Suture Training System for Students: Knot Classification Model

Team Members: Lucy Hockerman, Presley Hansen, Maddie Michels, Kate Hiller, Sadie Rowe
Clients: Dr. Margene Anderson, Dr. Sara Colopy, Dr. Paul Merkatoris
Advisor: Dr. Walter Block
BME 400, Fall 2025

College of Engineering
University of Wisconsin–Madison

Problem Statement

- Novice veterinary students struggle to judge proper suture tension, often tying knots too loose or too tight.
- Knot security depends on achieving the correct tension in the final throw, yet no established tension values exist, as required force varies with suture material and thickness.
- A data-driven feedback system is needed to help students learn proper knot tension, improving training efficiency and surgical outcomes.

Background & Impact

Need: System to train the correct tension needed for a secure knot.

- No system on the market
- Students currently learn by "feel" with instructor feedback after the knot is tied.

Too Loose: suture will unravel, healing complications.

Too Tight: risk of tissue damage, material failure, and wound dehiscence.

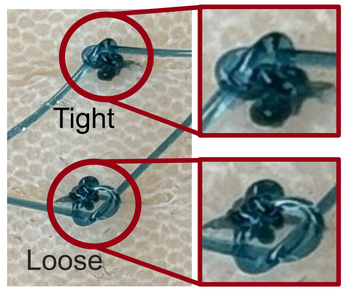

Figure 1: Comparison of loose and tight knots.

Impact

- Each stitch costs $1.75 - $1.83, contributing to high material use and waste. [1]
- Future expansion of use application.

Figure 2: Suture material failure from over tightening [2].

Design Specifications

Model Shall...

- Classify a variety of suture sizes, colors, and materials
- Classify knots with an overall **accuracy ≥ 80%** [4]
- Achieve **≥ 80% precision** for "tight" class to minimize false positives [4]
- Attain **≥ 80% recall** for "loose" class to minimize false negatives [4]
- Have a **F1 score of ≥ 0.8** [4]

System Shall...

- Withstand repeated use in training
- Be minimally disruptive to the suture process

Concept Prototyping

Initial Idea:

Greater finger pressure/force → greater suture tension → tighter knot

Limitations:

- Force not directly proportional to tension [5]
  - Direction of force does not match direction of suture tension
- Does not detect plastic deformation
- High variability with suture-tying techniques

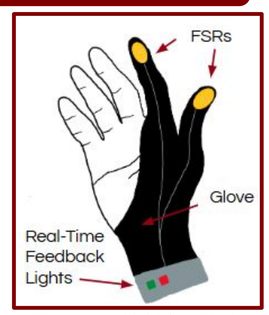

Figure 3: Initial Glove Design Measuring Force.

ML Model Design & Testing

	TensorFlow Model				RoboFlow Model
	Side and Top		Side Only		Side and Top
Model Number	1	2	3	4	5
Model Type	ResNet	VGG	ResNet	VGG	ResNet
Layers	50	16	50	16	50
Dataset Size	424	424	309	309	424
Tight : Loose	1.24	1.24	1.22	1.22	1.24

Table 1: Model structures and dataset sizes, types, proportionalities.

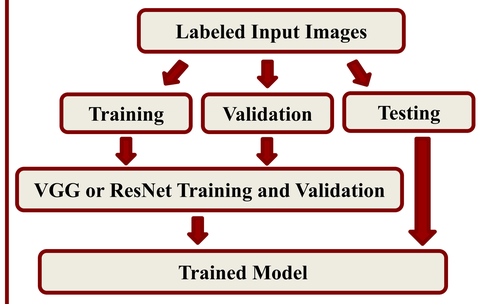

Figure 5: Model training pathway.

Figure 4: Top (left) and side (right) knot examples.

Testing

- Tested on 40 tight and 40 loose new images
  - Metrics collected: F1 score, accuracy, recall, and precision
  - Confusion matrices

Results

			Model 1: ResNET (Side & Top)		Model 2: VGG (Side & Top)		Model 3: ResNET (Side Only)		Model 4: VGG (Side Only)		Model 5: Roboflow (Side & Top)	
	Actually Tight	Actually Loose										
Predicted Tight	True Positive (TP)	False Positive (FP)	34	**4**	36	11	36	15	40	20	30	**4**
Predicted Loose	False Negative (FN)	True Negative (TN)	6	36	4	29	4	25	0	20	10	36

Table 2: Confusion matrices for each model. Here, a true positive is defined as a tight knot correctly predicted as tight.

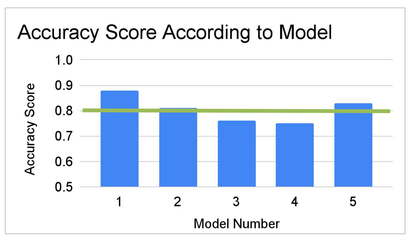

Figure 6: Accuracy of each model, representing the proportion of all predictions that were correct.

