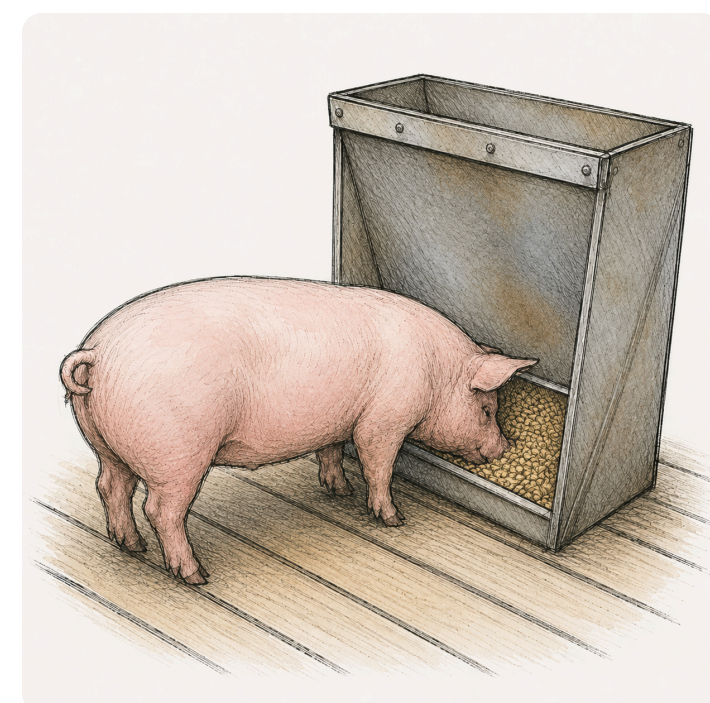


MOTIVATION Why look under the skin?

Feed is the dominant cost in pork production, and how much a sow eats should track her body condition — the reserves of backfat and loin muscle that drive lactation performance and piglet survival. Yet the on-farm measures used to manage body condition — visual scoring and calipers — correlate poorly with the fat and muscle actually under the skin, so producers can't reliably measure what matters at the scale of a working barn.



TODAY Manual ultrasound

Ultrasound is the reference for backfat & loin at the **last rib** — but it needs a skilled operator and animal handling, so it **can't scale**.

GOAL One overhead camera

PigFormer reads the same three traits from raw depth alone — non-contact, automatic, and **continuous** at herd-scale.

2.43mm Backfat MAE — within 2× the noise floor of repeated manual ultrasound (1.30 mm)

-22%/-39% Lower overall MAE than the ResNet-18 (-22%) and ViT-small (-39%) single-stage depth baselines

