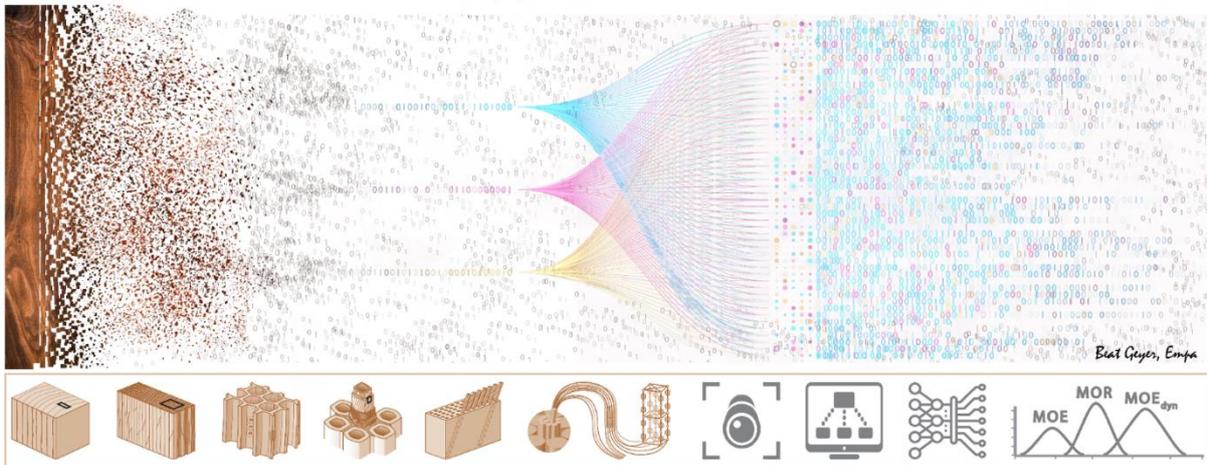


## Abschlussbericht WHFF-CH 2021.09:

# Bestimmung der Festigkeit von Holz mit Computer Vision



**Auftraggeber** BAFU/KWL

**Projektleitung** Dr. Mark Schubert & Dr. Mirko Luković

**Projektlaufzeit** 18 Monate

**Datum** Dezember 2023

# Inhaltsangabe

VORWORT & DANKSAGUNG.....	3
ZUSAMMENFASSUNG.....	4
<b>1 PROJEKTGRUNDLAGEN .....</b>	<b>6</b>
1.1 Problembeschreibung.....	6
1.2 Forschungsarbeiten und Wissen an der Empa.....	7
1.3 Projektinformationen .....	7
1.3.1 Ziele.....	7
1.3.2 Bedeutung des Projektes für Forschung und Praxis.....	8
1.3.3 Thematische Eingliederung in die Schwerpunkte des WHFF-CH.....	8
<b>2 DATENMANAGEMENT .....</b>	<b>9</b>
2.1 Datenerfassung .....	9
2.2 Daten-Vorverarbeitung .....	9
2.2.1 Bildsegmentierung.....	9
2.3 Datenerweiterung.....	10
2.3.1 Einfluss der Datenerweiterung.....	11
2.4 Schlussfolgerung .....	12
<b>3 COMPUTER VISION MIT CONVOLUTIONAL NEURAL NETWORKS (CNNs) .....</b>	<b>13</b>
3.1 Trainingsprozess .....	13
3.2 BigTransfer Modell.....	15
3.2.1 Performance Evaluation .....	15
3.3 Schlussfolgerung.....	17
<b>4 EXPERIMENTELLE VALIDIERUNG DES CNN-MODELLS.....</b>	<b>18</b>
4.1 Schlussfolgerung.....	18
<b>5 FAZIT.....</b>	<b>18</b>
5.1 Umsetzung .....	19
5.2 Offene Fragen / Weiterer Forschungsbedarf.....	20
<b>LITERATURVERZEICHNIS .....</b>	<b>21</b>
<b>ANHANG</b>	

## Vorwort & Danksagung

Der vorliegende Bericht dokumentiert die Ergebnisse des WHFF-CH Projektes 2021.09 „Bestimmung der Festigkeit von Holz mit Computer Vision“.

Folgende Organisationen und Personen waren massgeblich an der Realisierung des Projektes beteiligt:

- Empa, Gruppe WoodTec: M. Schubert, M. Lukovic, G. Müller, R. Elsener
- ETH (Studenten): L. Ciernik, D. Kluser, T. Pham
- Derix: A. Kassenbrock
- Schilliger Holz AG: E. Schilliger
- Pfeifer Group: R. Eller
- Wald- und Holzforschungsförderung Schweiz WHFF-CH: C. Suter

An dieser Stelle möchte sich die Empa, Gruppe WoodTec bei allen Personen und Institutionen für das Engagement, wertvollen Empfehlungen und Anmerkungen sowie die geleisteten Arbeiten, die zum Gelingen des Projektes beigetragen haben, bedanken.

Basierend auf den vielversprechenden Ergebnissen werden Teile des Projektes in folgender Publikation veröffentlicht:

Lukovic, M. Ciernik, L., Müller, G., Kluser, D., Pham, T., Burgert, I., Schubert, M. (2023). Probing wood complexity with computer vision: From pixels to properties. *Journal of The Royal Society Interface*. Accepted

### **Kontakt:**

Dr. Mark Schubert

Empa – Swiss Federal Laboratories for Materials Science and Technology  
Laboratory for Cellulose & Wood Materials  
Group 'WoodTec'  
Überlandstrasse 129  
CH-8600 Dübendorf  
+41 (0)58 7657624  
[mark.schubert@empa.ch](mailto:mark.schubert@empa.ch)  
[www.empa.ch](http://www.empa.ch)

## Zusammenfassung

Holz stellt einen ausgezeichneten Baustoff dar, dank seiner herausragenden Gewichts- und Leistungseigenschaften, nachhaltigen Verfügbarkeit und allgemeinen Wohlfühlaspekten. Holz wird als Hauptbauteil für verschiedene technische Konstruktionen genutzt. Dennoch könnte seine Effizienz und Anwendung in anspruchsvolleren Szenarien weiter gesteigert werden. Die inhärente heterogene Materialstruktur auf verschiedenen Längenskalen führt zu einer beträchtlichen Variabilität der Holzeigenschaften, was grosse Sicherheitsfaktoren erfordert und die Materialnutzungseffizienz begrenzt.

In der Vergangenheit wurden Optimierungen in der Holzsortierung durchgeführt, wobei neben der etablierten visuellen Inspektion auch maschinelle Sortiersysteme entwickelt wurden. Durch die Entwicklung von maschinellen Sortiersystemen können mehr Parameter, einschliesslich mechanischer Eigenschaften, schneller und quantitativ gemessen werden. Allerdings erfolgt die Korrelation zwischen den gemessenen Sortierparametern und den strukturellen Eigenschaften im Allgemeinen mit einfachen linearen oder multilinearen Korrelationen, ausseracht lassend, dass die komplexe Holzstruktur und oft nicht-lineare Zusammenhänge aufweist. Techniken des maschinellen Lernens sind besonders geeignet, um diese Herausforderungen bei der Mustererkennung in komplexen Systemen zu bewältigen. Die Kombination von maschinellem Lernen mit zerstörungsfreien optischen und digitalen Bildanalysemethoden stellt daher eine vielversprechende Strategie dar um die Effizienz der Sortierprozesse zu steigern.

Das Ziel dieser Machbarkeitsstudie war es daher ein auf maschinellem Lernen basierenden Algorithmus zu entwickeln der die Bestimmung der mechanischen Eigenschaften von Holzlamellen, allein anhand von Bildern als Eingabe, erlaubt. Durch die Umsetzung eines solchen maschinellen Lernansatzes für die Holzsortierung könnten nicht nur die Effizienz und Genauigkeit der Prozesse gesteigert, sondern auch Ressourcen effektiver genutzt werden. Dieser innovative Ansatz könnte somit einen bedeutenden Beitrag zur Weiterentwicklung und Optimierung von Sortierverfahren im Holzbereich leisten.

### Projektziele

#### **1. Erstellung eines Trainings-, Validierungs- und Testdatensatzes mit Data Augmentation:**

*Ziel:* Zusammenstellung eines umfassenden Datensatzes von hochqualitativen Holzbildern, die verschiedene Texturen und Zustände repräsentieren. Integration von Data Augmentation-Techniken, um die Robustheit der Modelle zu verbessern.

