
            room_b = room_mask[b].detach().cpu().numpy()
            wall_b = wall_mask[b].detach().cpu().numpy()

            # Find connected components (rooms)
            try:
                from scipy import ndimage
                labeled_rooms, num_rooms = ndimage.label(room_b > 0.5)

                # Create adjacency matrix
                adj_matrix = np.zeros((max(num_rooms, 1), max(num_rooms, 1)))
                room_centroids = []

                for i in range(1, num_rooms + 1):
                    room_i_mask = (labeled_rooms == i)
                    if np.sum(room_i_mask) > 0:
                        centroid = ndimage.center_of_mass(room_i_mask)
                        room_centroids.append(centroid)

                        # Check connectivity to other rooms through walls
                        for j in range(i + 1, num_rooms + 1):
                            room_j_mask = (labeled_rooms == j)
                            if np.sum(room_j_mask) > 0:
                                # Check if rooms are connected via wall adjacency
                                connectivity = GraphTopologyExtractor._check_room_connectivity(
                                    room_i_mask, room_j_mask, wall_b
                                )
                                adj_matrix[i-1, j-1] = connectivity
                                adj_matrix[j-1, i-1] = connectivity

                # Convert to tensor
                adj_tensor = torch.from_numpy(adj_matrix).float().to(device)
                centroids_tensor = torch.from_numpy(np.array(room_centroids) if room_centroids else np.zeros

            except ImportError:
                # Fallback if scipy not available
                adj_tensor = torch.zeros((1, 1), device=device)
                centroids_tensor = torch.zeros((0, 2), device=device)

            adjacency_matrices.append(adj_tensor)
            room_features.append(centroids_tensor)

        return {
            "adjacency_matrices": adjacency_matrices,
            "room_features": room_features
        }
```

```
    @staticmethod
    def _check_room_connectivity(room1_mask, room2_mask, wall_mask):
        """Check if two rooms are connected through walls"""
        try:
            from scipy.ndimage import binary_dilation

            # Dilate room masks to check wall adjacency
            dilated1 = binary_dilation(room1_mask, iterations=2)
            dilated2 = binary_dilation(room2_mask, iterations=2)

            # Check overlap through wall areas
            wall_overlap = (dilated1 & dilated2) & (wall_mask > 0.3)
            return float(np.sum(wall_overlap) > 0)
        except ImportError:
            # Simple distance-based fallback
            return 0.0


class ResearchGradeLoss(nn.Module):
    """
    Multi-task loss combining:
    - Traditional losses (segmentation, SDF, attributes, polygons, voxels, topology)
    - NEW: Cross-modal latent consistency
    - NEW: Graph-based topology constraints
    - NEW: Dynamic loss weighting via GradNorm
    - NEW: Conditional geometric losses based on run_full_geometric flag
    """

    def __init__(
        self,
        seg_weight: float = 1.0,
        dice_weight: float = 1.0,
        sdf_weight: float = 0.5,
        attr_weight: float = 1.0,
        polygon_weight: float = 1.0,
        voxel_weight: float = 1.0,
        topology_weight: float = 0.5,
        latent_consistency_weight: float = 0.5,
        graph_constraint_weight: float = 0.3,
        enable_dynamic_weighting: bool = True,
        gradnorm_alpha: float = 0.12,
        device: str = 'cuda',
        weight_update_freq: int = 10,
        weight_momentum: float = 0.9,
    ):
        super().__init__()

        # Store initial weights
        self.initial_weights = {
            'seg': float(seg_weight),
            'dice': float(dice_weight),
            'sdf': float(sdf_weight),
            'attr': float(attr_weight),
            'polygon': float(polygon_weight),
            'voxel': float(voxel_weight),
            'topology': float(topology_weight),
            'latent_consistency': float(latent_consistency_weight),
            'graph': float(graph_constraint_weight)
        }

        # Current weights (will be dynamically updated)
        self.weights = self.initial_weights.copy()

        # Core losses
        self.ce_loss = nn.CrossEntropyLoss()
        self.mse_loss = nn.MSELoss()
        self.l1_loss = nn.L1Loss()
        self.cosine_loss = nn.CosineEmbeddingLoss()

        # Dynamic weighting
        self.enable_dynamic_weighting = enable_dynamic_weighting
        if enable_dynamic_weighting:
```

```
            self.loss_weighter = DynamicLossWeighter(
                list(self.initial_weights.keys()), alpha=gradnorm_alpha, device=device,
            )
            self.loss_weighter.update_freq = weight_update_freq
            self.loss_weighter.momentum = weight_momentum

        self.device = device

    def update_loss_weights(self, new_weights: Dict[str, float]):
        """Update loss weights (called by trainer for curriculum scheduling)"""
        for key, value in new_weights.items():
            if key in self.weights:
                self.weights[key] = float(value)

    def forward(self, predictions: dict, targets: dict, shared_parameters=None, run_full_geometric=True):
        """
        Compute multi-task loss with conditional geometric computation and dynamic weighting.

        Args:
            predictions: Model predictions dict
            targets: Ground truth targets dict
            shared_parameters: Model parameters for GradNorm (optional)
            run_full_geometric: Whether geometric losses should be computed

        Returns:
            tuple: (total_loss, individual_losses_dict)
        """
        # Input validation - check for NaN/Inf values
        for name, tensor in predictions.items():
            if torch.is_tensor(tensor) and (torch.isnan(tensor).any() or torch.isinf(tensor).any()):
                print(f"WARNING: NaN/Inf in predictions[{name}] - zeroing out")
                predictions[name] = torch.zeros_like(tensor)

        for name, tensor in targets.items():
            if torch.is_tensor(tensor) and (torch.isnan(tensor).any() or torch.isinf(tensor).any()):
                print(f"WARNING: NaN/Inf in targets[{name}] - zeroing out")
                targets[name] = torch.zeros_like(tensor)

        device = self._get_device_from_inputs(predictions, targets)
        losses = {}
        total_loss = torch.tensor(0.0, device=device)

        # ---- 1) Core losses (always computed - lightweight) ----
        if "segmentation" in predictions and "mask" in targets:
            seg_pred = predictions["segmentation"]
            seg_target = targets["mask"].long()

            ce_loss = self.ce_loss(seg_pred, seg_target)
            losses["ce"] = ce_loss
            losses["seg"] = ce_loss  # alias for dynamic weighting

            dice_loss = self._dice_loss(seg_pred, seg_target)
            losses["dice"] = dice_loss

        if "sdf" in predictions and "mask" in targets:
            sdf_pred = predictions["sdf"]
            sdf_pred = torch.clamp(sdf_pred, -1.0, 1.0)   # FIX: prevent huge values
            sdf_target = self._mask_to_sdf(targets["mask"])
            sdf_target = sdf_target.to(sdf_pred.device).type_as(sdf_pred)
            losses["sdf"] = self.mse_loss(sdf_pred, sdf_target)

        if "attributes" in predictions and "attributes" in targets:
            pred_attr = predictions["attributes"].float()
            tgt_attr = targets["attributes"].float().to(pred_attr.device)
            losses["attr"] = self.l1_loss(pred_attr, tgt_attr)

        # ---- 2) Conditional geometric losses (heavy operations) ----
        if run_full_geometric:
            # Polygon loss (only if model produced polygons)
            if ("polygons" in predictions and predictions["polygons"] is not None and
                "polygons_gt" in targets):
                losses["polygon"] = self._polygon_loss(predictions, targets["polygons_gt"])
            else:
```

```
                # Zero loss if polygons not available
                losses["polygon"] = torch.tensor(0.0, device=device)

            # Voxel loss (only if model produced voxels)
            if ("voxels_pred" in predictions and predictions["voxels_pred"] is not None and
                "voxels_gt" in targets):
                pred_vox = predictions["voxels_pred"].float()
                tgt_vox = targets["voxels_gt"].float().to(pred_vox.device)
                losses["voxel"] = self._voxel_iou_loss(pred_vox, tgt_vox)
            else:
                # Zero loss if voxels not available
                losses["voxel"] = torch.tensor(0.0, device=device)

            # Cross-modal latent consistency (only if embeddings available)
            if ("latent_2d_embedding" in predictions and "latent_3d_embedding" in predictions and
                predictions["latent_2d_embedding"] is not None and predictions["latent_3d_embedding"] is no
                consistency_loss = self._latent_consistency_loss(
                    predictions["latent_2d_embedding"],
                    predictions["latent_3d_embedding"]
                )
                losses["latent_consistency"] = consistency_loss
            else:
                losses["latent_consistency"] = torch.tensor(0.0, device=device)
        else:
            # When geometric computation is skipped, use zero losses
            losses["polygon"] = torch.tensor(0.0, device=device)
            losses["voxel"] = torch.tensor(0.0, device=device)
            losses["latent_consistency"] = torch.tensor(0.0, device=device)

        # ---- 3) Independent auxiliary losses (always computed if enabled) ----
        # Traditional topology loss
        if "segmentation" in predictions:
            losses["topology"] = self._topology_loss(predictions["segmentation"])

        # Graph-based topology constraints
        if "segmentation" in predictions:
            graph_loss = self._graph_topology_loss(predictions["segmentation"])
            losses["graph"] = graph_loss

        # ---- 4) Apply weighting ----
        if self.enable_dynamic_weighting and shared_parameters is not None:
            # Only include differentiable losses for GradNorm
            task_losses = {
                name: loss for name, loss in losses.items()
                if name in self.weights and isinstance(loss, torch.Tensor) and loss.requires_grad
            }

            dynamic_weights = self.loss_weighter.update_weights(task_losses, shared_parameters)

            # Apply weights (dynamic for diff losses, static for non-diff losses)
            for name, loss in losses.items():
                if name in self.weights:
                    if name in dynamic_weights:
                        weight = dynamic_weights[name]
                    else:
                        weight = self.weights[name]
                    total_loss = total_loss + weight * loss
        else:
            # Static weights
            for name, loss in losses.items():
                if name in self.weights:
                    total_loss = total_loss + self.weights[name] * loss

        # Final NaN/Inf guard
        for k, v in list(losses.items()):
            if torch.isnan(v).any() or torch.isinf(v).any():
                print(f"[Warning] {k} loss is NaN/Inf ? zeroed out")
                losses[k] = torch.tensor(0.0, device=device)

        losses["total"] = total_loss
        return total_loss, losses

    def __call__(self, predictions: dict, targets: dict, shared_parameters=None, run_full_geometric=True):
```

```
        """
        Convenience method for trainer compatibility

        Args:
            predictions: Model predictions dict
            targets: Ground truth targets dict
            shared_parameters: Model parameters for GradNorm (optional)
            run_full_geometric: Whether to compute geometric losses

        Returns:
            tuple: (total_loss, individual_losses_dict)
        """
        return self.forward(predictions, targets, shared_parameters, run_full_geometric)

    def _get_device_from_inputs(self, predictions, targets):
        """Helper to determine device from inputs"""
        for pred_dict in [predictions, targets]:
            for value in pred_dict.values():
                if torch.is_tensor(value):
                    return value.device
        return self.device

