ScrapAR: Democratizing Woodworking Education Through Contemporary Computational Advancements

Isaac Martinotti

Carnegie Mellon University

School of Architecture

Committee Members:

Juney Lee
Daniel Cardoso Llach
Vernelle Noel

Primary Advisor: Juney Lee

Pittsburgh, PA
May 7, 2025

CARNEGIE MELLON UNIVERSITY

School of Architecture

College of Fine Arts

Thesis

Submitted in Partial Fulfillment of the requirements for the degree of

[Master of Science in Computational Design]

TITLE:

ScrapAR: Democratizing Woodworking Education Through Contemporary Computational Advancements

AUTHOR:

Isaac Martinotti

ACCEPTED BY ADVISORY COMMITTEE:

Juney Lee	Principal Advisor	May 7th, 2025 DATE
Daniel Cardoso Llach	Secondary Advisor	May 6 2025 DATE
Vernelle Noel	Secondary Advisor	May 7th, 2025 DATE

Abstract

This thesis introduces a computational workflow that integrates computer vision (CV) and augmented reality (AR) to democratize woodworking education for beginners. Grounded in ethnographic interviews with professionals, iterative prototyping, and systematic testing, the proposed approach enables users to catalog and repurpose scrap materials as volumetric "ingredients" within AR assisted recipes. Leveraging machine-learning-based object detection and accessible AR game engines, the system delivers interactive user-interface overlays – including tool tutorials, nested part visualizations, and step-by-step assembly guidance – for constructing a pre-designed chair. By automating tasks such as scrap management, spatial planning, and measurement, the workflow reduces barriers to entry to woodworking as a form of craft and fosters creative reuse of existing materials. To evaluate its effectiveness, users – ranging from a novice to an experienced woodworker – completed an illustrative project. Their case studies demonstrate improvements in user confidence, material efficiency, and exploratory learning. The findings suggest that integrating CV and AR can enhance craft-education accessibility and support innovative, sustainable practices in alternative learning environments.

Keywords: Augmented Reality, Computer Vision, Woodworking Education, Material Reuse, Circular Economies, Sustainable Design, Craft

Contents

1 Introduction **3**
- 1.1 Inspiration . . . 4
- 1.2 Springboard Projects . . . 6

2 State of the Art **11**
- 2.1 Woodworking & Craft Pedagogy . . . 11
- 2.2 Augmented Reality and Computer Vision in Learning and Craft . . . 11
- 2.3 Existing Tools and Digital Fabrication Approaches . . . 12
- 2.4 Sustainability and Circular Economies . . . 13
- 2.5 Machine Learning in Design . . . 13

3 Thesis Overview **15**
- 3.1 Research Gap . . . 15
- 3.2 Objectives . . . 16
- 3.3 Scope . . . 17

4 Field Investigation **19**
- 4.1 Ethnographic Interviews of Woodworking Educators . . . 19
- 4.2 Augmented Reality Environment Testing . . . 22
- 4.3 Monocular Object Detection . . . 23

5 Technical Implementation **26**
- 5.1 Overview . . . 26
- 5.2 Kit of Parts . . . 26
- 5.3 On-boarding . . . 28
- 5.4 Inventory Creation . . . 31
- 5.5 Main User Journey . . . 42
- 5.6 Home Screen . . . 42
- 5.7 Making a Recipe . . . 45

6 Results **49**
- 6.1 Study Parameters . . . 49
- 6.2 Study Outcome . . . 49
- 6.3 Method of Evaluation . . . 52

7 Conclusion **55**
- 7.1 Interpretation of Findings . . . 55
- 7.2 Contributions . . . 57
- 7.3 Limitations . . . 58
- 7.4 Future Research Directions . . . 59

1 Introduction

Woodworking occupies a curious place in contemporary craft culture. Popular media present it as a hands-on antidote to screen-based labor – slow, tactile, and richly rewarding – yet in practice the craft is hard to enter. Effective instruction remains tied to physical workshops, beginner-level tutorials often presume access to expensive or dangerous machinery, and the tacit, experience-driven knowledge can often only be gained through apprenticeships that exclude many would-be makers. At the same time, considerable volumes of perfectly serviceable lumber, whether off-cuts from construction, discarded furniture panels, or short lengths of dimensional stock, continue to travel from cradle to landfill without significant opportunities for reuse.

This thesis argues that recent advances in commercial computer vision (CV) pipelines and gamified augmented reality (AR) frameworks can combine to ease two long-standing bottlenecks at once: first, the difficulty novices face when they try to measure, plan, and cut unfamiliar stock with confidence; and second, the scarcity of situated, step-wise feedback that can traditionally be found only in a classroom. Over the past year, interviews have been conducted with shop managers, digital-fabrication instructors, and self-taught hobbyists to understand where current practice breaks down. Through these conversations and scholarly exploration, a proof-of-concept pipeline containing three technical explorations – henceforth referred to as technical vignettes – is developed to take an unfamiliar user through a simple project using easily gatherable scrap materials and a commonly available single-view camera – such as a tablet or webcam. This exercise not only incorporates the capabilities of contemporary machine learning (ML), but also serves as a stepping stone toward creating a workflow that could be extended in a future holistic package.

Complementing these technical vignettes is the development of an AR overlay in Unity [1] – an application that loosely guides users through the process of assembling a design using a step-based sequence of 3D models and annotations. By overlaying digital templates onto physical objects, the system seeks to alter the typical experience of furniture construction away from an underwhelming Ikea chore into a playful, accessible activity. In doing so, it intensifies the ownership of the construction process by investing the users time and energy, and tackles two central challenges: lowering the barriers to entry for novice makers and creating a new use-case for material waste that can abundantly be found on street corners and free online forums [2] [3].

Seeking to enhance accessibility and promote sustainability in craft, this thesis introduces ScrapAR, a prototype workflow using AR and CV to bridge digital advancements with traditional woodworking education. It empowers individuals lacking formal training or ample time to engage with craft while simultaneously finding value in discarded materials, directly addressing the research question: How can AR and CV provide craft education that reduces waste and enhances woodcraft accessibility? This research explores the feasibility of integrating object perception, AR instruction, and waste-conscious nesting using widely accessible hardware.

The subsequent chapters detail the technical implementation and evaluate the system through a user study, ultimately arguing that computational design need not replace hand skills but can effectively scaffold them. By demonstrating that object recognition, gamified guidance, and efficient material use can coexist on current devices, ScrapAR positions these technologies as tools to broaden woodworking's reach and foster more sustainable making practices.

Here Section 2 surveys adjacent work in AR craft learning, generative design for circular economies, and monocular six-degree-of-freedom pose estimation – the ability for a single-camera to track the position and location of an object in 3D space; Section 3.1 distills those strands into a precise problem statement and clarifies the original contribution claimed here; Section 4 walks through the technical implementation, from Unity scenes to a lightweight depth-proxy pipeline for single-view scaling; Section 6 presents an exploratory user study in which a user of divergent skill turns campus scrap into an Enzo Mari–inspired chair; and Section 7 reflects on present limitations and sketches a path toward a community-driven recipe library that could extend the idea beyond woodworking into other material streams.

1.1 Inspiration

The inspiration for this project was found during a previous opportunity as a research assistant for Daniel Cardoso Llach titled "A Multi-scalar and Multi-modal Approach to Architectural Heritage Documentation: An Interactive Digital Representation of the St. Nicholas Chapel" [4]. Parts of this Deep Time Architectural Data (DTAD) [5] project are included in the subsequent section under Nucleation Projects to show their direct connection and ideological parallels. Additionally, during the previous summer and in the final months of 2024, small proof-of-concept explorations were implemented and included in the Methodology section Inventory Creation. This springboard provides a better understanding of the limits of various recently-published CV implementations, from both contemporary computer science literature as well as pertinent historical literature surrounding this field – it was during this exploration that the original scope of work was set (see Figure 1).

In addition to engaging with emerging literature, two CMU classes were chosen to further the practical skill-building of these techniques alongside meetings with specific faculty to ensure feasibility. The *notably difficult* Introduction to Deep Learning with Professor Bhiksha Raj and the Creative Technical Sprints (CTS) course with Joshua Horowitz each allowed an exploration of this potential inquiry, coming from two distinct directions. First, the ability to go through the process of designing and iteratively testing a neural network, and particularly a convolutional neural network, framework should be implemented somewhere in this project, if not for the pure ability to say that the author has the capability to do so. Secondly, the ability to quickly and iteratively work on simple augmented and computational projects directly reinforces the feasibility of this exploration, as well as the possibility of requesting consultation and assistance from various connections gathered through Horowitz and CTS.

In recent years, the rapid proliferation of mobile AR and the continually improving capabilities of everyday camera sensors have redefined how we interact with our physical surroundings [6], and the fidelity of information that we can overlay atop it. Technologies once confined to professional / industrial applications or specialized educational settings are now beginning to transform traditional handcraft such as woodworking and carpentry [7][8], along with more mature examples in creative fields such as ceramics [9] and the culinary arts [10]. This push is occurring in conjunction with an unprecedented excitement for generative design methodologies in industrial and architectural fields, as is explored in companion guides such as Sculpting Spaces of Possibility [11].

Historically, woodworking has relied on tacit knowledge transmitted through hierarchical apprenticeships – a system rich in tradition but limited in scope and accessibility, particularly for those outside established networks or shared demographics [12], such as those in Figure 2. This project addresses this gap by proposing a gamified AR approach that combines CV, intuitive user interfaces,

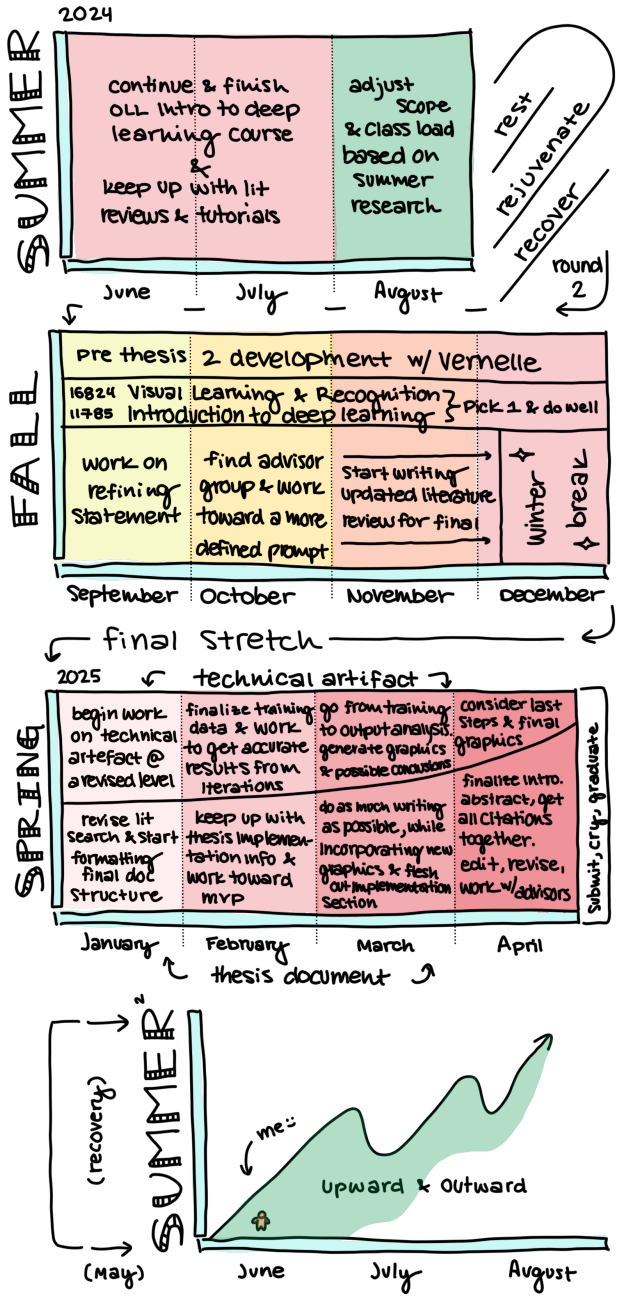


Figure 1: Overall timeline of the thesis process including required / optional classes and a high-level sectional overview [Credit to the author].

and digital overlays to democratize woodworking as a craft practice, making it accessible to a broader, more diverse audience. At its core, the project explores how digital innovations can breathe new life into traditional practices by challenging conventional learning models. Drawing on both literary and ethnographic investigation – and informed by the author's own personal journey as a self-taught internet-enabled woodworker – an explicit effort is put into understanding the emerging precedent for

learning woodworking skills through shared digital libraries of tutorials and non-trade instructional content.



Figure 2: The Carpenter's Apprentice, an 1841 woodblock print illustrating traditional male-to-male apprenticeship structures [13].

1.2 Springboard Projects

The methods explored in this thesis were significantly inspired and directly informed by three preceding projects. These works, each featuring key contributions from the author, encompassed two computational experiments and an earlier independent study, collectively providing the specific technological experience and foundational insights for the current exploration.

A significant early experience was the project titled "A Multi-scalar and Multi-modal Approach to Architectural Heritage Documentation" [4]. This group effort, involving four research assistants and two primary investigators, focused on documenting the St. Nicholas Chapel in Beaver, Pennsylvania. Through a multi-month exploration, diverse datasets were captured and created – including photogrammetry, LiDAR (Light Detection and Ranging) scans, deep learning form analyses, and ethnographic interviews – with the goal of creating an interactive, web-based cultural heritage experience. The development of a prototype interactive document (an example of which is shown in Figure 3) was particularly instructive. It showcased how to effectively combine these varied datasets into a cohesive platform, enabling broader audiences to virtually explore the chapel's historical context, architectural features, and sociocultural significance without direct interaction with the physical site.

This process not only provided the author with familiarity in several contemporary web-based rendering strategies, but also underscored the value of integrating an ethnographic layer – first-person accounts from local practitioners and community members highlighted the often unseen social and cultural context underpinning the chapel's construction, bridging purely geometric data with human narratives and experiences. The project blended large-scale "distant" analysis of Carpathian-style



Figure 3: Screenshot from the St. Nicholas Chapel prototype document, showcasing a Gaussian Splatting based visualization and overlaid interactive information [Credit to the thesis author].

wooden churches (see Figure 4, Left), based on an earlier thesis [14], with a "close reading" of the chapel's geometry, materials, and personal narratives (see Figure 4, Right). This approach, heavily inspired by the prior work of P.I. Daniel Cardoso Llach [15], was invaluable.

Although centered on preserving a physical space, this process underscored the value of ethnographic interviews and field study – an approach that directly influenced the desire to use on-campus experts as a sounding board in the current thesis. Furthermore, the exploration of an interactive platform demonstrated the utility of accessible 3D content for engaging larger audiences and offered valuable lessons in interface design, the non-linearity of scholarly materials, and building user-centered explorations, all highly relevant to the current project.

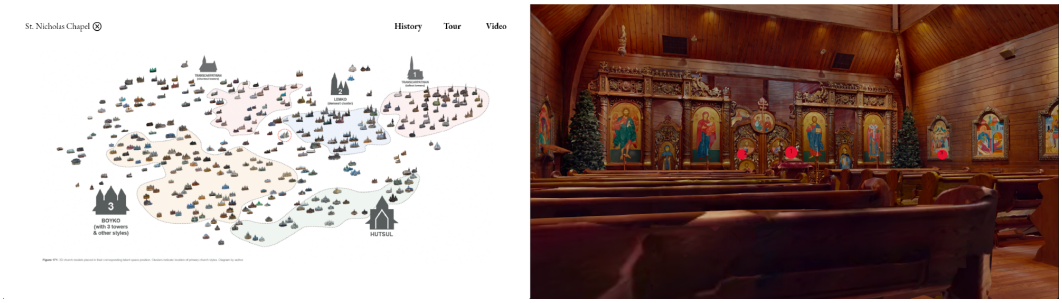


Figure 4: **Left:** Prototype interactive map of historic Carpathian churches, clustered using a ML algorithm [Credit to M. Hasey & I. Martinotti]. **Right:** User-controlled immersive tour of the Beaver, PA. chapel made from photogrammetry [Credit to the thesis author].

The experience gained in developing interactive and user-centered digital experiences was further honed in a fast-paced industry collaboration with Netflix's Marketing Innovation Manager, Annatruus Bakker. This project, "Netflix Presents Hansel & Gretel – A Walk in the Park" [16], originated within Joshua Horowitz's Creative Technology Sprints course. Over a three-week sprint, our team developed an AR concept prototype for a proposed *Hansel & Gretel* Netflix Original Film tie-in,

responding to a brief that emphasized holistic experiential marketing and narrative immersion. Using Unity and ARFoundation, we delivered an intuitive and highly generalizable alternate reality game (ARG) experience. The early prototype (see Figure 5, Left for an interaction example and Right for the user journey) featured an AR "scavenger hunt," allowing users to collect 3D objects and unlock exclusive content. This exercise demonstrated the feasibility of rapidly prototyping location-based AR experiences on consumer devices.

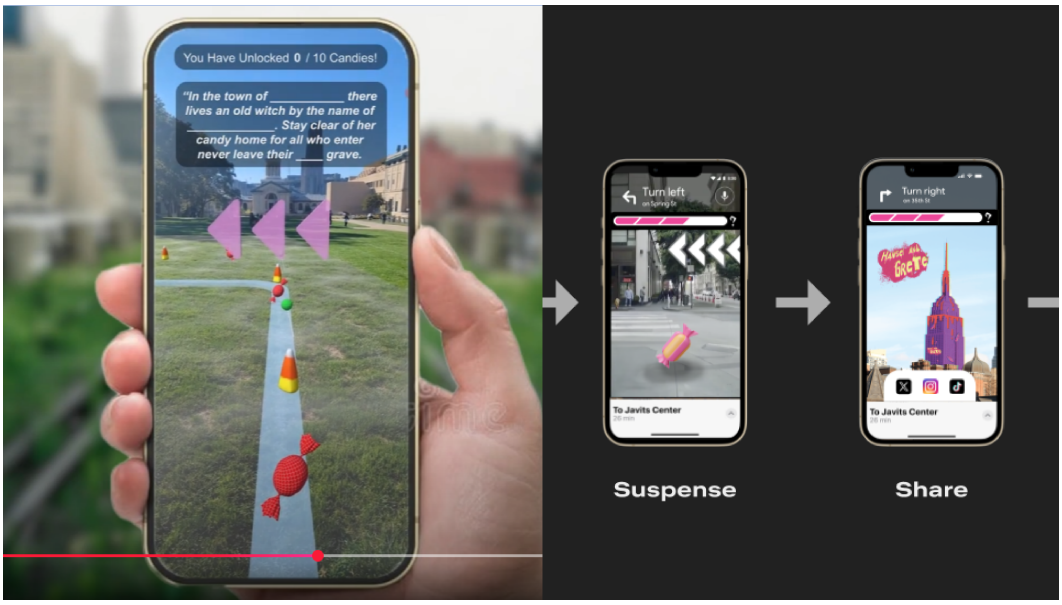


Figure 5: **Left**) Pre-recorded demo screenshot in Unity AR showcasing "candy pickup" interactions [Credit to the thesis author]. View the full video here. **Right**) User journey diagram showing the AR path, pickup points, and shareable reward loops [Credit to I. Martinotti & S. Chen].

Key insights included the power of a gamified user journey with reward loops to encourage social sharing and repeat engagement – mirroring contemporary ARG design and principles from educational applications like Duolingo [17]. The focus on user-centric design, ensuring accessibility and intuitive mechanics for AR novices, along with the use of AR overlays and on-screen guidance, provided practical lessons. While inherently geared toward entertainment, this sprint underscored the value of quick prototyping and real-world testing, and how game mechanics can create engaging experiences that lower barriers to participation – a crucial consideration for addressing potential user disengagement in learning services, a concern highlighted in preliminary presentations of the current thesis.

Complementing these explorations in digital heritage and AR entertainment, a more personal experiment, "ARRecipes: Prototype Pipeline" [18], offered direct preliminary work in CV-based dimensioning and automated nesting for material reuse. Originally conceived to reduce wood waste, this rough prototype leveraged the Sony digital time-of-flight (d-ToF) sensor in an Apple iPad Pro to capture 3D geometry of scrap lumber (an inventory of scans is shown in Figure 6). By placing a known-scale marker in each scene, a user could calibrate the point-cloud scans for relatively accurate dimensioning. The project then employed Rhino/Grasshopper and GHPython scripts, including open-source tools like OpenNest [19], to fit predefined furniture parts – inspired by Enzo Mari pieces [20] – onto the inventoried geometries (see Figure 7 for example instructions and nesting).

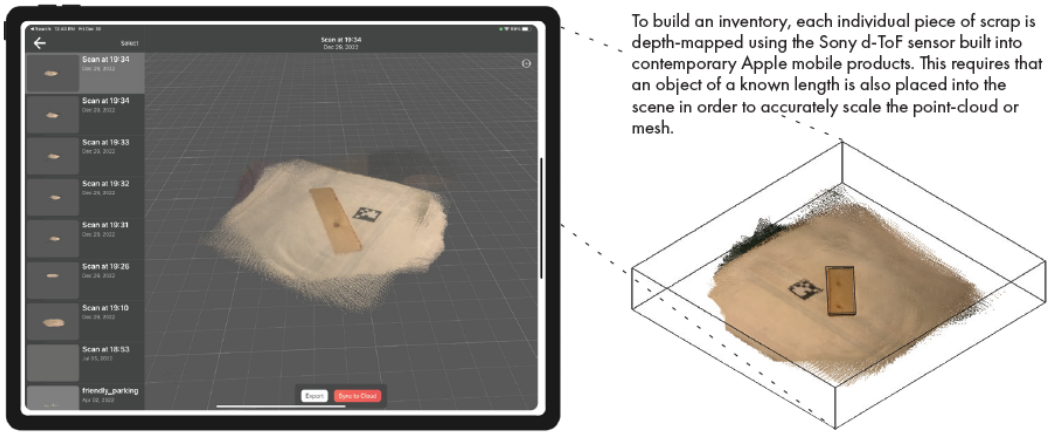


Figure 6: Inventory of scans utilized to scale geometry bounding boxes.

Though a proof-of-concept, ARRecipes illustrated how a low-overhead workflow, combining consumer AR scanning with visual coding, could support a "recipe" approach in woodworking. It reinforced the potential to promote interest in craft by streamlining steps like dimension capture, database entry, and part layout – precisely the sorts of automated, user-centric features central to the current thesis's goal for the ScrapAR system.

