# Technical Documentation: Advanced Merging of SDXL Checkpoints

Version: 1.0

Date: July 14, 2025

Author: Award-Winning Technical Writer

## Abstract

This document provides a comprehensive technical guide to the process of merging Stable Diffusion XL (SDXL) model checkpoints, with a specific focus on navigating the challenges presented by architecturally divergent models. It details the foundational architecture of SDXL 1.0, analyzes key fine-tuned variants (Illustrious, NoobAI, RouWei), and explains the root causes of merging incompatibilities such as "Tensor Shape Mismatch" errors. Practical, step-by-step instructions are provided for performing both standard block-weighted merging and advanced surgical merging techniques using the ComfyUI-DareMerge extension within the ComfyUI environment. The document concludes with best practices for evaluation, a summary of potential outcomes, and strategic recommendations for achieving successful, coherent model hybrids.

## 1. Foundational Concepts: The SDXL 1.0 Architecture

A prerequisite for successful model merging is a thorough understanding of the baseline SDXL 1.0 architecture. Its components are a highly interdependent system, and their compatibility is the primary determinant of a merge's success.

### 1.1. The UNet Backbone

The core of SDXL is its UNet, a deep convolutional neural network that performs the iterative denoising process.

* **Parameter Count:** The SDXL UNet contains approximately 2.6 billion parameters, a nearly threefold increase over the 860 million in Stable Diffusion v1.5, enabling higher fidelity and a native 1024x1024 resolution.1
* **Hybrid Architecture:** It strategically integrates transformer blocks at the mid-levels of the UNet, enhancing its ability to interpret complex prompt-to-image relationships without excessive computational overhead.1
* **Block Structure:** The UNet is divided into input blocks, a middle block, and output blocks, which form the basis for block-weighted merging techniques.5

### 1.2. Dual Text Encoder System

SDXL utilizes two distinct text encoders to achieve a more nuanced understanding of prompts.2

1. **OpenAI CLIP ViT-L:** Provides a 768-dimensional embedding vector, maintaining some continuity with SDv1.5 prompting styles.2
2. **OpenCLIP ViT-bigG:** A larger encoder that outputs a 1280-dimensional embedding vector.1

The embeddings from both are concatenated, creating a 2048-dimensional conditioning vector that is fed into the UNet's cross-attention layers.2

### 1.3. Pooled Embeddings and Global Context

In addition to per-token embeddings, the OpenCLIP ViT-bigG model generates a single 1280-dimensional "pooled" vector representing the entire prompt's global theme. This provides a secondary contextual signal to the UNet, improving overall compositional coherence.1

### 1.4. The Optional Refiner Model

The full SDXL pipeline is a two-stage "ensemble of experts" process.3

* **Base Model:** Generates the initial latent image from noise.7
* **Refiner Model:** An optional, separate diffusion model trained to operate on low-noise latents, adding high-frequency details and correcting minor artifacts.1 While not directly involved in the base model merging process, its function is integral to the complete SDXL workflow.

## 2. Analysis of Divergent Checkpoints

The open-source community has produced numerous fine-tuned SDXL models. The checkpoints under review—Illustrious, NoobAI, and RouWei—represent a distinct lineage of anime-style models with significant architectural divergences.

### 2.1. SDXL Illustrious

* **Lineage:** Built upon the kohaku-xl-beta5 checkpoint, not directly from SDXL 1.0 base.9
* **Training Data:** Heavily trained on the Danbooru dataset, making it highly proficient with tag-based prompting.11
* **Architectural Divergence:**
  + **Native High Resolution:** Supports native generation up to 1536x1536 pixels, a significant departure from SDXL's 1024x1024 native resolution, suggesting modifications to internal layers.12
  + **Hybrid Prompting:** Fine-tuned to understand both Danbooru tags and natural language, indicating sophisticated adjustments to the text encoders.12
* **Compatibility:** Noted as "not fully compatible with 'regular' SDXL LoRAs, ControlNets, etc.".9