*Begründung:* Die Qualität und Vielfalt des Datensatzes sind entscheidend für die Robustheit und die Leistungsfähigkeit der Machine-Learning-Modelle. Die Anwendung von Data Augmentation während des Trainingsprozesses ermöglicht es, den Datensatz zu erweitern und verbessert die Fähigkeit der Modelle, auf unterschiedliche Bedingungen zu generalisieren.

## **2. Entwicklung von Convolutional Neural Netzwerk-Modellen (CNNs):**

*Ziel:* Trainieren und Testen von CNN-Modellen zur Vorhersage der mechanischen Eigenschaften auf der Grundlage des aufgebauten Datensatzes.

*Begründung:* CNN-Modelle ermöglichen die automatisierte Analyse grosser Datenmengen (Bilder) und können komplexe Beziehungen zwischen Merkmalen und Eigenschaften lernen.

## **3. Validierung der CNN-Modelle:**

*Ziel:* Validierung der entwickelten Modelle mit unabhängigen Datensätzen oder experimentellen Ergebnissen.

*Begründung:* Die Zuverlässigkeit und Genauigkeit der Modelle müssen überprüft werden, um die Generalisierungsfähigkeit sicherzustellen.

## **Schlussfolgerung**

In diesem Projekt wurde erfolgreich ein umfassender Trainings- und Testdatensatz mit hochqualitativen Holzbildern erstellt, indem Data Augmentation-Techniken integriert wurden. Dieser Datensatz repräsentiert verschiedene Texturen und Zustände der Holzlamellen und ermöglichte das erfolgreiche Trainieren und Testen von Convolutional Neural Network (CNN)-Modellen.

Die CNN-Modelle zeigten vielversprechende Ergebnisse, indem sie eine hohe Genauigkeit bei der Vorhersage der mechanischen Eigenschaften von Holzlamellen erreichten. Die Effektivität der Modelle im Training sowie ihre Generalisierungsfähigkeit auf unterschiedliche Bedingungen wurden durch eine sorgfältige Validierung mit Testdaten und eigenen Experimenten erfolgreich nachgewiesen.

Die Ergebnisse dieses Projekts zeigen die Machbarkeit die mechanischen Eigenschaften von Holz allein durch die Analyse von Bildern vorherzusagen. Diese Erkenntnis kann nicht nur zur Weiterentwicklung des Sortierprozesses beitragen, sondern fördert auch das Verständnis und die Anwendung von maschinellem Lernen im Holzbereich.

# 1. Projektgrundlagen

## 1.1 Problembeschreibung

Wälder spielen eine besondere Rolle in den Strategien zur Minderung des globalen Klimawandels, da Bäume aktiv CO<sub>2</sub> im Holz speichern, das gleichzeitig eine Schlüsselressource zur nachhaltigeren Gestaltung des Bausektors darstellt [1-3]. Herausfordernd ist jedoch, nachhaltige Forstwirtschaft und die Bereitstellung von Holz im grossen Massstab zu kombinieren, und diese Situation könnte sich aufgrund des Klimawandels potenziell verschärfen. Längere Dürreperioden und extreme Wetterereignisse setzen Bäume zunehmendem Stress aus und erfordern spezifische Massnahmen, um die Widerstandsfähigkeit und Stabilität der Waldökosysteme zu erhalten und manchmal wiederherzustellen, einschliesslich der Verhinderung weiterer Verschlechterungen der Biodiversität. Ohne Zweifel erfordert ein nachhaltigerer Bausektor, der von einer verstärkten Nutzung nachhaltiger Holzmaterialien profitiert, die Bewirtschaftung und Pflege nachhaltiger Wälder. Angesichts des sich ändernden Klimas und des Verlusts an Biodiversität bauen diese widerstandsfähigen und nachhaltigen Wälder höchstwahrscheinlich auf einer Mischung von Holzarten auf, anstatt Monokulturen zu sein [4,5].

Ein Hauptproblem besteht darin, dass derzeit nur sehr wenige Baumarten für die Bauindustrie genutzt werden. Dabei handelt es sich hauptsächlich um Nadelhölzer wie die Fichte (*Picea abies*). Abgesehen davon, dass dies in Bezug auf Ressourcen begrenzend ist, begünstigt dieser Mangel an Vielfalt die Bildung von Monokulturen, die im Vergleich zu natürlichen Wäldern weniger widerstandsfähig sind und ein höheres Risiko für Krankheiten aufweisen. Darüber hinaus werden Nadelhölzer auch in vielen Teilen der Welt aufgrund von trockeneren und wärmeren Wetterbedingungen, die durch den Klimawandel verursacht werden, knapper [6]. Folglich muss man beginnen, sich auf eine breitere Vielfalt von Baumarten zu verlassen, von denen einige resistenter gegen die erwarteten Klimaveränderungen sind [4,5,7].

Eine veränderte Baumartenzusammensetzung erfordert angepasste Sortierprozesse. Diese Sortierprozesse müssen nicht nur effektiv und präzise sein, sondern auch kostengünstig und flexibel, um den vielfältigen Anforderungen der unterschiedlichen Holzarten gerecht zu werden. Dies erfordert möglicherweise unterschiedliche Parameter und Kriterien für die Sortierung, um sicherzustellen, dass die Holzprodukte den spezifischen Anforderungen der Bauindustrie entsprechen. Gleichzeitig ist es notwendig, die Kosten im Auge zu behalten, um die Wirtschaftlichkeit der Holzverarbeitung aufrechtzuerhalten und den Einsatz nachhaltiger Holzressourcen zu fördern.

Insgesamt müssen die Sortierprozesse an die veränderte Baumartenzusammensetzung angepasst werden, um eine effiziente, kostengünstige und flexible Holzverarbeitung zu gewährleisten, die den Herausforderungen des sich wandelnden Klimas und der veränderten Verfügbarkeit von Holzarten gerecht wird.

Techniken des maschinellen Lernens sind besonders geeignet, um Herausforderungen bei der Mustererkennung in komplexen Systemen zu bewältigen. Die Kombination von maschinellem Lernen mit zerstörungsfreien optischen und digitalen Bildanalysemethoden stellt daher eine vielversprechende Strategie dar um die Effizienz der Sortierprozesse zu steigern.

## 1.2 Forschungsarbeiten und Wissen an der Empa

Die Empa widmet sich seit vielen Jahren intensiv der Erforschung von Holz. Von den anfänglichen Untersuchungen zu den Eigenschaften Schweizer Holzsorten hat sich die Forschung zu einem breiten Spektrum ausgedehnt, das von Grundlagenforschung über die Holzstruktur bis hin zu modifizierten Hölzern, innovativen Holzprodukten und Digitalisierungstechnologien reicht. Das interdisziplinäre Team von Wissenschaftlern der WoodTec-Gruppe an der Empa und ETH verfügt über vielfältige Kompetenzen in Holztechnologie, Datenwissenschaft und insbesondere im Bereich Machine Learning. So wurden bereits in Zusammenarbeit mit der Holzindustrie für diverse Anwendungen erfolgreich Machine Learning Algorithmen entwickelt und programmiert [8,9,10,11,12]. Um das vorgeschlagene Projekt in der kurzen Laufzeit von 18 Monaten zu realisieren wurden diverse Vorarbeiten im Vorfeld durchgeführt. Es wurden in Python Pipelines programmiert, welche den Datenimport, das Trainieren und Evaluieren der maschinellen Lernalgorithmen sowie das Einstellen der Hyperparameter erleichtern.

## 1.3 Projektinformationen

### 1.3.1 Ziele

Das Ziel dieser Arbeit bestand darin, einen Algorithmus auf Basis von maschinellem Lernen zu entwickeln, der es ermöglicht, die mechanischen Eigenschaften von Holzlamellen ausschliesslich anhand von Bildern zu bestimmen. Folgende Projektziele sollten im Laufe des Projektes erreicht werden:

#### **1. Erstellung eines Trainings-, Validierungs- und Testdatensatzes mit Data Augmentation:**

**2. Ziel:** Zusammenstellung eines umfassenden Datensatzes von hochqualitativen Holzbildern, die verschiedene Texturen und Zustände repräsentieren. Integration von Data Augmentation-Techniken, um die Robustheit der Modelle zu verbessern.