    # ---- NEW: Cross-modal latent consistency loss ----
    def _latent_consistency_loss(self, embedding_2d: torch.Tensor, embedding_3d: torch.Tensor) -> torch.Ten
        """
        Ensure 2D floorplan embeddings match 3D voxelized structure embeddings
        embedding_2d: [B, D] - 2D floorplan embeddings
        embedding_3d: [B, D] - 3D structure embeddings
        """
        if embedding_2d.shape != embedding_3d.shape:
            # Project to same dimension if needed
            min_dim = min(embedding_2d.shape[-1], embedding_3d.shape[-1])
            embedding_2d = embedding_2d[..., :min_dim]
            embedding_3d = embedding_3d[..., :min_dim]

        # Cosine similarity loss (maximize similarity)
        target = torch.ones(embedding_2d.shape[0], device=embedding_2d.device)
        cosine_loss = self.cosine_loss(embedding_2d, embedding_3d, target)

        # L2 consistency loss
        l2_loss = F.mse_loss(embedding_2d, embedding_3d)

        return 0.7 * cosine_loss + 0.3 * l2_loss

    # ---- NEW: Graph-based topology constraints ----
    def _graph_topology_loss(self, segmentation_logits: torch.Tensor) -> torch.Tensor:
        """
        Graph-based topology constraints on room connectivity
        segmentation_logits: [B, C, H, W]
        """
        try:
            # Extract graph structure
            graph_data = GraphTopologyExtractor.extract_room_graph(segmentation_logits)
            device = segmentation_logits.device

            total_graph_loss = torch.tensor(0.0, device=device)
            batch_size = segmentation_logits.shape[0]

            for b in range(batch_size):
                if b < len(graph_data["adjacency_matrices"]):
                    adj_matrix = graph_data["adjacency_matrices"][b]
                    if adj_matrix.numel() == 0:
                        continue

                    # Connectivity constraint: encourage reasonable connectivity
                    # Penalize isolated rooms (degree 0) and over-connected rooms
                    degrees = adj_matrix.sum(dim=1)

                    # Isolation penalty (rooms should have at least 1 connection)
                    isolation_penalty = torch.exp(-degrees).mean()

                    # Over-connection penalty (rooms shouldn't connect to everything)
                    max_reasonable_connections = min(4, adj_matrix.shape[0] - 1)
```

```
                    over_connection_penalty = F.relu(degrees - max_reasonable_connections).mean()

                    # Graph smoothness (connected rooms should have similar features)
                    if b < len(graph_data["room_features"]) and graph_data["room_features"][b].numel() > 0:
                        room_features = graph_data["room_features"][b]
                        if room_features.shape[0] > 1:
                            feature_distances = torch.cdist(room_features, room_features)
                            # Weight by adjacency - connected rooms should be similar
                            smoothness_loss = (adj_matrix * feature_distances).sum() / (adj_matrix.sum() + 1
                        else:
                            smoothness_loss = torch.tensor(0.0, device=device)
                    else:
                        smoothness_loss = torch.tensor(0.0, device=device)

                    batch_graph_loss = (0.4 * isolation_penalty +
                                      0.3 * over_connection_penalty +
                                      0.3 * smoothness_loss)
                    total_graph_loss = total_graph_loss + batch_graph_loss

            return total_graph_loss / batch_size

        except Exception as e:
            # Fallback to zero loss if graph extraction fails
            return torch.tensor(0.0, device=segmentation_logits.device)

    # ---- Existing helper methods (preserved) ----
    def _dice_loss(self, pred: torch.Tensor, target: torch.Tensor, smooth: float = 1e-6) -> torch.Tensor:
        """Dice loss implementation"""
        pred_soft = F.softmax(pred, dim=1)
        B = pred_soft.shape[0]
        C = pred_soft.shape[1]

        dice_losses = []
        for c in range(C):
            pred_c = pred_soft[:, c, :, :]
            target_c = (target == c).float().to(pred_c.device)
            intersection = (pred_c * target_c).view(B, -1).sum(dim=1)
            union = pred_c.view(B, -1).sum(dim=1) + target_c.view(B, -1).sum(dim=1)
            dice = (2.0 * intersection + smooth) / (union + smooth)
            dice_losses.append((1.0 - dice).mean())

        return torch.stack(dice_losses).mean()

    def _mask_to_sdf(self, mask: torch.Tensor) -> torch.Tensor:
        """Convert mask to SDF with performance warning"""
        device = mask.device if torch.is_tensor(mask) else None
        if not torch.is_tensor(mask):
            mask = torch.tensor(mask, device=device)

        B, H, W = mask.shape
        sdf = torch.zeros((B, 1, H, W), dtype=torch.float32, device=device)

        # FIX: Add performance warning for CV2 bottleneck
        if B > 8:  # Warn for large batches
            print(f"[Performance Warning] SDF conversion with batch_size={B} uses CPU cv2 - consider GPU imp

        for b in range(B):
            mask_np = mask[b].detach().cpu().numpy().astype(np.uint8)  # FIX: explicit detach
            try:
                dist_inside = cv2.distanceTransform((mask_np > 0).astype(np.uint8), cv2.DIST_L2, 5)
                dist_outside = cv2.distanceTransform((mask_np == 0).astype(np.uint8), cv2.DIST_L2, 5)
                sdf_np = dist_inside.astype(np.float32) - dist_outside.astype(np.float32)
                sdf_np = np.tanh(sdf_np / 10.0).astype(np.float32)
                sdf[b, 0] = torch.from_numpy(sdf_np)
            except Exception:
                # Fallback if cv2 fails
                sdf[b, 0] = torch.zeros_like(mask[b].float())

        return sdf

    def _polygon_loss(self, predictions: dict, targets: dict) -> torch.Tensor:
        """Polygon/DVX loss (preserved from original)"""
        pred_polys = predictions.get("polygons")
```

```
        tgt_polys = targets.get("polygons")
        valid_mask = targets.get("valid_mask")

        if pred_polys is None or tgt_polys is None:
            return torch.tensor(0.0, device=pred_polys.device if pred_polys is not None else self.device)

        pred_polys = pred_polys.float()
        tgt_polys = tgt_polys.float().to(pred_polys.device)

        point_loss = self.mse_loss(pred_polys, tgt_polys)

        pred_valid = predictions.get("polygon_validity")
        if pred_valid is None or valid_mask is None:
            validity_loss = torch.tensor(0.0, device=pred_polys.device)
        else:
            pred_valid = pred_valid.float().to(pred_polys.device)
            valid_mask_f = valid_mask.float().to(pred_polys.device)
            validity_loss = self.mse_loss(pred_valid, valid_mask_f)

        smoothness_loss = self._polygon_smoothness(pred_polys)
        rect_loss = self._rectilinearity_loss(pred_polys)

        return point_loss + 0.1 * validity_loss + 0.05 * smoothness_loss + 0.1 * rect_loss

    def _polygon_smoothness(self, polygons: torch.Tensor) -> torch.Tensor:
        """Polygon smoothness loss (preserved)"""
        if polygons is None or polygons.numel() == 0:
            return torch.tensor(0.0, device=polygons.device if polygons is not None else self.device)

        p1 = polygons
        p2 = torch.roll(polygons, -1, dims=2)
        p3 = torch.roll(polygons, -2, dims=2)
        curvature = torch.norm(p1 - 2.0 * p2 + p3, dim=-1)
        return curvature.mean()

    def _rectilinearity_loss(self, polygons: torch.Tensor) -> torch.Tensor:
        """Encourage axis-aligned structure (preserved)"""
        if polygons is None or polygons.numel() == 0:
            return torch.tensor(0.0, device=polygons.device if polygons is not None else self.device)

        edges = torch.roll(polygons, -1, dims=2) - polygons
        edge_norms = torch.norm(edges, dim=-1, keepdim=True)
        edges_normalized = edges / (edge_norms + 1e-6)

        edge1 = edges_normalized
        edge2 = torch.roll(edges_normalized, -1, dims=2)

        cos_angles = (edge1 * edge2).sum(dim=-1)
        cos2 = cos_angles ** 2
        perp_penalty = cos2
        parallel_penalty = (cos2 - 1.0) ** 2
        angle_penalty = torch.minimum(perp_penalty, parallel_penalty)
        return angle_penalty.mean()

    def _voxel_iou_loss(self, pred_voxels: torch.Tensor, target_voxels: torch.Tensor) -> torch.Tensor:
        """3D voxel IoU loss (preserved)"""
        pred_prob = torch.sigmoid(torch.clamp(pred_voxels, -10.0, 10.0))  # FIX: safe sigmoid range
        target = target_voxels.float().to(pred_prob.device)

        intersection = (pred_prob * target).view(pred_prob.shape[0], -1).sum(dim=1)
        union = (pred_prob.view(pred_prob.shape[0], -1).sum(dim=1) +
                 target.view(target.shape[0], -1).sum(dim=1) - intersection)

        iou = (intersection + 1e-6) / (union + 1e-6)
        return (1.0 - iou).mean()

    def _topology_loss(self, segmentation_logits: torch.Tensor) -> torch.Tensor:
        """Traditional topology loss (preserved)"""
        seg_soft = F.softmax(segmentation_logits, dim=1)
        C = seg_soft.shape[1]
        device = seg_soft.device

        walls = seg_soft[:, 1] if C > 1 else torch.zeros_like(seg_soft[:, 0])
```

```
        doors = seg_soft[:, 2] if C > 2 else torch.zeros_like(walls)
        windows = seg_soft[:, 3] if C > 3 else torch.zeros_like(walls)

        door_wall_overlap = doors * walls
        window_wall_overlap = windows * walls

        door_penalty = torch.maximum(doors - door_wall_overlap, torch.zeros_like(doors))
        window_penalty = torch.maximum(windows - window_wall_overlap, torch.zeros_like(windows))

        connectivity_loss = self._connectivity_loss(walls)

        return door_penalty.mean() + window_penalty.mean() + 0.1 * connectivity_loss

    def _connectivity_loss(self, wall_prob: torch.Tensor) -> torch.Tensor:
        """Connectivity loss for walls (preserved)"""
        if wall_prob is None or wall_prob.numel() == 0:
            return torch.tensor(0.0, device=wall_prob.device if wall_prob is not None else self.device)

        kernel = torch.ones((1, 1, 3, 3), device=wall_prob.device, dtype=wall_prob.dtype) / 9.0
        neighbors = F.conv2d(wall_prob.unsqueeze(1), kernel, padding=1).squeeze(1)

        isolation_penalty = wall_prob * torch.exp(-neighbors)
        return isolation_penalty.mean()


class LossScheduler:
    """Manages curriculum-based loss weight scheduling"""

    def __init__(self, config):
        self.config = config
        self.loss_schedules = config.loss_schedule

    def get_scheduled_weights(self, current_stage: int, current_epoch: int,
                              stage_epoch: int, total_stage_epochs: int,
                              base_weights: Dict[str, float]) -> Dict[str, float]:
        """
        Calculate loss weights based on curriculum schedule