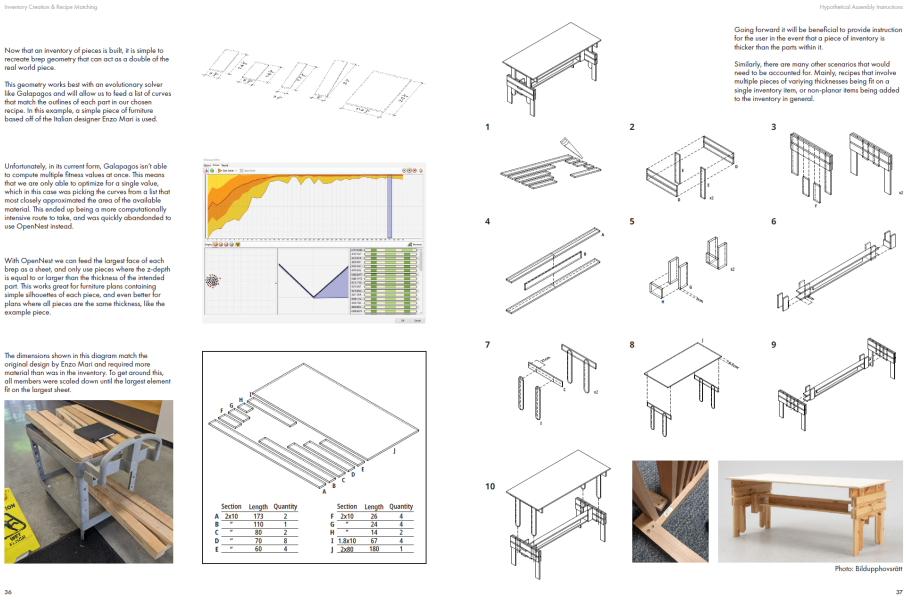


Figure 7: Example project instructions & nesting using parts from an Enzo Mari table [Layout credit to the author].

In essence, these three projects collectively provided the essential experiential bedrock for this thesis. The St. Nicholas project established a methodology integrating ethnographic insights with multi-modal digital representation, while the Hansel & Gretel sprint honed skills in rapid AR prototyping and gamified user engagement within Unity. Finally, the ARRecipes experiment offered direct preliminary work in CV-based dimensioning and automated nesting for material reuse. This combined foundation in AR/CV implementation, user interface considerations, and iterative design

directly informed the technical approach and research direction pursued in the development and evaluation of ScrapAR detailed in the subsequent chapters.

2 State of the Art

Recent significant advances in AR and photogrammetry have improved the accuracy, accessibility, and intuitiveness of computationally driven physical workflows across many industries. In tandem, developments in CV and ML are opening new avenues for understanding and interacting with the physical world. This section reviews key areas of research and technological development relevant and connected to the goals of ScrapAR, including woodworking education, the application of AR/CV in craft and learning, existing digital tools for fabrication, and the broader context of sustainable design and circular economies.

2.1 Woodworking & Craft Pedagogy

Traditionally, woodworking knowledge has been transmitted through apprenticeships and hands-on workshop instruction, a system rich in inherent knowledge but often limited in accessibility [12]. This model, historically male-dominated, can present barriers to entry for many potential learners due to time, cost, or access to facilities. The literature on craft pedagogy highlights the importance of material understanding and experiential learning. For instance, alumnus such as Chi's work emphasizes cultivating material knowledge through physical / digital interfaces, aiming to bridge tactile craft with digital tools at all scales [21]. Tsai's exploration of "Making Mistakes" forces us to re-frame error as an emergent and desired property in craft making, suggesting that embracing imperfections can be a vital part of the learning and design process, particularly when integrating digital tools [22] – a computerized kintsugi.

The challenges in traditional instruction, such as limited one-on-one time and the difficulty of conveying complex spatial concepts, are areas where new technologies might offer significant benefits. Ethnographic interviews conducted for this thesis with woodworking educators like Jon Holmes and Terry Hritz (Sections 4.1 and 4.1) further underscored these challenges, highlighting issues of time constraints for comprehensive instruction, student retention of safety and techniques, and the hurdles novices face with both hand tools and complex digital fabrication workflows. In conjunction, the rise of the "maker movement" and increased interest in Do-It-Yourself (DIY) practices also signal a shift in how craft skills are acquired and valued [23]. This movement often leverages online resources and community-based learning, creating a demand for more accessible and democratized forms of instruction. Shorey's research on the "Handmade Future" investigates innovation through making and craft, analyzing how public discourse and computational access shapes perceptions of these practices [24].

2.2 Augmented Reality and Computer Vision in Learning and Craft

Augmented Reality has shown considerable potential in educational contexts by making abstract concepts more tangible and enhancing engagement. Ibrahim et al. demonstrated AR's utility in vocabulary learning by providing visual and interactive representations [25], a concept that can likely be extended to visualizing 3D designs and fabrication processes. Research by Chi et al. surveyed trends in AR applications for architecture, engineering, and construction (AEC) industries, highlighting its potential for on-site information overlay and training [26]. Specifically in the context of craft and vocational training, AR also offers promising avenues. Lee et al. explored the use of virtual reality (VR) for learning woodworking in batch furniture production, indicating that immersive technologies can enhance students' understanding of manufacturing processes [7]. More recently,

Lee and Lin further investigated integrating AR and VR to bridge the gap between a highly-traditional form of classroom woodworking education and more atypical expectations of industrial furniture production [27].

CV techniques are increasingly sophisticated in their ability to interpret and interact with the physical world. The development of methods for 3D reconstruction from single or multiple images, such as Manni et al.'s 'Snap2CAD' for indoor environments [28], is crucial for digitizing physical objects and spaces for AR applications. In the domain of woodworking and material understanding, CV can be applied to tasks like wood species identification – Yang et al. on wood classification using deep learning [29], and Verly Lopes et al. on North American hardwoods [30] – or attempting to capture the structural integrity of in-situ objects [31]. A recent work by Yang et al. even explores sensing tacit knowledge and the quality of specific tool-actions in woodworking through a multi-layered sensing system and a traditional hand plane [32].

The intersection of AR and CV for direct guidance in fabrication is an incredibly active area of research. Burkhardsmeier's thesis on "Augmented Woodworking" explored line-of-sight AR for tasks like part layout and table saw setup [33], in addition to hypothetical interactions that the user could employ to control a head-mounted AR system. Settimi et al. have been actively developing "TSLAM," a tag-based object-centered monocular navigation system for augmented manual woodworking, aiming to provide robust camera pose for overlaying digital twins onto physical workpieces during fabrication [34]. Their related work on "Augmented Carpentry" further explores these concepts [35], including an ML-based visual pose detector named "TTool" [34]. Dreesbach et al. evaluated AR-based learning in vocational woodworking training, comparing it with traditional teacher-instructed units and finding comparable learning outcomes [36]. Similarly, Beltrán Alvarado studied the implementation of visual AR in a carpentry course, noting its effectiveness, particularly for lower-performing learners [37]. Cousin et al. explore mixed-reality construction for nonstandard "wild wood" gridshells, showcasing AR's role in complex assembly tasks as well as in the pursuit to better utilize undesirable materials [38].

Monocular 3D object detection and pose estimation are critical for CV systems that operate without specialized depth sensors. Research in this area, often driven by autonomous driving – such as Chen et al. [39] – is increasingly relevant for accessible AR applications. Works like Barabanau et al. on geometric reasoning on keypoints [40] and Tang et al. on Center3D for joint depth understanding [41] explore methods to derive 3D information from single images while simultaneously inferring higher dimensions of object characteristics. The development of foundation models like FoundationPose – by Wen et al. [42] – for 6D pose estimation of novel objects and techniques like OnePose++ – by He et al. [43] – for keypoint-free one-shot object pose estimation are pushing the boundaries of what's possible with CV for object tracking in AR. Recent advancements such as SplatPose (Yang et al., 2025) [44] leverage even more contemporary 3D gaussian splatting for high-efficiency pose estimation from a single RGB image, further indicating the rapid progress in this field.

2.3 Existing Tools and Digital Fabrication Approaches

The landscape of digital tools for design and fabrication is evolving rapidly. While traditional CAD software provides powerful modeling capabilities, the integration of fabrication awareness and material intelligence is an ongoing area of development. The "Carpentry Compiler" [45] aims to optimize cutting plans for dimensional lumber – with a high focus on ease of use – showcasing computational approaches to both material and labor efficiency. Open-source initiatives and platforms

for sharing designs are also influencing how people approach making. Projects like AtFAB offer open-source furniture designs as digital files that can be locally fabricated [46], democratizing access to designs and manufacturing. This aligns with the broader trend of using pattern languages [47] and shape grammars [48] to define design systems that can be adapted and utilized by a wider audience.

In academic research, computational design methods are being explored to address specific fabrication challenges. Reisach et al. present a design-to-fabrication workflow for free-form timber structures using offcuts, highlighting methods for handling non-standard geometries [49] and building upon the concept of upcycling timber [50]. Capunaman's work explores CAM as a tool for creative expression by informing digital fabrication through human interaction, proposing a feedback loop where users actively manipulate artifacts during the fabrication process [51]. Sunshine's research on "Inventory: CAD for Medium Resolution Materials" (2023) [52] investigates how CAD tools can better account for the realities of material sourcing and variability, moving beyond idealized digital representations. The development of specific AR/VR tools for woodworking training is also underway. Besides the work of Lee and Lin mentioned earlier, The National Taipei University of Technology has explored AR for spatial skills learning in carpentry [53]. Yang and Chen (2024) applied paper prototyping to design an AR teaching tool for novice woodworkers focusing on component arrangement and material cutting [54], further highlighting the iterative design process for such educational tools.

2.4 Sustainability and Circular Economies

The imperative for sustainable design practices is a significant driver for this research. Traditional construction and manufacturing processes often generate substantial material waste [55]. This thesis aligns with contemporary efforts to promote creative reuse and environmental responsibility [56]. The concept of circular economies, where materials are kept in use for as long as possible, is gaining traction in design and manufacturing. Poteschkin et al. discuss the recycling of cross-Laminated Timber (CLT) production waste, an example of industry-specific efforts towards circularity [57]. The Buy Nothing Project and similar digital gift economies [2, 3] also reflect a societal shift towards valuing reuse and community-based resource sharing, creating a receptive cultural context for tools that facilitate the repurposing of materials.

Digital tools can play a crucial role in enabling material reuse. Yu et al. propose a framework for sustainable building design and construction using off-cut wood, involving 3D scanning for material reclamation and generative systems for structural assessment [58]. The ability to digitize irregular scrap materials, optimize their use through nesting algorithms (such as OpenNest [19] or meta-heuristic solvers [59]), and visualize their transformation into functional objects using AR is a key aspect of ScrapAR.

2.5 Machine Learning in Design

Machine learning is increasingly applied to analyze and generate design iterations. Architectural style classification using CNNs, as explored by Yoshimura et al. [60] and Han and Yin [61], demonstrates ML's ability to recognize complex visual patterns. Kim and Lee applied deep learning for the stochastic detection of interior design styles [62]. While this thesis does not focus on design generation in the same vein as Rhee and Kim's work on high-rise building forms [63] or design space interpretation [64, 65], the underlying principles of pattern recognition and data-driven analysis are relevant. For ScrapAR, ML is leveraged for object detection as well as for providing feedback – with

the ability of ML to "see" and interpret visual information being discussed by Witt in "Through a Machine Lens" [66], underpinning many of the CV functionalities in this project.

This State of the Art review showcases a convergence of cross-disciplinary advancements in AR/CV, digital fabrication, craft pedagogy, and sustainable design. While each area has seen significant progress, the specific application of these technologies in an integrated workflow to democratize novice woodworking education using scrap materials addresses a distinct and valuable niche.

3 Thesis Overview

3.1 Research Gap

Despite these advancements, there are notable gaps in the literature concerning the application of AR, 3D perception, and computational pipelines specifically in woodworking and furniture production (see Figure 8). While AR has been employed in educational contexts and architectural visualization, its potential to facilitate hands-on design and fabrication processes in carpentry has only recently been explored in limited ways. Most existing studies focus on professional applications like assembly lines, or educational tools for trade schools rather than practical situations for hobbyists and novices trying to produce something at home [67, 37]. Moreover, the integration of material waste utilization with AR technologies presents an area in need of extensive research. Existing studies have addressed sustainable design and waste reduction separately from user-centric technological applications. The potential for a distributed intervention that focuses on repurposing locally available waste within the community, is overshadowed by less efficient and lofty goals involving centralized facilities. This potential for AR to visualize how irregular or discarded materials can be transformed into functional furniture is underdeveloped and underappreciated.

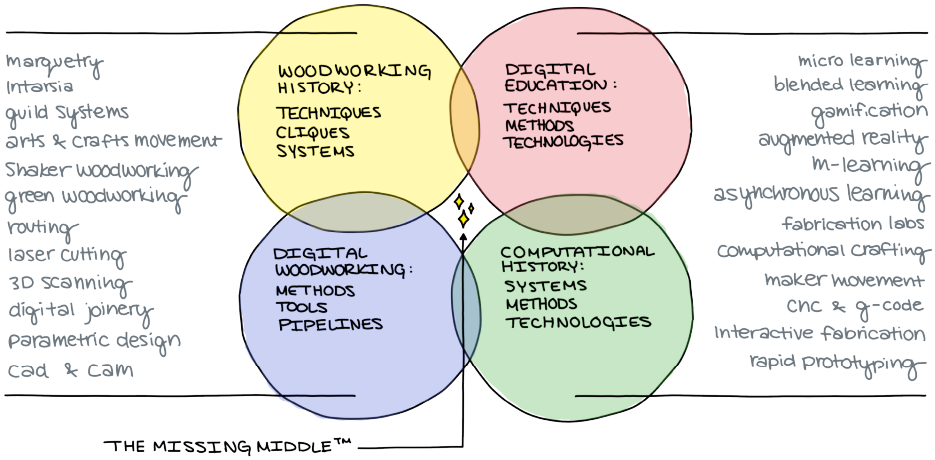


Figure 8: General fields of prior knowledge – including explored methodologies (Below Titles) and recurring terms of note discovered and explored throughout this project's research (Left and Right).

Additionally, while deep learning and other ML practices have been applied to architectural style classification and visual generation, their use in creating accessible tools for non-professionals interested in crafting is limited. There is an opportunity to develop systems that do not analyze or generate designs, but guide users through the fabrication process with a custom and individualized experience. The intersection of ML-driven design suggestions with user-friendly interfaces could revolutionize how individuals approach DIY projects, even if the scope of this project is limited to a very small piece of it [51]. Furthermore, the existing design based algorithms [49] and material optimization tools [58] that are out of the reach of normal users is an incredibly untapped resource. Nesting methods that optimize existing materials, and discovery resources that promote the sharing of pre-designed 'recipes' could be adapted away from closed ecosystems [46], proliferating and increasing in adoption as well as efficiency. The challenge lies in creating a pipeline that can handle

the variability and unpredictability of waste materials while providing practical and intuitive assembly solutions.

To explore this line of inquiry, the background and objectives take the form of several adjacent interests:

- What are the current barriers that prevent casual users from engaging in woodworking, and how can technology mitigate these challenges?
- How applicable are contemporary computer-scientific discoveries in unintended and under-explored fields?
- In what ways can pattern-based design approaches be integrated with UI/UX strategies to facilitate the intuitive creation furniture recipes by unfamiliar participants?
- What impact could potential digital workflows have on material waste reduction and the democratization of craft education across many avenues?

3.2 Objectives

The overarching objective of this thesis is to go through a holistic exploration of the process for designing, prototyping, and evaluating a computational workflow. By bringing together multiple fields of influence, and focusing on repurposing otherwise wasted materials, the system would aim not only to reduce waste but also to increase the number of people capable of enacting this kind of change within their own circles. This process could then assist interested participants in not only reducing the time spent on learning skills, but also reduce the barrier of entry for those individuals who might otherwise lack the confidence, time, or resources to learn traditional woodworking skills.

More specifically, the first objective centers on developing a CV-based pipeline for dimensioning and cataloging scrap wood pieces. By leveraging both novel and established published works, consumer products, and publicly maintained CV libraries – methods for estimating the size and shape of wooden offcuts through a standard smartphone or webcam monocular system are documented and tested.

A second objective focuses on designing an AR-based interface that guides a learner through the process of cutting scrap material into pre-defined parts, and culminates with a step-by-step assembly process. This includes overlaying cut lines onto key-pose tracked stock, and utilizing a secondary ML based object tracking package to identify and augment assistive 3D models to show the positioning of subsequent pieces during the assembly process. Rather than delivering static diagrams or printed text-based instructions, these overlays aim to blend interactive 3D elements and short animations, increasing the learner's understanding of core tasks – such as measuring, cutting, and nailing – without requiring specialized knowledge of design software or access to woodworking machinery.

A third objective involves collaborating with woodworking educators and iterating on the user-experience design to ensure that the system effectively addresses common misconceptions, hangups, and safety issues. By incorporating this on-the-ground feedback from instructors, shop managers, and novices alike – the workflow can ideally evolve to better reflect real-world needs as opposed to hypothetical or purely scholarly accounts. This section culminates with a final technical vignette that provides one possible method for utilizing multi-modal ML models as a proxy for direct instructional feedback based on the simple task of cutting a straight line with a hand saw.

Finally, this thesis briefly evaluates the impact of this workflow on both learning outcomes and materials usage. Through a series of user-tests with individuals who have little to no woodworking

background, the study examines whether AR-based guidance improves skill acquisition and reduces frustration or rework. In parallel, by utilizing pre-gathered scrap material from the Carnegie Mellon campus through these tests, a cause and effect relationship is used to validate the ability for the system to create something from what would otherwise end up as nothing.

3.3 Scope

Despite the broad ambitions of this research, certain boundaries inevitably shape what can be achieved. On the technical front, the current prototype prioritizes a monocular camera approach for object dimensioning and inventory creation. This approach is selected for its broad accessibility, and the system is designed to operate within the inherent precision characteristics of monocular vision, which can be influenced by factors such as lighting conditions, camera calibration, the color complexity of tracked items, and the framing provided by the end user. The current phase focuses on the core functionality, with the development of specific guidelines for optimal usage seen as a subsequent step. In addition, the project deliberately focuses on monocular technology to maximize accessibility, rather than incorporating more specialized technologies like LiDAR or stereo cameras, which, while potentially offering higher accuracy, would narrow the user base due to hardware costs and setup adaptability.

The educational content is specifically curated to cover fundamental hand-tool operations suitable for novices constructing basic furniture, focusing on foundational skills rather than advanced joinery, power-tool usage, or finishing techniques which are considered beyond the introductory nature of this project. Consequently, the tutorial content is designed to serve as a foundational stepping stone, equipping users with initial skills before they progress to more advanced projects or traditional learning resources. Ideally, this will act as a starting point for someone to feel comfortable continuing their journey with more traditional learning content.

The usability and user-testing protocols are defined by the parameters suitable for a thesis project, focusing on core concept validation. The testing phase is designed to highlight major usability aspects with a specific target user group, rather than encompass the full spectrum of all potential user scenarios, such as experienced woodworkers or those entirely new to digital learning tools. This focused approach means that further iterative development will be necessary to adapt the AR interface and workflow for broader application in larger educational contexts or diverse maker communities.

The project's scope also encompasses specific material considerations. The system is designed to perform optimally with softwoods and manageable hardwood offcuts in regular rectangular profiles – materials such as extremely warped, wet, or difficult-to-cut timbers fall outside the current design parameters for dimensioning, nesting, and assembly guidance. Users dealing with these materials will encounter challenges and inaccuracies in dimensioning and assembly instructions and the direct instructional scope of this thesis does not extend to the specific safety or toxicity considerations of working with reclaimed or treated materials.

Finally, integration with external resources remains beyond the immediate scope. While the interface and back-end is structured to accommodate additional features -like a crowd-sourced recipe library or advanced parametric design tools – these expansions, while anticipated in the system's architecture, are defined as out-of-scope for the current phase to maintain focus on core feasibility, as they necessitate a broader digital platform and community engagement. Overall, these defined parameters

establish a realistic boundary for the practical application of the proposed AR workflow in its current stage.

Overall, these defined parameters establish a realistic boundary for the practical application of the proposed AR workflow in its current stage. By prioritizing accessible hardware, fundamental hand-tool skills, and a guided learning path to assembly, the project aims to maximize its utility for novices, establishing a foundation upon which a more cohesive and fully connected system could be developed in future iterations.

4 Field Investigation

Woodworking and digital fabrication is ultimately a human, not a purely technical, problem. This study adopts a mixed-methods strategy in which empirical software experimentation is paired with qualitative inquiry. Rapid-prototyping cycles reveal what the computer-vision pipeline can do, while conversations with shop staff clarify what it *should* do to be trusted in real living rooms, classrooms, or maker-spaces. In the following pages, ethnographic work anchors the subsequent engineering decisions in practitioners' lived experiences – from time crunches to mental gymnastics – culminating with the translation from tacit professional wisdom into the explicit design requirements for the AR workflow implemented in the following Chapter 5.

4.1 Ethnographic Interviews of Woodworking Educators

While the technical viability of a computationally-assisted woodworking workflow can be demonstrated through prototyping alone, its educational relevance hinges on how well it addressed the lived realities of the instructors who supervise novices every day. Ethnographic interviews with shop managers, fabrication-lab directors, and entry-level instructors surface precisely the tacit expertise that rarely appears in textbooks or safety manuals: the subtle cues they watch for when a student grips a saw incorrectly, the compromises they make when a 30 minute orientation must cover both life-threatening hazards and basic etiquette, and the motivational levers they pull to keep learners engaged across steep software learning curves. Capturing these situated perspectives ensures that the design space for AR interventions is framed not only be technological possibility, but also by authentic pedagogical insight – time scarcity, retention gaps, and the perpetual tension between safety and creative freedom.

Incorporating practitioner voices also strengthens both the internal and external validity of this research. Internally, the interviews act as a needs-based assessment, aligning prototype features with clearly articulated frustrations. Externally, they provide a credible narrative for various stakeholders who must see that proposed solutions reflect more than just assumptions. Moreover, triangulating these qualitative insights with a mixed-methods approach, demonstrates that educational technology is most effective when engineered solutions are iteratively co-designed with the practitioners who would likely be tasked to deploy them.

Jon Holmes



Our first conversation focused on common instructional challenges in woodworking education, especially for students with minimal prior experience. We explored Jon's insights regarding new-student orientations, his opinions on barriers to woodworking seen in students, and the hurdles encountered while teaching basic hand tool and power tool usage.

During the interview, we first discussed the primary focii that they reinforce during new-student orientation-whether the emphasis tends to be on knowledge, form, or practice. We next discussed potential barriers that might prevent casual woodworkers from feeling comfortable or confident in the shop. Finally, discussed the most common learning hurdles, particularly which concepts or tools students typically struggle to master under his instruction, as well as some of the hurdles.