*Begründung:* Die Qualität und Vielfalt des Datensatzes sind entscheidend für die Robustheit und die Leistungsfähigkeit der Machine-Learning-Modelle. Die Anwendung von Data Augmentation während des Trainingsprozesses ermöglicht es, den Datensatz zu erweitern und verbessert die Fähigkeit der Modelle, auf unterschiedliche Bedingungen zu generalisieren.

### **3. Entwicklung von Convolutional Neural Netzwerk-Modellen (CNNs):**

*Ziel:* Trainieren und Testen von CNN-Modellen zur Vorhersage der mechanischen Eigenschaften auf der Grundlage des aufgebauten Datensatzes.

*Begründung:* CNN-Modelle ermöglichen die automatisierte Analyse grosser Datenmengen (Bilder) und können komplexe Beziehungen zwischen Merkmalen und Eigenschaften lernen.

### **4. Validierung der CNN-Modelle:**

*Ziel:* Validierung der entwickelten Modelle mit unabhängigen Datensätzen oder experimentellen Ergebnissen.

*Begründung:* Die Zuverlässigkeit und Genauigkeit der Modelle müssen überprüft werden, um die Generalisierungsfähigkeit sicherzustellen.

#### *1.3.2 Bedeutung des Projektes für Forschung und Praxis*

Die Einführung eines Machine-Learning-Systems (Computer Vision), das die Bestimmung der mechanischen Eigenschaften von Holzlamellen anhand von Bildern ermöglicht, könnte zur Effizienzsteigerung beitragen und eine optimierte Nutzung von Holzressourcen ermöglichen. Die Entwicklung eines solchen Holzsortierungsalgorithmus könnte einen bedeutenden Fortschritt für Forschung und Praxis darstellen, indem er innovative, kosteneffiziente und ressourcenschonende Holzverarbeitungsverfahren ermöglicht.

#### *1.3.3 Thematische Eingliederung in die Schwerpunkte des WHFF-CH*

Das vorgeschlagene Projekt befasst sich mit der Entwicklung eines ML Algorithmus für die Holzsortierung mittels Computer Vision und bezieht sich somit insbesondere auf den WHFF-CH Schwerpunkt 2 'Optimierung von Prozessen und Produktionsmethoden'.

## 2. Datenmanagement

### 2.1 Datenerfassung

Die verwendeten Daten bestanden aus Bildern von 15.104 Proben von Fichtenholz (*Picea abies*) zusammen mit Werten des statischen Elastizitätsmoduls (MOE), des dynamischen Elastizitätsmoduls (MOE<sub>dyn</sub>) und des Biegemoduls (MOR) für jede Probe. Jedes Rohbild hatte eine Auflösung von 1000x562 Pixeln und enthielt Farbfotos von vier Seiten der Holzlamellenprobe. Die Bilder und Daten wurden durch das Scannen jeder Lamelle mit einer Microtec GoldenEye-Holzsortiermaschine [13] erhalten. Die Holzsortierung ist der nicht destruktive Prozess der Zuweisung einer Festigkeitsklasse zu Lamellen basierend auf gemessenen physikalischen Eigenschaften und standardisierten Verfahren. Da alle Eigenschaften, die mit der Holzfestigkeit zusammenhängen, zwischen verschiedenen Baumproben stark variieren, ist das Messen jeder einzelnen Lamelle notwendig. Die Daten wurden in drei Sätze unterteilt, die für das Training, die Validierung und das Testen der Modelle im Verhältnis 80:10:10 verwendet wurden.

### 2.2 Daten-Vorverarbeitung

#### 2.2.1 Bildsegmentierung

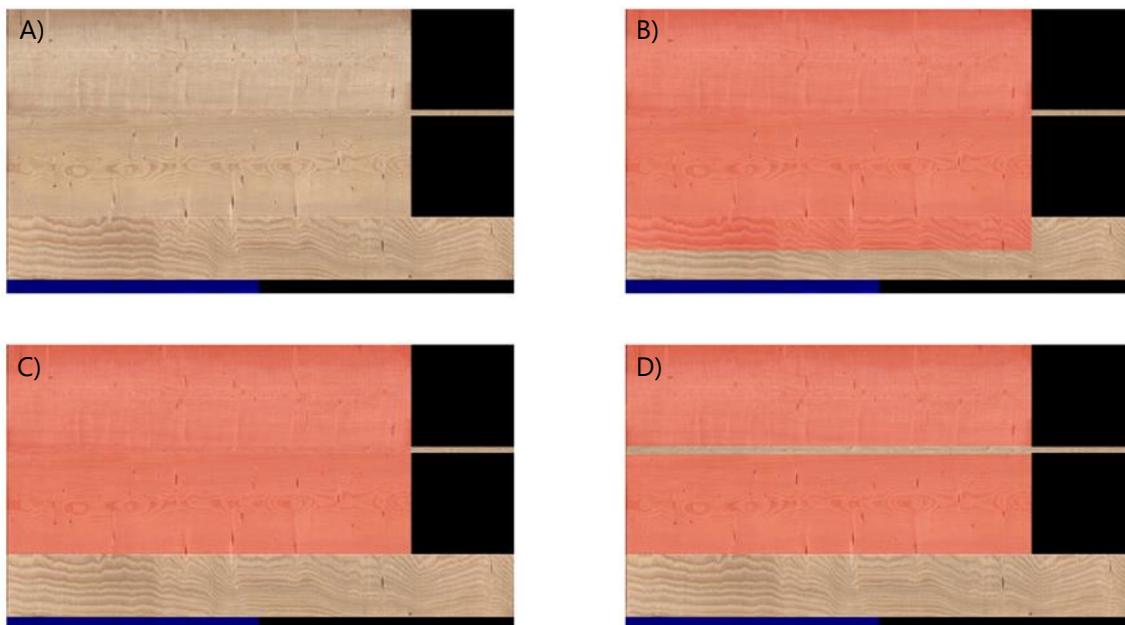
Bevor die Bilder verwendet werden konnten, mussten diese segmentiert werden, um nur die Teile zu extrahieren, die die notwendigen Informationen über die Holzproben enthielten. Wie in Abbildung 1 ersichtlich wurden 3 unterschiedliche Segmentierungsmethoden (Basic, Symmetric und Double crop) ausprobiert. Die Dimensionen der Bilder nach dem Zuschneiden waren wie folgt:

Crop	Size in pixels
Basic	535 x 797
Symmetric	410 x 797
Double	2 x 195 x 797

**Basic Crop:** Der rote Bereich in Abbildung 1 wurde als Eingabe für unseren Lernalgorithmus verwendet. Der unterste Teil der Bildkomposition wurde in der Mitte abgeschnitten (Abbildung 1B). Obwohl in jedem Bild nicht sichtbar, enthielten einige der Bilder Artefakte im unteren Teil (am unteren Bildrand), die man entfernen musste.

**Symmetric Crop:** Bei dieser Segmentationstechnik wurde der unterste Teil des Bildes ausgeschnitten (siehe den roten Bereich in Abbildung 1C). Obwohl der Lernalgorithmus in diesem Fall weniger Informationen für das Training zur Verfügung hatte, waren die Segmente in Bezug auf die Skala passender.

**Double Crop:** Wie in Abbildung 1D dargestellt wurde dem Modell nur Informationen von derselben Skala zur Verfügung gestellt. Die Regressions-Schicht (siehe Kapitel 3) kombinierte die extrahierten Merkmale von beiden Seiten der Lamelle. Der Effekt war zweifach: Das Modell konnte auf Bildern derselben Skala lernen, musste jedoch auch über die untere und obere Oberfläche generalisieren.



**Abbildung 1:** A) Original. B) Basic Crop. C) Symmetric Crop. D) Double Crop.

Je Konfiguration haben wir ein Modell 60 Epochen lang trainiert und den Bestimmtheitskoeffizient ( $R^2$ ) nach jeder Epoche auf dem zurückgehaltenen Validierungsset bestimmt. Obwohl die Unterschiede gering waren, zeigte sich, dass der 'Basic Crop' in allen Epochen die anderen Methoden dominierte. Daher wurde diese Methode in den folgenden Untersuchungen als Standard verwendet.