≈7ms End-to-end per frame (UNet Stage 1) — **real-time** on a single A100

319 annotated instances

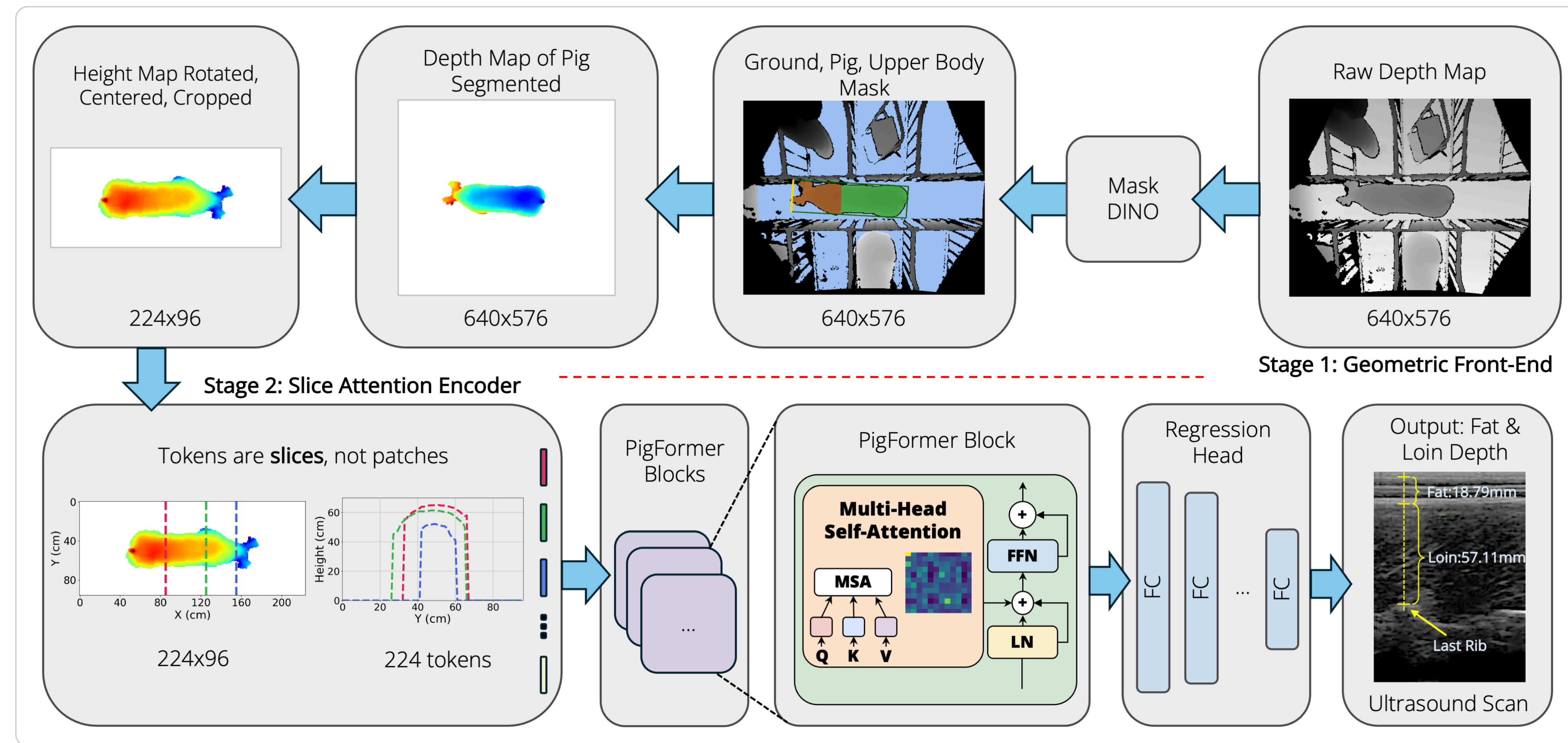
6,705 depth frames

2 commercial sites

WHY IT MATTERS Contributions

- First to cast RGB-D sow body-condition as a fully automated, end-to-end **multi-target regression** of backfat, loin & total tissue.
- PigFormer: a **geometric front-end** (SAM3→MaskDINO distillation, ground-plane removal, orientation normalization) feeding a **Slice Attention Encoder** (RoPE + dual pooling) that regresses all three jointly.
- 3.87 mm overall MAE** on 319 instances across two farms — **22% & 39%** below ResNet-18 and ViT-small.

METHOD PigFormer pipeline



RESULTS Held-out test — 79 unseen sows & gilts

METHOD (INPUT)	FAT ↓	LOIN ↓	TOTAL ↓	OVERALL ↓
ViT-small — raw depth	4.98	3.57	7.29	6.34
ResNet-18 — raw depth	2.88	2.88	6.10	4.93
PigFormer — MaskDINO	2.43	5.01	4.19	3.87
PigFormer — pruned MaskDINO	2.34	5.27	4.20	3.94
PigFormer — UNet (16× faster)	2.40	5.20	4.26	3.95
Human ultrasound std (ref)	1.30	2.02	2.29	1.87

Mean absolute error in **mm**, lower is better. All three PigFormer Stage 1 variants land within **0.08 mm** overall — the geometric inductive bias, not the segmenter's size, drives accuracy. Loin gains come from Stage 1's heading alignment placing the rib site at a canonical location.



Scan for code
github.com/iambashar/Pigformer