Time constraints often force the shop to prioritize demonstrations of the most dangerous objects, in particular power tools during orientation rather than covering hand tools or specific techniques (see Figure 9 A.). Practice and feedback loops are considered critical, with Jon highlighting the importance of hands-on repetition-such as proper sawing motions and hammer grips-and suggesting that AR or a guided digital workflow could create a secondary method of instruction that might offset some of the timing issues. Tool maintenance and setup also emerged as a frequent issue; the misuse of shared tools like hand saws and chisels underscore the need for best practices on use & storage, safe material setup, and awareness that many students lack proper woodworking spaces at home. Lastly, retention concerns stood out as an issue overall, as a single workshop onboarding session often fails to ensure that students remember every demonstrated technique, pointing to the potential value of an on-demand reference system (like AR instructions) to reinforce learning.

These insights highlight the need for concise tutorials focusing on critical tasks, such as safe cutting, measuring, and material handling. They also emphasize the value of real-time feedback to replicate some form of in-person guidance that is often unavailable outside a formal shop setting. Furthermore, the limited time students receive instruction suggests the importance of a digital "quick-reference" resource that allows novices to revisit essential steps independently and repeatedly, bridging any gaps left by restricted in-person training.

Future interviews will build on these findings by collecting additional perspectives on common pitfalls as well as the functionality of prototypical versions of this workflow. Jon's observations seemed both interested and dubious of this kind of instruction properly educating students or end users, but he did seem to believe that something was indeed better than nothing.

Steven Sontag



Steven Sontag, a recent graduate now instructing the introductory digital fabrication course at Carnegie Mellon, offered valuable insights during our conversation about the unique challenges faced by new students navigating complex digital fabrication workflows under tight time constraints. Steven noted the inherent difficulty students encounter in internalizing advanced software skills-such as learning Rhino, independently operating a laser cutter after only a single supervised session, creating properly formatted meshes for 3D print farms, and successfully configuring CNC toolpaths with RhinoCAM-all within a half-semester course structure (see Figure 9 B.).

Drawing from his recent experience as a student himself, Steven empathized deeply with his students' struggles, highlighting the disconnect between the ambitious curricular expectations and the limited time for guided practice. He emphasized how critical additional in-class time would be for enabling students to move beyond superficial proficiency and genuinely internalize essential fabrication skills. Steven's insights underscore the need for supplementary digital support systems or augmented instructional approaches to reinforce learning outside the constraints of limited classroom hours.

Terry Hritz



A subsequent conversation took place with Terry Hritz, Manager of the Design Fabrication Lab at the School of Architecture. While Jon Holmes's interview emphasized traditional woodworking challenges, Terry's perspective focuses on advanced fabrication tools such as CNC routers, laser cutters, 3D printers, and the software pipelines that power them – Figure 9 B, C, D. We discussed the most common technical hurdles that new users face when getting familiar with 3D environments, going between digital and physical objects, and the instructional strategies Terry employs to maintain enthusiasm and skill progression.

During the interview, we delved into the learning curves associated with computer-aided manufacturing (CAM) software and the challenges of juggling multiple digital platforms – e.g., Rhino, Ultimaker's Cura, or Grasshopper – before ever attempting to construct something in the real world. Terry noted that some students come with existing CAD skills but struggle to convert their designs into fully realizable models, often misunderstanding key concepts of constructability such as overhangs, planning for support structures, or the difference between raster, vector, and even 2D vs 3D datatypes paths for laser cutting. Additionally, he highlighted how novices may overcomplicate geometry in ways that are simply impractical once they need to factor in physical tooling sizes, the relation between time and detail, and the connection between digital and physical build-volume limitations.

One recurring difficulty lies in software complexity. While many students can sketch or model adeptly, fewer grasp the deeper manufacturing parameters – such as process simulation, order of operations, and model simplification. Terry sees this as a knowledge gap that arises partly from the steepness of CAM learning curves and partly from a lack of practice on real equipment. Furthermore, students often underestimate the importance and complexity of proper planning: they may design intricate forms without considering the difficulties of working at high resolution, or to conceptually orient their ideas in 3D space to minimize wasted material and time. This disconnect can translate into mismatched or misfitting parts, wasted stock, and frequent mid-job fixes or completely restarted projects.

The conversation then shifted to bridging 3D design and physical production. Terry underscored how critical it is for students to visualize each stage of fabrication, from cutting raw stock to finishing surfaces, and factoring in the capabilities of their particular tools. In his experience, novices rarely account tolerances leading to assemblies that don't fit together properly, or designs that break under the stress of production. A well-intentioned project can quickly become impractical if the user doesn't have support when grasping these real-world nuances.

A final hurdle that Terry emphasized is retention and motivation. Much like Jon's observations in the woodshop, one-off workshops / orientations or tutorials can spark initial interest but often fail to instill lasting knowledge. Students learn best with hands-on repetition and iterative feedback loops – yet digital fabrication often requires dedicated time slots, careful scheduling of machine use, and significant materials costs. As a result, novices may get frustrated if they can't quickly see tangible results, leading to a drop-off in engagement or creative ambition.

These insights mirror some of the challenges noted in traditional woodworking but extend them into the realm of digital fabrication, where software literacy and tool constraints compound the learning

curve. An AR-assisted workflow might work to mitigate both design and production pitfalls by visually guiding students in the correct setup of stock materials, the order of operations of assemblies, and contextualize software parameters as a project enters the real world. Additionally, incorporating short, repeatable tutorials within an instructional medium, covering everything from basic grip location to skill building activities, could offer a more "on-demand" method of skill reinforcement.

Future discussions with Terry will focus on integrating these AR or guided workflow concepts into tangible actionable steps. By creating a tight feedback loop between virtual design and physical assembly, the hope is to reduce wasted material and repetitive user errors. Terry's observations also underscore the importance of fostering a culture of iterative experimentation, especially one that supports novices at all points along the assembly process – reducing the chance of losing momentum or motivation along the way.

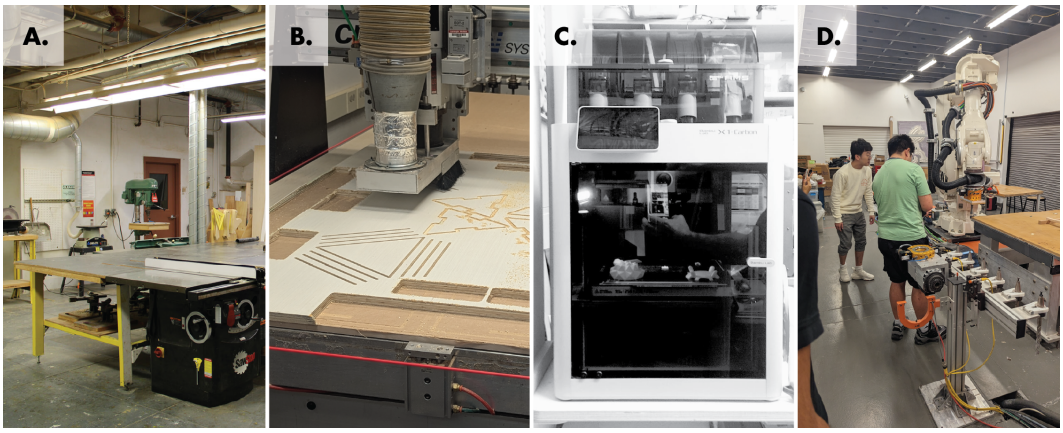


Figure 9: **A.**) CMU Architecture Woodshop. **B.**) dFAB CNC Router
C.) Bambu Labs X1Carbon 3D Printer **D.**) School of Architecture ABB Robotic Arm

4.2 Augmented Reality Environment Testing

Given the project's emphasis on AR overlays, identifying a robust and flexible AR platform became a central priority early in the prototyping process. A variety of frameworks were evaluated based on ease of integration, cross-device compatibility, available CV plugins, and overall spatial accuracy. Niantic's 8th Wall, a browser based-AR framework, initially showed promise due to its minimal hardware requirements and user-friendly web editor. However, its simplified interface and limited performance under more demanding conditions – such as detailed location detection and rich 3D asset management (see Figure 10 Right) – proved unsuitable for complex prototyping. This platform appeared more tailored to physics-based games and promotional content rather than flexible development scenarios.

Google's ARCore also provided compelling capabilities, including stable plane detection, plugins for cloud-based recognition (see Figure 10 Left), and robust motion tracking optimized for Android devices. Its extensive documentation was advantageous – however, the debugging challenges stemming from Android's native programming language Kotlin posed significant barriers, particularly given the project's objective for broad user adoption and rapid iteration.

Apple's ARKit, tested through Xcode on a Mac Pro 6,1 and prototyped on an iPad Pro 5th generation, offered superior out-of-the-box object tracking, precise spatial mapping, and comprehensive resources for developers. Although inherently platform-restrictive, ARKit provided valuable insights into direct



Figure 10: **Left**) Demo project in ARCore showing the ability to share camera and telemetry information with cloud-based ML applications. Credit to Dereck Bridie [68]. **Right**) Demo project in Niantic 8th Wall showing spatial capabilities and object tracking of a simple scene.

mobile API utilization, essential for transitioning towards a unified mobile application, as discussed in Section 7.4. Despite ARKit's current feature parity gaps compared to ARCore, these discrepancies are anticipated to narrow as Android continues expanding its AR capabilities.

To overcome platform-specific constraints and simplify development, Unity's ARFoundation, combined with Vuforia, emerged as the optimal solution. ARFoundation's ability to abstract the differences between ARKit and ARCore facilitated seamless cross-platform compatibility and accelerated prototyping within a single visual environment. Vuforia further enhanced this framework by introducing robust keypoint-based tracking ideal for recognizing objects with subtle textures and uniform tones, like timber components. Unity's well-supported ecosystem with extensive plugins, tutorials, and community assistance-solidified this combined approach as versatile, intuitive, and adaptable.

Ultimately, leveraging both native and cross-platform environments across Mac OS, Windows, Android, and iOS provided the necessary flexibility for extensive experimentation. This hybrid methodology allowed comprehensive exploration of AR capabilities, aligning closely with the project's iterative development goals and maximizing practical application within the final asynchronous system. In the following sections, a combination of Unity (by extension ARFoundation & Vuforia) and native ARKit development through Xcode and Swift is employed, allowing development and testing of specific vignettes across the most applicable operating system – including MacOS, Windows, Android, and iOS.

4.3 Monocular Object Detection

Monocular object detection is a CV technique that estimates an object's position, scale, and orientation using data from a single-lens camera (see Figure 11). This technology has garnered significant attention recently, partially due to its application in the pursuit of autonomous driving – where accurately gauging the distance between objects and their general size is critical. A key aspect of

monocular detection is six-degree-of-freedom (6DoF) tracking, encompassing translation along the X, Y, and Z axes and their individual rotation, vital for AR systems requiring precise alignment of digital overlays onto physical objects. Despite substantial advancements, existing implementations remain fragmented across different use cases. In exploring practical solutions, three prominent commercial frameworks were evaluated based on performance, usability, and cost-effectiveness.

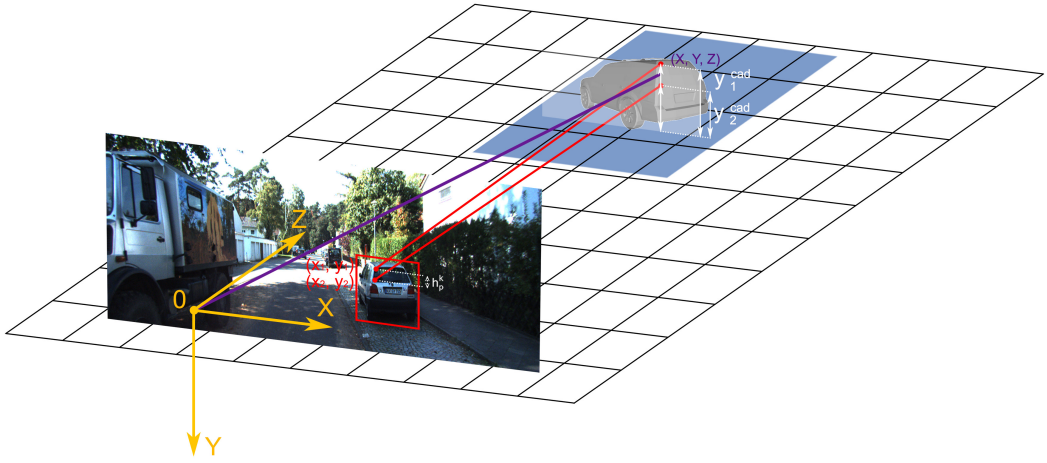


Figure 11: Diagrammatic process of the computation required to infer a 3D pose from a 2D image [40].

EasyAR Sense 4.0 (Figure 12 A.) offers an affordable but notably limited monocular detection solution relying primarily on a key-point algorithm to detect and track objects. While economically appealing, tests quickly revealed significant performance drawbacks, particularly in its ability to reliably detect less geometrically complex objects. Its algorithms struggled to recognize and maintain tracking for items lacking distinct, complex features, which severely restricted the range of usable applications. Although suitable for basic educational or hobbyist experiments, EasyAR Sense proved inadequate for consistent and robust project implementations.

Visometry's VisionLib (Figure 12 B.) strikes a better balance between performance and usability, integrating traditional key-point detection methods with machine-learning-enhanced frame-to-frame tracking. This hybrid approach improved stability and tracking continuity but introduced its own set of challenges. VisionLib occasionally struggled with accurately adjusting scale or re-detecting objects after temporary occlusions or significant viewpoint changes. Despite these issues, the framework demonstrated reasonable performance for moderately complex scenarios, making it useful in contexts requiring moderate accuracy without excessive licensing costs.

Vuforia Engine 11 (Figure 12 C.) provided the most comprehensive and reliable solution among the evaluated frameworks, employing advanced ML algorithms to quickly and robustly detect, track, and anchor digital content to physical objects. It demonstrated rapid and consistent object detection and orientation handling, even under significant rotation and varied environmental conditions. However, this superior performance came at a substantial cost, making it potentially prohibitive for low-budget or student-led projects. Nevertheless, its accuracy and ease of integration into Unity workflows validated its position as the best-in-class option for rigorous, professional-grade AR applications.



Figure 12: Screenshots of each tested commercial package tracking the same object. All tests were recorded on the same date and at similar times – all visual differences are package based [69, 70, 71].

Collectively, these high-level investigations – spanning ethnographic insights into pedagogical needs, evaluation of augmented reality development frameworks, and exploration of monocular object detection techniques – provided the foundational knowledge and technical direction necessary for the subsequent design and implementation of the following workflow. The findings from this phase directly informed the specific architectural choices and user experience considerations throughout the implementation and user evaluation – acting as a best-practices overview of the various neighboring fields.

5 Technical Implementation

Having established the research context, identified pertinent gaps, and defined the overarching objectives, this section provides a comprehensive account of the technical development and practical implementation of the explored prototype. It translates the conceptual framework discussed previously into a tangible system, detailing the specific methodologies, software tools, and algorithms employed to realize the vision of an AR and CV-assisted craft woodworking companion. The subsequent subsections guide through each stage of a user's interaction with the system, from initial setup and material input to the guided creation of a complete project, illustrating the functional application of the explored technologies.

5.1 Overview

This chapter details the end-to-end pipeline that guides a first-time woodworker from initial device calibration to the completion of a full furniture recipe (see Figure 13). The section is arranged as a staged workflow so the reader can follow the same steps a user experiences inside the application:

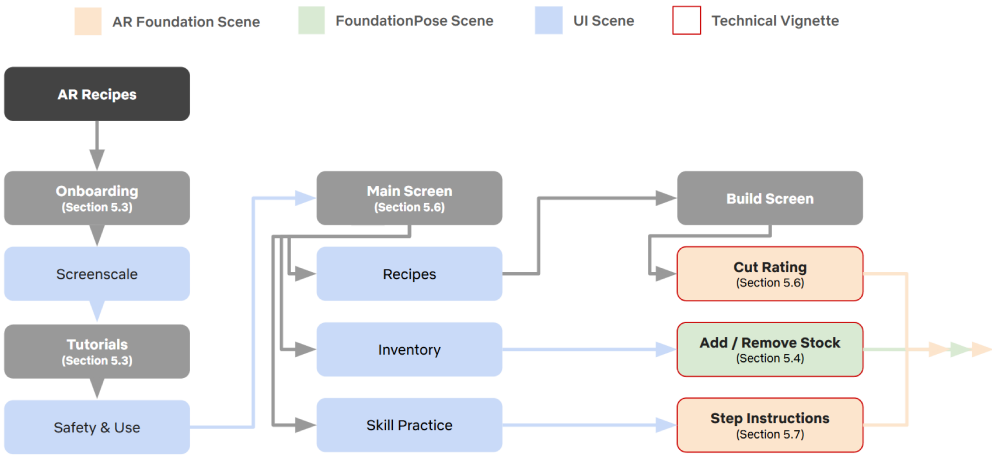


Figure 13: Flow chart of the proposed implementations – color coded by method and ordered by user flow.

5.2 Kit of Parts

To successfully prototype and test the ScrapAR workflow, a clearly defined toolkit was assembled, consisting of accessible and commonly owned tools intended to support new learners with minimal experience and financial investment in woodworking or fabrication. This set of required and optional items was carefully selected to balance affordability, availability, and ease of use, reflecting the project's goals of catering to an uninvested and transient user base (see Figure 14). Each tool plays a specific role, either in facilitating basic woodworking tasks, supporting accurate tracking, or serving as a supplement to safety or assembly.

Required items:

IKEA – TRIXIG 15-piece Tool Set:
Chosen for its affordability and its place as likely one of the most common tool sets in the United

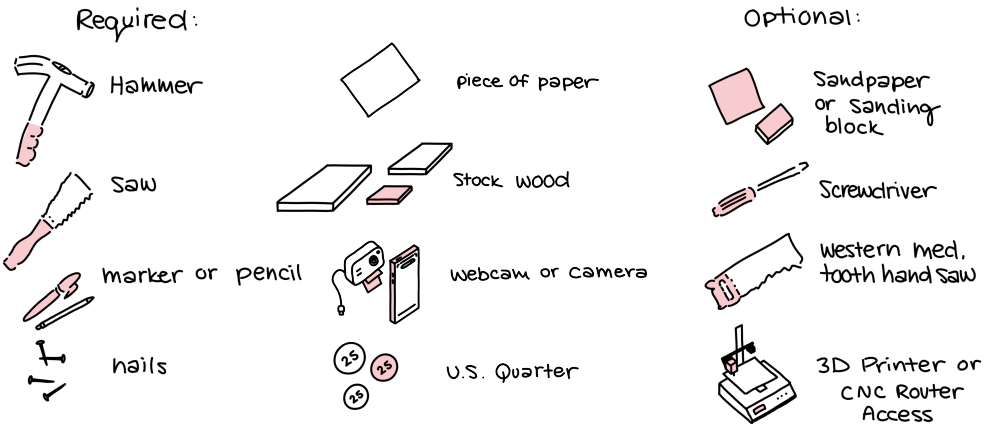


Figure 14: Required and optional tools for the proposed workflow.

States, the TRIXIG kit provides a compact hammer with a rubber cover, pliers, and other basics in an easy to find package. In the context of this system, the hammer is the most important element, acting as an object of known scale and geometry, making it useful for tracking and scaling with various CV implementations. In general, this kit tends to be one's first foray into small-scale construction tasks – and is often recommended on forums [72, 73, 74] for those looking to get into home improvement without a large investment in professional equipment.

Nail Assortment Kit:
Featuring multiple sizes and lengths, a nail assortment kit allows for versatile assembly across various projects. A simple kit that can be found for $6 online ensures that users can securely attach pieces of wood without needing to visit traditional hardware or home improvement stores [75] – despite Wolf et al. and Lowe's estimating that women initiate 80% of home-improvement projects, and are "highly involved" in 90% of home-improvement decisions [76].

Japanese Pull Saw:
A two-sided Japanese pull saw was selected for its safety and ease of use. Unlike western push-style saws, these saws cut on the pull stroke, requiring less force while providing more control. Its thin blade allows for precise cuts, which is especially important when honing skills, working with irregular off-cuts, and when building without a purpose-built workstation.

Monocular Camera (Phone or Webcam):
Central to this system's CV approach, a smartphone camera or basic webcam enables the user to scan and measure scrap materials. By capturing images or streaming video, the camera is used for estimating dimensions, identifying tools and parts, and providing a canvas on which to overlay "recipe" instructions in real time. This workflow also reduces the need for specialized equipment like stereo or depth-of-field (DoF) cameras utilized in other approaches [77][78], making the system widely accessible to most people without needing to purchase a new device.

Pencil:
A pencil is expected for marking cut lines or guidelines on the wood once provided object dimensions. The act of drawing a line along a measured dimension helps a user confidently align their saw and ensures a more accurate cut than attempting to cut through the augmented display itself, something

tested in Making A Recipe. Finally, pencil marks can be easily erased or sanded away, maintaining a clean finished surface.

Paper:
A standard sheet of paper is expected for use as a speed square. During the tutorial steps, a small graphic is shown that prompts the user to fold this paper in a way that ensure it will always have a 90 degree angle.

Table or Large Surface:
A sturdy, flat work surface is critical for safe and accurate woodworking, whether you're hammering nails, cutting boards, or measuring angles. By laying materials on a stable table, the end-user can minimize wobbling and reduce the likelihood of mistakes or accidents.

Optional Items:

Sandpaper or a Sanding Block:
Due to the unknown quality of available scraps – and the likely imperfect quality of a user's first few cuts – having a method of smoothing finishes or rounding edges might be a good quality-of-life addition to the toolkit. This is currently not going to be mentioned as it would just add to the scope creep.

Screwdriver:
Support for this tool isn't planned, but it is included in the TRIXIG toolkit. This could be an avenue for alternative fasteners into the future.

Western Saw:
Could be implemented quite easily using similar methodology of the existing pull-saw. This isn't going to be included as it is traditionally seen as a more dangerous style of saw.

3D Printer or CNC Router Access:
This method is briefly explored in Section 7.4. It would be interesting to have a method for generating parametric parts that could be cut on CNC or printed on a 3D printer. If possible, it would be interesting to have a proof-of-concept object to end the final presentation, acting as possible next steps.

5.3 On-boarding

The following subsections guide the user through the onboarding process (see Figure 15): (1) calibrating the application using a coin-based scaling technique, (2) selecting and scaling their tracked tool, and (3) completing a brief animated safety module for three primary categories: general, hammer, and saw.

Screen Calibration To make full use of the stock measurement in Section 5.4, users must first calibrate their screen size to a known object. This relies on placing a quarter atop a circle image displayed on the user's device and adjusting the scale of the on-screen prompt until the circle perfectly matches the physical quarter.