## 2.3 Datenerweiterung

Data-Augmentationsmethoden wurden implementiert mit dem Ziel den vorhandenen Datensatz zu erweitern und das Training des tiefen neuronalen Netzwerks zu verbessern. Eine der grössten Herausforderungen bestand darin, Methoden zu entwickeln, die 'labelschonend' sind. Zum Beispiel beeinflusst das Vorhandensein von Ästen drastisch

die mechanische Festigkeit der Lamellen [14]. Daher könnte die Anwendung einer gängigen Augmentationsmethode die Leistung beeinträchtigen, da wir Äste entfernen könnten, ohne die Labels entsprechend zu aktualisieren. Folglich wurde besonders sorgfältig der Einfluss der Data-Augmentationstechniken untersucht.

**Experimenteller Aufbau:** In den durchgeführten Data-Augmentations-Experimenten wurden ausschliesslich niedrig aufgelöste Bilder verwendet, die zudem einfach zugeschnitten waren. Als Modell diente das BigTransfer-Modell, wie im Kapitel 3 beschrieben. Die Entscheidung für niedrig aufgelöste Daten trug dazu bei, die Hyperparameter-Suche zu beschleunigen. Respektive wurde das ResNet 21k für jedes Experiment über 60 Epochen hinweg mit einer Batch-Grösse von 16 feinabgestimmt. Der BiT Hyper-Rule kam dabei als Optimierer und Zeitplaner zum Einsatz. Die Trainings- und Evaluierungseinstellungen sind in Tabelle 1 aufgeführt. Zur Gewährleistung von Zuverlässigkeit und Aussagekraft wurde jede Konfiguration fünfmal durchgeführt.

**Tabelle 1:** Trainingsparameter, die für alle Datenaugmentations-Experimente verwendet wurden

Training and evaluation settings overview	
Data crop	Basic crop
Data resolution	Low-resolution (188 x 280 pixels)
Model	BigTransfer model
Optimizer and scheduler	BiT Hyper-Rule
Number of epochs	60
Train and test batch size	16

### 2.3.1 Einfluss der Datenerweiterung

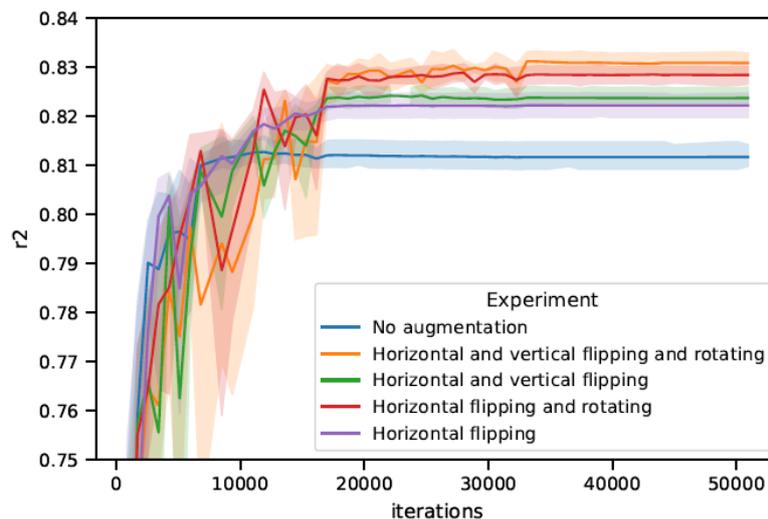
Aufgrund der vergleichsweise geringen Anzahl verfügbarer Bilder wurden erhebliche Anstrengungen in die Datenanreicherung unternommen. Dazu gehörten verschiedene Techniken wie zufälliges Spiegeln, zufällige Rotation, zufälliges Zuschneiden sowie zufällige Anpassungen von Kontrast, Sättigung und Helligkeit. Darüber hinaus wurde ein Gauss'scher Weichzeichnungsfilter angewendet, und den Bildern wurde Gauss'sches Rauschen hinzugefügt. Detaillierte Informationen zu den Anreicherungsverfahren finden sich im Anhang.

Die Tabelle 2 zeigt die höchsten erreichten Trainings- und Validierungs-Scores  $R^2$ . Für jedes Experiment wurde der Durchschnittsverlauf über die Anzahl der Iterationen berechnet und den maximalen  $R^2$  des Durchschnittsverlaufs bestimmt (Abbildung 2). Dies wurde sowohl für die Trainings- als auch für die Validierungs-Scores durchgeführt. Je

mehr Augmentationsmethoden angewendet wurden, desto besser wurde der  $R^2$  (Validierung), was gleichbedeutend ist, dass die Generalisierungsfähigkeit des Modells zugenommen hat.

**Tabelle 2:** Training und Validierung der Data-Augmentation Experimente

Training and validation scores		
Experiment	Training $R^2$	Validation $R^2$
No augmentation	0.996	0.813
Randomly flip horizontally and rotate by max. three degree (Baseline)	0.971	0.828
Baseline, Contrast, Saturation and Brightness (Colouring)	0.965	0.831
Baseline, Cropping	0.944	0.834
Baseline, Cropping and Colouring	0.933	<b>0.842</b>



**Abbildung 2:** Die Abbildung zeigt den Validierungs- $\langle R^2 \rangle$  Die Linie entspricht dem Durchschnitt über die fünf Durchläufe. Der schattierte Bereich repräsentiert eine empirische Standardabweichung.

## 2.4 Schlussfolgerung

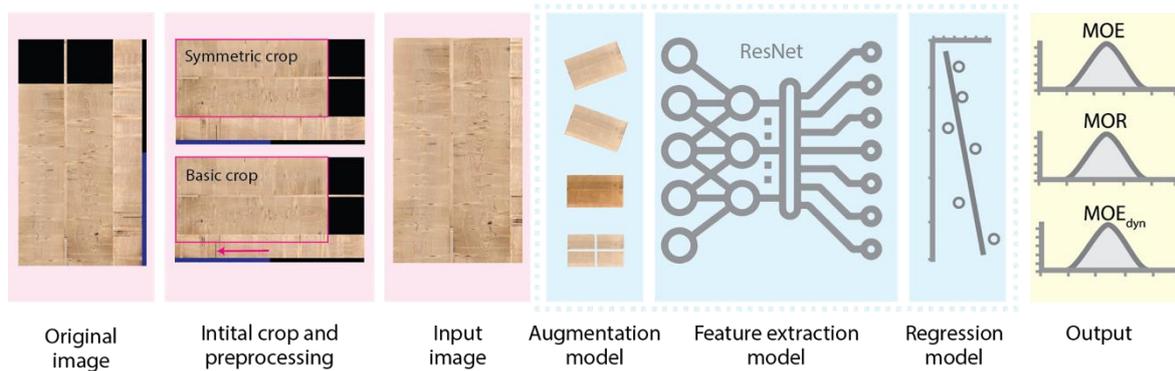
Die Entwicklung einer geeigneten Segmentierungsmethode sowie die Implementierung eines optimierten Daten-Augmentationsverfahrens erlaubte den Aufbau eines soliden und robusten Datensatzes, der für das Trainieren und Testen der CNN-Modelle geeignet war. Die Daten, die vom industriellen Partner zur Verfügung gestellt wurden, wurden vollständig verarbeitet und sind in einer Datenbank gespeichert, die in zukünftigen Projekten verwendet und kontinuierlich erweitert wird.

### 3. Computer Vision mit Convolutional Neural Networks (CNNs)

In diesem Abschnitt wird ein datenzentrierter Ansatz im Bereich des Deep Learning präsentiert. Dieser Ansatz weist sowohl praktische als auch theoretische Vorteile im Vergleich zu bestehenden statistischen und mathematischen Modellen auf. Er ist schnell, zuverlässig und kostengünstig und funktioniert, ohne auf vorbestehende mathematische Modelle und ad hoc-Parameter angewiesen zu sein. Es wurden tiefe neuronale Netzwerke, die typischerweise in der Bildverarbeitung verwendet werden, programmiert und getestet, um ein Modell zu entwickeln, welches die makroskopischen Muster und Merkmale auf der Oberfläche von gesägtem Holz den mechanischen Eigenschaften verknüpfen kann. Dabei wurde Bildverarbeitung mit Transfer Learning verwendet, indem einige der neueren residualen neuronalen Netzwerke angepasst wurden. Aus theoretischer Sicht war das Ziel, Holz als Beispiel für ein komplexes System zu verwenden und einen rein datenzentrierten Ansatz zu nutzen, um seine Eigenschaften zu charakterisieren und eine Verbindung zwischen Informationen und komplexer Struktur herzustellen. Aus praktischer Sicht wurde versucht eine vielseitige Methode einzuführen, die es der Holzindustrie, insbesondere kleinen und mittelständischen Unternehmen, ermöglicht, Holz schnell und einfach zu sortieren.