        Args:
            current_stage: Current training stage (1, 2, 3)
            current_epoch: Global epoch count
            stage_epoch: Epoch within current stage
            total_stage_epochs: Total epochs planned for current stage
            base_weights: Base weight configuration
        """
        scheduled_weights = base_weights.copy()

        for loss_name, schedule_type in self.loss_schedules.items():
            if loss_name not in scheduled_weights:
                continue

            base_weight = scheduled_weights[loss_name]

            if schedule_type == "static":
                # Keep original weight
                continue

            elif schedule_type == "progressive":
                # Gradually increase throughout training
                if loss_name == "topology":
                    start_weight = self.config.topology_start_weight
                    end_weight = self.config.topology_end_weight
                    ramp_epochs = self.config.topology_ramp_epochs
                    progress = min(current_epoch / ramp_epochs, 1.0)
                    scheduled_weights[loss_name] = start_weight + progress * (end_weight - start_weight)

            elif schedule_type == "linear_ramp":
                # Linear increase within current stage
                progress = stage_epoch / max(total_stage_epochs, 1)
                scheduled_weights[loss_name] = base_weight * progress

            elif schedule_type == "exponential":
                # Exponential increase
```

```
                progress = stage_epoch / max(total_stage_epochs, 1)
                scheduled_weights[loss_name] = base_weight * (progress ** 2)

            elif schedule_type == "early_decay":
                # Decay after Stage 1 (for SDF loss)
                if current_stage > 1:
                    scheduled_weights[loss_name] = base_weight * 0.3

            elif schedule_type == "staged_ramp":
                # Ramp up during specific stage (polygon in Stage 2)
                if current_stage == 2:
                    progress = stage_epoch / max(total_stage_epochs, 1)
                    scheduled_weights[loss_name] = base_weight * progress
                elif current_stage < 2:
                    scheduled_weights[loss_name] = 0.0

            elif schedule_type == "late_ramp":
                # Ramp up in Stage 3 (voxel loss)
                if current_stage == 3:
                    progress = stage_epoch / max(total_stage_epochs, 1)
                    scheduled_weights[loss_name] = base_weight * progress
                elif current_stage < 3:
                    scheduled_weights[loss_name] = 0.0

            elif schedule_type == "mid_ramp":
                # Activate mid-training (latent consistency)
                if current_stage >= 2:
                    if current_stage == 2:
                        progress = min(stage_epoch / (total_stage_epochs * 0.5), 1.0)
                        scheduled_weights[loss_name] = base_weight * progress
                    else:  # Stage 3
                        scheduled_weights[loss_name] = base_weight
                else:
                    scheduled_weights[loss_name] = 0.0

            elif schedule_type == "delayed_ramp":
                # FIX: gentler ramp for graph constraints
                if current_epoch >= self.config.graph_start_epoch:
                    epochs_since_start = current_epoch - self.config.graph_start_epoch
                    ramp_duration = 50  # FIX: slower ramp (was 20)
                    progress = min(epochs_since_start / ramp_duration, 1.0)
                    scheduled_weights[loss_name] = self.config.graph_end_weight * progress
                else:
                    scheduled_weights[loss_name] = 0.0

        return scheduled_weights
```

■ File: training\trainer.py

==

```
"""
Dynamic Multi-stage training system with adaptive curriculum learning
Implements novel training strategies: dynamic stage transitions, topology-aware scheduling,
multi-objective optimization, and cross-modal consistency learning
"""

import torch
import torch.nn.utils

# training/trainer.py - Fixed AMP imports
from torch.amp import autocast, GradScaler
import time
import numpy as np
import random
from pathlib import Path
from tqdm import tqdm
from typing import Dict, List, Optional, Tuple
from collections import deque

from .losses import ResearchGradeLoss, LossScheduler
from config import DEFAULT_TRAINING_CONFIG, DEFAULT_LOSS_CONFIG, StageTransitionCriteria

```

```python

class CurriculumState:
    """Tracks curriculum learning state and metrics"""

    def __init__(self, config):
        self.config = config

        # Loss history for plateau detection
        self.loss_history = {
            "stage1": deque(maxlen=config.plateau_detection_window * 2),
            "stage2": deque(maxlen=config.plateau_detection_window * 2),
            "stage3": deque(maxlen=config.plateau_detection_window * 2),
        }

        # Component loss tracking
        self.component_losses = {
            "segmentation": deque(maxlen=20),
            "dice": deque(maxlen=20),
            "polygon": deque(maxlen=20),
            "voxel": deque(maxlen=20),
            "topology": deque(maxlen=20),
            "latent_consistency": deque(maxlen=20),
            "graph": deque(maxlen=20),
        }

        # Gradient magnitude tracking for dynamic weighting
        self.gradient_norms = {
            name: deque(maxlen=config.gradient_norm_window)
            for name in self.component_losses.keys()
        }

        # Stage transition tracking
        self.epochs_without_improvement = 0
        self.best_val_loss = float("inf")
        self.stage_transition_epochs = []

        # Dynamic weights history
        self.weight_history = []

    def update_loss_history(self, stage: str, val_loss: float):
        """Update validation loss history for plateau detection"""
        if stage in self.loss_history:
            self.loss_history[stage].append(val_loss)

        # Update improvement tracking
        if val_loss < self.best_val_loss:
            self.best_val_loss = val_loss
            self.epochs_without_improvement = 0
        else:
            self.epochs_without_improvement += 1

    def update_component_losses(self, loss_components: Dict[str, float]):
        """Update individual loss component history"""
        for name, loss_val in loss_components.items():
            if name in self.component_losses:
                self.component_losses[name].append(loss_val)

    def should_transition(self, current_stage: int) -> bool:
        """Check if should transition to next stage"""
        if current_stage == 1:
            val_losses = list(self.loss_history["stage1"])
            return StageTransitionCriteria.should_transition_from_stage1(
                [], val_losses, self.config
            )
        elif current_stage == 2:
            polygon_losses = list(self.component_losses["polygon"])
            return StageTransitionCriteria.should_transition_from_stage2(
                polygon_losses, self.config
            )

        return False


```

```
class AdaptiveMultiStageTrainer:
    """
    Advanced multi-stage trainer with dynamic curriculum learning:
    - Adaptive stage transitioning based on performance plateaus
    - Topology-aware loss scheduling
    - Multi-objective optimization with dynamic weighting
    - Cross-modal latent consistency learning
    - Graph-based topology constraints
    """

    # Class constant for rolling checkpoint path
    ROLLING_CHECKPOINT = "latest_checkpoint.pth"

    def __init__(self, model, train_loader, val_loader, device=None, config=None):
        if config is None:
            config = DEFAULT_TRAINING_CONFIG

        self.model = model.to(device or config.device)
        self.train_loader = train_loader
        self.val_loader = val_loader
        self.device = device or config.device
        self.config = config

        # Initialize curriculum state
        self.curriculum_state = CurriculumState(config.curriculum)
        self.loss_scheduler = LossScheduler(config.curriculum)

        # Training state tracking for resume functionality
        self.current_stage = 1
        self.current_epoch = 0
        self.global_epoch = 0
        self.stage_epoch = 0
        self.stage_start_time = None
        self.epoch_times = []

        # Add AMP and optimization settings - Updated for new PyTorch API
        self.use_amp = getattr(config, "use_mixed_precision", True)
        self.scaler = GradScaler("cuda", enabled=self.use_amp)
        self.accumulation_steps = getattr(config, "accumulation_steps", 1)
        self.dvx_step_freq = getattr(config, "dvx_step_freq", 1)
        self.voxel_size_stage = getattr(config, "voxel_size_stage", None)
        self.image_size_stage = getattr(config, "image_size_stage", None)
        self._step = 0

        # Enhanced optimizers with better hyperparameters
        self.optimizer_2d = torch.optim.AdamW(
            list(self.model.encoder.parameters())
            + list(self.model.seg_head.parameters())
            + list(self.model.attr_head.parameters())
            + list(self.model.sdf_head.parameters()),
            lr=config.stage1_lr,
            weight_decay=config.stage1_weight_decay,
            betas=(0.9, 0.999),
        )

        self.optimizer_dvx = torch.optim.AdamW(
            self.model.dvx.parameters(),
            lr=config.stage2_lr,
            weight_decay=config.stage2_weight_decay,
            betas=(0.9, 0.999),
        )

        self.optimizer_full = torch.optim.AdamW(
            self.model.parameters(),
            lr=config.stage3_lr,
            weight_decay=config.stage3_weight_decay,
            betas=(0.9, 0.999),
        )

        # Enhanced learning rate schedulers
        if config.use_cosine_restarts:
            self.scheduler_2d = torch.optim.lr_scheduler.CosineAnnealingWarmRestarts(
                self.optimizer_2d, T_0=20, T_mult=1
```

```
            )
            self.scheduler_dvx = torch.optim.lr_scheduler.CosineAnnealingWarmRestarts(
                self.optimizer_dvx, T_0=15, T_mult=1
            )
            self.scheduler_full = torch.optim.lr_scheduler.CosineAnnealingWarmRestarts(
                self.optimizer_full, T_0=30, T_mult=1
            )
        else:
            self.scheduler_2d = torch.optim.lr_scheduler.CosineAnnealingLR(
                self.optimizer_2d, T_max=config.max_stage1_epochs
            )
            self.scheduler_dvx = torch.optim.lr_scheduler.CosineAnnealingLR(
                self.optimizer_dvx, T_max=config.max_stage2_epochs
            )
            self.scheduler_full = torch.optim.lr_scheduler.CosineAnnealingLR(
                self.optimizer_full, T_max=config.max_stage3_epochs
            )

        # Enhanced loss function with dynamic weighting
        base_loss_kwargs = {
            k: v
            for k, v in DEFAULT_LOSS_CONFIG.__dict__.items()
            if k != "enable_dynamic_weighting"
        }
        self.loss_fn = ResearchGradeLoss(
            **base_loss_kwargs,
            enable_dynamic_weighting=bool(config.curriculum.use_gradnorm),
            gradnorm_alpha=float(config.curriculum.gradnorm_alpha),
            device=self.device,
        )

        self.history = {
            "stage1": {"train_loss": [], "val_loss": [], "component_losses": []},
            "stage2": {"train_loss": [], "val_loss": [], "component_losses": []},
            "stage3": {"train_loss": [], "val_loss": [], "component_losses": []},
            "stage_transitions": [],
            "dynamic_weights": [],
            "curriculum_events": [],
        }

    def _get_eta_string(self, epoch, total_epochs):
        """Calculate and format ETA string"""
        if len(self.epoch_times) == 0:
            return "ETA: calculating..."