Hammer Selection and Scene Scaling Next, users identify whether they will use the recommended TRIXIG hammer or an alternative model. For the latter, they must input the hammer's overall length to adjust the CV object tracking scale, allowing the digital overlay to appropriately match the real



Figure 15: Splash screen presented to a user upon launch.

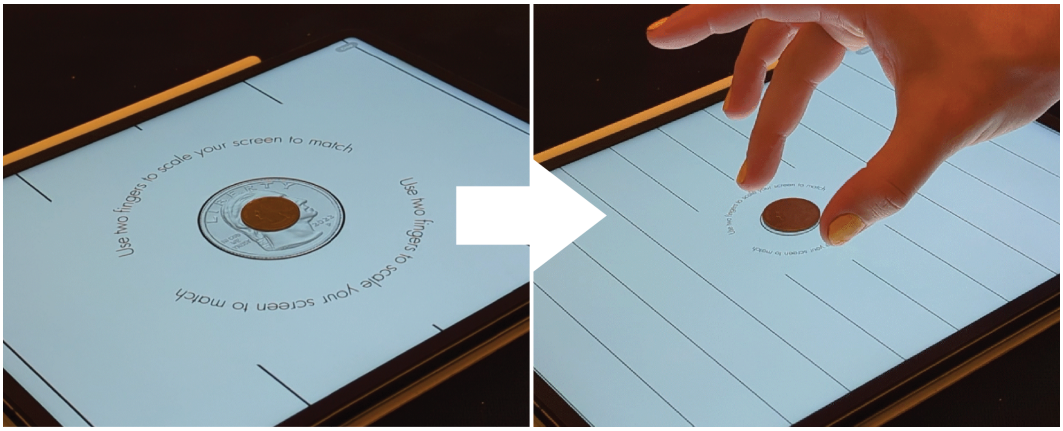


Figure 16: UI of screen calibration – the circle matches the dimensions of a U.S. Quarter and scales a visual ruler for use in Section 5.4.

world. By defining the size of a key reference tool the workflow warps and scales the scene to discern the relative size of other objects for Section 5.4.

Tool Tutorials The following tutorials were created in conjunction with Jon Holmes in Section 4.1 and by referencing publicly available beginner-focused woodworking resources on YouTube [79][80][81] and in print publications [82][83]. Their primary aim is to lower the learning curve for first-time users by providing a short but structured initiation into the core actions and safe usage practices that underpin a beginner's project.

Before attempting any 'recipe,' the users must complete three module groups covering the hammer, saw, and a broader "general" tips – each module being split into three short animations that focus on core hand-tool and beginner competencies.

General Tutorials

This trio of tutorials is intended to ensure that each user has the bare minimum in terms of space and work surface, a small list of expectations to ensure that their tools are fully functional, and a small paper folding instruction to ensure that they have a right angled object for making 90° part outlines.

- *Folding a Paper Speed Square:* Illustrates how to create a makeshift right-angle guide using a simple piece of paper. This ensures novice woodworkers can quickly mark accurate angles without specialized measuring tools.
- *Choosing a Work Surface:* Advises on picking a sturdy, flat table or bench to help maintain control and reduce the risk of materials slipping. A stable surface also improves AR tracking accuracy.
- *Maintaining Tools:* Briefly mentions that looking up proper care instructions for particular tools can help extend their effective lifespan, reducing the need for purchasing new versions, sharpening existing tools, or requiring unnecessary force for use.

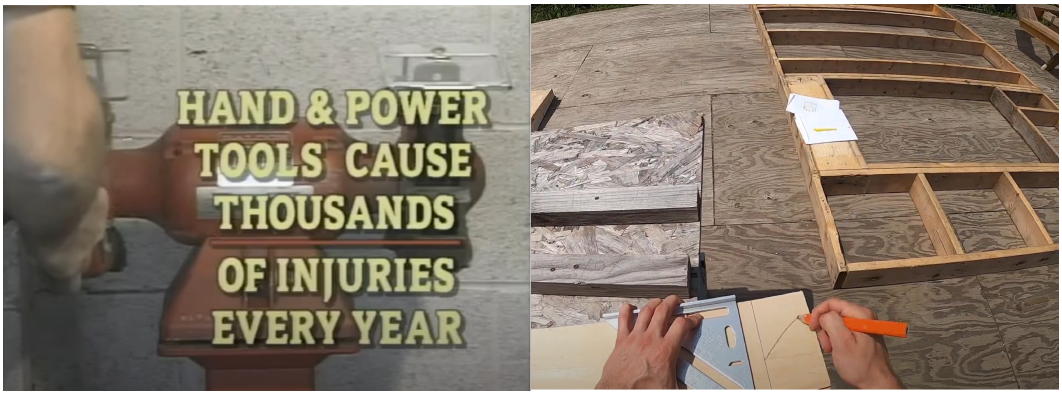


Figure 17: Example frames from the general tutorial clips. Left shows how improperly maintained tools are the most common cause of injury. Right shows how a paper or metal speed-square can be used to properly mark a cutline. Credit to Oklahoma State University and Dan Leipert [81, 80].

Module	Tutorial Description	References
General	Folding a Paper Speed Square	[84]
	Choosing a Work Surface	[80]
	Maintaining Tools	[84]
Hammer	Where to Grip	[79]
	How to Swing	[79, 81]
	Starting a Nail	[79]
Saw	Where to Grip	[80, 81]
	Where to Stand	[80]
	Use the Whole Blade	[82, 83]

Hammer Tutorials

These tutorials, gathered from several sources, attempt to preempt the most common issues with using a standard claw-hammer. While the Ikea TRIXIG hammer itself is technically non-standard, the prevalence and the ability to easily adapt these tutorials to different hammers retain its place as the tool of choice going forward.

- *Where to Grip:* Emphasizes how holding the hammer too close to the head sacrifices leverage, while gripping too far down can reduce accuracy. Proper grip fosters better control and minimizes user fatigue.

- *How to Swing:* Demonstrates a fluid, rhythmic motion rather than a stiff push. Correct swing mechanics prevent bending nails and reduce strain on wrists and elbows.
- *Starting a Nail:* Shows how to lightly tap a nail into the wood so it stands upright on its own. This improves accuracy when driving it the rest of the way and lowers the chance of slipping or self-injury.

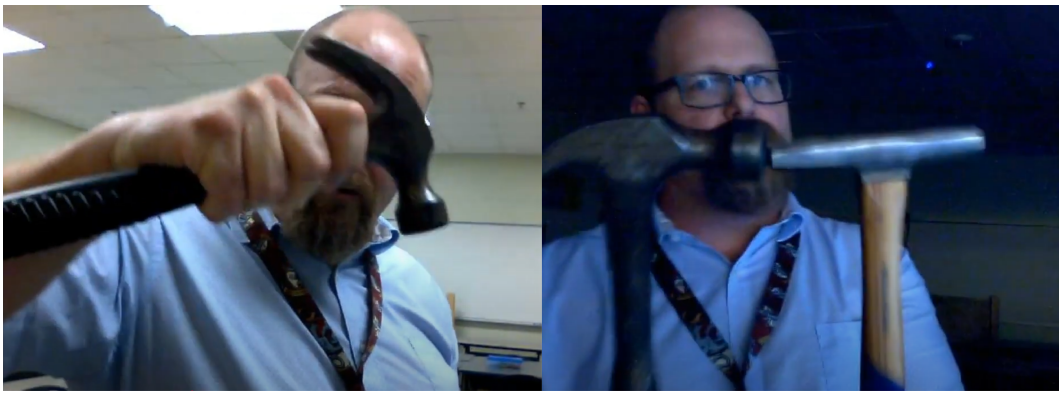


Figure 18: Example frames from the hammer tutorial clips. Left illustrates the use of leverage to drive the nail. Right shows the perpendicularity of the hammer against another hammer – acting as a proxy for a nail. Credit to Brandon Gilbert [79].

Saw Tutorials

While many saws could work for this system, the tutorials provided focus specifically on the Japanese pull saw as it has a lighter weight and better maneuverability than typical push saws. Similar to above, the general themes explored in these tutorials can be easily generalized to other tools as the user gains comfort with the movements and methods.

- *Where to Grip:* Encourages users to hold the handle firmly but not rigidly, allowing the blade to do the work. This reduces blade wobble and improves cut precision.
- *Where to Stand:* Guides the user on maintaining a stable, open stance-ideally with the saw blade in clear line-of-sight. Proper positioning reduces kickback and uneven cuts.
- *Use the Whole Blade:* Stresses the importance of smooth, full-blade strokes rather than forceful thrusting. This preserves the blade's sharpness and delivers more consistent, controlled cuts.

5.4 Inventory Creation

One of the largest hurdles in this project is creating a method to get the approximate dimensions of scrap materials from a photo or video feed. While there are many possible ways to accomplish this more accurately through traditional measurement, or through CV utilizing physical markers – the emerging field of monocular object scale detection has been making great strides, driven primarily by autonomous driving [39] and smartphone research [85]. In this section, three main methods were explored – ranging from a simple fully-connected neural network to calculate the distance of known objects, to a more contemporary method able to deduce an outline from an approximate 6-degree-of-freedom (6DoF) location of a hammer in a scene. While the current pipeline works with



Figure 19: Example frames from the saw tutorial clips. Left illustrates the light grip recommended on American and Japanese saws. Right shows how the angle of the saw changes over time and as the cut grows. Credit to Dan Leipert [80].

a high level of oversight, a more robust method is a top priority, such as that discussed in Figure 20, and in Section 7.4.

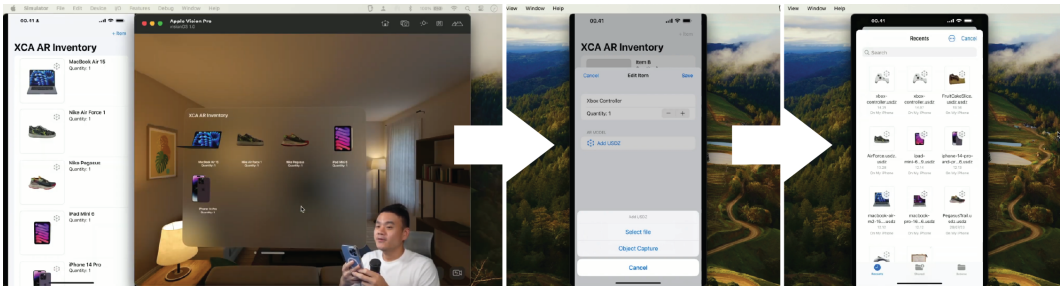


Figure 20: VisionOS & iOS AR Realtime Inventory app using SwiftUI, ARKit, RealityKit, and Google Firebase – Alfian Losari [86].

The initial approach explored, aligning with the simpler end of the spectrum mentioned, was a high-level reimplementation of an open-source model called NeuralDistance [87], which itself leverages an open-source CV project called YOLO [88]. Although it is a bit antiquated at this time, the original aim of NeuralDistance is to utilize a custom multi-layer-perceptron (MLP) on top of the existing object detection convolutional neural network to roughly predict the distance of recognized objects in the scene. Although it wouldn't be directly applicable to finding the 6DoF location of objects, finding the approximate size of an object due to its proximity to other known sized objects seemed like a reasonable first step. By potentially building upon this kind of model, another system could hypothetically output scaled bounding boxes based on their relative scales and distances to the camera. If functional, this approach would allow for a method to quickly assess and store information about complex scenes providing a low effort trial project.

The original NeuralDistance project employs a fairly common CNN called YOLOv3, so named for its 'You Only Look Once' processing architecture, to identify objects in images or videos. The novelty of this particular implementation is the additional layer that utilizes a simple parallax calculation and the output from the CNN in order to estimate the distance of objects in the scene from the camera. Thus, the reimplementation's goal is to further adapt this framework into the context of woodworking, focusing on detecting pieces of lumber and estimating the dimensions in something approximating real-time. Due to this constraint, the decision to run things locally was chosen for control and to

simplify the streaming of video frames into the network during inference. The language of choice was Python 3.12, set up in an Anaconda virtual environment to prevent overlapping dependencies – including PyTorch for neural network modeling, OpenCV for image processing, and Matplotlib for visualization. The hardware was a desktop computer with a 3080ti, 32gb of RAM, and a generic brand 1080p webcam for testing.

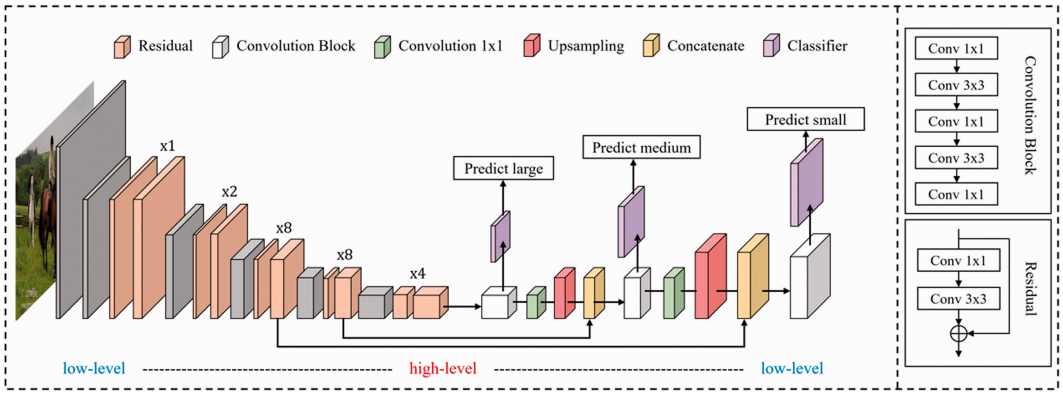


Figure 21: Breakdown of the connected layers in the YOLOv3 CNN [89, 90].

To further simplify things for this initial test, a new data set was not assembled. While it would be nice to augment the YOLOv3 pre-trained model to be hyper-sensitive to objects with known dimensions such as tools, or typical household objects, it was completely out of the purview of this learning exercise – and likely even of a thesis project in general. In a typical workflow, images would be captured from various angles and distances to ensure diversity, and then annotated using a tool such as LabelImg [91]. This would essentially create simplified bounding boxes around the reference objects and record their actual distances from the camera using captured measurements at the time of photographing, or a structure-from-motion technique to find an approximate value.

The annotations would then be stored in a local dataset that could be easily loaded, ensuring that it would encompass many different lighting conditions, backgrounds, and object sizes to improve the robustness. In addition to this work, it would be simple to add data augmentation strategies such as grayscale, cropping, noise, and other non-warping techniques. Without the data prepared, it wasn't necessary to finetune the YOLOv3 model, although there are many tools available to assist in this process. Using transfer learning with pre-trained weights would drastically reduce the training time and enhance the model's performance – as would adapting the YOLOv3 version of the implementation to a more contemporary model such as YOLOv11 [92]. This would need to be followed by adapting the existing neutral network for dimension estimation to align with the slightly different output format of the newer versions.

Currently, to calculate the distance of an object to the camera, a three-part system is used: First, the stock YOLO model is used to detect high-confidence objects in the scene that fit a known dimensional dataset. Second, the built-in bounding boxes available during inference are calculated as dimensional pixel values and the entire frame is scaled (but not warped) to their known dimension. Finally, with this scaling factor available for different objects, the approximate perspective of the camera can be triangulated, including where a ground plane and hypothetical depth map exist in a simplified 3D scene.

The integration of object detection and dimensional estimation into a cohesive pipeline was implemented using a VSCode notebook, similar to Jupyter. This contains the code to download and import modules, load the pre-trained YOLOv3 mode, and the secondary layer to process images and detect objects to approximate dimensions. Processed images were saved each 300 frames, approximating 10 seconds. This allowed for testing the pipeline quickly, and changing parameters on the fly without reloading entire datasets or reprocessing finished sequences. Despite the progress and uniqueness of this process, there were many challenges throughout the reimplementation. One significant issue is the complexity of the code base – many different libraries have been changed in the year since this model's creation, requiring adjustments in imports, refactoring, and completely rewriting sections that have been depreciated. Additionally, understanding and adapting the original NeuralDistance code demanded a deeper than expected understanding of the YOLO framework itself.

While the existing process for calculating the scaling factor of a scene was complex, it did provide a good amount of insight for how a similar system could be created to understand height and width indicators. As the scaling factor is dependent on the distance from the camera to the objects, any misunderstood differences in depth between the reference objects had the ability to drastically change the outputs. In the future, it will likely be necessary to coach users on how to photograph their offcuts to ensure a realistic capture. Likewise, this implementation shows the power of having known scale markings in the scene.

Overall, performance was also a concern, especially when attempting to run in real-time – a large issue if the intention of this thesis is to prototype a mobile friendly setup. While the YOLOv3 model once provided high accuracy at a relatively light weight, there are many models by Ultralytics and others that easily beat this version in both efficiency and accuracy, and even both at the same time. During the testing phase, YOLOv3-Tiny was also tested to improve inference speed – although it did so at an incredible drop in accuracy. Going forward, integration of this kind of method would also prove difficult. While it was pretty trivial to set up the local environment with specific dependencies, it would be much harder to get this kind of system running in an environment or larger script with many other programs. The dependencies have quickly become outdated, and having access to both the CNN model and the custom model ended up exhausting most of the 32gb of available RAM.

Despite the many challenges, this attempt was also an incredible learning experience. The reimplemented pipeline successfully functioned, and some of the provided depths were within 25% of their actual value. This system is definitely not capable of fully identifying the scale of anything, but it does provide a good framework and proof of concept for continued development of this segment of the proposal. There are many different implementations, many far more contemporary, that claim to handily outperform NeuralDistance.

Furthermore, this exploration provided valuable experience in addressing real-time performance challenges. There were many interesting tradeoffs that needed to be attempted, tested, and learned from, and that is something that just can't be recreated without physically trying something. For example, the method of using known reference objects to calculate the scaling factors proved effective and lightweight, but would need much more efficient methods of implementation if it were to run on a mobile device – or even generalized into a more robust system with a larger knowledge base.

So far, all of this work has aligned with the core goals of the proposed thesis, allowing individuals without traditional training to engage in woodworking. By providing a tool that enables quick assessment of scrap materials, they can make easier decisions about the worthiness of waste, fostering sustainability in their lives and potentially influencing others'. In conclusion, the reimplementation

of NeuralDistance not only demonstrates the capabilities of neural networks in detecting objects and estimating their dimensions, but also illustrates their potential application in non-computer-science tasks. By finding the approximate size and shape of scrap lumber through known objects, this approach offers a glimpse into streamlining a process that dissuades users and promotes waste. The project's outcomes and lessons learned pave a personal way for further and research opportunities in this domain, with the potential to make a meaningful impact on sustainable practices and the democratization of craft.

To address the limitations of the previous approach and advance the investigation, a more robust model capable of inferring the full 6DoF pose of an object was required. Although 6DoF pose estimation is inherently more complex than simpler distance estimation tasks like those handled by NeuralDistance (and thus potentially more susceptible to failure and inaccuracy), the research field dedicated to this challenge is considerably more mature, offering sophisticated solutions. OnePose++ represents a state-of-the-art, keypoint-free method for monocular object pose estimation without relying on pre-existing CAD models [43], making it particularly suitable for applications involving arbitrary household objects. The technique builds upon the concept of one-shot object pose estimation, requiring only a short video sequence with annotated object poses for initial setup. Unlike traditional methods, OnePose++ reconstructs a semi-dense point cloud through a Structure-from-Motion process, eliminating the need for detecting repeatable image keypoints, and thus, significantly improving performance on low-textured objects – something that will be important given the homogeneity of wood products. This keypoint-free feature matching leverages a multi-stage, coarse-to-fine pipeline, providing an effective and adaptable framework for object localization in AR.

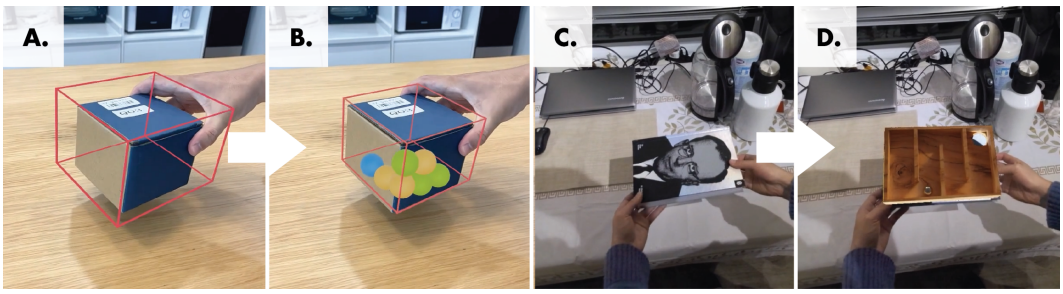


Figure 22: **A.**) Bounding box overlay on a simply textured boxy object. **B.**) Bounding box used to superimpose a physics based interactive experience. **C.**) Rectangular object tracked from a head-mounted camera rig. **D.**) Object used to interact with a physics based augmented game as the primary controller. [43]

Implementing OnePose++ involves first capturing a sequence of annotated reference images from multiple viewpoints. Using a coarse reconstruction phase, a rough point cloud is initially generated using semi-dense correspondences from these images. A subsequent refinement phase improves point accuracy, optimizing the reconstructed cloud with sub-pixel precision correspondences. Finally, during real-time application, it employs a 2D-3D matching network that directly matches camera imagery to this optimized 3D model, establishing accurate object poses rapidly and reliably without the cumbersome requirement of generating a detailed CAD model beforehand.

Despite its substantial improvements in accuracy and robustness-particularly with low-textured objects – OnePose++ faces significant limitations in practical implementation. Most notably, the computational resources required are substantial, demanding large amounts of RAM that exceed typical local computing capabilities. This high memory requirement makes OnePose++ impractical

for local deployment, particularly on mobile or lightweight AR devices. Consequently, its use is currently constrained to scenarios where extensive computational resources, such as cloud-based processing or high-performance workstations, are readily available.

While OnePose++ demonstrated advanced 6DoF capabilities, its substantial computational demands highlighted the need for a more practically implementable solution. Drawing from the insights gained from both the initial NeuralDistance experiments – which underscored challenges in basic scaling – and the exploration of sophisticated models like OnePose++, a new proof-of-concept pipeline was developed. This approach aimed to integrate several promising open-source algorithms to balance robustness with manageable resource requirements, moving towards a more feasible 'final prototype'.

A central component of this new proof-of-concept pipeline was FoundationPose, a unified framework created by Nvidia for 6D pose estimation and the spatial tracking of unseen objects. It employs a scoring network and a pose refinement network to initialize the object pose in the first frame and subsequently track it in sequenced frames. Although the algorithm itself is extensible to work from many different data-types – During the initialization stage, the framework loads an OBJ file-along with any textures and UV data – computing an oriented bounding box for visualization, as well as several debug point clouds and diagrams. Part of the strength of this pipeline is that it can utilize either valid depth images or placeholders; then applying an intrinsic matrix to project between 2D and 3D space, which is crucial for accurate pose estimation – which can be complex to calculate.