#### 3.1 Trainingsprozess

Bei der Entwicklung des CNN-Modells wurden zwei Teilziele verfolgt. Zum einen sollte eine hohe Genauigkeit und Generalisierungsfähigkeit des Modells erreicht werden. Zum anderen stand die Programmierung einer End-to-End-Pipeline im Fokus, die eine spätere Implementierung in den industriellen Sortierungsprozess erleichtern soll. Die entwickelte Pipeline ist in Abbildung 3 schematisch dargestellt und besteht aus drei Teilen, von denen jeder eine spezifische Aufgabe hat. Der erste Teil ist für die Datenvorbereitung verantwortlich. Der zweite Teil bildet das Rückgrat der gesamten Pipeline und besteht aus einem modifizierten ResNet-Modell. Der letzte Teil ist das Regressionsmodell, welches die Werte der mechanischen Eigenschaften vorhersagt. Bei der Ausgestaltung des Regressionsmodells wurde ein 'Multi-Target-Ansatz' verwendet, der es ermöglicht, mehrere Zielvariablen (MOE, MOR,  $MOE_{dyn}$ ) mit nur einem trainierten Modell vorherzusagen (ohne in der Genauigkeit Kompromisse eingehen zu müssen). Dies hat den Vorteil, dass der Modelltrainingsprozess deutlich effizienter und robuster war.



**Abbildung 3:** Diese schematische Darstellung zeigt die entwickelte End-to-End-Pipeline. Bildsegmentierung mit anschließender Data-Augmentation. Wichtige Features werden unter Verwendung des vortrainierten ResNet-Modells extrahiert. Die Bestimmung der mechanischen Eigenschaften der Holzlamelle erfolgt mittels eines Regressionsmodells.

Besonders wichtig für das erfolgreiche trainieren des CNN-Modells war 'Transfer Learning'. Transfer Learning ist eine Methode aus dem 'Deep Learning', mit der ein vortrainiertes künstliches neuronales Netz für die Lösung neuer Problemstellungen genutzt wird. Dazu wird der Lernfortschritt des bestehenden Modells transferiert. Dadurch ergeben sich Vorteile, wie: schnellere Erstellung, bessere Modellqualität und weniger Ressourceneinsatz. In dieser Arbeit wurde Transfer-Learning mit einer modifizierten Version des ResNet50-Modells, das auf dem Image-Net-21k-Datensatz [15, 16] vortrainiert wurde verwendet. Die Wahl erfolgte aufgrund seiner guten Leistung sowie der Benutzerfreundlichkeit und Verfügbarkeit [17, 18]. Es wurde sowohl Transfer-Learning mit als auch ohne Feinabstimmung der ResNet50-Architektur durchgeführt. Bei der Feinabstimmung werden alle originalen Gewichte des vortrainierten Modells 'aufgetaut', um sie während der Schulungsphase mit den domänenspezifischen Daten, in diesem Fall Holzbildern, zu modifizieren.

Das Modell wurde mit zwei unterschiedlichen Ansätzen trainiert. Zum einen wurde 'BiT Hyper-Rule' [15] angewendet, der auf stochastischem Gradientenabstieg basiert und einen Lernratenplaner verwendet. Zum anderen wurde der Adam-Optimierer [18] getestet, bei dem vor allem die Lernrate angepasst werden musste. Als Verlustfunktion wurde der mittlere quadratische Fehler (MSE) verwendet, da in dem Datensatz wenige Ausreisser vorkommen und dieser durch seine Einfachheit leicht zu interpretieren ist. Sind in einem Datensatz viele Ausreisser, kann man Verlustfunktionen verwenden, die weniger empfindlich gegenüber diese sind, wie der mittlere absolute Fehler (MAE) oder die Huber-Verlustfunktion [19].

## 3.2 BigTransfer-Modell

Im Folgenden werden die Ergebnisse der Festigkeitsvorhersage des ResNet50-Modells präsentiert, das auf dem ImageNet 21k-Datensatz vortrainiert wurde und dem eine vollständig verbundene Schicht hinzugefügt wurde, um die Regressionsziele (MOE, MOR, MOEdyn) zu bestimmen. Im Folgenden wird es, in Anlehnung an die Referenz [15], als das BigTransfer-Modell bezeichnet. Zum Zweck des Vergleichs wurde zudem ein einfaches neuronales Netzwerk mit fünf 'Convolutional layers' (CNN5-Modell) trainiert. Es diente als Referenz für die Leistungsvalidierung des BigTransfer-Lernalgorithmus. Für beide Modelle wurde die gleiche Pipeline und daher auch die gleiche Vorverarbeitungsmethode verwendet.

### 3.2.1 Performance Evaluation

In Tabelle 3 ist der durchschnittliche Bestimmtheitskoeffizient ( $R^2$ ) des BigTransfer-Modells für verschiedene Hyperparameter-Kombinationen dargestellt. Das beste Ergebnis wurde erreicht bei der Verwendung des BigTransfer-Modells, welches mit dem Adam-Optimierer, einer Lernrate von  $10^{-5}$  und der Verwendung von augmentierten Daten trainiert wurde.

**Tabelle 3:** Genauigkeit des BigTransfer-Modells. Die  $R^2$ -Werte basieren auf den Ausgabewerten, die durch die Anwendung des Modells auf den Validierungssatz generiert wurden.

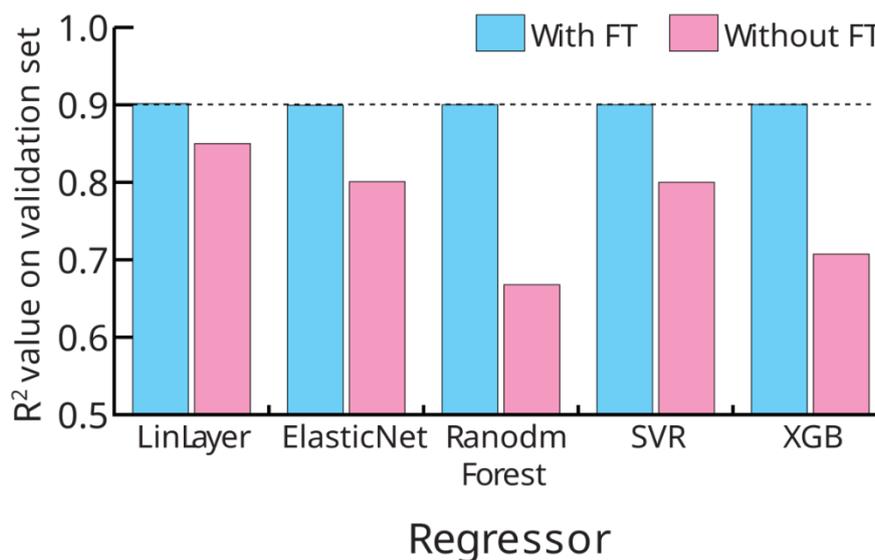
Model Accuracy in terms of the coefficient of determination based on the validation dataset				
Experiment	$\langle R^2 \rangle_5$ ( Var[ $R^2$ ] )			$\langle \langle R^2 \rangle \rangle_{IP}$ ( Var[ $\langle R^2 \rangle$ ] )
BC, NA, BiT Hyper Rule	[0.8481 (0.004), 0.8486 (0.003), 0.8359 (0.005)]			0.844 (0.004)
SC, NA, BiT Hyper Rule	[0.8543 (0.001), 0.8521 (0.002), 0.8413 (0.002)]			0.849 (0.002)
BC, A, BiT Hyper Rule	[0.8782 (0.001), 0.8729 (0.002), 0.8681 (0.002)]			0.873 (0.002)
SC, A, BiT Hyper Rule	[0.8801 (0.001), 0.8738 (0.001), 0.8704 (0.002)]			0.875 (0.001)
SC, A, Adam lr=0.0001	[0.8858 (0.006), 0.8793 (0.004), 0.8751 (0.005)]			0.880 (0.005)
BC, A, Adam lr=0.0001	[0.8868 (0.002), 0.8801 (0.002), 0.8765 (0.001)]			0.881 (0.002)
SC, A, Adam lr=0.00001	[0.897 (0.013), 0.8913 (0.012), 0.8881 (0.013)]			0.892 (0.013)
BC, A, Adam lr=0.00001	[0.9074 (0.017), 0.9015 (0.015), 0.8989 (0.020)]			<b>0.903 (0.017)</b>

Die erste Spalte enthält einen Identifikator für jedes Experiment. Er besteht aus einer Abkürzung für den Ausschnitt: BC für Basic crop und SC für den symmetric crop; gefolgt von A, wenn Augmentationsmethoden verwendet wurden, und NA, wenn keine Augmentationsmethoden verwendet wurden. Der letzte Teil des Identifikators (lr) gibt den Optimierer und die Lernrate an. Bestimmtheitsmass  $R^2$ , gemittelt über fünf unabhängige Durchläufe.