        avg_epoch_time = sum(self.epoch_times) / len(self.epoch_times)
        remaining_epochs = total_epochs - epoch - 1
        eta_seconds = avg_epoch_time * remaining_epochs

        if eta_seconds < 60:
            return f"ETA: {int(eta_seconds)}s"
        elif eta_seconds < 3600:
            return f"ETA: {int(eta_seconds // 60)}m {int(eta_seconds % 60)}s"
        else:
            hours = int(eta_seconds // 3600)
            minutes = int((eta_seconds % 3600) // 60)
            return f"ETA: {hours}h {minutes}m"

    def _get_shared_parameters(self):
        """Get shared parameters for GradNorm weighting"""
        # Return encoder parameters as shared across tasks
        return list(self.model.encoder.parameters())

    def _update_loss_weights_for_curriculum(
        self, current_stage: int, stage_epoch: int, total_stage_epochs: int
    ):
        """Update loss weights based on curriculum schedule"""
        base_weights = {
            "seg": self.loss_fn.initial_weights["seg"],
            "dice": self.loss_fn.initial_weights["dice"],
            "sdf": self.loss_fn.initial_weights["sdf"],
            "attr": self.loss_fn.initial_weights["attr"],
            "polygon": self.loss_fn.initial_weights["polygon"],
```

```python
            "voxel": self.loss_fn.initial_weights["voxel"],
            "topology": self.loss_fn.initial_weights["topology"],
            "latent_consistency": self.loss_fn.initial_weights["latent_consistency"],
            "graph": self.loss_fn.initial_weights["graph"],
        }

        scheduled_weights = self.loss_scheduler.get_scheduled_weights(
            current_stage,
            self.global_epoch,
            stage_epoch,
            total_stage_epochs,
            base_weights,
        )

        self.loss_fn.update_loss_weights(scheduled_weights)

        # Log weight changes
        self.history["dynamic_weights"].append(
            {
                "epoch": self.global_epoch,
                "stage": current_stage,
                "weights": scheduled_weights.copy(),
            }
        )

    def _train_epoch(self, mode="stage1"):
        """Enhanced training epoch with AMP, gradient accumulation, and DVX gating"""
        self.model.train()
        total_loss = 0
        component_loss_sums = {}

        # Select appropriate optimizer based on mode
        if mode == "stage1":
            optimizer = self.optimizer_2d
        elif mode == "stage2":
            optimizer = self.optimizer_dvx
        else:  # stage3
            optimizer = self.optimizer_full

        # Progress bar for training batches
        train_pbar = tqdm(
            self.train_loader, desc=f"Training {mode.upper()}", leave=False, ncols=120
        )

        batch_count = 0
        epoch_start_time = time.time()

        for batch_idx, batch in enumerate(train_pbar):
            self._step += 1
            batch = {
                k: v.to(self.device, non_blocking=True) if torch.is_tensor(v) else v
                for k, v in batch.items()
            }

            # Gate heavy DVX/extrusion: only run full forward every dvx_step_freq steps
            run_full_geometric = (self.dvx_step_freq <= 1) or (
                (self._step % self.dvx_step_freq) == 0
            )

            # First-batch profiling (optional timing helper)
            if batch_idx == 0 and self.global_epoch == 0:
                torch.cuda.synchronize()
                t0 = time.time()
                with autocast("cuda", enabled=self.use_amp):
                    out = self.model(batch["image"], run_full_geometric=True)
                    # Prepare targets for loss computation
                    targets = self._prepare_targets(batch, mode)
                    shared_params = (
                        self._get_shared_parameters()
                        if self.config.curriculum.use_gradnorm
                        else None
                    )
                    l, _ = self.loss_fn(
```

```
                        out, targets, shared_params, run_full_geometric=True
                    )
                torch.cuda.synchronize()
                print(f"First-batch forward+loss time: {time.time() - t0:.3f}s")

            with autocast("cuda", enabled=self.use_amp):
                # Forward pass with geometric gating
                predictions = self.model(
                    batch["image"], run_full_geometric=run_full_geometric
                )

                # Add latent embeddings if model supports it
                if hasattr(self.model, "get_latent_embeddings"):
                    latent_2d, latent_3d = self.model.get_latent_embeddings(
                        batch["image"]
                    )
                    predictions["latent_2d_embedding"] = latent_2d
                    predictions["latent_3d_embedding"] = latent_3d

                # Prepare targets based on training mode
                targets = self._prepare_targets(batch, mode)

                # Get shared parameters for dynamic weighting
                shared_params = (
                    self._get_shared_parameters()
                    if self.config.curriculum.use_gradnorm
                    else None
                )

                # Compute loss with dynamic weighting and geometric gating
                loss, loss_components = self.loss_fn(
                    predictions,
                    targets,
                    shared_params,
                    run_full_geometric=run_full_geometric,
                )

                # Scale loss for gradient accumulation
                loss = loss / self.accumulation_steps

            # Scale and backward pass
            self.scaler.scale(loss).backward()

            # Gradient accumulation step
            if ((batch_idx + 1) % self.accumulation_steps) == 0:
                # Unscale and clip gradients
                self.scaler.unscale_(optimizer)

                # Apply gradient clipping
                torch.nn.utils.clip_grad_norm_(
                    self.model.parameters(), self.config.grad_clip_norm
                )

                # Optimizer step with scaler
                self.scaler.step(optimizer)
                self.scaler.update()
                optimizer.zero_grad()

            current_loss = loss.item() * self.accumulation_steps
            total_loss += current_loss

            # Track component losses
            for name, component_loss in loss_components.items():
                if name != "total":
                    loss_val = (
                        component_loss.item()
                        if torch.is_tensor(component_loss)
                        else component_loss
                    )
                    if name not in component_loss_sums:
                        component_loss_sums[name] = 0
                    component_loss_sums[name] += loss_val

```

```
                batch_count += 1

                # Occasional lightweight logging
                if (batch_idx + 1) % 50 == 0:
                    elapsed = time.time() - epoch_start_time
                    avg_time_per_batch = elapsed / (batch_idx + 1)
                    current_weights = {
                        k: f"{v:.3f}" for k, v in self.loss_fn.weights.items() if v > 0.001
                    }
                    print(
                        f"[Epoch {self.global_epoch}] Batch {batch_idx+1}/{len(self.train_loader)} | "
                        f"avg batch {avg_time_per_batch:.3f}s | loss {total_loss/batch_count:.4f}"
                    )

                # Update progress bar
                current_weights = {
                    k: f"{v:.3f}" for k, v in self.loss_fn.weights.items() if v > 0.001
                }
                train_pbar.set_postfix(
                    {
                        "loss": f"{current_loss:.4f}",
                        "weights": str(current_weights)[:50] + "..."
                        if len(str(current_weights)) > 50
                        else str(current_weights),
                    }
                )

            # Final epoch timing
            epoch_time = time.time() - epoch_start_time
            avg_loss = total_loss / batch_count
            print(
                f"Epoch {self.global_epoch} finished in {epoch_time/60:.2f} min. avg loss: {avg_loss:.4f}"
            )

            # Average component losses
            avg_component_losses = {
                name: loss_sum / batch_count
                for name, loss_sum in component_loss_sums.items()
            }

            return avg_loss, avg_component_losses

        def _prepare_targets(self, batch, mode):
            """Prepare targets based on training mode"""
            if mode == "stage1":
                return {"mask": batch["mask"], "attributes": batch["attributes"]}
            elif mode == "stage2":
                return {
                    "polygons_gt": {
                        "polygons": batch["polygons_gt"]["polygons"].to(self.device),
                        "valid_mask": batch["polygons_gt"]["valid_mask"].to(self.device),
                    }
                }
            else:  # stage3
                return {
                    "mask": batch["mask"],
                    "attributes": batch["attributes"],
                    "voxels_gt": batch["voxels_gt"],
                    "polygons_gt": {
                        "polygons": batch["polygons_gt"]["polygons"].to(self.device),
                        "valid_mask": batch["polygons_gt"]["valid_mask"].to(self.device),
                    },
                }

        def _validate(self, mode="stage1"):
            """Enhanced validation with detailed metrics and AMP support"""
            self.model.eval()
            total_loss = 0
            component_loss_sums = {}

            val_pbar = tqdm(
                self.val_loader, desc=f"Validating {mode.upper()}", leave=False, ncols=120
            )
```

```

        batch_count = 0
        with torch.no_grad():
            for batch in val_pbar:
                batch = {
                    k: v.to(self.device, non_blocking=True) if torch.is_tensor(v) else v
                    for k, v in batch.items()
                }

                with autocast("cuda", enabled=self.use_amp):
                    # Always run full geometric computation during validation
                    predictions = self.model(batch["image"], run_full_geometric=True)

                    # Add latent embeddings if available
                    if hasattr(self.model, "get_latent_embeddings"):
                        latent_2d, latent_3d = self.model.get_latent_embeddings(
                            batch["image"]
                        )
                        predictions["latent_2d_embedding"] = latent_2d
                        predictions["latent_3d_embedding"] = latent_3d

                    targets = self._prepare_targets(batch, mode)

                    loss, loss_components = self.loss_fn(
                        predictions, targets, run_full_geometric=True
                    )

                current_loss = loss.item()
                total_loss += current_loss

                # Track component losses
                for name, component_loss in loss_components.items():
                    if name != "total":
                        loss_val = (
                            component_loss.item()
                            if torch.is_tensor(component_loss)
                            else component_loss
                        )
                        if name not in component_loss_sums:
                            component_loss_sums[name] = 0
                        component_loss_sums[name] += loss_val

                batch_count += 1
                val_pbar.set_postfix({"loss": f"{current_loss:.4f}"})

        avg_component_losses = {
            name: loss_sum / batch_count
            for name, loss_sum in component_loss_sums.items()
        }

        return total_loss / batch_count, avg_component_losses

    def train_stage_adaptive(self, stage: int, max_epochs: int, min_epochs: int):
        """
        Train a stage with adaptive termination based on curriculum learning

        Args:
            stage: Stage number (1, 2, 3)
            max_epochs: Maximum epochs for this stage
            min_epochs: Minimum epochs before considering transition
        """
        print("=" * 60)
        print(f"STAGE {stage}: Adaptive Training with Dynamic Curriculum")
        print("=" * 60)

        self.current_stage = stage
        self.stage_start_time = time.time()

        # Only reset if not resuming
        if not hasattr(self, "epoch_times") or self.epoch_times is None:
            self.epoch_times = []

        start_epoch = int(self.stage_epoch or 0)
```

```python

        # Set parameter gradients for current stage
        self._configure_stage_parameters(stage)

        mode_name = f"stage{stage}"

        for epoch in range(start_epoch, max_epochs):
            epoch_start_time = time.time()
            self.stage_epoch = epoch
            self.global_epoch += 1

            # Update loss weights based on curriculum
            self._update_loss_weights_for_curriculum(stage, epoch, max_epochs)

            print(
                f"\nStage {stage} - Epoch {epoch+1}/{max_epochs} (Global: {self.global_epoch})"
            )