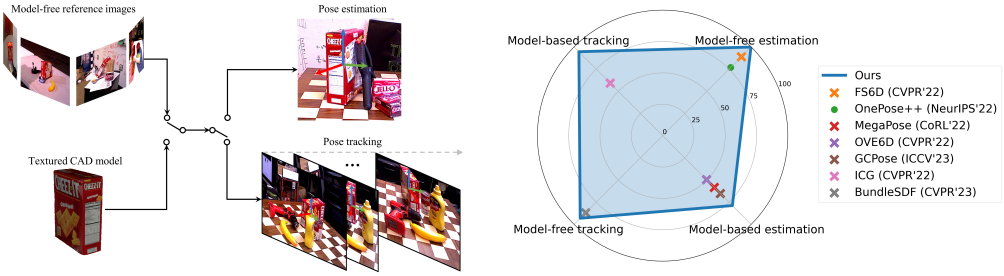


Figure 23: **Left**) Diagram of adjustable input / output connections. **Right**) Diagram of input/output pairs compared to non-unified algorithms of each type. [42]

Prior to loading the model into FoundationPose, the OBJ file must contain valid UV coordinates if it is intended to include textures – requiring specific embedding and material linking workflows. Despite the models for this exploration being created in Blender – it became evident that exporting the model from meshlab produced a more refined final object and better texture linking. The MTL file should point to an image texture without an alpha channel; relying solely on a material color without an actual texture can cause the code to fail when it attempts to load non-existent (NoneType) UV data. In scenarios that require placeholder depth images, the images must be single-channel (16 bit, greyscale) and have non-zero values to provide meaningful information to the depth-based refinement step – with a domain of 1mm per change of 1 in the intensity of the pixel. Otherwise, the registration process may fail to find enough valid depth pixels to estimate even the initial pose.

Accurate camera intrinsics are also essential for consistently mapping pixel coordinates to 3D points, as can be seen in the exaggerated example from Figure 24. This data is stored in a file such as cam_K.txt, which contains the camera's focal lengths and principal point. If the bounding box appears disproportionately large or the object seems to "spin" erratically, common causes include

an oversized or off-center mesh and misalignments in either the coordinate frame or the camera intrinsics. Verifying the mesh scale, ensuring proper centering, and confirming that the initial registration converges are critical steps to maintain alignment.

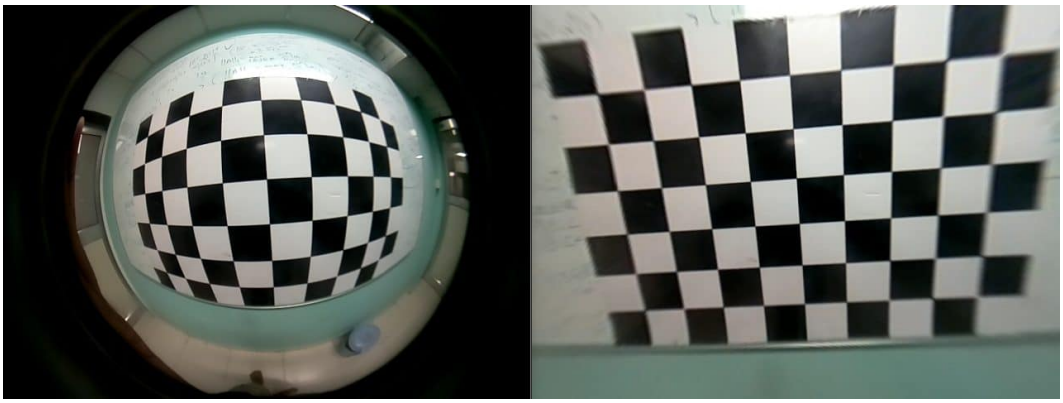


Figure 24: Before and after images of a frame being treated with a camera calibration file [93].

Implementation involves following the official instructions from the FoundationPose repository. Environment setup was pre-defined through a Docker container, after which the scoring and refinement weights are placed in new directories, and the relevant test data is organized under a predefined demo folder. Executing python run_demo.py demonstrates a sample object being registered and tracked across a sequence of images. Replacing the example mesh with a custom object involves providing the corresponding OBJ file, single-channel or genuine depth images, and a correct camera intrinsic matrix. Although not applicable in this particular exploration, it seems that cases where the novel object differs substantially from the examples in terms of geometry or appearance, retraining or fine-tuning is downright necessary for any usable results.

Given that FoundationPose can utilize depth data for enhanced performance – and a key project constraint is the use of standard monocular cameras – a method for generating approximate depth maps from 2D images became an important consideration for this pipeline. To address this, DepthPro, a software tool designed to infer depth from ordinary images or short video sequences, was investigated as a complementary component.

DepthPro is a software tool designed to generate approximate depth maps from ordinary 2D images or short video sequences, bridging the gap for workflows that rely on depth data but lack dedicated sensors. Internally, it employs a combination of monocular depth-estimation models and photometric refinements to infer per-pixel distance values [85]. Although DepthPro can work with diverse image sources, it by default produces a false-color depth map, which requires modification to output 16-bit single-channel data compatable with downstream processes. This step is crucial for applications like FoundationPose, which rely on consistent depth data for bounding-box refinement and alignment. For best results each input frame is well-lit, free of excessive motion blur, and framed so the scene of interest is clearly visible. DepthPro seems to be highly sensitive to image quality and can produce noisy or artifact-ridden depth maps if the scene is too dark, if the camera moves too quickly, or if textures in the environment cause confusion.

Setting up DepthPro involves installing its required deep-learning libraries – in this case PyTorch – and placing the pretrained weights in a prescribed directory. Setup is automated for WSL and can be quickly loaded into a Conda environment, with it up and running in only a few minutes.

Although it can be a useful shortcut for generating depth maps from standard 2D input, this version of DepthPro is not currently a long-term replacement for true depth sensors. In scenarios where the target object significantly differs from DepthPro's training data retraining or fine-tuning the underlying depth network may be necessary to achieve stable performance – but this was thankfully not necessary during this exploration. As a proof-of-concept, however, DepthPro seems to be sufficiently capable of generating approximate depth maps so that software like FoundationPose can operate in simple environments without a dedicated depth camera.

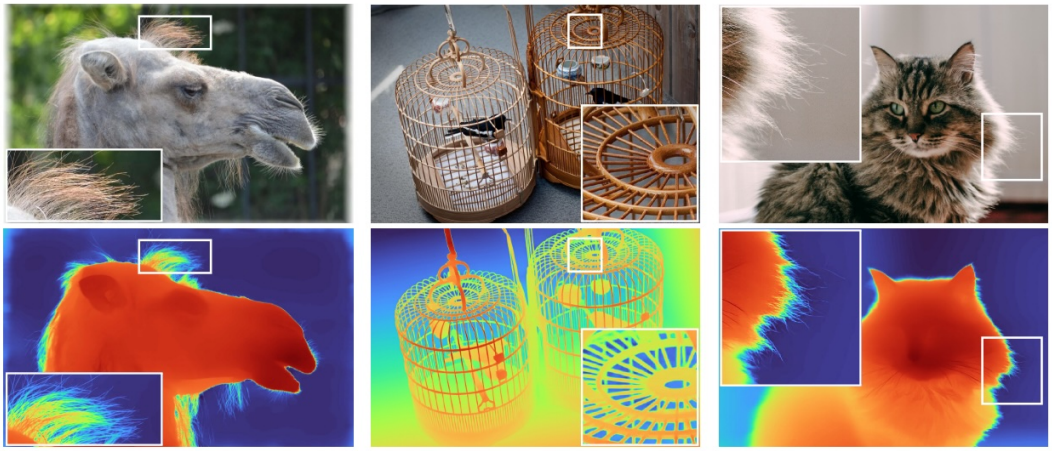


Figure 25: Various examples of DepthPro working on complex images [85].

Building upon these individual explorations, the next phase involved attempting to combine key components – specifically FoundationPose for 6DoF tracking and DepthPro for depth approximation – into a partially-integrated prototype workflow. The primary objective of this experimental pipeline is to isolate a vector outline from a piece of scrap wood, using a single video frame and referencing a known-size object (the Ikea TRIXIG hammer). It is important to note that while this approach

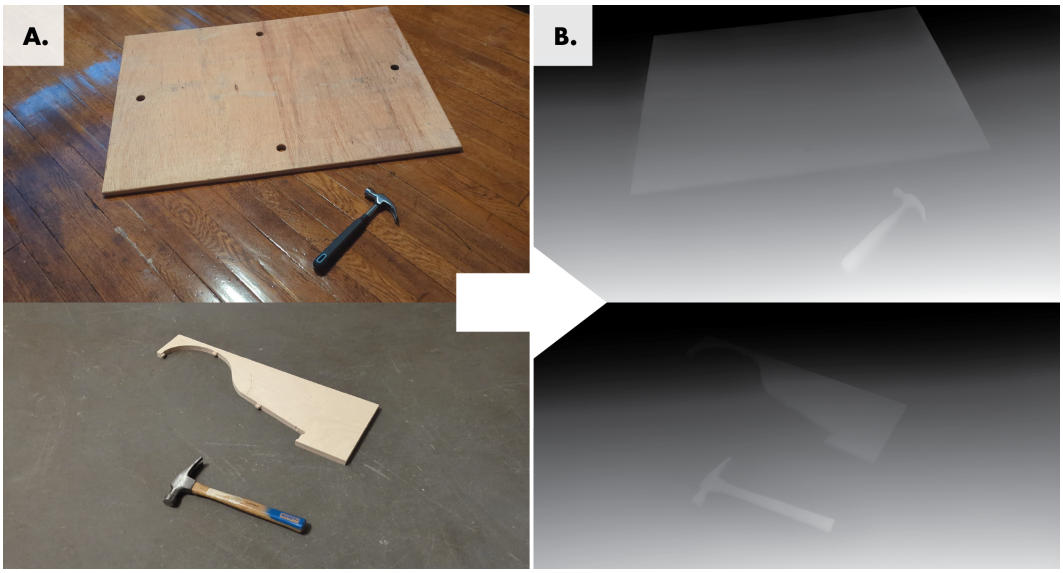


Figure 26: **A.**) Before images captured through a single lens. **B.**) Depthmaps created through DepthPro – color corrected to show detail.

demonstrates feasibility in principle, its current implementation is very much a proof-of-concept, exhibiting inaccuracies and requiring carefully chosen images and controlled conditions. The subsequent paragraphs detail each major step in this integrated pipeline, highlighting current limitations and potential areas for improvement.

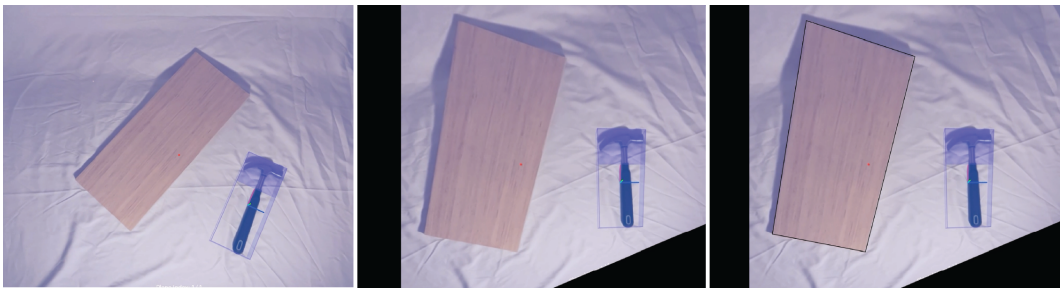


Figure 27: Overall method from raw captured video frame to final outline of scrap wood.

Because FoundationPose expects a real depth sensor, we approximate depth by running DepthPro on each frame from the video. DepthPro is extremely sensitive to lighting and image quality, so only a subset of frames are clean enough for the bounding box to work as intended. In practice, a good deal of cherry-picking is required to find frames that produce usable depth maps. For instance, frames with stable lighting, minimal motion blur, and clear color contrast around the hammer generally yield better results (see Figure 26).

The process continues by feeding individual frames from a short video clip, along with the newly created depth map, into FoundationPose. The 3D model of the TRIXIG hammer is used as a reference .obj – after a long process of properly formatting it to be compatible with trimesh, letting FoundationPose return a rectangular 3D bounding box that encloses the entire hammer – including its head. The system is tested on video captured by two devices (the main camera of a 5th generation iPad Pro and the main camera of a Google Pixel 7 Pro), each requiring a separate intrinsics file for correct calibration (see Figure 24). While this bounding-box identification typically holds steady – as the hammer is not symmetrical enough to cause directionality issues – it still depends heavily on reliable depth input and this part of the process is the most likely to fail.

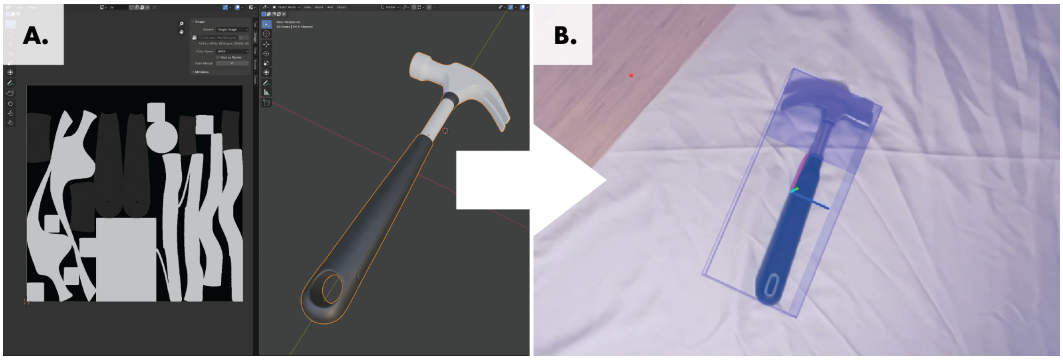


Figure 28: **A.**) Reference model for FoundationPose, scaled and textured in Blender, and based off of [94] **B.**) Bounding-box superimposed through ARKit for clarity.

Once a frame has been chosen, the image is unwarped based on the known intrinsics for either the iPad Pro or the Pixel 7 Pro. This step corrects lens distortion so that subsequent bounding-box coordinates align properly with the 2D image.

FoundationPose returns four bottom-corner coordinates of the hammer's bounding box; these correspond to the plane on which the hammer rests. We know the real width and height of the hammer, so OpenCV's perspective transform can warp those four corners to an idealized rectangle matching that size. After this homography, the entire scene appears as a flattened top-down isometric view scaled to the hammer's known dimensions. Because the hammer is consistently black and silver, it stands out against the backdrop-especially given our high camera angle.

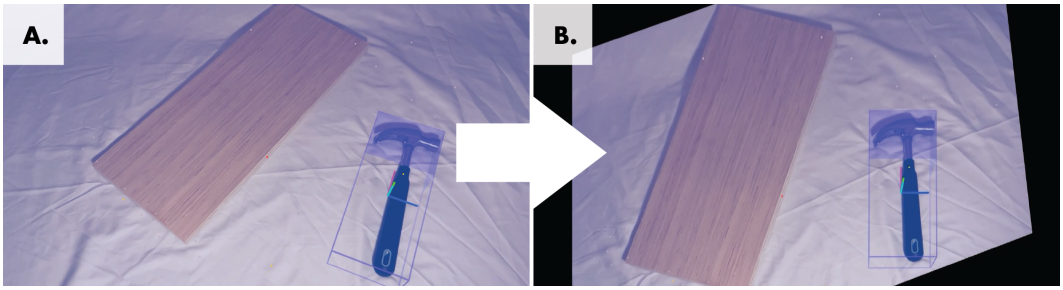


Figure 29: **A.**) Raw output before OpenCV perspective warp. **B.**) Parallel-perspective warped output.

With the warped frame, we easily identify and mask out the hammer by its distinct grayscale and known geometry. There's little risk of overlap between the hammer and the scrap wood because each video is shot from above at a steep angle, ensuring the hammer does not visually intersect with the wood piece. Eliminating the hammer leaves a high-contrast view of the scrap wood's edges against the background.

Finally, Potrace [95] is run on the masked image to produce a closed vector outline of the scrap wood. Because the wood is always photographed against a contrasting background, even small color variations can be customly thresholded effectively. Potrace then converts the silhouette into a polyline vector, giving us a simplified 2D outline that relatively reflects the real-world dimensions established by the hammer's reference measurements. This step does limit the pipeline to scenes where the backdrop is significantly different from the wood itself, but it seems sufficient as a proof-of-concept.

Overall, this pipeline successfully demonstrates the feasibility of extracting a scale-accurate outline of an unknown scrap piece from a single 2D frame. As a conceptual prototype, it meets the objective of dimensioning leftover wood for reuse (see also Sections 5.4 and 5.7), but it is far from a turnkey solution. Refinements in depth sensing, less restrictive backgrounds, or improved classification and masking techniques could make it more robust if future iterations are created.

To complete the dimensional profile of the scrap piece beyond its 2D outline, its thickness must also be gathered. This is achieved through a simple user-provided measurement, leveraging the device's calibrated screen (Figure 30). Users are prompted to align their screen with the material's depth and adjust two guides to match its edges. The system then converts the distance from this alignment, established during on-boarding (Section 5.3), into a physical dimension. One early concession of this style of measurement system is that it is only particularly applicable for flat, unwarped stock. For simplicity the measurements are currently rounded to the nearest quarter-inch, a high enough resolution for the common stock found around Carnegie Mellon's wood shops and fabrication labs.

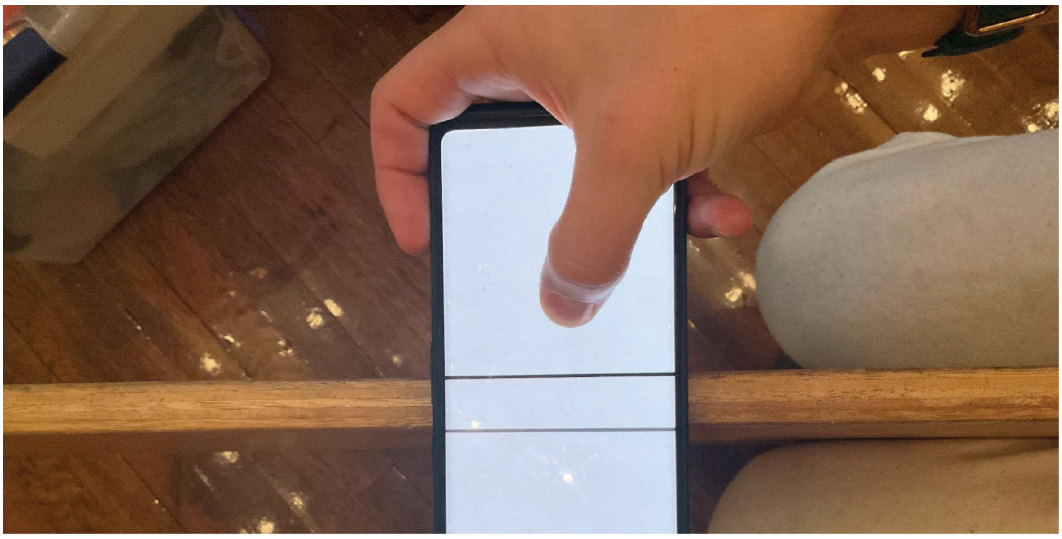


Figure 30: Process of aligning two lines on an example device to signify the thickness of the stock.

5.5 Main User Journey

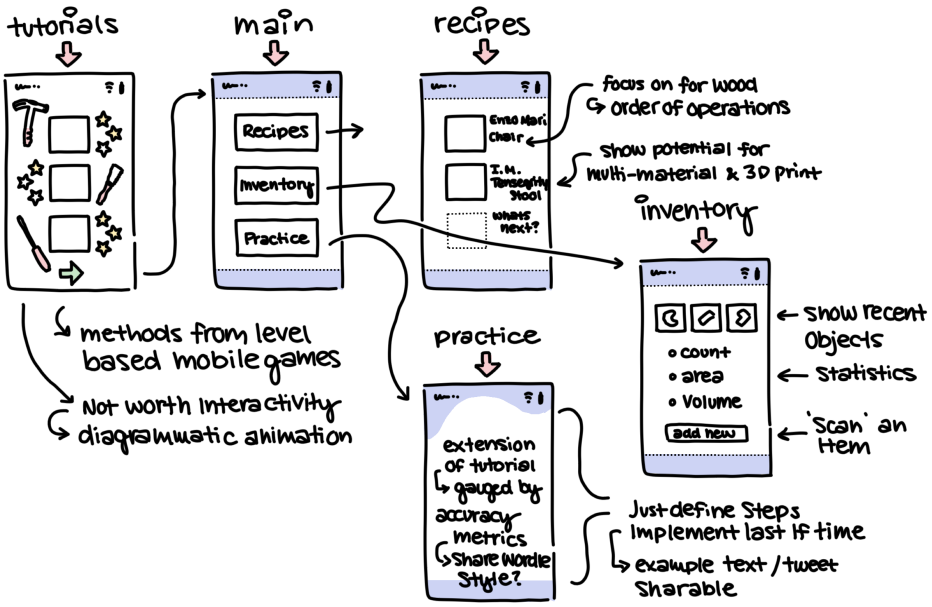


Figure 31: Sketch version of the user's journey, showing direct UI connections between elements on a screen.

5.6 Home Screen

Once users complete the onboarding steps and calibrations (as described in Section 5.3), they brought to the Home Screen, which serves as the proverbial central hub for all ScrapAR functionalities. This is where the Unity scenes can be started: viewing a recipe, viewing your inventory, practicing your skills, as well as accessing general options and links for more information (see Figure 31). The goal is to minimize confusion, enabling new or unfamiliar users to navigate the process with ease.

Stands

It became quickly apparent that holding the device was getting in the way of otherwise accessing the camera and working on the pieces – thus the following stands were designed and tested as temporary solutions to this issue.



Figure 32: Prototype view of the main application screen – some elements are implemented in SwiftUI, and based on an Illustrator layout. Some placeholder information and buttons.