Mit dem Ziel, die Genauigkeit des BigTransfer-Modells zu verbessern, wurden verschiedene Regressionsmodelle anstelle der letzten linearen Schicht des BigTransfer-Modells implementiert (siehe Abbildung 4). Es wurden fünf der gängigeren Regressionsmodelle ausprobiert, mit und ohne Feinabstimmung. Es wurden folgende Modelle verwendet:

Support Vector Regression (SVR) mit radialer Basisfunktion (RBF)-Kernel, SVR mit einem Polynomkernel vom Grad 3, ElasticNet, Random Forest und XGBoost Regression. Die wichtigsten Hyperparameter der Algorithmen wurden unter Verwendung von Optuna [20], des  $R^2$ -Wertes und des Validierungsdatensatzes optimal eingestellt.

Selbst ohne Feintuning scheint die Repräsentation, die durch das vortrainierte ResNet50 bereitgestellt wird, für die vorliegende Aufgabe ausreichend zu sein, wie durch den  $R^2$ -Testwert von ElasticNet und SVR - RBF ohne Feintuning ersichtlich wird. Allerdings zeigen die Ergebnisse in Abbildung 4, dass die Leistung für jeden Regressor signifikant besser ist, wenn er aus der feinabgestimmten Version des Modells extrahiert wird. Dies ist zu erwarten, da das Feintuning es ermöglicht, die Gewichte des neuronalen Netzwerks in Reaktion auf die Lamellenbilder zu modifizieren. Interessanterweise übertrifft keiner der Regressoren das BigTransfer-Modell mit Feintuning, das eine All-to-All-Schicht als Regressor verwendet.

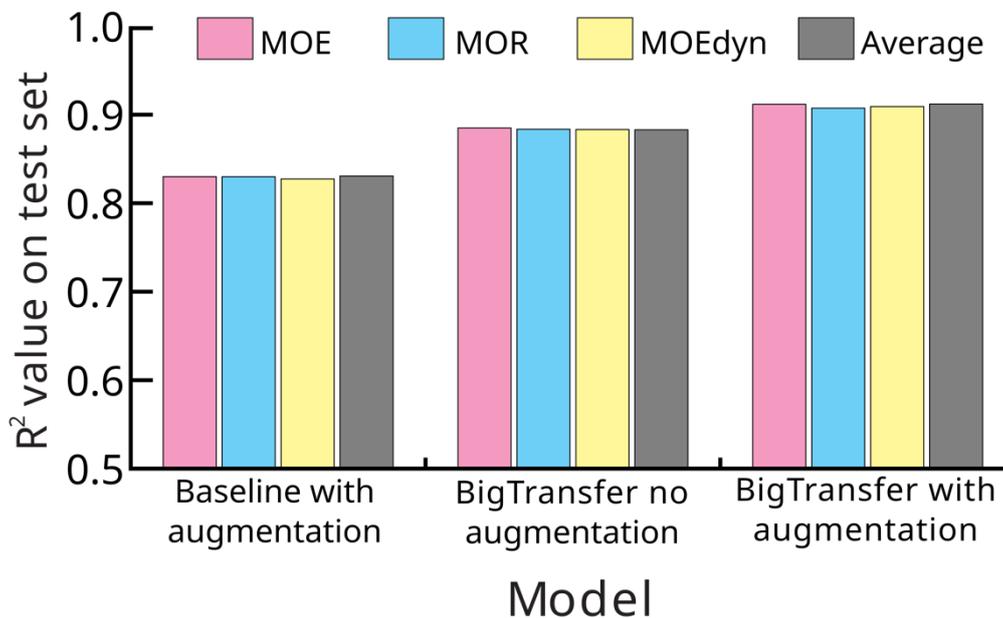


**Abbildung 4:** Modellgenauigkeit mit verschiedenen Regressoren auf dem Validierungsdatensatz.

In Abbildung 5 zeigen wir eine Zusammenfassung unserer Ergebnisse für das BigTransfer-Modell zusammen (mit und ohne Datenanreicherung) mit dem Basismodell CNN5. Es ist ersichtlich, dass das BigTransfer-Modell mit Data-Augmentation die besten Ergebnisse erzielte ( $R^2$ -Wert > 90%).

Formal ausgedrückt bedeutet dieser  $R^2$ -Wert, dass das BigTransfer-Modell mehr als 90% der Variabilität in den Daten erklären oder berücksichtigen kann. Die Qualität dieses Ergebnisses ist schwer zu beurteilen, da kein Vergleichsreferenzpunkt vorliegt. Selbst erfahrensten Sortierer ist es nahezu unmöglich, die mechanischen Festigkeiten

von Holzlamellen allein durch visuelle Inspektion zu bestimmen. Dennoch ist die Leistung unserer Methode vergleichbar mit der in Referenz [21], obwohl deutlich weniger Ressourcen und eine viel einfachere Methode verwendet wurde.



**Abbildung 5:** Übersicht über die Leistung des BigTransfer-Modells auf Testdaten im Vergleich zum Basismodell. Die farbigen Balken entsprechen Experimenten, bei denen nur ein IP vorhergesagt wurde. Der Wert ist ein Durchschnitt über 5 verschiedene Experimente. Der graue Balken entspricht dem Durchschnitt über die drei IPs.

Diese Ergebnisse deuten darauf hin, dass vor allem der Faserverlauf ein entscheidender Faktor ist um die mechanischen Eigenschaften zu bestimmen. Dieses Ergebnis steht im Einklang mit den Emergenz-Eigenschaften von faserbasierten Materialien, wie sie von Jeronimidis diskutiert wird [22], und den Ergebnissen, die in einer Studie von Ehrhart et al. erzielt wurden [11]. Die Autoren verwendeten eine Bildanalysetechnik mit vordefinierten Indikatoren um die Richtungen der Holzfaser zu bestimmen und damit die Festigkeit von Buchenbrettern abzuleiten.

### 3.3 Schlussfolgerung

Es konnte gezeigt werden, dass es möglich ist, Computer Vision zu verwenden, um die mechanische Festigkeit von Fichtenlamellen nur aus Bildern ihrer Oberfläche abzuleiten. Der Hauptbeitragsfaktor ist die Faserrichtung und die daraus resultierenden Muster und Holzstrukturen. Dies war, wie bereits erwähnt, auch die Schlussfolgerung von Erhart und Kollegen [11], die Buchenholz untersuchten. Die Autoren charakterisierten die lokalen Faserausrichtungen in Brettern mithilfe eines einzigen numerischen Bewer-

tungsparameters, der dann mit der Zugfestigkeit parallel zu den Brettern mittels verschiedener maschineller Lernverfahren korreliert wurde. Aus diesem Grund könnte Computer Vision auch bei anderen Holzarten funktionieren, unter der Voraussetzung, dass der Faserverlauf im Bild erkennbar ist. Um diese Hypothese zu verifizieren sollte eine systematische Studie an einer Vielzahl von Holzarten durchgeführt werden.