### 2.2. SDXL NoobAI

* **Lineage:** A direct fine-tune of an early version of Illustrious (Illustrious-xl-early-release-v0).15
* **Training Data:** Trained on a massive dataset of ~13 million images from Danbooru and e621, making it an expert in tag-based prompting for anime and furry content.15
* **Architectural Divergence:**
  + **Prediction Type:** Released in two incompatible versions: eps-prediction (standard) and v-prediction. These two types cannot be merged directly as their UNets are trained to predict different targets.19
  + **Extreme Specialization:** The text encoders and UNet are heavily optimized for a structured, symbolic language (Danbooru tags).15
* **Compatibility:** Largely incompatible with the standard SDXL ecosystem and requires its own set of compatible tools.19

### 2.3. SDXL RouWei

* **Lineage:** Exists within the same anime-focused family, often used as a base for further merges.22
* **Architectural Features:**
  + It is an eps-prediction model, making it compatible with Illustrious and the eps version of NoobAI.22
  + Its text encoder configuration is consistent with the standard OpenCLIP component of SDXL.23
  + Supports native high-resolution generation (up to 1536px) and has a baked-in VAE.22

### 2.4. Comparative Analysis Table

| Model Name | Base Model/Lineage | Key Training Data | Native Resolution | Prediction Type | Text Encoder Characteristics | Core Strengths | Known Incompatibilities |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **SDXL 1.0 Base** | N/A | Broad, general-purpose dataset | 1024x1024 3 | Epsilon (eps) | Dual CLIP (ViT-L + OpenCLIP ViT-bigG) 1 | Versatility, good prompt understanding, photorealism | Incompatible with SD1.5/2.x LoRAs and ControlNets 24 |
| **SDXL Illustrious** | SDXL (via kohaku-xl-beta5) 9 | Danbooru dataset (anime/illustration tags) 12 | Up to 1536x1536 12 | Epsilon (eps) | Fine-tuned for hybrid tag/natural language prompting 12 | High-quality anime style, high resolution, good anatomy | Not fully compatible with standard SDXL LoRAs/ControlNets 9 |
| **SDXL NoobAI (eps)** | Illustrious-xl-early-release-v0 15 | Danbooru & e621 datasets (~13M images) 15 | Up to 1344x768 (recommended) 15 | Epsilon (eps) | Heavily optimized for Danbooru tag structure 15 | Excellent tag adherence, character/artist knowledge, furry content | Incompatible with standard SDXL ecosystem; requires own ControlNets 19 |
| **SDXL NoobAI (v-pred)** | Illustrious-xl-early-release-v0 21 | Danbooru & e621 datasets 17 | Up to 1344x768 (recommended) 21 | V-prediction (v) | Same as eps version but with different prediction target | Same as eps version; may offer different quality trade-offs | Incompatible with eps-prediction models, including base SDXL and Illustrious 20 |
| **SDXL RouWei** | Illustrious/NoobAI lineage 22 | 35k+ artist styles, gacha game images 22 | Up to 1536x1536 22 | Epsilon (eps) | Uses Illustrious/NoobAI prompting style 22 | Vibrant colors, aesthetic stability, good anatomy | Shares incompatibilities with the Illustrious/NoobAI family |

## 3. Core Merging Challenges & Artifacts

Model merging is frequently obstructed by technical barriers that manifest as visual artifacts or outright failures.