            # Training and validation
            train_loss, train_components = self._train_epoch(mode_name)
            val_loss, val_components = self._validate(mode_name)

            # Record epoch time
            epoch_time = time.time() - epoch_start_time
            self.epoch_times.append(epoch_time)

            if len(self.epoch_times) > 10:
                self.epoch_times.pop(0)

            # Update curriculum state
            self.curriculum_state.update_loss_history(mode_name, val_loss)
            self.curriculum_state.update_component_losses(val_components)

            # Store training history
            self.history[mode_name]["train_loss"].append(train_loss)
            self.history[mode_name]["val_loss"].append(val_loss)
            self.history[mode_name]["component_losses"].append(val_components)

            # Update learning rate
            if stage == 1:
                self.scheduler_2d.step()
            elif stage == 2:
                self.scheduler_dvx.step()
            else:
                self.scheduler_full.step()

            # Display comprehensive results
            self._display_epoch_results(
                epoch,
                max_epochs,
                train_loss,
                val_loss,
                train_components,
                val_components,
                epoch_time,
            )

            # Check for adaptive stage transition
            if epoch >= min_epochs:
                should_transition = self.curriculum_state.should_transition(stage)
                if should_transition:
                    print(
                        f"\n? ADAPTIVE TRANSITION: Stage {stage} converged after {epoch+1} epochs"
                    )
                    print(
                        "   Detected performance plateau - transitioning to next stage"
                    )

                    self.history["stage_transitions"].append(
                        {
                            "from_stage": stage,
                            "epoch": epoch + 1,
                            "global_epoch": self.global_epoch,
```

```
                            "reason": "performance_plateau",
                        }
                    )

                    self.history["curriculum_events"].append(
                        {
                            "type": "stage_transition",
                            "stage": stage,
                            "epoch": self.global_epoch,
                            "details": f"Converged after {epoch+1} epochs",
                        }
                    )
                    break

            # Save rolling checkpoint
            if (epoch + 1) % self.config.checkpoint_freq == 0:
                self._save_rolling_checkpoint()

        print(f"\nStage {stage} completed after {epoch+1} epochs")

    def _configure_stage_parameters(self, stage: int):
        """Configure which parameters require gradients for each stage"""
        # First freeze everything
        for param in self.model.parameters():
            param.requires_grad = False

        if stage == 1:
            # Stage 1: Segmentation + Attributes (2D only)
            for param in self.model.encoder.parameters():
                param.requires_grad = True
            for param in self.model.seg_head.parameters():
                param.requires_grad = True
            for param in self.model.attr_head.parameters():
                param.requires_grad = True
            for param in self.model.sdf_head.parameters():
                param.requires_grad = True

        elif stage == 2:
            # Stage 2: DVX training (polygon fitting) - keep encoder frozen initially
            for param in self.model.dvx.parameters():
                param.requires_grad = True
            # Optionally unfreeze encoder in later epochs
            if self.stage_epoch > 10:
                for param in self.model.encoder.parameters():
                    param.requires_grad = True

        else:  # stage == 3
            # Stage 3: End-to-end fine-tuning (all parameters)
            for param in self.model.parameters():
                param.requires_grad = True

    def _display_epoch_results(
        self,
        epoch: int,
        total_epochs: int,
        train_loss: float,
        val_loss: float,
        train_components: Dict,
        val_components: Dict,
        epoch_time: float,
    ):
        """Display comprehensive epoch results with curriculum information"""
        eta_str = self._get_eta_string(epoch, total_epochs)

        print(f"Train Loss: {train_loss:.4f}, Val Loss: {val_loss:.4f}")
        print(f"Epoch time: {epoch_time:.1f}s, {eta_str}")

        # Show significant component losses
        significant_components = {
            k: v
            for k, v in val_components.items()
            if v > 0.01
            and k
```

```
            in [
                "seg",
                "dice",
                "polygon",
                "voxel",
                "topology",
                "latent_consistency",
                "graph",
            ]
        }
        if significant_components:
            comp_str = ", ".join(
                [f"{k}: {v:.3f}" for k, v in significant_components.items()]
            )
            print(f"Components: {comp_str}")

        # Show current loss weights for active components
        active_weights = {k: v for k, v in self.loss_fn.weights.items() if v > 0.001}
        if active_weights:
            weight_str = ", ".join([f"{k}: {v:.3f}" for k, v in active_weights.items()])
            print(f"Weights: {weight_str}")

        # Show curriculum status
        plateau_epochs = self.curriculum_state.epochs_without_improvement
        if plateau_epochs > 0:
            print(f"Plateau: {plateau_epochs} epochs without improvement")

    def _save_rolling_checkpoint(self):
        """Enhanced checkpoint saving with curriculum state, RNG state, and scaler state"""
        checkpoint = {
            "model_state_dict": self.model.state_dict(),
            "optimizer_2d_state_dict": self.optimizer_2d.state_dict(),
            "optimizer_dvx_state_dict": self.optimizer_dvx.state_dict(),
            "optimizer_full_state_dict": self.optimizer_full.state_dict(),
            "scheduler_2d_state_dict": self.scheduler_2d.state_dict(),
            "scheduler_dvx_state_dict": self.scheduler_dvx.state_dict(),
            "scheduler_full_state_dict": self.scheduler_full.state_dict(),
            "scaler_state_dict": self.scaler.state_dict(),  # Add AMP scaler state
            "loss_fn_state": {
                "weights": self.loss_fn.weights,
                "initial_weights": self.loss_fn.initial_weights,
            },
            "history": self.history,
            "config": self.config,
            "current_stage": self.current_stage,
            "current_epoch": self.current_epoch,
            "global_epoch": self.global_epoch,
            "stage_epoch": self.stage_epoch,
            "epoch_times": self.epoch_times,
            "step_counter": self._step,  # Save step counter for DVX gating
            "curriculum_state": {
                "loss_history": dict(self.curriculum_state.loss_history),
                "component_losses": dict(self.curriculum_state.component_losses),
                "epochs_without_improvement": self.curriculum_state.epochs_without_improvement,
                "best_val_loss": self.curriculum_state.best_val_loss,
                "stage_transition_epochs": self.curriculum_state.stage_transition_epochs,
            },
            "rng_state": {
                "torch": torch.get_rng_state(),
                "cuda": torch.cuda.get_rng_state_all()
                if torch.cuda.is_available()
                else None,
                "numpy": np.random.get_state(),
                "python": random.getstate(),
            },
        }

        checkpoint_path = self.ROLLING_CHECKPOINT
        torch.save(checkpoint, checkpoint_path)
        print(f"Rolling checkpoint saved: {checkpoint_path}")

    def load_checkpoint(self, filename):
        """Enhanced checkpoint loading with curriculum state restoration and device handling"""
```

```
        checkpoint = torch.load(filename, map_location=self.device)

        self.model.load_state_dict(checkpoint["model_state_dict"])
        self.optimizer_2d.load_state_dict(checkpoint["optimizer_2d_state_dict"])
        self.optimizer_dvx.load_state_dict(checkpoint["optimizer_dvx_state_dict"])
        self.optimizer_full.load_state_dict(checkpoint["optimizer_full_state_dict"])

        # Load scaler state for AMP
        if "scaler_state_dict" in checkpoint:
            self.scaler.load_state_dict(checkpoint["scaler_state_dict"])

        # Safer scheduler loading
        for sched_key, sched_obj in [
            ("scheduler_2d_state_dict", self.scheduler_2d),
            ("scheduler_dvx_state_dict", self.scheduler_dvx),
            ("scheduler_full_state_dict", self.scheduler_full),
        ]:
            if sched_key in checkpoint:
                sched_obj.load_state_dict(checkpoint[sched_key])

        # Load loss weights with proper device handling
        if "loss_fn_state" in checkpoint:
            loaded_weights = checkpoint["loss_fn_state"]["weights"]
            if isinstance(loaded_weights, dict):
                self.loss_fn.weights = {
                    k: (v.to(self.device) if torch.is_tensor(v) else v)
                    for k, v in loaded_weights.items()
                }
            else:
                self.loss_fn.weights = loaded_weights
            self.loss_fn.initial_weights = checkpoint["loss_fn_state"][
                "initial_weights"
            ]

        if "history" in checkpoint:
            self.history = checkpoint["history"]

        # Restore training state
        if "current_stage" in checkpoint:
            self.current_stage = checkpoint["current_stage"]
        if "current_epoch" in checkpoint:
            self.current_epoch = checkpoint["current_epoch"]
        if "global_epoch" in checkpoint:
            self.global_epoch = checkpoint["global_epoch"]
        if "stage_epoch" in checkpoint:
            self.stage_epoch = checkpoint["stage_epoch"]
        if "epoch_times" in checkpoint:
            self.epoch_times = checkpoint["epoch_times"]
        if "step_counter" in checkpoint:
            self._step = checkpoint["step_counter"]

        # Restore curriculum state
        if "curriculum_state" in checkpoint:
            cs = checkpoint["curriculum_state"]
            for key, history in cs["loss_history"].items():
                self.curriculum_state.loss_history[key] = deque(
                    history, maxlen=self.config.curriculum.plateau_detection_window * 2
                )
            for key, history in cs["component_losses"].items():
                self.curriculum_state.component_losses[key] = deque(history, maxlen=20)
            self.curriculum_state.epochs_without_improvement = cs.get(
                "epochs_without_improvement", 0
            )
            self.curriculum_state.best_val_loss = cs.get("best_val_loss", float("inf"))
            self.curriculum_state.stage_transition_epochs = cs.get(
                "stage_transition_epochs", []
            )

        # Restore RNG states
        if "rng_state" in checkpoint:
            rs = checkpoint["rng_state"]

            # --- Torch RNG (CPU) ---
```

```
            try:
                torch_state = rs.get("torch", None)
                if torch_state is not None:
                    # If it's already a torch tensor with uint8 dtype, use directly
                    if torch.is_tensor(torch_state) and torch_state.dtype == torch.uint8:
                        torch.set_rng_state(torch_state)
                    else:
                        # Convert lists / numpy arrays / other tensors to uint8 torch tensor
                        torch.set_rng_state(torch.tensor(torch_state, dtype=torch.uint8))
            except Exception as e:
                print(f"Warning: could not restore torch RNG state ({e}), skipping.")

            # --- CUDA RNG (all devices) ---
            try:
                cuda_state = rs.get("cuda", None)
                if cuda_state is not None and torch.cuda.is_available():
                    # cuda_state might be a list of states (one per device)
                    cuda_tensors = []
                    for s in cuda_state:
                        if torch.is_tensor(s) and s.dtype == torch.uint8:
                            cuda_tensors.append(s)
                        else:
                            cuda_tensors.append(torch.tensor(s, dtype=torch.uint8))
                    torch.cuda.set_rng_state_all(cuda_tensors)
            except Exception as e:
                print(f"Warning: could not restore CUDA RNG state ({e}), skipping.")