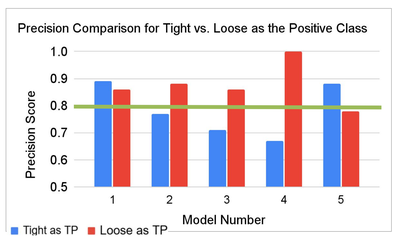

Figure 7: Precision score of each model, defined as the proportion of predicted positives that were correct.

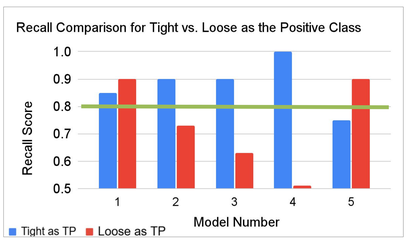

Figure 8: Recall score of each model, representing the proportion of actual positives the model correctly identified.

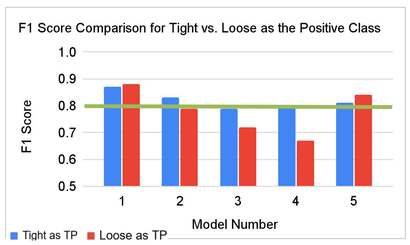

Figure 9: F1 score of each model, combining precision and recall into a single performance measure.

Discussion

- Based on results, continue to train and improve the **TensorFlow ResNET: Side and Top** and **RoboFlow** models

Successes	Improvements
• Overall accuracy ≥ 80% • F1-score ≥ 0.8 • Achieve ≥ 80% precision for tight as true positive	• Achieve accuracy and precision above 90% • Minimize model differences for identifying tight vs. loose knots

- **Sources of Error:**
  - Suture type variability: diameter, color, sample size imbalances
  - Inconsistent photo quality: lighting, background color, zoom
  - Model inability to locate top knot
- **Limitations:**
  - RoboFlow has limited ability to adjust parameters compared to TensorFlow
  - Limited computing power to train models with large image sets

Future Work

1.) **Model Improvement**
- Retrain using a larger, more consistent image set of suture types
- Meet with expert to fine-tune model

2.) **Incorporate Model in a User-Friendly System**
- Design and build camera stand
  - Improve lighting and camera angle consistency
  - Incorporate a digital camera to improve image quality
- Investigate real-time feedback systems
  - App, website, or microcontroller
  - Tight vs. loose knot visual display

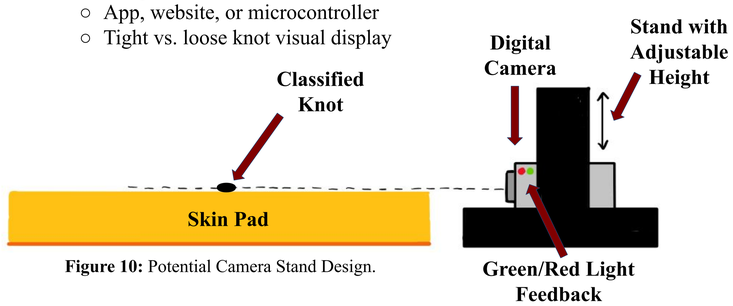

Figure 10: Potential Camera Stand Design.

Acknowledgements

- Dr. Wally Block
- Dr. Margene Anderson
- Dr. Sara Colopy
- Dr. Paul Merkatoris
- Dr. Amit Nimunkar
- BME Department
- UW-Madison School of Veterinary Medicine

References

[1] R. K. Elmallah et al., "Economic evaluation of different suture closure methods: barbed versus traditional interrupted sutures," Ann. Transl. Med., vol. 5, no. Suppl 3, p. S26, Dec. 2017, doi: 10.21037/atm.2017.08.21.
[2] "The ideal surgical knot (according to science)," *Healthcare-in-europe.com*, Jun. 12, 2023. http://healthcare-in-europe.com/en/news/ideal-surgical-knot.html (accessed Oct. 01, 2025).
[3]"Is 80% accuracy good in machine learning?," *Milvus.io*, 2025. https://milvus.io/ai-quick-reference/is-80-accuracy-good-in-machine-learning (accessed Dec. 01, 2025).
[4] "F1 Score in Machine Learning." Accessed: Oct. 07, 2025. [Online]. Available: https://encord.com/blog/f1-score-in-machine-learning/
[5] K.-T. von Trotha *et al.*, "Surgical sutures: coincidence or experience?," *Hernia*, vol. 21, no. 4, pp. 505–508, Aug. 2017, doi: 10.1007/s10029-017-1597-8.