### 3.1. The "Tensor Shape Mismatch" Error

This error is the most common symptom of architectural incompatibility. Model merging is a mathematical operation on tensors, which assumes that corresponding layers in both parent models have identical tensor dimensions (shapes).5 Intensive fine-tuning can alter these shapes, making a direct merge impossible.9

**Example Diagnosis:** An error like size mismatch for model.diffusion\_model.input\_blocks.1.1.transformer\_blocks.0.attn2.to\_v.weight:... shape torch.Size()... shape in current model is torch.Size() indicates that the cross-attention layers of the two models are built to accept different-sized text conditioning vectors (1024 vs. 768 dimensions), a fundamental incompatibility.30

Another common mismatch occurs when merging standard models with inpainting models. The UNet of an inpainting model is architecturally modified to accept 9 input channels (4 for latent, 5 for masked image and mask) instead of the standard 4, causing a shape mismatch at the very first layer.32

### 3.2. Common Visual Artifacts and Failures

* **"Fried" Models:** A catastrophic failure resulting in noisy, desaturated, or chaotic images. Often caused by merging incompatible architectures or using overly aggressive ratios. A common sign is the appearance of purple or blue splotches on skin tones.34
* **Style Bleed:** Undesirable elements from one model's style contaminate the other, leading to an aesthetically messy result.
* **Concept Confusion:** The merged model's understanding of concepts becomes muddled, leading to poor prompt adherence and bizarre hybrid objects.35
* **Inherited Flaws:** The merged model will inherit the biases and artifacts (e.g., watermarks) present in its parent models.22

## 4. Merging Methodologies in ComfyUI

ComfyUI provides a node-based interface for both basic and advanced model merging.

### 4.1. Standard Block-Weighted Merging

This technique offers granular control by adjusting the influence of the three main sections of the UNet.

**Essential Nodes:**

* Load Checkpoint: Two nodes to load the parent models.
* ModelMergeBlocks: The primary node for block-weighted merging.5
* KSampler: To generate test images.
* Save Checkpoint: To save the final merged model.39

**Block Heuristics:**

* **Input Blocks:** Tend to control overall composition, structure, and color palette.5
* **Middle Block:** Affects both composition and finer details.
* **Output Blocks:** Have a strong impact on final details, textures, lighting, and subject style.5

**Procedural Walkthrough (Illustrious + RouWei):**

1. **Setup:** Load Illustrious as Model A and RouWei as Model B into two Load Checkpoint nodes. Connect them to a ModelMergeBlocks node.
2. **Systematic Test:** Generate a test grid using a fixed prompt and seed. Create an image for each of the eight primary permutations of the block weights (e.g., IN=0, MID=0, OUT=1; IN=0, MID=1, OUT=0; etc.) to isolate the effect of each block.5
3. **Analysis:** Compare the eight images to identify which blocks from each model contribute the desired characteristics.
4. **Refinement:** Move from binary (0.0/1.0) weights to fractional weights (e.g., 0.7) to create a more subtle blend.
5. **Save:** Once a desirable result is achieved, use the CheckpointSave node to save the new model.

### 4.2. Advanced Surgical Merging with ComfyUI-DareMerge

For highly divergent models, the ComfyUI-DareMerge extension provides advanced algorithms for more intelligent integration.40

DARE-TIES Methodology:

This approach, based on recent research, treats a fine-tune not as a block of weights, but as a set of skills or changes applied to a base model.40

* **DARE (Drop and Rescale):** Randomly drops a fraction of the weight differences (deltas) between models and rescales the remainder to prevent conflicting edits.
* **TIES-Merging (Trim, Elect, and Merge):** A more advanced method that trims low-impact deltas, resolves sign conflicts, and merges only the most significant, non-conflicting changes.41

**Key DareMerge Nodes:**

* Model Merger (Advanced/DARE): Implements the full DARE-TIES algorithm, ideal for incompatible models.40
* Magnitude Masker: Creates a mask by identifying the parameters with the largest changes between a fine-tuned model and its base. This mask represents the "essence" of the fine-tune and can be used to protect critical parameters during a merge.40

**Procedural Walkthrough (Transferring NoobAI skills to Illustrious):**

1. **Define Roles:**
   * **Host Model:** Illustrious.safetensors (the model to preserve).
   * **Donor Model:** NoobAI-eps.safetensors (the model providing the skills).
   * **Reference Model:** A common ancestor (e.g., kohaku-xl-beta5) to calculate the deltas.
2. **Create Protective Mask:** Use a Magnitude Masker node with the Host and Reference models to create a mask that identifies and protects the most critical parameters of Illustrious.
3. **Perform Surgical Merge:** Use a Model Merger (Advanced/DARE) node. Connect the Host (model1), Donor (model2), and the protective mask. This instructs the algorithm to apply changes from the Donor *only* to the unprotected areas of the Host.
4. **Iterate and Evaluate:** The DARE process is stochastic. Experiment with different seeds and the density parameter to find the optimal result.40 Test the final model with prompts that specifically target the skills you intended to transfer.