            # --- numpy RNG ---
            try:
                if "numpy" in rs and rs["numpy"] is not None:
                    np.random.set_state(rs["numpy"])
            except Exception as e:
                print(f"Warning: could not restore numpy RNG state ({e}), skipping.")

            # --- python random RNG ---
            try:
                if "python" in rs and rs["python"] is not None:
                    random.setstate(rs["python"])
            except Exception as e:
                print(f"Warning: could not restore python RNG state ({e}), skipping.")

        # Restore DataLoader sampler states if available
        if "dataloader_state" in checkpoint:
            dl_state = checkpoint["dataloader_state"]
            if dl_state["train_sampler_state"] is not None and hasattr(
                self.train_loader.sampler, "__dict__"
            ):
                try:
                    self.train_loader.sampler.__dict__.update(
                        dl_state["train_sampler_state"]
                    )
                except Exception:
                    print("Warning: Could not restore train_loader sampler state")
            if dl_state["val_sampler_state"] is not None and hasattr(
                self.val_loader.sampler, "__dict__"
            ):
                try:
                    self.val_loader.sampler.__dict__.update(
                        dl_state["val_sampler_state"]
                    )
                except Exception:
                    print("Warning: Could not restore val_loader sampler state")

        print(f"Checkpoint loaded: {filename}")
        print(
            f"Resuming from Stage {self.current_stage}, Global Epoch {self.global_epoch}"
        )
        print(
            f"Curriculum state restored with {self.curriculum_state.epochs_without_improvement} epochs with
        )

    def train_all_stages(self):
        """
```

```
        Run complete adaptive multi-stage training pipeline

        This is the main entry point that orchestrates the dynamic curriculum learning
        """
        if Path(self.ROLLING_CHECKPOINT).exists():
            print(f"Found existing checkpoint: {self.ROLLING_CHECKPOINT}")
            print("Resuming adaptive training from checkpoint...")
            self.load_checkpoint(self.ROLLING_CHECKPOINT)
        else:
            print("Starting fresh adaptive training pipeline...")
            self.current_stage = 1
            self.current_epoch = 0
            self.global_epoch = 0

        print("\n" + "=" * 80)
        print("ADAPTIVE MULTI-STAGE TRAINING WITH DYNAMIC CURRICULUM")
        print("Novel Training Strategies:")
        print("? Adaptive Stage Transitioning (Dynamic Curriculum)")
        print("? Topology-aware Loss Scheduling")
        print("? Multi-objective Optimization with Dynamic Weighting")
        print("? Cross-modal Latent Consistency Learning")
        print("? Graph-based Topology Constraints")
        print("=" * 80)

        # Stage 1: Adaptive 2D training
        if self.current_stage <= 1:
            print("\n? STAGE 1: Adaptive 2D Segmentation + Attributes Training")
            self.train_stage_adaptive(
                stage=1,
                max_epochs=self.config.max_stage1_epochs,
                min_epochs=self.config.min_stage1_epochs,
            )
            self.current_stage = 2
            self.stage_epoch = 0
            print("\nStage 1 completed. Transitioning to Stage 2...")

        # Stage 2: Adaptive DVX training
        if self.current_stage <= 2:
            print("\n? STAGE 2: Adaptive DVX Polygon Fitting Training")
            self.train_stage_adaptive(
                stage=2,
                max_epochs=self.config.max_stage2_epochs,
                min_epochs=self.config.min_stage2_epochs,
            )
            self.current_stage = 3
            self.stage_epoch = 0
            print("\nStage 2 completed. Transitioning to Stage 3...")

        # Stage 3: Adaptive end-to-end fine-tuning
        if self.current_stage <= 3:
            print("\n? STAGE 3: Adaptive End-to-End Fine-tuning with Full Loss Suite")
            self.train_stage_adaptive(
                stage=3,
                max_epochs=self.config.max_stage3_epochs,
                min_epochs=self.config.min_stage3_epochs,
            )
            print("\nStage 3 completed!")

        print("\n" + "=" * 80)
        print("? ALL ADAPTIVE TRAINING STAGES COMPLETED!")
        print("=" * 80)

        # Generate training report
        self._generate_training_report()

        # Save final model
        self._save_checkpoint("final_adaptive_model.pth")

        # Clean up rolling checkpoint
        if Path(self.ROLLING_CHECKPOINT).exists():
            Path(self.ROLLING_CHECKPOINT).unlink()
            print(f"Cleaned up rolling checkpoint: {self.ROLLING_CHECKPOINT}")

```

```
        return self.history

    def _generate_training_report(self):
        """Generate comprehensive training report with curriculum insights"""
        print("\n" + "=" * 60)
        print("ADAPTIVE TRAINING REPORT")
        print("=" * 60)

        # Stage transition summary
        if self.history["stage_transitions"]:
            print("\n? Stage Transitions:")
            for transition in self.history["stage_transitions"]:
                print(
                    f"  ? Stage {transition['from_stage']} ? {transition['from_stage']+1}: "
                    f"Epoch {transition['epoch']} (Global: {transition['global_epoch']})"
                )
                print(f"    Reason: {transition['reason']}")

        # Dynamic weight evolution
        if self.history["dynamic_weights"]:
            print(
                f"\n??  Dynamic Weight Updates: {len(self.history['dynamic_weights'])} updates"
            )
            final_weights = self.history["dynamic_weights"][-1]["weights"]
            print("  Final loss weights:")
            for name, weight in final_weights.items():
                if weight > 0.001:
                    print(f"    {name}: {weight:.3f}")

        # Curriculum events
        if self.history["curriculum_events"]:
            print(
                f"\n? Curriculum Events: {len(self.history['curriculum_events'])} events"
            )
            for event in self.history["curriculum_events"][-5:]:  # Show last 5 events
                print(
                    f"  ? {event['type']} at global epoch {event['epoch']}: {event['details']}"
                )

        # Performance summary
        print("\n? Final Performance:")
        for stage_name, data in self.history.items():
            if isinstance(data, dict) and "val_loss" in data and data["val_loss"]:
                final_loss = data["val_loss"][-1]
                best_loss = min(data["val_loss"])
                print(
                    f"  ? {stage_name.upper()}: Final={final_loss:.4f}, Best={best_loss:.4f}"
                )

        print("\n? Training completed with novel adaptive curriculum strategies!")
        print("=" * 60)

    def _save_checkpoint(self, filename):
        """Save final training checkpoint"""
        checkpoint = {
            "model_state_dict": self.model.state_dict(),
            "optimizer_2d_state_dict": self.optimizer_2d.state_dict(),
            "optimizer_dvx_state_dict": self.optimizer_dvx.state_dict(),
            "optimizer_full_state_dict": self.optimizer_full.state_dict(),
            "scheduler_2d_state_dict": self.scheduler_2d.state_dict(),
            "scheduler_dvx_state_dict": self.scheduler_dvx.state_dict(),
            "scheduler_full_state_dict": self.scheduler_full.state_dict(),
            "scaler_state_dict": self.scaler.state_dict(),
            "loss_fn_state": {
                "weights": self.loss_fn.weights,
                "initial_weights": self.loss_fn.initial_weights,
            },
            "history": self.history,
            "config": self.config,
            "final_stage": self.current_stage,
            "total_epochs": self.global_epoch,
            "training_complete": True,
            "curriculum_summary": {
```

```
                "stage_transitions": len(self.history["stage_transitions"]),
                "weight_updates": len(self.history["dynamic_weights"]),
                "curriculum_events": len(self.history["curriculum_events"]),
            },
        }
        torch.save(checkpoint, filename)
        print(f"Final model saved: {filename}")


# Legacy compatibility class
class MultiStageTrainer(AdaptiveMultiStageTrainer):
    """
    Legacy wrapper for backward compatibility
    Redirects to the new adaptive trainer
    """

    def __init__(self, *args, **kwargs):
        super().__init__(*args, **kwargs)
        print("Note: Using enhanced AdaptiveMultiStageTrainer with dynamic curriculum")

    def train_stage1(self, epochs=None):
        """Legacy method - redirects to adaptive training"""
        max_epochs = epochs or self.config.max_stage1_epochs
        min_epochs = self.config.min_stage1_epochs
        return self.train_stage_adaptive(1, max_epochs, min_epochs)

    def train_stage2(self, epochs=None):
        """Legacy method - redirects to adaptive training"""
        max_epochs = epochs or self.config.max_stage2_epochs
        min_epochs = self.config.min_stage2_epochs
        return self.train_stage_adaptive(2, max_epochs, min_epochs)

    def train_stage3(self, epochs=None):
        """Legacy method - redirects to adaptive training"""
        max_epochs = epochs or self.config.max_stage3_epochs
        min_epochs = self.config.min_stage3_epochs
        return self.train_stage_adaptive(3, max_epochs, min_epochs)
```

■ File: utils\visualization.py

==

```
"""
Visualization and utility functions
"""

import matplotlib.pyplot as plt
import numpy as np
import cv2
import torch
from pathlib import Path
from evaluation.metrics import compute_iou


def plot_training_history(history, save_path="training_history.png"):
    """Plot training curves for all stages"""
    fig, axes = plt.subplots(1, 3, figsize=(15, 5))

    for idx, (stage, data) in enumerate(history.items()):
        if isinstance(data, dict) and "train_loss" in data and data["train_loss"]:  # Only plot if stage wa
            axes[idx].plot(data["train_loss"], label="Train", linewidth=2)
            axes[idx].plot(data["val_loss"], label="Validation", linewidth=2)
            axes[idx].set_title(f"{stage.upper()} Training")
            axes[idx].set_xlabel("Epoch")
            axes[idx].set_ylabel("Loss")
            axes[idx].legend()
            axes[idx].grid(True, alpha=0.3)

    plt.tight_layout()
    plt.savefig(save_path, dpi=300, bbox_inches="tight")
    plt.show()


```

```python
def plot_curriculum_analysis(history, save_path="curriculum_analysis.png"):
    """Plot curriculum learning analysis including stage transitions and adaptive behavior"""
    fig, axes = plt.subplots(2, 2, figsize=(15, 10))

    # Plot 1: Stage transition timeline
    if "stage_transitions" in history and history["stage_transitions"]:
        transitions = history["stage_transitions"]

        # Extract transition epochs and reasons
        transition_epochs = [t["epoch"] for t in transitions]
        transition_stages = [t["from_stage"] + " ? " + t["to_stage"] for t in transitions]
        transition_reasons = [t.get("reason", "threshold") for t in transitions]

        # Create timeline
        y_positions = range(len(transition_epochs))
        colors = ['red' if 'patience' in reason else 'green' for reason in transition_reasons]

        axes[0, 0].barh(y_positions, transition_epochs, color=colors, alpha=0.7)
        axes[0, 0].set_yticks(y_positions)
        axes[0, 0].set_yticklabels(transition_stages)
        axes[0, 0].set_xlabel("Epoch")
        axes[0, 0].set_title("Stage Transition Timeline")
        axes[0, 0].grid(True, alpha=0.3)