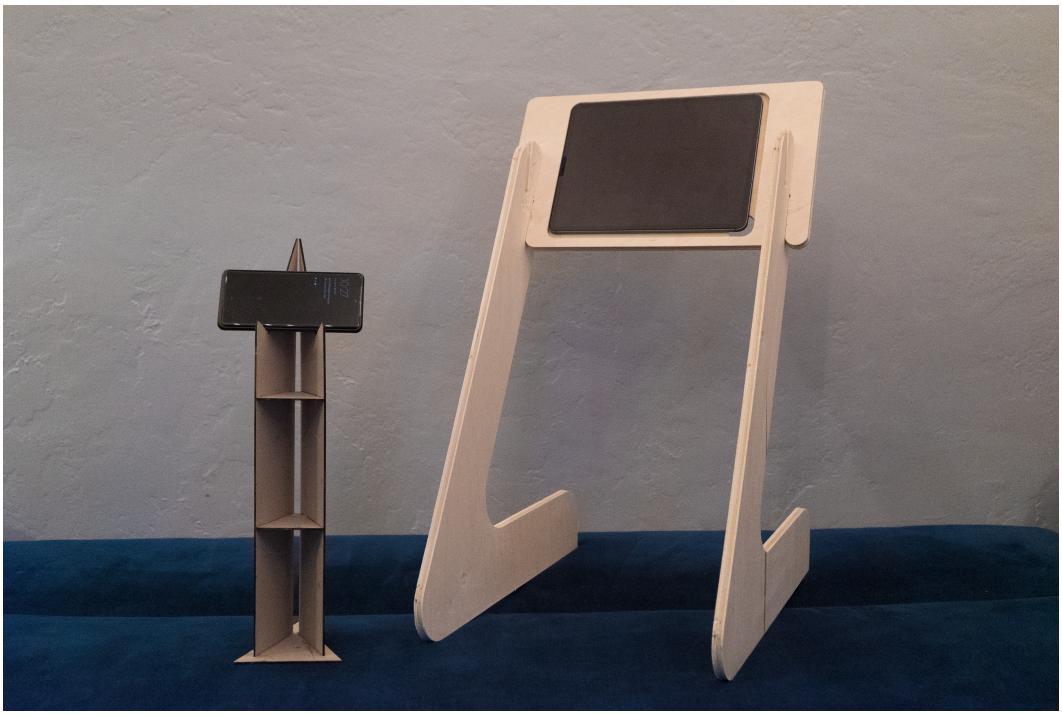


Figure 33: **Left**) A laser cut chipboard stand for a smartphone. **Right**) A CNC plywood holding an iPad Pro at different angles for different uses.

Practice Your Skills Another feature on the Home Screen is the "Practice Your Skills" mode, where users can refine specific hand-tool techniques outside of a larger, multi-step project. One of the earliest mini-games implemented here is a "straight-line cut" exercise, which prompts the user to cut a piece of scrap wood in half with a hand saw and then take a top-down photo of the resulting edge (see Figure 34). Instead of handling analysis entirely within a mobile environment (like an iOS app), this prototype was built on a Windows-based workflow: the picture is captured and then sent to a GPT-4 API call. The model is given a targeted prompt-asking for a 0–100 rating on the cut's straightness and quality, plus three bullet points of praise and three points for improvement-and returns a short text evaluation that is output in Figure 35, and could easily appear in a mini-game interface.

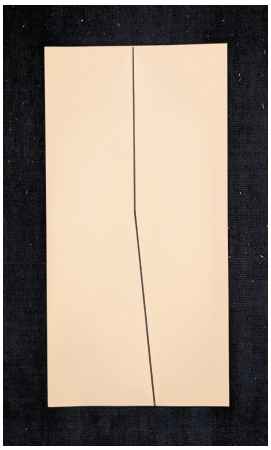


Figure 34: Example input for the MM-LLM showing a hypothetical cut piece of scrap.

By using a Windows-based approach, this prototype can accept a file upload or drag-and-drop submission from the user's desktop, which is then forwarded to the GPT-4 API. Once the server processes the image and prompt, it delivers a concise JSON or text-based response back to the application. This approach avoids the complexities of integrating GPT-4 directly into an Xcode environment or bundling large language model dependencies into a mobile build. Instead, the Windows prototype focuses on the upload and retrieval mechanisms, ensuring that communication with the server remains smooth and that results arrive promptly for the user's immediate feedback.

This mini-game supports the overarching ScrapAR goal-outlined in Sections 5.4 and 5.7-of lowering the barrier to woodworking by providing frequent, bite-sized opportunities for skill development. Short feedback loops encourage beginners to stay motivated and learn by doing, rather than waiting for in-person instruction or more formal evaluations. Moreover, by having GPT-4 provide a simple numerical score alongside specific tips, the system offers an easy entry point into self-assessment. Users can attempt another cut right away, steadily improving their technique through guided repetition and tying neatly into the project's broader aim of fostering confidence and craft literacy.

```
C:\Windows\System32\cmd.e
Microsoft Windows [Version 10.0.26100.3476]
(c) Microsoft Corporation. All rights reserved.

E:\Other computers\New Legion\Desktop\Thesis>more OpenAI_Output.txt
Rating: 70/100

Good aspects of the cut:

        Relatively straight overall, showing good initial control.

        Clean edges without significant splintering or chipping.

        Minimal deviation, indicating steady saw technique.

Areas to work on:

        Slight angle noticeable in the middle; focus on consistent saw alignment throughout the cut.

        Improve stability of the workpiece by securing it firmly to reduce wandering of the saw blade.

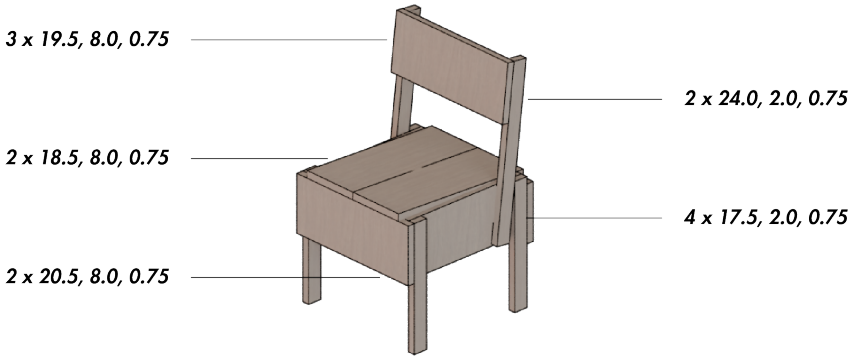
        Aim for consistent pressure and pace to eliminate minor variations in the cutting line.
```

Figure 35: Example output from OpenAI's GPT-4

5.7 Making a Recipe

Overview & The Chair

To demonstrate the complete ScrapAR workflow, this study culminates in the guided construction of Enzo Mari's 'Sedia 1' chair, a seminal piece from his 1974 *Autoprogettazione* project. This collection was conceived as an act of design democratization, providing open plans for furniture built from affordable, standard lumber using only the most basic tools – a saw, hammer, and nails. Mari's explicit goal was pedagogical: to empower users by demystifying the construction process through hands-on making. This spirit aligns perfectly with this project's mission to lower barriers and enhance accessibility in woodworking education. The Sedia 1 is particularly apt – its simple, rectangular components are well suited for nesting onto likely rectangular offcuts and demand only the fundamental skills covered in Sections 17–19. Its frequent use in introductory woodworking courses, including the author's own first experience building furniture, further validates its selection as an ideal test case for novice users guided by this system.



Enzo Mari – Sedia 1 Chair Autoprogettazione – Plans by Ian Anderson [96].

Each recipe includes:

- **Component List**: Required parts with adjustable dimensions.
- **Assembly Instructions**: Step-by-step guidance.
- **Material Compatibility**: Recommended wood thicknesses and rough dimensions.

Nesting Parts

To begin, outlines of each piece of scrap are imported into Rhino, allowing each required part to be nested within the the user's available materials. The system is run with a grasshopper script utilizing two open-source methods for nesting, including OpenNest [19] and a meta-heuristic solver [59] for rectangular forms, with custom parameters chosen to best minimize waste. The final instructions are then created in both vector and mesh formats – allowing for both static diagrams and dynamic 3D objects ready for AR overlay.

Once the nesting result is created, the user can begin marking stock. Here, the flattened pattern is mapped onto the glTF created in Section 5.4 using the silhouette and thickness of the real-world raw stock, allowing it to be tracked in a similar method to that outlined in Figure 27. When viewing through the development laptop's display, the scene continuously tracks the overlay so that cut

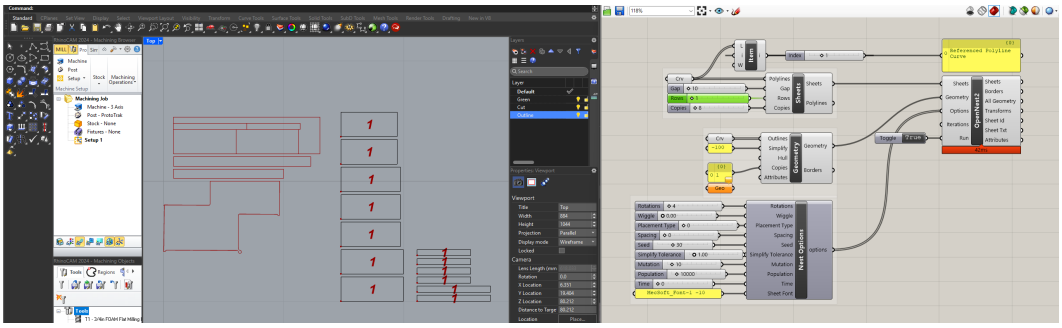


Figure 36: Rhino view of scrap outlines and parts to be nested, leaving a buffer around the objects for kerf, warping errors along the edges, and maximizing overlapping cuts to reduce cutting time.

lines cling to the live view, allowing for both the camera and the object to move independently and compensating for uncertainty. To guard against parallax and occlusion, the model is tracked with allowances for perspective warp – similarly, the training of the tracking model is intentionally untextured to create a color agnostic key-point setup, allowing for variation in the materiality of the physical stock while still maintaining tracking. The user is prompted to trace the projected outlines with a pencil, and utilize the paper speed-square created in Section 17, rather than attempting to saw parts while using the AR overlay. This two-step workflow keeps the user in full control of the blade, as well as better preparing them for future woodworking in a more traditional measure-cut methodology.

Cutting Parts

When it is time to cut, the user is encouraged to set the phone or tablet down on a table or the stand shown in Figure 33 and focus on the saw (see Figure 37). The screen can now be set to show a top-down view of the stock pieces and with outlines, acting as a digital cutting key that can be referred to without having to juggle the device in one hand and the saw in the other. Since the demo chair has several identical parts, there is opportunity here for the overlay to show multiple colors, each duplicate in the same hue (e.g. all leg blanks in blue, all seat-slat blanks in orange). Once the first copy of a duplicated piece is cut, the user can lay it on top of the remaining blanks and trace around it – the matching colors confirms they are pairing the correct shapes, helping ensure every copy follows the same transferred sketch lines before any further cuts begin. This particular proof-of-concept chair was partially chosen as the physical dimensions of individual pieces is less important than that duplicate pieces more closely match.

Final Assembly

With the individual parts cut, optionally sanded, and double checked against the provided parts list, the workflow shifts from 2D layout guidance to a live 3D assembly overlay streamed from a Unity 2022 project to the user's monitor or tablet. Each construction stage is packaged as its own Unity Scene. When a scene loads, the partial model the user has just completed is locked to the tracked object, a translucent "ghost" of the next piece is shown in a red overlay, and a small annotation shows where to drive nails. At any time the user can also be shown a recommended orientation of the current part to most easily attach the next piece. In future iterations, the user would press either the Prev or Next UI buttons to jump between steps, and a general progress bar is referenced at the top of the screen to show where they are within the entire assembly sequence (Figure 38) – currently the prototype can be manually swapped between stages using the Unity editor.

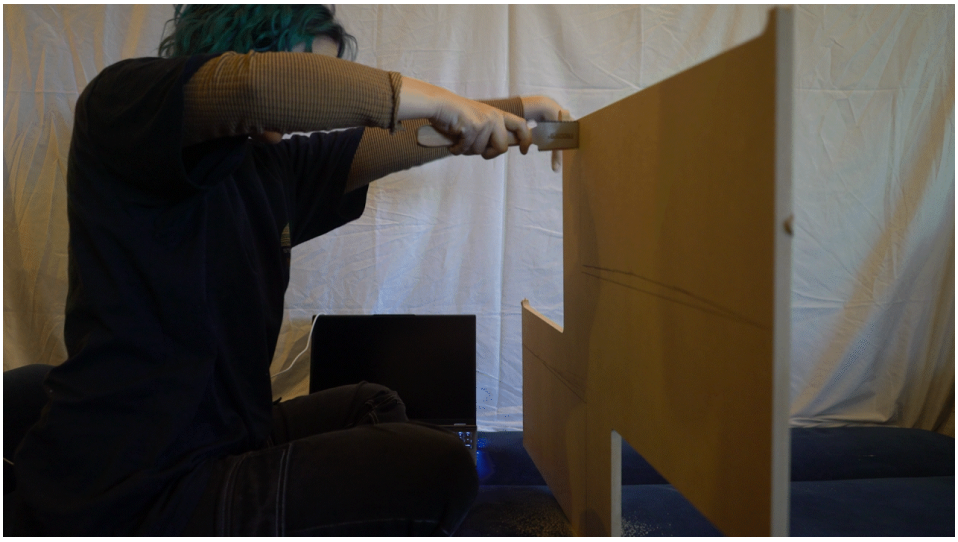


Figure 37: The slightly monotonous process of cutting each piece of the chair by handsaw.

Alignment in each scene is built upon a Vuforia model target generated with the Model Target Generator (MTG). Each intermediate assembly state was exported from Blender as a glTF and OBJ, with the untextured OBJ version being trained in MTG, and bundled into a lightweight database that is referenced into the Unity project. At runtime, ARFoundation swaps in the correct model from the database for the current step, Vuforia acquires the object as soon as its keypoints reach a set confidence threshold, and ARFoundation's object-manager hands back a continuous 6-DoF pose. To stop the overlay from "snapping" off the model whenever the piece is rotated or struck with a hammer, the raw pose stream is smoothed and the user is recommended to center the part within a large field of view to best maintain continuous tracking – the iPad stand allows for this to occur simultaneously while holding and manipulating the parts.

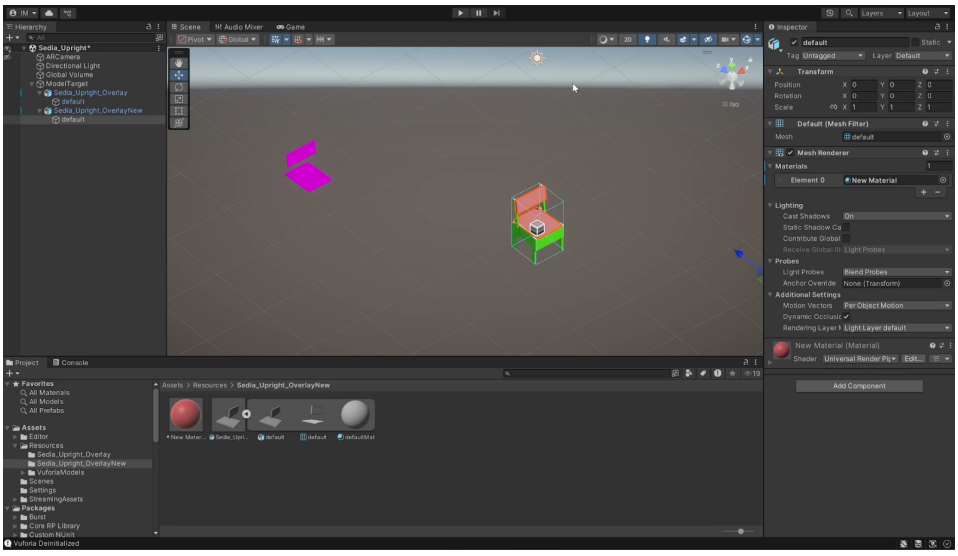


Figure 38: Scene outline for the Unity/Vuforia step-wise assembly overlays.

The build sequence itself is deliberately incremental. First the user aligns two of the side legs and the lower cross-rail on a work surface; when the algorithm detects that those three members sit within roughly within the expected pose, the overlay applies a semi-transparent green hue to the entire object. The rear back rest comes next (see Figure 39), followed by recreating the entire assembly mirrored. Once the two halves are combined with the front bridge, the growing geometry starts to self-occlude, so the latter steps rely on a less opaque overlay so the user can peek behind the part being added to see the context of the larger assembly. In the final "finishing check" scene the overall chair model is overlaid – in future iterations, there could be an extension of this system so that any vertex that deviates by more than a defined amount flashes red, flagging joints that need adjustment. At any time in the construction sequence, the user can also toggle a reference-only mode that shows a rendered version of the same model for reference.

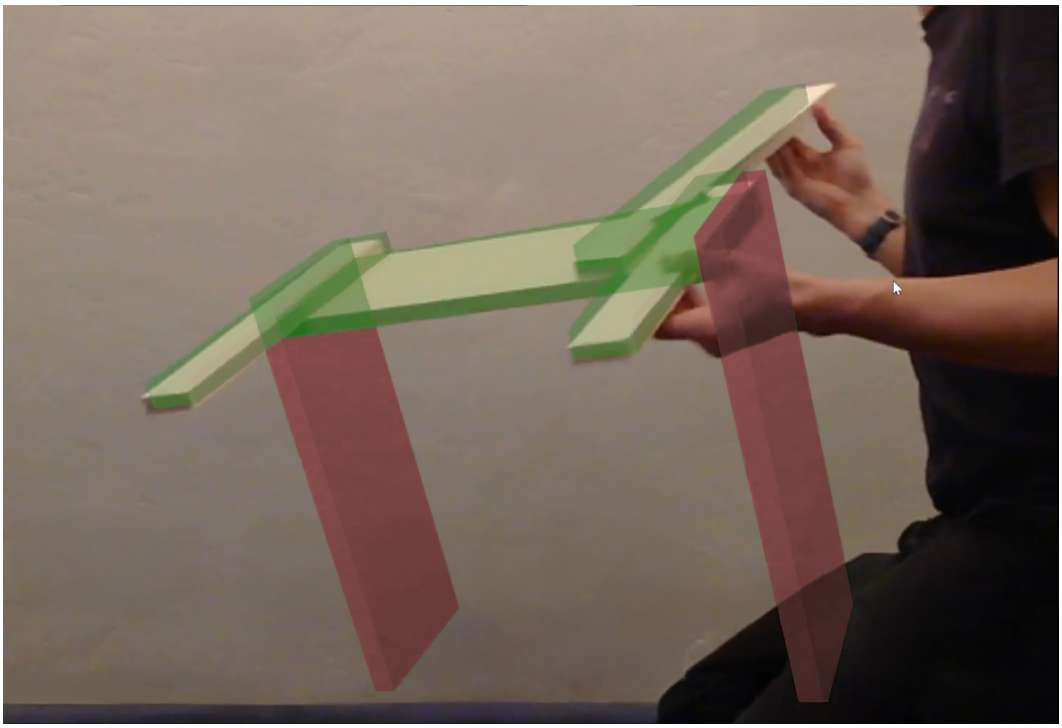


Figure 39: Live overlay during an early assembly step. Completed parts are shown tracked to the object, and the next piece(s) to add appear in translucent red.

6 Results

Following the detailed account of the system's design and technical implementation, this section presents the empirical findings through its application in a practical setting. It focuses on evaluating the system's efficacy and user experience, primarily through a user study involving a non-woodworker tasked with constructing an intentionally chosen piece of furniture. The subsections that follow describe the parameters of this study, narrate the user's journey and the tangible outcomes of their engagement with the workflow, and offer an evaluation based on direct feedback and systematic observations. These results are intended to provide critical insights into the system's usability, its impact on the learning process, and its potential to achieve the thesis's core objectives.

6.1 Study Parameters

To gauge whether the technical vignettes developed in Sections 5.4 to 5.7 hold up in novice hands, a single "cold" tester – Maddie, a post-graduate psychology student with no shop experience – was chosen to run through the majority of the pipeline: on-boarding, part nesting, AR stock marking, hand-saw cutting, and the step-wise assembly overlay. Maddie was chosen precisely because she represents the kind of learner the system hopes to empower: curious and active in crafting communities but wary of woodworking, ample access to scrap materials, and well versed in screen-based methods of learning without a desire for traditional apprenticeship. The session lasted an afternoon ($\tilde{2}$h 40m) in a non-workshop residential environment with the same iPad Pro (5th-gen) and Lenovo laptop (16ACHg6) used during the system development. Aside from mandatory safety oversight, no coaching was offered beyond what the application itself presented – but support was provided when launching different Unity scenes and debugging unexpected technical issues throughout the process.

6.2 Study Outcome

The complete user session was screen-captured and recorded with a nearby Sony A7s II camera to capture gestures, audio, and the general beats of the session to later be time-coded. What follows alternates narrative text with a images pulled from the footage.

Maddie began by working through the tool tutorials. She reviewed the saw tutorial a few times as it seemed to be where the majority of the apprehension was involved. She also practiced the general process of hammering on a piece of scrap materials, mentioning that it wasn't the most intuitive motion. She mentioned that she is more familiar with a drill and had only used a hammer on a select few occasions.

Inventory capture followed. While the fully-automated version of this system is not fully implemented, in fifteen minutes we walked through the scanning process for three pieces of scrap gathered from previous projects and the dumpster at CMU's digital-fabrication lab, ranging from a roughly 4' x 4' piece of Revolution 0.75" and 0.5" plywood to 3" linear elements of Hempwood leftover from previous research. While running the nesting algorithm on the rough outlines, she remarked that the process seemed a bit rough.

Once the outlines were saved and the nesting overlays were created and provided to Unity, she began the tracing process. During the marking stage she went back and forth between the tracked overlay (see Figure 40 A. & B.) to start each line, and placing the device on the provided stand to extend the lines with the paper square. The first piece of stock seemed to take the longest, especially in getting

familiar with moving between the overlay and using two hands to complete the projected cut lines in pencil. Duplicated outlines appeared in corresponding tones on the overlay, and – after a quick suggestion from the author – she realized that they could be visually compared to ensure that their general size was comparable.

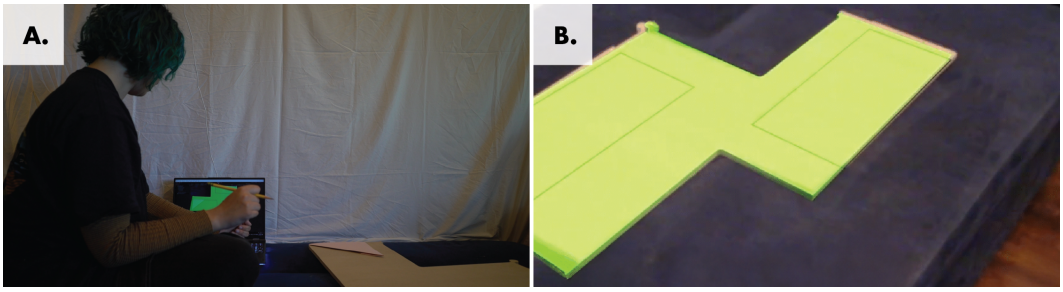


Figure 40: A.) User viewing live overlay atop stock material. B.) Raw generated overlay in green from the overhead locking onto the first stock material.