## 5. Best Practices and Evaluation Protocol

A systematic approach is crucial for producing and validating high-quality merged models.

### 5.1. Evaluation Protocol

1. **Isolate Variables:** When testing, change only one parameter at a time (e.g., a single merge ratio), keeping the prompt, seed, CFG, and sampler constant for a true comparison.5
2. **Comprehensive Prompting:** Test the merged model with a diverse range of prompts, including simple subjects, complex scenes, and abstract concepts, to assess its versatility.
3. **Stability Testing:** Use a fixed prompt and generate images with a large number of different seeds to check for consistency and stability.
4. **Visual Comparison:** Use a tool like PureRef to arrange and compare test outputs side-by-side to identify subtle differences and trends.5
5. **Prompt Engineering:** A merged model is a new entity. Expect to perform new prompt engineering to discover the keywords and structures that work best with its hybrid architecture.35

### 5.2. Strategic Recommendations

* **Research Lineage:** Prioritize merging models from the same "family" or with a close common ancestor.
* **Define a Clear Goal:** Articulate what you want to achieve (e.g., "combine style X with knowledge Y") before you begin.
* **Use LoRAs for Augmentation:** For adding a specific character or a minor style tweak, a LoRA is the appropriate, non-destructive tool.34
* **Use Merging for Hybridization:** Use merging when the goal is to create a fundamentally new base model with a hybrid core.

## 6. Conclusion

Merging SDXL models is a powerful technique for creating novel generative tools, but it demands a deep understanding of the underlying architectures and a methodical approach. While simple block merging is effective for closely related models, the architectural divergence in highly specialized checkpoints necessitates advanced techniques like DARE-TIES, available through extensions like ComfyUI-DareMerge. By respecting model incompatibilities, employing surgical merging methods, and adhering to a rigorous evaluation protocol, practitioners can successfully navigate the complexities of this process to produce unique, stable, and high-quality hybrid models. The continued development of such advanced merging algorithms points toward a future of increasingly modular and customizable generative AI.