Eine einfache, genaue und robuste Technik zur Vorhersage der Festigkeit von Lamellen nur aus Bildern allein würde zu einer erheblichen Steigerung der Ressourceneffizienz führen und eine effiziente Sortierung ermöglichen. Die Idee ist es die Festigkeit einer Holzprobe unabhängig von ihrer Form, Grösse und Art genau vorhersagen zu können und sofort den besten Verwendungszweck zu identifizieren.

## **4. Experimentelle Validierung des CNN-Modells**

Eine Revision der Golden Eye Sortiermaschine wurde durchgeführt. Nach der maschinellen Sortierung wurden die Lamellen von unabhängigen Experten experimentell geprüft, um die tatsächlichen mechanischen Eigenschaften zu bestimmen. Aufgrund des industriellen Charakters der Daten und ihrer vertraulichen Natur werden diese in diesem Bericht nicht veröffentlicht. Die Auswertung und Validierung unserer Modelle zeigen jedoch, dass sie äusserst präzise sind. Zudem ermöglicht Transfer Learning die Anpassung der Gewichte unserer Modelle, sodass sie die tatsächlichen mechanischen Eigenschaften präzise widerspiegeln.

### **4.1 Schlussfolgerung**

Die Auswertung und Validierung der Modelle zeigt, dass diese äusserst präzise sind. Besonders bemerkenswert ist die Fähigkeit des Transfer Learning, die Gewichte der Modelle so anzupassen, dass sie die realen mechanischen Eigenschaften genau widerspiegeln.

## **5. Fazit**

In diesem Projekt wurde erfolgreich ein umfassender Trainings-, Validierungs- und Testdatensatz mit hochqualitativen Holzbildern mittels Data Augmentation-Techniken erstellt. Dieser Datensatz repräsentiert verschiedene Texturen und Zustände der Holzlamellen und ermöglichte das erfolgreiche Trainieren und Testen von Convolutional Neural Network (CNN)-Modellen.

Die CNN-Modelle, insbesondere das entwickelte BigTransfer Modell, zeigten vielversprechende Ergebnisse, indem sie eine hohe Genauigkeit bei der Bestimmung der mechanischen Eigenschaften von Holzlamellen erreichten. Die Effektivität der Modelle im Training sowie ihre Generalisierungsfähigkeit auf unterschiedliche Bedingungen wurden durch eine sorgfältige Validierung mit Testdaten und durchgeführten Experimenten erfolgreich nachgewiesen.

Die Ergebnisse dieses Projekts zeigen, dass es möglich ist, die mechanischen Eigenschaften von Holz allein durch die Analyse von Bildern zu bestimmen. Die Ergebnisse können nicht nur zur Weiterentwicklung des Sortierprozesses beitragen, sondern fördern auch das Verständnis und die Anwendung von maschinellem Lernen in der Holztechnologie.

## **5.1 Umsetzung**

Es wurde eine ML-Pipeline in Python programmiert, die den Implementierungsprozess des entwickelten Modells erleichtern soll. Ein Schema dieser Pipeline ist in Abbildung 3 dargestellt. Das im Rahmen dieses Projekts erzielte Endmodell (BigTransfer Modell) wird in dem neuen Demonstrationszentrum implementiert. So ist die Installation eines semi-automatisierten Kamera-Systems in den Laboren der Empa für weiterführende Untersuchungen im Gange. Diese Einrichtung ermöglicht eine effiziente und präzise Erfassung von Bildmaterial für weitere Forschungszwecke. Das entwickelte Modell kann aktiv für weitere Forschungsvorhaben eingesetzt werden, wie sie in den gestellten Forschungsfragen näher definiert sind. Die Integration dieses Modells in den Laborbetrieb eröffnet neue Möglichkeiten zur automatisierten Analyse und Bewertung von Proben, insbesondere im Zusammenhang mit den physikalischen Eigenschaften von Holzlamellen.

Die in diesem Projekt erzielten Ergebnisse sind äusserst vielversprechend und belegen die Machbarkeit physikalische Eigenschaften von Holzlamellen direkt aus einem Bild zu bestimmen. Die erfolgreichen Ergebnisse legen nahe, dass das gewählte Konzept einen vielversprechenden Ansatz für die Anwendung in der Praxis darstellt. Daher ist in Planung die Forschung auf diesem Gebiet fortzusetzen und mit der Schweizer Holzindustrie umzusetzen.

## 5.2 Offene Fragen / Weiterer Forschungsbedarf

Aus den Erkenntnissen des vorliegenden Projektes ergeben sich folgende offenen Fragen und Ansatzpunkte für weitere Untersuchungen:

- **Übertragbarkeit des Modells:** Wie gut lässt sich das entwickelte Modell auf unterschiedliche Holzarten und –zustände (Dimensionen, Form, etc.) übertragen?
- Gibt es spezifische Herausforderungen bei der Anwendung des Modells auf verschiedene Holzproben (-arten)?
- **Integration in industrielle Prozesse:** Wie kann das Modell in bestehende Produktions- und Qualitätskontrollprozesse der Holzindustrie integriert werden?
- Welche Anpassungen sind erforderlich, um eine nahtlose Integration in industrielle Abläufe zu gewährleisten?
- **Einfluss auf die Holzproduktqualität:** Kann das Modell dazu beitragen, den Einsatz von Holzressourcen effizienter zu gestalten?
- **Nutzen für die Holzindustrie:** Wie können die Prozesse in der Holzproduktion durch die Anwendung von Machine Learning optimiert werden?
- **Langfristige Leistung und Aktualisierbarkeit:** Welche Strategien können verfolgt werden, um das Modell bei Bedarf zu aktualisieren und an neue Anforderungen anzupassen?

## Literaturverzeichnis

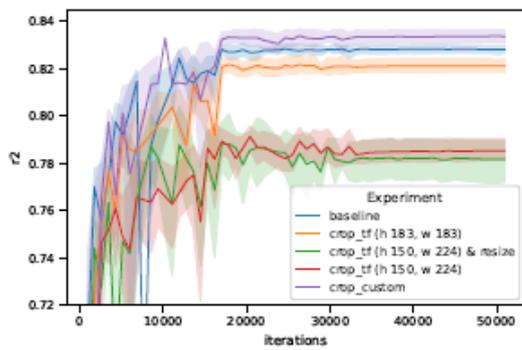
- [1] Lehmann, S., Sustainable Construction for Urban Infill Development Using Engineered Massive Wood Panel Systems. *Sustainability* 2012, 4 (10), 2707-2742
- [2] Sedjo, R. A., Forests. *Environment: Science and Policy for Sustainable Development* 1989, 31 (1), 14-20
- [3] Woodbury, P.; Wightman, J., Forest Management & Greenhouse Gas Mitigation Opportunities. *Information Sheet* 2017, 7
- [4] Pardos, M.; del Río, M.; Pretzsch, H.; Jactel, H.; Bielak, K.; Bravo, F.; Brazaitis, G.; Defosse, E.; Engel, M.; Godvod, K.; et al. The greater resilience of mixed forests to drought mainly depends on their composition: Analysis along a climate gradient across Europe. *For. Ecol. Manag.* 2021, 481, 118687
- [5] Spiecker, H. Silvicultural management in maintaining biodiversity and resistance of forests in Europe temperate zone. *J. Environ. Manage.* 2003, 67, 55–65
- [6] Wolf, S.; Paul-Limoges, E. Drought and heat reduce forest carbon uptake. *Nat Commun.* 2023, 14, 6217
- [7] Schubert, M.; Panzarasa, G.; Burgert, I., Sustainability in Wood Products: A New Perspective for Handling Natural Diversity. *Chem. Rev.* 2023, 123 (5), 1889–1924
- [8] Schubert, M.; Kläusler, O., Applying machine learning to predict the tensile shear strength of bonded beech wood as a function of the composition of polyurethane prepolymers and various pretreatments. *Wood Sci. Technol.* 2020, 54, 19-29
- [9] Schubert, M.; Lukovic, M.; Christen, H., Prediction of mechanical properties of wood fiber insulation boards as a function of machine and process parameters by random forest. *Wood Sci. Technol.* 2020, 54, 703– 713
- [10] Schubert, M.; Sonderegger, W.; Luković, M.; Kläusler, O., Semi-supervised learning for quality control of high-value wood products. *Wood Sci. Technol.* 2022, 56, 1439–1453
- [11] Ehrhart, T.; Palma, P.; Schubert, M.; Steiger, R.; Frangi, A. Predicting the strength of European beech (*Fagus sylvatica* L.) boards using image-based local fibre direction data. *Wood Sci. Technol.* 2022, 56, 123–146
- [12] Luković M, Zweifel R, Thiry G, Zhang C, Schubert M., Reconstructing radial stem size changes of trees with machine learning. *J. R. Soc. Interface* 2022, 19, 20220349.
- [13] Microtec, Official website, 2022.
- [14] Ritter, M. A., Timber bridges: Design, construction, inspection, and maintenance, Washington, DC, USA, 1990.