After tracing, cutting took around half an hour. Maddie placed the tablet and laptop to the side of the makeshift work surface so she could "concentrate without worrying about it." Although a bit nervous at first, she seemed to take up cutting pretty well, with a notable increase in accuracy and ease as the process continued. Some of the first cuts drifted a bit from her cut-lines, but it seemed like she started to better understand the general motion, the impact of forcing the saw into the stock, and how small adjustments in the angle of the blade impacted the overall cut throughout the process. Like before, she realized that the cut pieces could be used with her lines to ensure that they were similarly sized.



Figure 41: First parts laid out without an overlay due to issues tracking the symmetrical, low-keypointed profiles.

Assembly proceeded through the six Unity scenes described in Section 5.7. To start the first scene, she referred to the 3D model of the initial pieces, allowing for a complex enough object to be tracked in further steps (see Figure 41). As the most complex task, keeping tracking on the chair while working on it was difficult, but not impossible. While it seemed easy for the ARFoundation scenes to lose tracking, it was quite easy to gain it again by blocking the lens and allowing it to regain a lock before continuing. As the chair grew, it also became apparent that attempting to work on it from behind the screen wasn't possible, and Maddie began referencing the object through the screen before moving it to a more comfortable position before adding the next element – working in conjunction with the rendered 3D views to understand the ideal assembly orientation (see Figure 42).

Total elapsed build time – including the onboard and 'scans' – was 2h 38min, comfortably inside the anticipated three-hour target. The final object is notably visually similar to the control version shown in Figure 44, and seemed to provide a qualitative level of fulfillment while watching the process unfold (see Figure 43 B.). Throughout the process there were several issues with tracking and awkward points waiting for scenes to be updated, both (positively) creating moments for Maddie's self reflection as well as (negatively) lowering her momentum from completing the process.

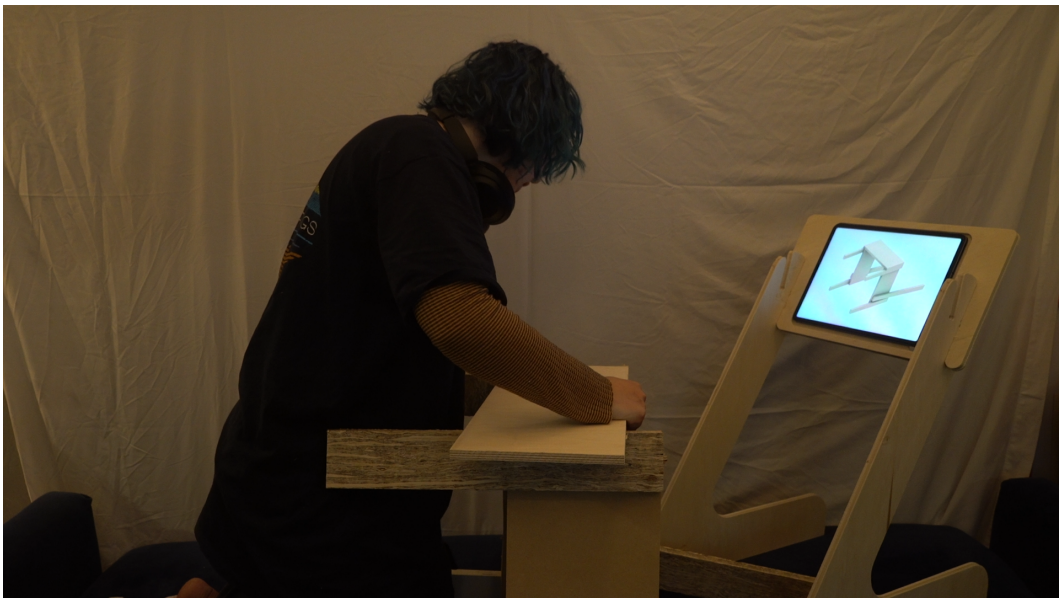


Figure 42: Reference screen consulted during assembly with 3D model of the current step in assembly.



Figure 43: **A.**) Final assembly of last cut parts. **B.**) Fully completed object.

Informal debrief notes captured immediately afterward show that Maddie's two pre-study concerns shifted markedly. She still worried about interpreting instructions "if things got really complicated," but rated overall safety and ease of use as "much better than expected," citing the cutting overlay and the tracked 3D instructions as the two most useful features. Those qualitative impressions are echoed in the post-study questionnaire analyzed in the next subsection.

6.3 Method of Evaluation

The original versions of this questionnaire can be found in Appendix 7.4

Pre-Study Questionnaire Immediately before touching any tools Maddie filled out a 10-item pre-study questionnaire.

Each item targeted one of three constructs:

- Finding the baseline familiarity with woodworking and tools.



Figure 44: **A.**) CNC routed control chair – also created from scrap stock. **B.**) Maddie's hand cut chair.

- Understanding preexisting comfort with AR and computer-based guidance.
- Understanding their attitude toward sustainable or waste-reducing practice.

The questionnaire mixed open-ended prompts (e.g. "Describe your past experience with hands-on projects") with 5-point Likert scales for confidence or familiarity. A similar post-study questionnaire – plus ten reflection items about the ScrapAR workflow – was administered after the chair was complete. Responses were captured digitally – without the author present to minimize the chance of influencing any aspect of the questionnaire.

Pre-study Baseline

The original versions of this questionnaire can be found in Appendix 7.4.

Maddie reported only fragmentary exposure to making: she has "minimal experience" beyond assembling IKEA furniture "under supervision" and once took a theater-set class in which she "stripped screws and broke several drills" (the author is unsure exactly how this is possible). She rated her confidence with basic hand tools at 2.5–3/5, squarely in the "somewhat unsure" range. She had never used AR or VR for learning and placed her familiarity with sustainable or waste-reduction practices at a flat 1/5. Asked how she typically learns hands-on skills, she preferred either formal classes or one-on-one mentorship over informal online videos. Her self-rated comfort with accurate measuring and cutting was again low (2/5). She had never repurposed scrap wood, citing "getting the materials" and "knowing what to do" as the two biggest barriers to starting a woodworking project. From ScrapAR she hoped mainly "to become more confident in the basics," while voicing two concerns: first, that directions might be hard to interpret without a teacher on hand, and second, that safety could be compromised if she misunderstood an AR prompt.

Taken together, these answers confirm the target-user profile the system was designed for: low domain knowledge, low tool confidence, sustainability in name only, but a strong motivation to learn if clear,

trustworthy instruction can be provided. They also supply a benchmark against which the following post-study reflections can be compared.

Post-Study Questionnaire

Comparing Maddie's self-ratings before and after the exercise, the biggest jump was her confidence in measuring and cutting: from 2/5 to 4/5. She attributed the gain to "seeing exactly how to cut the wood out of the larger slabs" and in an overall jump in comfort based on her gaining experience with the hand-saw. Her sense of assembly competence rose more modestly (2 to 3), and she remarked that without on-screen sequencing she "would not feel comfortable joining the pieces unaided."

Asked to rate overall ease of use, Maddie chose "three – [cutting] angles can be difficult." Her only recurrent friction point was finding a camera position that satisfied both the model-target lock and fitting comfortably within the limited amount of space during the demo. Nevertheless she declared herself "very likely" (4 / 5) to use an AR guide again, especially if given more movement freedom without needing to stay within the narrow view of the trial camera.

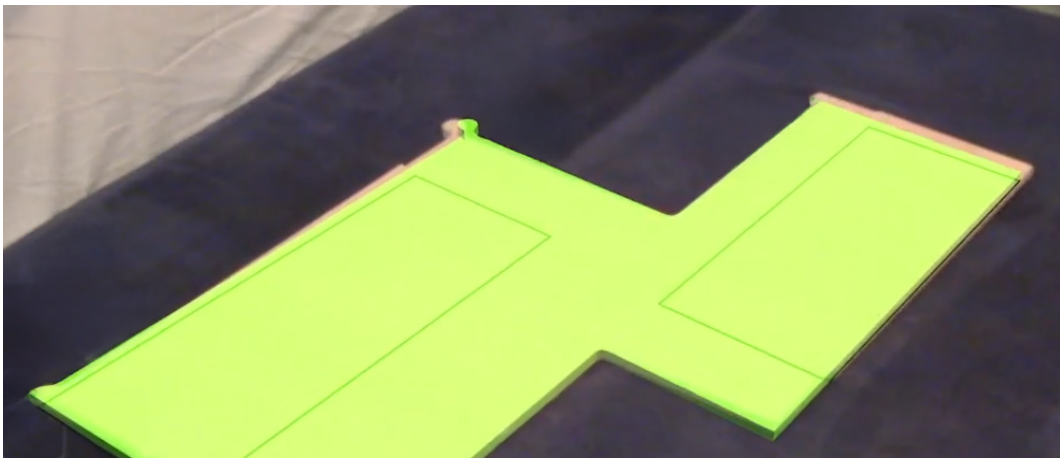


Figure 45: Example of the slightly askew marking due to the optics of the marking overlay, illustrating the "angle difficulty" comment.

In open-ended feedback she singled out the overlaid measurements as "the most helpful thing," noting that the fixed top-down view let her "double-check a line before committing the saw," which in turn reduced her anxiety about wasting stock. The idea of having real-time feedback during sawing wasn't mentioned and made the safety-forward decision seem correct.

In terms of whether the system helped her sustainability goals, Maddie answered "Yes! Lots of scrap wood was used so less waste was created," confirming that, at a minimum, the system made the reuse value of odd off-cuts visible and their reclamation viable (see Figure 46 A.). Although objective waste metrics were not captured for this single session, no additional lumber was needed: every part used in the chair was pulled from the inventoried scraps taken from Carnegie Mellon's Digital Fabrication Lab dumpster (see Figure 46 B.).

Throughout the process tracking rarely failed outright, but the overlays sometimes "had to be placed at specific angles" in order to properly identify the current step. If unexpected warping occurred, it was trivial to restart the Unity scene to quickly re-stabilize the stream in only a few seconds. Lighting in the residential environment proved adequate, and no particular hardware-specific issues were observed over the <3 hour run.

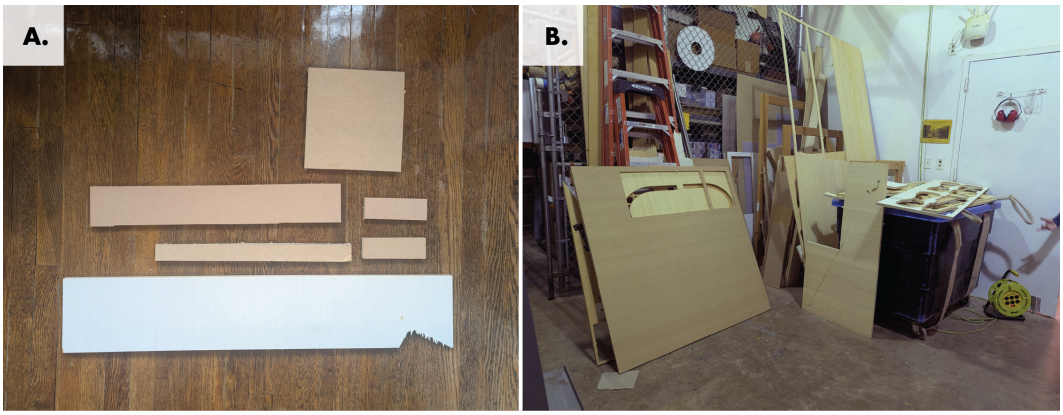


Figure 46: **A.**) Some of the remaining pieces from the original scraps nested and cut. **B.**) The dFAB scrap pile and trash bin while gathering scraps – in addition to other scraps pulled from the nearby dumpster.

Maddie reported that the work was "more of a workout" than originally anticipated and felt that the overall safety of the process wasn't as risky as she had originally thought. No near-miss incidents occurred, and the recommendations given at the beginning of the process seemed to be taken to heart, being continually referenced throughout the test.

7 Conclusion

This thesis set out to investigate how contemporary advancements in CV and AR could be integrated into a succinct workflow to supplement beginner woodworking education, enhance accessibility, and promote the creative reuse of scrap materials. From the beginning, the core research question defining this approach was simply: *How can AR and the growth of CV provide a method of craft education that works to reduce waste and enhance the accessibility of woodcrafts?* Due to the vagueness of this question, and the freeform exploration of this topic, many different avenues were explored to work toward an answer that could be adapted beyond a simple project (see Figure 47). The development process, documented through ethnographic insights, technical vignettes, and pipeline implementation, culminated in a user study where a novice successfully constructed an Enzo Mari 'Sedia 1' chair using the prototype system and locally reclaimed materials. The results presented in Section 6 offer preliminary validation of the proposed approach, but also highlight areas and holdups requiring further development.

7.1 Interpretation of Findings

The successful completion of the chair project by a user with very-minimal prior woodworking experience (Section 6.2) directly addresses the central research question. It demonstrates that the synergistic combination of CV-assisted inventorying (even in its current prototypical state, Section 5.4), AR-guided marking and assembly (Section 5.7), and structured tutorials (Section 5.3) can effectively scaffold the learning process for a novice – providing a chance for an unfamiliar party to attempt and otherwise unfamiliar process. The participant, Maddie, explicitly cited the AR overlays for measurement and assembly sequencing as the most helpful features, directly aligning with the project's goal of mitigating common barriers like measurement anxiety and procedural uncertainty, which were also echoed in the initial ethnographic interviews (Sections 4.1–4.1).



Figure 47: Final project exhibition including the two Sedia chairs, prototype holders, and workflow poster. Part of Carnegie Mellon University's currently-under-construction Rockwell Park exhibition hall.

Maddie's reported increase in confidence, particularly in measuring and cutting (from 2/5 to 4/5, Section 6.3), suggests that ScrapAR can foster a sense of competence and potentially lower the affective filter often associated with acquiring complex manual skills. This aligns with research on gamified learning and AR in education, which often emphasizes engagement and confidence-building [17, 25]. The system provided immediate, actionable guidance (e.g., overlaying cut lines, showing assembly steps) that compensated for the lack of tacit knowledge typically gained through apprenticeship [12]. Furthermore, the successful use of scrap materials gathered from campus waste streams directly supports the objective of promoting sustainable practices and demonstrating a tangible pathway for material reuse [56], addressing the participant's low initial familiarity with waste reduction (1/5).

In addition, the multi-modal feedback prototype using GPT-4 for cut analysis (Section 5.6) also points towards novel ways to provide asynchronous critique, potentially mimicking aspects of instructor feedback without requiring their physical presence. While rudimentary, and acting as a more exploratory component of this thesis, it hints at how the continued human-infatuation with ML could further personalize and scale craft education in unexpected ways.

However, the thesis – and more specifically the user study in general – also revealed nuances. While the AR overlays were valued, ponts of interaction friction arose, particularly concerning camera positioning and general tracking stability (Sections 6.2, 6.3). The slight "angle difficulty" noted by the participant highlights the challenges of translating precise 2D digital instructions onto 3D physical objects via a single camera view, suggesting that calibration and perspective correction remain critical areas for refinement – not to mention the advantages of stereo and depth capable setups. The reliance on specific hardware (iPad Pro & Lenovo laptop) and software (Unity, Vuforia, & open-source CV libraries) also underscores the persistent tension between leveraging advanced capabilities and ensuring broad accessibility.

7.2 Contributions

Despite its prototypical nature, the ScrapAR workflow offers several contributions to the fields of computational design, craft education, sustainable practices, and applied human-computer interaction. Primarily, ScrapAR presents a novel, integrated pipeline concept specifically designed to democratize traditional craft for novice users. This end-to-end system intentionally connects material digitization of scrap (Section 5.4), computational planning through nesting (Section 5.7), structured skill development via initial tutorials (Section 5.3), and guided fabrication using AR overlays for both marking and assembly (Section 5.7). While individual elements like AR assembly guides [7] or CV measurement techniques [97] exist, their synthesis within ScrapAR – explicitly aimed at empowering beginners to repurpose scrap materials – represents a unique application focus.

A key technical contribution lies in its advancement of applied CV for irregular material assessment using openly accessible tools. The exploration in Section 5.4, particularly the proof-of-concept pipeline combining FoundationPose with a known reference object (the TRIXIG hammer) for scale, DepthPro for depth approximation, and Potrace for vector outline extraction from a single image, demonstrates a tangible approach to digitizing and dimensioning unstructured scrap wood. This offers a practical pathway for understanding and valuing otherwise discarded materials without requiring specialized scanning hardware.

Furthermore, the project makes a distinct contribution through its exploration of AR for full-process fabrication guidance and potential error reduction. The AR overlays for projecting cut lines directly onto physical stock and providing step-by-step 3D assembly instructions (Section 5.7) illustrate how AR can deliver situated, real-time support. This direct visual guidance can mitigate common novice errors in measurement and spatial understanding, moving beyond the limitations of static 2D plans and enhancing the physical making process, as reflected in the user study (Section 6).

The system's design was consciously user-centered, grounded in direct pedagogical needs identified through ethnographic insights from experienced woodworking educators (Holmes, Sontag, Hritz, as detailed in Sections 4.1–4.1). This foundational research helped ensure that developed features, such as the animated tutorials and clear visual prompts, directly addressed real-world teaching challenges, including limited instructor availability, student knowledge retention, and the initial intimidation factor often associated with craft.

Crucially, the workflow offers a practical framework for material reuse and enhances sustainable design practices through digital enablement. ScrapAR explicitly designs for scrap utilization, and by providing computational tools like nesting optimization (Section 5.7), it makes the process of transforming "waste" into valuable objects more systematic and achievable for individuals. This offers a tangible model for applying computational design to support circular economy principles within craft. The user study (Section 6) provided encouraging evidence that such a system can effectively foster confidence and reduce specific barriers for novices. The participant, Maddie, reported a significant increase in her confidence for core tasks like measuring and cutting, attributing this to the clarity of the AR guidance. This directly addressed her pre-study concerns and demonstrates the system's potential to improve not just skill but also the learner's affective experience and motivation for further engagement.

Finally, viewed more broadly, ScrapAR contributes a prototype pedagogical framework for integrating digital tools into traditional craft education. It showcases how various digital interventions – CV for material understanding, AR for process guidance, pre-defined tutorials for foundational skills, and

even emerging AI like LLMs for practice feedback (Section 5.6) – can be thoughtfully combined to supplement and enrich, rather than replace, hands-on learning. This offers a replicable model for educators in other craft domains seeking to modernize curricula and extend learning opportunities. These contributions, taken together, underscore the potential of computationally-assisted workflows like ScrapAR to make craft more approachable, sustainable, and engaging for a new generation of makers.

7.3 Limitations

While promising, the current iteration of ScrapAR is subject to several limitations that temper its immediate applicability and highlight critical areas for future work. The accuracy and robustness of the CV pipeline remain a significant challenge, as detailed in Section 5.4. The performance of the monocular dimensioning and pose estimation is highly variable, influenced by environmental factors like lighting and background contrast, as well as object properties such as texture, making the system sensitive and currently reliant on controlled conditions. This is compounded by the computational cost of the underlying models (e.g., FoundationPose, DepthPro), which necessitated a capable laptop in the prototype phase and may limit deployment on less powerful mobile devices. Additionally, the CV pipeline is not yet fully integrated, requiring manual steps that disrupt a seamless user experience, and the manual method for thickness measurement, while intuitive, has inherent accuracy limits and assumes idealized material flatness.

AR interaction and ergonomics present another set of challenges. Although Vuforia provided functional tracking, stability issues arose, particularly during forceful actions or when view angles were suboptimal, interrupting the user's workflow (Sections 5.7, 6.2). The physical ergonomics of simultaneously handling tools, materials, and the AR device also proved awkward, indicating a need for better device mounting solutions or alternative interaction paradigms beyond direct screen manipulation or perhaps beyond held devices all together. Moreover, while designed for ease of use, the AR interface still possesses a learning curve, requiring users to familiarize themselves with interpreting and interacting with the 3D overlays (Section 6.2).

The scope of the current system is also intentionally constrained. It focuses on a limited toolset (hand saw, hammer) and basic butt-joinery, omitting power tools, complex techniques, finishing processes, and comprehensive safety instruction beyond the initial tutorials. The workflow's effectiveness was primarily demonstrated through a single project (Sedia 1), and its applicability to more intricate designs or different material types requires further investigation. The system implicitly assumes relatively well-behaved scrap materials, without specific mechanisms to handle severely warped, knotted, or fiber-based wood products.

Finally, the evaluation itself was limited. The primary user study involved only one novice participant, restricting the generalization of the findings. The assessment relied heavily on qualitative feedback and self-reported measures, lacking systematic quantitative data on efficiency, waste reduction, or objective skill improvement compared to traditional methods. The dependence of the GPT-based feedback feature on cloud connectivity also presents a practical limitation in potentially offline workshop environments.

7.4 Future Research Directions

Despite these limitations, the ScrapAR project carries implications for practice, education, and theory. Practically, it serves as a blueprint for tools that could empower individuals in DIY contexts, maker spaces, or remote learning scenarios, potentially broadening participation in craft while promoting resourcefulness through material reuse. Educationally, it demonstrates how AR and CV can supplement traditional pedagogy, offering personalized, on-demand guidance suitable for self-paced learning or situations with limited instructor access. Theoretically, it contributes to human-computer-interaction research in fabrication, illustrating how computational systems can function as cognitive buoys – reducing the weight of complex physical tasks – and prompting further inquiry into digital mediation of tacit knowledge and skill acquisition.

Maddie's sole feature request was in supporting a wider camera or lens – she felt that being able to more easily fit objects into the view without needing to go back and forth to the tablet and laptop would solve several frustrations. She also expressed interest in a way to log cut accuracy and seemed interested in the long-term gamification goals of this project – specifically mentioning it would be cool to have progress levels and watch her score climb over multiple projects.

Building upon this foundation and addressing the identified challenges, future research should pursue several interconnected avenues, consistent with the directions noted above. A primary focus must be on improving the robustness and accessibility of the CV pipeline, perhaps by exploring more advanced or lightweight monocular depth and scale estimation techniques, integrating simple physical markers for reliable calibration, and fully consolidating the workflow into a seamless application. Concurrently, enhancing the AR interaction is crucial, involving investigation into more resilient tracking methods, potentially leveraging onboard sensors, and exploring alternative interaction modalities such as touch, voice, or gesture control to improve ergonomics. Expanding the system's capabilities is another key direction, requiring the development of a broader recipe library – or connecting to existing libraries, incorporating tutorials for additional tools and techniques, refining nesting, adding gamified customization features, and potentially integrating CV-based wood recognition or defect detection. To validate these developments, rigorous user studies involving diverse participant groups and employing both qualitative and quantitative metrics are essential to thoroughly assess usability, learning effectiveness, and real-world impact.