#### Cytowane prace

1. Deep Dive into Stability AI's Generative Models - Stable ... - GitHub, otwierano: lipca 14, 2025, <https://github.com/neobundy/Deep-Dive-Into-AI-With-MLX-PyTorch/blob/main/deep-dives/008-stable-diffusion-sdxl/README2.md>
2. The SDXL Model Pipeline, otwierano: lipca 14, 2025, <https://www.xta0.me/2025/01/20/GenAI-Stable-Diffusion-SDXL.html>
3. Stable Diffusion XL: Everything You Need to Know - Magai, otwierano: lipca 14, 2025, <https://magai.co/stable-diffusion-xl-1-0/>
4. Best settings for Stable Diffusion SDXL 1.0 - Generative AI Blog by Segmind, otwierano: lipca 14, 2025, <https://blog.segmind.com/sdxl-1-0-settings-guide/>
5. A First Timer's Guide to Cooking Up a Fine Tuned SDXL Model - Rob Laughter, otwierano: lipca 14, 2025, <https://roblaughter.medium.com/a-first-timers-guide-to-cooking-up-a-fine-tuned-sdxl-model-446652d018f4>
6. Checkpoint merge blocks : r/comfyui - Reddit, otwierano: lipca 14, 2025, <https://www.reddit.com/r/comfyui/comments/1b0xdwf/checkpoint_merge_blocks/>
7. Ultimate Guide to Stable Diffusion's SDXL Model (2024) | Shakker AI, otwierano: lipca 14, 2025, <https://wiki.shakker.ai/en/stable-diffusion-sdxl-guide>
8. Adventures with Stable Diffusion XL | by Hitoruna | Jul, 2025 - Medium, otwierano: lipca 14, 2025, <https://medium.com/@hitorunajp/adventures-with-stable-diffusion-xl-7617fdb80f43>
9. What is Illustrious/IllustriousXL? : r/StableDiffusion - Reddit, otwierano: lipca 14, 2025, <https://www.reddit.com/r/StableDiffusion/comments/1jjubsr/what_is_illustriousillustriousxl/>
10. Illustrious Model Series Guide created with SeaArt AI, otwierano: lipca 14, 2025, <https://www.seaart.ai/articleDetail/ctuib45e878c73ekrc90>
11. Illustrious XL-Modular WF v1.0 - MimicPC, otwierano: lipca 14, 2025, <https://www.mimicpc.com/workflows/illustrious-xl-modular-wf-v10>
12. Comprehensive Guide of Illustrious XL - Tensor.Art, otwierano: lipca 14, 2025, <https://tensor.art/articles/831123524065191393>
13. illustrious-xl-v1.0 - PromptLayer, otwierano: lipca 14, 2025, <https://www.promptlayer.com/models/illustrious-xl-v10>
14. Illustrious XL - v1.0 - v1.0 | Stable Diffusion Model - CHECKPOINT | Tensor.Art, otwierano: lipca 14, 2025, <https://tensor.art/models/828830887698374324>
15. NOOBAI XL | SeaArt Guide, otwierano: lipca 14, 2025, <https://docs.seaart.ai/guide-1/6-permanent-events/high-quality-models-recommendation/noobai-xl>
16. Noobai XL 1.0 · Models - Dataloop, otwierano: lipca 14, 2025, <https://dataloop.ai/library/model/laxhar_noobai-xl-10/>
17. NoobAI-XL (NAI-XL) - Epsilon-pred 1.0-Version - Civitai, otwierano: lipca 14, 2025, <https://civitai.com/models/833294?modelVersionId=1022833>
18. WAI-NSFW-illustrious-SDXL VS NoobAI-XL (NAI-XL) which one is better? - Reddit, otwierano: lipca 14, 2025, <https://www.reddit.com/r/comfyui/comments/1j45dqm/wainsfwillustrioussdxl_vs_noobaixl_naixl_which/>
19. Base Models | Krita AI Handbook - Interstice.cloud, otwierano: lipca 14, 2025, <https://docs.interstice.cloud/base-models/>
20. NoobAI-XL ControlNet - eps-canny - Civitai, otwierano: lipca 14, 2025, <https://civitai.com/models/929685?modelVersionId=1040650>
21. Noobai XL Vpred 0.75 Cyberfix · Models - Dataloop, otwierano: lipca 14, 2025, <https://dataloop.ai/library/model/panchovix_noobai-xl-vpred-075-cyberfix/>
22. JANKU v4 NSFW Trained + NoobAI EPS + RouWei Illustrious XL ..., otwierano: lipca 14, 2025, <https://civitai.com/models/1277670/janku-v4-nsfw-trained-noobai-eps-rouwei-illustrious-xl>