- [15] Kolesnikov, A., Beyer, L., Zhai, X., Puigcerver, J., Yung, J., Gelly S., Houlsby, N. Big Transfer (BiT): General Visual Representation Learning. In Computer Vision – ECCV 2020, Cham
- [16] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K.; Fei-Fei, L., ImageNet: A large-scale hierarchical image database. In 2009 IEEE Conference on Computer Vision and Pattern Recognition
- [17] He, K., Zhang, X., Ren, S., Sun, J., Deep Residual Learning for Image Recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
- [18] Kingma, D. P., Ba, J., Adam: A Method for Stochastic Optimization. In 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings
- [19] Huber, P. J., Robust Estimation of a Location Parameter. In Springer Series in Statistics, Springer New York, 1992, p. 492–518.
- [20] Akiba, T., Sano, S., Yanase, T., Ohta T., Koyama M., Optuna. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019, 2623–2631
- [21] Huber, J. A. J., Broman, O., Ekevad, M., Oja, J., Hansson, L., A method for generating finite element models of wood boards from X-ray computed tomography scans, *Computers & Structures*, 2022, 260, 106702.
- [22] Jeronimidis, G., Vincent, J. F. V., Connective tissue matrix, D. W. L. Hukins, Ed., Macmillan, London, 1984.

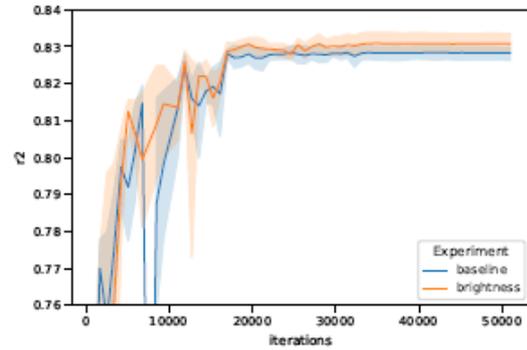
## Anhang

Die verwendeten Augmentationsmethoden waren:

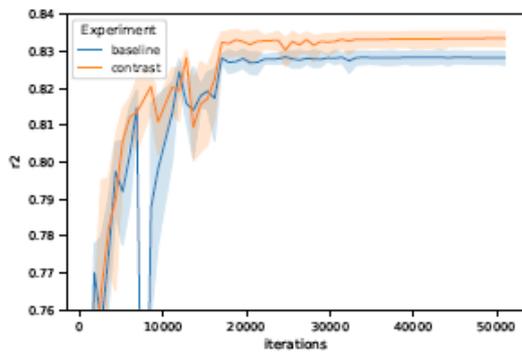
- **Random flipping:** Zufälliges horizontales, vertikales oder kombiniertes Spiegeln aller Bilder in einem Batch.
- **Random rotation:** Zufälliges Drehen der Bilder in einem Batch um höchstens  $\pm 3$  Grad.
- **Random cropping (cropCustom):** Diese Methode schneidet die Bilder zufällig für 30% der Batches zu. Sie wählt zufällig gleichmässig aus, welche spezifischen Batches zugeschnitten werden sollen. Die Zuschneidegrösse wird gemäss des gewünschten Faktors von Höhe und Breite berechnet. Wir haben diesen Faktor auf 80% gesetzt. Dann wird die Zuschneideposition zufällig ausgewählt. Ausserdem wird das zugeschnittene Bild auf die ursprüngliche Eingabebildgrösse skaliert.
- **Random cropping (cropTF):** Diese Methode schneidet die Bilder zufällig entsprechend einer festen Grösse. Wir haben folgende Zuschneidegrössen in Betracht gezogen: 80% der Höhe und Breite (d.h. Höhe 150, Breite 224) und ein quadratisches Format  $183 \times 183$ . Die letztere Wahl wurde durch die Tatsache inspiriert, dass ImageNet-Bilder Quadrate von etwa dieser Grösse sind.
- **Random contrast:** Zufällige Anpassung des Kontrasts der Bilder in einem Batch. Der Kontrastfaktor wird zufällig zwischen einer unteren (l) und oberen (u) Grenze gewählt, d.h.  $l = 0.5, u = 1$ .
- **Random brightness:** Zufällige Anpassung der Helligkeit der Bilder in einem Batch. Der Helligkeitsunterschied wird zufällig zwischen  $[-\Delta, \Delta]$  abgetastet, wobei  $\Delta$  dem maximalen Wert entspricht, den wir den Pixeln hinzufügen. Wir haben  $\Delta$  auf 0.1 festgelegt.
- **Random saturation:** Zufällige Anpassung der Sättigung der Bilder in einem Batch. Der Sättigungsfaktor wird zufällig zwischen einer unteren (l) und oberen (u) Grenze gewählt, d.h.  $l = 0.5, u = 2$ .
- **Random sharpening:** Diese Methode erfordert zwei Parameter für das zufällige Schärfen eines Bildes: den Schärfungsfaktor und die Wahrscheinlichkeit. Der erste beschreibt die Menge, mit der ein Bild geschärft wird. Der zweite gibt den Anteil der Trainingsbatches an, die geschärft werden sollen. Wir haben den Schärfungsfaktor auf 2 gesetzt und 50% geschärft.
- **Injection of Gaussian noise:** Diese Augmentation fügt den Bildern Gaussisches Rauschen hinzu. Wir haben Standardabweichungen von 0.001 und 0.0001 in Betracht gezogen.
- **Application of Gaussian blur:** Diese Augmentation weicht Bilder mit einem Gausschen Filter auf. Es nimmt zwei Parameter: den Prozentsatz der Batches, auf die die Augmentation angewendet werden soll, und Sigma, die Standardabweichung des Gauss-Filters. Wir haben Wahrscheinlichkeiten von 0.25, 0.5 und 1 sowie Sigmas von 0.1 und 0.5 getestet.



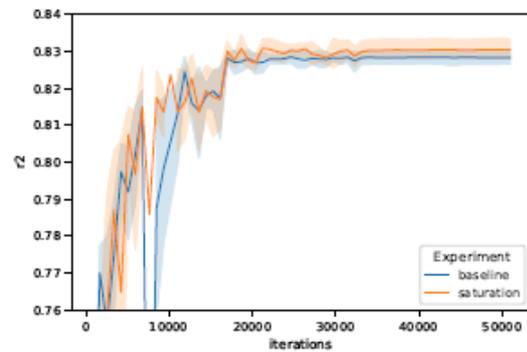
(a) **Random cropping:** consider crop\_tf with cropping sizes (150, 224) and (183, 183), crop\_tf with additional resizing and crop\_custom.



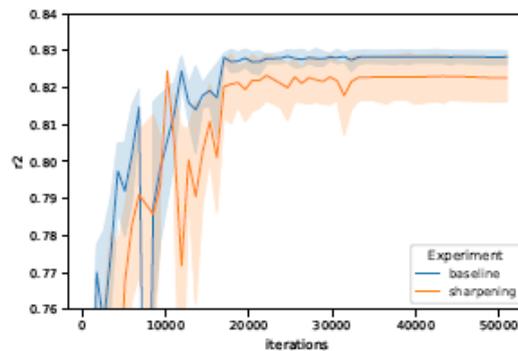
(b) **Random brightness:** Adjust image brightness by addition of  $\delta$  randomly chosen from  $[-0.1, 0.1]$



(c) **Random contrast:** Adjust image contrast by multiplying  $\delta$  randomly chosen from  $[0.5, 1]$



(d) **Random saturation:** Adjust image contrast by multiplying  $\delta$  randomly chosen from  $[0.5, 2]$



(e) **Random sharpening:** Adjust image contrast by adding  $\delta$  randomly chosen from  $[0, 2]$



Mark Schubert



Mirko Lukovic