        # Add legend
        axes[0, 0].legend(['Patience-based', 'Threshold-based'], loc='lower right')
    else:
        axes[0, 0].text(0.5, 0.5, "No stage transitions recorded",
                       ha='center', va='center', transform=axes[0, 0].transAxes)
        axes[0, 0].set_title("Stage Transition Timeline")

    # Plot 2: Loss component evolution
    if "dynamic_weights" in history and history["dynamic_weights"]:
        weight_data = history["dynamic_weights"]
        epochs = [entry["epoch"] for entry in weight_data]

        # Plot each loss component weight
        weight_names = list(weight_data[0]["weights"].keys()) if weight_data else []
        for weight_name in weight_names[:5]:  # Limit to top 5 for readability
            weights = [entry["weights"].get(weight_name, 0) for entry in weight_data]
            if any(w > 0.001 for w in weights):  # Only plot significant weights
                axes[0, 1].plot(epochs, weights, label=weight_name, linewidth=2, marker='o', markersize=3)

        axes[0, 1].set_xlabel("Global Epoch")
        axes[0, 1].set_ylabel("Loss Weight")
        axes[0, 1].set_title("Dynamic Loss Weight Evolution")
        axes[0, 1].legend()
        axes[0, 1].grid(True, alpha=0.3)
    else:
        axes[0, 1].text(0.5, 0.5, "No dynamic weights recorded",
                       ha='center', va='center', transform=axes[0, 1].transAxes)
        axes[0, 1].set_title("Dynamic Loss Weight Evolution")

    # Plot 3: Curriculum progress indicators
    if "curriculum_events" in history and history["curriculum_events"]:
        events = history["curriculum_events"]
        event_types = {}

        for event in events:
            event_type = event.get("type", "unknown")
            if event_type not in event_types:
                event_types[event_type] = []
            event_types[event_type].append(event["epoch"])

        # Plot event timeline
        y_offset = 0
        for event_type, epochs in event_types.items():
            axes[1, 0].scatter(epochs, [y_offset] * len(epochs),
                              label=event_type, s=50, alpha=0.7)
            y_offset += 1

        axes[1, 0].set_xlabel("Epoch")
        axes[1, 0].set_ylabel("Event Type")
```

```
        axes[1, 0].set_title("Curriculum Learning Events")
        axes[1, 0].legend()
        axes[1, 0].grid(True, alpha=0.3)
    else:
        axes[1, 0].text(0.5, 0.5, "No curriculum events recorded",
                       ha='center', va='center', transform=axes[1, 0].transAxes)
        axes[1, 0].set_title("Curriculum Learning Events")

    # Plot 4: Stage performance comparison
    stage_names = ["stage1", "stage2", "stage3"]
    stage_performance = {}

    for stage_name in stage_names:
        if stage_name in history and isinstance(history[stage_name], dict):
            stage_data = history[stage_name]
            if "val_loss" in stage_data and stage_data["val_loss"]:
                stage_performance[stage_name] = {
                    "final_loss": stage_data["val_loss"][-1],
                    "best_loss": min(stage_data["val_loss"]),
                    "epochs": len(stage_data["val_loss"])
                }

    if stage_performance:
        stages = list(stage_performance.keys())
        final_losses = [stage_performance[s]["final_loss"] for s in stages]
        best_losses = [stage_performance[s]["best_loss"] for s in stages]

        x = np.arange(len(stages))
        width = 0.35

        axes[1, 1].bar(x - width/2, final_losses, width, label='Final Loss', alpha=0.8)
        axes[1, 1].bar(x + width/2, best_losses, width, label='Best Loss', alpha=0.8)

        axes[1, 1].set_xlabel("Training Stage")
        axes[1, 1].set_ylabel("Validation Loss")
        axes[1, 1].set_title("Stage Performance Comparison")
        axes[1, 1].set_xticks(x)
        axes[1, 1].set_xticklabels([s.upper() for s in stages])
        axes[1, 1].legend()
        axes[1, 1].grid(True, alpha=0.3)

        # Add epoch count annotations
        for i, stage in enumerate(stages):
            epochs = stage_performance[stage]["epochs"]
            axes[1, 1].text(i, max(final_losses) * 0.9, f'{epochs} epochs',
                           ha='center', va='bottom', fontsize=9)
    else:
        axes[1, 1].text(0.5, 0.5, "No stage performance data",
                       ha='center', va='center', transform=axes[1, 1].transAxes)
        axes[1, 1].set_title("Stage Performance Comparison")

    plt.tight_layout()
    plt.savefig(save_path, dpi=300, bbox_inches="tight")
    plt.close()

    print(f"Curriculum analysis saved to {save_path}")


def visualize_predictions(image, predictions, targets=None, save_path=None):
    """Visualize model predictions"""
    fig, axes = plt.subplots(2, 3, figsize=(15, 10))

    # Original image
    if len(image.shape) == 4:
        img_np = image[0].permute(1, 2, 0).cpu().numpy()
    else:
        img_np = image.permute(1, 2, 0).cpu().numpy()

    axes[0, 0].imshow(img_np)
    axes[0, 0].set_title("Input Image")
    axes[0, 0].axis('off')

    # Predicted segmentation
```

```
    if "segmentation" in predictions:
        seg_pred = torch.argmax(predictions["segmentation"], dim=1)[0].cpu().numpy()
        axes[0, 1].imshow(seg_pred, cmap='tab10')
        axes[0, 1].set_title("Predicted Segmentation")
        axes[0, 1].axis('off')

    # Ground truth segmentation (if available)
    if targets and "mask" in targets:
        gt_mask = targets["mask"][0].cpu().numpy()
        axes[0, 2].imshow(gt_mask, cmap='tab10')
        axes[0, 2].set_title("Ground Truth Segmentation")
        axes[0, 2].axis('off')

    # SDF prediction
    if "sdf" in predictions:
        sdf_pred = predictions["sdf"][0, 0].cpu().numpy()
        im = axes[1, 0].imshow(sdf_pred, cmap='RdBu', vmin=-1, vmax=1)
        axes[1, 0].set_title("Predicted SDF")
        axes[1, 0].axis('off')
        plt.colorbar(im, ax=axes[1, 0])

    # Polygon visualization
    if "polygons" in predictions:
        poly_vis = visualize_polygons(
            predictions["polygons"][0],
            predictions["polygon_validity"][0],
            image_size=(256, 256)
        )
        axes[1, 1].imshow(poly_vis)
        axes[1, 1].set_title("Predicted Polygons")
        axes[1, 1].axis('off')

    # 3D voxel slice
    if "voxels_pred" in predictions:
        voxels = torch.sigmoid(predictions["voxels_pred"][0]).cpu().numpy()
        # Show middle slice
        mid_slice = voxels[voxels.shape[0]//2]
        axes[1, 2].imshow(mid_slice, cmap='viridis')
        axes[1, 2].set_title("3D Voxels (Mid Slice)")
        axes[1, 2].axis('off')

    plt.tight_layout()

    if save_path:
        plt.savefig(save_path, dpi=300, bbox_inches="tight")

    plt.show()


def visualize_polygons(polygons, validity, image_size=(256, 256), threshold=0.5):
    """Visualize predicted polygons"""
    vis_img = np.zeros((*image_size, 3), dtype=np.uint8)

    for poly_idx, (polygon, valid_score) in enumerate(zip(polygons, validity)):
        if valid_score > threshold:
            # Convert to image coordinates
            points = polygon.cpu().numpy() * np.array(image_size)

            # Remove zero-padded points
            valid_points = points[points.sum(axis=1) > 0]

            if len(valid_points) >= 3:
                points_int = valid_points.astype(np.int32)

                # Different colors for different polygons
                color = plt.cm.tab10(poly_idx % 10)[:3]
                color = tuple(int(c * 255) for c in color)

                cv2.polylines(vis_img, [points_int], True, color, 2)

                # Add polygon index
                center = points_int.mean(axis=0).astype(int)
                cv2.putText(vis_img, str(poly_idx), tuple(center),
```

```
                              cv2.FONT_HERSHEY_SIMPLEX, 0.5, (255, 255, 255), 1)

    return vis_img


def save_model_outputs(predictions, output_dir, sample_id):
    """Save all model outputs for detailed analysis"""
    output_dir = Path(output_dir)
    output_dir.mkdir(exist_ok=True)

    sample_dir = output_dir / sample_id
    sample_dir.mkdir(exist_ok=True)

    # Save segmentation
    if "segmentation" in predictions:
        seg_pred = torch.argmax(predictions["segmentation"], dim=1)[0].cpu().numpy()
        cv2.imwrite(str(sample_dir / "segmentation.png"), seg_pred * 50)

    # Save SDF
    if "sdf" in predictions:
        sdf_pred = predictions["sdf"][0, 0].cpu().numpy()
        sdf_normalized = ((sdf_pred + 1) * 127.5).astype(np.uint8)
        cv2.imwrite(str(sample_dir / "sdf.png"), sdf_normalized)

    # Save attributes
    if "attributes" in predictions:
        attrs = predictions["attributes"][0].cpu().numpy()
        np.save(sample_dir / "attributes.npy", attrs)

    # Save polygons
    if "polygons" in predictions:
        polygons = predictions["polygons"][0].cpu().numpy()
        validity = predictions["polygon_validity"][0].cpu().numpy()

        np.save(sample_dir / "polygons.npy", polygons)
        np.save(sample_dir / "polygon_validity.npy", validity)

    # Save voxels
    if "voxels_pred" in predictions:
        voxels = torch.sigmoid(predictions["voxels_pred"][0]).cpu().numpy()
        np.save(sample_dir / "voxels.npy", voxels)


def create_comparison_grid(input_images, predictions, targets=None, num_samples=4):
    """Create a comparison grid showing inputs, predictions, and targets"""
    fig, axes = plt.subplots(num_samples, 4, figsize=(16, 4 * num_samples))

    for i in range(min(num_samples, len(input_images))):
        # Input image
        img = input_images[i].permute(1, 2, 0).cpu().numpy()
        axes[i, 0].imshow(img)
        axes[i, 0].set_title(f"Sample {i+1}: Input")
        axes[i, 0].axis('off')

        # Predicted segmentation
        seg_pred = torch.argmax(predictions["segmentation"][i], dim=0).cpu().numpy()
        axes[i, 1].imshow(seg_pred, cmap='tab10')
        axes[i, 1].set_title("Predicted Seg")
        axes[i, 1].axis('off')

        # Ground truth segmentation (if available)
        if targets and "mask" in targets:
            gt_mask = targets["mask"][i].cpu().numpy()
            axes[i, 2].imshow(gt_mask, cmap='tab10')
            axes[i, 2].set_title("GT Segmentation")
        else:
            axes[i, 2].text(0.5, 0.5, "No GT", ha='center', va='center',
                               transform=axes[i, 2].transAxes)
            axes[i, 2].set_title("GT Segmentation")
        axes[i, 2].axis('off')