Ultimately, realizing the vision of a collaborative community platform where users can share scrap inventories, recipes, and modifications could significantly amplify the project's reach and foster a collective engagement with sustainable making. Furthermore, adapting the core ScrapAR concepts to other craft domains like textiles or cooking could demonstrate the broader applicability of this approach to computationally-assisted, resource-conscious fabrication. In essence, the path forward involves technological refinement, content expansion, rigorous evaluation, and community building to fully realize the potential of systems like ScrapAR in transforming how we learn, make, and reuse.

Acknowledgments

My sincere gratitude goes to my primary advisor, Professor Juney Lee, for his incredible support, patience, and willingness to provide detailed feedback throughout this thesis journey. It's been a wild year of anxiety and exploration, and I deeply appreciate his graciousness, especially in his commitment to working until the early hours on our exhibition setups. These are memories I will cherish, right alongside the distinct smell of routing HempWood.

I am also deeply indebted to Professor Daniel Cardoso Llach, who was the first faculty member I met and interviewed me for CMU. I am immensely grateful that he took a chance on me, offering me an RA position before we had even met in person. That opportunity was pivotal, leading directly to the explorations in this thesis, and I thank him for the significant time and mentorship he has generously provided over these past two years.

I extend my heartfelt thanks to Professor Vernelle Noel for her invaluable assistance on my committee. Your insightful perspectives, your willingness to acknowledge the work of others, and your remarkable ability to share knowledge while empowering individual agency have been truly inspiring and crucial to this process. These are lessons I will carry with me well into the future.

I am also grateful to Terry Hritz and Jon Holmes for generously sharing their time and expertise. Terry, who was the first person I met physically on campus, has provided indispensable technical assistance throughout my time at CMU – I particularly appreciate him taking a chance on hiring me before I even set foot in Pennsylvania and for trusting me with the dFAB resources (which, to my knowledge, remain unburnt!). Jon's insights from the woodshop were equally crucial in shaping this research. Their willingness to sit down and explain the nuances of their work provided invaluable real-world context for this journey.

Finally, I am profoundly grateful for the unwavering support of my family, loved ones, and friends. Their encouragement, emotional regulation, and inspiration have been more than I could have asked for, and I appreciate it nonetheless. To my partner, Maddie, thank you for taking an incredible leap in moving to Pittsburgh with me, sight unseen and futures uncertain. To my mother and father, thank you for keeping me grounded through all the growth and uncertainty of these years. And to my cherished friends, both near and far, including those digitally that have been a source of community since I was a teen, your support has meant the world to me.

Appendices

Appendix A: Technical Documentation

Detailed system architecture, code snippets, and other technical information can be found by clicking here. Or by visiting: https://www.mrmartinotti.com/technical-documentation/

Appendix B: User Study Materials

Raw feedback, questionnaires, and other user study materials can be found by clicking here. Or by visiting https://www.mrmartinotti.com/user-study/

References

[1] Unity Technologies. Unity, September 2024.

[2] Constance Paras, Danielle Kincaid, and Patty Tran. Buy Nothing Squirrel Hill North, Pittsburgh, PA | Facebook, February 2025.

[3] Madeline Taylor and Paige Street. Learning to (not quite) buy nothing: Digital gift economies and household innovations reducing consumption. *Environmental Innovation and Societal Transitions*, 53:100929, December 2024.

[4] Federica Joe Gardella, Zhixin Luo, Isaac Martinotti, David Troetschel, Michael Hasey, and Daniel Cardoso Llach. A Multi-scalar and Multi-modal Approach to Architectural Heritage Documentation: An Interactive Digital Representation of the St. Nicholas Chapel. *Florence Heri-Tech*, August 2024.

[5] Daniel Cardoso Llach. Research | CodeLab, 2025.

[6] Julie Carmigniani, Borko Furht, Marco Anisetti, Paolo Ceravolo, Ernesto Damiani, and Misa Ivkovic. Augmented reality technologies, systems and applications. *Multimedia Tools and Applications*, 51(1):341–377, January 2011.

[7] I-Jui Lee. Applying virtual reality for learning woodworking in the vocational training of batch wood furniture production. *Interactive Learning Environments*, 31(3):1448–1466, April 2023.

[8] Ryan Arisandi, Mai Otsuki, Asako Kimura, Fumihisa Shibata, and Hideyuki Tamura. Virtual Handcrafting: Building Virtual Wood Models Using ToolDevice. *Proceedings of the IEEE*, 102(2):185–195, February 2014. Conference Name: Proceedings of the IEEE.

[9] Peiyu (Steve) Hu. *Bridging the Gap: Exploring a Design-via-Making Approach in Architecture through Digital Guidance Systems*. Master of Science in Computational Design, Carnegie Mellon University, July 2024.

[10] Isaias Majil, Mau-Tsuen Yang, and Sophia Yang. Augmented Reality Based Interactive Cooking Guide. *Sensors*, 22(21):8290, October 2022.

[11] Daniel Cardoso Llach. Sculpting spaces of possibility. In Imdat As and Prithwish Basu, editors, *The Routledge Companion to Artificial Intelligence in Architecture*, pages 13–28. Routledge, Abington, Oxon; New York: Routledge, 2021., 1 edition, April 2021.

[12] Deirdre Visser. *Joinery, Joists and Gender: A History of Woodworking for the 21st Century*. Routledge, London, 1 edition, January 2022.

[13] Mary Evans. Carpenter's Apprentice - Child Labor in Carpenter's Shop, 1841.

[14] Michael Hasey. *The 3D Form Analysis of Regional Architecture Using Deep Learning: A Case Study of Wooden Churches from the Carpathian Mountains*. Master of Science in Computational Design, Carnegie Mellon University, August 2022.

[15] Cecilia Ferrando, Niccolò Dalmasso, Jiawei Mai, and Daniel Cardoso Llach. Architectural Distant Reading. *CAAD Futures 2019*, June 2019.

[16] Isaac Martinotti and Sarah Chen. Netflix Presents Hansel & Gretel – A Walk in the Park, 2024.

[17] Mitchell Shortt, Shantanu Tilak, Irina Kuznetcova, Bethany Martens, and Babatunde Akinkuolie. Gamification in mobile-assisted language learning: a systematic review of Duolingo literature from public release of 2012 to early 2020. *Computer Assisted Language Learning*, 36(3):517–554, March 2023. Publisher: Routledge _eprint: https://doi.org/10.1080/09588221.2021.1933540.

[18] Isaac Martinotti. ARRecipes. Independent Study, University of Oregon, Eugene, Oregon, December 2022.

[19] Petras Vestartas. petrasvestartas/OpenNest, November 2024. original-date: 2018-11-16T12:53:26Z.

[20] Ian Anderson. Self Assembly - Collection, 2021.

[21] Wei Wei Chi. 3D Microscopic Texture Interface in CAD: Cultivating Material Knowledge for the Practiced Digital Hand. *CMU KiltHub*, June 2019.

[22] Mary Tsai, Jonathan Chapman, and Mark Baskinger. *Making Mistakes: Error as Emergent Property in Craft Practice How might one find a more human connection into computational forms of making?* Master of Science in Computational Design, Carnegie Mellon University, July 2019.

[23] Laura Wolf-Powers, Marc Doussard, Greg Schrock, Charles Heying, Max Eisenburger, and Stephen Marotta. The Maker Movement and Urban Economic Development. *Journal of the American Planning Association*, 83(4):365–376, October 2017. Publisher: Routledge _eprint: https://doi.org/10.1080/01944363.2017.1360787.

[24] Samantha Shorey. *Handmade Future: A Field-based Inquiry of Innovation through Making and Craft*. Doctor of Philosophy, University of Washington, October 2019.

[25] Adam Ibrahim, Brandon Huynh, Jonathan Downey, Tobias Höllerer, Dorothy Chun, and John O'Donovan. ARbis Pictus: A Study of Language Learning with Augmented Reality. *IEEE Trans Vis Comput Graph*, November 2017.

[26] Hung-Lin Chi, Shih-Chung Kang, and Xiangyu Wang. Research trends and opportunities of augmented reality applications in architecture, engineering, and construction. *Automation in Construction*, 33:116–122, August 2013.

[27] I-Jui Lee and Yu-Ting Lin. Bridging the gap: integrating AR and VR production line training to enhance furniture carpentry students' understanding of production processes. *Interactive Learning Environments*, pages 1–22, May 2024.

[28] Alessandro Manni, Damiano Oriti, Andrea Sanna, Francesco De Pace, and Federico Manuri. Snap2cad: 3D indoor environment reconstruction for AR/VR applications using a smartphone device. *Computers & Graphics*, 100:116–124, November 2021.

[29] Sang-Yun Yang, Ohkyung Kwon, Yonggun Park, Hyunwoo Chung, Hyunbin Kim, Se-Yeong Park, In-Gyu Choi, and Hwanmyeong Yeo. Application of neural networks for classifying softwood species using near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, 28(5-6):298–307, October 2020.

[30] Dercilio Junior Verly Lopes, Greg W. Burgreen, and Edward D. Entsminger. North American Hardwoods Identification Using Machine-Learning. *Forests*, 11(3):298, March 2020. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.

[31] Kemal Hacıefendioğlu, Selen Ayas, Hasan Basri Başağa, Vedat Toğan, Fatemeh Mostofi, and Ahmet Can. Wood construction damage detection and localization using deep convolutional neural network with transfer learning. *European Journal of Wood and Wood Products*, 80(4):791–804, August 2022.

[32] Jialuo Yang, Yilin Elaine Liu, and Michael Nitsche. Looking for Tacit Knowledge in Woodworking: A Pilot Study. In *Adjunct Proceedings of the 2024 Nordic Conference on Human-Computer Interaction*, pages 1–5, Uppsala Sweden, October 2024. ACM.

[33] Derek Burkhardsmeier. *Augmented Woodworking*. Master of Design, University of Washington, December 2018.

[34] Andrea Settimi, Chutisilp Naravich, Zholmagambetova Nazgul, and Aymanns Florian. Software of TTool: a supervised AI-assisted Visual Pose Detector for AR Wood-working, January 2024.

[35] Andrea Settimi, Yves Weinand, and Julien Gamerro. Augmented carpentry, January 2021.

[36] Tobias Dreesbach, Henning Gosling, Jan Schulte to Brinke, and Martin Kasselmann. AR-BASED LEARNING IN VOCATIONAL TRAINING – EVALUATION OF LEARNING OUTCOME COMPARING AUGMENTED REALITY AND TEACHER-INSTRUCTED LEARNING UNITS. *ECIS 2024 Proceedings*, 10, May 2024.

[37] Gonzalo Alfonso Beltrán Alvarado. Implementation of a Visual Augmented Reality Method in a Carpentry Course: A Case Study. *Electronic Journal of e-Learning*, 22(3):141–159, December 2024.

[38] Tim Cousin, Latifa Alkhayat, Natalie Pearl, Christopher B. Dewart, and Caitlin Mueller. Wild Wood Gridshells: Mixed-Reality Construction of Nonstandard Wood. *Technology|Architecture + Design*, 7(2):216–231, July 2023.

[39] Xiaozhi Chen, Kaustav Kundu, Ziyu Zhang, Huimin Ma, Sanja Fidler, and Raquel Urtasun. Monocular 3D Object Detection for Autonomous Driving. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2147–2156, 2016.

[40] Ivan Barabanau, Alexey Artemov, Evgeny Burnaev, and Vyacheslav Murashkin. Monocular 3D Object Detection via Geometric Reasoning on Keypoints, May 2019. arXiv:1905.05618 [cs].

[41] Yunlei Tang, Sebastian Dorn, and Chiragkumar Savani. Center3D: Center-based Monocular 3D Object Detection with Joint Depth Understanding, May 2020. arXiv:2005.13423 [cs].

[42] Bowen Wen, Wei Yang, Jan Kautz, and Stan Birchfield. FoundationPose: Unified 6D Pose Estimation and Tracking of Novel Objects, March 2024. arXiv:2312.08344 [cs].

[43] Xingyi He, Jiaming Sun, Yuang Wang, Di Huang, and Xiaowei Zhou. OnePose++: Keypoint-Free One-Shot Object Pose Estimation without CAD Models. *NeurIPS 2022*, 2022.

[44] Linqi Yang, Xiongwei Zhao, Qihao Sun, Ke Wang, Ao Chen, and Peng Kang. SplatPose: Geometry-Aware 6-DoF Pose Estimation from Single RGB Image via 3D Gaussian Splatting, March 2025. arXiv:2503.05174 [cs].

[45] Chenming Wu, Haisen Zhao, Chandrakana Nandi, Jeffrey I. Lipton, Zachary Tatlock, and Adriana Schulz. Carpentry compiler. *ACM Transactions on Graphics*, 38(6):1–14, December 2019.

[46] Filson & Rohrbacher. AtFAB Open Source Furniture, 2008.

[47] Christopher Alexander, Sara Ishikawa, Murray Silverstein, Max Jacobson, Ingrid Fiksdahl-King, and Shlomo Angel. *A Pattern Language: Towns, Buildings, Construction*. Oxford University Press, New York, January 1977.

[48] G Stiny. Introduction to shape and shape grammars. *Environment and Planning B: Planning and Design*, 7(3):343–351, 1980.

[49] Dominik Reisach, Stephan Schütz, Jan Willmann, and Sven Schneider. A Design-to-Fabrication Workflow for Free-Form Timber Structures Using Offcuts. In Michela Turrin, Charalampos Andriotis, and Azarakhsh Rafiee, editors, *Computer-Aided Architectural Design. INTERCONNECTIONS: Co-computing Beyond Boundaries*, pages 361–375, Cham, 2023. Springer Nature Switzerland.

[50] Dominik Reisach. *Upcycle Timber*. Master of Sciences in Architecture, Bauhaus-Universitat Weimar, 2024.

[51] Ozguc Bertug Capunaman. *CAM as a Tool for Creative Expression: Informing Digital Fabrication through Human Interaction*. Master of Science in Computational Design, Carnegie Mellon University, Bangkok, Thailand, October 2019.

[52] Gil Sunshine. Research Connections | Inventory: CAD for Medium Resolution Materials, May 2023.

[53] I-Jui Lee, Ten-Li Chen, and Meng-Cong Zheng. The Application of AR Technology to Spatial Skills Learning in Carpentry Training. *International Journal of Information and Education Technology*, 9(1):56–60, 2019.

[54] Chia-lin Yang and Chien-Hsu Chen. Applying Paper Prototyping to Design an AR Teaching Tool for Novice Woodworkers Practicing Components Arrangement in Material Cutting. In *15th International Conference on Applied Human Factors and Ergonomics (AHFE 2024)*, 2024.

[55] Philip A Araman, D P Hindman, and M F Winn. Characterization and potential recycling of home building wood waste. *Processing Technologies for the Forest and Biobased Products Industries*, January 2010.

[56] Emma Whittaker and Polly Macpherson. *The Green Maker Initiative Book*. University of Plymouth Press, May 2023.

[57] Viktor Poteschkin, Jürgen Graf, Stefan Krötsch, and Wenchang Shi. Recycling of Cross-Laminated Timber Production Waste. In *Research Culture in Architecture: Cross-Disciplinary Collaboration*, pages 101–112. Birkhäuser, December 2019. Section: Research Culture in Architecture.

[58] Boyuan Yu, Jianing Luo, Yi Shi, Mingming Zhao, Adam Fingrut, and Lei Zhang. Framework for sustainable building design and construction using off-cut wood. *npj Materials Sustainability*, 1(1):2, October 2023.

[59] Ernesto Lo Valvo. Meta-heuristic Algorithms for Nesting Problem of Rectangular Pieces. *Procedia Engineering*, 183:291–296, 2017.

[60] Yuji Yoshimura, Bill Cai, Zhoutong Wang, and Carlo Ratti. Deep Learning Architect: Classification for Architectural Design Through the Eye of Artificial Intelligence. In Stan Geertman, Qingming Zhan, Andrew Allan, and Christopher Pettit, editors, *Computational Urban Planning and Management for Smart Cities*, volume Computational Urban Planning and Management for Smart Cities of *Lecture Notes in Geoinformation and Cartography*, pages 249–265. Springer International Publishing, Cham, 2019. Series Title: Lecture Notes in Geoinformation and Cartography.

[61] Qing Han and Chao Yin. Architectural style classification of the Chinese traditional settlements using deep learning. In *International Conference on Geographic Information and Remote Sensing Technology (GIRST 2022)*, volume 12552, pages 784–790. SPIE, February 2023.

[62] Jinsung Kim and Jin-Kook Lee. Stochastic Detection of Interior Design Styles Using a Deep-Learning Model for Reference Images. *Applied Sciences*, 10(20):7299, January 2020. Number: 20 Publisher: Multidisciplinary Digital Publishing Institute.

[63] Jinmo Rhee and Eddy Man Kim. Deeprise: A Novel Method to Model and Design a High-rise Building Form through Deep Neural Network. *CAAD Futures 2021*, 1465, 2020.

[64] Jinmo Rhee, Daniel Cardoso Llach, and Ramesh Krishnamurti. Context-rich Urban Analysis Using Machine Learning. *CITY INFORMATION MODELLING AND GIS*, 3, 2019.

[65] Jielin Chen and Rudi Stouffs. Deciphering the noisy landscape: Architectural conceptual design space interpretation using disentangled representation learning. *Computer-Aided Civil and Infrastructure Engineering*, 38(5):601–620, 2023. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/mice.12908.

[66] Andrew Witt. Through a Machine Lens: Case Studies of Computer Vision and Machine Learning in Design Methodology. *Loci Communes*, 1(2):1–16, June 2022.

[67] Mirko Karič, Manja Kitek-Kuzman, and Jože Kropivšek. Possibilities of Using Artificial Intelligence in Furniture/Woodworking Industry. *Drvna industrija*, 75(3):371–380, September 2024. Publisher: Sveučilište u Zagrebu Fakultet šumarstva i drvne tehnologije.

[68] Dereck Bridie. googlesamples/arcore-ml-sample, May 2021.

[69] EasyAR. free ar sdk – EasyAR Sense, 2025.

[70] Visometry GmbH. VisionLib – Augmented Reality Tracking for Industries, July 2024.

[71] Vuforia. Model Targets - Vuforia Engine Library, February 2025.

[72] seeeeeeeeeeeeeed. TRIXIG - Is this toolbox enough for all Ikea assemblies and dismantles?, November 2024.

[73] petalmasher. Ikea tool kit?, August 2022.

[74] level1hero. Not quite BIFL but I've abused this for nearly 20 years, May 2024.

[75] Eileen Fischer, Brenda Gainer, and Julia Bristor. The sex of the service provider: Does it influence perceptions of service quality? *Journal of Retailing*, 73(3):361–382, 1997.

[76] Marco Wolf, Pia A. Albinsson, and Cherylynn Becker. Do-It-Yourself Projects as Path toward Female Empowerment in a Gendered Market Place. *Psychology & Marketing*, 32(2):133–143, February 2015.

[77] Charles R. Qi, Wei Liu, Chenxia Wu, Hao Su, and Leonidas J. Guibas. Frustum PointNets for 3D Object Detection from RGB-D Data, April 2018. arXiv:1711.08488 [cs].

[78] Peiliang Li, Xiaozhi Chen, and Shaojie Shen. Stereo R-CNN Based 3D Object Detection for Autonomous Driving. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7636–7644, Long Beach, CA, USA, June 2019. IEEE.

[79] Brandon Gilbert. Hammer safety, August 2018.

[80] Dan Leipert. Hand Saw Use and Safety, August 2020.

[81] EHS OkState. HAND & POWER TOOL SAFETY, September 2016.

[82] Albert Jackson, David Day, and Simon Jennings. *The complete manual of woodworking*. Alfred A. Knopf, New York, 2019.

[83] Alan Bridgewater, Gill Bridgewater, Glyn Bridgewater, Colin Eden-Eadon, Sally Francis, John Lloyd, Jonathan Tibbs, Chris Ward, and J. M. Wilkie. *Woodworking: the complete step-by-step manual*. DK Publishing, New York, NY, new edition edition, 2024. OCLC: on1385297394.

[84] Jon Holmes. Initial Onboarding Session with Jon Holmes, January 2025.

[85] Aleksei Bochkovskii, Amaël Delaunoy, Hugo Germain, Marcel Santos, Yichao Zhou, Stephan R. Richter, and Vladlen Koltun. Depth Pro: Sharp Monocular Metric Depth in Less Than a Second, October 2024. arXiv:2410.02073 [cs].

[86] Alfian Losari. Build visionOS & iOS AR Realtime Inventory App | SwiftUI | RealityKit | Firebase | Full Tutorial, August 2023.

[87] Ananda Sai A. anandasaia/NeuralDistance, July 2024. original-date: 2023-06-13T19:19:49Z.

[88] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection, May 2016. arXiv:1506.02640 [cs].

[89] Qi-Chao Mao, Hong-Mei Sun, Yan-Bo Liu, and Rui-Sheng Jia. Mini-YOLOv3: Real-Time Object Detector for Embedded Applications. *IEEE Access*, 7:133529–133538, 2019. Conference Name: IEEE Access.

[90] Seokyong Shin, Hyunho Han, and Sang Hun Lee. Improved YOLOv3 with duplex FPN for object detection based on deep learning. *International Journal of Electrical Engineering & Education*, 62(2):127–143, April 2025. Publisher: SAGE Publications Ltd STM.

[91] Tzuta Lin. HumanSignal/labelImg, March 2025. original-date: 2015-09-17T01:33:59Z.

[92] Glenn Jocher, Jing Qiu, and Ayush Chaurasia. Ultralytics YOLO, January 2023. original-date: 2022-09-11T16:39:45Z.

[93] Kaustubh Sadekar and Satya Mallick. Camera Calibration using OpenCV | LearnOpenCV #, February 2020.

[94] Rino Falstad. IKEA Toolbox - The Hammer - Download Free 3D model by cyberreno, May 2023.

[95] Peter Selinger. Potrace, September 2019.

[96] Ian Anderson. Self Assembly - Sedia 1 Chair Autoprogettazione by Enzo Mari, October 2019.

[97] Andrea Settimi, Hong-Bin Yang, Julien Gamerro, and Yves Weinand. TSLAM: a tag-based object-centered monocular navigation system for augmented manual woodworking. *Construction Robotics*, 8(1):4, June 2024.