23. text\_encoder\_2/config.json · John6666/rouwei-v080-epsilon-sdxl at main - Hugging Face, otwierano: lipca 14, 2025, <https://huggingface.co/John6666/rouwei-v080-epsilon-sdxl/blob/main/text_encoder_2/config.json>
24. WARNING: shape mismatch when trying to apply embedding, embedding will be ignored 768 1280 : r/StableDiffusion - Reddit, otwierano: lipca 14, 2025, <https://www.reddit.com/r/StableDiffusion/comments/15gpqw3/warning_shape_mismatch_when_trying_to_apply/>
25. Exception using ControlNet model control\_v11f1e\_sd15\_tile not compatible with SDXL models. · Issue #139 - GitHub, otwierano: lipca 14, 2025, <https://github.com/lllyasviel/ControlNet-v1-1-nightly/issues/139>
26. How to Merge Models in Stable Diffusion Checkpoint Merger - MimicPC, otwierano: lipca 14, 2025, <https://www.mimicpc.com/learn/how-to-merge-models-in-stable-diffusion-checkpoint-merger>
27. Merging four models at an even 25% with Comfy builtins returns no ..., otwierano: lipca 14, 2025, <https://www.reddit.com/r/comfyui/comments/1bxrcxt/merging_four_models_at_an_even_25_with_comfy/>
28. Merging Models for SDXL - Stable Diffusion Art, otwierano: lipca 14, 2025, <https://stable-diffusion-art.com/forums/topic/merging-models-for-sdxl/>
29. Merging SD in XL model : r/StableDiffusion - Reddit, otwierano: lipca 14, 2025, <https://www.reddit.com/r/StableDiffusion/comments/15roz9a/merging_sd_in_xl_model/>
30. [Bug]: Size mismatch for model.diffusion\_model.output\_block · Issue #5116 · AUTOMATIC1111/stable-diffusion-webui - GitHub, otwierano: lipca 14, 2025, <https://github.com/AUTOMATIC1111/stable-diffusion-webui/issues/5116>
31. Timeline for SDXL Controlnet? · lllyasviel ControlNet · Discussion #468 - GitHub, otwierano: lipca 14, 2025, <https://github.com/lllyasviel/ControlNet/discussions/468>
32. Why is SDXL inpainting so trash ? : r/StableDiffusion - Reddit, otwierano: lipca 14, 2025, <https://www.reddit.com/r/StableDiffusion/comments/1bh53d6/why_is_sdxl_inpainting_so_trash/>
33. [Bug]: Error merging checkpoints: The size of tensor a (4) must match the size of tensor b (9) at non-singleton dimension 1 · Issue #7472 · AUTOMATIC1111/stable-diffusion-webui - GitHub, otwierano: lipca 14, 2025, <https://github.com/AUTOMATIC1111/stable-diffusion-webui/issues/7472>
34. What is the point of the endless model merges? : r/StableDiffusion - Reddit, otwierano: lipca 14, 2025, <https://www.reddit.com/r/StableDiffusion/comments/11kau9d/what_is_the_point_of_the_endless_model_merges/>
35. What's the best way to merge Pony and non-Pony models together ..., otwierano: lipca 14, 2025, <https://www.reddit.com/r/StableDiffusion/comments/1dzp7rg/whats_the_best_way_to_merge_pony_and_nonpony/>
36. Is It Possible to Merge Pony Model and Illustrious XL for Ultimate ..., otwierano: lipca 14, 2025, <https://www.reddit.com/r/StableDiffusion/comments/1ho8r08/is_it_possible_to_merge_pony_model_and/>
37. Model Merge Blocks | SeaArt ComfyUI WIKI, otwierano: lipca 14, 2025, <https://docs.seaart.ai/seaart-comfyui-wiki/core-nodes/advanced/model-merging/model-merge-blocks>
38. Model Merge Blocks - ComfyUI Wiki, otwierano: lipca 14, 2025, <https://comfyui-wiki.com/en/comfyui-nodes/advanced/model-merging/model-merge-blocks>
39. Model Merging Examples | ComfyUI\_examples - GitHub Pages, otwierano: lipca 14, 2025, <https://comfyanonymous.github.io/ComfyUI_examples/model_merging/>
40. ComfyUI-DareMerge detailed guide - RunComfy, otwierano: lipca 14, 2025, <https://www.runcomfy.com/comfyui-nodes/ComfyUI-DareMerge>
41. 54rt1n/ComfyUI-DareMerge: ComfyUI powertools for SD1.5 ... - GitHub, otwierano: lipca 14, 2025, <https://github.com/54rt1n/ComfyUI-DareMerge>