        # Polygon overlay
        poly_vis = visualize_polygons(
```

```
            predictions["polygons"][i],
            predictions["polygon_validity"][i]
        )
        axes[i, 3].imshow(poly_vis)
        axes[i, 3].set_title("Predicted Polygons")
        axes[i, 3].axis('off')

    plt.tight_layout()
    return fig


def analyze_failure_cases(predictions, targets, threshold_iou=0.5):
    """Analyze failure cases for debugging"""
    failure_indices = []

    for i, (pred_seg, gt_mask) in enumerate(zip(predictions["segmentation"], targets["mask"])):
        seg_pred = torch.argmax(pred_seg, dim=0)
        iou = compute_iou(seg_pred, gt_mask)

        if iou < threshold_iou:
            failure_indices.append({
                "index": i,
                "iou": iou,
                "pred_classes": torch.unique(seg_pred).tolist(),
                "gt_classes": torch.unique(gt_mask).tolist()
            })

    return failure_indices


class ProgressiveVisualization:
    """Track and visualize training progress"""

    def __init__(self, save_dir="./training_progress"):
        self.save_dir = Path(save_dir)
        self.save_dir.mkdir(exist_ok=True)

    def log_epoch_results(self, epoch, stage, predictions, targets, sample_image):
        """Log results for a specific epoch"""
        epoch_dir = self.save_dir / f"{stage}_epoch_{epoch}"
        epoch_dir.mkdir(exist_ok=True)

        # Save prediction visualization
        fig = plt.figure(figsize=(12, 8))
        visualize_predictions(sample_image, predictions, targets)
        plt.savefig(epoch_dir / "predictions.png", dpi=150, bbox_inches="tight")
        plt.close()

        # Save individual outputs
        save_model_outputs(predictions, epoch_dir, "sample")

    def create_training_animation(self, stage, metric_name="total_loss"):
        """Create animated GIF showing training progress"""
        # This would create an animation of training progress
        # Implementation depends on having saved epoch results
        pass


def compute_architectural_metrics(predictions, image_size=(256, 256)):
    """Compute architecture-specific metrics"""
    metrics = {}

    if "segmentation" in predictions:
        seg_pred = torch.argmax(predictions["segmentation"], dim=1)[0]

        # Room count
        room_mask = (seg_pred == 0).cpu().numpy().astype(np.uint8)
        contours, _ = cv2.findContours(room_mask, cv2.RETR_EXTERNAL, cv2.CHAIN_APPROX_SIMPLE)
        room_count = len([c for c in contours if cv2.contourArea(c) > 100])
        metrics["room_count"] = room_count

        # Wall connectivity
        wall_mask = (seg_pred == 1).cpu().numpy().astype(np.uint8)
```

```
        wall_components = cv2.connectedComponents(wall_mask)[0] - 1  # Subtract background
        metrics["wall_components"] = max(0, wall_components)

        # Door and window counts
        door_pixels = (seg_pred == 2).sum().item()
        window_pixels = (seg_pred == 3).sum().item()
        metrics["door_pixels"] = door_pixels
        metrics["window_pixels"] = window_pixels

    if "polygons" in predictions:
        validity = predictions["polygon_validity"][0]
        valid_polygons = (validity > 0.5).sum().item()
        metrics["valid_polygon_count"] = valid_polygons

        # Average polygon area
        polygons = predictions["polygons"][0]
        areas = []
        for poly_idx, (polygon, valid) in enumerate(zip(polygons, validity)):
            if valid > 0.5:
                # Compute polygon area using shoelace formula
                points = polygon.cpu().numpy() * np.array(image_size)
                valid_points = points[points.sum(axis=1) > 0]
                if len(valid_points) >= 3:
                    area = compute_polygon_area(valid_points)
                    areas.append(area)

        metrics["avg_polygon_area"] = np.mean(areas) if areas else 0.0

    return metrics


def compute_polygon_area(points):
    """Compute polygon area using shoelace formula"""
    if len(points) < 3:
        return 0.0

    x = points[:, 0]
    y = points[:, 1]

    # Shoelace formula
    area = 0.5 * abs(sum(x[i] * y[i+1] - x[i+1] * y[i] for i in range(-1, len(x)-1)))
    return area


def create_model_summary_report(model, sample_input, save_path="model_summary.txt"):
    """Create detailed model summary report"""
    with open(save_path, "w") as f:
        f.write("Neural-Geometric 3D Model Generator - Model Summary\n")
        f.write("=" * 60 + "\n\n")

        # Model architecture
        f.write("MODEL ARCHITECTURE:\n")
        f.write("-" * 20 + "\n")

        total_params = sum(p.numel() for p in model.parameters())
        trainable_params = sum(p.numel() for p in model.parameters() if p.requires_grad)

        f.write(f"Total parameters: {total_params:,}\n")
        f.write(f"Trainable parameters: {trainable_params:,}\n")
        f.write(f"Model size: {total_params * 4 / 1024 / 1024:.2f} MB\n\n")

        # Component breakdown
        f.write("COMPONENT PARAMETERS:\n")
        f.write("-" * 25 + "\n")

        encoder_params = sum(p.numel() for p in model.encoder.parameters())
        seg_params = sum(p.numel() for p in model.seg_head.parameters())
        attr_params = sum(p.numel() for p in model.attr_head.parameters())
        sdf_params = sum(p.numel() for p in model.sdf_head.parameters())
        dvx_params = sum(p.numel() for p in model.dvx.parameters())
        ext_params = sum(p.numel() for p in model.extrusion.parameters())

        f.write(f"Encoder: {encoder_params:,} ({encoder_params/total_params*100:.1f}%)\n")
```

```
        f.write(f"Segmentation Head: {seg_params:,} ({seg_params/total_params*100:.1f}%)\n")
        f.write(f"Attribute Head: {attr_params:,} ({attr_params/total_params*100:.1f}%)\n")
        f.write(f"SDF Head: {sdf_params:,} ({sdf_params/total_params*100:.1f}%)\n")
        f.write(f"DVX Module: {dvx_params:,} ({dvx_params/total_params*100:.1f}%)\n")
        f.write(f"Extrusion Module: {ext_params:,} ({ext_params/total_params*100:.1f}%)\n\n")

        # Forward pass analysis
        f.write("FORWARD PASS ANALYSIS:\n")
        f.write("-" * 25 + "\n")

        model.eval()
        with torch.no_grad():
            predictions = model(sample_input)

            for key, value in predictions.items():
                if torch.is_tensor(value):
                    f.write(f"{key}: {list(value.shape)} - {value.dtype}\n")
                else:
                    f.write(f"{key}: {type(value)}\n")

    print(f"Model summary saved to {save_path}")


def debug_gradient_flow(model, loss):
    """Debug gradient flow through the model"""
    print("Gradient Flow Analysis:")
    print("-" * 30)

    total_norm = 0
    component_norms = {}

    for name, param in model.named_parameters():
        if param.grad is not None:
            param_norm = param.grad.norm().item()
            total_norm += param_norm ** 2

            # Group by component
            component = name.split('.')[0]
            if component not in component_norms:
                component_norms[component] = 0
            component_norms[component] += param_norm ** 2

    total_norm = total_norm ** 0.5

    print(f"Total gradient norm: {total_norm:.4f}")
    print("Component gradient norms:")

    for component, norm in component_norms.items():
        norm = norm ** 0.5
        print(f"  {component}: {norm:.4f} ({norm/total_norm*100:.1f}%)")


def create_3d_visualization(voxels, output_path="3d_preview.png"):
    """Create 3D visualization of voxel prediction"""
    try:
        import matplotlib.pyplot as plt
        from mpl_toolkits.mplot3d import Axes3D

        # Convert to binary
        if isinstance(voxels, torch.Tensor):
            voxels = voxels.cpu().numpy()

        binary_voxels = voxels > 0.5

        # Get occupied voxel coordinates
        occupied = np.where(binary_voxels)

        if len(occupied[0]) == 0:
            print("No occupied voxels to visualize")
            return

        # Create 3D plot
        fig = plt.figure(figsize=(10, 8))
```

```
        ax = fig.add_subplot(111, projection='3d')

        # Plot occupied voxels
        ax.scatter(occupied[0], occupied[1], occupied[2],
                  c=occupied[2], cmap='viridis', s=1, alpha=0.6)

        ax.set_xlabel('X')
        ax.set_ylabel('Y')
        ax.set_zlabel('Z')
        ax.set_title('3D Voxel Occupancy')

        plt.savefig(output_path, dpi=150, bbox_inches="tight")
        plt.close()

        print(f"3D visualization saved to {output_path}")

    except ImportError:
        print("3D visualization requires matplotlib with 3D support")
```

■ File: vortex.py

==

```
import cv2
import numpy as np
from pathlib import Path
import json

# Base dataset path
data_root = Path("./data/floorplans")

def is_valid_mask(mask_file):
    m = cv2.imread(str(mask_file), 0)
    return m is not None and np.sum(m) > 0

def is_valid_voxel(voxel_file):
    try:
        data = np.load(str(voxel_file))
        key = "voxels"  # Use the correct key in your npz files
        if key not in data.files:
            print(f"?? Key '{key}' not found in {voxel_file}, available keys: {data.files}")
            return False
        v = data[key]
        return not np.isnan(v).any()
    except Exception as e:
        print(f"?? Error reading {voxel_file}: {e}")
        return False

def is_valid_polygon(polygon_file):
    try:
        with open(polygon_file, "r") as f:
            json.load(f)
        return True
    except Exception as e:
        print(f"?? Invalid polygon {polygon_file}: {e}")
        return False

def check_split(split="train"):
    split_dir = data_root / split
    mask_files = list(split_dir.rglob("mask.png"))
    voxel_files = list(split_dir.rglob("voxel_GT.npz"))
    polygon_files = list(split_dir.rglob("polygon.json"))

    print(f"\n? Checking split: {split}")

    # Check mask files
    total_masks = len(mask_files)
    bad_masks = 0
    for f in mask_files:
        if not is_valid_mask(f):
            bad_masks += 1
            print(f"?? Invalid mask: {f}")

```

```
    # Check voxel files
    total_voxels = len(voxel_files)
    bad_voxels = 0
    for f in voxel_files:
        if not is_valid_voxel(f):
            bad_voxels += 1
            print(f"?? Invalid voxel: {f}")

    # Check polygon files
    total_polygons = len(polygon_files)
    bad_polygons = 0
    for f in polygon_files:
        if not is_valid_polygon(f):
            bad_polygons += 1

    # Summary
    print(f"\n? Summary for split: {split}")
    print(f"Total mask files checked: {total_masks}, Invalid: {bad_masks}")
    print(f"Total voxel files checked: {total_voxels}, Invalid: {bad_voxels}")
    print(f"Total polygon files checked: {total_polygons}, Invalid: {bad_polygons}")

def main():
    for split in ["train", "val", "test"]:
        check_split(split)

if __name__ == "__main__":
    main()
```
